

Comparison of Scores Obtained from Different Depression Scales: A Test Equating Study on University Students

Akif AVCU, Marmara University, ORCID ID: 0000-0003-1977-7592

Şaziye Senem BAŞGÜL, Hasan Kalyoncu University, ORCID ID: 0000-0002-9843-8048

Abstract

This study aims to equate the scores of three commonly used depression scales used with university students and to create a cross-walk table. The data were collected from 1038 undergraduate students. Beck Depression Inventory, Epidemiological Research Center-Depression Scale and Depression-Anxiety-Stress Scale were used to measure the depression level of students. The prerequisites of the equating studies, were examined. Then, unidimensionality and local independence assumptions were examined for applicability of equating analysis based on Item Response Theory Models. Finally, equipercentile equating with single group design was conducted to align the scores to a common metric. Throughout the analysis process. R statistical program was used. Findings show that assumptions of equating are met. On the other hand, the findings suggest that the data are not suitable for equating using Item Response Theory models. Accordingly, a score conversion table was created based on single group equating design and with equipercentile equating method. The cutoff score of mild depression Beck Depression Inventory was found to be comparable with the cutoff score of Depression Anxiety Stress Scale while higher than the that of Center for Epidemiological Research Center Depression Scale. It implies that the cutoff scores determined for depression measurement tools do not perfectly overlap. In addition, the depression severity spectrum measured by the depression scales does not overlap. These findings suggest the necessity of the use of cross-walk tables which enable researchers and professionals to compare their scores with the scores obtained from different depression scales.

Keywords: test equating, depression, students



Inonu University
Journal of the Faculty of
Education
Vol 24, No 1, 2023
pp. 208-224
[DOI](#)
10.17679/inuefd.915613

Article Type
Research Article

Received
14.04.2021

Accepted
30.01.2023

Suggested Citation

Avcu, A. & Başgöl, Ş.S. (2023). Comparison of Scores Obtained from Different Depression Scales: A Test Equating Study on University Students, *Inonu University Journal of the Faculty of Education*, 24(1), 208-224. DOI: 10.17679/inuefd.915613

EXTENDED ABSTRACT

Introduction

Depression is defined as an emotional state manifested in the form of loss of pleasure in daily activities, sadness, guilt and worthlessness, anorexia, insomnia, and a decrease in sexual desire. A recent study (WHO, 2017) showed that 4.4% of the global population have depression. Many problems such as high risk of obesity and overeating (Stunkard, Faith, & Allison, 2003) and heart attack (Penninx et al., 2015) accompany people with depression. Therefore, various depression scales were developed and used by researchers in different fields to assess the depression levels of individuals belonging to the normal population who do not need to be diagnosed as depressive disorders. However, an unwanted result of the development of many different tools is that scores obtained from them cannot be directly compared (Fischer et al., 2011). One commonly used technique to obtain comparable scores from different scales is test equating. In this technique, the scores obtained by individuals are equated with a mathematical function.

Purpose

The purpose of this study is to obtain a cross-walk table that will enable the comparison of scores by equating three different depression scales widely used. The main reason for choosing university students as study group is that there are many studies conducted with university students to examine the level of depression (e.g. Tezel et al., 2009; Bozkurt, 2004; Günay et al., 2018; İskender et al., 2018) and depression is highly common among them (Ulas et al., 2016).

Method

The sample group consists of 1038 Turkish university students. During the data collection process, the authors reached out to university students studying at different departments of five universities through their faculty members and collected data using online data collection tools. The convenient sampling technique was used in the data collection process. The sample consists of 210 males (20.2%) and 828 (79.8) female students. Beck Depression Inventory, which was first developed by Beck et al. (1961) and adapted to Turkish by Hisli (1989); Center of Epidemiological Research Depression Scale which was developed by Radloff (1977) and adapted to Turkish by Tatar and Saltukoğlu (2010) and, Depression Anxiety Stress Scales which was developed by Lovibond and Lovibond (1996) and adapted into Turkish by Akin and Çetin (2007) were used. A pooled item pool was created by combining items from different depression scales, and also, the measurement tools were examined separately using Explanatory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) to obtain evidence of unidimensionality. While performing EFA, the ratio of the eigenvalue of the first dimension to the second dimension was used in deciding on unidimensionality. It is stated that this ratio of 4 can be used as proof of unidimensionality (Slocum-Gori & Zumbo, 2011). After finishing the preliminary analysis, equipercentile equating was performed with log-linear smoothing and the cross-walk table was obtained. For the equating of the scale metric of Beck Depression Inventory was used as reference and other were transformed into that metric. During the analysis process, the R program (2020) was used extensively.

Findings

The analysis started with the examination of test equating assumptions. The first assumption examined is that the scales must be equated need to be strongly correlated. Correlation coefficients between the scales were calculated accordingly. According to the findings, it was determined that they are highly correlated. Another assumption is that the scales to be equated needs to measure the same structure. Although the results obtained with CFA do not support that the data is unidimensional, EFA results show that the responses can be explained by an underlying dominant factor and can be regarded as a sign of unidimensionality. These findings are evidence that the second assumption is partially met. In the next step, the equating process was carried out using log-linear smoothing. While performing this analysis, the metric of Beck Depression Inventory was taken as reference and the estimated scores of the other two scales were transferred onto that metric. In this way, a cross-walk table for score comparison was obtained. When the table was examined, it was seen that equating relationships between the scores of depression scales were not linear. For example, when the relationship between Beck Depression Inventory and Depression Anxiety Stress Scale scores were examined it was seen that after the 52nd point observed for Beck Depression Inventory, Depression Anxiety Stress Scale scores reach the score of 42, which is the highest value that can be observed, and does not increase afterwards.

Discussion & Conclusion

In this study, a cross-walk table was created using the equipercetile equating technique. In this way, researchers were able to compare the scores obtained from different depression scales. Also, it is advantageous to use a single group equating design because it is more robust to sample-related variances. In addition, the application of all measurement tools to all participants is a factor that increases the precision of the equating functions (Dorans et al., 2010). When the score comparisons of the depression measurement tools are examined, it is seen that the cut-off point of "mild" level of depression determined for the Center for Epidemiological Research Depression Scale was found to correspond to a "severe" level of violence in the Beck Depression Inventory. That is, the cut-off scores determined for depression measurement tools do not overlap. Based on these findings, it can be said that the depression severity spectrum measured by depression scales does not overlap. Especially, Center for Epidemiological Research Depression Scale can be said to be more suitable for high depression severity. With the use of conversion tables, researchers can find the opportunity to compare their scores with the scores obtained in studies using different scales. In particular, the inclusion of studies using different scales in meta-analysis studies will increase the flexibility of researchers in the study selection process and it will be possible to obtain more accurate results. It should be kept in mind that it is not possible to generalize the obtained equalization functions to different populations.

