Original Article

Study of Validity and Reliability of the GADOT Personality Types Determination Scale Developed Based on Enneagram

Zeynep Gungormus, PhD
Gaziantep Islamic Science and Technology University, Faculty of Health Sciences, Department of Nursing, Gaziantep, Turkey

Iknur Aydin Avci, PhD
Ondokuz Mayıs University, Faculty of Health Sciences, Department of Nursing, Samsun, Turkey

Hasan Dagıdır, PhD Student
Samsun Ilkadim Yesilyurt Demir Celik Primary School, Samsun, Turkey

Oya Sevcan Orak, PhD
Ondokuz Mayis University, Faculty of Health Sciences, Department of Nursing, Samsun, Turkey

Necla Afyonkale Talay, PhD
Private Psikoodak Psychology and Family Counseling Center, Samsun, Turkey

Correspondence: Dr. Oya Sevcan Orak. Ondokuz Mayıs University, Faculty of Health Sciences, Department of Nursing, Department of Psychiatric Nursing, Turkey. E-mail: oysev@hotmail.com

Abstract

Objective: The aim of this research is to conduct an Arabic validity and reliability study of the "GADOT Personality Types Scale" developed based on Enneagram to determine the personality types of those belonging to the Arab culture of immigrants living in Turkey.

Method: In this methodological study, data were collected between 1 June and 1 September 2021. The universe of the research consists of immigrant individuals living in Turkey. Using the snowball sampling method, one of the most improbable sampling methods, 500 immigrant individuals who met the inclusion criteria formed the sample of the study. The developed scales are of the 5-point likert type and consist of 64 items. 1.type 9 substance, 2. and 4. type 5 items, 3., 5. and 7.types 7 items, 6., 8. and 10.it is in the form of type 8 substance. The height of the scores taken from the scales shows that the personality type measured by the relevant factor is strong in the person. IBM SPSS 22 and AMOS 22 programs were used to evaluate the data. In order to determine the validity of the scales; scope validity, criterion validity and structure validity (exploratory and confirmatory) were performed.

Results: As a result of the explanatory factor analysis, it was found that all items showed a significant correlation with each other in the nine sub-dimensions of the GADOT Personality Type Determination Scale (p<0.05). Kaiser-Meyer-Olkin (KMO) and Bartlett's Test results, which indicate the appropriateness of the GADOT Scale for factor analysis and sample adequacy, were found to be significant, since p(sig)=0.000 on all scales.

Conclusions and Recommendations: In this research, a personality types identification scale has been developed that evaluates nine personality types based on Enneagram separately. As a result of the analyzes, it was found that the scale models are suitable for their conceptual structure, valid and reliable.

Keywords: Enneagram, Personality Types, Scale Development

Introduction

Personality is an established thought, feeling and behavior pattern that is consistent with lifestyle developed in a conscious and unconscious way through the experiences, development and structural characteristics of a person through their social life. With all of these characteristics, personality reflects the characteristics, beliefs and behaviors that differentiate individuals from others (Guney, 2018; Acarkan, 2019). Thus, even though people have similar biological structures, their reactions and behaviors may differ. Scientists have conducted studies and developed a range of classification systems to explain people's similarities and differences and to enable

individuals to understand themselves and each other (Jung, 1971; Tieger et al., 2014). In this way, they have facilitated the process of understanding and measuring personality by collecting thousands of personality traits under a smaller number of main factors (Piskin, 2011). Even though the number of personality types changes in the classification systems that the researchers do, what they have in common is that the system allows understanding the behavior, limitations, capabilities, preferences and trends of the individual with any personality type (Yildiz et al., 2012). Each type of personality has strengths and weaknesses and no type of personality is better than the other and describes behavioral patterns of all types of individuals (Piskin, 2011). Knowing the type of personality also makes it easier to predict the future behavior of individuals.

One of the models generated to determine personality traits in literature is the ''Enneagram Personality Model.'' An Enneagram Personality Model is a well-known and ancient personality typology that features nine different personality types (Matise, 2019; Sahin et al., 2020). The conducted studies have highlighted the parallels of the Enneagram system with other contemporary psychological theories, such as the humanitarian, psychodynamic, cognitive-behavioral, developmental approaches of psychology and the field of neuroscience (Matise, 2019; Heuertz, 2020; Hook et al., 2021). It is noted that the number of forms based on the Enneagram Model on the Internet is large, while the number of studies in which validity and reliability studies are conducted is more limited. The use of Enneagram in psychometric evaluations has not become widespread due to lack of appropriate measurement tools, and in recent years, the use of Enneagram-based scale development has contributed to this field (Demir, 2020; Demir et al., 2020a; Demir et al., 2020b; Galves et al., 2021; Kastelein, 2021; Tastan, 2019). After examining the literature, it is concluded that there are various scales that can be accessed in foreign sources, that are based on the Enneagram Model and that have been reviewed for validity (Demir et al., 2020; Galvez et al., 2021). Similarly, among the available resources, the scales developed based on the Enneagram Model, whose sample consists of Turkish-speaking participants, for which a validity and reliability study has been conducted,

have been increasing in recent years (Tastan, 2019; Sirin, 2020; Demir et al., 2020).

Due to the recent conditions in the Middle East, Turkey has been hosting many immigrants. Immigrants experience a series of difficulties before, during and after migration, in areas such as unemployment, change of social roles, environmental interaction, violence, uncertainty, acceptance, asylum, cultural harmony and need support according to their individual differences in countries they migrate as a risk group in terms of spiritual problems (Kirmayer et al., 2011). Immigrants are mostly referred to by the host society as ''foreign'' and individual differences related to personality types can be ignored (Nizam & Gul, 2019).

