Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban
DOI: https://doi.org/10.15575/jpba.v8i1.35023

Ta'lim al-'Arabiyyah
JURNAL PENDIDIKAN BAHASA ARAB & KEBAHASAARABAN
P-ISSN: 2721-5237 | E-ISSN: 2549-208X // Vol. 8 No. 1 | 1-17
https://journal.uinsgd.ac.id/index.php/Talim/index

INVESTIGATING THE VALIDITY AND RELIABILITY OF WTC CONSTRUCT IN L2 ARABIC: A CASE STUDY

Muhamad Fahri[1*], Fajar Pinastia Arrahman[2], Azkia Muharom Albantani[3], Muhammad Rizal Mahfuzo[4], Muhamad Ihwanul Muslimin[5]
[1,2,3,5] UIN Syarif Hidayatullah Jakarta, Indonesia
[4] University of Michigan, United States

Article Information
Article History:
Received : 15-April-2024
Revised : 23-Mei-2024
Accepted : 26-Juni-2024
Published : 30-Juni-2024

Keywords:
Instrument;
Learning Arabic as a second language;
Online communication;
Willingness to communicate

Articles Available Online:

ABSTRACT
The main goal of learning Arabic is to develop practical linguistic skills among learners centered around a willingness to communicate (WTC). WTC refers to learners' strong desire and readiness to speak in specific contexts or situations. A critical challenge is accurately measuring WTC in Arabic language learning using a valid and reliable instrument. Understanding WTC is crucial as it indicates learners' motivation to use Arabic, enhances their speaking proficiency, and creates a supportive linguistic environment. Research on WTC in Arabic language learning is currently limited, necessitating further exploration. Consequently, this study aims to develop a valid and reliable WTC measurement instrument for Arabic language learners using a quantitative approach, incorporating both Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) methods. The participants for this research are 204 Arabic language learners, and a survey was utilized for data collection. The WTC instrument employed in this study is an adaptation of existing instruments, modified to align with the context of Arabic language learning and the specific conditions of Indonesian learners. The findings reveal that the WTC construct has yielded valid and reliable items that can be used to measure WTC among learners in Arabic language learning. Furthermore, this construct has three main dimensions: online, classroom, and offline context. Thus, it is hoped that the WTC construct developed can be utilized to measure aspects of WTC among Arabic language learners, thereby enabling the determination of further steps in realizing linguistic behaviors.

CORRESPONDING AUTHOR:
Muhamad Fahri
Departement of Arabic Education,
Universitas Islam Negeri Syarif Hidayatullah Jakarta,
Jl. Ir. H. Djuanda No. 95, Ciputat, Tangerang Selatan, Banten 15412, Indonesia
Email: m.fahri@staff.uinjkt.ac.id

How to Cite:
Fahri, M., Arrahman, P. F., Albantani, A. M., Mahfuzo, M. R., & Muslimin, M. I. (2024). Investigating the Validity and Reliability of WTC Construct in L2 Arabic: A Case Study. *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban, 8*(1), 1–17. https://doi.org/10.15575/jpba.v8i1.35023

INTRODUCTION

An ideal Arabic language learning environment should effectively cultivate learners' strong willingness and language skills, ultimately shaping language behaviours among Arabic learners. An essential aspect of shaping language behaviors is the Willingness to Communicate (WTC), which is defined as an individual's readiness to engage in conversation with one or more people at a given time using the second language (Macintyre et al., 1998; Zarrinabadi & Tanbakooei, 2016). WTC is also recognized as a determinant of language behaviour (Macintyre et al., 1998). Research on WTC in second language learning has been extensively conducted, focusing on factors influencing WTC among language learners, including communicative competence, language anxiety, and learners' motivation (Elahi Shirvan et al., 2019). Understanding WTC requires examining all contributing aspects, not solely focusing on learners' cognitive aspects but also considering their affective aspects tailored to specific communication contexts, whether in class, outside class, or online communication contexts (Lee & Chen Hsieh, 2019).

Significant attention should be devoted to the precise measurement of WTC, as researchers may hold their perceptions and perspectives when assessing WTC among second language learners. Particularly in Arabic language learners, limited research on WTC is available. Mahmoodi & Moazam (2014) is one of the few studies, albeit developing their measurement without prior testing of its validity and reliability. A valid and reliable WTC measurement construct will accurately depict language learners' conditions, thus serving as a crucial foundation for creating more effective learning environments, particularly in Arabic learning.

Examining its inception, the measurement construct of WTC was initially formulated by (McCroskey & Baer, 1985), comprising 20 items with response types being percentage-based (0 = Never – 100 = Always). In their validation, they posited the unidimensional nature of the construct, with 12 items demonstrating the highest factor analysis scores and exhibiting high-reliability values as well (McCroskey & Baer, 1985). A limitation of the WTC construct from McCroskey & Baer's (1985) study is that it fails to account for how other dimensions may influence an individual's communicative willingness. For instance, if a teacher poses a question to their student using a second language, not all students may desire or be capable of responding due to other factors such as the mental state, abilities, and skills of the student, as well as the circumstances and conditions that may render a student uncomfortable in engaging in communication (Macintyre et al., 1998). Thus, WTC is understood as a concept with multidimensional facets shaping it.

Moreover, upon reviewing the findings of Zarrinabadi & Tanbakooei (2016), which qualitatively explored the study of WTC, it is noted that several studies related to WTC employ a diverse array of instruments, such as semi-structured interviews, stimulated recalls, open-ended questionnaires, focused essay techniques, and classroom observations conducted by researchers such as (Y. Cao, 2011; Y. K. Cao, 2014; Freiermuth & Jarrell, 2006; Kang, 2005; Peng, 2012; Zarrinabadi, 2014). However, a primary concern arises regarding the measurement constructs of WTC, as the multitude of variations has yet to adequately capture WTC across different learner profiles and language learning conditions. In other words, there has not been a conclusive determination regarding a valid and reliable measurement construct for assessing WTC among second language learners, particularly those studying Arabic.

