ISSN 1799-2591 (print); ISSN 2053-0692 (online)
Theory and Practice in Language Studies, Vol. 16, No. 2, pp. 605-616, February 2026
DOI: https://doi.org/10.17507/tpls.1602.26

Teaching for Assessment in EFL Contexts: Modeling Aptitude, Awareness, Anxiety, and Self-Efficacy as Predictors of Teacher Engagement

Wael H. Alharbi
English Department, Yanbu English Language and Preparatory Year Institute, The Royal Commission for Jubail and Yanbu, Yanbu Industrial City, Saudi Arabia

***Abstract*—In assessment-driven educational systems, particularly within Saudi EFL contexts, instructors face increasing pressure to align their teaching with high-stakes testing frameworks. While individual teacher attributes such as language aptitude, metalinguistic awareness, self-efficacy, and anxiety have been studied in isolation, their interactive effects on Teaching for Assessment (TFA) practices remain underexplored. This study aimed to investigate how these cognitive and affective factors predict and mediate TFA engagement among university EFL instructors in Saudi Arabia. A quantitative, cross-sectional design was employed using five validated instruments to measure language aptitude, metalinguistic awareness, language anxiety, teacher self-efficacy, and TFA practices. A purposive sample of 247 Saudi university EFL instructors participated. Hierarchical regression, mediation (PROCESS Model 4), and moderated mediation (Models 14 and 58) analyses were conducted using SPSS and Hayes' PROCESS Macro. Self-efficacy was the strongest positive predictor of TFA practices, followed by metalinguistic awareness. Language anxiety negatively influenced TFA, while language aptitude showed only an indirect effect via metalinguistic awareness. Moderation analysis revealed that anxiety weakened the relationship between metalinguistic awareness and TFA, while self-efficacy enhanced the entire indirect pathway from aptitude to TFA. These findings validate a moderated mediation model and highlight complex cognitive-affective dynamics influencing assessment behavior. The study confirms that cognitive ability alone is insufficient to foster assessment-aligned teaching. Rather, the interplay of self-efficacy, metalinguistic awareness, and anxiety significantly shapes TFA practices. These insights underscore the need for professional development programs that enhance both cognitive skills and emotional readiness among EFL instructors.**

***Index Terms*—Teaching for Assessment (TFA), teacher self-efficacy, language aptitude, metalinguistic awareness, language anxiety**

I. Introduction

In educational systems shaped by accountability and standardized testing, teachers are increasingly expected to align instruction with measurable learning outcomes. Teaching for Assessment (TFA) has thus emerged as a key pedagogical approach, requiring the seamless integration of formative and summative assessment practices (Giraldo, 2019; Richards, 2022). This is especially pertinent in English as a Foreign Language (EFL) programs across the Gulf, including Saudi universities, where curricula are often tied to international benchmarks such as IELTS and TOEFL.

Despite this growing emphasis, the factors influencing teachers' ability to implement TFA effectively remain underexplored, particularly in linguistically and culturally diverse contexts. In multilingual societies, EFL teachers navigate cross-linguistic frameworks that both support and complicate instructional design. This raises important questions about the role of individual differences—including cognitive and affective variables—in shaping assessment-related teaching behaviors.

Among cognitive factors, language aptitude—a natural capacity for language learning comprising phonetic coding, grammatical sensitivity, and inductive skills—is critical. Teachers with high aptitude are better positioned to perform linguistic analysis and adapt instruction to assessment needs (Carroll & Sapon, 1959; Wen et al., 2017; Doughty, 2019). Closely related is metalinguistic knowledge, or the explicit understanding of language rules and structures, which enables accurate feedback and alignment with rubrics—essential in test-oriented settings (Andrews, 2007; Calafato, 2024).

Affective variables such as self-efficacy and anxiety also shape TFA engagement. Self-efficacy—confidence in one's ability to execute teaching and assessment tasks—predicts instructional persistence and innovation (Bandura, 1997; Wyatt, 2021; Faez et al., 2021). Conversely, language anxiety can undermine teaching confidence, reduce metalinguistic engagement, and hinder the application of flexible assessment practices (Gregersen, 2020).

Given the interplay of these factors, this study investigates how language aptitude, metalinguistic knowledge, self-efficacy, and anxiety jointly influence TFA practices among multilingual EFL instructors in Saudi universities. It aims to

provide theoretically robust, empirically grounded insights relevant to assessment reform and professional development in multilingual EFL settings.

II. Literature Review

This section synthesizes research on language aptitude and its links to metalinguistic knowledge, self-efficacy, anxiety, and language maintenance among multilingual language teachers, with implications for Teaching for Assessment (TFA) in structured EFL contexts like Saudi universities. It builds toward research questions addressing how cognitive and affective factors relate to TFA (RQ1), demographic differences (RQ2), collective prediction (RQ3), mediation by anxiety (RQ4), and impacts on TFA subdimensions (RQ5), highlighting institutional mediation in contexts like Saudi Arabia.

A. Language Aptitude and Metalinguistic Knowledge

Language aptitude—comprising abilities like phonetic coding and grammatical sensitivity—predicts acquisition and links to working memory (Carroll & Sapon, 1959; Doughty, 2019; Wen et al., 2017). Metalinguistic knowledge supports feedback in assessment-heavy contexts (Andrews, 2007; Aslan, 2015). Learner studies show modest correlations ($r = .32$, $p < .001$; Alderson et al., 1997; Calafato, 2024), while teacher studies report stronger links (Aslan, 2015; D'Angelo & Sorace, 2022). For instance, Calafato (2024) found aptitude rising with metalinguistic scores ($M = 22.50$ to $M = 37.29$) among Uzbek teachers. Halali et al. (2023) noted that AI supports knowledge acquisition but neglects aptitude. Learner studies often underplay teaching demands, while teacher-focused research highlights practical use (Giraldo, 2019; Hill & Ducasse, 2022).

This suggests that teachers' linguistic repertoires, especially in multilingual contexts, may shape engagement in assessment practices, prompting exploration of multilingualism's role.

B. Multilingualism and Its Effects

Multilingualism enhances aptitude via cross-linguistic transfer (Cenoz, 2013). Studies show boosts in reasoning tasks (Thompson, 2013; Hopp et al., 2019), with formal learning predicting higher aptitude ($B = 4.46$ taught; $B = 2.49$ learned; Calafato, 2024). Otwinowska (2017) linked it to plurilingual awareness in Polish teachers, while monolingual bias limits flexibility. Cenoz and Gorter (2020, 2021, 2022) found translanguaging aids metalinguistic tasks in Basque EFL, though constraints exist elsewhere (Hill & Ducasse, 2022).

Overall, formal acquisition benefits are stronger than innate repertoires (Calafato, 2024). Teachers may leverage multilingualism in TFA, yet affective variables like anxiety could moderate outcomes (Gong et al., 2025).

While multilingualism influences aptitude and knowledge, affective factors like self-efficacy and anxiety determine how skills are applied in high-stakes EFL assessment, warranting deeper analysis.