These problems bring along many cultural, social, health, education and economic issues regarding migrants, which are not independent in an interactive process. It is important that many migrants are employed and able to find work in order to meet economic needs. Awareness and recognition of one's own qualities is an important factor in an individual's ability to choose a profession/job within the current conditions. Familiarity with individuals is also important in selecting individuals that fit the job description in the institutions and organizations that will provide employment (Kale & Shrivasta, 2003). We believe that choosing and/or employing a profession/job suitable for one's own personality traits will contribute in the avoidance of losing economic investments such as work performance, job satisfaction, career development, employee training required for the job and purchasing equipment. Piskin (2011) states that individuals have personality traits that they require in their occupations and that the overlapping of personality traits required by the occupation/jobs will increase both individual productivity and job satisfaction.

Even though determining the types of personality has beneficial effects for everyone, migrants make up a more special group on the importance of this issue. It is believed that identifying the types that represent the personalities of migrants will provide self-awareness, help society to make sense of the differences, and provide benefits in many areas such as healthcare services, education and employment. The most reliable way of making this definition is to develop a validated and reliable measuring tool that can identify the personality types of the target group.

The purpose of this study is to provide Arabic validity and reliability studies of the ''GADOT Scale of Personality Types'', which is intended to determine the personality types of migrants living in Turkey from Arab culture.

Method

The Target Group of the Study and Sampling of the Research: The target group of this methodological research is comprised of the migrants living in Turkey. Data were collected from June 1 to September 1, 2021 by using the snowball sampling method, which is one of the improbable sampling. 500 migrants were included into the target group of the study. The inclusion criteria are those individuals who accept to participate in the study, who are 18-65 years of age, who have the technological equipment to access Google forms, and who live in Turkey. Literature states that individuals should be 5-10 times more than the scale items in scale development studies (Erdogan et al., 2015). As there are 66 items in this scale, the aim was to reach 330-660 migrants and the study was completed with 500 people who agreed to participate in the study. The research consisted of the stages of generation of a item pool, preparation of a draft scale and submitting it to expert opinion, conducting a pilot study of the scale, collecting and analyzing data after a review of the literature on the generation of scale items.

Generation of the Item Pool: During the generation phase of the item pool, the researchers reviewed domestic and international measurement tools developed based on Enneagram, characteristics of the personality types in Enneagram, strengths and weaknesses, studies on Enneagram, and theoretical information on the characteristics of migrants. The generation of the item pool was carried out under the supervision of a senior researcher who had previously conducted studies on immigrants, having methodological research experience. In generating the scale items, measurement tools and items based on Enneagram were reviewed (Demir, 2020; Demir et al., 2020a; Demir et al., 2020b; Galves et al., 2021; Kastelein, 2021; Tastan, 2019). As a result, 66-items, five-point likert-type scale expressions were generated to determine the Enneagram-based personality types of immigrant individuals.

Data Collection Tools

Introductory information form: The questionnaire contains 14 questions on the age, gender, marital status, educational status, employment status, economic status, and the socio-demographic information of migrants. The questions also include information about the years they have lived in Turkey, their current legal status, and their state of psychological problem after they come to Turkey.

GADOT identification scale of personality types: The scale was originally made up of 66 items and there was no need to make any changes in the scale items after the pilot application. The scale consisting of 66 items was applied to 500 people, 2 items with a factor load below .30 were removed from the scale and a scale of 9 dimensions and 64 items was obtained to determine the 9 personality types. The scale, prepared in a five-likert type, is scored as follows: 1=It does not define me at all, 2= It defines me a little, 3= It defines me moderately, 4= It defines me quite a lot, 5= It defines me completely. The scale has no overall score. The total score of each dimension is evaluated on the scale. Each sub-dimension represents a type of personality, and each type of personality can be used together or separately. The developed form for determining the personality types specified in the Enneagram is classified as follows and the characteristics of these types are defined. Type 1: The Reformist/Perfectionist (9 items), Type 2 : The Helpful (5 items); Type 3: The Achiever (7 items), Type 4: Individualist(5 items); Type 5: Researcher (7 items); Type 6: Inquisitive (8 items); Type 7: Curious (7 items); Type 8: Challenging (8 items) and Type 9: Peaceful (8 items) (Riso & Hudson, 2018; Hook et al., 2021). Total scores are evaluated for each dimension of nine personalities. Each of the nine items at Enneagram represents a personality type. This type of personality describes the individual better than others and is called the basic personality type of the individual. The basic type of personality of an individual is a result of all the factors that make up the personality, including the genetic factors. Individuals do not change from one basic type of personality to another, the definition of personality types are universal and gender-independent. Not everything in the personality type is always accurate because every personality has healthy, normal and unhealthy situations. Everyone is a unique blend of the basic personality type and the two personality types adjacent to it at Enneagram. The two personality types adjacent to the basic personality type are called ''wing''. The wings complete the basic type of personality and constitute a ''second aspect" of personality (Riso & Hudson,

2018). The main type is the dimension where the highest score is obtained from the sub-dimensions indicating the personality types on the scale, and the other two personality types with the highest score are wing personality types. There is no total score of the scale and the score that is obtained from the dimensions assigned to each personality type is used for evaluation.

Data Collection: Data collection was performed in two stages. A pilot application was done in the first stage, and a validity-reliability study was conducted in the second stage. In the pilot application, a scale was applied by meeting with 50 migrants face-to-face, and then the second stage of the validity-reliability study was initiated as there was no need for revision. At this stage, 500 volunteers filled the scale. The scale form was prepared in Arabic and converted into Google Form. In terms of Arabic language, the forms were sent to three language specialists for evaluation. One of these linguists is a Syrian national academic and the other two are Saudi Arabian nationals. The announcement of the study was made via social media networks such as WhatsApp, Facebook, Instagram and Twitter, and those who met the selection criteria of the migrants who would participate in the study completed the study by filling out the form prepared with the approval of experts.