In the context of Arabic language learning, only two studies examine WTC. Firstly, Mahmoodi & Moazam (2014) investigated the relationship between WTC and achievement or success in second language learning (Arabic) among Bu-Ali Sina University, Iran students. This study utilized and adopted the WTC measurement by Baghaei & Dourakhshan (2012) within English language learning, comprising a measurement construct of 20 items, all considered to measure L2 Arabic WTC within a one-dimensional framework. Neither Mahmoodi & Moazam (2014) nor Baghaei & Dourakhshan (2012) presented validity and reliability testing results for the WTC measurement construct in their papers. As previously noted, WTC can be multidimensional and influenced by various factors, leading to overlap among items in the instrument. The second study related to L2 Arabic WTC was conducted by Kholis et al. (2023), examining WTC in both in-class, out-of-class, and digital contexts among students at State Islamic University in Indonesia. Their study adopted the WTC measurement construct by Lee & Chen Hsieh (2019) within the context of English language learning. Notably, Kholis et al. (2023) only adopted the WTC construct by Lee & Chen Hsieh (2019) within the in-class context, reducing the original 8 items to 6.

Meanwhile, in out-of-class and digital contexts, they introduced their innovations. Additionally, neither Kholis et al. (2023) nor Lee & Chen Hsieh's (2019) studies provided valid and reliable testing results for their instruments. Based on this premise, this study develops and evaluates the WTC construct, focusing on its validity and reliability, thereby providing a comprehensive understanding of WTC itself.

Understanding the aspects of WTC among Arabic language learners is crucial for comprehending their behavior and enhancing their language competence, particularly their speaking proficiency. This understanding is expected to enable Arabic language instruction to create learning conditions that provide supportive environments, fostering learners' desire to realize functional language use in communicative and interactive activities. Ultimately, this can shape effective language behavior, creating a more effective learning environment for Arabic language learners.

Considering that the concept of WTC serves as a fundamental basis in shaping language behaviour, which is a primary objective in Arabic language learning, along with the issues outlined regarding the measurement instruments for WTC, this research aims to construct a WTC L2 Arabic instrument for Arabic language learners, taking into account the results of validity and reliability testing of the construct by using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Thus, it is anticipated that this construct can serve as a valid and reliable instrument in measuring the extent of learners' willingness to communicate in Arabic.

METHOD

Instrument and Participants

This study constitutes developmental research aiming to construct a measurement instrument tested under a quantitative paradigm. The instrument in question is the WTC construct, anticipated to gauge the extent of Arabic language learners' willingness to communicate in specific situations and conditions. The participants in this study comprise 204 Arabic language learners at Syarif Hidayatullah State Islamic University Jakarta. The sample comprises 60 male students (29.4%) and 144 female students (70.6%). Additionally,

the participants are distributed across various semesters: 146 students (71.6%) in the second semester, 29 students (14.2%) in the fourth semester, 11 students (5.4%) in the sixth semester, 10 students (4.9%) in the eighth semester, and 8 students (3.9%) in the tenth semester. This sample is deemed adequate for examining a construct concerning a particular behaviour, according to Sürücü et al. (2022).

The measurement construct developed in this study is an adapted and modified construct derived from several studies, such as those by (Lee & Drajati, 2020; Lee & Lu, 2023; Mahmoodi & Moazam, 2014), tailored to the conditions and situations of the Arabic language learning in Indonesia. The measurement of L2 WTC by Lee & Drajati (2020) provides insights into the digital (online) context, termed *Informal Digital Learning of English* (IDLE). It encompasses four items regarding how language learners utilize Facebook and online games to communicate in English online through various features. This study adapted these items to four items expressing the desire to communicate in Arabic online via social media. Additionally, the measurement by Lee & Lu (2023) offers an understanding of English communication within a classroom setting, represented by five items (e.g., "*When you have a chance to make a presentation in front of a large group*"). The current study adapted it into four items that reflect classroom communication.

The construct by Mahmoodi & Moazam (2014) provides insights into real-life situations in Arabic communication, initially comprising 20 items. It was adapted in this study to four items illustrating real-life scenarios in Arabic communication. Consequently, the measurement of L2 WTC in Arabic in this study consists of a total of 12 items. Thus, 12 items were formulated as follows:

Table 1. WTC Construct for L2 Arabic

Item	Item	Respond
WTC 1	If I encounter a native Arabic speaker on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to engage in conversation with them.	(1 – 5)
WTC 2	If I meet someone who speaks Arabic (non-native) on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to converse with them.	(1 – 5)
WTC 3	If someone introduces me to native Arabic speakers, I ... to attempt to communicate in Arabic with them to test my language skills.	(1 – 5)
WTC 4	I ... to accompany native Arabic speakers and serve as their tour guide.	(1 – 5)
WTC 5	I ... to ask questions in Arabic during lectures.	(1 – 5)
WTC 6	I ... to deliver presentations in Arabic in front of the class.	(1 – 5)
WTC 7	I ... to give presentations in Arabic at seminars.	(1 – 5)
WTC 8	I ... to communicate in Arabic with my campus friends.	(1 – 5)
WTC 9	I ... to communicate with native Arabic speakers through online media consistently.	(1 – 5)
WTC 10	I ... to engage in chat/direct messages in Arabic with my colleagues regularly.	(1 – 5)
WTC 11	I ... to provide comments in Arabic on digital platforms.	(1 – 5)
WTC 12	I ... to post status updates on social media using Arabic..	(1 – 5)

Respond: (1 = Definitely not willing, 2 = Not willing, 3 = Neutral, 4 = Willing, 5 = Definitely willing)

The data collection technique employed in this study is an online survey, wherein respondents voluntarily complete the structured instrument via an online survey link.

