The Adaptation of the CES-Depression Scale into Turkish through the use of Confirmatory Factor Analysis and Item Response Theory and the Examination of Psychometric Characteristics

Arkun Tatar[1], Gaye Saltukoglu[1]

ÖZET:
CES-Depresyon Ölçeği'nin doğrulayıcı faktör analizi ve madde cevap kuramı kullanımı ile Türkçe'ye uyarlanması ve psikometrik özelliklerinin incelenmesi

Amaç: Bu araştırmada CES-Depresyon Ölçeğinin Türkçe'ye uyarlanması ve psikometrik özelliklerinin incelenmesi amaçlanmıştır.
Yöntem: Ölçeğin uyarlanması için toplamda 1143 kişilik veri kullanılmıştır. Değişik yollarla toplanan veriler, farklı geçerlik ve güvenirlik çalışmaları için kullanılmıştır.
Bulgular: Türkçe'ye uyarlanan ölçeğin güvenirlik ve geçerlik bulguları literatürle oldukça benzerdir. Ölçeğin iç tutarlık katsayısı 0,75 ile 0,90 arasında, Guttman iki yarım test güvenirliği 0,89, iki hafta arayla yapılan test-tekrar test güvenirlik katsayısı 0,69 olarak bulunmuştur. Madde Cevap Kuramı ile incelenen ölçek maddelerinden 7. madde en düşük ayırt edicilik gösteren madde olmuştur. Bu maddeyi sırasıyla 11, 15 ve 13. maddeler izlemiştir. Ayırt ediciliği en yüksek maddeler ise sırasıyla 6, 3 ve 8. maddelerdir. Ölçeğin açıklayıcı ve doğrulayıcı faktör analizi sonuçları da diğer çalışmalara oldukça benzerdir. Ölçeğin dört faktörlü yapısının test edildiği Doğrulayıcı Faktör Analizi sonucunda iyi uyum katsayısı 0,84 olarak elde edilmiştir. Ölçeğin Beck Depresyon envanteri ile bağıntı düzeyi 0,77 bulunmuştur. Ölçeğin hasta ve hasta olmayan grupları birbirinden ayırt etmede %81,7 oranında etkili olduğu görülmüştür.
Sonuç: Türkçe'ye uyarlanan CES-Depresyon Ölçeği'nin tarama amaçlı grup uygulamalarında kullanılabileceği görülmüştür.

Anahtar sözcükler: Depresyon ölçeği, uyarlama, geçerlik, güvenirlik

ABSTRACT:
The adaptation of the CES-Depression scale into Turkish through the use of confirmatory factor analysis and item response theory and the examination of psychometric characteristics

Objective: The purpose of this study was to adapt the CES-Depression scale into Turkish and to investigate the psychometric characteristics of the scale.
Method: The data were collected from 1143 subjects for the validity and reliability analyses through different procedures.
Results: The reliability and validity results of the scale found in this study were quite similar to the results in the literature. The internal consistency coefficient was between 0.75 and 0.90. The Guttman split-half coefficient was 0.89. The test-retest reliability by two weeks was 0.69. The analysis of items by item response theory showed item 7 to be the least discriminant item, and consecutively item 11, 15, and 13 followed. The most discriminant items happened to be 6, 3, and 8. The factor loadings of the items of the scale, both by the explanatory and confirmatory factor analyses, led to similar results found in other studies. The four factor structure of the scale was tested and the GFI was 0.84 through confirmatory factor analysis. The correlation coefficient between the CES-D and Back Depression Inventory was 0.77. We observed that the scale discriminated between patient and non-patient groups effectively at 81.7%.
Conclusion: The Turkish version of the CES-D scale can be used for the screening purposes during the assessment process of groups.

Key words: Depression scale, adaptation, validity, reliability

[1,2]Assistant Professor, Haliç Üniversitesi Fen Edebiyat Fakültesi Psikoloji Bölümü öğretim üyesi, Istanbul-Turkey

Yazışma Adresi / Address reprint requests to:
Arkun Tatar, Haliç Üniversitesi Fen Edebiyat Fakültesi Psikoloji Bölümü Kaptanpaşa Mh. Darülaceze Cd. No:14 Okmeydanı, İstanbul-Turkey

Telefon / Phone: +90-212-220-9696

Elektronik posta adresi / E-mail address:
arkuntatar@yahoo.com

INTRODUCTION

CES- Depression Scale (The Center for Epidemiologic Studies Depression Scale, (1)), was developed by The American National Mental Health Institute (2), though it has not been advised as a tool for individual diagnosis (3), it is a short self report scale developed for screening purposes and widely used in scientific studies for the assessment of depressive symptoms in general population (4,5). The item pool of the CES-D Scale, which has been found sensitive in measuring depressive symptoms (6), was obtained from other valid depression scales at the time of the development process. The basic depressive symptoms were chosen from the clinical literature and from the factor analytical studies (1). The assessments regarding the last week were asked and the answers were scored on a four point likert scale; 0 (never-rarely) and 3 (mostly-usually). The scale is comprised of 20 items, of which four scored in reverse order (7). The scale has also many short versions (8-14).

The score range of the long form differs between 0 and 60, and high scores indicate depression (15).

Although, different cut-off scores are reported in different languages and different cultures (3), a score of 16 which is a cut-off point or above, is accepted to be a symptom of clinical depression according to the American norms (4,16).

The scale was frequently examined in different constructs; which were one dimension, three and four dimensions (17,18). It was tried to measure depressive symptoms, positive affect, somatic symptoms and difficulties in interpersonal relationships through the four dimensions construct (17,19,20). Although sub-dimensions were not scored separately, in many studies, where the four dimension constructs were preferred (1,2,17,19-22), it was reported that the explanatory factor analytical results supported the four factors construct (2,23). Besides there are many studies that support the above mentioned construct using confirmatory factor analysis (18,24, 25).

The scale was translated into many languages (3,26-32) and has been used frequently in different age groups (33-40), in different patient groups (41-44), and in different ethnic groups (45-49). In the reviews of the articles, where depression was investigated using the CES-D Scale, the sample size was found to be wider than 7300 in four articles (50). In another study, where twenty eight articles were reviewed, the sample size was found to be 22340 (22). In many studies the scale was used to compare people, whether they have depression diagnosis or not (5,51-54).

The scale has been shown to have good psychometric qualities (25,55), high reliability and validity (56-58), acceptable test-retest results (59), high correlation with clinical depression diagnosis, and sensitivity to life events (60,61). The internal consistency coefficients in different studies were reported to change between 0.72 and 0.92 (62-64).

Depression is a disease that weakens the person physically and mentally and reduces efficiency in the social environment. Outpatient diagnosis and treatment is possible but it is a widespread and common disease (65). Therefore, it becomes crucial to diagnose depression. It has been reported that there are many scales used in Turkish (66). The advantage of the CES-D Scale is that it is free to use for all investigators (i.e. no copyright payment is necessary) and for the same reason it's a widely used instrument in different cultures, facilitating comparative studies between different cultures. Because the scale was developed as a screening device, it differs from other depression scales in Turkish.