Scope Validity: The scale was originally formulated by using 66-items expressions, and was then sent to specialists including three psychologists, three psychiatry doctors, two public health nursing professors, one Enagram consultant and two psychiatric nursing professors. There was no reduction in the number of items after receiving a notification from experts, and the contents of the statements were revised. The revised form was translated into Arabic and submitted to expert opinion. No revision was needed in Arabic. Compliance between experts was calculated as CVI 0.80.

Data Assessment: The IBM SPSS 22 and AMOS 22 programs were used in the evaluation of the data. Scope validity, criterion validity, and structure validity (descriptive and confirmatory) were used to determine the validity of the scales. Scope validity was based on CVI value above 0.80. For criterion validity, internal criterion validity was assessed through sub-top group comparisons and item-total correlations. For structure validity, Kaiser Meyer Olkin and Barlett tests, Eiguen values, Total variance explained and component matrix were examined in the exploratory factor analysis. In the

confirmatory factor analysis, the factor loads of items, X2/sd, NFI, TLI, CFI, IFI, RFI, GFI, AGFI, RMSEA, and RMR compliance values were reviewed. In order to determine the reliability of the scales, we performed internal consistency reliability coefficients (cronbach alpha), the average of correlation coefficients between the items, item-total score correlation (item separation index), scoring consistency (Intra-class correlation), SEM (certainty of scales in measurement), bottom-top effect analysis (homogeneity of scale) and Hotelling's T^2 tests.

Research Ethics: Prior to starting the research, the permission of the ethics committee was obtained from the Health Sciences Department of Non-Interventional Research Ethics Committee with decision number 2021/44. The decision about accepting to participate in the study was questioned as the first question of the questionnaire. Additionally, the first page of the Google form contains general information on the purpose of the study, confidentiality of the information, and participation in the research, and we asked for the consent of the participants. The individuals participating in the research completed the other sections of the form stating their acceptance. Thus, the research was carried out on the basis of the approval in terms of voluntariness and receiving consent. The research and publication ethics principles were followed at all stages throughout the study. All phases of the research were conducted in accordance with the ethical principles stated in the Helsinki Declaration (2013).

Results

Structural Validity

Exploratory factor analysis: Due to the fact that the value of Kaiser-Meyer-Olkin (KMO) and Bartlett's Test results, which indicate the suitability of the GADOT Scale for factor analysis and adequate sampling is p(sig)=0.000, it found to be significant for all scales (Table 1). On the correlation matrix, all items were determined to correlate significantly with each other (p<0.05). **Confirmatory factor analysis:** As a result of the exploratory factor analysis of GADOT Scale, the structure validity of the scales was tested by confirmatory factor analysis. The values obtained from the compliance indices are shown in Table 2. Acceptable fit values in the literature are as follows (Civelek, 2018; Yaslıoglu, 2017; Weng et al., 2021; Karagoz, 2016; Simsek, 2007; Ozdamar, 2016):

$0 < X^2/SD \leq 3, \leq 5$; $0.90 < NFI \leq 1.0$;

$0.90 \leq TLI$ (NNFI)≤ 1.0; $0.90 \leq CFI \leq 1.0$; $0.90 \leq IFI \leq 1.0$; $0.85 < RFI \leq 1.0$;

$0.90 <$, $0.85 \leq GF \leq 1.0$; $0.85 <$, $0.70 <$, $0.80 \leq AGFI \leq 1.0$; $0.00 < RMSEA < 0.10, < 0.08$;

$0.00 < SRMR < 0.10$.

Type 1: Figure 1 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.33 and 0.75 (p=0.000) (Figure 1).

Type 2: Figure 11 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.49 and 0.81 (p=0.000) (Figure 1).

Type 3: Figure 12 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.40 and 0.81 (p=0.000) (Figure 1).

Type 4: Figure 13 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.36 and 0.77. There is a weak relationship (r=.23) between e1 and e2 items (p=0.000). The relationship between e1 and e2 highlights a similar characteristic in that it reflects the personality trait ''which cares about the emotions of its own and those around him/her'' (Figure 2).

Type 5: Figure 14 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.48 and 0.65. There is a weak relationship (r=.37) between e1 and e2 items (p=0.000). The relationship between e1 (Others trust my information) and e2 (I especially care about mastering on the field of my interest to the fullest) highlights the similar characteristic due to the fact that it reflects the ''knowledgeable'' personality trait (Figure 2).

Type 6: Figure 15 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary

significantly between 0.45 and 0.67 (p=0.000) (Figure 2).

Type 7: Figure 16 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.43 and 0.67 (p=0.000) (Figure 3).

Type 8: Figure 17 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.52 and 0.72. There is a weak relationship (r=.30) between e3 and e4 items (p=0.000). The relationship between e3 (I am a dominant person who is listened to in the environments I enter) and e4 (I focus on the works that make me feel strong and practical) emphasizes a similar trait in terms of reflecting the ''dominant'' and ''strong'' personality traits (Figure 3).

Type 9: Figure 18 indicates the first level diagram of the model resulting from the DFA analysis of the scale. In the DFA analysis, the correlation coefficients of the items vary significantly between 0.49 and 0.72 (p=0.000) (Figure 3).

