

Likert Tipi Ölçeklerde Rasch Modelinin Kullanımı: Olumsuz Değerlendirilme Korkusu Ölçeği-Öğrenci Formu (ODKÖ-ÖF) Üzerinde Bir Uygulama

The Use of Rasch Model in Likert Types Scales: An Application on the Fear of Negative Evaluation Scale-Student Form (FNE-SF)

Mustafa İLHAN¹, Neşe GÜLER²

Öz: Bu çalışmada, Likert tipi ölçeklerin psikometrik özelliklerinin incelenmesinde ve bu ölçeklerden alınan puanlar üzerinden gerçekleştirilen farka dayalı istatistiklerde Rasch modelinin kullanımına örnek olabilecek bir çalışmanın alanyazına kazandırılması amaçlanmıştır. Araştırma, 367 ortaöğretim öğrencisinden oluşan bir çalışma grubu üzerinde yürütülmüştür. Araştırmanın verileri Olumsuz Değerlendirilme Korkusu Ölçeği-Öğrenci Formu (ODKÖ-ÖF) ile Akademik Beklentilere İlişkin Stres Envanteri (ABSE) kullanılarak toplanmıştır. Çalışma kapsamında toplanan veriler FACETS paket programından yararlanılarak Rasch modeline göre analiz edilmiştir. Araştırmadan elde edilen bulgular; ODKÖ-ÖF'ye ilişkin 16 maddeden oluşan tek boyutlu modelde öğrenciler ve maddelerin yüksek güvenirlikte birbirinden ayırt edildiğini, ölçekte kullanılan beşli derecelendirmenin etkin bir biçimde çalıştığını, uyum istatistiklerinin kabul edilebilir sınırlar içerisinde kaldığını ve gözlenen ile beklenen test karakteristik eğrilerinin büyük ölçüde örtüştüğünü göstermiştir. Çok yüzeyli Rasch analizi sonucunda, öğrencilerin sınıf ortamında öğretmenleri ya da arkadaşları tarafından olumsuz değerlendirileceklerine dair yaşadıkları korkunun ABSE puanlarına göre farklılaştığı saptanmış ve bu bulgu ODKÖ-ÖF'nin ölçüt geçerliğine yönelik bir kanıt olarak yorumlanmıştır.

Anahtar sözcükler: Rasch modeli, sıralama ölçekli model, çok yüzeyli Rasch analizi, olumsuz değerlendirilme korkusu.

Abstract: This study aims to introduce a new study into the literature that serves as an example of an investigation into psychometric characteristics of Likert-type scales, as well as an example of the utilization of the Rasch model for determining statistics based on the differences in the scores obtained by using these scales. The study was conducted on a study group who consisted of 367 secondary school students. Data were collected using Fear of Negative Evaluation Scale-Student Form (FNE-SF) and Academic Expectations Stress Inventory (AESI). The collected data were analyzed in accordance with the Rasch model using the FACETS packet program. The findings obtained in the study show that students and items could be distinguished from one another with a high reliability, the 5-point structure of the scale worked effectively, the fit indexes fell within acceptable limits and the observed and expected test characteristic curves were overlap to a considerable degree. As a result of many facet Rasch analysis it was determined that significant differentiation among students' fears of negative evaluations according to their AESI scores and this finding was considered as evidence for criterion validity of FNE-SF.

Keywords: Rasch model, rating scale model, many facet Rasch model, fear of negative evaluation

EXTENDED ABSTRACT

Introduction

This study aims to introduce a new study into the literature that serves as an example of an investigation into psychometric characteristics of Likert-type scales, as well as an example of the

¹ Dr. Öğr. Üys., Dicle Üniversitesi, e-posta: mustafailhan21@gmail.com

² Doç. Dr., İzmir Demokrasi Üniversitesi, e-posta: gnguler@gmail.com

Cite this article as:

İlhan, M., & Güler, N. (2018). Likert Tipi Ölçeklerde Rasch Modelinin Kullanımı: Olumsuz Değerlendirilme Korkusu Ölçeği-Öğrenci Formu (ODKÖ-ÖF) Üzerinde Bir Uygulama. *Trakya Eğitim Dergisi*, 8(4), 756-775.

utilization of the Rasch model for determining statistics based on the differences in the scores obtained by using these scales. Fear of Negative Evaluation Scale-Student Form (FNE-SF) was developed within this study and the Rasch-model analysis aided in determining the psychometric characteristics of the scale. Additionally, in accordance with the criterion validity of FNE-SF, a Rasch analysis was undertaken to determine whether students' scores in the scale differed according to the stress variable of academic expectations.

Method

The study was conducted with 367 secondary school students, 200 males (54.50%) and 167 females (45.50%). Data were collected using the Academic Expectations Stress Inventory (AESI) and FNE-SF. The AESI was developed by Ang and Huan (2006) and adapted into Turkish by Kelecioğlu and Bilge (2009). The scale consisted of nine 5-point Likert-type items and included two factors: Expectations of Parents/Teachers, and Expectations of Self. A high point obtained in the sub-dimensions of the scale and an overall high point in the scale indicates a high-stress level related to academic expectations. In this study, the scales overall scores—rather than scores from the AESI sub-scales—were processed and the Cronbach Alpha internal consistency found to be .81 for the overall AESI.

The FNE-SF was developed by the researchers by taking the Brief Fear of Negative Evaluation-General Form (BFNE-GF), as developed by Watson and Friend (1969), and updating it per the Brief Fear of Negative Evaluation-Short Form (BFNE-ShF), as devised by Leary (1983). An item pool consisting of 20 items was created by formulating BFNE-GF and BFNE-ShF items so to better fit them to an academic context and classroom environment and by benefitting from the literature. The theoretical structure regarding the fear of negative evaluation predicts a single factorial model; hence the item pool of BFNE-SF was created with reference to a single factorial model. Subsequently, expert opinions were consulted to ensure the content and face validities of the scale. Based on expert opinion, several items were reformulated, one item was excluded from the scale, and a new item introduced and included in the study. Hence, a draft scale was devised that consisted of 20 items with 5-point Likert-type ranging from Strongly Agree (5) to Strongly Disagree (1).

Data were collected in November 2015 and subsequently analyzed using the FACETS software program, and in accordance with the Rasch model. While a two-facet design involving students and items was used to determine the psychometric characteristics of FNE-SF, a three-facet design—including students, items, and the AESI scores—was employed for those analyses investigating whether students' scores within the scale differed according to the stress variable regarding academic expectations.

Results

First, a two-facet Rasch analysis was conducted to investigate the psychometric characteristics of FNE-SF. The analysis revealed that three of 20 items in the draft scale negatively influenced the model-data fit, and that one item had a low value regarding the point biserial correlation. Consequently, these four items were excluded from the scale and the Rasch analysis repeated with 16 items; this analysis indicated acceptable fit indexes regarding these items and that the item correlations were found to exceed the .30 limit. Subsequently, the analysis reports on the items and person facets were subject to an in-depth examination using the variables map and the categories' statistics. A single factorial model consisting of 16 items was developed so that individuals and items could be distinguished from one another with a high reliability. The category statistics table was able to determine that the 5-point structure of the scale worked effectively. The fit indexes of both items and person facets fell within a .5–1.5 range, which is recommended for such statistics. Finally, observed and expected test characteristic curves were matched to a considerable degree.

Once the two-facet Rasch analysis was conducted, the hypothesis on the relationship between students' fear of negative evaluation and their AESI scores was tested. This served as an example of the Rasch model used to determine the statistics based on differences among the scores obtained per Likert-type scales, as well as to find evidence regarding the criterion validity of FNE-SF. First, a cluster analysis was conducted to categorize students into homogenous groups according to their AESI scores. The cluster analysis categorized the 367 students into two groups: The first groups consisted of 265 students and had a mean AESI score of 4.09; consequently, these

students' AESI scores were described as *High*. The second group consisted of 102 students with a mean AESI score of 2.73; consequently, these students' AESI scores were described as *Low*. Subsequently, these two clusters were included in the Rasch analysis as a variable; the Rasch analysis was conducted per a three-facet design including students, items, and AESI scores. The results revealed significant differentiation among students' fears of negative evaluations according to their AESI scores.

Conclusions

A scale was developed in this study to measure student's fear of negative evaluations by their teachers or peers within a classroom setting. It is hoped that this survey will inform existing Turkish literature in the field and serve as an example investigation into the psychometric characteristics of Likert-type scales, as well as for the utilization of the Rasch model to determine statistics-based differentiation per the scores obtained using these scales.

1. GİRİŞ

Bu araştırma, Likert tipi ölçeklerin psikometrik özelliklerinin incelenmesinde ve bu ölçeklerden alınan puanlar üzerinden gerçekleştirilen farka dayalı istatistiklerde Rasch modelinin kullanımına örnek olabilecek bir çalışmayı alanyazına kazandırma düşüncesiyle ortaya çıkmıştır. Araştırmada, Olumsuz Değerlendirilme Korkusu Ölçeği-Öğrenci Formu (ODKÖ-ÖF) geliştirilmiş ve ölçeğin psikometrik özelliklerinin belirlenmesinde Rasch analizinden yararlanılmıştır. Ayrıca ODKÖ-ÖF'nin ölçüt geçerliği kapsamında, öğrencilerin ölçekten aldıkları puanların akademik beklentilere ilişkin stres değişkenine göre farklılık gösterip göstermediği yine Rasch analizi kullanılarak test edilmiştir. Araştırmanın alanyazına sağlayacağı katkıyı arttırmak amacıyla, Rasch analizinde alanyazındaki mevcut ölçeklerden biri yerine, araştırmacılar tarafından geliştirilen ODKÖ-ÖF kullanılmıştır. Rasch modeli ve olumsuz değerlendirilme korkusu ile ilgili kuramsal bilgilere aşağıda yer verilmiştir.

1.1. Rasch Modeli

Rasch modeli, madde tepki kuramı çatısı altında yer alan bir parametrelili modeldir (Hasford & Bradley, 2011). Yalnızca madde güçlük parametresini içeren bu model, başlangıçta doğru/yanlış şeklinde puanlanan iki kategorili (dichotomous) maddeler için geliştirilmiştir (Haiyang, 2010). Daha sonra, Andrich (1978) sıralama ölçekli modeli (rating scale model) geliştirerek Rasch modelini Likert tipi ölçek verilerinin analizinde kullanılabilecek şekilde genişletmiştir. Likert tipi ölçeklerde Rasch modelinin kullanılması, klasik test kuramına dayalı yöntemlere ilişkin çeşitli sınırlılıkların üstesinden gelmeye yardımcı olmaktadır.