Research Analysis Methods

Two analytical techniques were employed to assess the validity of the WTC L2 Arabic measurement instrument in this study. Firstly, Exploratory Factor Analysis (EFA), a multivariate statistical method that has become fundamental in developing and validating psychological theories and measurement instruments (Watkins, 2018). The primary goal of EFA is to summarize data for more straightforward interpretation and to understand the relationships or patterns among observed variables in a measurement instrument, where observed variables are grouped into limited clusters. Through such clustering, researchers can focus on fewer items that can explain the variable structure rather than considering too many items that may not have significant impact or even importance in a study. Ultimately, items falling into meaningful categories are likely to be used in studies related to variables (Sürücü et al., 2022). EFA in this research aims to identify the factors or dimensions measured in the WTC L2 Arabic construct among Arabic language learners, along with reliability values based on the criteria of KMO and Bartlett's Test, where the significance value of Bartlett's Test should be < 0.05 (Sürücü et al., 2022), and KMO values based on Kaiser (1960) criteria as follows:

Table 2. KMO Ranges and Description

KMO	Description
> 0.90	Marvellous
> 0.80	Meritorious
> 0.70	Middling
> 0.60	Mediocre
> 0.50	Miserable
< 0.50	Unacceptable

Upon confirming the acceptance of *KMO* and *Bartlett's Test*, the subsequent step is to analyze the factors or dimensions that the items in the instrument can explain. In conducting the EFA analysis, this study utilized the assistance of SPSS software version 25.

Secondly, following identifying the dimensions measured by the WTC instrument through EFA, the next step is to conduct a Confirmatory Factor Analysis (CFA). The use of CFA begins with defining latent variables to be measured based on substantive theory or prior knowledge (such as factor findings in EFA), which are then constructed into dimensions to be observed in measuring these variables through the items in the instrument (Jöreskog et al., 2016). The objective of CFA in this research is to test the validity and reliability of the instrument in measuring communication tendencies among Arabic language learners. Item validity is assessed based on the loading factor values for each item, where an item is considered valid if the loading factor value > 0.60 (MacCallum et al., 2001) and if the t-value > 1.96 (Umar & Nisa, 2020). For reliability assessment, Cronbach's alpha is utilized, a widely used reliability measure in social sciences, education, and organizational studies (Bonett & Wright, 2015). The criteria for interpreting Cronbach's alpha values are 0.93-0.94 (excellent); 0.91-0.93 (firm); 0.84-0.90 (reliable); 0.81 (robust); 0.76-0.95 (reasonably high); 0.73-0.95 (high); 0.71-0.91 (good); 0.70-0.77 (relatively high); 0.68 (slightly low); 0.67-0.87 (reasonable); 0.64-0.85 (adequate); 0.61-0.65 (moderate); 0.58-0.97 (satisfactory); 0.45-0.98 (acceptable); 0.45-0.96 (sufficient); 0.4-0.55 (not adequate); 0.11 (low) (Taber, 2018). Additionally, to support item reliability, the inter-item correlation results are also examined. The correlation analysis at this stage utilizes the Jamovi software.

Once the items are deemed valid and reliable, the subsequent step involves detecting and modifying model fit by examining the Root Mean Square Error of Approximation (RMSEA) in the path model according to the following criteria (Browne & Cudeck, 1992; Kline, 2016; Xia & Yang, 2019):

Table 3. RMSEA Criteria

RMSEA	Criteria
≥ 0.10	poor fit
≥ 0.08	reasonable
≥ 0.05	close approximate fit
≤ 0.05	fit
≤ 0.03	good fit

After the construct model meets the abovementioned criteria, the final step is to organize the WTC L2 Arabic construct items according to the existing dimensions. In conducting the CFA analysis, this study utilized Lisrel software.

RESULT AND DISCUSSION

Exploratory Factor Analysis (EFA) for WTC Construct L2 Arabic

In the EFA analysis, the first consideration is the results of Bartlett's Test of Sphericity and the KMO test. It is stipulated that if the significance value of Bartlett's Test of Sphericity is < 0.05, it is advisable to increase the sample size or remove items causing the scattered correlation model from the analysis (Sürücü et al., 2022). With the WTC Construct L2 Arabic consisting of 12 items, the results of Bartlett's Test of Sphericity and the KMO test are as follows:

Table 4. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		**0.884**
Bartlett's Test of Sphericity	Approx. Chi-Square	1135.648
	df	66
	Sig.	**0.000**

Based on Table 4, it can be seen that the significance value of Bartlett's Test of Sphericity is 0.000, indicating that the observed correlation matrix significantly differs from a single correlation matrix. Thus, sufficient evidence suggests that the observed items are not linearly uncorrelated. It supports linear combinations among the WTC Construct L2 Arabic items. This study's results endorse conducting factor structure extraction from the data. Additionally, the KMO value is found to be 0.884, which, according to the pre-established criteria, falls into the category of Meritorious. In other words, the WTC Construct L2 Arabic demonstrates good adequacy for factor analysis. It is because a KMO value approaching 1 indicates a strong correlation pattern among the items in the dataset, making them suitable for factor analysis. Hence, it can be inferred that the existing factors can explain most inter-item variability, rendering the data ideal for EFA analysis.

Based on the analysis, the WTC Construct L2 Arabic data appears highly suitable for exploratory factor analysis (EFA). The significance value of Bartlett's Test of Sphericity, which is 0.000, indicates that the correlation matrix derived from the observed data significantly departs from an identity matrix where variables are uncorrelated. It suggests

substantial correlations among the items, which is essential for identifying underlying factors influencing the observed variables. Additionally, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.884, falling within the "Meritorious" range, indicating a solid correlation pattern among the items. A high KMO value near 1 implies that a significant portion of variance among variables can be explained by common factors, supporting the dataset's appropriateness for factor analysis. Consequently, these findings collectively endorse conducting EFA on the WTC Construct L2 Arabic. Bartlett's Test of Sphericity and the KMO measure affirm that the variables are interconnected and their variance structure is conducive to factor extraction and interpretation.

Next, we examine the results of commonalities, which provide information regarding the variance of each item that can be explained by the factors extracted in an analysis (Watkins, 2018). In other words, commonalities output describes how well the WTC Construct L2 Arabic items contribute to forming the existing factors.