Farklı Depresyon Ölçeklerinden Alınan Puanların Karşılaştırılması: Üniversite Öğrencileri Üzerine Bir Test Eşitleme Çalışması

Akif AVCU, Marmara Üniversitesi, ORCID ID: 0000-0003-1977-7592

Şaziye Senem BAŞGÜL, Hasan Kalyoncu University, ORCID ID: 0000-0002-9843-8048

Öz

Bu çalışma üniversite öğrencileri ile gerçekleştirilen çalışmalarda yaygın olarak kullanılan üç farklı depresyon ölçeğini eşitleyerek bir puan dönüşüm tablosu oluşturmayı amaçlamaktadır. Veriler üniversitede öğrenim gören 1038 lisans öğrencisinden toplanmıştır. Öğrencilerin depresyon düzeyini ölçme için Beck Depresyon Envanteri, Epidemiyolojik Araştırmalar Merkezi-Depresyon Ölçeği ve Depresyon-Kaygı-Stres Ölçeği kullanılmıştır. Eşitleme işlemi gerçekleştirilmeden önce bu işlem için önerilen ön koşullar incelenmiştir. Sonrasında, eşitlemenin madde tepki kuramı modelleri ile analiz uygulanabilirliğini değerlendirmek için tek boyutluluk ve yerel bağımsızlık varsayımları incelenmiştir. Son olarak tek grup deseni ile toplanan veriler eşit yüzdelikli eşitleme yöntemi ile ortak bir metriğe aktarılmıştır. Gerçekleştirilen bu analizler için R istatistik programı kullanılmıştır. Elde edilen bulgular eşitleme varsayımlarının karşılandığını göstermiştir. Diğer taraftan, verilerin Madde Tepki Kuramı modelleri kullanılarak eşitlenmesi için uygun olmadığı belirlenmiştir. Bu doğrultuda eşit yüzdelikli eşitleme yöntemi için gerçekleştirilen analizlere dayalı olarak puan dönüşüm tablosu oluşturulmuştur. Sonuç olarak, Beck Depresyon Envanteri için hafif düzey depresyon belirti düzeyi için belirlenen kesme puanının ile Depresyon Kaygı ölçeğinin depresyon alt boyutu için belirlenen kesme puanının örtüştüğü; Epidemiyolojik Araştırmalar Merkezi Depresyon Ölçeği için belirlenen kesme puanının ise Beck Depresyon Envanterinde şiddet olarak "hafif" düzeye karşılık geldiği görülmüştür. Yani, depresyon ölçüm araçları için belirlenen kesme puanları örtüşmemektedir. Ayrıca, depresyon ölçeklerinin ölçtükleri depresyon şiddet yelpazesinin örtüşmediği söylenebilir. Nihai olarak, oluşturulan dönüşüm tablolarının kullanılması ile birlikte, araştırmacı ve profesyonellerin elde ettikleri puanları farklı depresyon ölçeklerin kullanıldığı araştırmalarda elde edilen puanlar ile de karşılaştırmalarının gerekliliği görülmüştür.

Anahtar Kelimeler: test eşitleme, depresyon, öğrenciler



Inönü Üniversitesi
Eğitim Fakültesi Dergisi
Cilt 24, Sayı 1, 2023
ss. 208-224
[DOI](https://doi.org/10.17679/inuefd.915613)
10.17679/inuefd.915613

Makale Türü
Araştırma Makalesi

Gönderim Tarihi
14.04.2021

Kabul Tarihi
30.01.2023

Önerilen Atıf

Avcu, A. (2023). Farklı Depresyon Ölçeklerinden Alınan Puanların Karşılaştırılması: Üniversite Öğrencileri Üzerine Bir Test Eşitleme Çalışması. *Inönü Üniversitesi Eğitim Fakültesi Dergisi*, 24(1), 208-224. DOI: 10.17679/inuefd.915613

Farklı Depresyon Ölçeklerinden Alınan Puanların Karşılaştırılması: Üniversite Öğrencileri Üzerine Bir Test Eşitleme Çalışması

Depresyon, günlük aktivitelerde haz kaybı, üzüntü, suçluluk ve değersizlik, iştahsızlık, uykusuzluk ve cinsel istekteki azalma şeklinde kendini gösteren duygusal bir durum olarak tanımlanmaktadır (Davidson ve Neale, 1994). Blackburn'e (2005) göre depresyon, birçok duygusal durumun (üzüntü, karamsarlık, hayal kırıklığı vb.) bir arada var olduğu bir durumdur. Albrecht ve Herrick (2006), depresyonun yaşam boyu yaygınlık oranını % 15 olarak ortaya koymuştur. Ayrıca, yakın zamanda yürütülen bir araştırma (WHO, 2017) küresel nüfusun % 4,4'ünün depresyona sahip olduğunu göstermiştir. Daha da ötesinde, majör depresyon yaygınlığının yıllar içinde belirgin bir şekilde arttığını ortaya koymuştur (Compton ve diğ., 2006).

Depresyon, bireyleri farklı birçok şekilde etkileyebilmektedir. Örneğin, yakın zamanda yapılan araştırmalar depresyonlu bireylerin yürütsel işlevlerinde bozulma ile karşı karşıya olduğunu (Otte ve diğ., 2015), yüksek obezite ve aşırı yeme riski taşıdıklarını (Stunkard ve diğ., 2003), kalp krizi (Penninx ve diğ., 2015) riskine sahip olduklarını göstermiştir. Bu ikincil etkilerin bir sonucu olarak depresif bireylerin günlük yaşamları da olumsuz etkilenmektedir. Örneğin, depresif öğrencilerin not ortalamaları daha düşüktür ve ders çalışmak için daha az zaman harcarlar (Field ve diğ., 2001). Ayrıca depresyonlu bireylerde yaşam kalitesinin (Gao ve diğ., 2019) ve yaşam doyumunun (Koivumaa-Honkanen ve diğ., 2004) azaldığı bilinmektedir.

Depresyonun yaygınlığı ve insanlar üzerindeki olumsuz etkileri sebebiyle, farklı birçok evrende insanların depresyon düzeyini belirlemeye yönelik çalışmalar yürütülmüştür (örn. Bahar ve diğ., 2005; Ören ve Gençdoğan, 2007; Tezel ve diğ., 2009; Softa Kaçan, 2013). Günümüzde; Beck Depresyon Envanteri, Zung Depresyon Ölçeği, Hamilton Depresyon Derecelendirme Ölçeği, Montgomery-Asberg Depresyon Derecelendirme Ölçeği, Epidemiyolojik Araştırmalar, Merkezi-Depresyon Ölçeği ve Depresyon-Kaygı-Stres Ölçeği profesyoneller ve araştırmacılar tarafından yaygın olarak kullanılan ölçüm araçları arasındadır. Bu ölçüm araçlarından bazıları katılımcılar tarafından yanıtlanmaktadır. Bu nedenle, depresif bozukluk olarak teşhis edilmesi gerekmeyen normal evrene ait bireylerin depresyon düzeylerini değerlendirmek çeşitli alanlardaki araştırmacılar tarafından ilgi görmektedir. Bununla birlikte, çalışmalarda çok fazla depresyon ölçeği kullanılmasının istenmeyen bir sonucu ise farklı uygulamalardan elde edilen puanların doğrudan karşılaştırılamamasıdır (Fischer ve diğ., 2011).