The studies in the literature related to the use of the scale on Turkish samples are those that were conducted in Holland and in these studies the Turkish immigrants and the immigrants from North Africa in Germany were compared (67-69). The scale were applied to elder Turkish immigrants in the presence of the translators (67) or by forms that were previously translated into Turkish by the translation agencies (68,69). All results demonstrated that the Turkish form of the CES-D Scale could be a preferable measuring instrument for the investigators in this field. Consequently this study aimed to obtain a reference for the investigators who would want to use the instrument as well as to adapt the scale into Turkish, and to analyze its psychometric properties comprehensively.

METHOD

Preliminary Study: Study of Translation and the Specialist View for the Language Equivalence Validity

In this study at first, the scale was translated into Turkish by the investigator. Afterwards two assistant professors from the department of translation and the department of American culture and literature were asked separately to control and correct the translation. The two different corrected forms were combined, and a second opinion was seeked. Upon the consensus of the given opinion the next step was taken.

The items in the two forms, in Turkish and in English, were given to four associate professors in psychology department all of whom speak both languages and who are qualified enough in the area to be able to assess the language equivalence issue. The experts were asked to assess the item equivalence by scoring the items in both forms of the scale. Kendall Goodness of Fit analysis were applied to the assessment scorings of the experts and it was determined that there were no opinion differences between the items of the scale written in different languages (Kendall W= 0.169; $\chi2(19) = 12.832$; p> 0.05).

Main Study

Participants

In the process of adaptation of the CES-D Scale different procedures were applied in different studies. After the preliminary study regarding the four different validity and reliability procedures and a data collection application, a total of five data collection studies were

Table I. Below under the heading of "Procedure" besides the procedures carried on, the other characteristics of the participants were explained; under the heading of "Findings" to which statistical procedures were applied; each case of data were indicated. Data of 720 people in the main application study, the combined total data, and all of the analyses showed very similar results. So this data were not assessed separately in the next section, but were included in the combined data.

Table 1: Demographic Characteristics of the all studies					
	Linguistic Equivalence for Adaptation Validity	Test Retest Reliability	Criterion Validity	Discriminative Validity	All Participants
	n= 60	n= 54	n= 233	n= 76	n= 1143 (423+720)
Characteristic	Student (Normals)	Student (Normals)	Adult (Normals)	Adult (Depressed)	Student+Adult (Normals+Depressed)
Age (years)					
Min-Max	18-31	18-26	20-64	15-64	17-85
Mean - St. dev.	22.22±2.00	21.89±1.60	36.02±11.11	35.16±11.88	33.80±12.71
Sex					
Women	44 (73.3%)	43 (79.6%)	118 (50.6%)	43 (56.6%)	724 (63.3%)
Men	16 (26.7%)	11 (20.4%)	115 (49.4%)	33 (43.4%)	419 (36.7%)
Education Level					
Elementary School	-	-	13 (5.6%)	10 (13.2%)	93 (8.1%)
Middle School	-	-	16 (6.9%)	9 (11.8%)	69 (6.0%)
High School	-	-	76 (32.6%)	31 (40.8%)	300 (26.2%)
University	60 (100%)	54(100.0%)	128 (54.9%)	26 (34.2%)	681 (59.6%)
Marital Status					
Single	58 (96.7%)	52 (96.3%)	95 (40.8%)	27 (35.5%)	605 (52.9%)
Married	2 (3.3%)	2 (3.7%)	128 (54.9%)	44 (57.9%)	476 (41.6%)
Widowed	-	-	5 (2.1%)	5 (6.6%)	35 (3.1%)
Divorced	-	-	5 (2.1%)	-	27 (2.4%)
Annual Income					
Low	5 (8.3%)	4 (7.4%)	18 (7.7%)	17 (22.4%)	103 (9.0%)
Middle	27 (45.0%)	23 (42.6%)	134 (57.5%)	39 (51.3%)	756 (66.1%)
High	27 (45.0%)	26 (48.1%)	81 (34.8%)	20 (26.3%)	272 (23.8%)
Unspecified	1 (1.7%)	1 (1.9%)	-	-	12 (1.0%)

performed. In these studies 60 students for language equivalence in adaptation validity study, 54 students for test-retest reliability study, 233 adults for criterion validity study, 76 adults diagnosed as depressive for discriminative validity study participated, and in addition to the data, 720 adults participated in the main study; which totals up to 1143 participants of the ages 17-85 (724 female, 419 male). The characteristics of the four participating groups separately and the combined characteristics of all the participants were summarized in

Data Collection Tools

In this study, which was performed in order to adapt the CES-D Scale into Turkish, apart from the Turkish form of the scale, a five-item questionnaire was administered of which the results were indicated in Table I. Other tools are the English form of the scale which were used for the language equivalence in the adaptation validity study, and the Beck Depression Inventory which were used to obtain external criteria material for the criterion validity study.

Beck Depression Inventory: In this study, for the criterion validity study along with the CES-D Scale the Beck Depression Scale (BDS) was been used. The Scale had been developed in 1961 by Beck et all. (65). It then was adapted into Turkish by two independent studies (66). In this study the version adapted by Hisli was used (70, 71).

Procedure

All of the participants accepted to participate in a scientific study and to answer the questionnaires voluntarily during all the studies. The procedures in different data collection studies were shown below.

60 undergraduate students who speak both of the languages from the department of Translation and the Department of American Culture and Literature served as subjects in the study to determine the adaptation validity for linguistic equivalence of the scale. The students who participated in the study answered the questions on Turkish and English forms of the CES-D Scale individually or in groups with one week interval. A one week interval (in the instructions of the scale, it has been stated that the replies must be given regarding the last week (see appendix)) was decided on since it is a long enough period that subjects would not remember their previous answers nor there would be significant changes in the characteristics to be measured since it is a short enough period as well. Half of the students who participated in the study were given the English form before the Turkish form and the other half of the participants were given the Turkish form before the English form in order to compensate for the effects of priority. Only the data from the Turkish form were included in all participants data set and the English data set were excluded.

Test-retest reliability study was conducted on 54 participants easily accessible from three different schools; Engineering, Literatur, and Letters and Sciences. The applications were carried out partly individually and partly in groups by two week intervals. Only the pre-application data were included in all the participants data set, post-application data were excluded.

Beck Inventory was used with the CES-D Scale to provide external criteria and to conduct criterion validity study. For this purpose, 233 adult participants were included in the study. The participants, who the researcher could reach randomly, completed both the scales and the questionnaire individually in an average of 10-20 minutes. The data of this study that were completed within a period of approximately six months and it was included as a whole in all participants data set.

76 adult patients diagnosed with depression by the psychiatrist in the psychiatry clinics of a total of twelve public and private hospitals in Bursa and Istanbul were included for discriminant validity study. Applications were conducted by a psychologist individually on the patients being diagnosed with depression and on those who came for check up. The data of this study were included as a whole in all participants data set.

Lastly, the study data that were explained above were not considered sufficient for the adaptation of the CES-D Scale into Turkish and would not suffice to examine the psychometric properties of the scale. Therefore 720 adult participants were included in the study to increase the number of participants and to diversify the attendees profile. Participants filled out the questionnaire used in the other studies with the Turkish form of the CES-D Scale. Applications were completed in a time interval of 5-10 minutes, due to the limited number of items. The study has been carried out approximately for ten months. The data of this study were not considered separately and, were combined with other study data (n=423) and it included 'all participants' data.