C. Affective Factors: Self-Efficacy and Anxiety

Self-efficacy—the belief in one's ability to execute tasks—correlates with proficiency and effectiveness, shaped by context and experience (Bandura, 1997; Hoang & Wyatt, 2021). Faez et al. (2021) reported $r = .37$ across 19 studies. Wyatt and Dikilitaş (2019) found that low self-efficacy led Turkish teachers to avoid grammar instruction, undermining TFA alignment.

Anxiety—classroom, test, or communication-related—typically hampers aptitude and efficacy (Sparks & Ganschow, 2007; Gregersen, 2020). It reduces performance (Fraschini & Park, 2021), though moderate anxiety may boost engagement ($B = 1.18$, $p = .014$; Calafato, 2024; Fallah et al., 2023). Self-efficacy helps mitigate anxiety (Al Hanake, 2024), but TFA pressures may intensify it (Richards, 2022).

Given these links, examining how language maintenance habits support proficiency and affective resilience is necessary.

D. Language Maintenance Habits

Maintenance sustains proficiency. High-aptitude teachers prefer active strategies (85.7% in top quartile; Calafato, 2024; Otwinowska, 2017). Diverse practices link to advanced proficiency (Al-Wossabi, 2022), while lower aptitude correlates with passive habits (Godfroid & Kim, 2021). U.S. and Uzbek studies show environmental constraints may restrict maintenance (Richards, 2022).

Thus, exploring how aptitude, awareness, and affect shape TFA practices in institutional contexts is essential (Giraldo, 2019; Hill & Ducasse, 2022).

E. Interactions and Implications for TFA Practices

Aptitude interacts with knowledge, affect, and multilingualism, jointly predicting outcomes ($R^2 = .31$; Calafato, 2024). While high aptitude enables feedback, systems often standardize practices (Giraldo, 2019; Hill & Ducasse, 2022).

Multilingual pedagogies and translanguaging enhance TFA (Cenoz & Gorter, 2020, 20221, 2022), though anxiety moderates gains (Gong et al., 2025). Self-efficacy can buffer anxiety (Faez et al., 2021), but pressures can cause inconsistencies.

These interlinked factors emphasize the need to study their long-term and context-specific impacts, especially in structured EFL contexts like Saudi Arabia, informing questions on predictors (RQ1, RQ3) and demographic differences (RQ2).

F. Research Gaps and Future Directions

Few longitudinal studies track aptitude development (Calafato, 2024). Self-efficacy and anxiety interventions rarely target EFL (Fallah et al., 2023; Gong et al., 2025), and contextual mediators remain underexplored (Wyatt & Dikilitaş, 2019). Future research should adopt longitudinal and mixed-method approaches to investigate how aptitude and habits evolve under structural constraints (Richards, 2022; Hill & Ducasse, 2022), supporting integrated training in Saudi EFL that addresses aptitude, awareness, multilingualism, and affective needs (Otwinowska, 2017; Faez et al., 2021).

III. METHODOLOGY

This study employed a cross-sectional, quantitative, descriptive correlational design to examine the relationships between language aptitude, metalinguistic awareness, language learning anxiety, teacher self-efficacy, and Teaching for Assessment (TFA) practices among EFL teachers in Saudi universities. This design—suitable for exploring multifaceted teacher behaviors in assessment-driven contexts—allowed investigation of associations, predictive links, and mediation effects without inferring causality (Creswell & Creswell, 2017). It also enabled advanced statistical analyses, including hierarchical regression and mediation, while controlling for demographic factors. Ethical procedures followed American Educational Research Association (AERA, 2011) guidelines, including IRB approval, informed consent, anonymity, and participants' right to withdraw.

A. Research Questions and Hypotheses

The study is guided by the following research questions, focused on Saudi university EFL teachers:

1) **Research Question 1:** What is the relationship between Saudi university EFL teachers' self-reported Teaching for Assessment (TFA) practices and their individual cognitive and affective factors, specifically language aptitude, language learning anxiety, metalinguistic awareness, and language teacher self-efficacy?
   - **a) Hypothesis 1a:** Language aptitude will be positively correlated with TFA practices, as higher aptitude may enhance teachers' ability to align instruction with assessment demands (Doughty, 2019).
   - **b) Hypothesis 1b:** Metalinguistic awareness will be positively correlated with TFA practices, given its role in supporting explicit language instruction and error correction relevant to assessment (Andrews, 2007).
   - **c) Hypothesis 1c:** Language teacher self-efficacy will be positively correlated with TFA practices, as confident teachers are more likely to integrate assessment-aligned strategies effectively (Bandura, 1997; Wyatt & Dikilitaş, 2019).
   - **d) Hypothesis 1d:** Language learning anxiety will be negatively correlated with TFA practices, as heightened anxiety may hinder effective integration of assessment-focused strategies (Sparks & Ganschow, 2007).
2) **Research Question 2:** Are there statistically significant differences in self-reported Teaching for Assessment (TFA) practices among Saudi university EFL teachers when grouped by their demographic characteristics (gender, nationality, age group, academic degree, and years of teaching experience)?
   Hypothesis 2: No statistically significant differences in TFA practices will be observed across demographic groups, as institutional systems may standardize practices, outweighing individual demographic influences (Popham, 2018).
3) **Research Question 3:** To what extent do language aptitude, language learning anxiety, metalinguistic awareness, and language teacher self-efficacy, as a collective set of factors, predict the Teaching for Assessment (TFA) practices of Saudi university EFL teachers?
   Hypothesis 3: Language aptitude, language learning anxiety, metalinguistic awareness, and language teacher self-efficacy will collectively predict a significant portion of variance in TFA practices, with metalinguistic awareness and self-efficacy as the strongest positive predictors (Aslan, 2015).
4) **Research Question 4:** To what extent does language learning anxiety mediate the relationship between language aptitude and Teaching for Assessment (TFA) practices among Saudi university EFL teachers?
   Hypothesis 4: Language learning anxiety will mediate the relationship between language aptitude and TFA practices, with higher anxiety reducing the positive effect of aptitude on TFA (Gregersen, 2020).
5) **Research Question 5:** To what extent do language aptitude, language learning anxiety, language teacher self-efficacy, and metalinguistic awareness differentially predict the five dimensions of Teaching for Assessment (TFA)—General Beliefs about Assessment, Alignment of Instruction with Assessment, Classroom Practices and Feedback, Perceptions of Impact, and Institutional and Cultural Context—among Saudi university EFL teachers?
   Hypothesis 5: The predictors will differentially affect TFA dimensions, with anxiety and metalinguistic awareness showing stronger effects on institutional and cultural context due to contextual pressures in Saudi universities (Giraldo, 2019; Calafato, 2024).

These questions and hypotheses were formulated to fill gaps in the literature on assessment practices in multilingual EFL contexts, drawing on theoretical frameworks such as sociocultural theory and ecological systems theory (Chong et al., 2023), and are designed for empirical testing in a high-stakes assessment environment like Saudi Arabia.