İlk olarak; madde tepki kuramı çatısı altında yer alan tüm modellerde olduğu gibi Rasch modelinde de bireylerin yetenek düzeyleri, ölçme aracındaki madde örneklemeden bağımsız olarak kestirilebilir ve madde parametreleri, testin uygulandığı gruptaki bireylerin yetenek düzeylerinden bağımsız olarak hesaplanabilir (Engelhard, 2013). İkinci olarak, klasik test kuramında bireylerin yetenek kestirimleri, ölçek maddelerine verdikleri yanıtların toplanmasıyla elde edilir. Bu yaklaşım, yetenek kestirimi hesaplanırken ölçek maddelerinin güçlük düzeyleri (herhangi bir ölçek maddesinde katılma ya da katılmama olasılığı) arasındaki potansiyel farklılıkların göz ardı edilmesine neden olur. Klasik test kuramında ölçekteki tüm maddelerin güçlük düzeylerinin benzer olduğu varsayılır. Ancak ölçek maddelerinden herhangi birinin, ölçülmesi hedeflenen özelliğin farklı bir yönünü temsil etmesi halinde, tüm maddelerin ölçekten alınacak toplam puana eşit katkı sağladığı şeklindeki bir varsayım yanıltıcı olabilmektedir (Anshel, Weatherby, Kang & Watson, 2009). Üçüncü olarak, Likert tipi ölçek verileri sıralama düzeyinde olmasına rağmen klasik test kuramında bu veriler, eşit aralık düzeyindeymiş gibi işlem görür ve ölçekten alınan puanlar toplanarak parametrik testlerle analiz edilir. Sıralama ölçeğindeki veriler toplanabilir olmadığından, toplanan ölçek puanları ve bu puanlara uygulanan parametrik testler yanlış sonuçlar üretebilir (Brinthaup & Kang, 2014). Rasch modeli, sıralama ölçeğindeki orijinal verileri eşit aralıklı logit ölçeğine dönüştürerek bu sınırlılığın aşılmasını sağlar (Wright, & Masters, 1982). Rasch modelinin bir başka avantajı, ölçek kategorilerinin ne derece iyi çalıştığına belirlenmesine olanak tanınmasıdır (Linacre, 2014). Rasch analizi çıktılarında yer alan kategori istatistikleri tablosu, ölçekte kullanılan derecelendirmenin etkin bir biçimde çalışıp çalışmadığına ilişkin bir kanıt oluşturur. Rasch analiz sonucunda rapor edilen logit cetvel incelenerek, ölçeğe eklenmesi gereken madde olup olmadığına tespit edilebilmesi Rasch modelinin güçlü yönlerinden bir diğeridir. Rasch analizinde, maddelerin ve bireylerin logit cetvel üzerinde gösterdikleri dağılım,

maddeler ile örneklemin ne kadar iyi eşleştiği ve ölçeğe eklenecek daha kolay ya da daha güç maddelerin (işaretlenme olasılığı daha yüksek ya da düşük) ölçme aracının ayırt ediciliğini nasıl etkileyeceği hakkında bir ipucu niteliği taşır (Brinthaup & Kang, 2014). Son olarak, Rasch modelinde madde yüzeyinin yanı sıra birey yüzeyine ilişkin güvenilirlik katsayısı da rapor edilmekte ve bu sayede ölçeğin uygulandığı gruptaki bireylerin hangi güvenilirlikte birbirinden ayırt edildiği belirlenebilmektedir (Güler, İlhan, Güneyle & Demir, 2017). Oysa klasik test kuramında birey yüzeyindeki değişkenliklerin tümüyle katılımcılar arasındaki bireysel farklılıklardan kaynaklandığı kabul edilmekte ve ölçek maddelerini yanıtlayan kişilerin hangi güvenilirlikte ayırt edildiğine dair bir bilgi sunulmamaktadır (Taşdelen Teker, Güler & Kaya Uyamık, 2015). Ölçeği yanıtlayan bireylerin hangi güvenilirlikte ayırt edildiği hakkında bilgi vermesi, Rasch modelinin klasik test kuramına göre bir diğer üstün yönüdür.

1.2. İki Yüzeyle ve Çok Yüzeyle Rasch Modeli

Rasch modeli ölçme sonuçlarını etkileyebilecek değişkenlik kaynaklarının sayısına göre, iki yüzeyle ve çok yüzeyle model şeklinde iki başlıkta incelenebilir. İki yüzeyle modelde, bireyler ve maddeler şeklinde iki değişkenlik kaynağı bulunur. İki yüzeyle Rasch modeline ilişkin eşitlik aşağıda gösterilmiştir.

$$\log \frac{P_{nik}}{P_{nik-1}} = B_n - D_i - F_k$$

Bu eşitlikte, P_{nik} ; n numaralı öğrencinin i numaralı maddeye k kategorisine karşılık gelen cevap verme olasılığını göstermektedir. P_{nik-1} ise n numaralı öğrencinin i numaralı maddeye $k-1$ kategorisine karşılık gelen cevap verme olasılığını temsil etmektedir. Eşitlikteki, B_n n numaralı öğrencinin ölçülen örtük özelliğe ilişkin yetenek düzeyini, D_i i numaralı maddenin güçlük düzeyini ve F_k $k-1$ kategorisinden k kategorisine geçişin güçlük düzeyini ifade etmektedir (Linacre, 2002). Araştırmada, ODKÖ-ÖF'nin psikometrik özelliklerinin belirlenmesinde iki yüzeyle Rasch modeli kullanılmıştır.

Çok yüzeyle Rasch modelinde ise birey ve madde yüzeylerinin yanı sıra ölçme sonuçlarını etkileyebilecek diğer değişkenlik kaynakları da modele dâhil olmakta ve ölçme modeline eklenen her bir değişkenlik kaynağının test puanlarını sistematik bir şekilde etkilediği kabul edilmektedir (Eckes, 2009). Örneğin, akademik beklentilere ilişkin stres değişkeninin ODKÖ-ÖF'den alınan puanları etkileme potansiyelinin olduğu düşünülüyorsa, bu değişken de Rasch analizine eklenmekte ve üç yüzeyle bir model söz konusu olmaktadır. Aşağıdaki eşitlik; birey, madde ve Akademik Beklentilere İlişkin Stres Envanteri (ABSE) puanları şeklinde üç yüzeyle içeren bir Rasch modeline karşılık gelir.

$$\ln \frac{P_{nik}}{P_{nik-1}} = B_n - D_i - \alpha_j - F_k$$

Eşitlikteki α_j akademik beklentilere ilişkin stres değişkenini temsil eder. Görüldüğü gibi bu eşitlikte iki yüzeyle modelden farklı olarak, ABSE puanları da Rasch modeline girmiştir. Araştırmada, öğrencilerin olumsuz değerlendirme korkularının ABSE puanlarına göre farklılık gösterip göstermediği üç yüzeyle bu modelden yararlanılarak test edilmiştir. Böylece ODKÖ-ÖF'nin ölçüt geçerliğine ilişkin kanıt elde edilmesi hedeflenmiştir.

1.3. Olumsuz Değerlendirilme Korkusu

Olumsuz değerlendirilme korkusu ilk olarak 1969 yılında Watson ve Friend tarafından kavramsallaştırılmış olup (Crawford, Leuzinger, Brannon, Hamner, 2015) kişinin başkaları tarafından kötüleyici ve düşmanca eleştirileceğine dair yaşadığı kaygıyı ifade etmektedir (Leary, 1983). Devamlılık taşıyan bu yüksek düzeyli kaygıdan dolayı; olumsuz değerlendirilme korkusu yaşayan kişiler başkalarının düşüncelerine çok fazla önem verir, diğer insanların olumlu düşüncelerini kazanmak ve kazandıkları olumlu düşünceleri korumak için yoğun gayret sarf ederler (Utschig, Presnell, Madeley & Smits, 2010). Diğer bir ifadeyle, olumsuz değerlendirilme korkusu yaşayan bireyler, diğer insanların onayını kazanma düşüncesiyle motive olur ve başkaları tarafından reddedilmemek için çalışırlar (Leary & Kowalski, 1995). Bu durum, yüksek düzeyde

olumsuz değerlendirilme korkusu yaşayan kişinin *i*) diğer bireylerin kendilerinden çok yüksek başarı beklentisi içinde olduklarını düşünmesine, *ii*) yeteneği ve performansı ile ilgili kuşku duymasına (Rapee & Heimberg, 1997), *iii*) sosyal çevrenin beklentileri ve toplumsal standartlar ile uyumlu olma konusunda yüksek baskı hissetmesine (Utschig vd., 2010), *iv*) bir hata yapması halinde alay konusu olacağı şeklinde koşullu inançlar geliştirmesine (Lubetkin & Oumano, 1991) ve *v*) tehdit olarak algıladıkları kişisel girişimlerden kaçınarak güvenlik sağlayıcı davranışlara yönelmesine neden olmaktadır (Leary, 1983; Rapee & Heimberg, 1997). Güvenlik sağlayıcı bu davranışların en tipik örneği, kişinin başka bireylerin değerlendirmesiyle karşı karşıya kalacağı pozisyonlardan uzak durmasıdır (Weeks, Heimberg & Rodebaugh, 2008). Ancak, bazen bireyin değerlendirileceği ortamlardan uzak kalma yönündeki çabası sonuçsuz kalabilmekte ve başkaları tarafından değerlendirilmesi kaçınılmaz olmaktadır. Bu durumda, olumsuz değerlendirilme korkusu değerlendirme sürecinde yaşanan yüksek düzeydeki tedirginlik ve gerginlik şeklinde kendisini göstermektedir (Moeni & Nejhad, 2014).