RESULTS

The results obtained from this study were not arranged in the order of the afore mentioned study but were classified according to the content of the study and are presented below. Similarly, first the reliability findings then the validity findings of the scale were placed.

Reliability Results

Internal Consistency

The total average scores, standard deviations and Cronbach Alfa internal consistency coefficients of the CES-D Scale in all the studies are shown in Table 2. The internal consistency of the scale was found between 0.75 and 0.90 in different parts of the study.

The item analysis of all the participants included in the study are shown in Table 3. Though the Cronbach Alfa internal consistency coefficient of 0.89 was observed

Table 2: Means, standard deviations and internal consistencies in different studies

	n	Min-Max	Mean	St. dev.	Cronbach's Alpha
Linguistic Equivalence for Adaptation Validity					
English Form	60	2-43	19.60	9.67	0.87
Turkish Form	60	0-45	19.27	10.80	0.87
Test Retest Reliability					
Test	54	5-42	19.56	9.00	0.84
Retest	54	0-45	19.42	11.08	0.88
Criterion Validity					
CES-D	233	0-60	15.53	12.27	0.90
Beck	233	0-63	8.88	9.09	0.91
Discriminative Validity					
(Depressed)	76	7-50	36.01	7.60	0.75
All Participants*	1143	0-60	18.10	11.63	0.89

*The data of the linguistic equivalence for adaptation validity study of the English form and the last data for the test-retest study are not included in the total group data consisted of all the participants of the CES-D Scale (See "Participants and Procedure").

Table 3: Summary of the Internal Reliability Analysis of the CES-Depression Scale

Items n=1143	Mean	Variance	r	Alpha if Item Deleted
madde01	17,25	120,95	0,58	0,88
madde02	17,47	124,64	0,46	0,88
madde03	17,47	120,04	0,66	0,88
madde04	17,00	123,18	0,43	0,89
madde05	17,25	121,77	0,55	0,88
madde06	17,28	117,81	0,71	0,88
madde07	15,96	133,83	0,02	0,90
madde08	17,20	122,82	0,49	0,88
madde09	17,52	122,07	0,60	0,88
madde10	17,45	122,57	0,55	0,88
madde11	16,92	128,15	0,22	0,89
madde12	17,11	119,97	0,64	0,88
madde13	17,20	125,23	0,41	0,89
madde14	17,23	119,06	0,65	0,88
madde15	16,97	124,91	0,38	0,89
madde16	17,16	122,17	0,53	0,88
madde17	17,57	123,58	0,54	0,88
madde18	17,20	118,88	0,69	0,88
madde19	17,53	125,14	0,44	0,89
madde20	17,12	120,79	0,59	0,88

in the data for the whole scale, the results of the item analysis show that the correlation between item 7 and the item-total scale score was low. The Cronbach Alfa internal consistency coefficients were calculated as 0.86 for the Depressive Symptoms subscale, 0.78 for the Positive Emotional State subscale, 0.62 for the Somatic Symptoms subscale and 0.36 for the Interpersonal Relations subscale.

Split Half Reliability

The split half reliabilities of the scale were 0.80 for

the first and the second half, while Guttman Split-half and Equal Length Spearman-Brown coefficients were obtained as 0.89, the correlation level between the two halves were 0.80.

Test Retest Reliability

Pearson Correlation Analysis was performed for the two applications with two week intervals on 54 people, both to the total scale scores and to the single items. The correlation level of the total scale scores between the two applications were determined as $r=0.69$; $p<0.001$. The analysis for the items between the two applications showed the lowest correlation ($r=0.09$) for item 11 "My sleep was restless" and the highest correlation ($r= 0.57$) for item 10 "I felt fearful".

Investigation of Items by Item Response Theory

The items of CES-D were been investigated by the classic test theory under the headings of "internal consistency and split half reliability". In the framework of this theory the state of the scale and its items were described above. Because the classic test theory is not sufficient to obtain information for the investigation of the scale and its items (72), here, Item Response Theory (IRT) was used. IRT is defined as "Modern Test Theory" which distinguishes it from the classic test theory (73). When it was decided to use the modern test theory for the investigation of the items of the CES-D Scale, the important problem of which program would be used and which IRT model would be chosen to estimate the item and person parameters rose.

Collins et al. (74) reported that the Parscale program is used for likert type personality scale items. Raju et al. (75) indicated that they used the Parscale computer program and the Samejima's Graded Response Model for the IRT calibration. Reise and Waller (76), in their study where they investigated whether The Logistic Model with Two Parameters (PLM) or The Logistic Model with Three Parameters (3PLM) were more appropriate, reported that there was not a very significant difference between the two models. A study in Turkish (77) used the Parscale computer program and 2PLM for the estimation of the item and person parameters.

In this study for the afore mentioned reasons 2PLM and

PARSCALE computer statistic program version 4.1 (78) were used for the estimation of the parameter estimates and fit statistics. The program analyzes dichotomous, multiple choice and likert type (polytomous) items by IRT.

Table 4 shows parameter estimates and fit statistics of the items according to the two parameters model estimated by the Parscale program. The results indicate that item 7, that was observed as having a weak correlation level between item-scale total score in the previous item analysis, has the lowest item discrimination level and is shown by the letter "a". Because CES-D Scale attempts to measure "depression level", the discrimination parameter indicates how well individuals with low and high depression levels are discriminated (73). Respectively, items 11, 15 and 13 are the other low discriminative items, and 6, 3 and 18 are the high discriminative items.

The other technical specification obtained by IRT is the item difficulty level shown by the letter "b". The difficulty or the power of the item defines the function of the item, and where it stands on the measured characteristic. An item defined as difficult, i.e. having high "b" value, functions among the test applied to people with high levels of depression, and an easy item (item with low b level) functions among people with low levels of depression. For this reason the difficulty of the item is a situation index upon the depression level (79). In other words, the difficulty parameter shows the location of the item on the depression level. But the value of item 7, $b= 0.000$, shows no location on the depression level. While items 4, 12, 18 function relatively at low levels on the measured characteristics, consecutively items 19, 2 and 17 are the first three items indicating a high level of function among the measured characteristics. Items 4, 7, 11, 13, 16 and 19 are the weakest items in their integration to the model.

Validity Findings

Linguistic Equivalence for Adaptation Validity

As a result of the analysis on the data of 60 students, the Pearson Correlation Coefficient calculated as the fitness measure between the English and the Turkish forms, was found to be $r=0.70$; $p<0.001$.