Çoğu depresyon ölçüm aracının geliştirilmesi ve değerlendirilmesi, gerçek puanlar, gözlenen puanlar ve ölçüm hatası kavramlarına odaklanan Klasik Test Kuramına (KTK) dayanmaktadır. Bu ölçüm araçlarının tamamı aynı veya benzer bir yapıyı ölçse ve hatta maddelerinin oluşturulması aynı belirtiler kümesine dayalı olsa da, KTK yaklaşımı ile elde edilen toplam puanların karşılaştırılması mümkün değildir. Çünkü toplam puanlar maddelerin psikometrik özelliklerine bağlıdır (Reise ve Waller, 2009). Bu nedenle bireylerin elde ettikleri puanlar testin güçlük (ya da maddelerin ölçtüğü belirtilerin şiddet düzeyi) düzeyine göre değişir (Lord ve Novick, 2008). Ayrıca, farklı depresyon ölçüm araçlarındaki puan farklılıklarının diğer sebebi madde sayılarının ve yanıt seçeneklerinin aynı olmamasıdır. Yani, depresyon ölçme araçları içerik ve istatistiksel özellikler açısından birbirine benzemesine rağmen, bunların birbirleriyle tam olarak eşdeğer olması beklenemez ve formlar en azından güçlük ve madde sayısı bakımından farklılık gösterebilir.

Yanıt verenlerin gözlenen puanlarının yüzdelerle dilimlere veya standart z puanlarına dönüştürülmesi puanları karşılaştırılabilir hale getirebilir. Bununla birlikte, bu ortak metrik, örneklem grubunun özelliklerine ve puan dağılımlarına karşı duyarlı olacaktır (Baguley, 2009). Yani bu şekilde gerçekleştirilen bir puan karşılaştırması yanıltıcı olabilir. Farklı ölçeklerden karşılaştırılabilir puanlar elde etmek için yaygın olarak kullanılan fakat daha karmaşık olan yöntem ise test eşitlemedir. Bu yaklaşımda, bireylerin elde ettikleri puanlar matematiksel bir fonksiyonla eşitlenir.

Kolen ve Brennan (2014), test eşitlemeyi, benzer içerik ve düzeylerde geliştirilen farklı test formlarından elde edilen puanların, puanlar arasındaki farkları ayarlayarak, birbirinin yerine kullanılabilmesine izin veren bir dizi istatistiksel işlemler dizisi olarak tanımlamışlardır. Test formlarının denkleştirilmesi, bir formun metrik sisteminin diğer formun metrik sistemine dönüştürülmesiyle yapılır (Angoff, 1996). Bu şekilde gerçekleştirilen dönüşümden sonra farklı formlar eşdeğer olacaktır ve bir ölçüm aracından elde edilen puanın aynı özelliği ölçen başka bir ölçüm aracındaki puan karşılığı bilinecektir.

Bu bilgiler ışığında, gerçekleştirilecek olan bu çalışmanın amacı araştırmalarda ve profesyoneller tarafından yaygın olarak kullanılan üç farklı depresyon ölçeğini eşitleyerek puan karşılaştırmasını mümkün kılacak bir dönüşüm tablosu elde etmektir. Çalışma grubunun üniversite öğrencilerinden seçilmesinin temel sebebi, üniversite öğrencileriyle yürütülmüş ve depresyon düzeyinin incelenmesine yönelik çok sayıda çalışma olması (örn. Tezel ve diğ., 2009; Bozkurt, 2004; Günay Öztürk ve Arslantaş, 2018; İskender ve diğ., 2018) ve üniversite öğrencilerinde depresyonun yaygın olarak gözlenmesidir (Ulas ve diğ., 2016).

Yöntem

Araştırma Deseni

Gerçekleştirilen bu çalışmada depresyonu ölçen farklı ölçüm araçlarının üniversite öğrencilerine yönelik eşitlenerek puan dönüşüm tablosu elde edilmesi amaçlandığından dolayı betimsel araştırma modeli kullanılmıştır. Betimsel araştırmalar var olan durumu olduğu gibi ortaya çıkarılmasının amaçladığı amaçlar için uygun bir araştırma desendir (Karasar, 2005). Ayrıca, gerçekleştirilen bu çalışmada veriler mevcut eşitleme desenleri içerisinde yer alan tek grup eşitleme deseni kullanılarak elde edilmiştir.

Çalışma Grubu

Çalışma grubu Türk üniversite öğrencilerinden oluşmaktadır. Veri toplama sürecinde yazarlar tarafından beş üniversitenin farklı bölümlerinde öğrenim gören üniversite öğrencilerine ders aldıkları öğretim üyeleri aracılığıyla ulaşılarak çevrimiçi veri toplama araçları ile veriler toplanmıştır. Pandemi salgını sebebiyle veri toplama sürecinde yüz yüze görüşmeler gerçekleştirilmemiştir. Çalışma grubunda yer alan öğrenciler sekiz farklı fakültede öğrenim görmektedir. Araştırmacı, tanıdığı öğretim üyelerinin ders verdiği sınıflarda öğrenim gören öğrencilere ulaştığından, veri toplama sürecinde elverişli örnekleme tekniğinin kullanıldığı söylenebilir. Çalışmaya 1038 öğrenci dâhil edilmiştir. Katılımcıların demografik özellikleri Tablo 1'de yer almaktadır. Çalışma grubu 210 erkek (% 20.2) ve 828 (79.8) kadın öğrenciden oluşmaktadır.

Tablo 1*Çalışma Grubuna Ait Demografik Özellikler*

	<i>Frekans (f)</i>	<i>Yüzde (%)</i>
<i>Cinsiyet</i>		
<i>Erkek</i>	210	20.2
<i>Kadın</i>	828	79.8
<i>Fakülte</i>		
<i>Dişçilik Fakültesi</i>	85	8.2
<i>Eğitim Fakültesi</i>	520	50.1
<i>Fen Edebiyat Fakültesi</i>	150	14.5
<i>Güzel Sanatlar Fakültesi</i>	70	14.5
<i>İktisadi ve İdari Bilimler Fakültesi</i>	127	6.8
<i>Sağlık Bilimleri Fakültesi</i>	42	4.0
<i>Mühendislik Fakültesi</i>	48	3.7
<i>Tıp Fakültesi</i>	6	0.6
<i>Yaş</i>		
<i>17-20 yaş</i>	436	42.4
<i>21-25 yaş</i>	474	45.7
<i>26-30 yaş</i>	73	7.0
<i>+30 yaş</i>	55	5.1

Veri Toplama Araçları

Bu araştırmada kullanılan veri toplama araçları ve sahip oldukları psikometrik özellikler bu kısımda tanıtılmıştır.

Beck Depresyon Envanteri (BDE)

İlk olarak Beck ve arkadaşları (1961) tarafından geliştirilen BDE, bireylerin depresyon düzeylerini ölçmek için kullanılmaktadır. Ölçek, katılımcılar tarafından 0-3 arasında puanlanan 21 adet 4'lü Likert tipi maddeden oluşmaktadır. Ölçeğin Türkçe'ye psikiyatrik bozukluğu olmayan üniversite öğrencileri için adaptasyon çalışması Hisli (1989) tarafından yapılmıştır. Ölçeğin iç tutarlılık katsayısı 0.80, test-tekrar test güvenilirlik değeri 0.74 olarak bulunmuştur. Bulgulara dayanarak, Hisli ölçeğin üniversite öğrencileri ile kullanım için güvenilir ve geçerli olduğu sonucuna varmıştır.