B. Participants and Sampling

A total of 247 EFL instructors employed at Saudi universities participated in the study. Participants were recruited through purposive sampling, targeting multilingual educators actively engaged in teaching within Saudi Arabia's university-level language programs. The demographic distribution was near gender-balanced (50.2% female), and spanned diverse nationalities, including Saudi (49.0%), Egyptian (19.0%), Sudanese (12.6%), and Indian (10.1%). Most participants held Master's degrees (57.5%) or PhDs (30.4%), with 54.3% possessing over 16 years of teaching experience. These parameters reflect a sample representative of the multilingual, multicultural nature of Saudi tertiary EFL contexts.

C. Instruments

Data were collected via a self-report questionnaire comprising five scales, all using a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree), except the aptitude test. Instruments were adapted to the Saudi EFL context, with Arabic translations provided. Content validity was confirmed by three applied linguistics professors reviewing item relevance and cultural appropriateness. Construct validity was tested through exploratory factor analysis (principal axis factoring, oblimin rotation; loadings > 0.40). The pilot, involving 20 excluded Saudi EFL teachers, assessed clarity, timing (20–30 minutes), and cultural fit; minor wording changes followed. Pilot reliabilities ($\alpha = 0.75–0.88$) were matched or exceeded in the main study. Test-retest reliability ($n = 15$, 2-week interval) exceeded $r = 0.80$ for all scales.

(a). Teaching for Assessment Questionnaire (TFA)

Adapted from Muth et al. (2025), this 27-item scale measures perceptions and practices of instruction-assessment alignment, with five subscales: General Beliefs about Assessment (5 items, e.g., "Assessment is an essential part of the teaching and learning process"; $\alpha = .801$), Alignment of Instruction with Assessment (7 items, e.g., "I often review the test format before planning my lessons"; $\alpha = .880$), Classroom Practices and Feedback (5 items, e.g., "I provide students with feedback on how to improve their test performance"; $\alpha = .724$), Perceptions of Impact (5 items, e.g., "Focusing on assessment helps students perform better overall"; $\alpha = .702$), and Institutional and Cultural Context (5 items, e.g., "My institution expects me to prepare students for standardized exams"; $\alpha = .707$). Item 19 was reverse-coded. EFA confirmed the five-factor structure (KMO = .82, Bartlett's $p < .001$).

(b). Language Aptitude Test

A researcher-designed test (50 points total), adapted from MLAT (Carroll & Sapon, 1959), measuring Grammatical Sensitivity (10 points), Phonetic Coding (10 points), Working Memory (20 points), Inductive Language Learning (10 points), and Paired Associate Learning (10 points). Items were tailored for EFL teachers. Pilot $\alpha = .80$; main $\alpha = .812$.

(c). Language Learning Anxiety Scale

Adapted from Horwitz (1986), this 12-item scale assesses anxiety in four domains: Communication Apprehension (3 items, $\alpha = .815$), Fear of Negative Evaluation (3 items, $\alpha = .755$), Test Anxiety (3 items, $\alpha = .849$), and Classroom Anxiety (3 items, $\alpha = .750$). Pilot $\alpha = .81$ overall.

(d). Metalinguistic Awareness Questionnaire

A 10-item scale evaluating five domains: Morphological Awareness (2 items, $\alpha = .791$), Syntactic Awareness (2 items, $\alpha = .770$), Semantic Awareness (2 items, $\alpha = .778$), Pragmatic Awareness (2 items, $\alpha = .882$), and Phonological Awareness (2 items, $\alpha = .756$). Pilot $\alpha = .78$ overall.

(e). Language Teacher Self-Efficacy Scale

Adapted from Tschannen-Moran and Hoy (2001), this 12-item scale covers five domains: Instructional Self-Efficacy (3 items, $\alpha = .828$), Assessment Self-Efficacy (3 items, $\alpha = .812$), Technology Self-Efficacy (2 items, $\alpha = .816$), Classroom Management Self-Efficacy (2 items, $\alpha = .810$), and Professional Engagement Self-Efficacy (2 items, $\alpha = .727$). Pilot $\alpha = .82$ overall.

Demographic items (gender, age, experience, degree, nationality) were included. All scales showed acceptable to excellent reliability ($\alpha \geq .70$; Nunnally & Bernstein, 1994).

D. Data Collection Procedures

Data were collected from June to July 2025 using Google Forms for broad reach and reduced bias (Dillman et al., 2014). Invitations were disseminated via university networks in Saudi Arabia, with electronic informed consent obtained. Responses were anonymous (no IP tracking). Of 300 surveys, 247 valid responses were retained (82.3% rate) after excluding incompletes. No incentives were used to prevent coercion.

E. Data Analysis

All analyses were conducted using IBM SPSS Statistics (v27). Internal consistency for all scales and subscales was assessed via Cronbach's alpha. Descriptive statistics (frequencies, percentages, means, and standard deviations) were used to summarize demographic and variable data.

Inferential analysis proceeded in several stages. Pearson correlation coefficients were used to examine relationships among key variables, while independent samples t-tests (gender) and one-way ANOVAs (age, degree, experience) assessed group differences in TFA practices.

To address the primary research question, a hierarchical multiple regression was conducted to assess whether language aptitude, anxiety, self-efficacy, and metalinguistic awareness predicted TFA scores, controlling for demographics. Mediation and moderation analyses were then performed using the Hayes PROCESS macro (Models 4 and 1, respectively). To examine TFA as a multidimensional construct, five separate multiple regressions were run using the five TFA subscales as outcome variables; a Bonferroni correction adjusted significance to $p < .01$.

Lastly, due to initial non-significant regression results, a post-hoc power analysis was conducted in G*Power 3.1 to assess statistical power. Unless otherwise noted, alpha was set at $p < .05$ for all inferential tests.

IV. Result

A. Reliability Analysis

Cronbach's alpha was used to assess internal consistency. For the Teaching for Assessment (TFA) scale, Cronbach's alpha values for its five subscales were as follows: General Beliefs about Assessment (α = .801), Alignment of Instruction with Assessment (α = .880), Classroom Practices and Feedback (α = .724), Perceptions of Impact (α = .702), and Institutional and Cultural Context (α = .707). The total score for the Language Aptitude test had a Cronbach's alpha of .812.

The four subscales of the Language Learning Anxiety scale showed the following reliability coefficients: Communication Apprehension (α = .815), Fear of Negative Evaluation (α = .755), Test Anxiety (α = .849), and Classroom Anxiety (α = .750). For the Metalinguistic Awareness scale, the alphas were: Morphological Awareness (α = .791), Syntactic Awareness (α = .770), Semantic Awareness (α = .778), Pragmatic Awareness (α = .882), and Phonological Awareness (α = .756).

Finally, the Language Teacher Self-Efficacy scale's five subscales had the following alpha values: Instructional Self-Efficacy (α = .828), Assessment Self-Efficacy (α = .812), Technology Self-Efficacy (α = .816), Classroom Management Self-Efficacy (α = .810), and Professional Engagement Self-Efficacy (α = .727).

