



IMAM KHOMEINI
INTERNATIONAL UNIVERSITY



Print ISSN: 2676-5587
Online ISSN: 2676-5985

Developing and Validating an Online Diagnostic Test of L2 Pragmatic Competence: A Meaning-Oriented Model

Malihe Ketabdar¹, Mohammad Hashamdar^{2*}, Gholamhassan Famil Khalili³

¹Ph.D. Candidate, Department of Teaching English and Translation, Ka.C., Islamic Azad University, Karaj, Iran. malihe.ketabdar@iau.ac.ir

^{2*} (Corresponding author) Assistant Professor, Department of Teaching English and Translation, Ka.C., Islamic Azad University, Karaj, Iran.

mohammad.hashamdar@iau.ac.ir

³Assistant Professor, Department of Teaching English and Translation, Ka.C., Islamic Azad University, Karaj, Iran. khalili@kiaa.ac.ir

Article info

Abstract

Article type:
Research article

Received:
2025/10/06

Accepted:
2025/12/30

Research on second language (L2) pragmatics has predominantly focused on speech acts, leaving broader dimensions of pragmatic competence relatively unexplored. To address this gap, the present study developed and validated an Online Diagnostic Test of L2 Pragmatic Competence (DTPC) informed by Purpura's (2017) meaning-oriented model. Using a mixed-methods design, the researchers developed DTPC through several stages, including literature and expert review, a written discourse completion task (DCT), test construction, piloting, final administration, and follow-up interviews. Initially, a total of 369 Iranian EFL learners were selected through purposive sampling to participate in the study. Following the proficiency test, the researchers excluded 11 advanced learners to ensure sample homogeneity, and data from 358 participants were retained for the final statistical analyses. After item screening and refinement, 34 of the 75 DTPC items were retained based on the results of item analysis, exploratory factor analysis, and confirmatory factor analysis. The analyses confirmed a four-factor structure comprising sociolinguistic, sociocultural, psychological, and situational meanings, with high reliability and acceptable model fit. Sociolinguistic and psychological meanings showed the most decisive influence on learners' pragmatic competence. Thematic analysis of interviews indicated learners' positive perceptions of the DTPC's diagnostic value and online delivery. The findings provide empirical validation of Purpura's meaning-oriented model in the Iranian EFL context. In this context, the DTPC provides diagnostic insights into EFL learners' pragmatic challenges, thereby supporting more targeted and responsive instruction.

Keywords: meaning-oriented model, online diagnostic test of pragmatic competence (DTPC), pragmatic meaning, second language pragmatics

Cite this article: Ketabdar, M., Hashamdar, M., & Famil Khalili, G. (2026). Developing and validating an online diagnostic test of L2 pragmatic competence: A meaning-oriented model. *Journal of Modern Research in English Language Studies*, 13(2), 125-152.

DOI: [10.30479/jmrels.2025.21965.2514](https://doi.org/10.30479/jmrels.2025.21965.2514)

©2026 by the authors. Published by Imam Khomeini International University.

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0) <https://creativecommons.org/licenses/by/4.0>



1. Introduction

Pragmatic competence, defined as the ability to use language appropriately in social contexts, is a cornerstone of second language (L2) proficiency (Taguchi & Roever, 2020). Although research on assessing L2 pragmatics has expanded considerably since the 1990s, high-stakes proficiency tests have been indirect and limited due to practical constraints, insufficient pragmatics test literacy among test-takers and developers, and challenges associated with testing pragmatics (Bardovi-Harlig, 2020; Roever, 2024; Timpe-Laughlin & Choi, 2017). Therefore, EFL learners who achieve high test scores still face pragmatic difficulties in real-world interactions (Ikeda, 2021). Despite the emergence of functional approaches to language testing since the 1980s, L2 pragmatics, particularly intercultural competence, has remained comparatively underrepresented in mainstream assessment research (Kecskes, 2014; Mootoosamy & Aryadoust, 2024).

There is extensive research on L2 pragmatics and its assessment, developed through two main traditions: one grounded in speech act and politeness theory, and another informed by conversation analysis (Roever, 2024). Many empirical studies in the Iranian EFL context have focused on specific speech acts (Derakhshan et al., 2021, 2024). Little is known about broader pragmatic dimensions (Beltrán, 2019; Laughlin et al., 2015).

Although Iranian EFL learners now have greater access to authentic pragmatic input via digital platforms, pragmatic instruction and materials remain underdeveloped in classroom practices (Savadkouhi & Mostafaei Alaei, 2022). Addressing practical barriers such as large class sizes, insufficient teacher training, and exam-oriented curricula (Bardovi-Harlig, 2020; Derakhshan & Shakki, 2023), the integration of explicit, authentic input-based instruction is essential for fostering L2 pragmatic development (Shakki et al., 2021).

To fill this gap, a valid and theory-driven diagnostic assessment is needed to identify EFL learners' pragmatic strengths and weaknesses, and inform pragmatic targeted instruction. While pragmatic assessment instruments have been validated in ESL contexts (Grabowski, 2013; Roever, 2006), a critical void remains in the empirical validation of comprehensive frameworks like Purpura's (2017) model in the Iranian EFL context. In this regard, the current study seeks to develop and validate an Online Diagnostic Test of L2 Pragmatic Competence (DTPC) based on Purpura's (2017) meaning-oriented model, which conceptualizes pragmatic competence across multiple pragmatic meanings. Accordingly, four research questions framed the scope of the study:

1. What are the underlying components of the Online Diagnostic Test of L2 Pragmatic Competence (DTPC)?
2. What are the psychometric features of the DTPC?

3. To what extent does DTPC's structural model fit Purpura's (2017) meaning-oriented model?
4. How do Iranian EFL learners perceive and experience the DTPC?

2. Literature Review

Leech's (2014) dichotomy between pragmalinguistics (linguistic resources) and sociopragmatics (sociocultural norms) has established the foundation for pragmatic competence, highlighting the interdependence of language and context (Purpura, 2017). Building on Chomsky's linguistic competence, subsequent communicative competence models (Bachman & Palmer, 2010) incorporated sociocultural and strategic dimensions, emphasizing that pragmatic competence extends beyond grammar to the contextually appropriate language use (Taguchi & Roever, 2020). Recent research has highlighted the dynamic interplay between linguistic resources and authentic contextual appropriateness in shaping L2 pragmatic competence (Roever, 2024).

In L2 pragmatic assessment, authenticity refers to test tasks that simulate real-world communication aligned with the target-language use (TLU) domains (Purpura, 2017). Authentic tasks integrate natural language features such as connected speech and spoken grammar, and contextualized responses (O'Grady, 2023). The data sources range from naturally occurring discourse (corpus-based data) to controlled elicitation formats like multiple-choice discourse completion tasks (Leech, 2014). Pragmatic testing seeks to balance standardization and contextual realism.

L2 Pragmatic testing refers to standardized instruments such as written discourse completion tasks, multiple-choice discourse completion tests, oral discourse completion tasks, role-plays, and self-assessments, designed to elicit measurable responses (Ishihara & Cohen, 2021; Roever, 2024). However, pragmatic assessment represents a broader formative and summative evaluation of learners' pragmatic ability (Laughlin et al., 2015). While often used interchangeably, testing functions as a methodological tool within assessment (Laughlin et al., 2015; Taguchi & Roever, 2020).