Olumsuz değerlendirilme korkusu başlangıçta bireyin günlük yaşamdaki davranışlarını etkileyen psikolojik bir yapı olarak alanyazına girmiştir. Ancak daha sonra yapılan araştırmalar, olumsuz değerlendirilme korkusunun aynı zamanda, öğrencilerin öğrenme sürecindeki davranışlarının şekillenmesinde etkili olan duyuşsal bir özellik olduğunu göstermiştir (Çetin, İlhan & Yılmaz, 2014). Sınıfın öğrencilerin sürekli olarak değerlendirmelere tabi tutulduğu bir ortam olması (Özgüngör, 2006), olumsuz değerlendirilme korkusunun öğrenme sürecini etkileyen duyuşsal bir özellik olarak ele alınmasına kaynaklık eden en önemli nedenlerden biridir. Olumsuz değerlendirilme korkusunun öğrenme sürecini etkileyen duyuşsal bir özellik olarak kabul edilmesiyle birlikte, öğrenme sürecine ilişkin çeşitli değişkenler ile olumsuz değerlendirilme korkusu arasındaki ilişkilerin incelendiği çok sayıda araştırma yapılmıştır. Bu araştırmalardan elde edilen bulgular; olumsuz değerlendirilme korkusunun; akademik beklentilere ilişkin stres (Kelecioğlu & Bilge, 2009), ders ve sınav kaygısı (Aydın, Yavuz & Yeşilyurt, 2006; Çam, Sevimli & Yerlikaya, 2010), başarısızlık korkusu (Dinnel vd., 2002), performans yönelimi (VandeWalle, 1997), akademik erteleme (Berber Çelik, 2014) ile pozitif; motivasyon, ilgi, tutum (Çekirdek, 2014), benlik saygısı (Dinnel vd., 2002), öğrenme yönelimi (McKinney, 2003), akademik risk alma (Çetin, İlhan & Yılmaz, 2014) ve akademik başarı (Alkan, 2015; Çekirdek, 2014; Uzun, 2017) ile negatif anlamlı ilişki içerisinde olduğunu ortaya koymuştur. Buna göre, olumsuz değerlendirilme korkusunun öğrenme çıktılarında farklılaşmaya neden olan önemli bir değişken olduğu söylenebilir.

Alanyazındaki çalışmaların bir kısmında ise öğrencilerin olumsuz değerlendirilme korkuları çeşitlik demografik değişkenler açısından incelenmiştir. Bu çalışmalardan biri, Ünal, Arık ve Uzun'un (2016) sınıf ve sosyal bilgiler öğretmenliği bölümlerinde öğrenim gören üniversite öğrencileriyle yürüttükleri araştırmadır. Ünal, Arık ve Uzun (2016), 588 öğrencinin katılımıyla gerçekleştirdikleri çalışmada, sınıf ve sosyal bilgiler öğretmenliği programlarında okuyan öğrencilerin olumsuz değerlendirilme korkuları arasında anlamlı fark olduğunu belirlemiş; cinsiyet, anne ile baba eğitim durumu ve sınıf düzeyi değişkenlerinin ise olumsuz değerlendirilme korkusu üzerinde anlamlı bir etkiye sahip olmadığı sonucuna varmışlardır. Öğrencilerin olumsuz değerlendirilme korkularının demografik değişkenler açısından ele alındığı ikinci bir araştırma Alkan (2015) tarafından yapılmıştır. Alkan'ın (2015) yaptığı çalışmada ortaokul öğrencilerinin akademik ortamlarda olumsuz değerlendirilme korkuları; anne-baba eğitim durumu, anne-baba mesleği, kardeş sayısı, kişilik tipi, tercih edilen çalışma türü, cinsiyet, sınıf düzeyi, dershaneye gitme ve özel ders alma durumlarına göre test edilmiştir. Araştırma sonucunda; öğrencilerin olumsuz değerlendirilme korkularının anne-baba eğitim durumu, anne baba mesleği, kardeş sayısı, kişilik tipi ve tercih edilen çalışma türüne göre anlamlı farklılık gösterdiği tespit edilmiştir. Cinsiyet ve sınıf düzeyi değişkenleri ile dershaneye gitme ve özel ders alma durumunun ise öğrencilerin olumsuz değerlendirilme korkularında anlamlı bir farklılık yaratmadığı saptanmıştır.

Alanyazın gözden geçirildiğinde öğrencilerin olumsuz değerlendirilme korkularının belirlenmesi amacıyla yürütülen çalışmalarda, genel olarak Watson ve Friend (1969) tarafından geliştirilen Olumsuz Değerlendirilme Korkusu Ölçeği Genel Formu'nun (ODKÖ-GF) ya da bu ölçeğin Leary (1983) tarafından revize edilmesiyle elde edilen kısa formun (ODKÖ-KF) kullanıldığı görülmektedir. Ancak; olumsuz değerlendirilme korkusu öğrenme sürecine ilişkin duyuşsal bir özellik olarak ele alındığında; bu ölçekler yerine, doğrudan öğrencilerin sınıf ortamında olumsuz değerlendirilme korkularını ölçmek amacıyla geliştirilmiş bir ölçeğin

kullanılmasının daha doğru bir tercih olacağı düşünülmektedir. Öğrenciler sınıf ortamında sürekli olarak değerlendirme süreçlerinden geçtiğinden (Özgüngör, 2006) ve öğrenciler hakkında alınan kararlar bu değerlendirmelerin sonuçlarına göre şekillendiğinden, günlük hayatta olumsuz değerlendirilme korkusu yaşamayan bir öğrenci sınıf ortamında olumsuz değerlendirilme korkusu yaşayabilmektedir. Dolayısıyla, öğrencilerin sınıf ortamında olumsuz değerlendirilme korkularının ölçülmesinde kullanılabilir bir ölçek geliştirmenin alanyazına katkı sunacağı tahmin edilmektedir.

1.4. Araştırmanın Amacı ve Önemi

Bu araştırmanın, Likert tipi ölçek geliştirmede Rasch modelinin kullanımına ilişkin bir örnek olması amaçlanmıştır. Bu amaç doğrultusunda, öncelikle iki yüzeyle (birey ve madde) Rasch analizi uygulanarak ODKÖ-ÖF'nin psikometrik özellikleri incelenmiştir. Ardından öğrencilerin ODKÖ-ÖF'den aldıkları puanların akademik beklentilere ilişkin stres değişkenine göre farklılaşmış farklılaşmadığı çok yüzeyle (birey, madde ve ABSE puanları) Rasch modeli ile test edilmiştir.

Araştırmanın alanyazına önemli katkılarının olacağı tahmin edilmektedir. Uluslararası alanyazın incelendiğinde, Rasch analizinin Likert tipi ölçeklerin psikometrik özelliklerinin belirlenmesinde yaygın biçimde kullanıldığı görülmektedir. Özellikle, çok yüzeyle Rasch modelinin geliştirilmesinden sonra Rasch analizi klasik test kuramına göre daha fazla tercih edilmiş (Meads & Bentall, 2008) ve Likert tipi ölçeklerde Rasch modelinin kullanımında büyük artış olmuştur. Örneğin Anshel (2009) tarafından yapılan çalışmada tek boyutlu Mükemmeliyetçilik Ölçeği'nin; Brinthaup ve Kang'ın (2014) yaptığı çalışmada Öz Konuşma Ölçeği'nin; Behizadeha ve Engelhard (2014) tarafından yapılan çalışmada Yazmada Algılanan Otantiklik Ölçeği'nin; Ricketts, Engelhard ve Chang'ın (2015) yaptığı çalışmada Matematikte Akademik Dayanıklılık Ölçeği'nin; Rameli ve Kosnin'in (2017) yürüttüğü çalışmada Ortaokul Öğrencileri için Matematik Kaygısı Ölçeği'nin; Milliken, Ludlow, DeSanto-Madeya ve Grace'in (2018) yaptığı çalışmada Ahlaki Farkındalık Ölçeği'nin (2018) psikometrik özellikleri Rasch analizi ile incelenmiştir. Likert tipi ölçeklerde Rasch modelinin kullanımı, ölçeklerin psikometrik özelliklerinin belirlenmesine yönelik çalışmalar ile sınırlı kalmamıştır. Likert tipi ölçeklerden alınan puanlar üzerinden gerçekleştirilen farka dayalı istatistiklerde de Rasch modelinden yararlanılmıştır. Söz gelimi, Behizadeha ve Engelhard (2014) tarafından yapılan çalışmada, öğrencilerin yazmada algılanan otantiklik düzeylerinin cinsiyet değişkenine göre farklılık gösterip göstermediği çok yüzeyle Rasch modeli ile test edilmiştir. Benzer şekilde, Ricketts, Engelhard ve Chang'ın (2015) yaptığı çalışmada öğrencilerin matematikte akademik dayanıklılık düzeylerinin cinsiyetlerine, akademik başarılarına ve sosyo-ekonomik düzeylerine göre farklılaşmış farklılaşmadığını ortaya koymak için çok yüzeyle Rasch modelinden faydalanılmıştır. Türkçe alanyazında ise Likert tipi ölçeklerden alınan puanlar üzerinden uygulanan farka dayalı istatistiklerde Rasch analizinin kullanıldığı bir araştırmaya rastlanmadığı gibi madde tepki kuramına dayalı Likert tipi ölçek geliştirme çalışmalarının da (Demirtaşlı, Yalçın & Ayan, 2016) son derece sınırlı olduğu görülmüştür. Bu durum Türkçe alanyazın açısından bir eksiklik olarak düşünülmekte ve dolayısıyla bu araştırmanın alanyazına katkı sağlayacağı öngörülmektedir.