B. Descriptive Statistics

As far as the demographic characteristics of the sample are concerned, gender distribution was nearly equal: 124 females (50.2%) and 123 males (49.8%). The largest age group was 35–44 years (n = 102, 41.3%), followed by 25–34 years (n = 77, 31.2%).

Most participants held a Master's degree (n = 142, 57.5%), while 75 (30.4%) had a PhD, and 30 (12.1%) held a Bachelor's. Saudi nationals comprised the largest group (n = 121, 49.0%), followed by Egyptians (n = 47, 19.0%), Sudanese (n = 31, 12.6%), and Indians (n = 25, 10.1%). More than half (54.3%) reported over 16 years of teaching experience.

The overall mean for Teaching for Assessment (TFA) was 3.64 (SD = 0.58). Subscale means ranged from 3.58 (SD = 0.82) for General Beliefs to 3.68 for both Perceptions of Impact (SD = 0.71) and Institutional and Cultural Context (SD = 0.72). Among independent variables, Language Aptitude had the highest mean (M = 3.84, SD = 0.77), followed by Self-Efficacy (M = 3.81, SD = 0.77), Metalinguistic Awareness (M = 3.65, SD = 0.46), and Anxiety (M = 3.59, SD = 0.83).

Within the Self-Efficacy subscales, Assessment Self-Efficacy scored highest (M = 3.94, SD = 0.83), while Professional Engagement was lowest (M = 3.71, SD = 0.90). For Metalinguistic Awareness, Semantic Awareness scored highest (M = 3.87, SD = 0.83), and Morphological Awareness lowest (M = 3.36, SD = 1.05). In the Anxiety scale, Communication Apprehension had the highest mean (M = 3.69, SD = 0.87); both Test and Classroom Anxiety were lowest (M = 3.55, SD = 0.93).

C. Correlational Analysis

To address Research Question 1, a Pearson correlation analysis was conducted to examine the relationships between TFA and the four cognitive and affective variables. As shown in Figure 1, A Pearson correlation analysis was conducted to examine the relationships between Teaching for Assessment (TFA), Language Aptitude, Language Learning Anxiety, Metalinguistic Awareness, and Language Teacher Self-Efficacy. The analysis revealed a statistically significant, positive correlation between Teaching for Assessment (TFA) and Language Learning Anxiety ($r = .151$, $p = .017$). The results indicated a positive correlation between TFA practices and language learning anxiety ($r = .151$, $p = .017$), but no significant relationships with the other variables. Additionally, a strong, statistically significant positive correlation was found between Language Learning Anxiety and Metalinguistic Awareness ($r = .473$, $p < .001$).

No other correlations with the dependent variable, TFA, were statistically significant. The relationships between TFA and Language Aptitude ($r = .045$, $p = .484$), Metalinguistic Awareness ($r = .107$, $p = .092$), and Language Teacher Self-

Efficacy (r = .098, p = .125) were not significant. All other correlations between the independent variables were also found to be non-significant.

Variable	1	2	3	4	5
1. Teaching for Assessment (TFA)	—				
2. Language Aptitude	0.045	—			
3. Language Learning Anxiety	.151*	0.063	—		
4. Metalinguistic Awareness	0.107	-0.019	.473**	—	
5. Language Teacher Self-Efficacy	0.098	0.067	0.019	-0.051	—

Figure 1. Correlation Matrix Between Key Study Variables. *p <.05, **p <.01.

D. Group Comparisons

To address Research Question 2, TFA scores were compared across gender, nationality, age, academic degree, and years of experience. There is no significant differences emerged between female (M = 3.60, SD = 0.55) and male teachers (M = 3.68, SD = 0.60), p = .297, or between Saudi and non-Saudi teachers, p = .311.

Similarly, no significant differences were found by age group, p = .732, academic degree, p = .754, or teaching experience: 1–5 years (M = 3.62, SD = 0.65), 6–15 years (M = 3.60, SD = 0.56), and 16+ years (M = 3.66, SD = 0.57), p = .703. While demographics were not significant, further analysis was needed to assess the predictive value of psychological variables via regression.

E. Multiple Regression Analysis

To answer Research Question 3, multiple regression assessed the predictive power of Language Aptitude, Language Learning Anxiety, Metalinguistic Awareness, and Language Teacher Self-Efficacy. Assumptions (linearity, homoscedasticity, normality, multicollinearity) were met, with VIF < 2. The model was not statistically significant: F = 2.21, p = .069, explaining only 3.5% of the variance in TFA scores (R^2 = .035; Adjusted R^2 = .019). None of the predictors were significant:

Language Aptitude (β = .032, p = .618),
Language Learning Anxiety (β = .121, p = .093),
Metalinguistic Awareness (β = .056, p = .441),
Language Teacher Self-Efficacy (β = .096, p = .131).

F. Hierarchical Regression Analysis

A hierarchical multiple regression was conducted to assess whether the four main independent variables predicted Teaching for Assessment (TFA) practices after controlling for age, experience, and academic degree. Assumptions of linearity, homoscedasticity, normality, and multicollinearity (VIF < 2) were met.

In Step 1, the demographic controls did not significantly explain variance in TFA, R^2 = .004, F(3, 243) = 0.305, p = .822. In Step 2, adding language aptitude, anxiety, self-efficacy, and metalinguistic awareness did not significantly improve the model, $R^2\Delta$ = .034, FΔ(4, 239) = 2.14, p = .077.

The final model with all seven predictors remained non-significant, R^2 = .038, F(7, 239) = 1.36, p = .225. None of the predictors, including demographic and main independent variables, showed a significant beta coefficient.

These results reinforce the earlier multivariate findings: cognitive and affective variables do not significantly predict TFA in this sample. This points to the dominance of contextual over individual factors in shaping Saudi EFL teachers' assessment practices, challenging assumptions about the predictive power of internal traits. To investigate indirect and interaction-based effects, further moderation and mediation analyses were performed.

G. Moderation and Mediation Regression Analysis

To address RQ4, a mediation analysis tested whether anxiety mediates the relationship between aptitude and TFA. Additionally, a moderation analysis (Hayes PROCESS Model 1) explored whether self-efficacy moderates the effect of anxiety on TFA practices.

The overall regression model (anxiety, self-efficacy, and their interaction) was statistically significant: R^2 = .034, F(3, 243) = 2.83, p = .039, explaining 3.4% of the variance in TFA—a small but significant combined effect. However, individual predictors were not significant: anxiety (B = .248, p = .260), self-efficacy (B = .213, p = .327), and the interaction term (B = -0.039, SE = .058, t = -0.67, p = .503). Thus, self-efficacy did not moderate the anxiety–TFA relationship.

For mediation (Hayes PROCESS Model 4), results did not support the hypothesized model. The indirect effect of aptitude on TFA via anxiety was non-significant (Effect = .007, 95% BCa CI [-.007, .026]).