2.1. Experimental Background

Early L2 pragmatic tests, such as Hudson et al.'s (1992) test battery, emphasized speech acts but overlooked the multifaceted nature of pragmatic competence (Roever, 2024). However, empirical research has progressively advanced the design and development of L2 pragmatic tests. Roever (2006) validated a 36-item web-based test in combined ESL and EFL contexts and found that speech act and implicature knowledge increased with proficiency, while routine knowledge depended on L2 exposure. Jianda (2007) developed a multiple-choice pragmatic test for Chinese EFL learners. Similarly, Birjandi

and Rezaei (2010) developed a Multiple-choice Discourse Completion Test (MDCT) on the speech acts of requests and apologies for Iranian EFL learners. Subsequent research has further adapted MDCT instruments for Iranian EFL contexts. However, these studies have typically examined a narrow range of speech acts, and psychometric evidence, such as construct validity across pragmatic dimensions, has been inadequate, despite its importance in cross-cultural pragmatic assessment (Arabmofrad & Mehdiabadi, 2022).

In the adult ESL context, Grabowski (2007, 2013) confirmed that reciprocal written tasks elicit more authentic, contextually grounded responses than traditional methods. Beltrán (2019) validated a meaning-based MDCT in an ESL context and showed that contextualized, discourse-grounded items effectively assess learners' pragmatic knowledge.

Contemporary studies integrate sociocognitive and affective dimensions into pragmatic test design. Sedaghatgoftar et al. (2019) validated a L2 pragmatics aptitude test (SLPAT) for Iranian EFL learners, confirming pragmatic aptitude as a unified construct.

In a systematic review, Tajeddin and Khanlarzadeh (2024) investigated individual differences in L2 pragmatics among EFL learners in study abroad contexts and emphasized the need for diverse research designs. Complementing this line of inquiry, recent studies have explored how technology has transformed pragmatic instruction and assessment through AI-driven chatbots (Saeedi & Soltani, 2025) and computerized dynamic assessment (Alavi et al., 2020). These innovations provide real-time feedback, adaptive scaffolding, and contextualized tasks.

Despite advances in L2 pragmatic test development, much of the existing literature has relied on instruments targeting a limited range of speech acts or isolated pragmatic functions. As a result, L2 pragmatic competence has frequently been operationalized in fragmented ways. Moreover, comprehensive, theory-driven assessments that systematically include multiple pragmatic meanings remain scarce, particularly in EFL contexts such as Iran. These limitations highlight the need for a more integrative, meaning-oriented model to guide L2 pragmatic assessment.

2.2. Theoretical Framework

Recent models of communicative competence emphasize the interplay between grammatical, sociocultural, socio-cognitive, and pragmatic dimensions (Laughlin et al., 2015; Mao & He, 2021; Purpura, 2017). However, this multidimensionality is not always reflected in pragmatic assessment practices. To address these limitations, the present study adopted Purpura's (2017) meaning-oriented model **of pragmatic competence**, which conceptualizes pragmatic competence as the integration of form, function, and appropriateness across seven pragmatic meanings. This model provides a

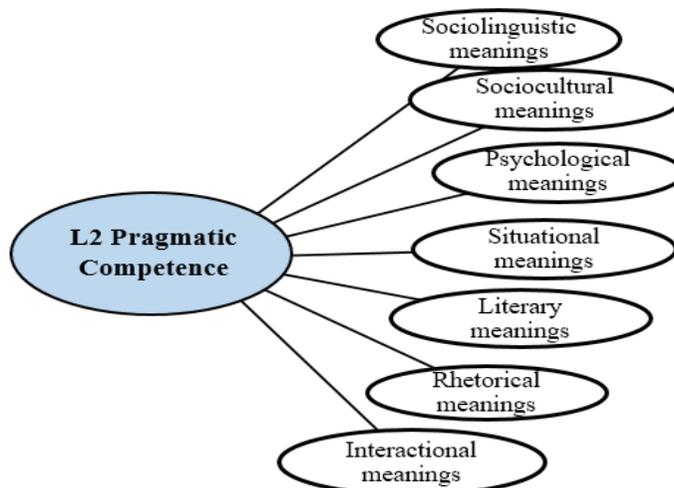
principled basis for operationalizing pragmatic competence as a coherent, meaning-oriented construct suitable for assessment purposes. The model also distinguishes between functional knowledge (intended meaning) and implicational knowledge (implied meaning), encompassing the following types of pragmatic meanings:

- **Situational (contextual) meaning** is derived from contextual cues, such as reference, implicature, association, and figurative language.
- **Sociolinguistic meaning** concerns register, politeness, and directness appropriate to social norms.
- **Sociocultural (intercultural) meaning** reflects culturally bound conventions, including humor, irony, and topic management.
- **Psychological meaning** expresses speakers' attitudes, emotions, and interpersonal stances.
- **Rhetorical meaning** relates to discourse organization, coherence, and genre awareness.
- **Interactional meaning** manages conversational flow through turn-taking, sequencing, and repair strategies.
- **Literary meaning** encompasses imaginative and stylistic language use, such as metaphor and exaggeration.

Grounded in Purpura's (2017) model, the Diagnostic Test of Pragmatic Competence (DTPC) was developed to examine how Iranian EFL learners comprehend L2 pragmatic competence through the interplay of linguistic, sociocultural, psychological, and contextual dimensions (see Figure 1). Purpura's model thus provides a comprehensive foundation for assessing L2 pragmatic competence as a dynamic, meaning-oriented construct.

Figure 1

L2 Pragmatic Competence (Adapted from Purpura, 2017)



3. Method

This research implemented an explanatory sequential mixed-methods approach (Creswell & Creswell, 2018) to develop and evaluate the Diagnostic Test of L2 Pragmatic Competence (DTPC). The quantitative phase was conducted. Then, qualitative data collection was conducted through semi-structured interviews to enrich the interpretation of results. The process comprised four stages: instrument development, piloting, online administration, and semi-structured interviews.

3.1. Participants

Three groups were involved in this research, including Iranian EFL learners, natives, and near-native speakers.

3.1.1. Iranian EFL Learners

In the first phase, 30 Iranian undergraduates majoring in English Translation completed a Written Discourse Completion Task (WDCT).

3.1.2. Native English and Near-native Speakers

In the WDCT phase, the study also included 19 native English speakers (L1 English, born and educated in English-speaking countries) and 17 near-native speakers with varied L1s who studied or worked in English-dominant countries for at least 15 years. These participants were university students working on research projects. They participated via an online survey. L1 history was verified via adopted items of the Language Experience and Proficiency Questionnaire (Marian et al., 2007) integrated into the online DTPC. Table 1 summarizes the characteristics of non-native EFL learners, natives, and near-natives in the WDCT first phase.