Depresyon Kaygı Stres Ölçekleri (DAS-Ö)

DAS-Ö, Lovibond ve Lovibond (1996) tarafından geliştirilmiştir. Her biri 14 maddeden oluşan üç boyuttan (depresyon, kaygı ve stres) oluşmaktadır. Ölçek maddeleri 4'lü Likert türünde (0=asla-3=çoğu zaman) puanlanır ve puan aralığı her alt boyut için 0 ile 42 arasındadır. Yüksek puanlar daha şiddetli depresyon, kaygı ve / veya stres düzeylerine karşılık gelir. Ölçek, Akın ve Çetin (2007) tarafından Türkçe'ye uyarlanmış ve iyi psikometrik özelliklere sahip olduğu belirlenmiştir (toplam ölçek için $\alpha = .89$, depresyon alt ölçeği için $\alpha = .90$ ve stres ve kaygı alt ölçekleri için $\alpha = .92$). Mevcut çalışma için veri toplama sürecinde katılımcılara sadece depresyon alt ölçeği verilmiştir.

Epidemiyolojik Arařtırmalar Merkezi Depresyon Ölçeđi (EAMD-Ö)

EAMD-Ö Radloff (1977) tarafından geliştirilmiř ve depresif belirtileri ölçmek için tasarlanmıřtır. Genel evren ve klinik gruplara uygulanabilen kısa bir ölçektir. Genel depresyon tarama aracı olarak sıklıkla kullanılmaktadır. Ölçeđin maddeleri depresyonla iliřkili belirtilerden oluřmaktadır. EAMD-Ö 20 maddeden oluřur ve her bir madde 0-3 puan arasında puanlanabilir. Ölçeđin 4., 8., 12. maddeleri negatif olarak oluřturulmuřtur ve toplam puanlar hesaplanırken ters olarak kodlanmalıdır. Olası puan aralıđı 0 ila 60 arasındadır ve daha yüksek puanlar, yanıtlayıcının gösterdiđi daha depresif belirtiler olarak deđerlendirilmiřtir. Ölçek, Türkçe'ye Tatar ve Saltukođlu (2010) tarafından uyarlanmıřtır. Ölçeđin kabul edilebilir (Cronbach alfa=0.89) düzeyde güvenilirliđe sahip olduđunu belirtilmiřtir.

Verilerin Analizi

Eřitleme çalıřmaları, veri toplama deseninin belirlenmesi ile bařlar. Literatürde genel olarak benimsenen ve yaygın olarak kullanılan tasarımlar tek grup deseni, eřdeđer grup deseni, dengelenmiř gruplar deseni ve ortak madde eřdeđer olmayan grup desenidir. İlgili okuyucular için bu desenlerin detaylı açıklamaları von Davier'de (2010) yer almaktadır. Bu çalıřma için tüm depresyon ölçüm araçları katılımcıların tamamı tarafından yanıtlandıđı için tek grup deseni kullanılmıřtır. Daha sonrasında, eřitlemenin temel varsayımlarının karřılanıp karřılanmadıđına bakılır.

Varsayımların incelenmesinden sonra hangi eřitleme yönteminin kullanılacađına karar verilir. Geleneksel olarak, eřitlemede kullanılan yöntemler iki geniř yöntem ailesine ayrılır: KTK'ya dayalı yöntemler ve Madde Tepki Kuramına (MTK) dayanan yöntemler. KTK'ya dayalı yöntemler ortalama eřitleme, dođrusal eřitleme ve eřit yüzdelikli eřitleme olmak üzere üç tanedir. Özellikle, eřit yüzdelikli eřitleme, eřitlenecek test formlarının puanları arasında dođrusal olmayan iliřkilerin varlıđına izin vermesi bakımından üstünlüđe sahiptir (Kolen ve Brennan, 2014).

Günümüzde eřitleme çalıřmalarına yönelik genel kabul gören bir anlayıř çok yöntemli yaklařımın benimsenmesidir. Bu yaklařıma göre bir KTK temelli yöntem (genellikle eřit yüzdelikli eřitleme tercih edilir) ve bir MTK'ya dayalı eřitleme yöntemi eř zamanlı olarak kullanılır (Kolen ve Brennan, 2014).

MTK'ya dayalı eřitlemelerdeki önemli sorunlardan bir tanesi, MTK paradigmasında eřitleme yapılmadan önce iki temel varsayımın kontrol edilmesinin gerektiđidir. Bu varsayımlar tek boyutluluk ve yerel bađımsızlıktır. Tek boyutluluk varsayımı, tek ve baskın bir boyutun katılımcıların maddelere verdikleri yanıtları açıkladıđına iřaret etmektedir. Yerel bađımsızlık ise baskın faktörün etkisi kontrol edildikten sonra bireyin farklı maddelere verdikleri yanıtlar arasında anlamlı bir iliřki olmaması gerektiđini ima eder (Reise ve Waller, 2009). Bu çalıřma için ise MTK'ya dayalı eřitleme yaklařımı kullanılmamıřtır çünkü kullanılan depresyon ölçüm araçlarının yerel bađımsızlık varsayımının sađlanmadıđı görülmüřtür.

Eřit yüzdelikli eřitlemede ise sırasıyla X ve Y adlı eřitlenecek iki test formumuzun olduđunu varsayalım. Eřit yüzdelik eřitlemede yapılacak ilk şey, her form için yüzdelik satırlarını hesaplamaktır. İkinci ařamada ise yıđılmalı frekanslarda, aynı yüzde sırasına karřılık gelen noktalar eřitlenir (Kolen, 1984). Yani, eřitleme fonksiyonu hesaplanırken, X formundan elde edilen puanların dađılımı Y formundan elde edilen puanların dađılımına dönüřtürülür. X

formunda belirli bir yüzdeye karşılık gelen puan ve Y formunda aynı yüzdeye karşılık gelen puan eşit kabul edilir. Eşit yüzdeliği eşitleme puanlarının dağılımı hakkında herhangi bir varsayımda bulunmaz. Bu anlamda, eşit yüzdeliği eşitleme yöntemi farklı dağılımlara sahip verilerde kullanılabilir (van der Linden, 2000).

Bununla birlikte, X formundaki bir ham puanın, Y formundaki ham bir puanla aynı yüzde derecesine karşılık gelmesi nadiren mümkündür. Bu nedenle, aynı yüzde derecesine karşılık gelen ham puanları bulmak için puanların dağılımı ayrı olan dağılımları sürekli yapılmalıdır. Bu amaçla çeşitli düzeltme (smoothing) yöntemleri kullanılır (Livingston, 2014). Düzeltme, özellikle sıfır frekanslı skorlar olduğunda ve daha az rastgele hata üreten denklem fonksiyonlarının elde edilmesi için gereklidir. İki farklı düzeltme tekniği vardır: ön düzeltme ve sonradan düzeltme. Ön düzeltmede öncelikli olarak puanların dağılımı ilk adım olarak düzeltilir eşitleme sonradan yapılır. Sonradan düzeltmede ise, öncelikli olarak eşitleme yapılır ve sonradan eşitlenmiş puanlar düzeltilir (Kolen ve Brennan, 2014).