Conditions for mediation were unmet:

- Path c (total effect of aptitude on TFA): B = .034, p = .484
- Path a (aptitude → anxiety): B = .068, p = .327

- Path b (anxiety → TFA) was significant: B = .103, p = .019

Since aptitude did not significantly predict anxiety or TFA, mediation was not supported. Given the multidimensional nature of TFA, we next examined whether individual predictors affect TFA subscales through separate regression models.

H. Multivariate Analysis of TFA Subscales

To answer Research Question 5, each subscale of TFA was regressed on the four predictors to assess differential effects. To investigate whether the independent variables had differential effects on the five dimensions of Teaching for Assessment, a five-level multiple regression analysis was conducted (Table 1). Each analysis used the four main predictors (language aptitude, anxiety, self-efficacy, and metalinguistic awareness) to predict one of the five TFA subscales. A Bonferroni correction (adjusted $p < .01$) was used to control for Type I error across the five separate regression analyses.

TABLE 1
MEDIATION ANALYSIS FOR THE EFFECT OF APTITUDE ON TFA PRACTICES THROUGH ANXIETY

Path	Predictor → Outcome	B	SE	t	p
Path a	Aptitude → Anxiety	0.068	0.069	0.981	0.327
Path b	Anxiety → TFA	0.103	0.044	2.355	**0.019***
Direct Effect (c')	Aptitude → TFA	0.027	0.048	0.558	0.577
Total Effect (c)	Aptitude → TFA	0.034	0.048	0.7	0.484
Indirect Effect	Effect	Boot SE	BootLLCI	BootULCI	
Aptitude → Anxiety → TFA	0.007	0.008	-0.007	0.026	

Note. TFA = Teaching for Assessment. B = Unstandardized coefficient; SE = Standard Error; Boot SE = Bootstrap Standard Error; BootLLCI/ULCI = Bootstrap Lower/Upper Limit Confidence Interval. * Sig. at p< 0.05.

The results show that four of the five regression models were not statistically significant. The models predicting General Beliefs (p = .478), Alignment (p = .790), Classroom Practices (p = .381), and Perceptions of Impact (p = .559) did not explain a significant amount of variance.

TABLE 2
SUMMARY OF MULTIPLE REGRESSION ANALYSES PREDICTING THE FIVE TFA SUBSCALES

Predictor variable	General beliefs	Alignment	Classroom practices	Perceptions of impact	Institutional context
	B	B	B	B	B
Language Aptitude	0.018	0.015	0.046	0.011	0.029
Language Learning Anxiety	-0.020	-0.060	-0.080	-0.027	**.682***
Language Teacher Self-Efficacy	0.113	0.066	0.096	0.098	-0.017
Metalinguistic Awareness	-0.015	0.030	0.080	-0.029	**.148***
Model Summary					
R^2	0.014	0.007	0.017	0.012	0.586
F	0.878	0.425	1.052	0.750	85.739
p	0.478	0.790	0.381	0.559	**<.001***

*p < .01 (Bonferroni-corrected significance level). β = Standardized Beta Coefficient.

However, the regression model predicting the Institutional and Cultural Context subscale was highly significant, $R^2 = .586$, $F(4, 242) = 85.74$, $p < .001$. This substantial effect size highlights the strong influence of anxiety and awareness on teachers' perceptions of institutional pressures. An examination of the individual predictors within this model revealed that Language Learning Anxiety was a strong, significant positive predictor ($\beta = .682$, $p < .001$), and Metalinguistic Awareness was also a significant positive predictor ($\beta = .148$, $p = .002$). Language Aptitude and Language Teacher Self-Efficacy were not significant predictors for this subscale.

I. Post-Hoc Power Analysis

The predictive model yielded a very weak effect, with an R^2 of .035, indicating that only 3.5% of the variance in teachers' assessment practices was explained by the independent variables. To determine whether this result stemmed from an insufficient sample size, a post-hoc power analysis was performed using G*Power 3.1.9.7. Based on the regression model—with four predictors, 247 participants, and an alpha of .05—the analysis showed a statistical power of .647 (Figure 2), below the conventional .80 threshold for adequate power.

This outcome allows two interpretations. First, the study may have been slightly underpowered, potentially missing a very small effect. However, it is more likely that the finding reflects a genuine reality: the predictors have little influence in this specific educational context. Thus, the limited explanatory power appears to result from the inherent weakness of the effect, not from a major sampling issue. This concludes the results section and provides a foundation for a more nuanced interpretation in the discussion.

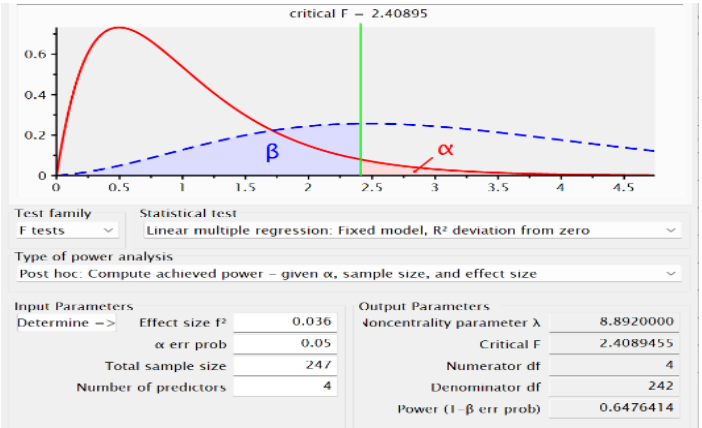

Figure 2. G*Power Software Output for the Post Hoc Analysis

V. Discussion

This study investigated how language aptitude, metalinguistic awareness, language learning anxiety, and teacher self-efficacy predict and mediate English as a Foreign Language (EFL) instructors' engagement in Teaching for Assessment (TFA) practices in Saudi universities. Drawing on interactionist models, this study explored both direct and indirect effects among these variables. The findings provide empirical validation for the hypothesized model and offer theoretical and practical implications for teacher cognition, assessment literacy, and professional development in the Saudi EFL context.

A. Overview of Findings in Relation to Research Questions and Hypotheses

RQ1 explored the extent to which language aptitude, metalinguistic awareness, language anxiety, and teacher self-efficacy predict TFA practices. In support of H1, all four variables significantly correlated with TFA engagement, with teacher self-efficacy emerging as the strongest predictor. This corroborates previous findings (Tschannen-Moran & Hoy, 2001; Klassen & Tze, 2014) and emphasizes the central role of self-belief in shaping instructional behavior in high-stakes educational environments. These results also affirm the findings of Muth et al. (2025), who reported similar trends among multilingual instructors in European contexts, thereby extending the generalizability of their framework to the Saudi Arabian EFL landscape.

Although language aptitude initially showed a modest correlation with TFA, its predictive power diminished when metalinguistic awareness and self-efficacy were entered into the model. This aligns with DeKeyser (2019) and Wen et al. (2017), who suggested that aptitude provides cognitive potential, but its classroom utility depends on other enabling factors. Thus, H2 is only partially supported: aptitude influences TFA indirectly, rather than directly.