Table 1*Demographic Information of the Participants in the First Phase*

First Phase	Non-Native EFL Learners		Native English Speakers		Near-Native English Speakers	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Gender						
Female	23	77	16	84.2	15	88.2
Male	7	23	3	15.8	2	11.8
First Language						
British			7	36.8		
American			8	42.1		
Canadian			3	15.8		
Australian			1	5.3		
German					2	11.8
Norwegian					1	5.9
Bulgarian					2	11.8
Arabic					1	5.9
Vietnamese					1	5.9
Malay					2	11.8
Filipino					1	5.9
Afrikaans					3	17.6
Persian	30	100			4	23.5
Proficiency						
Low	14	47				
Intermediate	12	40				
Advanced	4	13				
Total Phase 1		30	19		17	

Note. Near-native participants had lived in English-speaking countries for more than 15 years. All participants were in the age range of 18–35.

Two applied linguistics professors from Azad University of Karaj rated the WDCT's responses for appropriateness, politeness, and naturalness, using Grabowski's (2007) 5-point scale (1 = inappropriate, 5 = highly appropriate). The use of natives' responses served as a baseline for contextual appropriateness, acknowledging multilingual norms rather than an idealized native benchmark (Taguchi & Li, 2020).

In the final administration, 369 Iranian university students were purposively selected to participate in this study. All participants were English majors enrolled in Translation Studies and English Literature programs at the Islamic Azad Universities of Karaj and Qazvin, Iran. These participants were native Persian speakers with no study-abroad experience, and were between 19 and 35 years old. Participants' proficiency was assessed using the Oxford Quick Placement Test (OQPT). Those who scored within the intermediate proficiency range (30–47 out of 60) were retained, resulting in a sample of 358 participants for the final statistical analysis. Because the number of advanced

participants was small ($n = 11$), they were excluded to ensure sample homogeneity.

Table 2 exhibits the features of the 358 participants in the final phase. Sampling via online Google Forms. All participants in the quantitative (WDCT, piloting, and administration) and qualitative phases (interview) provided their informed consent to contribute to the study. The confidentiality of the data was assured in line with standard ethical research practices.

Table 2

Demographic Information of the 358 Test-takers in the Final Phase

		<i>n</i>	%
Age	18– 22	252	66.5
	23– 35	117	30.9
Gender	Female	201	53
	Male	168	44.3
Degree	Undergraduate	358	94.5
	Postgraduates	11	2.9
English-Major	Literature	125	33.5
	Translation	244	64.4
First Language	Persian /Farsi	353	95.5
	Turkish	13	3.5
	Kurdish	3	0.8
Total		358	100

Note. All participants were intermediate EFL learners as determined by the Quick Oxford Placement Test.

3.2. Materials and Instruments

3.2.1. Oxford Quick Placement Test (OQPT)

OQPT (Cambridge ESOL and Oxford University Press, 2001) was administered online via Google Forms to measure participants' English proficiency and ensure they are homogeneous. The OQPT consists of 60 multiple-choice items that assess vocabulary and grammar. The test outlines three proficiency bands based on total score: low (17–30), intermediate (30–47), and advanced (48–60), as established in prior studies. Participants ($n = 358$) who scored in the *intermediate* proficiency range (30–47 out of 60) were retained for the final statistical analysis, and the advanced participants ($n = 11$) were excluded to minimize proficiency-related variance.

3.2.2. Written Discourse Completion Task (WDCT)

A WDCT was developed based on Purpura's (2017) meaning-oriented model to elicit pragmatic production through written responses. AI-assisted writing tools (Grammarly and Google Forms) were used for grammatical editing, typographical correction, and response management, while final

scenarios and items generation were reviewed by two applied linguistics professors to ensure cultural sensitivity and content validity.

3.2.3. Online Diagnostic Test of Pragmatic Competence (DTPC)

The DTPC comprised 75 multiple-choice items mapped to seven pragmatic meanings including sociolinguistic, sociocultural, psychological, situational, literary, interactional, and rhetorical. Items (questions) were reviewed by applied linguistics experts and refined (see Table 3). For online administration, all 75 items were randomized via Google Forms to avoid potential construct-blocking and minimize order effects. (see Appendix A for sample items)

Table 3

Conceptual mapping of Components and Themes of DTPC

Pragmatic Meanings	Theme	No. of Items
Sociolinguistic	Politeness, registers	10
Sociocultural	Cultural norms	10
Psychological	Humor, emotion, stance	10
Situational	figurative speech, implicatures	10
Literary	Aesthetic imagination, fantasy	10
Rhetorical	Textual structuring, genres	12
Interactional	Turn-taking and repair	13

3.2.4. Semi-structured Interviews

Ten intermediate-level participants were interviewed via Google Meet to see EFL learners' perceptions towards the DTPC's authenticity, usefulness, and challenges. Each 15–30-minute session was recorded with consent, transcribed, and analyzed thematically (Braun & Clarke, 2012).

3.3. Procedure

The DTPC was developed through a multi-stage process involving literature review, expert validation, piloting, final administration, and qualitative follow-up.

3.3.1. Scenario Generation

Some scenarios were adapted from established pragmatic studies (Grabowski, 2007; Roever, 2006). However, many other situations were developed to reflect daily conversations (e.g., classrooms, offices). The authenticity and frequency of the situations were cross-checked by the Michigan Corpus of Academic Spoken English (MICASE). Two applied linguistics professors also reviewed all scenarios for contextual appropriateness.

3.3.2. *WDCT Administration*

30 Iranian EFL learners, 19 native, and 17 near-native speakers completed the WDCT. These participants were students in the second semester of Translation Studies. They were purposively selected from those who had completed basic courses in their majors. All agreed to take part in the study, provided informed consent, after they were assured that data would be preserved privately. The same WDCT was administered to native and near-native speakers. They were selected through purposive sampling via online form shared on professional social media platforms (LinkedIn, Facebook, research exchange groups).

3.3.3. *Test Development of DTTPC*

Native and learner responses were analyzed to inform item construction through contrastive pragmatic analysis (Taguchi & Roever, 2020). Two applied linguistics professors independently evaluated responses with an appropriateness scale (1 = highly inappropriate; 5 = highly pragmatically appropriate). Following Grabowski's (2007) pragmatic rubric and achieving interrater reliability ($\kappa = .84$), the researchers retained EFL learners' responses who were rated 2–3 as distractors (representing plausible, non-target-like interlanguage realizations). The representation of pragmatically appropriate use aligned with communicative effectiveness and contextually appropriate meaning rather than native-speaker norms (Taguchi & Li, 2020). Two applied linguistics professors then reviewed all items to ensure linguistic clarity, cultural sensitivity, face, and construct validity.

3.3.4. *DTTPC Pilot Testing*

The newly developed multiple-choice DTTPC was piloted with another group of 30 Iranian EFL learners and four native speakers, yielding satisfactory reliability ($\alpha = .82$). Revisions were made accordingly.

3.3.5. *DTTPC Administration*

The revised 75-item DTTPC was administered online to Iranian EFL learners. Items were randomized to reduce order effects. Collected response options were dichotomously coded (1 = highly appropriate, 0 = inappropriate) before being exported for statistical analysis.