Araştırmayı önemli kılan bir başka özellik, ODKÖ-ÖF'nin geliştirilecek olmasıyla ilgilidir. Araştırmada, Rasch analizleri ODKÖ-ÖF üzerinden gerçekleştirilecek ve çalışma sonucunda lise öğrencilerin sınıf ortamında öğretmenleri ya da arkadaşları tarafından olumsuz değerlendirileceklerine dair yaşadıkları korkuyu ölçmek amacıyla kullanılabilir bir ölçek geliştirilmiş olacaktır. Alanyazına bakıldığında; tükenmişlik, kendini sabotaj ve risk alma gibi psikolojik yapıların öğrencilerin öğrenme sürecindeki davranışlarını etkileyen değişkenler olduğunun belirlenmesinin ardından, bu değişkenleri ölçmeye yönelik ölçme araçlarının öğrenci formlarının geliştirildiği görülmektedir (Aypay, 2011; Beghetto, 2009; Cocoradă, 2011; İlhan & Çetin, 2013; Schaufeli, Martinez, Marques-Pinto, Salanova & Bakker, 2002). Örneğin; tükenmişliğin hemşireler, doktorlar, öğretmenler, emniyet görevlileri, psikologlar vb. farklı mesleklerden bireylerde görülebildiği gibi öğrencilerde de görülebilen bir özellik olduğunun belirlenmesiyle Schaufeli vd. (2002) tarafından Maslach Tükenmişlik Envanteri'nin Öğrenci Formu geliştirilmiştir. Yine, öğrencilerin öğrenme sürecindeki risk alma davranışının genel risk alma davranışından ayrı olarak ele alınması düşüncesinin bir ürünü olarak Clifford (1991) tarafından Akademik Risk Alma Ölçeği geliştirilmiştir. Midgley ve Urda (2001) benzer bir düşünceyle, kendini sabotaj ölçeğinin öğrenci formunu geliştirmiş ve geliştirdikleri ölçme aracı Akademik Benliği Sabotaj Ölçeği olarak adlandırılmışlardır. Sıralanan araştırmalar, benzer bir

yaklaşımın olumsuz değerlendirilme korkusu ölçeğinde de benimsenmesi gerektiğine işaret etmektedir. Nitekim Alkan (2015), akademik bağlamdaki olumsuz değerlendirilme korkusunun genel anlamdaki olumsuz değerlendirilme korkusundan ayrı olarak ele alınması gerektiğini savunmuş ve akademik ortamlarda olumsuz değerlendirilme korkusu ölçeğini geliştirmiştir. Alkan (2015), ortaokul öğrencilerine yönelik geliştirdiği bu ölçeğin psikometrik özelliklerini klasik test kuramı çerçevesinde incelemiştir. Bu çalışmada ise Alkan'ın (2015) çalışmasından farklı olarak lise öğrencilerine yönelik bir olumsuz değerlendirilme korkusu ölçeği geliştirilecek ve ölçeğin psikometrik özellikleri madde tepki kuramına dayalı modeller ailesinin bir üyesi olan Rasch analizi ile test edilecektir. Bu bakımdan, çalışmanın özgün bir nitelik taşıdığına inanılmaktadır.

2. YÖNTEM

2.1. Çalışma Grubu

Araştırma, 167'si (%45.50) kız ve 200'ü (%54.50) erkek olmak üzere toplam 367 lise öğrencisi üzerinde yürütülmüştür. Öğrencilerin, 84'ü (%22.90) 9. sınıfa, 133'ü (36.20) 10. sınıfa, 83'ü (%22.60) 11. sınıfa ve 67'si (%18.30) 12. sınıfa devam etmektedir. 100 ile 200 katılımcıdan oluşan bir çalışma grubu Rasch analizi için yeterli görülmektedir (Castello, 2008; DeMars, 2010). Dolayısıyla çalışma grubunun kabul edilebilir büyüklükte olduğu söylenebilir.

2.2. Veri Toplama Araçları

Araştırmanın verileri, ODKÖ-ÖF ile ABSE'den yararlanılarak toplanmıştır.

Olumsuz Değerlendirilme Korkusu Ölçeği-Öğrenci Formu (ODKÖ-ÖF): ODKÖ-ÖF'de yer alacak maddelerin hazırlanmasına yönelik basamaklar işlem başlığı altında açıklanmış, ölçeğin psikometrik özelliklerinin incelenmesinde kullanılan istatistiksel tekniklere veri analizi başlığı altında yer verilmiş; geçerlik ve güvenilirlik çalışması doğrultusunda uygulanan Rasch analizine ilişkin sonuçlar ise bulgular bölümünde sunulmuştur.

Akademik Beklentilere İlişkin Stres Envanteri (ABSE): ABSE, Ang ve Huan (2006) tarafından geliştirilmiş, Kelecioğlu ve Bilge (2009) tarafından Türkçeye uyarlanmıştır. Ölçek beşli Likert tipi bir derecelendirmeye sahip dokuz maddeden oluşmaktadır. Ölçekte, Aile/Öğretmen Beklentileri (AÖB) ve Kendine İlişkin Beklentiler (KİB) şeklinde iki faktörlü bir yapı söz konusudur. Alt boyutlardan ve ölçeğin genelinden alınan yüksek puanlar akademik beklentilere ilişkin yüksek stresi ifade etmektedir. ABSE'nin orijinal ve Türkçe formunun geçerlik ile güvenilirlik analizine ilişkin elde edilen bulgular Tablo 1'de özetlenmiştir.

Tablo 1. ABSE'nin orijinal ve Türkçe formundan elde edilen geçerlik ile güvenilirlik bulguları

Boyut	Madde Sayısı	Açıkladığı Varyans		Cronbach Alfa		Test Tekrar-Test	
		Orijinal Form	Türkçe Form	Orijinal Form	Türkçe Form	Orijinal Form	Türkçe Form
AÖB	5	%53.74	%40.84	.85	.81	.79	.64
KİB	4	%12.18	%13.60	.84	.65	.77	.49

Bu çalışmada; ABSE'nin alt boyutları üzerinden değil ölçekten alınan toplam puan üzerinden işlem yapılmıştır. Çalışmada, ABSE'nin geneli için hesaplanan iç tutarlılık katsayısı .81 olarak bulunmuştur. İç tutarlılık katsayısı .70 ve üzerinde olan ölçümlerin güvenilir olduğu kabul edilmektedir (Büyüköztürk, 2010). Buna göre, ABSE kullanılarak elde edilen ölçümlerin yeterli derecede güvenilir olduğu ifade edilebilir.

2.3. İşlem

Ölçek geliştirme sürecinde ilk olarak, ölçülmek istenen özelliğin kuramsal tanımından yola çıkılarak ölçeğin ne tür gözlenebilir ve ölçülebilir uyarıcıları oluşturacağını tanımlanması gerekir (Tezbaşaran, 1997). Ölçülmesi hedeflenen yapının işevuruk tanımının yapıldığı bu süreçte genellikle nitel çalışmalardan yararlanılmakta ve nitel çalışmada ulaşılan bulgular ölçek maddelerinin yazımı konusunda bilimsel bir kaynak işlevi görmektedir. Alanyazında yer alan olumsuz değerlendirilme korkusu ölçekleri, bu tür bir korkunun sınıf ortamında ne gibi gözlenebilir etkileri oluşturacağı konusunda yeterli bilgiyi verdiğinden, bu çalışmada sözü edilen türde nitel bir çalışmaya ihtiyaç duyulmamıştır. ODKÖ-ÖF geliştirilirken, nitel bir çalışmayla başlangıç yapmak yerine; Watson ve Friend (1969) tarafından geliştirilen ODKÖ-GF ile bu ölçeğin Leary

(1983) tarafından yenilenmesiyle elde edilen ODKÖ-KF'den faydalanılmıştır. ODKÖ-GF ile ODKÖ-KF'de yer alan maddeler akademik bağlama ve sınıf ortamına daha iyi uyacak şekilde formüle edilerek ve ilgili alanyazından yararlanılarak 20 maddeden oluşan bir madde havuzu elde edilmiştir. Örneğin, ODKÖ-GF'deki yer alan “*Komik görünmekten ya da aptal durumuna düşmekten çoğunlukla korkarım*” maddesinden hareketle, ODKÖ-ÖF'nin madde havuzuna “*Sınıfta tartışılan bir konu hakkında yanlış şeyler söyleyip alay konusu olmamak için sessiz kalmayı tercih ederim*” ve “*Derste anlatılan konu ile ilgili anlayamadığım noktaları sormaktan çekinirim, çünkü bunun beni komik durumuna düşürmesinden korkarım*” maddeleri eklenmiştir. Benzer şekilde, ODKÖ-KF'de bulunan “*Diğer insanların hakkımdaki düşünceleri beni rahatsız etmez*” maddesinden yola çıkılarak ODKÖ-ÖF'ye “*Derslerdeki performansım hakkında, başkalarının ne düşündüğü beni rahatsız etmez*” şeklinde bir madde dâhil edilmiştir. ODKÖ-ÖF'deki bazı maddeler ise olumsuz değerlendirilme korkusuna ilişkin kuramsal bilgiler ışığında yazılmıştır. Söz gelimi, diğer insanların kendisinden yüksek başarı beklentisi içerisinde olduğunu düşünmesi olumsuz değerlendirilme korkusu yaşayan bireylerin karakteristik özelliklerinden biri olarak ifade edildiğinden (Rapee & Heimberg, 1997) ODKÖ-ÖF'de “*Öğretmenlerimin benden çok yüksek bir başarı beklentisi içerisinde olduklarını düşünürüm*” şeklinde bir maddeye yer verilmiştir.

Ölçekte yer alan ifadeler için *Kesinlikle Katılıyorum* (5), *Katılıyorum* (4), *Kararsızım* (3), *Katılmıyorum* (2) ve *Kesinlikle Katılmıyorum* (1) şeklinde beşli Likert tipi bir derecelendirme kullanılmıştır. Daha sonra ODKÖ-ÖF'nin kapsam ve görünüş geçerliğini sağlamak için ölçme ve değerlendirme, eğitim programları ve öğretim ile psikolojik danışmanlık ve rehberlik alanlarından birer uzman olmak üzere üç uzmandan görüş alınmıştır. Uzmanlardan alınan görüşler sonucunda bazı maddelerde değişikliğe gidilmiştir. Örneğin, “*Sınıfta başarısız bir öğrenci olarak görünme konusunda kaygılarım var*” şeklinde ifade edilen madde “*Sınıfta başarısız bir öğrenci olarak görünmekten korkarım*” biçiminde değiştirilmiştir. Yine uzman görüşleri doğrultusunda “*Öğretmenlerime karşı mahcup olma korkusundan dolayı sınıfta tutukluk yaşarım*” maddesi “*Öğretmenlerime karşı mahcup olma korkusundan dolayı sınıfta kendimi rahat ifade edemem*” şeklinde yeniden düzenlenmiştir. Ayrıca üç uzmandan ikisi “*Öğretmenlerimin nasıl bir öğrenci olduğum konusundaki düşünceleri beni endişelendirir*” şeklinde ifade edilen maddenin ölçtüğü özelliği belirlemeye yönelik başka maddelerin ölçme aracında yer aldığını dile getirmiş ve bu maddenin ölçekten çıkarılmasını önermiştir. Dolayısıyla söz konusu madde ölçekten çıkarılmıştır. Uzman görüşleri sonucunda bir madde ölçekten çıkarılırken, “*Beni ders çalışmaya motive eden en önemli faktör başkalarının (aile/öğretmen/arkadaşlar) takdirini kazanma düşüncesidir*” şeklindeki başka bir madde ölçme aracına eklenmiştir.