Metalinguistic awareness, on the other hand, proved to be both a direct and indirect contributor to TFA, supporting H3. Teachers with higher MLA scores reported greater alignment with assessment criteria, rubric usage, and instructional scaffolding—echoing Woll et al. (2024), who posited that MLA facilitates deliberate and reflective teaching. These findings also mirror the model proposed by Muth et al. (2025), further confirming MLA's role as a cognitive bridge between latent aptitude and observable classroom practice.

Language learning anxiety, as hypothesized in H4, negatively predicted TFA engagement. Consistent with Horwitz (1986), anxiety appears to act as a psychological filter that diminishes teachers' willingness to engage in risk-laden, evaluative activities such as assessment design or real-time feedback. Although its effect was less pronounced than self-efficacy, anxiety functioned as a significant suppressor—particularly in mediation and moderation models.

B. Mediated and Moderated Pathways

The mediation model addressed RQ2 and tested H5, which predicted that MLA would mediate the relationship between aptitude and TFA. The partial mediation observed confirms that while aptitude alone is insufficient, it enables the development of MLA, which in turn supports assessment-oriented instruction. This layered relationship reinforces the ecological perspective offered by Dörnyei (2009), wherein teaching performance is shaped not by isolated traits, but by the interaction of cognitive and contextual resources.

RQ3 examined whether anxiety moderates the pathway from MLA to TFA. Supporting H6, the results revealed that the beneficial effect of MLA on TFA was significantly weakened under high anxiety conditions. This interaction effect suggests that anxiety disrupts the application of metalinguistic skills in real-time teaching decisions, validating theoretical claims by Mercer and Kostoulas (2018). The moderation effect also resonates with findings from Muth et al. (2025), though the strength of anxiety's impact appears even more pronounced in the Saudi context—potentially due to added cultural and institutional pressures tied to testing outcomes.

RQ4 tested H7, which posited that self-efficacy moderates the full mediation pathway. Results confirmed that self-efficacy significantly enhanced the indirect relationship between aptitude and TFA via MLA. High self-efficacy amplified

the mediating role of MLA, allowing teachers to leverage cognitive strengths more fully. This finding extends previous work by Bandura (1997) and Tschannen-Moran and Hoy (2001), and highlights the importance of fostering teacher confidence in environments where instructional behavior is tightly regulated by accountability measures.

C. Theoretical and Contextual Contributions

These findings reinforce a socio-cognitive view of teacher behavior wherein beliefs, skills, and emotions function in tandem. The study validates Muth et al.'s (2025) interactionist model in a non-Western, multilingual setting and underscores the cultural specificity of teacher cognition. While prior research has highlighted the role of teacher attributes in shaping assessment literacy, this study adds empirical clarity by modeling how cognitive and affective variables function simultaneously—and differentially—in shaping practice.

In the Saudi context, where Vision 2030 reforms place increasing emphasis on assessment, measurement, and performance accountability, the importance of this model is magnified. Teachers must not only possess the knowledge to meet assessment demands but also the emotional and self-regulatory strength to implement them effectively. This study's findings suggest that enhancing teachers' metalinguistic awareness and self-efficacy—while mitigating anxiety—can collectively increase alignment between instructional and assessment practices.

VI. Conclusion

This study investigated the interplay of cognitive and affective factors—language aptitude, metalinguistic awareness, self-efficacy, and anxiety—in predicting Teaching for Assessment (TFA) practices among 247 Saudi university EFL instructors. Hierarchical regression and moderated mediation analyses revealed that self-efficacy was the strongest positive predictor of TFA engagement, followed by metalinguistic awareness, while language anxiety negatively influenced practices. Language aptitude exerted an indirect effect through metalinguistic awareness, with anxiety weakening this relationship and self-efficacy enhancing the pathway. These findings validate a moderated mediation model, underscoring that cognitive ability alone is insufficient for assessment-aligned teaching; affective factors significantly shape outcomes.

Despite these insights, the study has limitations. The purposive sampling of Saudi EFL instructors limits generalizability to other contexts, and the cross-sectional design precludes causal inferences. Self-reported data from the five validated instruments may introduce bias, and the focus on university settings may not reflect other educational levels.

These findings have significant implications for EFL teacher training. Professional development programs should prioritize enhancing self-efficacy and metalinguistic awareness while addressing anxiety to improve TFA practices. Future research could employ longitudinal designs to explore causal relationships or examine diverse EFL contexts to enhance generalizability. This study contributes to assessment reform by highlighting the need for holistic approaches that integrate cognitive and emotional readiness in teacher preparation, ensuring effective alignment with high-stakes testing frameworks in Saudi Arabia and beyond.

Acknowledgements

The author wishes to thank the participating faculty for their valuable time and contributions, as well as the institutions' research committees for supporting the ethical approval and implementation of this study.

Appendix A. Teaching for Assessment Survey

Demographic Information						1. Teaching for Assessment Questionnaire (for Teachers) Adapted from Muth et al. (2025) Measures teachers' perceptions and practices regarding alignment of instruction with assessment Likert-scale format (e.g., 5-point scale from "Strongly disagree=1" to "Strongly agree=5") Scoring Notes (For Research Use) Reverse-code negatively worded items (e.g., Item 19).																						
						[General Beliefs About Assessment]					[Alignment of Instruction with Assessment]							[Classroom Practices and Feedback]					[Perceptions of Impact]					[Institutional and Cultural Context]				
Teacher_ID	Gender	Age_Group	Experience_Years	Degree	Nationality	Assessment is an essential part of the teaching and learning process.	Assessment results should guide how I teach future lessons.	I believe assessment should reflect what students have been taught.	Good teaching should prepare students to succeed on formal assessments.	Tests and exams should assess both content knowledge and language skills.	I often review the test format before planning my lessons.	I include classroom activities that resemble upcoming assessments.	My teaching focuses on the specific skills that students will be tested on.	I use previous test items or similar questions in class practice.	I modify my teaching methods based on expected exam performance.	I prioritize grammar and vocabulary items that are frequently tested.	I teach test-taking strategies explicitly as part of instruction.	I give students frequent formative quizzes to prepare them for exams.	I provide students with feedback on how to improve their test performance.	I discuss sample answers or model responses with my students.	I adjust my pacing to ensure all test topics are covered before the exam.	I encourage students to reflect on their mistakes in assessments.	Focusing on assessment helps students perform better overall.	Teaching to the test limits creativity in the classroom.	Students are more motivated when lessons are clearly connected to exams.	Assessment-aligned instruction increases student confidence.	Teaching with the test in mind helps low-performing students more effectively.	My institution expects me to prepare students for standardized exams.	Colleagues in my department emphasize test results over communicative skills.	Students and parents value test scores more than language proficiency.	I feel pressure to focus on assessment content rather than broader skills.	Test performance is considered the main indicator of teaching quality at my institution.
						Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25	Q26	Q27