Bu varsayımlar şu şekildedir: tüm formlar aynı yapıyı ölçmelidir; tüm formlar benzer güvenilirlik düzeylerine sahip olmalıdır; eşitleme denklem fonksiyonları simetrik olmalıdır; ham puanlar ve eşitleme sonucunda kestirilen puanla aynı olmalıdır ve denklem fonksiyonları farklı alt gruplarda aynı olmalıdır (Kolen ve Brennan, 2014). Dorans ve diğ. (2010) son iki varsayımın, ilk üç varsayıma bağlı olduğunu belirtmiştir. Bu sebeple bu çalışma için sadece ilk üç varsayım test edilmiştir.

Çalışmada farklı depresyon ölçeklerinin eşitlenmesini gerçekleştirmeden önce, eşitlenecek depresyon ölçüm araçlarının aynı kavramı ölçtüğünden emin olmak gerekir. Farklı depresyon ölçeklerindeki maddeler bir araya getirilerek birleştirilmiş bir madde havuzu oluşturulursa ve bu madde havuzunun tek boyutlu bir yapıya sahip olduğu belirlenirse, eşitlenmesi amaçlanan ölçüm araçlarının aynı özelliği ölçtüğüne dair güçlü bir kanıt elde edilir. Ayrıca, ölçüm araçları arasındaki korelasyonlar, tek boyutluluk için ayrıca kanıt sağlayabilir. Ayrıca Açıklayıcı Faktör Analizi (AFA) ve Doğrulayıcı Faktör Analizi (DFA) kullanılarak da birleştirilmiş madde havuzunun tek boyutluluğuna dair kanıtlar elde edilmesi mümkündür. AFA gerçekleştirilirken, birinci boyutun öz değerinin ikinci boyutun öz değerine oranı tek boyutluluğa yönelik karar vermede kullanılmıştır. Bu oranın 4 olmasının tek boyutluluğun kanıtı olarak kullanılabilmesi belirtilmiştir (Slocum-Gori ve Zumbo, 2011). Ayrıca, birleştirilmiş madde havuzu için tek boyutluluğun incelenmesi için tek faktörlü DFA gerçekleştirilmiştir. Modelin uyumu yaygın olarak kullanılan uyum indeksleri ile değerlendirilmiştir: Comparative Fit Index (CFI: > .90), Tucker Lewis endeksi (TLI: > .90) ve root mean square error approximation (RMSEA: < .10) ve Standartized root mean residual (SRMR: < .10) (Lance, Butts ve Michels, 2006).

Ayrıca, verilerin Madde Tepki Kuramı ile analize uygunluğunu değerlendirmek için yerel bağımsızlığına da bakılmıştır. Bu doğrultuda madde çiftleri arasında Q3 istatistikleri hesaplanmıştır (Yen, 1984). Eğer herhangi bir madde çifti için Q3 istatistiği 0.3'ten fazlaysa, bu maddeler için yerel bağımlılık söz konusudur.

Son aşamada ise ön analizi bitirdikten sonra log-lineer öz düzeltme ile eşit yüzdeliği eşitleme gerçekleştirilmiştir ve puan dönüşüm tablosu oluşturulmuştur. Her bir denklem analizi için, EAMD-Ö ve DAS-Ö değerleri BDE puanlarının yer aldığı ham puan ölçeğine dönüştürülmüştür. Yani, EAMD-Ö ve DAS-Ö puanlarının BDE ölçeği için karşılık gelen puanları hesaplanmıştır.

Analiz süreci boyunca R (version 4.0.0) programından (2020) yoğun olarak faydalanılmıştır. Analizlerin farklı aşamalarında farklı paketler kullanılarak analizler gerçekleştirilmiştir: AFA analizi “psych” paketi (Revelle, 2018) ile gerçekleştirilmiştir; DFA ise “lavaan” paketi (Rosseel, 2012) ile yapılmıştır; Q3 yerel bağımlılık istatistikleri “mirt” paketi (Chalmers, 2012) ile hesaplanmıştır; ve eşit yüzdelliği eşitleme ise “equate” paketi (Albano, 2016) ile yapılmıştır.

Bulgular

Analizler test eşitleme varsayımlarının incelenmesi ile başlamıştır. İncelenen birinci varsayım eşitlenecek ölçeklerin güçlü bir şekilde ilişkili olmasıdır. Buna doğrultuda ölçekler arasındaki korelasyon katsayıları hesaplanmıştır. Elde edilen bulgulara göre BDE ($r=.74$, $p<.01$) ile DAS-Ö ($r=.75$, $p<.01$) ve EAMD-Ö arasında anlamlı düzeyde güçlü ilişki olduğu belirlenmiştir. Ayrıca, EAMD-Ö ve DAS-Ö arasındaki ilişki de anlamlıdır ($r=.79$, $p<.01$). Hesaplanan bu değerler birinci varsayımın karşılandığını göstermektedir.

Eşitleme için kabul edilen bir diğer varsayım ise eşitlenecek ölçeklerin aynı yapıyı ölçtüğüdür. Bu varsayımın incelenmesi için BDE, EAMD-Ö ve DAS-Ö maddelerinin tamamı ortak bir madde havuzunda toplanarak bu maddelerin tek boyutluluk gösterip göstermediği AFA ve DFA ile incelenmiştir. Gerçekleştirilen bu analize ait bulgular Tablo 2’de yer verilmiştir. Buna göre, DFA analizi ile elde edilen sonuçlar verinin tek boyutlu olduğunu desteklememektedir.

Diğer taraftan, AFA analizlerinden elde edilen sonuçlara göre birinci faktör toplam varyansın %34.17’sini açıkladığı ve birinci özdeğer ile ikinci özdeğer oranının 8.33 olduğu belirlenmiştir. Buna göre, DFA için elde edilen sonuçlar verinin tek boyutlu bir yapı göstermediğini fakat AFA için elde edilen sonuçlar ise verilen yanıtların altta yatan dominant bir faktör tarafından açıklanabileceğini göstermektedir.

Elde edilen bu bulgular, ikinci varsayımın kısmi olarak karşılandığına kanıttır. Son olarak, bir diğer varsayım olarak ölçeklerin benzer güvenilir düzeyinde olup olmadıklarına bakılmıştır. Tablo 2’de görüldüğü üzere, ölçeklerin güvenilirlik düzeyleri .77 ile .93 aralığındadır ve tamamı kabul edilebilir düzeyde güvenilirlerdir.

Daha sonrasında ise MTK temelli test eşitleme tekniklerinin kullanılabilirliğinin incelenmesi için tek boyutluluk ve yerel bağımlılık incelenmiştir. Tek boyutluluğun incelenmesi için her bir ölçek için ayrı ayrı AFA ve DFA gerçekleştirilmiştir. Elde edilen bulgular Tablo 2’de yer almaktadır.

Elde edilen bulgular, toplam madde havuzu için gerçekleştirilen analizlerde olduğu gibi tek boyutluluğu kısmen doğrulamaktadır. Diğer taraftan yerel bağımsızlığın incelenmesi için madde çiftleri için arasında hesaplanan Q3 istatistikleri ise kritik eşit değeri olan 0.3’ten büyük olduğu görülmüştür ve ölçeklerin hepsinin yerel bağımsızlık şartını sağlamadığı belirlendi. Bu sebepten ötürü MTK temelli eşitleme tekniklerinden yararlanılamamıştır ve sadece eş yüzdelliği eşitleme tekniği kullanılmıştır. Bir sonraki aşamada ise log-linear düzgünleştirme kullanılarak eşitleme işlemi gerçekleştirilmiştir. Elde edilen bulgulara Tablo 3’de yer verilmiştir.