Table 4: IRT item parameter estimates, standard errors, and fit statistics for the CES-Depression Scale

Items n=1143	Slope (a)	S.E.	Location (b)	S.E.	χ^2	D.F.	p
Item No 1	1.488	0.066	1.044	0.055	14.26002	11	0.22
Item No 2	1.175	0.059	1.611	0.074	19.03986	12	0.09
Item No 3	2.479	0.119	1.052	0.045	8.95548	8	0.35
Item No 4	1.125	0.045	0.627	0.065	53.24862	11	0.00
Item No 5	1.510	0.064	0.981	0.053	23.59669	11	0.02
Item No 6	2.573	0.122	0.815	0.042	11.78377	9	0.23
Item No 7	0.116	0.005	0.000	0.520	390.10599	13	0.00
Item No 8	1.278	0.057	0.878	0.062	22.67220	11	0.02
Item No 9	1.937	0.095	1.272	0.052	3.30237	10	0.97
Item No 10	1.674	0.075	1.271	0.056	8.18566	10	0.61
Item No 11	0.540	0.022	0.936	0.110	45.39807	12	0.00
Item No 12	1.815	0.084	0.662	0.048	17.26570	10	0.07
Item No 13	0.986	0.038	1.304	0.072	34.82866	11	0.00
Item No 14	2.075	0.093	0.807	0.046	24.88217	9	0.01
Item No 15	0.883	0.035	0.942	0.074	34.17760	12	0.01
Item No 16	1.423	0.057	0.831	0.057	38.49033	11	0.00
Item No 17	1.735	0.087	1.406	0.058	5.86768	10	0.83
Item No 18	2.367	0.096	0.782	0.042	8.79257	9	0.46
Item No 19	1.261	0.058	1.617	0.074	36.15842	12	0.00
Item No 20	1.609	0.065	0.823	0.050	12.40692	11	0.33
Total					813.41876	213	0.00

Construct Validity

The construct validity of the scale was investigated by two different ways and was presented below under separate titles.

Explanatory Factor Analysis

The construct validity of the scale was first tested by the explanatory factor analysis (EFA). The Principal Components Analysis, Varimax Rotation were applied to the data of all the participants and the results are shown in Table 5. In this study the scale which was investigated by explanatory factor analysis as four dimensional, and as a result, the dimensions of "depressive symptoms", "positive affect" and "interpersonal relations" were found as expected. In the dimension of "Somatic Symptoms", the items 1, 5 and 20 showed higher factor loadings from the depressive symptoms dimension and item 7 had a higher factor loading from the interpersonal relations dimension. The four factor construct of the scale explained the 49.903% of the total variance.

The correlation coefficients of the items in the four dimensions of the scale were calculated. The highest correlation was between the "Depressive Symptoms" and the "Somatic Symptoms" dimensions as 0.61; the lowest correlation was found between the "Positive Affect" and the "Interpersonal Relations" dimensions as -0.29. At the same time, the correlation observed between the "Depressive Symptoms" and the "Positive Affect" was -0.53, the correlation between the "Depressive Symptoms" and the" Interpersonal Relations" was 0.45; the correlation between the "Positive Affect" and the " Somatic Symptoms" was -0.41 and the correlation between the "Somatic Symptoms" and the "Interpersonal Relations" was 0.42. The reason for the negative relations obtained was that the items of the positive affect dimension are the items that were scored in the negative direction.

Confirmatory Factor Analysis (CFA)

In the study performed with Amos 16.0, to reveal the construct validity of the scale, confirmatory factor analysis was applied to all participants using four factor hierarchical models. Factors, that are the sub-dimensions of the scale in accordance with the literature (25) were taken as unrelated with each other. The factor loadings for each item obtained by CFA and by the specified model are presented in Table 5 in order to make comparisons with EFA results. One of the widely used fitness statistics of the model developed for the CES-D Scale, i.e. the Goodness of Fit Index is determined as 0.84 and the

Table 5: Factor Loadings from confirmatory factor analysis and exploratory factor analysis of original four-factor model

	Negative Affect		Positive Affect		Somatic Symptoms		Interpersonal Problems	
	EFA	CFA	EFA	CFA	EFA	CFA	EFA	CFA
Item No 3	0.67	0.90						
Item No 6	0.64	1.00						
Item No 9	0.59	0.77						
Item No 10	0.65	0.79						
Item No 14	0.52	0.95						
Item No 17	0.72	0.73						
Item No 18	0.66	0.97						
Item No 4			0.64	0.92				
Item No 8			0.73	0.84				
Item No 12			0.73	0.89				
Item No 16			0.74	0.86				
Item No 1	0.58				0.21	0.96		
Item No 2					0.65	0.78		
Item No 5	0.51				0.06	0.86		
Item No 7					0.07	0.51	0.55	
Item No 11					0.67	0.71		
Item No 13					0.52	0.81		
Item No 20	0.51				-0.03	0.93		
Item No 15							0.71	0.97
Item No 19							0.36	0.74
Eigenvalue	6.373		1.544		1.078		0.986	
% of Variance	20.850		13.718		7.914		7.421	
Cumulative % of Variance	20.850		34.568		42.483		49.903	

Root Mean Square Error of Approximation (RMSEA), is determined as 0.10 (χ2 (276)=3597.377; p<0.001).

Criterion Validity

In order to determine the criterion validity of the CES-D Scale, the referred scale and the Beck Depression Inventory had been applied and the similarity level of the total scores of the two scales was calculated by Pearson Correlation Analysis as r= 0.77; p<0.001.

Discriminative Validity

Another procedure carried out in the study was the discriminative validity study done in order to determine how efficient the total score of the CES-D Scale was in the differentiation between the patient and normal groups. Both the analyses of the Receiver Operating Characteristics (ROC) and the Discriminate Analysis were used to discriminate the 76 patients diagnosed as depressive by the psychiatrist from the people described as normal.

The results of the ROC Analysis showed that the field value under the curve was 0.91, the sensitivity ratio value was 0.97, and the 1- specificity value was 0.43 for the cut-off point of 16 of the CES-D Scale. The discriminative analysis results showed that the 81.4% of the normal group and the 88.2% of the patient group diagnosed as depressive were identified by the total score of the CES-D Scale. The correct classification ratio for the whole group was 81.9%. The result of the t- test that compared the average total CES-D scores of the two groups revealed a statistically significant difference t (97.923) = 20.58; p<0.001.

The Comparison of the Total Scores of the CES-D Scale of the Participating Sub-groups

The total scores of the participants were examined in terms of gender, education, marital status, and income. No statistically significant differences were found between male and female participants' average total scale scores by the one-way analysis of variance, F(1, 1141)=1.59; p>0.05. There were statistically significant differences between the four different education groups (F(3, 1139) =26.45; p<0.001), among the four different marital status groups (F(3, 1139)=3.20; p<0.05) and among the three different income status groups (F(2, 1128)=34.30;

p<0.001). The Tukey-B multiple comparison tests conducted in order to examine the differences among the groups revealed that the CES-D Scale average total scores of people who have a university degree were significantly lower than the those of the of other education groups. While the average scores of widowed and divorced groups were higher than those of the groups of single and married, no differences were observed between the latter groups. The group that described their income as low had significantly higher average scores than those of the people in the other two income groups.

DISCUSSION

This study that was conducted in order to adapt the CES-D Scale into Turkish, had very similar results to the literature in regards to the reliability and validity of the scale. The internal consistency of the scale in this study took place between 0.75 and 0.90. For the whole group data as in Edman et al. (64) study, it was found to be 0.89. It was observed that in some studies the internal consistency of the scale was found lower than this one (0.72 (63)) and in the others higher than this one (0.92 (55)). For this reason, it may be concluded that in different application groups and cultures, the internal consistency of the scale actualizes around the values specified. While the internal consistency coefficients of the sub-dimensions of the scale were found between 0.44 and 0.87 in the study of Knight et al. (25), in this study it was between 0.36 and 0.86. The internal consistency coefficients of the sub-dimensions of the scale differ from each other depending on the number of the items. While the split-half reliability of the scale in Radloff's study (1) was found to be 0.76 and 0.77 in the normal group, 0.85 in the patient group and 0.83 in the Bush et al. (62) study, in this study both the Guttman split-half reliability coefficient and the Spearman-Brown coefficient were found to be 0.89.