Validity of the criteria: In order to examine the validity of the scales, the internal validity of the criteria was evaluated with bottom-top group comparison and item - total correlations. 27% Bottom -27% (n=135) Top group comparison and the average difference between the slices of 27% with the lowest and 27% with the highest score of the distribution in the total score is given in table 4. It was determined that the bottom groups of the scales had significantly higher averages than their top groups. These scales were observed to be the scales that differentiate between top groups of 27 %, and 27 %, respectively (Table 3).

Reliability: When the internal consistency reliability coefficients of the scales (cronbach alpha) are examined, the relevant values were determined as follows: Type 1; α=.842, type 2; α=.832, type 3; α=.821, type 4; α=.695, type 5; α=.749, type 6; α=.795, type 7; α=.762, type 8; α=.835, type 9; α=.825 (Table 4). The mean of the correlation coefficients between the items was found as follows: Type 1; .373, type 2; .501, type 3; .399, type 4; .314, type 5; .301, type 6; .328, type 7; .314, type 8; .388, type 9; .373 (Table 4). When the item-total score correlation

(Item separation index) values are examined, it is concluded as follows: Type 1; between .326-.658, type 2; .454-.708 between, type 3; .358-.708 between, type 4; .375-.550 between, type 5; .407-.530 between, type 6; .412-.584 between, type 7; .398-.560 between, type 8; .474-.620 between, type 9; .452-.623. It was determined that the correlation coefficients were significant for all scales in terms of criterion validity (p=0.000).

Scoring consistency: When the intra-class correlation is evaluated, the relevant values were determined as follows: Type 1; .842, type 2; .832, type 3; .821, type 4; .695, type 5; .749, type 6; .795, type 7; .762, type 8; .835, type 9; .832 (Table 4).

Standard error: SEM is equal to the square root of the difference of the reliability coefficient to one multiplied by the standard deviation [SEM= (ss x $\sqrt{(1-\alpha)}$]. The fact that the SEM is less than or equal to half of the standard deviation is taken as an acceptable measure of accuracy in the measurement (Pontes & Griffiths, 2015). In our study, it was taken into account that the recommended SEM≤S/2 in the accuracy of the scales in terms of the measurement and their compliance with the equation was observed: Type 1; .3.14<3.95, type 2; 2.09<2.55, type 3; 2.83<3.35, type 4; 2=2, type 5; .2.9=2.9, type 6; 3.12<3.45, type 7; 2.93<3.01, type 8; 2.88<3.56, type 9; 3.48<4.25'tir.

Bottom-top effect analysis: It was determined that the scales did not exceed 15 %of the bottom-top effect, and the homogeneity of the scales was ensured (Table 4).

Hotelling's T2 test values indicate that the difference between the averages of the items is significant (p<0.05) (Table 4). The number of items of the GADOT scale, alpha values, mean and standard deviations of scale, the covariance and correlations between items, Hotelling's T^2 values and bottom and top impact percentages are presented in Table 4.

Table 1. Findings of KMO ve Bartlett's Test of GADOT Personality Types Scales based on Enneagram

Personality Types	KMO	Bartlett's Test[*]	Eigenvalues	Explained Variance	Component matrix scores
Type 1: The Reformer	.904	1372.404	4.063	45.145	.411-.771
Type 2: The Helper	.833	949.552	3.032	60.634	.606-.838
Type 3: The archiever	.869	1091.108	3.455	49.357	.479-.822
Type 4: The individualist	.738	399.845	2.271	45.428	.584-.769
Type 5: The investigator	.800	675.639	2.822	45.428	.556-.704
Type 6: The Loyalist	.853	924.460	3.327	41.584	.576-.830
Type 7: The Enthusiast	.821	696.337	2.905	41.493	.545-.724
Type 8: The Challenger	.869	1195.866	3.735	46.682	.598-.733
Type 9: The Peacemaker	.864	1122.986	3.630	45.377	.574-.744

Table 2. Confirmatory Factor Analysis of GADOT Personality Types Scales based on Enneagram