3.3.6. *Semi-structured Interviews*

Following quantitative analysis, ten intermediate EFL learners were interviewed to explore how they perceived the DTTPC in terms of effectiveness and challenges.

3.4. Data Analysis

The study employed both quantitative and qualitative data to evaluate the DTPC. The data from 358 intermediate EEL learners were analyzed using SPSS (version 25) and AMOS. Item difficulty and discrimination indices were computed, and the reliability of the DTPC was estimated using Cronbach's $\alpha = .93$. In addition, exploratory and confirmatory factor analyses were performed to explore factor structure and establish construct validity. In the qualitative phase, recorded interviews were analyzed using Braun and Clarke's (2012) thematic analysis to examine intermediate learners' metapragmatic reflections on the utility, engagement, and challenges of DTPC.

4. Results and Discussion

4.1. Quantitative Results

4.1.1. Item Analysis

Item analysis confirmed that most DTPC items demonstrated acceptable difficulty and discrimination indices. Consistent with Kline's (2023) threshold (.30-.80), most items fell within the acceptable range. Point-biserial correlations indicated that 72 items showed positive discrimination ($r_{pb} = .27-.53$, $p < .01$), while three items (Q11, Q16, Q53) with weak discrimination were removed.

4.1.2. Reliability of the DTPC

Item-total correlation analysis was performed to assess the items' reliability. Items 1 and 2 with correlations below .30 were excluded, resulting in 70 items with item-total correlations from .32 to .51. The communality values ranged from 0.60 to 0.75, indicating that the factor analysis was adequate. After removing five items, Cronbach's $\alpha = .93$ showed strong reliability (Table 4).

Table 4

Reliability of the DTPC

(α)	N of Items
0.930	70

4.1.3. Exploratory Factor Analysis (EFA)

An Exploratory Factor Analysis (EFA) employing principal axis factoring with Promax rotation was carried out on the 70-item DTPC to identify its underlying factor structure. A KMO value of .86 and Bartlett's test confirmed that the data were adequate for EFA (Table 5).

Table 5*KMO and Bartlett's Test*

Kaiser–Meyer–Olkin (KMO)		.863
Bartlett's Test of Sphericity	Approx. Chi-Square	7657.882
	df	2775
	Sig.	.000*

Parallel analysis, using the Monte Carlo simulation, identified four factors with eigenvalues exceeding random values (Table 6). These four factors explained approximately 28.85% of the variance, as supported by the scree plot (Figure 2). Although 24 components initially had eigenvalues greater than 1 (Table 7), only four accounted for a meaningful proportion of variance.

Table 6*Optimum Number of Factors (Monte Carlo PA)*

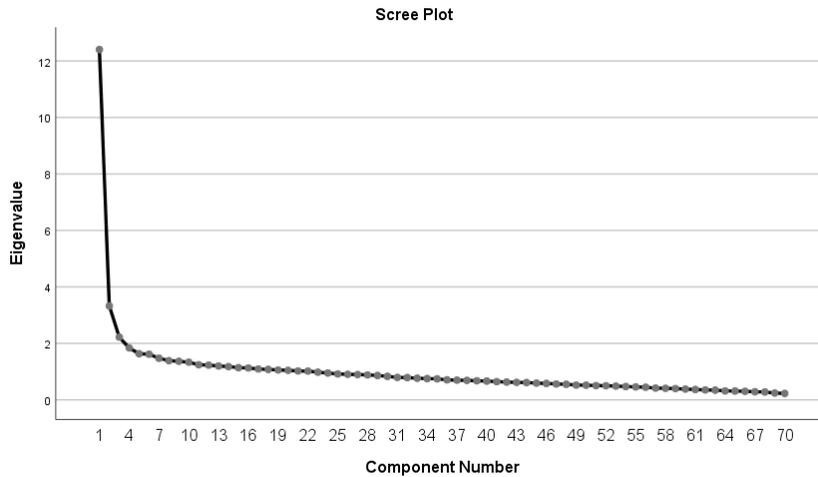
Factor	Random Eigenvalue	Observed
1	2.0323	12.576
2	1.9485	3.574
3	1.8879	2.312
4	1.8357	2.105

Table 7*Total Variance Explained*

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			
	Components	Total	Variance %	Cumulative %	Total	Variance %	Cumulative %
1		12.576	16.768	16.768	12.576	16.768	16.768
2		3.574	4.766	21.533	3.574	4.766	21.533
3		2.312	3.082	24.616	2.312	3.082	24.616
4		2.105	2.478	28.859	1.858	2.478	27.094

Extraction Method: Principal Axis Factoring

Figure 2
Scree Plot for Factors in the DTPC



After expert review and pilot testing, three items were removed due to low discrimination (Q11, Q16, Q53), and two items due to item-total correlations below 0.30 (Q1, Q2), and 70 items remained for factor analysis. The component matrix identified the items contributing meaningfully to each factor (Table 8). Following Promax rotation, items with loadings ($\geq .30$) were retained. Items with cross-loadings and low loadings ($n = 30$) were excluded. The remaining 40 items proceeded to CFA.

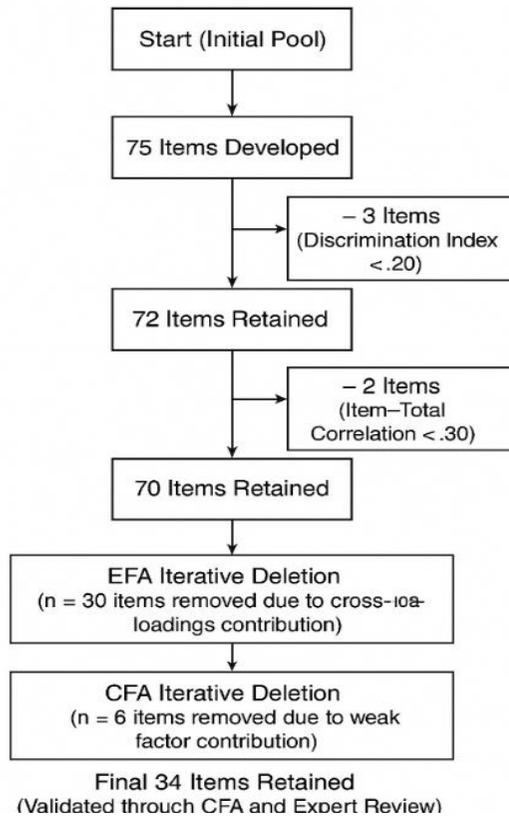
Six additional items were removed during CFA, resulting in 34 items loading on four interpretable factors of sociolinguistic (9 items), sociocultural (10 items), psychological (9 items), and situational (6 items) meanings. The DTPC item screening and retention process is shown in Figure 3.

Table 8
Structure Matrix of the Factors of the DTFC

	Components			
	Sociolinguistic	Sociocultural	Psychological	Situational
32	.611			
50	.585			
36	.548			
34	.511			
19	.482			
30	.477			
42	.464			
57	.457			
39	.444			
72		.600		
24		.599		
8		.572		
17		.567		
22		.566		
71		.496		
60		.493		
18		.453		
27		.436		
20		.418		
21			.570	
28			.530	
12			.512	
26			.490	
6			.455	
15			.454	
56			.452	
7			.450	
70			.436	
46				.539
38				.524
75				.489
37				.488
40				.455
48				.431

Extraction Method: Principal Axis Factoring.