Test-retest reliability of the scale by two weeks had been reported as 0.51 in the first publication of the scale (1). Hann et al. (59) reported the test-retest reliability of the scale as 0.51 and 0.57 in two different groups in their study, as for the study of Cook et al. (80), it was observed between 0.51 and 0.67. Yasuda et al. (81) reported a reliability coefficient of 0.84 in their test-retest application five days apart. In this study the test-retest

reliability coefficient in the test-retest interval of two weeks actualized as 0.69. When it is considered that the differences between the studies are not large enough, it can be said that the test-retest reliability of the scale in this study is comparable with other studies.

In this study, the results of the comparison of the individual test items in the test-retest application were found to be different than those specified in the study of Fountoulakis et al. (28). The researchers obtained the lowest correlation (0.45) for the item number 1, the highest correlation (0.95) for the item number 4 in test-retest interval of one or two days. In this study, the lowest correlation was found for item number 11 and the highest correlation was found for item number 10 in the test-retest interval of two weeks. Because there is no similar analysis, it has not been possible to make comparisons with other studies, but it seems that to deal with this issue in a separate study and to examine it in large groups may provide interesting information about the item reliabilities.

The study of Edelen and Reeve (15), where they investigated the items of the CES-D Scale by Item Response Theory (IRT), the items 4, 7, 8, 13, 15 and 19 were found as low discriminative items. The items 4, 10, 16, 17 and 20 were the weakest items in their fitness in the model. In this study some of these items attracted attention as having the lowest fitness in the model and the lowest discrimination. The items 4, 7, 13, 15, 16 and 19 being the weakest items of the scale exhibit similar findings as in the study of Edelen and Reeve (15). It is inevitable for any scale to have weak and strong items in terms of the measured characteristics however IRT in terms of determining the weak and strong items offers a more powerful method than does the classical theory. When many of the widely used tests and their items developed by the classic theory are examined from this perspective, the results will be seen as similar. For this reason, the situation observed in this study would not be accepted as a weakness.

The factor loadings of the items of the CES-D Scale in this study are quite similar to the results of other studies for both EFA and CFA (1,47,61,68,82). The study done by Croknett et al. (47) on 444 adolescences, item number 1 was loaded more from the depression sub-dimension rather than the factor; it belongs to the somatic symptoms as it was in this study. The items number 5 and 20 that

were loaded more from the depression sub-dimension in this study had higher loadings from the dimension they belong to, in the mentioned study but they also obtained high values in depression dimension. Whereas the items 11 and 13 that took place smoothly in their dimensions in this study have seemed as they were loaded more from the depressive symptoms dimension in the study of the mentioned researchers. These results indicate that there is a problem in the somatic symptoms dimension when the scale is treated as four dimensional. But on the other hand, it is acceptable because according to the results of EFA, some items may have lower loadings of their own dimensions, as EFA is an analysis that can give different results in different samples. Therefore it is not possible to say that the factor structure of the scale was corrupted. In fact, the factor loadings of the items in CFA analysis resemble the results of EFA. However, as a result of CFA analyses the factor loadings of the items seem to be clearer. As for the weak items, they reveal themselves more clearly in item analysis and in the both analyses CFA and EFA.

Item 7 seems not to work enough by its weak performance in this participating group. The content of the item (I felt that everything I did was an effort.) did not appear to be in the quality to differentiate between the depressives and the non depressives significantly. Though it was expected that the depressive people would answer as "more" days while non depressive people as "less" days, both groups answered in both directions. It was thought that answer alternatives were not responsible instead of "make an effort" or "felt the effort" expressions which did not stand out clearly as a symptom of depression. In other words, what has been stated in this item "the sense of effort about everything" may not be a good indicator of somatic symptoms of depression in the somatic symptoms sub-dimension of the scale, as it is observed that the item in question had not given different results in some other studies. For example, the study by Wong (61) done on a large sample showed that this item had a lower factor loading in comparison with the other items under the same sub-dimension. This situation constitutes the evidence that this item is the weakest item of the scale. The study of Wong also explains the situation of the item in this study at least partially. On the other hand, it is clear that the item needs further study.

As previously stated EFA are effected by different participating groups, i.e. by different applications. This situation leads to the deviations of the factor loadings of the items. However, seeing a core structure may be accepted as the capture of the basic structure. For this reason in many studies where the CES-D Scale is used, CFA was preferred instead of EFA (2,9,17,24,25,47,61,82-84). In all these studies, the four factor structure of the scale was investigated by CFA and by hierarchical model and GFI index was reported between 0.89 and 0.95. In this study GFI index was found slightly lower than the other studies, as 0.84. Whereas the GFI index, which had been reported between 0.56 and 0.88 in the other studies was found as 0.81 in this study. Root Mean Square Error of Approximation (RMSEA) was observed as 0.10, which was weak but consistent. These results show that the Turkish form of the CES-D Scale complies with the models presented in other studies though at low levels. Although, the somatic symptoms are a feature of depression, items in the somatic symptoms dimension have been distributed in part to the other dimensions, and this weakens the four factor structure of the scale partly. However, this does not create a problem in the use of the scale, because, as previously stated, the sub-dimensions had not been scored separately. Another point which was mentioned before was that the scale could be considered as single; three and four dimensional. In this study, the CFA results showed that the four factor structure of the scale was weak. In order not to expand the scope of the study, different structures were not tested, and the four dimensional structure was preferred, because it is the most commonly mentioned structure in the literature. However, the scale can also be considered as three and single dimensional.

The correlation levels of the sub-dimensions with each other of the examined four factor model are quite similar to other studies. The correlation levels among the dimensions were found as follows: between 0.45 and 0.96 in women, 0.37 and 0.88 in men in the study of Jennifer et al. (24); between 0.26 and 0.65 in the study of Pettit et al. (19); between 0.54 and 0.97 in the study of Wong (61); between 0.27 and 0.84 in the study of Knight et al. (25). As for the present study, the correlation between the dimensions was between 0.29 and 0.61. In all the studies mentioned, as well as in this study, the strongest correlation was between depressive symptoms and somatic symptoms, and the weakest correlation was

between positive affect and interpersonal relations.

The criterion validity coefficients of the CES-D Scale was reported as 0.73 with Zung Depression Scale (81), as 0.82 with Hamilton Depression Inventory (63), as 0.58 with MMPI-II (62) and as 0.77 with Depression Adjective Check List (85). Along with this, in many studies, (5, 56, 62, 86) the scale was used in conjunction with the Beck Depression Inventory, and the correlation level between the two scales was found to be between 0.67 and 0.81. In the present study, the correlation level between the two scales was 0.77.

The discriminate analysis of the total scores of the CES-D Scale classified 81.9% of the whole group, 88.2% of the patient group with a diagnosis of depression and 81.4% of the normal group correctly; the percentages seemed quite high. However, the percentage of correct classification, especially of the normal group, might be reduced due to the number of people (n=1143) in the group. Because no criterion was considered in the selection of this fairly large group, it is highly probable that there might be people diagnosed as depressive in the group. People with high scores whose possible diagnosis might had been depression, might be a factor that reduces the possibility of correct classification in the

normal group. On the other hand, the real discriminative power of the scale would be tested when it is used on people diagnosed with depression by a psychiatrist. In the present study, this percentage (88.2%) was higher than the percentage of correct classification of the normal group.