Table 5. Communalities

Item	Initial	Extraction
WTC 1	1.000	0.668
WTC 2	1.000	0.630
WTC 3	1.000	0.578
WTC 4	1.000	0.519
WTC 5	1.000	0.692
WTC 6	1.000	0.759
WTC 7	1.000	0.762
WTC 8	1.000	0.563
WTC 9	1.000	0.692
WTC 10	1.000	0.776
WTC 11	1.000	0.690
WTC 12	1.000	0.633

Extraction Method: Principal Component Analysis.

Based on Table 5, it is evident that the initial column (values before extraction) consistently shows a value of 1.000, indicating that each item is fully represented in the initial data. The extraction column, on the other hand, presents the commonalities, representing the variance proportion for each item that the extracted components can explain. The results reveal that item WTC1 has 66.8% of its variance explained by the extracted components. Item WTC2 has 63.0% of its variance explained, WTC3 has 57.8%, WTC4 has 51.9%, WTC5 has 69.2%, WTC6 has 75.9%, WTC7 has 76.2%, WTC8 has 56.3%, WTC9 has 69.2%, WTC10 has 77.6%, WTC11 has 69.0%, and WTC12 has 63.3% of their variances explained. The commonality values after extraction demonstrate how well the principal components explain the variance of each item. It is observed that WTC10 ("*I ... to engage in chat/direct messages in Arabic with my colleagues regularly*") has the highest communality value, indicating that the principal components are very effective in explaining the variance of this item. Conversely, WTC4 ("*I ... to accompany native Arabic speakers and serve as their tour guide*") has the lowest communality value, meaning that the principal components are less effective in explaining this item's variance than others. However, it is noteworthy that all item extraction values are above 0.5, indicating that all items generally represent the variance in measuring the WTC aspect.

Next, let us examine the output of the total variance explained below:

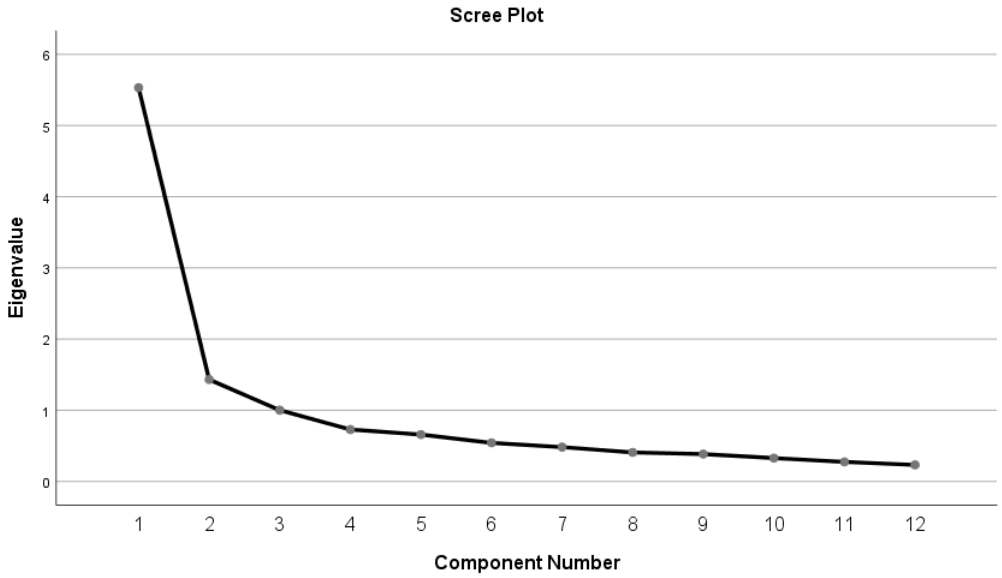
Table 6. Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total
1	5.531	46.092	46.092	5.531	46.092	46.092	2.939
2	1.431	11.928	58.02	1.431	11.928	58.02	2.552
3	1.001	8.338	66.358	1.001	8.338	66.358	2.472
4	0.729	6.076	72.434				
5	0.658	5.485	77.919				
6	0.542	4.514	82.433				
7	0.482	4.017	86.45				
8	0.407	3.395	89.845				
9	0.384	3.201	93.046				
10	0.328	2.734	95.78				
11	0.274	2.286	98.066				
12	0.232	1.934	100				

Extraction Method: Principal Component Analysis.

Based on the exploratory factor analysis (EFA) utilizing the Principal Component Analysis (PCA) extraction method, it is discerned that the WTC construct L2 Arabic can measure three dimensions, with eigenvalues of 5.531 for the first dimension, 1.431 for the second dimension, and 1.001 for the third dimension. Following the criterion stating that eigenvalues are considered significant if they exceed 1 (Kaiser, 1960), three dimensions emerge as substantial. It indicates three dimensions significantly measured by the WTC L2 Arabic instrument. The Scree Plot also illustrates this finding.

Figure 1. Scree Plot EFA for WTC Construct L2 Arabic



Based on the scree plot above, which is grounded in the previously mentioned eigenvalue scores, it becomes evident that the WTC construct L2 Arabic possesses three principal dimensions. These dimensions are elucidated as follows: the first dimension corresponds to 'online,' the second to 'classroom,' and the third to 'offline.' Following the

factor extraction process, which asserts these dimensions within a WTC construct L2 Arabic, the next step involves examining the outcomes of the rotated component matrix. Based on the Confirmatory Factor Analysis (CFA) and the alignment of the WTC instrument with the results of the rotated component matrix, the findings are as follows:

Table 7. Rotated Factor Component Coefficients

Item	Statement	F1 (Online)	F2 (Classroom)	F3 (Offline)
WTC 1	If I encounter a native Arabic speaker on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to engage in conversation with them.	0.295	0.289	**0.705**
WTC 2	If I meet someone who speaks Arabic (non-native) on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to converse with them.	0.307	0.027	**0.732**
WTC 3	If someone introduces me to native Arabic speakers, I ... to attempt to communicate in Arabic with them to test my language skills.	0.121	0.229	**0.714**
WTC 4	I ... to accompany native Arabic speakers and serve as their tour guide.	0.195	0.311	**0.620**
WTC 5	I ... to ask questions in Arabic during lectures.	-0.008	**0.790**	0.260
WTC 6	I ... to deliver presentations in Arabic in front of the class.	0.298	**0.795**	0.194
WTC 7	I ... to give presentations in Arabic at seminars.	0.314	**0.802**	0.140
WTC 8	I ... to communicate in Arabic with my campus friends.	0.207	**0.544**	0.474
WTC 9	I ... to communicate with native Arabic speakers through online media consistently.	**0.787**	0.196	0.187
WTC 10	I ... to engage in chat/direct messages in Arabic with my colleagues regularly.	**0.830**	0.208	0.209
WTC 11	I ... to provide comments in Arabic on digital platforms.	**0.772**	0.161	0.262
WTC 12	I ... to post status updates on social media using Arabic..	**0.756**	0.119	0.218