Uzman görüşlerinden hareketle, ölçek maddelerinde gerekli değişiklikler yapılmasının ardından ölçme aracının dil açısından anlaşılabilirliğini sağlamak için iki Türk Dili uzmanının görüşüne başvurulmuştur. Yazım kuralları ve noktalama işaretlerinin kullanımı ile ilgili uzmanlardan gelen görüşler doğrultusunda ölçek maddeleri gözden geçirilmiştir. Daha sonra ODKÖ-ÖF'de bulunan maddelerin anlaşılabilirliğine ve ölçme aracının uygulama süresine ilişkin dönüt almak amacıyla 10 lise öğrencisi üzerinde ön uygulama gerçekleştirilmiştir. Uygulamanın ardından ODKÖ-ÖF'yi yanıtlayan öğrencilerle görüşme yapılmıştır. Görüşmelerde, öğrencilerin ölçek maddelerinin anlaşılabilirliği hakkındaki fikirleri alınmıştır. Bunun yanı sıra öğrencilere; ölçeğin amacını, ölçekteki madde sayısını ve ölçeğin nasıl doldurulması gerektiğini açıklamak üzere ölçeğin başında sunulan yönerge hakkındaki düşünceleri sorulmuştur. Öğrencilerden gelen görüşler, ölçek maddelerinde ve yönergede herhangi bir değişikliğe ihtiyaç duyulmadığını ortaya koymuştur. Bu süreçlerin tamamlanmasıyla ölçek geniş bir katılımcı grubu üzerinde uygulamaya hazır hale gelmiştir.

2.4. Verilerin Toplanması ve Analizi

Araştırmanın verileri, 2015 yılı Kasım ayı içerisinde toplanmıştır. ODKÖ-ÖF sınıf ortamında öğrencilere uygulanmış ve öğrenciler 15 ile 20 dakika arasında değişen sürelerde ölçek maddelerini yanıtlamışlardır. Araştırma kapsamında toplanan veriler FACETS paket programından yararlanılarak Rasch modeline göre analiz edilmiştir. Analiz sonuçları yorumlanmadan önce Rasch modeline ilişkin varsayımların karşılanıp karşılanmadığı test edilmiştir.

Tek boyutluluk, yerel bağımsızlık ve model-veri uyumu Rasch analizinin gerçekleştirilebilmesi için karşılanması gereken varsayımlardır (DeMars, 2010). Her ne kadar Rasch analizine ilişkin üç varsayım bulunsa da bu varsayımlar ayrı ayrı test edilmemektedir. Çünkü model-veri uyumunun sağlanması tek boyutluluk varsayımının karşılandığı anlamına gelirken (Lee,

Peterson, & Dixon, 2010), tek boyutluluk varsayımının karşılanması da yerel bağımsızlık ile ilgili bir sorun olmadığını göstermektedir (Hambleton, Swaminathan & Rogers, 1991). Dolayısıyla Rasch analizinin uygulanabilmesi için kontrol edilmesi gereken temel varsayım, model-veri uyumunun sağlanıp sağlanmadığıdır (Güler vd., 2017). Rasch analizi çıktılarını bu doğrultuda incelendiğinde uyum-İçi ve uyum-Dışı istatistiklerinin bu istatistiklere ilişkin kabul edilebilir aralık içerisinde kaldığı görülmüş (bulgular bölümüne bakınız) ve model-veri uyumunun sağlandığı anlaşılmıştır.

Uyum istatistiklerinin yanı sıra araştırma verilerinin Rasch modeline uygunluğunu ortaya koymak için standartlaştırılmış artık (StRes) değerlerine de başvurulmuştur. Linacre'ye (2014) göre, model ile verilerin uyumlu olduğunun söylenebilmesi için ± 2 aralığı dışında kalan standartlaştırılmış artıklar çalışmadaki toplam veri sayısının %5'inden, ± 3 aralığı dışında yer alan standartlaştırılmış artıklar ise çalışmadaki toplam veri sayısının %1'inden fazla olmamalıdır. Araştırma 367 öğrencinin katılımıyla yürütüldüğünden, çalışmadaki toplam veri sayısı ODKÖ-ÖF'nin 20 maddelik ilk hali için 7340 iken (367×20); 16 maddelik nihai formu için 5872'dir (367×16). Buna göre, ± 2 ile ± 3 aralığı dışındaki standartlaştırılmış artık sayıları ODKÖ-ÖF'nin 20 maddelik ilk formunda sırasıyla 367 ve 73'ü; 16 maddelik nihai formunda ise sırasıyla 293 ve 59'u aşmamalıdır. Çalışmada ODKÖ-ÖF'nin 20 maddelik ilk halinde ± 2 aralığı dışındaki standartlaştırılmış artık sayısı 275 değeri ile 367'den, ± 3 aralığı dışındaki standartlaştırılmış artık sayısı da 53 değeri ile 73'ten küçük bulunmuştur. Benzer şekilde, ODKÖ-ÖF'nin 16 maddelik son formunda ± 2 aralığı dışındaki standartlaştırılmış artık sayısının 241 değeri ile 293'ten ve ± 3 aralığı dışındaki standartlaştırılmış artık sayısının 43 değeri ile 59'dan küçük olduğu saptanmıştır. Bu değerler, model-veri uyumunun sağlandığının yansıtmaktadır. Model-veri uyumu sağlandığından Rasch analizinin diğer iki varsayımı olan tek boyutluluk ve yerel bağımsızlığın da karşılandığı ifade edilebilir.

Varsayımların karşılandığı belirlendikten sonra, ilk etapta ODKÖ-ÖF'nin psikometrik özelliklerini belirlemeye yönelik olarak öğrenciler ve maddeler şeklinde iki yüzeyli bir desen ile gerçekleştirilen Rasch analizine ilişkin çıktılar yorumlanmıştır. Ardından öğrencilerin olumsuz değerlendirilme korkularının akademik beklentilere ilişkin yaşadıkları strese göre farklılık gösterip göstermediğini test etmek üzere çok yüzeyli Rasch analizi uygulanmıştır. Çok yüzeyli Rasch analizinde üç değişkenlik kaynağı söz konusu olmuştur. Bu değişkenlik kaynaklarından ilki öğrenciler, ikincisi ODKÖ-ÖF'deki maddeler ve üçüncüsü ise öğrencileri ABSE puanlarına göre sınıflandırmak üzere SPSS programında yapılan kümeleme analizinde ortaya çıkan gruplardır.

3. BULGULAR

Bu bölümde Rasch analizinden elde edilen bulgulara yer verilmiştir. Öncelikle, iki yüzeyli Rasch analizi gerçekleştirilerek ODKÖ-ÖF'nin psikometrik özellikleri incelenmiştir. Analiz sonucunda ölçekten çıkarılması gereken madde olup olmadığına karar verebilmek için madde yüzeyine ilişkin uyum istatistikleri ile korelasyon değerlerine bakılmış, ulaşılan bulgular Tablo 2'de sunulmuştur.

Tablo 2. ODKÖ-ÖF'ye ilişkin iki yüzeyli modelde madde uyum istatistikleri ve korelasyon değerleri

Madde r	İlk Analiz Sonuçları			İkinci Analiz Sonuçları			Üçüncü Analiz Sonuçları		
	Uyum İçi	Uyum Dışı	Corr. PtBis*	Uyum İçi	Uyum Dışı	Corr. PtBis	Uyum İçi	Uyum Dışı	Corr. PtBis
M1	1.11	1.06	.47	1.20	1.14	.48	1.28	1.18	.48
M2	.85	.82	.46	.94	.91	.46	.98	.98	.46
M3	.91	.90	.43	1.02	1.08	.44	1.07	1.14	.42
M4	.88	.87	.54	.96	.99	.54	.99	1.05	.54
M5	.94	.91	.58	.97	.96	.60	.99	.96	.61
M6	.70	.69	.55	.77	.78	.54	.80	.80	.54
M7	.88	.86	.40	.99	1.01	.41	1.02	1.04	.42
M8	.80	.81	.53	.84	.86	.56	.86	.90	.56
M9	.84	.84	.45	.95	.98	.45	1.01	1.07	.43
M10	.96	.92	.53	.98	.94	.56	1.00	.96	.57
M11	.85	.85	.52	.91	.93	.54	.95	.98	.53

M12	.90	.88	.54	.93	.90	.57	.95	.90	.58
M13	1.85	2.34	-.29	-	-	-	-	-	-
M14	.93	1.02	.15	-	-	-	-	-	-
M15	.76	.82	.46	.84	.96	.47	.88	1.02	.47
M16	.92	.90	.51	.94	.90	.56	.95	.92	.56
M17	1.11	1.12	.31	1.22	1.25	.34	1.24	1.26	.35
M18	.97	1.00	.40	1.07	1.32	.42	1.10	1.32	.43
M19	1.51	1.84	-.05	-	-	-	-	-	-
M20	1.39	1.33	.32	1.61	1.61	.30	-	-	-