Tablo 2*AFA ve DFA analizlerine ait bulgular*

	AFA			DFA				
	Alfa	1.faktör var.	1./2.özdeğ.	χ^2/df	CFI	TLI	RMSEA	SRMR
<i>BDE</i>	0.90	34.73	4.88	5.35	0.88	0.87	0.07	0.05
<i>DAS-Ö</i>	0.93	54.05	NA*	14.31	0.88	0.86	0.11	0.05
<i>EAMD-Ö</i>	0.77	36.99	4.18	8.97	0.83	0.81	0.09	0.07
<i>Birleşik havuz</i>	0.94	34.17	8.33	5.74	0.75	0.74	0.07	0.06

Not: * tek bir faktörlü yapı olduğu için birinci ve ikinci faktörün özdeğer oranı hesaplanmamıştır; N=maddes sayısı, = aritmetik ortalama; SS=standart sapma; Alfa = Cronbach Alfa değeri; 1.faktör var.=birinci faktör tarafından açıklanan varyans miktarı. 1./2.özdeğ.=birinci faktörün özdeğerinin ikinci faktörün özdeğerine oranı.

Bu işlem gerçekleştirilirken BDE ölçeği referans alınmış ve diğer iki ölçeğin eşitleme sonucunda kestirilen tahmini puanları BDE ölçeğine aktarılmıştır. BDE'nin olası her bir puanı için eşitleme sonucuna göre kestirilen DAS-Ö ve EAMD-Ö puanlarına ve her bir puana ait standart hata değerlerine parantez içinde yer verilmiştir.

Tablo 3 incelendiğinde genel olarak eşitleme ilişkilerinin doğrusal olmadığı görülmektedir. Örneğin, BDE ölçeği ile DAS-Ö ölçekleri arasındaki ilişki incelendiğinde BDE için gözlenen 52. puandan sonra DAS-Ö puanları gözlenebilecek en yüksek değer olan 42'ye ulaşmakta ve sonrasında artış göstermemektedir. Aynı şekilde BDE için gözlenen 48 puanından sonra EAMD-Ö puanları için tahmin edilen puanların artışı yavaşlamaktadır. Gözlenen bu durum, ölçek geliştirme yaklaşımlarındaki farklılaşmalardan, madde içeriklerinin farklı olmasından ya da yüksek depresyon düzeylerindeki ölçüm kesinliğinin düşük olmasından kaynaklanmış olabilir. Ayrıca, genel olarak EAMD-Ö için gözlenen hata değerlerinin daha yüksek olduğu söylenebilir.

Tablo 3*AFA BDE, EAMD-Ö ve DAS-Ö ölçekleri için elde edilen puan dönüşüm tablosu*

BDE	DAS-Ö	EAMD-Ö	BDE	DAS-Ö	EAMD-Ö
0	0.00 (0.14)	1.71 (0.77)	32	29.10 (1.06)	41.13 (1.33)
1	0.84 (0.22)	5.20 (0.95)	33	29.67 (0.84)	41.80 (1.01)
2	1.90 (0.51)	7.11 (0.56)	34	30.29 (0.78)	42.58 (1.58)
3	3.15 (0.37)	8.82 (0.44)	35	30.75 (0.74)	43.50 (1.50)
4	3.98 (0.28)	10.22 (0.61)	36	30.88 (0.72)	43.75 (1.46)
5	4.69 (0.35)	11.60 (0.63)	37	31.21 (0.69)	44.42 (1.39)
6	5.53 (0.32)	13.19 (0.74)	38	32.67 (2.51)	45.43 (1.08)
7	6.53 (0.37)	14.82 (0.57)	39	34.37 (1.69)	46.29 (0.96)
8	7.49 (0.38)	16.35 (0.60)	40	35.67 (2.05)	46.92 (1.01)
9	8.61 (0.47)	17.39 (0.47)	41	36.50 (1.45)	47.33 (0.95)
10	9.67 (0.55)	18.62 (0.52)	42	37.00 (1.34)	47.83 (1.79)
11	10.66 (0.44)	19.60 (0.55)	43	37.75 (1.20)	49.00 (2.41)
12	11.48 (0.45)	20.67 (0.76)	44	38.50 (4.22)	50.50 (2.11)
13	12.29 (0.41)	22.08 (0.70)	45	40.25 (1.77)	52.00 (3.56)
14	13.26 (0.57)	23.52 (0.69)	46	40.83 (1.02)	53.00 (1.54)
15	14.38 (0.55)	24.95 (0.77)	47	41.00 (0.96)	53.25 (1.47)
16	15.73 (0.64)	26.53 (1.06)	48	41.50 (0.82)	54.50 (2.45)
17	16.45 (0.63)	27.68 (0.82)	49	41.83 (0.61)	56.00 (2.00)
18	17.13 (0.59)	28.59 (0.69)	50	41.83 (0.61)	56.00 (2.00)
19	18.08 (0.74)	29.53 (0.67)	51	41.83 (0.61)	56.00 (2.00)
20	19.00 (0.56)	30.48 (0.65)	52	41.83 (0.61)	56.00 (2.00)
21	20.18 (1.04)	32.21 (0.93)	53	42.00 (0.50)	57.00 (1.66)
22	21.65 (0.73)	33.53 (1.13)	54	42.00 (0.38)	57.50 (1.41)
23	22.30 (0.71)	34.53 (0.91)	55	42.00 (0.38)	57.50 (1.41)
24	23.28 (0.97)	35.47 (0.87)	56	42.00 (0.38)	57.50 (1.41)
25	24.10 (0.99)	36.17 (0.82)	57	42.00 (0.38)	57.50 (1.41)
26	24.80 (0.96)	36.76 (0.84)	58	42.00 (0.38)	57.50 (1.41)
27	25.83 (1.13)	37.63 (0.90)	59	42.00 (0.38)	57.50 (1.41)
28	26.95 (1.16)	38.50 (0.92)	60	42.00 (0.38)	57.50 (1.41)
29	27.74 (0.71)	39.21 (0.86)	61	42.00 (0.38)	57.50 (1.41)
30	28.24 (0.67)	39.95 (1.14)	62	42.00 (0.38)	57.50 (1.41)
31	28.65 (1.10)	40.56 (1.38)	63	42.00 (0.25)	58.00 (0.87)

Tartışma, Sonuç ve Öneriler

Farklı birçok araştırmacı günümüze kadar depresyon puanlarının karşılaştırılabilmesi için ortak bir metriğe aktarılmasının gerekliliğini belirmiş ve farklı ülkelerde bu doğrultuda bir puan dönüşüm tablosu oluşturmaya yönelik çalışmaların yürütmüşlerdir (Fischer, 2011; Zhao Chan ve Lo, 2017). Gerçekleştirilen bu çalışma ise Türkiye’de yürütülen ilk çalışma olması sebebiyle var olan çalışmalardan farklılaşmaktadır. Bu doğrultuda, gerçekleştirilen eşitleme işlemi sonucunda iki farklı puan dönüşüm tablosu oluşturulmuştur. Bu sayede araştırmacıların, herhangi bir ölçme aracından elde edilen puanları diğer ölçekten elde edilen puanlarla karşılaştırabilmeleri sağlanmıştır. Ayrıca, gerçekleştirilen bu çalışmada tek grup eşitleme deseni kullanılması sebebiyle avantajlıdır çünkü ortak madde desenindeki ortak maddelerin ölçeğin minyatürü

olması gibi gereklilikler ve bu gerekliliğin sağlanmadığındaki oluşacak ek hata kaynakları tek grup deseni için geçerli değildir. Nitekim, ölçme araçlarının tamamının katılımcıların tamamına uygulanması oluşturulan eşitleme bağlantılarının kesinliğini arttıran bir faktördür (Dorans ve diğ., 2010).