Respond: (1 = Definitely not willing, 2 = Not willing, 3 = Neutral, 4 = Willing, 5 = Definitely willing)

Based on the description of the EFA analysis results above, it is evident that the WTC construct L2 Arabic comprises three dimensions: online, classroom, and offline. The 'online' dimension, serving as the first factor, encompasses items WTC 9, WTC10, WTC11, and WTC12. The 'classroom dimension, constituting the second factor, includes items WTC5, WTC6, WTC7, and WTC8. Meanwhile, the 'offline' dimension, the third factor, incorporates items WTC1, WTC2, WTC3, and WTC4. Once each item is classified according to its respective factor and dimension, the subsequent step involves arranging the construct items based on these dimensions, followed by conducting a CFA analysis to assess the validity and reliability of the constructs.

Confirmatory Factor Analysis (CFA) for WTC Construct L2 Arabic

Having identified the three dimensions within the WTC construct L2 Arabic (online, classroom, and offline) from the EFA analysis, the subsequent step is to confirm the validity

and reliability of these constructs through CFA analysis. Based on the results of the CFA analysis, the estimation of factor loading coefficients, standard error, t-values, and notes for the validity of the item is as follows:

Table 8. Validity of WTC L2 Arabic items

Factor (Dimension)	Item	Estimate	S.E.	t-value	Note
Online	WTC9	0.83	0.040	8.01	Valid
	WTC10	0.90	0.070	15.46	Valid
	WTC11	0.82	0.073	13.60	Valid
	WTC12	0.73	0.073	11.87	Valid
Classroom	WTC5	0.74	0.066	7.20	Valid
	WTC6	0.89	0.13	9.37	Valid
	WTC7	0.83	0.12	8.93	Valid
	WTC8	0.91	0.14	8.78	Valid
Offline	WTC1	0.80	0.054	6.71	Valid
	WTC2	0.63	0.071	11.04	Valid
	WTC3	0.70	0.091	9.66	Valid
	WTC4	0.71	0.090	9.86	Valid

Note. Valid = t-value > 1.96

Table 8 reveals that each item within the WTC construct for L2 Arabic exhibits factor loading coefficients exceeding 0.60, with the item WTC2 having the lowest factor loading coefficient at 0.63. In contrast, the highest coefficient is observed in WTC8 at 0.91, and the remaining items fall within the range of 0.70 to 0.90, indicating that this construct is reliable across its constituent items. Additionally, since the t-values for each item in the construct exceed 1.96, it can be inferred that all items are valid. It implies that each item independently verifies its effectiveness in measuring the intended construct, specifically WTC L2 Arabic.

For reliability, Cronbach's alpha values are as follows:

Table 9. Reliability WTC L2 Arabic

Scale	Mean	SD	Cronbach's alpha
WTC	3.51	0.592	0.893

Table 9 shows that the Cronbach's alpha value is 0.893, indicating strong internal consistency among the WTC L2 Arabic instrument items. According to the reliability criteria outlined by Taber (2018) in educational measurement studies, this result falls under the 'reliable,' affirming the instrument's ability to consistently measure willingness to communicate in Arabic as a second language. Refer to the correlation matrix in the following image to explore the specific correlations between items within the construct.

Figure 2. Correlation Heatmap of WTC Items

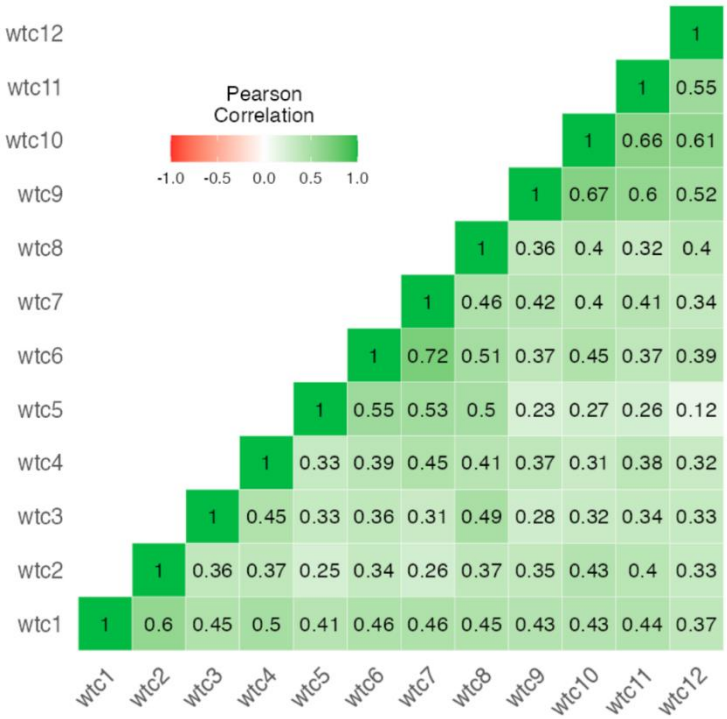


Figure 2 illustrates the Pearson correlation coefficients for each item within the construct (WTC1 through WTC12), where a correlation value of 1 indicates a perfect positive correlation, 0 indicates no correlation, and -1 indicates a perfect negative correlation. Based on Figure 2, it is evident that some items exhibit high correlations, such as WTC6 with WTC7 (correlation value of 0.72), WTC9 with WTC10 (correlation value of 0.67), and WTC6 with WTC10 (correlation value of 0.66). Additionally, many items within the construct show moderate correlations, such as WTC1 with WTC2 (correlation value of 0.6) and WTC9 with WTC10 (correlation value of 0.6). Meanwhile, a few item pairs display low correlations, including WTC1 with WTC12 (correlation value of 0.37) and WTC5 with WTC8 (correlation value of 0.23). The items with high and moderate correlations significantly contribute to the overall reliability of the construct, indicating that the WTC construct in this study is highly reliable due to the predominance of items with high and moderate correlations.