* Point Biserial Correlation

Tablo 2'ye göre; ODKÖ-ÖF ile toplanan veriler üzerinden gerçekleştirilen ilk analizde maddelere ilişkin uyum içi istatistikleri .70 ile 1.85 arasında sıralanmakta, uyum dışı istatistikleri ise .81 ile 2.34 arasında değişmektedir. Uyum istatistikleri 0 ile ∞ arasında uzanan değerler alabilmektedir. Bununla birlikte, uyum istatistikleri için istenilen değer 1.00'dir (Eckes, 2009). Bu değer, model ile veri arasında mükemmel bir uyum olduğunu yansıtır. Fakat gerçek ölçme durumlarında model ile veri arasında mükemmel bir uyum bulunması olasılığı düşüktür (Brentari & Golia, 2008). Bu sebeple uyum istatistiklerine ilişkin kabul edilebilir aralığın ne olduğu sorusunun yanıtlanması gerekir. Wright ve Linacre (1994) uyum istatistikleri için ölçüt olarak alınması gereken aralıkları; yüksek risk içeren sınavlarda .80 ile 1.20, Likert tipi ölçeklerde .60 ile 1.40 ve klinik gözlemlerde .50 ile 1.70 olarak belirtmiştir. Bununla birlikte, yaptıkları simülasyon çalışmalarından ve çok sayıda veri setinin analizi sonucunda edindikleri tecrübelerden yola çıkarak .50 ile 1.50 arasındaki uyum istatistiklerinin kabul edilebilir uyuma; bu aralığın dışında kalan değerlerin ise zayıf uyuma işaret ettiği şeklinde genel bir değerlendirme yapmışlardır (Wright & Linacre, 1994). Bu genel değerlendirme dikkate alınarak araştırmada uyum istatistikleri için .50 ile 1.50 arasındaki değerler kabul edilebilir uyum ölçütü olarak yorumlanmıştır. Tablo 2'deki değerlere bakıldığında, ilk analizde madde 13 (*Derslerdeki performansım hakkında, başkalarının ne düşündüğü beni rahatsız etmez*) ile madde 19'a (*Tüm öğrencilerin okul çalışmaları sırasında hatalar yapabileceğini, bunun için üzülmemek gerektiğini düşünürüm*) ait uyum istatistiklerinin bu aralığın dışında kaldığı tespit edilmiştir. Bu nedenle, belirtilen iki maddenin ölçekten çıkarılması gerektiğine karar verilmiştir. Yine ilk analiz sonuçlarında, madde 14'e (*Öğretmenlerimin benden çok yüksek bir başarı beklentisi içerisinde olduklarını düşünürüm*) ait nokta çift serili korelasyon değerinin oldukça düşük olduğu saptanmıştır. Rasch analizi sonuçlarında rapor edilen nokta çift serili korelasyon katsayısı; herhangi bir yüzeyin bütün bileşenlerinin aynı yönde çalışıp çalışmadığı hakkında bilgi vermektedir (Linacre, 2014). Dolayısıyla Rasch analizi çıktılarında yer alan bu korelasyon katsayısı için klasik test kuramındaki madde ayırt ediciliğinde olduğu gibi .30 değerinin ölçüt olarak alınabileceği belirtilmektedir (Güler vd., 2017). Madde 14'e ilişkin nokta çift serili korelasyon katsayısının .30 ölçütünü karşılamaması, bu maddenin ODKÖ-ÖF'deki diğer maddeler ile aynı yönde çalışmadığı şeklinde yorumlanmış ve maddenin ölçekten çıkarılması uygun görülmüştür.

Madde 13, madde19 ile madde 14 ölçekten çıkarıldıktan sonra ODKÖ-ÖF'de kalan 17 madde üzerinden ikinci bir Rasch analizi gerçekleştirilmiş ve bu analizde madde 20 için hesaplanan uyum istatistiklerinin .50 ile 1.50 kabul edilebilir aralığı (Wright & Linacre, 1994) dışında kaldığı belirlenmiştir. Bu sebeple, madde 20 ölçekten çıkarılıp Rasch analizi bir kez daha tekrarlanmıştır. Yinelene Rasch analizine ilişkin çıktılar, Tablo 2'nin üçüncü analiz sonuçları başlığı altında sunulmuştur. Tablo 2'de görüldüğü üzere, üçüncü analizde ODKÖ-ÖF'de kalan 16 maddenin tümü için uyum istatistikleri kabul edilebilir sınırlar içerisinde kalmakta ve nokta çift serili korelasyon değerleri .30 ölçütünün üzerinde yer almaktadır. ODKÖ-ÖF'ye ilişkin 16 maddelik formda model ile veri uyumunun sağlandığı belirlendikten sonra, Rasch analizi çıktıları detaylı olarak incelenmiş ve analiz sonucunda elde edilen logit cetvel Şekil 1'de sunulmuştur.

Measr	-Maddeler	+Birey	Ölçek
3			(5)
		.	
2			
		.	
		.	---
		*	
1			
		*	
	15	**	
		***	4
		**	
	13 7	****	
	10 12	*****	---
	14 9	*****	
* 0 *	11 16 3 8	* ***** *	* 3 *

	5 6	*****	
	4	*****	---
	1	***	
	2	****	
		***	2

-1			

		**	
		*	---
		.	
		*	
		.	
		*	
-2			
		.	
		.	
		*	
		.	
		*	
-3			
		.	(1)
Measr	-Maddeler	* = 4	Ölçek

Şekil 1. ODKÖ-ÖF'ye ilişkin logit cetvel

Şekil 1'in ilk sütunu logit ölçeğini temsil eder. Rasch analizine dâhil olan bütün yüzeyler bu ölçek üzerinde yer alır. Böylece yüzeyler birbirleri ile ilişkilendirilebilir ve yüzeyler arasında karşılaştırma yapılabilir. Logit cetvelin ikinci sütununda maddeler ve üçüncü sütununda bireyler yer alır. Birey yüzeyinin alt ucundan üst ucuna doğru gidildikçe öğrencilerin olumsuz değerlendirilme korkularında artış gözlenmektedir. Maddelerin logit cetvel üzerindeki sıralamaları, öğrencilerin katılmakta en çok zorlandıkları maddenin 15 numaralı madde (*Arkadaşlarım tarafından eleştirilme ihtimalinden dolayı grup çalışmalarından kaçınırım*) olduğunu yansıtmaktadır. Likert tipi ölçek maddeleri için kullanılan zor kelimesi; katılma durumunu bildiren seçeneklerin (kesinlikle katılıyorum ya da katılıyorum) işaretlenme düzeyinin düşük, katılmama durumunu bildiren seçeneklerin (katılmıyorum ya da kesinlikle katılmıyorum) işaretlenme düzeyinin ise yüksek olduğu anlamına gelir. Yine maddelerin logit cetvel üzerindeki sıralamasına göre, öğrenciler için katılması en kolay madde, 2 numaralı maddedir (*Öğretmenlerimin üzerinde*

nasıl bir izlenim bıraktığım konusunda endişelenirim). Diğer bir deyişle ODKÖ-ÖF’de, katılma durumunu bildiren seçeneklerin (kesinlikle katılıyorum ya da katılıyorum) işaretlenme düzeyinin en yüksek; katılmama durumu bildiren seçeneklerin (katılmıyorum ya da kesinlikle katılmıyorum) işaretlenme düzeyinin ise en düşük olduğu madde 2 numaralı maddedir. Logit cetvelin son sütunu ölçeğe ait derecelmeyi göstermektedir. Bu sütun incelendiğinde, ölçek kategorilerinin geniş bir aralığa sahip olduğu görülmektedir. Ölçek kategorilerinin gösterdiği bu dağılım, ölçekteki beşli derecelemenin etkin bir biçimde çalıştığına işaret etmektedir.

Logit cetvelden sonra, öğrenci ve madde yüzeylerine ilişkin ölçüm raporları incelenmiştir. Öğrenci yüzeyine ait ölçüm raporları Tablo 3’te sunulmuştur. Tablo 3’e göre, olumsuz değerlendirilme korkusuna ilişkin öğrenci puanları -1.17 logitlik bir ortalamaya sahiptir. Uyum içi ve uyum dışı istatistiklerine ilişkin ortalamalar birbirine eşit olup 1.03 logittir. Uyum istatistikleri, bu istatistiklerin istenilen değeri olan 1.00 ’e (Eckes, 2009) oldukça yakın olduğundan öğrenci yüzeyine ilişkin veriler ile model arasındaki uyumun yüksek olduğu söylenebilir.

Tablo 3. ODKÖ-ÖF’ye ilişkin iki yüzeyli modelde öğrenci yüzeyine ait ölçüm raporları

	Ölçüm	Standart Hata	Uyum içi	Uyum dışı
Ortalama	-.17	.24	1.03	1.03
Standart Sapma	.74	.13	.51	.53
Güvenirlilik = .86	Ayrırma Oranı = 2.47	Ki-Kare = 1884.80	sd = 366	p = .00

Tablo 3’e bakıldığında, öğrenci yüzeyine ilişkin güvenirlilik indeksinin $.86$ olduğu görülmektedir. Rasch analizinden elde edilen güvenirlilik indeksi 0 ile 1 arasında değişen değerler almaktadır. Veriler model ile uyum gösterdiği takdirde; bu güvenirlilik indeksi Cronbach alfa iç tutarlılık katsayısına benzer şekilde yorumlanmaktadır (Bond & Fox, 2012). Bu nedenle, Cronbach alfa iç tutarlılık katsayısında olduğu gibi Rasch analizinden elde edilen güvenirlilik indeksi için de $.70$ değerinin ölçüt olarak alınması önerilmektedir (Walker, Engelhard & Thompson, 2012). Araştırmada, öğrenci yüzeyi için hesaplanan güvenirlilik indeksi bu ölçütü aştığından, öğrencilerin ölçülen örtük özellik açısından başarılı bir şekilde ayırt edildiği söylenebilir. Birey yüzeyi için rapor edilen ayırma oranının 2 ’nin üzerinde (Linacre, 2012) ve ki-kare değerinin anlamlı olması da (ki-kare= 1884.80 , $sd=366$, $p<.001$), olumsuz değerlendirilme korkusu farklı olan öğrencilerin etkili bir biçimde ayırt edildiğini ortaya koymaktadır. Öğrenci yüzeyinin ardından madde yüzeyine ilişkin ölçümler incelenmiş ve elde edilen bulgular Tablo 4’te sunulmuştur.