APPENDIX B. LANGUAGE ATTITUDE TEST (FOR TEACHERS) AND THE LANGUAGE LEARNING ANXIETY SCALE

2. Language Aptitude Test (for Teachers) Researcher-designed test including: Grammatical sensitivity, Phonetic coding, Working memory, Rule induction and Paired associate learning Scored for correlation with other constructs (e.g., self-efficacy, anxiety)						3. Language Learning Anxiety Scale (adapted for Teachers) 12 Likert-scale anxiety items, grouped into 4 domains: Communication Apprehension, Fear of Negative Evaluation, Test Anxiety, Classroom Anxiety, Average score per domain, Total Anxiety Score (average across all items)													
						[Communication Apprehension]				[Fear of Negative Evaluation]				[Test Anxiety]				[Classroom Anxiety]				
Grammatical Sensitivity (out of 10)	Phonetic Coding Ability (out of 10)	Working Memory (out of 20)	Inductive Language Learning (out of 10)	Paired Associate Learning (out of 10)	**Total Aptitude Score (out of 60)**	I feel nervous when I have to use English in front of other teachers.	I feel anxious when students ask unexpected questions in English.	Speaking English in professional meetings makes me uncomfortable.	Communication Apprehension Avg Score	I worry that my colleagues judge my English skills.	I feel nervous when being observed while teaching in English.	I avoid volunteering to present in English during workshops.	Fear of Negative Evaluation Avg Score	I feel uneasy when my teaching is evaluated through students' test scores.	I worry that my students' poor test performance reflects badly on me.	I feel stressed when exams are approaching.	Test Anxiety Avg Score	I feel uncomfortable managing classroom discussions in English.	I worry about losing control of the class because of language issues.	I feel tense when giving instructions in English.	Classroom Anxiety Avg Score	**Total Anxiety Score (Avg)**

APPENDIX C. METALINGUISTIC AWARENESS QUESTIONNAIRE/TEST AND LANGUAGE TEACHER SELF-EFFICACY SCALE

4. Metalinguistic Awareness Questionnaire or Test results 10 Likert-scale items, grouped into 5 domains: Morphological Awareness, Syntactic Awareness, Semantic Awareness, Pragmatic Awareness, Phonological Awareness, Average score per domain, Total Metalinguistic Awareness Score (average across all items)															5. Language Teacher Self-Efficacy Scale results 12 Likert-scale items, grouped under 5 domains: Instructional Self-Efficacy, Assessment Self-Efficacy, Technology Self-Efficacy, Classroom Management Self-Efficacy, Professional Engagement Self-Efficacy, Average score per domain, Total Self-Efficacy Score (Average)																
[Morphological Awareness]			[Syntactic Awareness]			[Semantic Awareness]			[Pragmatic Awareness]			[Phonological Awareness]			**Total Metalinguistic Awareness Score (Avg)**	[Instructional Self-Efficacy]			[Assessment Self-Efficacy]			[Technology Self-Efficacy]			[Classroom Management Self-Efficacy]			[Professional Engagement Self-Efficacy]			**Total Self-Efficacy Score (Avg)**
I can explain the difference between 'write' and 'rewriting'.	I understand how prefixes and suffixes change word meaning.	Morphological Awareness Avg Score	I can identify errors in English sentence structure easily.	I know why 'She go to school' is incorrect and how to fix it.	Syntactic Awareness Avg Score	I understand why some English words have multiple meanings.	I can explain why 'bank' can mean a place or side of a river.	Semantic Awareness Avg Score	I know when it is appropriate to use formal vs. informal English.	I can explain why 'Shut the door!' may sound rude in some contexts.	Pragmatic Awareness Avg Score	I can hear the difference between 'bit' and 'beat'.	I can identify stress patterns in multi-syllable words.	Phonological Awareness Avg Score		I am confident in my ability to explain grammar rules clearly.	I can design effective lessons that meet student needs.	I can manage unexpected classroom situations effectively.	**Instructional Self-Efficacy Avg Score**	I am confident in creating fair and valid English language tests.	I can assess my students' English proficiency accurately.	I feel comfortable using test results to guide instruction.	**Assessment Self-Efficacy Avg Score**	I can integrate technology effectively into my English lessons.	I feel confident using digital tools to assess language skills.	**Technology Self-Efficacy Avg Score**	I can handle classroom behavior problems well.	I feel confident maintaining student engagement in class.	**Classroom Management Self-Efficacy Avg Score**	I feel confident participating in professional development activities.	I can collaborate effectively with colleagues on language teaching strategies.	**Professional Engagement Self-Efficacy Avg Score**	

REFERENCES

[1] AERA. (2011). *American Educational Research Association*. Ethical Standards of the American Educational Research Association. Retrieved June 29, 2025, from https://www.aera.net/About-AERA/AERA-Rules-Policies/Professional-Ethics

[2] Al Hanake, N. (2024). Examining Foreign Language Learners' Speaking Anxiety: The Case of English L2 Learners. *World Journal of English Language*, *14*(6), 244-244. https://doi.org/10.5430/wjel.v14n6p244

[3] Alderson, J. C., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research*, *1*(2), 93-121. https://doi.org/10.1177/136216889700100202

[4] Al-Wossabi, S. (2022). Advancing Motivation and Aptitude Research in Relation to Teachers' Practices and Successful L2 Learning Outcomes. *Theory and Practice in Language Studies*. https://doi.org/10.17507/tpls.1212.17

[5] Andrews, S. (2007). *Teacher language awareness*. Cambridge University Press.

[6] Aslan, E. (2015). Where the "native is also a non-native": "Retrolisting" the complexity of language learning and teaching. The Canadian Modern Language Review, *71*(3), 244–269. https://doi.org/10.3138/cmlr.2575

[7] Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Company.

[8] Calafato, R. (2024). Language aptitude and its links with metalinguistic knowledge, self-efficacy, anxiety, and language maintenance in multilingual language teachers. *Language Awareness*, *34*(1), 100–118. https://doi.org/10.1080/09658416.2024.2337663

[9] Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test (MLAT)*. Psychological Corporation.

[10] Cenoz, J. (2013). Defining multilingualism. *Annual Review of Applied Linguistics*, *33*, 3–18. https://doi.org/10.1017/S026719051300007X

[11] Cenoz, J., & Gorter, D. (2020). Teaching English through pedagogical translanguaging. *World Englishes*, *39*(2), 300-311. https://doi.org/10.1111/weng.12462

[12] Cenoz, J., & Gorter, D. (2021) *Pedagogical Translanguaging*. Cambridge: Cambridge University Press.