| Types	Compatibility values									
	X^2/SD	NFI	TLI	CFI	IFI	RFI	GFI	AGFI	RMSEA	SRMR
Type 1	2.660	.948	.956	.967	.967	.931	.967	.945	.05	.06
Type 2	4.312	.977	.965	.982	.985	.955	.984	.952	.08	.03
Type 3	3.746	.952	.946	.964	.965	.928	.970	.939	.07	.07
Type 4	1.515	.985	.987	.995	.995	.962	.995	.982	.03	.03
Type 5	3.850	.926	.909	.944	.944	.881	.972	.939	.07	.08
Type 6	4.463	.904	.893	.923	.924	.866	.955	.919	.08	.09
Type 7	4.836	.903	.882	.921	.922	.855	.963	.925	.08	.09
Type 8	4.405	.931	.919	.945	.945	.898	.958	.920	.08	.07
Type 9	4.979	.912	.899	.928	.928	.877	.952	.913	.08	.08
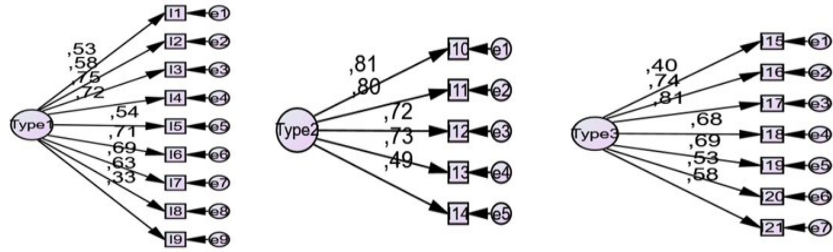
Figure 1. Type1, 2, and 3 CFA results
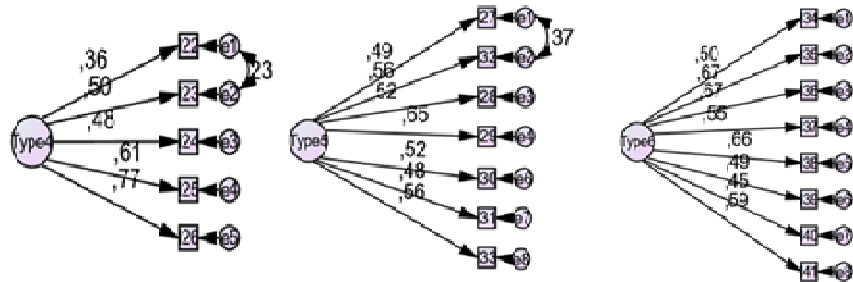
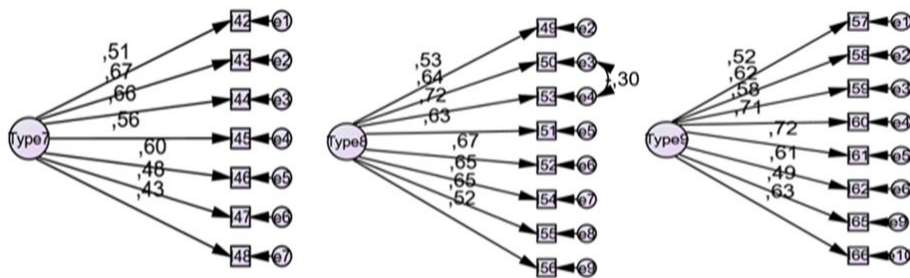Figure 2. Type 4, 5, and 6 CFA results

Figure 3. Type 7, 8, and 9 CFA results

Table 3. Comparison of the 27% lower and 27% upper groups of the scales

Scales	Scale Mean±SD	27% Upper group (Mean±SD)	27% Subgroup (Mean±SD)	t	p
Type 1	30.7±7.9	37.7±4.4	22.3±7.6	24.230	.000
Type 2	17.6±5.1	23.4±1.2	10.8±3.3	66.347	.000
Type 3	22.9±6.7	30.6±2.4	14.2±3.9	94.167	.000
Type 4	15.5±4.5	30.6±2.4	9.7±2.5	244.789	.000
Type 5	22.44±5.9	27.1±4.5	17.0±5.8	18.629	.000
Type 6	26.31±6.9	32.3±4.3	19.0±6.1	26.063	.000
Type 7	22.0±6.0	25.5±5.5	17.9±6.5	11.106	.000
Type 8	26.10±7.1	35.4±2.8	18.9±3.5	72.417	.000
Type 9	27.0±7.2	35.2±2.8	17.9±5.1	67.490	.000

Table 4. Findings on reliability

Scales	Cr α	Inter-Item Correlations	Intraclass correlation	Bottom (%)	Top (%)	Hotelling's T^2
Type 1	.842	.373	.842	3	1.8	195.877[*]
Type 2	.835	.501	.832	4.8	6.8	81.559[*]
Type 3	.821	.399	.821	3.4	2.8	90.142[*]
Type 4	.695	.314	.695	3.6	2.8	208.404[*]
Type 5	.749	.301	.749	3.2	1.4	127.714[*]
Type 6	.795	.328	.795	3.0	1.4	140.195[*]
Type 7	.762	.314	.762	3.6	2.0	84.305[*]
Type 8	.835	.388	.835	3.2	2.6	137.053[*]
Type 9	.825	.373	.832	3.4	3.6	98.610[*]

[*]p=0.000

Discussion

This study, conducted on the basis of methodological design, was developed for the first time as a total of nine scales determining each of the nine personality types based on Enneagram. As a result of the analyses, it was determined that the scale models are appropriate, valid and reliable according to the conceptual structure.

Discussion Related to the Findings on the Validity of the Scale: Validity is an indicator of how accurately the scale measures the intended characteristic without mixing it with other features. Scope validity, criterion validity and structure validity analyses are most commonly used in relation to validity (Buyukozturk et al., 2017). In this study, scope, structure and internal criterion validity were used.

Scope validity (SV): The validity of the scope determines whether the qualities of the items that make up the test are adequate in terms of quantity and quality in measuring the desired behavior. In this context, the academicians studying on the relevant field were contacted and an expert opinion was obtained so as to ensure the validity of the scale in terms of scope. Polit's and Beck's (2006), SV was used. According to this, items of scale are sent to at least 10 specialists for their opinions. Experts are asked to evaluate each item on the scale between 1 and 4 points. 1 point is assessed as ''not suitable'', 2 points is assessed as ''somewhat suitable and the item needs to be turned into a suitable structure'', 3 points is assessed as ''reasonably suitable, but small changes are required'' and 4 points is rated as ''very suitable''. Each of the items is expected to be scored as 3 points and 4 points.

The number of experts giving 3 or 4 points to each item is divided by the total number of experts and thus SV is calculated for both the item and the scale. The items deemed appropriate by the expert academicians were taken and the necessary arrangements were made on the items that the expert academicians expressed their opinion on changing or regulating. In addition to the scope validity, expert opinions were taken in order to ensure the suitability and comprehensibility of the items in terms of language and suitability to the target audience, and the items were revised in line with the recommendations of experts. The value must be 0.80 or over so as to ensure sufficient scale and item reliability (Polit & Beck, 2006). The study found that since SV values of all scales and items were 0.80 and above, the scales were valid in terms of reliability.