Tablo 4. ODKÖ-ÖF’ye ilişkin iki yüzeyli modele madde yüzeyine ait ölçüm raporları

Madde No	Madde	Ölçümler	Standart Hata	Uyum içi	Uyum dışı
M1	Sınıfta başarısız bir öğrenci olarak görünmekten korkarım.	-.51	.05	1.25	1.18
M2	Öğretmenlerimin üzerinde nasıl bir izlenim bıraktığım konusunda endişelenirim.	-.67	.05	.98	.98
M3	Okul çalışmalarındaki performansım hakkında başkalarının ne düşündüğü ile ilgili endişeler taşıyorum.	-.04	.05	1.07	1.14
M4	Öğretmenin sorduğu bir soruya yanlış cevap vereceğim düşüncesiyle kaygılanırım.	-.32	.05	.99	1.05
M5	Sınıfta bir soruya cevap verirken komik görünmekten veya aptal durumuna düşmekten korkarım.	-.22	.05	.99	.96
M6	Öğretmenlerin benim hakkımda ne düşündüğü ile ilgili çok fazla endişelendiğimi düşünüyorum.	-.25	.05	.80	.80
M7	Sınıf arkadaşlarımla yaptığım çalışmalarda, arkadaşlarımla eksiklerimi/hatalarımı bulmalarından korkarım.	.36	.05	1.02	1.04
M8	Öğretmenin, hazırladığım bir ödevi ya da bir soruya verdiğim cevabı onaylamamasından korkarım.	-.01	.05	.86	.90
M9	Sınıf arkadaşlarımla nasıl bir öğrenci olduğum konusundaki düşünceleri beni endişelendirir.	.15	.05	1.01	1.07

M10	Derste anlatılan konu ile ilgili anlayamadığım noktaları sormaktan çekinirim, çünkü bunun beni komik durumuna düşürmesinden korkarım.	.21	.05	1.00	.96
M11	Sınıf arkadaşlarımla başarısız bir öğrenci olduğumu düşünmesi beni endişelendirir.	-.04	.05	.95	.98
M12	Sınıfta tartışılan bir konu hakkında yanlış şeyler söyleyip alay konusu olmamak için sessiz kalmayı tercih ederim.	.19	.05	.95	.90
M13	Sınıf arkadaşlarımla hakkımda ne düşündüğü ile ilgili çok fazla endişelendiğimi düşünüyorum.	.31	.05	.88	1.02
M14	Öğretmenlerime karşı mahcup olma korkusundan dolayı sınıfta kendimi rahat ifade edemem.	.12	.05	.95	.92
M15	Arkadaşlarımla tarafından eleştirilme ihtimalinden dolayı grup çalışmalarından kaçınırım.	.70	.05	1.24	1.26
M16	Ders esnasında emin olmadığım fikirleri paylaşmaktansa sessiz kalmayı tercih ederim.	.01	.05	1.10	1.32
Ortalama		.00	.05	1.00	1.03
Standart Sapma		.34	.00	.12	.14
Güvenirlilik = .98	Ayrırma Oranı = 7.22	Ki-Kare = 728.60	sd = 15	p = .00	

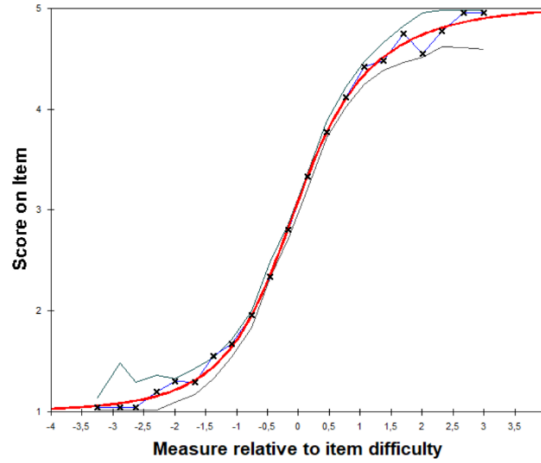
Tablo 4'e göre, ODKÖ-ÖF'deki maddelerin güçlük düzeyleri .70 logit ile -.67 logit arasında değişmekte olup maddelerin güçlük düzeylerine ilişkin aralık 1.37 logittir. Tablo 4'te görüldüğü üzere, uyum içi ve uyum dışı istatistiklerine ilişkin ortalamalar sırasıyla 1.00 ve 1.03 logit olarak bulunmuştur. Uyum istatistiklerine ilişkin istenen değerin 1.00 olduğu bilinmekte ve bu değere yakın olan uyum istatistikleri model ile verilerin uyumlu olduğunun göstergesi olarak kabul edilmektedir (Behizadeha & Engelhard, 2014). Buna göre, model ile veri arasında yüksek bir uyum bulunmaktadır. Rasch modeli tek boyutluluk varsayımına dayalı olduğundan, model ile veri arasında tespit edilen yüksek uyum ODKÖ-ÖF için öngörülen tek boyutlu yapının veriler tarafından doğrulandığının işareti olarak değerlendirilebilir.

Tablo 4'e bakıldığında madde yüzeyine ilişkin güvenirlilik indeksi .98 olarak görülmektedir. Rasch analizinde madde yüzeyine ilişkin güvenirlilik katsayısının yüksek olması, ölçme aracındaki maddelerin güçlük düzeylerinin birbirinden farklı olduğunu yansıtır (Haiyang, 2010; İlhan, 2016). Buna göre, ODKÖ-ÖF'de madde yüzeyi için hesaplanan güvenirlilik indeksi; ölçeğin öğrencilerin katılma olasılıkları farklı olan maddelerden oluştuğunu; bir başka deyişle, ölçek maddelerinin güçlük düzeyleri açısından değişkenlik gösterdiğini ifade etmektedir. Ki-kare testi sonuçlarının anlamlı olması (ki-kare=728.60, $sd=15$, $p<.001$), ODKÖ-ÖF'deki maddelerin güçlük düzeyleri arasında anlamlı fark bulunduğunu ortaya koyan bir başka istatistiksel göstergedir. Tablo 5'te, ODKÖ-ÖF'de kullanılan beşli derecelendirmenin ne derece etkin çalıştığına ilişkin bulgular sunulmuştur.

Tablo 5. ODKÖ-ÖF'ye ilişkin kategori istatistikleri

Kategori	Frekans	Yüzde	Ortalama Ölçüm	Beklenen Ölçüm	Uyum Dışı
1	1448	%25	-.73	-.74	1.0
2	1140	%20	-.42	-.40	.90
3	961	%16	-.11	-.11	1.0
4	1346	%23	.19	.18	1.0
5	945	%16	.56	.56	1.1

Puanlama ölçeğinin uygun biçimde çalıştığının söylenebilmesi için, ölçek kategorilerine ilişkin frekans değerlerinin düzenli olması ve ölçeğin her bir kategorisinde en az 10 gözlem bulunması gerekmektedir (Linacre, 2014). Tablo 5'teki frekans değerleri bu gerekliliği karşılamaktadır. Buna göre, ölçekte kullanılan beşli derecelendirmenin etkin bir şekilde çalıştığı ifade edilebilir. Tablo 5'te yer alan uyum istatistiklerinin .50 ile 1.50 kabul edilebilir aralığı içinde kalması ve puanlama ölçeğinin kategorileri arttıkça ortalama ölçümlerin de artması ODKÖ-ÖF'deki beşli derecelendirmenin etkin biçimde çalıştığını ortaya koyan diğer bulgulardır. Logit cetvel, öğrenci ve madde yüzeyine ilişkin ölçüm raporları ile kategori istatistiklerinin ardından test karakteristik eğrisi incelenmiştir. Test karakteristik eğrisi Şekil 2'de sunulmuştur.

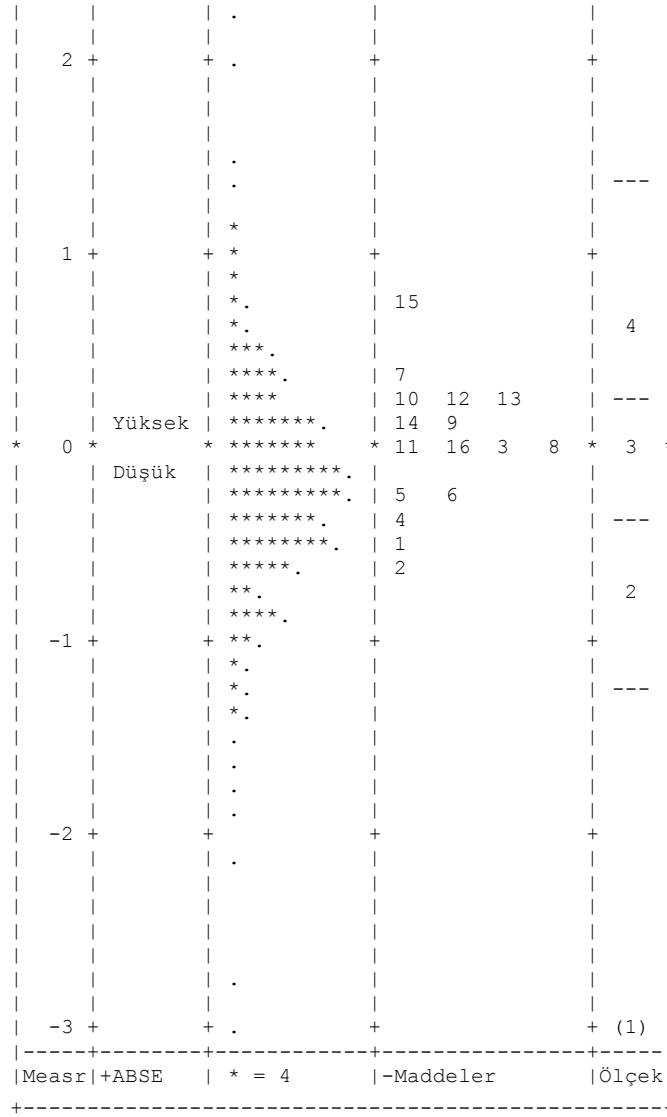


Şekil 2. Gözlenen ve beklenen test karakteristik eğrisi ile ODKÖ-ÖF için model-veri uyumu

Şekil 2’de yer alan kırmızı renkli düz çizgi, beklenen test karakteristik eğrisini gösterirken; üzerinde çarpı işaretlerinin bulunduğu mavi çizgi gözlenen test karakteristik eğrisini temsil etmektedir. Test karakteristik eğrisine ilişkin gözlenen ve beklenen eğrilerin örtüşmesi, model ile veri arasında kabul edilebilir uyum bulunduğunu ifade etmektedir. Ayrıca, Şekil 2’ye bakıldığında yetenek ölçeğinin orta noktalarında gözlenen ve beklenen test karakteristik eğrileri arasında tam bir uyumun bulunduğu fakat eğrinin uç noktalarına doğru gidildikçe bu uyumda düşüş yaşandığı görülmektedir. Bu durum; ODKÖ-ÖF’deki maddelere kıyasla, bireylerin yetenek ölçeği boyunca daha geniş bir dağılım göstermesi ile açıklanabilir. Çünkü madde tepki kuramına dayalı modellerde en doğru yetenek kestirimleri bireyin yetenek düzeyi ile maddenin güçlük düzeyinin eşleştiği ölçme koşullarında elde edilmektedir (Bond & Fox, 2015; Crocker, & Algina, 1986). Bireyin yetenek düzeyi ve maddenin güçlük düzeyi arasındaki fark arttığında ise ölçme işlemine karışan hata daha fazla olmaktadır. Logit cetvel incelendiğinde görüldüğü üzere, ODKÖ-ÖF’deki maddelerin güçlük düzeyleri ± 1 logit aralığında kalmaktadır. Bu yüzden ODKÖ-ÖF’de en hassas yetenek kestirimleri olumsuz değerlendirilme korkuları orta düzeyde olan bireyler için hesaplanmış; olumsuz değerlendirilme korkuları yüksek ya da düşük olan bireyler için ölçme işlemine karışan hata daha yüksek bulunmuştur. Örneklendirmek gerekirse; ölçme işlemine ilişkin standart hata değeri; olumsuz değerlendirilme korkusu -0.05 logit olan birey için $.20$ iken; olumsuz değerlendirilme korkusu -4.37 logit olan birey için ise 1.82 şeklindedir. Buna bağlı olarak, yetenek ölçeğinin orta noktalarında gözlenen ve beklenen test karakteristik eğrileri arasında bulunan mükemmel uyum; yetenek ölçeğinin alt ve üst ucunda kısmi bir düşüş göstermiştir.