Figure 3
Flowchart of DTPC Item Screening and Retention



Cronbach’s alpha ranged from .67 to .79 across factors, with overall reliability $\alpha = .93$, indicating acceptable internal consistency (Table 9). Items initially expected to convey rhetorical and literary meanings clustered with situational meaning, while interactional items merged with other sociolinguistic and sociocultural meanings.

Table 9
Final Factors and Reliabilities

Fators	Items	α	No.
Sociolinguistic	32-36-50-30-34-42-19-57-39	0.75	9
Sociocultural	8-17-22-72-24-71-60-27-18-20	0.79	10
Psychological	21-12-28-26-6-15-56-7-70	0.71	9
Situational	37-38-40-46-47-48	0.67	6
Total items of DTPC		0.93	34

The collapsing of rhetorical, literary, and interactional meanings into other factors is interpreted as a **context-specific empirical finding** rather than a methodological limitation. In EFL contexts such as Iran, learners' pragmatic interpretations often rely on shared contextual cues and surface-level inferencing, which can blur theoretically distinct pragmatic categories during test performance. Prior research has shown that learner variables, limited intercultural exposure, and L1-mediated inferential strategies frequently lead to construct convergence in pragmatic assessment (Tajeddin & Khanlarzadeh, 2024; Timpe-Laughlin & Choi, 2017). Such findings suggest that pragmatic meaning categories may not emerge as discrete dimensions in all instructional and sociocultural contexts.

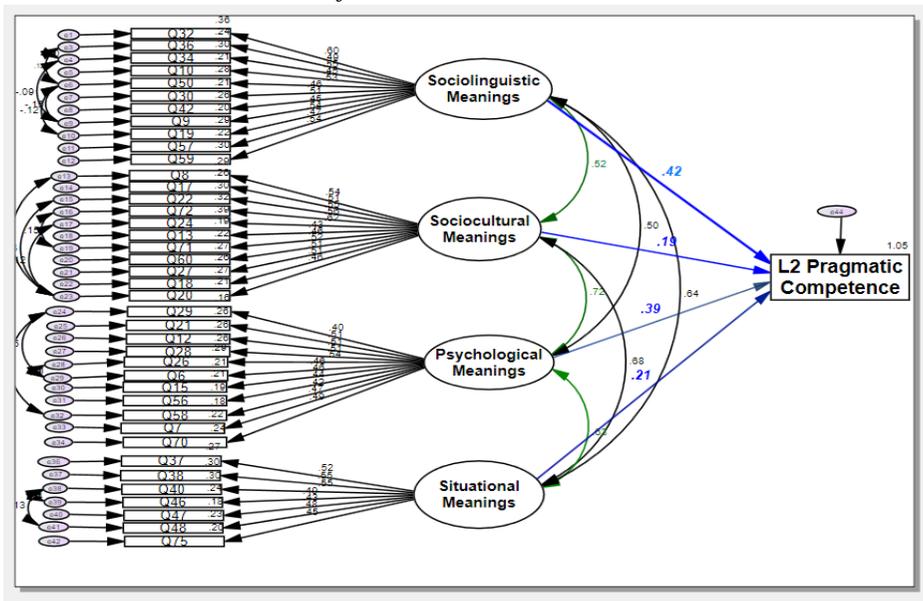
The situational meaning subscale demonstrated a Cronbach's alpha of .67, with marginal internal consistency. Given the limitations of α for context-sensitive constructs, McDonald's omega was calculated using standardized CFA loadings, yielding $\omega = .68$. This convergence between α and ω suggests stable, though modest, internal consistency for the situational subscale.

4.1.4. Confirmatory Factor Analysis (CFA)

A CFA was run using AMOS 24 to examine the fit of the DTPC model. The model showed moderate fit ($\chi^2/df = 1.40$, RMSEA = .033, GFI = .88, CFI = .92), but GFI (.88) was slightly under the threshold of .90 (Hu & Bentler, 1999). Following Browne and Cudeck's (1993) guidelines, model refinement was warranted (see Figure 4).

Figure 4

The First Structural Model of the DTPC



4.1.5. Second Structural Model's Goodness of Fit

After removing six misfitting items (Q9, Q10, Q13, Q29, Q58, Q59), the refined 34-item model demonstrated improved fit indices (Table 10) and confirmed the four-factor structure (Figure 5).

Table 10

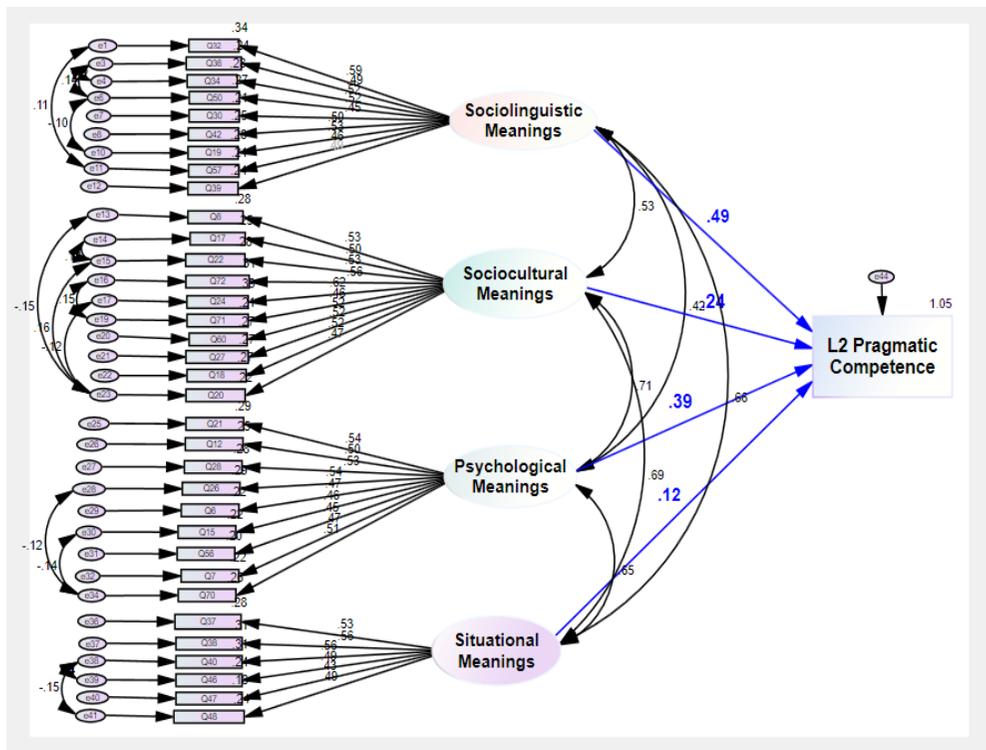
Final Model-Fit Indices for the DTPC

	χ^2 / df	RMSEA	GFI	IFI	CFI	TLI
Acceptable fit	1.33	0.030	0.90	0.94	0.99	0.94

Composite reliability (CR) values ranging from .65 to .79 exceeded the thresholds for all factors. Although average variance extracted (AVE) values were slightly low, this may be attributable to construct complexity and sample size effects (Alamer & Marsh, 2022). Nevertheless, all CV values above .60 supported discriminant validity.