In this present study, area under curve was found to be 0.91 by ROC analysis. When this value approaches 1.00 the discriminative power of the test increases. Though different cut-off scores were not investigated in this study, a cut-off point of 16 accepted in American norms seemed to be very effective in the Turkish form. Although the purpose of this study was not to determine the cut-off point, it is an issue and should be examined for the Turkish form.

As a result, when we consider the points presented in this study, it is seen that the CES-D Scale is an effective instrument to measure depression. At the same time, when the reliability and validity studies are evaluated as a whole, it is seen that the adapted Turkish form shows the same effectiveness of the original form. Due to this, the adapted Turkish form of the scale can be used for screening purposes in the assessment of groups.

References:

1. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. Appl Psych Meas 1977; 1: 385-401.

2. Sheehan TJ, Fifield J, Reisine S, Tennen H. The measurement structure of the Center for Epidemiologic Studies Depression Scale. J Pers Assess 1995; 64(3): 507-521.

3. Gonçalves B, Fagulha T. The Portuguese Version of the Center for Epidemiologic Studies Depression Scale (CES-D). Eur J Psychol Assess 2004; 20: 339-348.

4. Orme JG, Reis J, Herz E. Factorial and discriminant validity of the Center for Epidemiological Studies Depression (CESD) Scale. J Clin Psychol 1986; 42: 28-33.

5. Santor DA, Zuroff DC, Ramsay JO, Cervantes P, Palacios J. Examining scale discriminability in the BDI and CES-D as a function of depressive severity. Psychol Assessment 1995; 7: 131-139.

6. Hubley AM, Hemingway D, Michalos AC. A comparison of older informal caregivers and non-caregivers living in non-metropolitan areas. Soc Indic Res 2003; 61: 241-258.

7. Carpenter JS, Andrykowski MA, Wilson J, Hall LA, Rayens MK, Sachs B, Cunningham LL. Psychometrics for two short forms of the Center for Epidemiologic Studies-Depression Scale. Issues Ment Health Nurs 1998; 19: 481-94.

8. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994; 10: 77-84.

9. Cole JC, Rabin AS, Smith TL, Kaufman AS. Development and Validation of a Rasch-Derived CES-D Short Form. Psychol Assessment 2004; 16: 360-372.

10. Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). Arch Intern Med 1999; 159: 1701-1704.

11. Kohout FJ, Berkman LF, Evans DA, Cornoni-Huntley J. Two shorter forms of the CES-D depression symptoms index. J Aging Health 1993; 5: 179-193.

12. Luyckx K, Soenens B, Goossens L, Beckx K, Wouters S. Identity Exploration And Commitment In Late Adolescence: Correlates Of Perfectionism And Mediating Mechanisms On The Pathway To Well–Being. J Soc Clin Psychol 2008; 27: 336-361.

13. Santor DA ve Coyne JC. Shortening the CES-D to improve its ability to detect cases of depression. Psychol Assessment 1997; 9: 233-243.

14. Shrout PE, Yager TJ. Reliability and validity of screening scales: Effect of reducing scale length. J Clin Epidemiol 1989; 42: 69-78.

Klinik Psikofarmakoloji Bülteni, Cilt: 20, Sayı: 3, 2010 / Bulletin of Clinical Psychopharmacology, Vol: 20, N.: 3, 2010 - www.psikofarmakoloji.org

223

15. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res 2007; 16: 5-18.

16. Leventhal AM, Chasson GS, Tapia E, Miller EK, Pettit JW. Measuring Hedonic Capacity in Depression:A Psychometric Analysis of Three Anhedonia Scales. J Clin Psychol 2006; 62: 1545–1558.

17. Perreira KM, Deeb-Sossa N, Haris KM, Bollen K. What Are We Measuring? An Evaluation of the CES-D Across Race/Ethnicity and Immigrant Generation. Soc Forces 2005; 83: 1567–1602.

18. Phillips GA, Shadish WR, Murray DM, Kubik M, Lytle LA, Birnbaum AS. The Center for Epidemiologic Studies Depression Scale with a young adolescent population: A confirmatory factor analysis and implementation of a general hierarchical model. Multivar Behav Res 2006; 41: 147-163.

19. Pettit JW, Lewinsohn PM, Seeley JR, Roberts RE, Hibbard JH, Hurtado AV. Association between the Center for Epidemiologie Studies Depression Scale (CES-D) and mortality in a community sample: An artifact of the somatic complaints factor? IJCHP 2008; 8: 383-397.

20. Santor DA. The Centre for Epidemiologic Studies Depression Scale. A. E. Kazdin (Chief Editor) American Psychology Association Encyclopedia of Psychology. American Psychology Association Press. 2000.

21. Schein RL, Koenig HG. The Center for Epidemiological Studies-Depression (CES-D) Scale: Assessment of depression in the medically ill elderly. Int J Geriatr Psych 1997; 12: 436-446.

22. Shafer AB. Meta-analysis of the Factor Structures of Four Depression Questionnaires: Beck, CES-D, Hamilton, and Zung. J Clin Psychol 2006; 62:123-146.

23. Dick RW, Beals J, Keane EM, Manson SM. Factorial structure of the CES-D among American Indian adolescents. J Adolescence 1994; 17: 73-79.

24. Jennifer AB, McCreary DR, Wright KD, Asmundson GJG. Factorial validity of the center for epidemiologic studies-depression (CES-D) scale in military peacekeepers. Depress Anxiety 2003; 17(1): 19-25.

25. Knight RG, Williams S, McGee R, Olaman S. Psychometric properties of the centre for epidemiologic studies depression scale (CES-D) in a sample of women in middle life. Behav Res Ther 1997; 35: 373-380.

26. Azocar F, Arean P, Miranda J, Munoz RF. Differential item functioning in a Spanish translation of the Beck Depression Inventory. J Clin Psychol 2001; 57: 355-365.

27. Dershem LD, Patsiorkovski VV, O'Brien DJ. The use of the CES-D for measuring symptoms of depression in three rural russian villages. Soc Indic Res 1996; 39: 89-108.

28. Fountoulakis K, Iacovides A, Kleanthous S, Samolis S, Kaprinis SG, Sitzoglou K, Kaprinis GS, Bech P. Reliability, validity, and psychometric properties of the Greek translation of the Center for Epidemiological Studies-Depression (CES-D) Scale. BMC Psychiatry 2001; 1: 1-20.

29. Iwata N, Roberts RE. Age differences among Japanese on the Center for Epidemiologic Studies Depression Scale: An ethnocultural perspective on somatization. Soc Sci Med 1996; 43: 967-974.

30. Lin L, Shi-Zhong W, Changmin Z, Qifu F, Keqiang L, Golinag S. Psychological long-term effects of sterilization on anxiety and depression. Contraception 1996; 54: 345-357.

31. Noh S, Avison WR, Kaspar V. Depressive symptoms among Korean immigrants: Assessment of a translation of the Center for Epidemiologic Studies-Depression Scale. Psychol Assessment 1992; 4: 84-91.