After establishing the validity and reliability of the items and the construct, the next step is to perform a Confirmatory Factor Analysis (CFA). Based on the CFA results, the model specification for the WTC L2 Arabic construct is as follows:

Table 10. Construct Model of WTC L2 Arabic Specification

Aspect	Criteria Standard	Result	Comparison	Note
Chi-Square		53.04		
RMSEA (root mean square error of approximation)	$\leq$ 0.05	0.044	0.049 < 0.05	Fit
GFI (Goodness of Fit Index)	$\geq$ 0.90	0.96	0.96 > 0.90	Fit
NFI (Normal Fit Index)	$\geq$ 0.90	0.98	0.98 > 0.90	Fit
CFI (Comparative Fit Index)	$\geq$ 0.95	0.99	0.99 > 0.95	Fit

Based on the results of the path analysis from the CFA, the following findings were obtained:

Figure 3. CFA for WTC Construct L2 Arabic

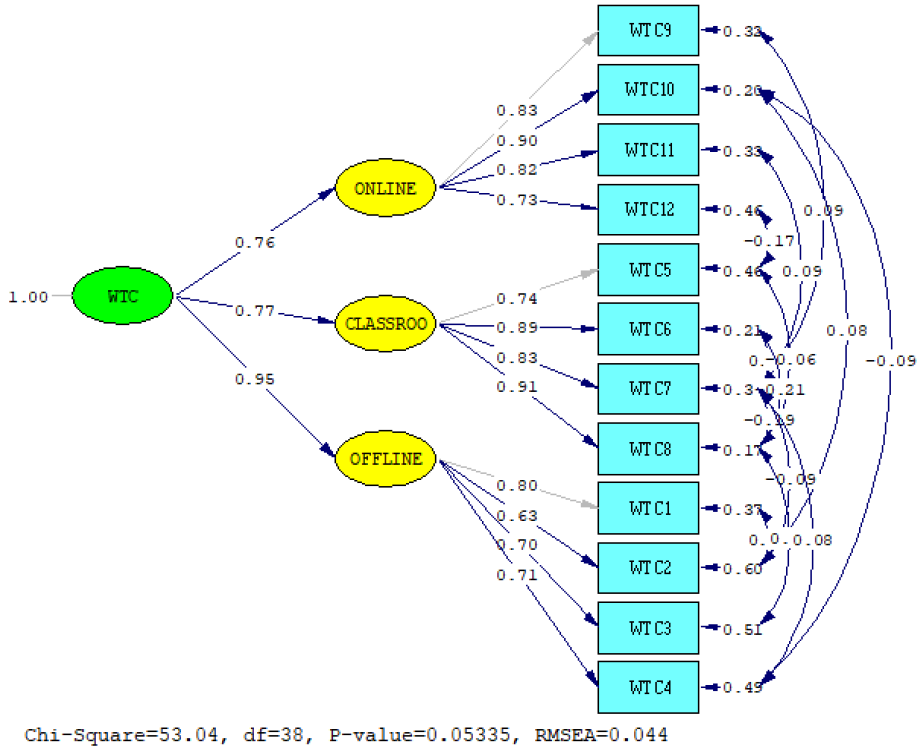


Table 10 and Figure 2 illustrate that the measurement model of WTC L2 Arabic has a Root Mean Square Error of Approximation (RMSEA) value of 0.044 < 0.05. Therefore, it can be concluded that the constructed model fits and aligns with the criteria previously presented (Browne & Cudeck, 1992; Kline, 2016; Xia & Yang, 2019). Additionally, the WTC construct model exhibits a Goodness of Fit Index (GFI) of 0.96, which is greater than 0.90; a Normal Fit Index (NFI) of 0.98, which exceeds 0.90; and a Comparative Fit Index (CFI) of 0.99, which surpasses 0.95. These indices indicate that the WTC construct model is well-fitted and capable of accurately measuring the three factors of the willingness to communicate in Arabic: online context, classroom context, and offline (real-world) context. Furthermore, the items are valid and reliable for assessing each factor.

Discussion

Willingness to communicate (WTC), as understood, denotes an individual's readiness with thorough preparation, both cognitively and affectively, to engage in communicative activities within specific situations and with specific individuals (Macintyre et al., 1998; Zarrinabadi & Tanbakooei, 2016). In Arabic language learning, WTC plays a crucial role in facilitating the development of linguistic behaviours among Arabic learners. Hence, there is a need for a valid measurement construct to assess WTC reliably among Arabic language learners.

Based on the analysis results presented in this study, a valid and reliable instrument construct was found for measuring WTC L2 Arabic. This construct consists of 12 items representing three dimensions: online, classroom, and offline (See Appendix 1). When compared with the WTC construct formulated by McCroskey & Baer (1985), the first WTC construct within a quantitative paradigm in English language communication, several

fundamental differences are evident. One of these differences lies in the type of response format within the construct. While McCroskey & Baer (1985) employed 20 items with a percentage scale response ranging from 0 to 100 (0 representing 'never,' and 100 representing 'always'), the WTC L2 Arabic construct prefers a 5-point Likert scale response format (with 1 = not willing to 5 = willing). This choice is argued based on the concept that WTC measures participants' desire and willingness, and thus, it would significantly differ when using a percentage scale from 'never' to 'always' as it primarily represents the experiences of second language learners.