Figure 5

The Second Structural Model of the DTPC



As noted by Cheung et al. (2024), reliability and validity should be interpreted holistically when item loadings and theoretical coherence are satisfactory (see CR in Table 11).

Table 11

Composite Reliability (CR), Average Variance Extracted (AVE), Maximum Shared Variance (MSV), and Average Shared Variance (ASV)

	CR	AVE	MSV	ASV	Sociocultural	Psychological	Sociolinguistic	Situational
Sociocultural	0.785	0.256	0.500	0.419	0.506			
Psychological	0.759	0.241	0.500	0.362	0.707	0.491		
Sociolinguistic	0.773	0.255	0.430	0.296	0.529	0.420	0.505	
Situational	0.650	0.225	0.476	0.439	0.690	0.641	0.656	0.475

4.2. Qualitative Results

Semi-structured interviews were conducted with 10 intermediate EFL learners to explore their perceptions of the DTTPC, including its effectiveness and challenges. Learners were purposefully selected from the DTTPC sample. Relying on their DTTPCs and OQPTs' marks, researchers ensured that interviewees possessed sufficient linguistic ability to articulate metapragmatic reflections. A sample size of ten was considered appropriate because thematic saturation was reached at $n = 10$. A thematic analytic method (Braun & Clarke, 2012) was used to code transcripts. Finally, two overarching themes emerged: the effectiveness and challenges of DTTPC (Appendix B provides interview questions, themes, and frequencies).

4.2.1. Effectiveness

Learners highlighted the DTTPC's diagnostic value and instructional relevance. Learner 3 reflected, "Even when my grammar was correct, I realized my answers did not sound natural. The DTTPC showed me why." Learner 5 commented, "It was my level, but it made me think harder."

Participants appreciated the online administration, immediate scoring, authentic task test, and reduced test anxiety. As Learner 7 explained, "I liked that the test was scored right away, so I knew what to improve."

Learners recognized the DTTPC's comprehensive construct coverage compared to traditional language exams. Learner 4 observed, "Other tests only check grammar or vocabulary, but this one checked how I use language in context." The communicative nature of the DTTPC was engaging, as Learner 7 found, "It reminded me of speaking situations, unlike grammar exams." These

reflections support the DTPC's validity in identifying pragmatic gaps often overlooked in conventional instructions and assessments.

4.2.2. Challenges

Learners reported L1 transfer and intercultural unfamiliarity as primary difficulties. Learner 5 admitted, "I thought of the Persian answer and just translated it, but it sounded strange in English." Learner 7 noted, "I didn't understand why the English answer was better, because I don't know much about English culture."

Additional barriers included task ambiguity and curriculum limitations. Learner 2 commented, "Sometimes I didn't know the point of the question, so I just guessed", Learner 9 observed, "In our books, there are grammar and vocabulary exercises, but nothing about how to answer in real social situations." These findings underscore the need to incorporate explicit pragmatics instruction into formal EFL curricula and assessment.

4.3. Discussion

Drawing on Purpura's (2017) meaning-oriented model, the study designed and validated a Diagnostic Test of Pragmatic Competence (DTPC) in Iranian EFL education. The outcomes confirmed the reliability and validity of the newly designed DTPC with 34 loaded items on four factors: sociolinguistic, sociocultural, psychological, and situational meanings.

Sociolinguistic meaning showed the most decisive loading ($\beta = .49$), reaffirming its centrality to pragmatic competence (Grabowski, 2013). Recent studies further show that online pragmatic training and technology-mediated corrective feedback (Yousefi & Nassaji, 2024) enhance learners' sociolinguistic development.

Sociocultural or intercultural meaning showed a modest contribution ($\beta = .24$), reflecting learners' limited exposure to real-life intercultural communication. Consistent with (Culpeper et al., 2018), these findings highlight that pragmatic competence is culturally grounded. While corpus-based and context-rich materials (Zhang, 2023) can improve sociocultural competence, the complexity of cross-cultural norms can often lead to pragmatic failure (Savadkouhi & Mostafaei Alaei, 2022). Explicit intercultural instruction is thus needed in Iranian EFL curricula.

Psychological meanings emerged as a moderately strong factor ($\beta = .39$). Despite being neglected in prior taxonomies of L2 pragmatic competence (Purpura, 2017), this result aligns with Grabowski's (2013) integration of emotional dimensions into L2 pragmatic test design. Research studies have mostly correlated L2 pragmatics with identity-related factors (Derakhshan et al., 2021). Malmir and Derakhshan (2020) found that learners' identity processing styles predict pragmatic knowledge through sociopragmatic, lexico-pragmatic, and cognitive strategies. By operationalizing psychological

meaning in a test design, the DTTPC extends Purpura's conceptualization and broadens how L2 pragmatic competence is assessed in the Iranian EFL context.

Situational meanings exhibited the lowest loading ($\beta = .12$), indicating learners' challenges with implicature and figurative language. This aligns with prior research (Ishihara & Cohen, 2021). Though moderately low, this factor confirms the DTTPC's diagnostic potential for identifying nuanced pragmatics.

The relatively lower reliability of the situational meaning subscale is likely attributable to its inherently context-dependent nature. Unlike form-based constructs, situational meanings require learners to integrate multiple contextual cues, sociocultural knowledge, and learners' experience rather than uniform linguistic features (Roever, 2024; Taguchi & Roever, 2020). This complexity may reduce internal consistency in psychometric measurement.

The merging of rhetorical, literary, and interactional meanings into other factors should be interpreted as a context-specific empirical finding rather than a methodological limitation. This pattern reflects how EFL learners often rely on overlapping discourse, inferential, and situational cues under constrained instructional and exposure conditions (Purpura, 2017; Tajeddin & Khanlarzadeh, 2024). Similarly, interactional meanings may not emerge as a distinct factor because turn-taking and discourse management are often embedded within broader communicative purposes (Beltrán, 2019). Consistent with prior factor-analytic research, pragmatic abilities may function as a 'family of skills' whose boundaries shift based on task and communicative goals (Wilson & Bishop, 2021). Overall, this pattern of construct convergence highlights the importance of validating pragmatic assessment models within context-specific EFL settings. Theoretically distinct pragmatic meaning categories may not always emerge as separate constructs in empirical data.

The structural model accounted for 28.85% of the variance in the DTTPC, demonstrating the context-dependence of L2 pragmatics. This result can be shaped by cognitive factors such as motivation and language mindset (Derakhshan et al., 2021). According to complex dynamic systems theory, this variability indicates that L2 pragmatic competence emerges through nonlinear interactions among linguistic, sociocultural, and affective subsystems (Taguchi, 2022)

In qualitative findings, EFL learners' reflections demonstrated the DTTPC's diagnostic effectiveness, real-life task design, and instructional usefulness despite cultural and linguistic challenges (Roever, 2021; Taguchi, 2019). Although the interview sample was small ($n = 10$), thematic saturation was achieved after the 10th participant, meeting the adequacy of the sample size for triangulating the quantitative findings.