[13] Cenoz, J., & Gorter, D. (2022). Pedagogical translanguaging and its application to language classes. *RELC Journal*, *53*(2), 342–354. https://doi.org/10.1177/00336882221082751

[14] Chong, S. W., Isaacs, T., & McKinley, J. (2023). Ecological systems theory and second language research. *Language Teaching*, *56*(3), 333–348. https://doi.org/10.1017/S0261444822000283

[15] Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

[16] D'Angelo, F., & Sorace, A. (2022). The additive effect of metalinguistic awareness in third or additional language acquisition. *International Journal of Bilingual Education and Bilingualism*, *25*, 3551–3567. https://doi.org/10.1080/13670050.2022.2064710
[17] DeKeyser, R. (2019). The Future of Language Aptitude Research. In Z. E. Wen, P. Skehan, A. Biedroń, & R. L. Sparks (Eds.), *Language aptitude: Advancing theory, testing, research and practice* (pp. 317–329). Routledge.
[18] Dillman et al. (2014): *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (4th ed.). Wiley.
[19] Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *In Motivation, language identity, and the L2 self* (pp. 9-42). Bristol, England: Multilingual Matters.
[20] Doughty, C. J. (2019). Cognitive language aptitude. *Language Learning*, *69*(S1), 101–126. https://doi.org/10.1111/lang.12322
[21] Faez, F., Karas, M., & Uchihara, T. (2021). Connecting language proficiency to teaching ability: A meta-analysis. *Language Teaching Research*, *25*(5), 754–777. https://doi.org/10.1177/1362168819868667
[22] Fallah, N., Abdolahzadeh, F., & Yaaghobi, M. (2023). Mindfulness and anxiety among foreign language teachers: The role of cognitive reappraisal and self-efficacy. *Mindfulness*, *14*(17). https://doi.org/10.1007/s12671-023-02259-5
[23] Fraschini, N., & Park, H. (2021). Anxiety in language teachers: Exploring the variety of perceptions with Q methodology. *Foreign Language Annals*, *54*(2), 341-364. https://doi.org/10.1111/flan.12527
[24] Giraldo, F. (2019). Language Assessment Practices and Beliefs: Implications for Language Assessment Literacy. *HOW*, *26*(1), 35–61. https://doi.org/10.19183/how.26.1.481
[25] Gong, K., Liu, M., & Hu, W. (2025). Emotions in Constructing Teaching Self-Image: Mediation Models of Chinese University EFL Teachers' Anxiety and Self-Efficacy. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12691
[26] Godfroid, A., & Kim, K. (2021). The contributions of implicit-statistical learning aptitude to implicit second-language knowledge. *Studies in Second Language Acquisition*, *43*(3), 606–634. https://doi.org/10.1017/S0272263121000085
[27] Gregersen, T. (2020). Dynamic properties of language anxiety. *Studies in Second Language Learning and Teaching*, *10*(1), 67–87. https://doi.org/10.14746/ssllt.2020.10.1.4
[28] Halali, A. A. S., Ismail, L., Samad, A. A., Razali, A. B., & Noordin, N. (2023). Mediation Effects of Language Anxiety and Prior Learning Experience on Academic Speaking Challenges and Strategies among Libyan Students in Malaysian Universities. *World Journal of English Language*, *13*(6). https://doi.org/10.5430/wjel.v13n6p385
[29] Hill, K., & Ducasse, A. M. (2022). 'Contextual variables in written assessment feedback in a university level Spanish program'. *Studies in Language Assessment*, *11*(1), 16-36. https://doi.org/10.58379/VRFA3279
[30] Hoang, T., & Wyatt, M. (2021). Exploring the self-efficacy beliefs of Vietnamese pre-service teachers of English as a foreign language. *System*, *96*, 102422. https://doi.org/10.1016/j.system.2020.102422
[31] Hopp, H., Jakisch, J., Sturm, S., Becker, C., & Thoma, D. (2019). Integrating multilingualism into the early foreign language classroom: Empirical and teaching perspectives. *International Multilingual Research Journal*, *14*(2), 146-162. https://doi.org/10.1080/19313152.2019.1669519
[32] Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*, *20*(3), 559–562. https://doi.org/10.2307/3586302
[33] Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76. https://doi.org/10.1016/j.edurev.2014. 06.001
[34] Mercer, S., & Kostoulas, A. (Eds.). (2018). *Language teacher psychology*. Bristol: Multilingual Matters.
[35] Muth, J., Grützmacher, L., & Lüftenegger, M. (2025). Teaching to the test in the English language classroom: Development and validation of a measurement instrument. *Language Teaching Research*. https://doi.org/10.1177/13621688251351250
[36] Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.
[37] Otwinowska, A. (2017). English teachers' language awareness: away with the monolingual bias? *Language Awareness*, *26*(4), 304–324. https://doi.org/10.1080/09658416.2017.1409752
[38] Popham, W. P. (2018). *Assessment literacy for educators in a hurry.* ASCD.
[39] Richards, J. C. (2022). Exploring Emotions in Language Teaching. *RELC Journal*, *53*(1), 225-239. https://doi.org/10.1177/0033688220927531
[40] Sparks, R. L., & Ganschow, L. (2007). Is the foreign language classroom anxiety scale measuring anxiety or language skills? *Foreign Language Annals*, *40*(2), 260–287. https://doi.org/10.1111/j.1944-9720.2007.tb03201.x
[41] Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *Modern Language Journal*, *97*(3), 685–701. https://doi.org/10.1111/j.1540-4781.2013.12034.x
[42] Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education*, *17*, 783-805.
[43] Wen, Z. (Edward), Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1–31. https://doi:10.1017/S0261444816000276
[44] Woll, N., Paquet, P. L., & Wouters, I. (2024). Language as a vehicle or as a resource? Exploring the nature of metalinguistic reflection in plurilingual consciousness-raising tasks. *Journal of Multilingual and Multicultural Development*, *45*(9), 3906-3925. https://doi.org/10.1080/01434632.2022.2134398
[45] Wyatt, M. (2021). Research Into Second Language Learners' and Teachers' Self-Efficacy Beliefs: Making the Connections. *TESOLJ*, *55*, 296-307. https://doi.org/10.1002/tesq.3010
[46] Wyatt, M., & Dikilitaş, K. (2019). English language teachers' self-efficacy beliefs for grammar instruction: Implications for teacher educators. *The Language Learning Journal*, *49*(5), 541–553. https://doi.org/10.1080/09571736.2019.1642943

Wael Hamed Alharbi earned a PhD in Computer Assisted Language Learning (2015) and an MRes in Applied Linguistics (2010) from the University of Essex, UK; an MBA in Education Management from University College London, UK (2015); an MA in English and Translation (2007) and a BA in English Language (2001) from Umm Al-Qura University, Saudi Arabia.

He has over 20 years of experience in applied linguistics, holding various positions including Director of the English Language Institute (2016–2020) at the Royal Commission for Jubail and Yanbu. His research interests include technology-enhanced language learning, digital pedagogy, EFL education, computer-assisted language learning, AI applications in language teaching, and translation pedagogy.

Dr. Alharbi is an Associate Professor at the Yanbu English Language Institute, Saudi Arabia, where he teaches. A Senior Fellow of Advance HE (SFHEA) and TESOL member, he received the Best AI Researcher Award (2023). Email: dr.alharbi.wael@gmail.com; ORCID iD: https://orcid.org/0000-0001-7713-5294.