Another difference lies in the dimensions measured by each construct. In the WTC construct by McCroskey & Baer (1985), various communication conditions were claimed as indicators, including public, meeting, group, dyad, stranger, acquaintance, and friend (McCroskey & Baer, 1985), where these indicators served as classifications rather than dimensions. In contrast, in this study's WTC L2 Arabic construct, dimensions were derived from EFA testing, resulting in three dimensions: online, classroom, and offline context. These dimensions are measured by four valid items each, which effectively gauge their respective dimensions.

Furthermore, compared to the WTC L2 Arabic construct utilized by Mahmoodi & Moazam (2014), it is noted that their construct comprises 20 items without distinct dimensions or factors differentiating them. Additionally, the responses utilized in their construct also take the form of a 5-point Likert scale (ranging from 1 = Strongly Disagree to 5 = Strongly Agree). This construct differs from that of the WTC L2 Arabic in this study. The rationale behind this deviation is that the responses in the WTC construct of this study do not align with those of Mahmoodi & Moazam (2014) as they primarily confirm a statement rather than representing the desire and willingness to communicate in a second language. Conversely, Mahmoodi & Moazam's (2014) construct includes statements that differentiate between expectations and desires in communication situations (see the following two items: 1. "*If I encountered some native speakers of Arabic in the street, restaurant, hotel, etc., I hope an opportunity would arise, and they would talk to me*; 2. *If I encountered some native speakers of Arabic in the street, restaurant, hotel, etc., I would find an excuse and would talk to them*.") (Mahmoodi & Moazam, 2014). Upon scrutiny, the statements "*I hope an opportunity*" and "*I would find*" emphasize different levels of effort in communication. Additionally, Mahmoodi & Moazam's (2014) construct distinguishes between native and non-native speakers, and a distinction is also adapted in the WTC L2 Arabic construct of this study. Furthermore, this study's construct adopts several that are deemed appropriate for WTC measurement with differing responses.

The next WTC L2 Arabic construct to consider is that which has been proposed by Kholis et al. (2023), which comprises three dimensions (in the classroom, outside the school, and in the digital context). In the in-classroom dimension, they adopt the WTC construct by Lee & Chen Hsieh (2019) while they develop their constructs for the other two dimensions. A fundamental difference between this construct and the WTC L2 Arabic construct in this study is the basis for constructing response options. Despite both employing the same response type (a 5-point Likert scale ranging from definitely not willing to definitely willing), the constructs by Kholis et al. (2023) dan Lee & Chen Hsieh (2019) commence with a fundamental question: '*How much are you willing to communicate in Arabic/English in this situation?*' (Lee & Chen Hsieh, 2019). In contrast, the WTC L2 Arabic construct in this study does not

begin with a question because doing so would limit the scale of willingness and desire that respondents, namely second language learners, might express. Therefore, respondents can complete item statements solely using the provided response scale.

None of the three constructs outlined above underwent validation and reliability testing, allowing further development or reconstruction into valid and reliable constructs. Based on the findings presented above, the WTC L2 Arabic construct, which consists of 12 items in this study, is based on EFA testing and measures three main dimensions: online, classroom, and offline context. Moreover, based on CFA results, each item within the WTC L2 Arabic construct demonstrates validity by effectively measuring the intended constructs, namely individuals' willingness and desire to communicate in Arabic, as per the mentioned dimensions.

CONCLUSION

Based on the findings and the discussion presented above, the WTC L2 Arabic construct in this study has been proven reliable and valid through the use of Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) methods. It effectively measures the extent of Arabic language learners' willingness to communicate in various contexts, including online, classroom, and real-world communication. Additionally, the WTC L2 Arabic construct model meets the fit criteria based on CFA testing, confirming its capability to measure learners' willingness to communicate in Arabic accurately. Understanding the WTC conditions of Arabic learners is a crucial first step toward realizing Arabic language learning that emphasizes functional language use. WTC, understood as an individual's willingness to communicate, plays a vital role in achieving the primary function and utility of the Arabic language: communication and interaction. Thus, Arabic language learning should focus on linguistic structures and support learners in expressing themselves in communicative and interactive activities using Arabic. This approach can foster effective language behavior among Arabic learners, ultimately creating a conducive linguistic environment within Arabic language learning.

The findings of this study are anticipated to aid future research in measuring the extent of Arabic learners' WTC and identifying the factors that influence WTC. Furthermore, it explores how WTC impacts Arabic learners' language skills, mainly speaking proficiency. This study has several limitations, including the restriction on the number of respondents, totaling 204 participants. This limitation suggests that with a larger sample size, the testing results of the construct might yield even more favorable outcomes. Additionally, the analysis was confined to EFA and CFA in assessing validity and reliability, which prevents the determination of the most sensitive and reliable items for measuring WTC L2 Arabic. This limitation highlights the need for future research to conduct testing using *Rasch model analysis*.

ACKNOWLEDGMENT

Any party did not fund this study, and all collaboration was solely based on personal agreements among researchers for scholarly purposes. All data and findings are scientifically and openly accountable.

AUTHOR CONTRIBUTIONS STATEMENT

[MF], as the principal investigator, is responsible for designing the grand research framework, guiding the research direction, and interpreting the research findings. [FPA] is tasked with conducting data analysis and serving as the primary research assistant. [AMA] ensures the alignment of the research process and the methodologies employed. [MRM] contributes insights into the measurement context within Second Language Acquisition (SLA). [MIM] is responsible for distributing respondent data in the study and assists with data analysis techniques.