5. Conclusion and Implications

This study was an initial step in developing and validating the Online Diagnostic Test of L2 Pragmatic Competence (DTPC), grounded in Purpura's (2017) model. Quantitative evidence supported a four-factor model including sociolinguistic, sociocultural, psychological, and situational meanings, while qualitative insights confirmed the DTPC's diagnostic effectiveness.

The DTPC shows robust potential to provide teacher researchers and educators with an instrument for identifying L2 pragmatic gaps and designing explicit, targeted instruction in EFL classrooms (Derakhshan & Shakki, 2023). Moreover, the DTPC's standardized digital delivery and automated data collection highlight the affordances of technology-mediated instruments for pragmatic assessment (O'Grady, 2023; Zangoei et al., 2019).

The DTPC results can help instructors design targeted pragmatic instruction by identifying which pragmatic meaning categories require greater support. For example, if situational and sociolinguistic meanings show lower performance, teachers can incorporate context-rich awareness-raising tasks, guided noticing, and structured input activities that draw learners' attention to contextual cues and social variables (Taguchi & Roever, 2021). Learners struggling with psychological or sociocultural meanings may benefit from scenario-based role-plays, interactive simulations, and metapragmatic reflection tasks, which have been shown to enhance learners' sensitivity to speaker intentions and cultural norms (Bardovi-Harlig, 2020). Technology-mediated simulations could be used to draw learners' attention to discourse organization and implied meanings in authentic contexts (Derakhshan, 2023).

In addition, the DTPC profile can guide teachers in providing focused, form–function feedback, such as prompts, recasts, and online teacher or chatbox-mediated feedback strategies that draw learners' attention not only to linguistic accuracy but also to pragmatic appropriateness (Derakhshan et al., 2024; Fakher Ajabshir, 2025; Timpe-Laughlin & Choi, 2020). By aligning instructional techniques with specific areas of pragmatic difficulty identified through the DTPC, teachers can implement data-informed, individualized interventions that strengthen learners' pragmatic awareness and performance in contextually varied communicative tasks.

Despite these contributions, limitations must be noted. The findings should be interpreted cautiously within the boundaries of the Iranian EFL context. The study's intermediate Iranian EFL learners and purposive sampling method restrict the generalizability of the findings.

Although previous research has shown that learner-related variables such as proficiency, study-abroad experience, and cognitive style may influence L2 pragmatic development, these factors were not examined in the present study and therefore remain an important direction for future research.

Future researchers could replicate the DTPC across different proficiency levels and contexts. Researchers may also refine integrated dimensions (rhetorical, literary, interactional) as distinct components. Pretest–posttest designs could be applied to track instructional gains. Further longitudinal and teacher-training courses are needed to evaluate the DTPC’s classroom integration and applicability.

Acknowledgements

The authors thank all professors, natives, and EFL learners who contributed to this research

References

- Alamer, A., & Marsh, H. (2022). Exploratory structural equation modeling in second language research: An applied example using the dualistic model of passion. *Studies in Second Language Acquisition*, 44(5), 1477-1500. <https://doi.org/10.1017/S0272263121000863>
- Alavi, S. M., Shahsavari, M., & Norouzi, M. H. (2020). Diagnosing EFL learners' development of pragmatic competence implementing computerized dynamic assessment. *Issues in Language Teaching*, 9(1), 117-149.
- Arabmofrad, A., & Mehdiabadi, F. (2022). Developing a multiple-choice discourse completion test for Iranian EFL learners: The case of the four speech acts of apology, request, refusal, and thanks. *Language Related Research*, 13(4), 1-34.
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- Bardovi-Harlig, K. (2020). Pedagogical linguistics: A view from L2 pragmatics. *Pedagogical Linguistics*, 1(1), 44-65.
- Beltrán, J. (2019). A meaning-based multiple-choice test of pragmatic knowledge: Does it work? *Studies in Applied Linguistics & TESOL*, 19(1), 42-71.
- Birjandi, P., & Rezaei, S. (2010). Developing a multiple-choice discourse completion test of interlanguage pragmatics for Iranian EFL learners. *ILLI Language Teaching Journal*, 6(1-2), 43-58.
- Braun, V., & Clarke, V. (2012). *Thematic analysis*. American Psychological Association. <https://doi.org/10.1037/13620-004>
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, 41(2), 745-783. <https://doi.org/10.1007/s10490-023-09871-y>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Culpeper, J., Mackey, A., & Taguchi, N. (2018). *Second language pragmatics: From theory to research*. Routledge.
- Derakhshan, A., R. Eslami, Z., & Shakki, F. (2024). Comparing compliments in face-to-face vs. online interactions among Iranian speakers of Persian. *Pragmatics and Society*, 15(2), 320-344.
- Derakhshan, A., Eslami, Z. R., & Ghandhari, N. (2021). Investigating the interplay of emotional intelligence and interlanguage pragmatic

- competence in Iranian lower-intermediate EFL learners. *Issues in Language Teaching*, 10(1), 37-66.
- Derakhshan, A., & Shakki, F. (2023). A meta-analytic study of instructed second language pragmatics: A case of the speech acts of request. *Journal of Research in Applied Linguistics*, 12(1), 15-32.
- Fakher Ajabshir, Z. (2025). A comparative study of teacher feedback and chatbot feedback on second language learners' pragmalinguistic and sociopragmatic competences. *International Journal of Human-Computer Interaction*, 41(16), 9881-9892. <https://doi.org/10.1080/10447318.2024.2429756>
- Grabowski, K. C. (2007). Reconsidering the measurement of pragmatic knowledge using a reciprocal written task format. *Studies in Applied Linguistics and TESOL*, 7(1), 1-48.
- Grabowski, K. (2013). Investigating the construct validity of a role-play test designed to measure grammatical and pragmatic knowledge at multiple proficiency levels. In S. J., Ross & G. Kasper (Eds.), *Assessing second language pragmatics* (pp. 149-171). Palgrave Macmillan. https://doi.org/10.1057/9781137003522_6
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Hudson, T., Detmer, E., & Brown, J. D. (1992). *A framework for testing cross-cultural pragmatics* (Vol. 2). National Foreign Language Resource Center.
- Ikeda, N. (2021). Assessing L2 learners' pragmatic ability in problem-solving situations at English-medium university. *Applied Pragmatics*, 3(1), 51-83.
- Ishihara, N., & Cohen, A. D. (2021). Assessment of L2 pragmatics in the classroom. In N. Ishihara & A. D. Cohen (Eds.), *Teaching and learning pragmatics* (2nd ed., pp. 259-296). Routledge.
- Jianda, L. (2007). Developing a pragmatics test for Chinese EFL learners. *Language Testing*, 24(3), 391-415.
- Kecskes, I. (2014). *Intercultural pragmatics* (Vol. 288). Oxford University Press.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.
- Laughlin, V. T., Wain, J., & Schmidgall, J. (2015). Defining and operationalizing the construct of pragmatic competence: Review and recommendations. *ETS Research Report Series*, 2015(1), 1-43. <https://doi.org/10.1002/ets2.12053>
- Leech, G. (2014). *The pragmatics of politeness*. Oxford University Press.