Depresyon ölçüm araçlarının puan karşılaştırmaları incelendiğinde BDE ölçeği için hafif düzey depresyon belirti düzeyi için belirlenen kesme puanının (10/63) DAS-Ö'nün depresyon alt boyutu için belirlenen kesme puanı (9.67/42) ile yaklaşık olarak örtüştüğü görülmektedir. Diğer taraftan, EAMD-Ö için belirlenen kesme puanının (16-60) ise BDE'de şiddet olarak "hafif" düzeye karşılık geldiği görülmüştür. Yani, depresyon ölçüm araçları için belirlenen kesme puanları örtüşmediği söylenebilir. Ayrıca, DAS-Ö depresyon alt boyutu için elde edilebilecek maksimum puan BDE'de 53 puana karşılık gelmektedir. Diğer taraftan, BDE için elde edilebilecek maksimum değer olan 63 puan EAMD-Ö'de elde edilebilecek maksimum puanın altındadır (58/60). Bu bulgulardan hareketle, depresyon ölçeklerinin ölçtükleri depresyon şiddet yelpazesinin örtüşmediği söylenebilir. Özellikle EAMD-Ö yüksek depresyon düzeyleri için daha geniş puan aralıkları sunduğu ve daha hassas ölçüm gerçekleştirilmesine olanak sağladığı söylenebilir.

Dönüşüm tablolarının kullanılması ile birlikte, araştırmacılar elde ettikleri puanları farklı ölçeklerin kullanıldığı araştırmalarda elde edilen puanlar ile de karşılaştırma fırsatı bulabilirler. Özellikle, meta analiz çalışmalarında farklı ölçekler kullanılarak gerçekleştirilen çalışmaların dâhil edilmesi araştırmacıların çalışma seçme sürecindeki esnekliğini arttıracak ve daha doğru sonuçlar elde edebilmeleri mümkün olacaktır.

Ayrıca, bu çalışmada mümkün olduğunca büyük ve heterojen bir çalışma grubuna ulaşılmaya çalışılmıştır. Bu sayede eşitlenen puanların uyumunun arttırılması ve çalışmanın gerçekleştirildiği çalışma grubuna bağlılığının azaltılması amaçlanmıştır. Yine de, puan karşılaştırma tablosu kullanılırken bu çalışma grubuna bağlılık durumu göz önünde bulundurulmalıdır. Nitekim hangi teknik kullanılırsa kullanılsın hiçbir eşitleme fonksiyonu tamamen çalışma grubundan bağımsız değildir. Bununla ilişkili olarak, gerçekleştirilen bu çalışmada üniversite öğrencileri çalışmaya dâhil edilmiştir ve elde edilen eşitleme fonksiyonlarının farklı evrenlere genellenebilmesi mümkün olmadığı unutulmaması gereken bir diğer durumdur. Ayrıca, genellenebilirliği etkilemesi olası bir diğer konu ise dengesiz cinsiyet dağılımıdır. Denk olmayan cinsiyet dağılımı, elde edilen eşitleme fonksiyonlarının genellenebilirliğini düşürmesi olasıdır.

Bu çalışmada sadece tek bir eşitleme yönteminin kullanılmıştır. Yerel bağımsızlığın sağlanmaması beklenen bir durumdur çünkü depresyon ölçeklerinin benzer içeriğe sahip olması ve maddelerin geniş kabul görmüş aynı belirtilere dayanılarak oluşturulmasından dolayı en azından bazı maddeler birbirlerine ölçülen özelliğin ötesinde başka faktörlerle bağlı olabilirler. Bu çalışmada, genel uygulayışın aksine tek bir yöntemin kullanılmasının bir sorun oluşturmayacağı düşünülmektedir çünkü farklı depresyon ölçeklerine ait veriler aynı bireylerden toplandığı için ve büyük bir çalışma grubu ile uygulama gerçekleştirildiğinden dolayı gruplar benzer yetenek dağılımına sahip olacaktır. Bu durumda ise eşit yüzdelikli eşitleme ve MTK'ya dayalı eşitleme yöntemleri benzer sonuçlar verecektir (Skaggs ve Lissitz, 1986). Fakat nihai olarak, eşit yüzdelikli eşitleme yöntemi ile elde edilen mevcut bulgular farklı bir yöntem ile elde edilecek bulgularla doğrulanamamıştır. Bundan sonraki çalışmalarda, özellikle MTK temelli eşitleme yöntemlerinin de kullanılması önerilmektedir. Aynı zamanda, gerçekleştirilen bu

çalışmada veriler yalnızca üniversite öğrencilerinden toplanmıştır. Gelecekte yürütülecek olan çalışmalarda, eşitleme fonksiyonlarının farklı evrenler için oluşturulması önerilmektedir.

Çıkar Çatışması Bildirimi

Yazar(lar), bu makalenin araştırılması, yazarlığı ve/veya yayınlanmasına ilişkin herhangi bir potansiyel çıkar çatışması beyan etmemiştir.

Destek/Finansman Bilgileri

Yazar(lar), bu makalenin araştırılması, yazarlığı ve / veya yayınlanması için herhangi bir finansal destek almamıştır.

Etik Kurul Kararı

Bu araştırma için Hasan Kalyoncu Üniversitesi Sosyal ve Beşeri Bilimler Etik Kurulu kurumdan (02.07.2020 tarih ve -804.01-2007020010 sayılı) etik izin alınmıştır.