Structural validity: Factor analysis is used to determine the structural validity. When the data is suitable for factor analysis for the validity of the structure, it is decided by the fact that the Kaiser-Meyer-Olkin (KMO) coefficient is higher than 0.60 and the Bartlett test is significant (Buyukozturk, 2018). According to the KMO coefficient, sample size is at very good levels (Karagoz, 2016). According to Barlett test values, there are high correlations between the variables in the scales, data comes from multiple normal distributions and sample size is sufficient for the scales. Accordingly, KMO and Barlett tests were determined as suitable for factor analysis (Buyukozturk, 2018).

Factor analysis is carried out in two ways as confirmatory and exploratory factor analysis.

DFA: This is the type of procedure that researchers use when they want to test a theory/model they have developed in their mind. In this research, the researchers wanted to test the nine-persons models developed based on the nine-person features of the Enneagram. The CMIN/DF (X^2/sd) value is the most basic measurement tool used to test the overall suitability of the model. If this value is less than or equal to 5, the model has goodness of fit (Yaslıoglu, 2017). This value is used to evaluate whether there is a difference between the sample covariance matrix and the covariance matrix adapted (modeled) by the model. The fact that the values of all the scales in this study are below five points suggests that there is general suitability of the models.

Apart from CMIN/DF, several different fit indexes are used to test the suitability of the model. The following are among the goodness of fit indexes used in the researches to evaluate the suitability of the model in the DFA researches : Normed Fit Index, , non-normed fit index in LISREL and NNFI-TLI specified in Tucker Lewis Index in AMOS ,Comparative Fit Index-CFI, Incremental Fit Index-IFI), Relative Fit Index- RFI, Goodness of Fit Index- GFI, Adjustment Goodness of Fit Index-AGFI), Root Mean Square Error of Approximation- RMSEA), Root Mean Square Residual- RMR. The model being tested needs to be ranked among the fit indexes stated in the literature (Table 2). The fit values of all scales were found to be in accordance with the literature (Civelek, 2018; Yaslıoglu, 2017). The models tested for these fit indexes were validated.

AFA: Since our scale is not an on the basis of total score index, but a profile inventory, each sub-scale was subjected to factor analysis within itself. Furthermore, the Varimax rotation was used in order to evaluate how many factors each sub-scale revealed. Inventory sub-scales were generated by considering items collected under a single factor, Eiegen values exceeding 1, and collected under a single factor. The component matrix values of the items of the sub- scales are .411 and above, and all items in the correlation matrix are significantly correlated with each other.

Criteria validity: Criteria validity is divided into two parts as internal criteria validity and external criteria validity. Since there is no similar scale in this study which separately researches Enneagram-based personality types, internal criteria validity was used. Within the scope of the internal criteria validity, the differentiability of the scale and the item-total score correlations with were analyzed with a 27% bottom - 27% top group comparison.

27% bottom - 27% top group comparison: since there is a significant difference between the item point averages of the bottom 27% and top 27% groups, it is expected that the scales will be able to distinguish the two opposite groups from each other (Erkus, 2019). In this context, it was evaluated that the scales are such scales that can distinguish the two end groups from each other.

Item-total score correlations (item separation index values) can be reviewed for validity of internal criteria. Item total correlations are the

relation between the total value of each item in the scale and the total value received from the entire scale (Erkus, 2019). The item-total correlation coefficient is desired to be minimum .20 or .25 (Tavsancıl, 2002). Buyukozturk (2018) determined that the items with a total correlation of 0.30 and over distinguish people well. Due to the fact that the item-total correlation coefficient values are above .30 in this study, all scales are considered as valid according to this analysis in terms of criteria.

Discussion of the Findings on the Reliability of the Scale: For the structural validity and reliability study of the scale, item analysis and reliability analysis of the items in test form as pilot study was carried out. Within this context, the values such as scale marker statistics, reliability coefficient in accordance with the data structure (Cronbach Alpha), reliability coefficient in case of deleting the question and total correlation of the item were examined.

Among the most widely used methods for assessing the reliability of the scales, internal consistency analysis (Cronbach's Alpha), correlations between items, item-total score correlation, scoring consistency, standard error, Hotelling's T^2 and bottom - top impact analysis were used.

Cronbach's alpha (α) coefficient indicates whether items measure the same property and whether items are related to the subject to be measured and participants circled the items of the scale by understanding. This value is expected to be as close to 1 as possible. The value (0.60) - (0.80) indicates that the scale is reliable and the value (0.80) - (1.00) indicates that the scale is highly reliable (Karagoz, 2016) Cronbach's alpha values of GADOT scales were determined to be between 0.695 and 0.835 (Table 4). Simsek (2007) stated that a reliability coefficient of 0.70 and above is generally an indication of the reliability of the scale. In this context, the scales that were analyzed in terms of reliability are concluded to be reliable in this study.

Average of item - total score correlation coefficients: It indicates internal consistency (Sencan, 2005). The overall item- total score correlation explains the relation between the scores obtained from the test items and the total score of the test. A high item-total correlation indicates that the items measure similar behaviors and the internal consistency of the test is high. In