REFERENCES

Baghaei, P., & Dourakhshan, A. (2012). The Relationship between Willingness to Communicate and Success in Learning English as a Foreign Language. *The Modern Journal of Applied Linguistics (MJAL)*, *4*(2). 53-67. Retrieved from https://www.semanticscholar.org/paper/The-relationship-between-willingness-to-communicate-Baghaei-Dourakhshan/dbcf4c3398e9490fe6bad7e516d5dfdf9f0c246b#extracted

Bonett, D. G., & Wright, T. A. (2015). Cronbach's Alpha Reliability: Interval Estimation, Hypothesis Testing, and Sample Size Planning. *Journal of Organizational Behavior*, *36*(1), 3–15. https://doi.org/10.1002/job.1960

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Cao, Y. (2011). Investigating Situational Willingness to Communicate within Second Language Classrooms from an Ecological Perspective. *System*, *39*(4), 468–479. https://doi.org/10.1016/j.system.2011.10.016

Cao, Y. K. (2014). A Sociocognitive Perspective on Second Language Classroom Willingness to Communicate. *TESOL Quarterly*, *48*(4), 789–814. https://doi.org/10.1002/tesq.155

Elahi Shirvan, M., Khajavy, G. H., MacIntyre, P. D., & Taherian, T. (2019). A Meta-analysis of L2 Willingness to Communicate and Its Three High-Evidence Correlates. *Journal of Psycholinguistic Research*, *48*(6), 1241–1267. https://doi.org/10.1007/s10936-019-09656-9

Freiermuth, M., & Jarrell, D. (2006). Willingness to Communicate: Can Online Chat Help?. *International Journal of Applied Linguistics*, *16*(2), 189–212. https://doi.org/10.1111/j.1473-4192.2006.00113.x

Jöreskog, K. G., Olsson, U. H., & Wallentin, F. Y. (2016). Confirmatory Factor Analysis (CFA). In K. G. Jöreskog, U. H. Olsson, & F. Y. Wallentin, *Multivariate Analysis with LISREL* (pp. 283–339). Springer International Publishing. https://doi.org/10.1007/978-3-319-33153-9_7

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kang, S.-J. (2005). Dynamic Emergence of Situational Willingness to Communicate in A Second Language. *System*, *33*(2), 277–292. https://doi.org/10.1016/j.system.2004.10.004

Kholis, M. N., Habibi, B. Y., Zaenuri, M., & Hasanah, L. Q. (2023). Willingness to Communicate (WTC) in Arabic for Students of State Islamic University. *Alsinatuna*, *9*(1), 51–64. https://doi.org/10.28918/alsinatuna.v9i1.2206

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.

Lee, J. S., & Chen Hsieh, J. (2019). Affective Variables and Willingness to Communicate of EFL Learners in In-Class, Out-Of-Class, and Digital Contexts. *System*, *82*, 63–73. https://doi.org/10.1016/j.system.2019.03.002

Lee, J. S., & Drajati, N. A. (2020). Willingness to Communicate in Digital and Non-Digital EFL Contexts: Scale Development and Psychometric Testing. *Computer Assisted Language Learning*, *33*(7), 688–707. https://doi.org/10.1080/09588221.2019.1588330

Lee, J. S., & Lu, Y. (2023). L2 Motivational Self System and Willingness to Communicate in The Classroom and Extramural Digital Contexts. *Computer Assisted Language Learning*, *36*(1–2), 126–148. https://doi.org/10.1080/09588221.2021.1901746

MacCallum, R. C., Widaman, K. F., Preacher, K. J., & Hong, S. (2001). Sample Size in Factor Analysis: The Role of Model Error. *Multivariate Behavioral Research*, *36*(4), 611–637. https://doi.org/10.1207/S15327906MBR3604_06

Macintyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing Willingness to Communicate in a L2: A Situational Model of L2 Confidence and Affiliation. *The Modern Language Journal*, *82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

Mahmoodi, M.-H., & Moazam, I. (2014). Willingness to Communicate (WTC) and L2 Achievement: The Case of Arabic Language Learners. *Procedia - Social and Behavioral Sciences*, *98*, 1069–1076. https://doi.org/10.1016/j.sbspro.2014.03.518

McCroskey, J. C., & Baer, J. E. (1985). Willingness to Communicate: The Construct and its Measurement. *Paper Presented at the Annual Convention of the Speech Communication Association*. https://files.eric.ed.gov/fulltext/ED265604.pdf

Peng, J.-E. (2012). Towards an Ecological Understanding of Willingness to Communicate in EFL Classrooms in China. *System*, *40*(2), 203–213. https://doi.org/10.1016/j.system.2012.02.002

Sürücü, L., Yikilmaz, İ., & Maslakçi, A. (2022). *Exploratory Factor Analysis (EFA) in Quantitative Researches and Practical Considerations*. World Peace University. https://doi.org/10.31219/osf.io/fgd4e

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Umar, J., & Nisa, Y. F. (2020). Uji Validitas Konstruk dengan CFA dan Pelaporannya. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia (JP3I)*, *9*(2), 1–11. https://doi.org/10.15408/jp3i.v9i2.16964

Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural Equation Modeling with Ordered Categorical Data: the Story They Tell Depends on the Estimation Methods. *Behavior Research Methods*, *51*(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Zarrinabadi, N. (2014). Communicating in a Second Language: Investigating the Effect of Teacher on Learners' Willingness to Communicate. *System*, *42*, 288–295. https://doi.org/10.1016/j.system.2013.12.014

Zarrinabadi, N., & Tanbakooei, N. (2016). Willingness to Communicate: Rise, Development, and Some Future Directions. *Language and Linguistics Compass*, *10*(1), 30–45. https://doi.org/10.1111/lnc3.12176

Appendix 1

WTC L2 Arabic Construct

Dimension	Item
Online	I ... to communicate with native Arabic speakers through online media consistently.
	I ... to engage in chat/direct messages in Arabic with my colleagues regularly.
	I ... to provide comments in Arabic on digital platforms.
	I ... to post status updates on social media using Arabic..
Classroom	I ... to ask questions in Arabic during lectures.
	I ... to deliver presentations in Arabic in front of the class.
	I ... to give presentations in Arabic at seminars.
	I ... to communicate in Arabic with my campus friends.
Offline	If I encounter a native Arabic speaker on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to engage in conversation with them.
	If I meet someone who speaks Arabic (non-native) on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to converse with them.
	If someone introduces me to native Arabic speakers, I ... to attempt to communicate in Arabic with them to test my language skills.
	I ... to accompany native Arabic speakers and serve as their tour guide.
	If I encounter a native Arabic speaker on the street, in a restaurant, hotel, or other public places, I ... to seek opportunities to engage in conversation with them.

Respond: (1 = Definitely not willing, 2 = Not willing, 3 = Neutral, 4 = Willing, 5 = Definitely willing)