Kaynakça/References

- Akgündüz, D., Aydeniz, M., Çakmakçı, G., Çavaş, B., Çorlu, M. S., Öner, T. ve Özdemir, S. (2015). *STEM eğitimi Türkiye raporu: Günün modası mı yoksa gereksinimi?* İstanbul: Aydın Üniversitesi.
- Akın, A., & Çetin, B. (2007). Depresyon Anksiyete Stres Ölçeği (DASÖ): Geçerlik ve Güvenirlik Çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 7(1), 241-268.
- Albano, A. D. (2016). equate: An R package for observed-score linking and equating. *Journal of Statistical Software*, 74(8), 1-36.
- Albrecht, A. T., & Herrick, C. (2010). *100 questions & answers about depression*. Jones & Bartlett Learning.
- Angoff, W. (1996). Scales, norms, and equivalent scores. *Educational Measurement: Theories and applications*, 121.
- Baguley, T. (2009). Standardized or simple effect size: What should be reported?. *British journal of psychology*, 100(3), 603-617.
- Bahar, A., Tutkun, H., & Sertbas, G. (2005). Huzurevinde yasayan yasli larin anksiyete ve depresyon duzeylerinin belirlenmesi. *Anadolu Psikiyatri Dergisi*, 6(4), 227.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry*, 4(6), 561-571.
- Blackburn, I. M. (1987). *Coping with depression*. Chambers Harrap Publishers Ltd.
- Bozkurt, N. (2004). Bir grup üniversite öğrencisinin depresyon ve kaygı düzeyleri ile çeşitli değişkenler arasındaki ilişkiler. *Eğitim ve Bilim*, 29(133), 52-59.
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of statistical Software*, 48(6), 1-29.
- Compton, W. M., Conway, K. P., Stinson, F. S., & Grant, B. F. (2006). Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. *American Journal of Psychiatry*, 163(12), 2141-2147.
- Davidson, G. C., & Neale, J. M. (1998). *Abnormal psychology*. John Wiley.
- Dorans, N. J., Moses, T. P., & Eignor, D. R. (2010). *Principles and practices of test score equating*. ETS Research Report Series, 2010(2), i-41.
- Field, T., Miguel, D., & Sanders, C. (2001). Adolescent depression and risk factors. *Adolescence*, 36(143), 491.
- Fischer, H. F., Tritt, K., Klapp, B. F., & Fliege, H. (2011). How to compare scores from different depression scales: Equating the Patient Health Questionnaire (PHQ) and the ICD-10-Symptom Rating (ISR) using item response theory. *International journal of methods in psychiatric research*, 20(4), 203-214.
- Gao, K., Su, M., Sweet, J., & Calabrese, J. R. (2019). Correlation between depression/anxiety symptom severity and quality of life in patients with major depressive disorder or bipolar disorder. *Journal of affective disorders*, 244, 9-15.
- Günay, O., Öztürk, A., & Arslantaş, E. E. (2018). Öğrencilerinde İnternet Bağımlılığı ve Depresyon Düzeyleri. *Neurological Sciences*, 31, 79-88.
- Hisli, N. (1989). Beck depresyon envanterinin üniversite öğrencileri için geçerliliği, güvenilirliği. (A reliability and validity study of Beck Depression Inventory in a university student sample). *Journal of Psychology*, 7, 3-13.
- İskender, H., Dokumacıoğlu, E., Kanbay, Y., & Kılıç, N. (2018). Üniversite öğrencilerinde sağlıklı yaşam ve depresyon puan düzeyleri ile ilgili faktörlerin belirlenmesi. *Acıbadem Üniversitesi Sağlık Bilimleri Dergisi*, 9(4), 414-423.
- Karasar, N. (2005). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.
- Koivumaa-Honkanen, H., Kaprio, J., Honkanen, R., Viinamäki, H., & Koskenvuo, M. (2004). Life satisfaction and depression in a 15-year follow-up of healthy adults. *Social psychiatry and psychiatric epidemiology*, 39(12), 994-999.

- Kolen, M. J. (1984). Effectiveness of analytic smoothing in equipercentile equating. *Journal of Educational Statistics*, 9(1), 25-44.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices*. Springer Science & Business Media.
- Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say?. *Organizational research methods*, 9(2), 202-220.
- Livingston, S. A. (2014). *Equating test scores (without IRT)*. Educational testing service.
- Lord, F. M., & Novick, M. R. (2008). *Statistical theories of mental test scores*. IAP.
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour research and therapy*, 33(3), 335-343.
- Otte, C., Wingenfeld, K., Kuehl, L. K., Kaczmarczyk, M., Richter, S., Quante, A., ... & Hinkelmann, K. (2015). Mineralocorticoid receptor stimulation improves cognitive function and decreases cortisol secretion in depressed patients and healthy individuals. *Neuropsychopharmacology*, 40(2), 386-393.
- Ören, N., Gençdoğan, B. (2007). Lise öğrencilerinin depresyon düzeylerinin bazı değişkenlere göre incelenmesi. *Kastamonu Eğitim Dergisi*, 15(1), 83-92.
- Penninx, B. W., Beekman, A. T., Honig, A., Deeg, D. J., Schoevers, R. A., Van Eijk, J. T., & Van Tilburg, W. (2001). Depression and cardiac mortality: results from a community-based longitudinal study. *Archives of general psychiatry*, 58(3), 221-227.
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385-401.
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual review of clinical psychology*, 5, 27-48.
- Revelle, W. (2018). *psych: Procedures for psychological, psychometric, and personality research*. R package version, 1(10).
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of statistical software*, 48(2), 1-36.
- Skaggs, G., & Lissitz, R. W. (1986). IRT test equating: Relevant issues and a review of recent research. *Review of Educational Research*, 56(4), 495-529.
- Slocum-Gori, S. L., & Zumbo, B. D. (2011). Assessing the unidimensionality of psychological scales: Using multiple criteria from factor analysis. *Social Indicators Research*, 102(3), 443-461.
- Softa Kaçan, H. (2013). Engelli çocuğa sahip ebeveynlerin depresyon düzeylerinin incelenmesi. *Kastamonu Üniversitesi Kastamonu Eğitim Dergisi*, 21(2), 589-600.
- Stunkard, A. J., Faith, M. S., & Allison, K. C. (2003). *Depression and obesity*. *Biological psychiatry*, 54(3), 330-337.
- Tatar, A., & Saltukoglu, G. (2010). 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. *Klinik Psikofarmakoloji Bülteni-Bulletin of Clinical Psychopharmacology*, 20(3), 213-227.
- Tezel, A., Arslan, S., Topal, M., Aydoğan, Ö., Çiğdem, K. O. Ç., & Şenlik, M. (2009). Hemşirelik öğrencilerinin problem çözme becerileri ve depresyon düzeylerinin incelenmesi. *Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi*, 12(4), 1-10.
- Ulas, B, Tatlibadem, B, Nazik, F, Sonmez, M, Uncu, F. (2016). Üniversite öğrencilerinde depresyon sıklığı ve ilişkili etmenler. *Celal Bayar Üniversitesi Sağlık Bilimleri Enstitüsü Dergisi*, 2(3), 71-75.
- van der Linden, W. J. (2000). A test-theoretic approach to observed-score equating. *Psychometrika*, 65(4), 437-456.

- von Davier, A. (Ed.). (2010). *Statistical models for test equating, scaling, and linking*. Springer Science & Business Media.
- WHO International Consortium in Psychiatric Epidemiology (2017, April14). *Depression and other common mental disorders global health estimates*. <http://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf>. adresinden alınmıştır.
- Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement*, 8(2), 125-145.
- Zhao, Y., Chan, W., & Lo, B. C. Y. (2017). Comparing five depression measures in depressed Chinese patients using item response theory: an examination of item properties, measurement precision and score comparability. *Health and Quality of Life Outcomes*, 15(1), 1-14.

İletişim/Correspondence

Doç. Dr. Akif AVCU

avcuakif@gmail.com

Prof. Dr. Şaziye Senem BAŞGÜL

senembasgul@gmail.com