general, the items with an item-total correlation of 0.25 and higher may be considered to distinguish individuals well (Karagoz, 2016). However, there are experts who stated that it would be appropriate to remove the items with an inverse correlation between them and the items with a total item correlation of 0.10 and below from the inventory of the scale (Ozdamar, 2016), and there are also experts who state that removing items from the inventory whose item total correlation values are too close to zero may be a more priority option. This analysis also examines the changes in the confidence coefficients (Cronbach's Alpha if item is deleted) of the scale. It can be decided that, if there is a more than 5 % increase in the Cronbach Alpha coefficient when the item is deleted from the inventory, it shall be appropriate to remove it from the inventory (Ozdamar, 2016). Internal consistency analysis performed in relation to the reliability of the scale in the item analysis conducted within the scope of this information revealed that the correlation coefficients are significant for all scales, in terms of criteria validity (p=0.000). The scale with the highest level of relationship between the total score of the items and the scale is the second type personality scale. This value can be considered as having sufficient internal consistency after the initial application. After determining that the internal consistency value is sufficient, the columns ''Corrected Item-Total Correlation'' and ''Cronbach's Alpha if Item Deleted'' were examined on the outputs obtained for item analysis. The items with a value of 0.20 and below and the items with negative correlation values were removed from the inventory. After these questions were removed, Cronbach's Alpha values were found to be within the range of 0.695-0.835 in the analysis. After the analysis, according to the results of the item-total correlation, it was observed that, there is not an item the value of which is below 0.20, and that there is not an item that can change the reliability coefficient by 5 % or more after removing the item. As a result, the items that were determined to have a low contribution to the test were removed from the test (items 63 and 64) and it was decided to continue the analyses.

Average of correlation coefficients between the items: The average correlation coefficients of the items indicate internal consistency. It shows the extent to which scales are related to each other (Sencan, 2005). The average of correlation

coefficients between the items was found as follows: Type 1; .373, type 2; .501, type 3; .399, type 4; .314, type 5; .301, type 6; .328, type 7; .314, type 8; .388, type 9; .373 (Table 4). The scale with the highest level of relationship of items with each other is the third type of personality scale.

Scoring consistency: For numeric measurements, it is assessed by examining the correlation within the class. It also provides information about reliability in terms of structural properties (Mehta et al., 2018). The consistency values between the two measurements indicate that the values between 0.60-0.80 mean good and the values over 0.80 is excellent in terms of reliability (Erkus, 2019). In the study, it was determined that the consistency of the scales was at a good level. The related values were determined as follows: Type 1; .842, type 2; .832, type 3; .821, type 4; .695, type 5; .749, type 6; .795, type 7; .762, type 8; .835, type 9; .832 (Table 4).

Standard error: It was considered to be an important measure by Pontes and Griffiths (2015), which reflect the level of stability of the scales as a result of measurement errors. The SEM equals the square root of the difference of the confidence coefficient to one multiplied by the standard deviation [SEM= (ss x (1-α)]. The fact that SEM is less than or equal to half the standard deviation, then this is considered as an acceptable accuracy measurement (Pontes and Griffiths, 2015). In our study, it was taken into account that the recommended SEM≤S/2 value is provided in terms of the accuracy of the scales in the measurement and it was decided that all the scales were stable in accordance with their compliance with the equation (Pontes & Griffiths, 2015).

Hotelling's T^2 analysis: When whether the students filled in the scales according to their own views or based on the pressure of the society or the researcher was examined through Hotelling's T^2 test (Sencan, 2005), it was determined that there was no response bias and the difference between the question averages was significant.

Bottom - top impact analysis: The bottom and top effects of the scales are examined by their percentage of the lowest and highest scores in total scores. According to Pontes and Griffiths (2015) it is undesirable that the ratio of the bottom and top be above 15%. This means that there is a bottom and top effect. When the distribution of the obtained scores is examined, it is concluded that the ratio of bottom and top effect of all scales is less than 15%, and that there is no bottom and top effect on the distribution of scores (Pontes & Griffiths, 2015).

Conclusion: As a result of the validity and reliability study of the GADOT Personality Types Determination Scale, it was determined that the scale is suitable for use in immigrants belonging to Arab culture. The developed scales are comprised of 64 items and 5 point likert scale of the following types: 1=It does not define me at all, 2= It defines me a little, 3= It defines me moderately, 4= It defines me quite a lot, 5= It defines me completely. 1. type consists of 9 items (min-max: 9-45), 2. and 4. type; 5 items (min-max: 5-25), 3., 5. and 7. types ; 7 items (min-max: 7-35), 6, 8. and 10. types; 8 items (min-max: 8-40). The degree of the scores obtained from the scales indicates that the personality type measured by the relevant factor is strong.

The use of the GADOT Personality Types Determination Scale developed within the scope of this research in immigrants with Arab culture will help to determine the more dominant personality types in individuals and to understand individual differences. Considering that the differences of immigrant individuals are associated with being a "foreigner", it is thought that the use of this scale will contribute to understanding the individual's behavior and providing individualized care, especially in psychosocial nursing practices.

References

Acarkan, I. (2019). Discover Yourself with Enneagram. 2nd Edition. Timas Publications: Istanbul.

Buyukozturk, S., Cakmak, E. K., Akgun, O. E., Karadeniz, S., & Demirel, F. (2017). Scientific Research Methods. Pegem Citation Index. Buyukozturk, S. (2018). Manual of data analysis for social sciences. Pegem Citation Index.

Civelek, M. (2018). Essentials of Structural Equation Modeling. Lincoln: University of Nebraska Lincoln-Zea Books.

Demir, A. (2020). Learning Styles of People According to the Enneagram. International Journal of Social Sciences, 4, (19): 13-29.

Demir, A., Rakhmanov, O., Tastan, K., Dane, S., & Akturk, Z. (2020a). Development and validation of the Nile personality assessment tool based on Enneagram. Journal of Research in Medical and Dental Science, 8 (4), 24-32.

Demir, A., Rakhmanov, O., & Dane, S. (2020b). Validity and reliability of the Nile personality assessment tool based on Enneagram for English-speaking people. Journal of Research in Medical and Dental Science, 8 (4):118-123.