32. Schroevers MJ Ranchor AV, Sanderman R. The role of age at the onset of cancer in relation to survivors' long-term adjustment: A controlled comparison over an eight-year period. Psycho-Oncol 2004; 13(10): 740-752.

33. Arreola S, Neilands T, Pollack L, Paul J, Catania J. Childhood Sexual Experiences and Adult Health Sequelae Among Gay and Bisexual Men: Defining Childhood Sexual Abuse. J Sex Res 2008; 45(3): 246-252.

34. Boey KW. Cross-validation of a short form of the CES-D in Chinese elderly. Int J Geriatr Psych 1999; 14(8): 608-17.

35. Fendrich M, Weissman MM, Warner V. Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for Children. Am J Epidemiol 1990; 131(3): 538-551.

36. Fichman L, Koestner R, Zuroff DC. Depressive Styles in Adolescence: Assessment, Relation to Social Functioning, and Developmental Trends. J Youth Adolescence 1994; 23(3): 315-330.

37. Garrison CZ, Addy CL, Jackson KLAB, McKeown RE, Waller JLBS. The CES-D as a screen for depression and other psychiatric disorders in adolescents. J Am Acad Child Psy 1991; 30: 636-641.

38. Gatz M, Hurwicz ML. Are old people more depressed? Cross sectional data on Center for Epidemiological Studies Depression Scale factors. Psychol Aging 1990; 5: 284-290.

39. Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES D) as a screening instrument for depression among community residing older adults. Psychol Aging 1997; 12: 277-287.

40. Mathet F, Martin-Guehl C, Maurice-Tison S, Bouvard MP. Prevalence of depressive disorders in children and adolescents attending primary care. A survey with the Aquitaine Sentinelle Network. Encephale 2003; 29: 391-400.

41. Badger T, Segrin C, Meek P, Lopez AM, Bonham E, Sieger A. Telephone interpersonal counseling for women with breast cancer: Symptom management and quality of life. Oncol Nurs Forum 2005; 32: 273-279.

42. Katz MR, Kopek N, Waldron J, Devins GM, Tomlinson G. Screening for depression in head and neck cancer. Psycho-Oncol 2004; 13: 269–280.

43. Kurtz ME, Kurtz JC, Stommel M, Given CW, Given B. Predictors of depressive symptomatology of geriatric patients with lung cancer-A longitudinal analysis. Psycho-Oncol 2002; 11(1): 12-22.

44. Lawrence F, Glasgow RE, Mullan JT, Skaff MM, Polonsky WH. Development of a brief diabetes distress screening instrument. Annals of Family Medicine 2008; 6: 246-252.

45. Beals J, Manson SM, Keane EM, Dick RW. Factorial Structure of the Center for Epidemiologic Studies-Depression Scale Among American Indian College Students. Psychological Assessment: J Consult Clin Psych 1991; 3, 4, 623-627.

46. Black SA, Markides KS. Depressive symptoms and mortality in older Mexican Americans. Ann Epidemiol 1999, 9: 45-52.

47. Crockett LJ, Randall BA, Shen Y, Russell ST, Driscoll AK. Measurement equivalence of the Center for Epidemiological Studies Depression scale for Latino and Anglo adolescents: A national study. J Consult Clin Psych 2005; 73: 47-58.

48. Golding JM, Aneshensel CS. Factor Structure of the Center for Epidemiologic Studies Depression Scale Among Mexican Americans and Non-Hispanic Whites. Psychological Assessment: J Consult Clin Psych 1989; 1: 163-168.

49. Manson SM, Ackerson LM, Dick RW, Baron AE, Fleming CM. Depressive symptoms among American Indian adolescents: Psychometric characteristics of the Center for Epidemiologic Studies Depression Scales (CES-D). Psychol Assessment: J Consult Clin Psych 1990; 2: 231-237.

50. Cuijpers P, Smit F. Excess mortality in depression: A meta-analysis of community studies. J Affect Disorders 2002; 72: 227-236.

51. Baker FM, Parker DA, Wiley C, Velli SA, Johnson JT. Depressive symptoms in African American medical patients. Int J Geriatr Psych 1995; 10: 9-14.

52. Kawada T, Suzuki S, Kubota K, Ohnishi N, Satoh K. Content and Cross Validity of the Todai Health Index Depression Scale in Relation to the Center for Epidemiologic Studies Depression Scale and the Zung Self-rating Depression Scale. J Occup Health 1999; 41: 154-159.

53. Santor DA, Coyne JC. Evaluating the continuity of symptomatology between depressed and nondepressed individuals. J Abnorm Psychol 2001a; 110: 216-225.

54. Santor DA, Coyne JC. Examining symptom expression as a function of symptom severity: Item performance on the Hamilton Rating Scale for Depression. Psychol Assessment 2001b; 13:127-139.

55. Segal DL, Coolidge FL, Cahill BS, O'Riley AA. Psychometric Properties of the Beck Depression Inventory-II (BDI-II) Among Community-Dwelling Older Adults. Behav Modif 2008; 32: 3-20.

56. Bonilla J, Bernal G, Santos A, Santos D. A Revised Spanish Version of the Beck Depression Inventory: Psychometric Properties with a Puerto Rican Sample of College Students. J Clin Psychol 2004; 60: 119-30.

57. Roberts RE, Andrews JA, Lewinsohn PM, Hops H. Assessment of depression in adolescents using the Center for Epidemiologic Studies Depression Scale. Psychological Assessment: J Consult Clin Psych 1990; 2: 122-128.

58. Van't Veer-Tazelaar PJ, van Marwijk HWJ, Jansen APD, Rijmen F, Kostense PJ, van Oppen P, van Hout HPJ, Stalman WAB, Beekman ATF. Depression in old age (75+), the PIKO study. J Affect Disorders 2008; 106: 295-299.

59. Hann D, Winter K, Jacobsen P. Measurement of depressive symptoms in cancer patients: Evaluation of the Center for Epidemiological Studies Depression scale (CES-D). J Psychosom Res 1999; 46: 437-443.

60. Morrison R, O'Connor RC. Predicting Psychological Distress in College Students: The Role of Rumination and Stres. J Clin Psychol 2005; 61: 447-460.

61. Wong YI. Measurement Properties of the Center for Epidemiologic Studies-Depression Scale in a Homeless Population. Psychol Assessment 2000; 12: 69-76.

62. Bush BA, Novack TA, Schneider JJ, Madan A. Depression Following Traumatic Brain Injury: The Validity of the CES-D as a Brief Screening Device. J Clin Psychol Med S 2004; 11: 195-201.

63. Dozois DJA. The Psychometric Characteristics of the Hamilton Depression. J Pers Assess 2003; 80: 31-40.

64. Edman JL, Danko GP, Andrade N, McArdle JJ, Foster J, Glipa J. Factor structure of the CES-D (Center for Epidemiologic Studies Depression Scale) among Filipino-American adolescents. Soc Psych Psych Epid 1999; 34: 211-215.

65. Wehner JS, Stoner SC. Tools and Techniques for Evaluating. J Pharm Pract 2001; 14: 448-452.

66. Ceyhun B. Depresyonun Psikolojik Değerlendirmesi, Duygu Durum Dizisi 2001; 4: 192-197.

67. Smits CHM, de Vries WM, Beekman ATF. The CIDI as an instrument for diagnosing depression in older Turkish and Moroccan labour migrants: an exploratory study into equivalence. Int J Geriatr Psych 2005; 20: 436-445.