- Malmir, A., & Derakhshan, A. (2020). The socio-pragmatic, lexicogrammatical, and cognitive strategies in L2 pragmatic comprehension: The case of Iranian male vs. female EFL learners. *Iranian Journal of Language Teaching Research*, 8(1), 1-23.
- Mao, T., & He, S. (2021). An integrated approach to pragmatic competence: Its framework and properties. *SAGE Open*, 11(2), 1–13 <https://doi.org/10.1177/21582440211011472>
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967.
- Mootoosamy, K., & Aryadoust, V. (2024). Transitioning from communicative competence to multimodal and intercultural competencies: A systematic review. *Societies*, 14(7), 115. <https://doi.org/10.3390/soc14070115>
- O'Grady, S. (2023). An AI-generated test of pragmatic competence and connected speech. *Language Teaching Research Quarterly*, 37, 188-203. <https://doi.org/10.32038/ltrq.2023.37.10>
- Purpura, J. E. (2017). Assessing meaning. In E. Shohamy, I. Or, & S. May (Eds.), *Language testing and assessment* (pp. 33–61). Springer.
- Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, 23(2), 229-256.
- Roever, C. (2021). *Teaching and testing second language pragmatics and interaction: A practical guide*. Routledge.
- Roever, C. (2024). Assessment and L2 pragmatics. *Applied Pragmatics*, 6(2), 117-130.
- Saeedi, Z., & Soltani, M. (2025). Developing an AI chatbot for language pragmatics instruction: from algorithms to dynamic assessment in an EFL context. *Computer Assisted Language Learning*, 1-22. <https://doi.org/10.1080/09588221.2025.2532014>
- Savadkouhi, Z. B., & Mostafaei Alaei, M. (2022). Cross-cultural pragmatic failure in English language classrooms: Development of a multi-dimensional model based on EFL teachers' and learners' perceptions. *Journal of Modern Research in English Language Studies*, 9(3), 49–72. <https://doi.org/10.30479/jmrels.2022.16141.1962>
- Sedaghatgoftar, N., Karimi, M. N., Babaii, E., & Reiterer, S. M. (2019). Developing and validating a second language pragmatics aptitude test. *Cogent Education*, 6(1), Article e1654650.
- Shakki, F., Naeini, J., Mazandarani, O., & Derakhshan, A. (2021). Instructed second language pragmatics for the speech act of apology in an Iranian EFL context: A meta-analysis. *Applied Research on English*

- Language*, 10(3), 77-104.
<https://doi.org/10.22108/are.2021.128213.1709>
- Tajeddin, Z., & Khanlarzadeh, N. (2024). The role of learner variables in pragmatic development during study abroad: A systematic review. *Languages*, 9(3), 96.
- Taguchi, N. (2019). Second language acquisition and pragmatics: An overview. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 1–14). Routledge.
- Taguchi, N., & Li, S. (2020). Contrastive pragmatics and second language (L2) pragmatics: Approaches to assessing L2 speech act production. *Contrastive Pragmatics*, 2(1), 1-23.
- Taguchi, N., & Roever, C. (2020). *Second language pragmatics*. Oxford University Press.
- Timpe-Laughlin, V., & Choi, I. (2017). Exploring the validity of a second language intercultural pragmatics assessment tool. *Language Assessment Quarterly*, 14(1), 19–35.
<https://doi.org/10.1080/15434303.2016.1256406>
- Wilson, A., & Bishop, D. (2021). A novel online assessment of pragmatic and core language skills: An attempt to tease apart language domains in children. *Journal of Child Language*, 49, 38-59. <https://doi.org/10.1017/s0305000920000690>
- Yousefi, M., & Nassaji, H. (2024). The impact of corrective feedback on L2 pragmatics production in face-to-face and technology-mediated settings. *Language Teaching Research Quarterly*, 39, 305–328.
- Zangoei, A., Zareian, G., Adel, S. M. R., & Amirian, S. M. R. (2019). The impact of computerized dynamic assessment on Iranian EFL learners' interlanguage pragmatic development. *Journal of Modern Research in English Language Studies*, 6(4), 139–165.
<https://doi.org/10.30479/jmrels.2019.11536.1433>
- Zhang, Y. (2023). The effects of interleaved and blocked corpus-based practice on L2 pragmatic development. *Studies in Second Language Acquisition*, 45(4), 812–837.

Appendix A

Sample Items of the DTFC

Please read the conversations and respond to situations by selecting the appropriate one.

Sociolinguistics Meanings

32- Your professor insists that students should listen and not ask any questions until the end of his lecture.

You: Can I ask a question?

Professor: No, you should not have forgotten what I told you.

You:

- a) I do apologize. It just slipped my mind. I shall wait until the end.
- b) I am sorry, I did not know that.
- c) My bad, Professor. I will not repeat it.
- d) If I don't ask, you look like you are going to talk forever.

Sociocultural Meanings

24- You are hosting a party and have prepared Iranian cuisine for your international friends. When they offer help, you say, "No, thanks, everything's on schedule". Then invite them to eat, you would say:

- a) Grubs up! Dig in and enjoy!
- b) Please, here you are.
- c) Food is served. You can start.
- d) Attention, everyone! Everything is ready! Help yourself, please.

Psychological Meanings

21- Your colleague has gotten a promotion. You were with her on many missions. You feel that you are qualified for the promotion. You see her at the office.

You: Hey. Hats off! You earned it!

Colleague: Thanks a lot! I thought you were the one who'd get it..

You:

- a) You deserve that
- b) God willing, it will be my turn.
- c) Yes, I work hard for it, so it should be for me.
- d) My time will come. Congrats to you!!

Appendix B

Questions, Themes, and Learners' Quotes from Semi-Structured Interviews

Interview Qs: <i>Samples of Learner's Quotes</i>	Theme	F
1. What do you think about the usefulness of the DTPC for your learning?		
<i>"Even when my grammar was correct, I realized my answers did not sound natural. The DTPC showed me why."</i>	Diagnostic value	5
	Instructional relevance	4
	Level-appropriate challenge	3
	Online usefulness	4
<i>"I liked that the test was scored right away, so I knew what to improve."</i>	Fast scoring	3
	Low-pressure environment	4
2. How did your background experiences help you answer DTPC questions?		
<i>"Because I watch films, play video games, and listen to English, it helps me answer in different situations."</i>	Exposure to authentic input	6
	Cultural awareness	4
3. What challenges did you face when answering the DTPC?		
<i>"I thought of the Persian answer and just translated it, but it sounded strange in English."</i>	L1 transfer/translation	5
	Lack of cultural knowledge	6
	Uncertainty/guessing	3
	Insufficient textbook coverage	4
4. How did DTPC compare to other language-learning English tests you have used?		
<i>"Other tests only check grammar or vocabulary, but this one checked how I use language in context."</i>	Broader construct coverage	8
	Engagement	6
	Communicative relevance	5