Erdogan S., Nahcivan, N, & Esin N. (2015) Research in Nursing (2nd edition). Nobel Bookstore, Ankara.

Erkus, A. (2019). Measurement and Scale Development in Psychology III, Pegem Academy, Ankara.

Galvez, C. N., Ramos-Vera, C. A., Barrientos, A. S., & Ogundokun, R. O. (2021). Adaptation of the Personality Type Inventory Based on Enneagram in Peruvian university students of health sciences. Journal of Research in Medical and Dental Science, 9(5), 10-15.

Guney, S. (2018). Behavioral Sciences. 11th edition, Ankara: Nobel Academic Publishing.

Heuertz, C. (2020). The Enneagram of Belonging: A Compassionate Journey of Self-Acceptance. Grand Rapids, MI: Zondervan.

Hook, J. N., Hall, T. W., Davis, D. E., Van Tongeren, D. R., & Conner, M. (2021). The Enneagram: A systematic review of the literature and directions for future research. Journal of Clinical Psychology, 77(4), 865-883. https://doi.org/10.1002/jclp.23097

Jung, C.G. (1971), Psychological Types: Collected Works, Volume 6, Routledge and Kegan Paul, London.

Kale, S. K., & Shrivastava, S. (2003). The Enneagram system for enhancing workplace spirituality. Journal of Management Development, 22 (4): 308-328. https://doi.org/10.1108/02621710310467596

Karagoz, Y. (2016). SPSS 23 and AMOS 23 Applied Statistical Analysis. Nobel Academic Publishing.

Kastelein, K. (2021). The Reliability and Validity of the Open Enneagram of Personality Scales. (Honors Program Projects). Illinois: Olivet Nazarene University.

Kirmayer, L.J., Narasiah, L., Munoz, M., Rashid, M., Ryder, A. G., Guzder, J., Ghayda, H., Rousseau, C., & Pottie, K. (2011). Common Mental Health Problems In Immigrants And Refugees: General Approach In Primary Care. Canadian Medical Association Journal, 183(12). https://doi.org/10.1503/cmaj.090292

Matise, M. (2019). The Enneagram: An Enhancement to Family Therapy. Contemporary Family Therapy. 41, 68-78. https://doi.org/10.1007/s10591-018-9471-0

Nizam, Ö. K., & Gul, S. S. (2019) A Critical Assessment on Addressing Syrians in Turkey as Sociological "Stranger". Suleyman Demirel University Faculty of Arts And Sciences Journal of Social Sciences, 2(47), 97-115.

https://dergipark.org.tr/tr/download/article-file/841915

Ozdamar, K. (2016). Scale and Test Development Structural Equation Modeling in Education, Health and Behavioral Sciences. April Bookstore.

Piskin, M. (2011). Factors Affecting Career Development Process. Yesilyaprak, E. (Ed.) Vocational Guidance and Career Counseling: From Theory to Practice. Ankara: Pegem Academy Publisher.

Polit, D. F., & Beck, C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 29(5), 489-497. https://doi.org/10.1002/nur.20147

Pontes, H. M., & Griffiths, M. D. (2015). Measuring DSM-5 Internet gaming disorder: Development and validation of a short psychometric scale. Computers in Human Behavior, 45, 137-143. https://doi.org/10.1016/j.chb.2014.12.006

Riso, D. R., & Hudson, R. (2018). Personality Analysis with Enneagram from Two of the World's Most Famous Experts. (Translation G. Aksoy). Istanbul: Boutique Publishing.

Sahin, I., Turan, H., & Yildiırim, A. (2020). Enneagram personality types and academic success (Kocaeli-Izmit District 7th and 8th Class Sample). Kocaeli Journal of Social Sciences Institutes, KOSBED, 39, 149-158. https://doi.org/10.35343/kosbed.526020

Sencan, H. (2005). Reliability and Validity in Social and Behavioral Measures (1st Edition).

Seckin Publishing Industry and Trade Inc., Ankara.

Simsek, O. F. (2007). Introduction to Structural Equation Modeling (Basic principles and LISREL applications). Ankara: Equinox.

Sirin, T. (2020). Enneagram personality inventory development and determination of turkey psychometric properties. Electronic Turkish Studies, 15(2). https://doi.org/0000-0002-4872-7386

Tastan, K. (2019). Development and validation of a personality type inventory based on Enneagram. Konuralp Medical Journal, 11(1), 112-118. https://doi.org/10.18521/ktd.462049

Tavsancil, E. (2005). Measuring Attitudes and Data Analysis with SPSS (2nd edition). Ankara: Nobel Publications.

Tieger, P. D., Barron, B., & Tieger, K. (2014). Do What You Are: Discover The Perfect Career For You Through The Secrets Of Personality Type. Hachette UK.

Weng, N., Li, K., Lan, H., Zhang, T., Zhang, X., Gui, Y., Xuefei, F., & Liu, Q. (2021). Evaluation of the reliability and validity of functional assessment of chronic illness therapy-spiritual well-being-expanded in elderly patients with chronic orthopaedic diseases. Psychogeriatrics, 21(1), 32-41. https://doi.org/10.1111/psyg.12624

Yaslioglu, M. M. (2017). Factor analysis and validity in social sciences: Application of exploratory and confirmatory factor analyses. Istanbul University Journal of The School of Business, 46, 74-85. https://dergipark.org.tr/tr/download/article-file/369427

Yildiz, S., Boz, I.T., & Yildirim, B.F. (2012). The relationship between personality type and positive social behavior: A research on Marmara University students. Journal of Economics and Administrative Sciences, 26(1), 215-233. https://dergipark.org.tr/tr/download/article-file/30257