68. Spijker J, Van der Wurff FB, Poort EC, Smits CHM, Verhoeff AP, Beekman ATF. Depression in first generation labour migrants in Western Europe: the utility of the Center for Epidemiologic Studies Depression Scale (CES-D). Int J Geriatr Psych 2004; 19, 6, 538-544.

69. Van der Wurff, FB, Beekman ATF, Dijkshoorn H, Spijker JA, Smits CHM, Stek ML, Verhoef A. Prevalence and risk factors for depression in elderly Turkish and Moroccan migrants in the Netherlands. J Affect Disorders 2004; 83: 33-41.

70. Hisli N. Beck Depresyon Envanteri'nin geçerliği üzerine bir çalışma. Türk Psikoloji Dergisi 1988; 6: 118-126.

71. Hisli N. Beck Depresyon Envanteri'nin üniversite öğrencileri için geçerliği, güvenirliği. Türk Psikoloji Dergisi 1989; 7: 3-13 (Turkish).

72. MacDonald P, Paunonen SV. A Monte Carlo comparison of item and person statistics based on item response theory versus classical test theory. Educ Psychol Meas 2002; 62: 921-943.

73. Hambleton RK, Robin F, Xing D. Item response models for the analysis of educational and psychological test data. In. HEA Tinsley ve SD Brown (Ed.), Handbook of Applied Multivariate Statistics and Mathematical Modeling. San Diego: Academic Press 2000; 553-581.

74. Collins WC, Raju NS, Edwards JE. Assessing differential functioning in a satisfaction scale. J Appl Psychol 2000; 85: 451-462.

75. Raju NS, Laffitte LJ, Byrne BM. Measurement equivalance: a comparison of methods based on confirmatory factor analysis and item response theory. J Appl Psychol 2002; 87: 517-529.

76. Reise SP, Waller NG. How many IRT parameters does it take to model psychopathology items? Psychol Methods 2003; 8(2): 164-184.

77. Somer O. Gruplararası Karşılaştırmalarda Ölçek Eşdeğerliğinin İncelenmesi: Madde ve Test Fonksiyonlarının Farklılaşması. Türk Psikoloji Dergisi 2004; 19: 69-82 (Turkish).

Klinik Psikofarmakoloji Bülteni, Cilt: 20, Sayı: 3, 2010 / Bulletin of Clinical Psychopharmacology, Vol: 20, N.: 3, 2010 - www.psikofarmakoloji.org

225

78. Du Toit M (Ed.). IRT from SSI: BILOG-MG, MULTILOG, PARSCALE, TESTFACT. Lincolnwood (IL): Scientific Software International 2003.

79. Rouse SV, Finger MS, Butcher JN. Advances in clinical personality measurement: An item response theory analysis of the MMPI-2 PSY-5 Scales. J Pers Assess 1999; 72: 282-307.

80. Cook JM, Elhai JD, Areán PA. Psychometric properties of the PTSD Checklist with older primary care patients. J Trauma Stress 2005; 18: 371-376.

81. Yasuda T, Lubin B, Kim J, Whitlock RV. The Japanese version of the Multiple Affect Adjective Check List-Revised: Development and Validation. J Clin Psychol 2003; 59: 93-109.

82. Thomas JL, Brantley PJ. Factor Structure of the Center for Epidemiologic Studies Depression Scale in Low-Income Women Attending Primary Care Clinics. Eur J Psychol Assess 2004; 20, 2, 106–115.

83. Helmes E, Nielson WR. An examination of the internal structure of the Center for Epidemiological Studies Depression scale in two medical samples. Pers Indiv Differ 1998; 25: 735-743.

84. Hertzog C, Alstine JV, Usala PD, Hultsch DF, Dixon R. Measurement Properties of the Center for Epidemiological Studies Depression Scale (CES-D) in Older Populations. Psychological Assessment: J Consult Clin Psych 1990; 2: 64-72.

85. Lubin B, Roth AV, Dean LM, Hornstra RK. Correlates of depressive mood among normals. J Clin Psychol 1978; 34: 650-653.

86. Spielberger CD, Ritterband LM, Reheiser EC, Brunner TM. The nature and measurement of depression. IJCHP 2003; 3, 2 209-234.

226

Klinik Psikofarmakoloji Bülteni, Cilt: 20, Sayı: 3, 2010 / Bulletin of Clinical Psychopharmacology, Vol: 20, N.: 3, 2010 - www.psikofarmakoloji.org

Appendix.

Center for Epidemiological Studies Depression Scale (Turkish Form)

Aşağıda duygu ve davranışlarınızla ilgili ifadeler yer almaktadır. Lütfen geçen hafta boyunca aşağıdakileri ne sıklıkla hissettiğinizi veya yaşadığınızı belirtin.

	Hiçbir Zaman - Nadiren (1 günden daha az)	Birazcık - Birkaç Kez (1-2 gün)	Arada Sırada - Bazen (3-4 gün)	Çokça - Çoğu Zaman (5-7 gün)
1. Genellikle canımı sıkmayan şeyler canımı sıktı.	☐	☐	☐	☐
2. Açlık hissetmedim, iştahım yerinde değildi.	☐	☐	☐	☐
3. Arkadaşlarım veya ailemin yardımına rağmen kötü ruh halinden kurtulamadım.	☐	☐	☐	☐
4. Ruh halimin diğer insanlar kadar iyi olduğunu hissettim.	☐	☐	☐	☐
5. Yaptığım işe odaklanmakta zorlandım.	☐	☐	☐	☐
6. Kendimi depresyonda hissettim.	☐	☐	☐	☐
7. Her şeye çaba harcamam gerektiğini hissettim.	☐	☐	☐	☐
8. Gelecek için umutlu hissettim.	☐	☐	☐	☐
9. Hayatımın bir başarısızlık olduğunu düşündüm.	☐	☐	☐	☐
10. Korktuğumu hissettim.	☐	☐	☐	☐
11. Huzursuz uyudum.	☐	☐	☐	☐
12. Mutluydum.	☐	☐	☐	☐
13. Her zamankinden az konuştum.	☐	☐	☐	☐
14. Kendimi yalnız hissettim.	☐	☐	☐	☐
15. İnsanlar arkadaş canlısı değildi.	☐	☐	☐	☐
16. Yaşamdan zevk aldım.	☐	☐	☐	☐
17. Ağlama nöbetleri geçirdim.	☐	☐	☐	☐
18. Kendimi üzgün hissettim.	☐	☐	☐	☐
19. İnsanların benden hoşlanmadığını hissettim.	☐	☐	☐	☐
20. İşler yolunda gitmedi.	☐	☐	☐	☐

PUANLAMA: "**Hiçbir Zaman-Nadiren (1 günden daha az)**" cevabı 0, "**Birazcık-Birkaç Kez (1-2 gün)**" cevabı 1, "**Arada Sırada-Bazen (3-4 gün)**" cevabı 2 ve "**Çokça-Çoğu Zaman (5-7 gün)**" cevabı 3 olarak puanlanmaktadır. 4, 8, 12 ve 16 numaralı maddeler ters yönlü olarak puanlanmaktadır. Toplam puan aralığı 0 ile 60 arasındadır.