Öğrenci ve madde şeklinde iki yüzeyden oluşan Rasch analizi uygulandıktan sonra, hem Likert tipi ölçeklerden alınan puanlar üzerinden gerçekleştirilen farka dayalı istatistiklerde çok yüzeyle Rasch modelinin kullanıldığı bir örnek sunmak hem de ODKÖ-ÖF’nin ölçüt geçerliğine ilişkin kanıt elde etmek amacıyla öğrencilerin olumsuz değerlendirilme korkularının ABSE puanları ile ilişkili olup olmadığı hipotezi test edilmiştir. Bu doğrultuda; ilk olarak öğrencileri ABSE puanları açısından görece türdeş gruplara ayırabilmek amacıyla kümeleme analizi uygulanmıştır. Kümeleme analizi sonucunda, çalışma grubundaki 367 öğrencinin ABSE puanlarına göre kendi içerisinde benzer (homojen) ve kendi aralarında farklı (heterojen) iki gruba ayrıldığı belirlenmiştir. 265 öğrenciden oluşan birinci grupta öğrencilerin ABSE puanlarına ilişkin ortalama 4.09 olarak bulunmuştur. Dolayısıyla bu kümenin ABSE puanları “*Yüksek*” olarak isimlendirilmiştir. 102 öğrencinin bulunduğu ikinci grupta öğrencilerin ABSE puanlarına ait ortalama 2.73 olarak hesaplanmıştır. Buna bağlı olarak, ikinci kümenin ABSE puanları “*Düşük*” olarak adlandırılmıştır. Daha sonra, oluşan bu iki küme bir değişkenlik kaynağı olarak Rasch analizine dâhil edilip; öğrenciler, maddeler ve ABSE puanları şeklindeki üç yüzeyle bir desen ile Rasch analizi gerçekleştirilmiştir. Analizden elde edilen logit cetvel Şekil 3’te sunulmuştur.

Measr	+ABSE	+Birey	-Maddeler	Ölçek
3	+	+	+	(5)



Şekil 3. ODKÖ-ÖF'ye ilişkin üç yüzeyli (öğrenciler, maddeler ve ABSE puanları) modele ait logit cetvel

Şekil 3'te, logit ölçeğinin negatif ucundan pozitif ucuna doğru gidildikçe öğrencilerin olumsuz değerlendirilme korkuları artmaktadır. Yine Şekil 3'e göre; ABSE puanları yüksek olan öğrenciler logit ölçeğinin pozitif ucunda, ABSE puanları düşük olan öğrenciler ise logit ölçeğinin negatif ucunda yer almaktadır. Dolayısıyla ABSE puanlarına göre öğrencilerin logit cetvel üzerindeki konumları, olumsuz değerlendirilme korkusu arttıkça akademik beklentilere ilişkin stresin de arttığına işaret etmektedir. Ancak, akademik beklentilere ilişkin stres ile olumsuz değerlendirilme korkusu arasında anlamlı bir ilişki olup olmadığını belirleyebilmek için ABSE puanları yüzeyine ait ölçüm raporlarının incelenmesi gerekmektedir. ABSE puanları değişkenine ilişkin ölçüm raporları Tablo 6'da sunulmuştur.

Tablo 6. ODKÖF'ye ilişkin üç yüzeyli modelde ABSE puanları değişkenine ait ölçüm raporları

	Ölçüm	Standart Hata	Uyum İçi	Uyum Dışı
ABSE puanı-Yüksek	.16	.01	.99	1.01
ABSE puanı-Düşük	-.16	.02	1.02	1.09
Ortalama	.00	.02	1.00	1.05
Standart Sapma	.22	.01	.02	.06
Güvenirlilik = .99	Ayrırma Oranı = 12.03	Ki-Kare = 145.70	sd = 1	p = .00

Tablo 6'ya göre, olumsuz değerlendirilme korkusuna ilişkin ortalamalar ABSE puanı yüksek olan öğrenciler için .16 logit iken, ABSE puanı düşük olan öğrenciler için -.16 logittir. Buna göre,

ABSE puanı düşük ve yüksek olan öğrencilerin olumsuz değerlendirilme korkuları arasında .32 logitlik bir fark bulunmaktadır. Hesaplanan ki-kare değeri ($ki-kare=145.70$, $sd=1$, $p<.01$), ABSE puanlarına göre öğrencilerin olumsuz değerlendirilme korkuları arasında gözlenen söz konusu farkın istatistiksel açıdan anlamlı olduğunu yansıtmaktadır. Güvenirlilik indeksine ilişkin .99 değeri, ABSE puanı yüksek ve düşük olan öğrencilerin araştırmada ölçülen örtük özellik (olumsuz değerlendirilme korkusu) bakımından yüksek güvenirlilikte birbirinden ayırt edildiğini göstermektedir.

4. TARTIŞMA ve SONUÇ

Bu araştırmada, ODKÖ-ÖF geliştirilmiş ve ölçeğin psikometrik özellikleri Rasch analizi ile incelenmiştir. Ayrıca çok yüzeyli Rasch modeli uygulanarak öğrencilerin olumsuz değerlendirilme korkularının ABSE puanlarına göre farklılık gösterip göstermediği test edilmiştir. ODKÖ-ÖF'nin geliştirilmesi sürecinde öncelikle ODKÖ-GF (Watson & Friend, 1969), ODKÖ-KF (Leary, 1983) ve ilgili alanyazından yararlanarak 20 madde içeren bir madde havuzu oluşturulmuştur. Olumsuz değerlendirilme korkusuna ilişkin kuramsal yapı tek faktörlü bir yapıyı öngördüğünden, ODKÖ-ÖF'nin madde havuzu bu tek faktörlü yapı referans alınarak hazırlanmıştır. Daha sonra, ölçeğin kapsam ve görünüş geçerliğini sağlamak için uzman görüşü alınmıştır. Uzman görüşleri doğrultusunda; bazı maddelerin ifade ediliş şekillerinde değişikliğe gidilmiş, bir madde ölçekten çıkarılmış ve uzmanlar tarafından önerilen bir madde ölçeye eklenmiştir. Böylece 20 maddelik taslak bir ölçme aracı elde edilmiştir. Ölçekteki maddeler, *Kesinlikle Katılıyorum(5)* → *Kesinlikle Katılmıyorum(1)* şeklinde beşli Likert tipi bir derecelendirme ile çalışma grubundaki öğrencilere uygulanmıştır.

Araştırmadan elde edilen bulgular; ODKÖ-ÖF'nin taslak formundaki 20 maddeden üçünün model ile veri arasındaki uyumu olumsuz yönde etkilediğini ve bir maddeye ait nokta çift serili korelasyon değerinin düşük olduğunu göstermiştir. Bu nedenle, söz konusu dört madde ölçekten çıkarılarak Rasch analizi tekrarlanmıştır. 16 madde üzerinden yinelenen Rasch analizinde, maddelere ilişkin uyum istatistiklerinin kabul edilebilir sınırlar içerisinde kaldığı ve madde korelasyonlarının .30 ölçütünü aştığı tespit edilmiştir. Bu tespitin ardından, logit cetvel ve kategori istatistikleri tablosu ile madde ve birey yüzeylerine ilişkin ölçüm raporları detaylı bir şekilde incelenmiştir. Buna göre, 16 maddeden oluşan ve ölçek geliştirme sürecinde temele alınan tek boyutlu modelde, bireyler ile maddelerin yüksek güvenirlilikte birbirinden ayırt edildiği belirlenmiştir. Kategori istatistikleri tablosu incelendiğinde, ölçekte kullanılan beşli derecelemenin etkin bir biçimde çalıştığı saptanmıştır. Gerek madde gerekse birey yüzeyine ait uyum istatistiklerinin, bu istatistikler için dikkate alınması önerilen .50 ile 1.50 aralığı içerisinde kaldığı sonucuna ulaşılmıştır. Son olarak, gözlenen ile beklenen test karakteristik eğrilerinin büyük ölçüde örtüştüğü gözlenmiştir. Uyum istatistikleri ve test karakteristik eğrisine ilişkin bu sonuçlar, model-veri uyumunun sağlandığını ortaya koymaktadır. Bu bulgulara dayanarak, araştırma verilerinin ölçek geliştirme sürecinde esas alınan tek boyutlu yapıyı doğruladığı şeklinde bir çıkarıma varılabilir.

İki yüzeyli Rasch analizinden elde edilen bulgular, ODKÖ-ÖF'nin geçerli ve güvenilir ölçümler üreten bir ölçme aracı olduğunu yansıtmaktadır. Ancak Rasch analizinde ulaşılan bulgular her zaman bu yönde olmayabilir. Bazen rapor edilen çıktılar; madde ya da birey yüzeyine ilişkin güvenirlilik katsayısının düşük olduğunu ve ölçeğin psikometrik özelliklerinin iyileştirilmesi gerektiğini ortaya koyabilir. İstenilen düzeyde geçerli ve güvenilir ölçümlerin elde edilemediği bu durumlarda, ölçeye eklenecek maddeler ile ölçeğin psikometrik özelliklerinin iyileştirilmesi mümkündür. Rasch analizinin klasik test kuramının ötesinde sağladığı en önemli avantajlarından biri bu noktada karşımıza çıkmaktadır. Çünkü Rasch analizi sonucunda, ölçeğin ayırt ediciliğinin artırılabilmesi için ölçme aracına yaklaşık olarak hangi güçlük düzeyinde maddelerin eklenmesi gerektiği tespit edilebilmektedir. Örneğin; ölçek maddelerinin büyük bir kısmının logit cetvelin negatif ucunda ve görece az sayıda maddenin logit cetvelin pozitif ucunda bulunması halinde, katılımcıların birbirinden daha iyi ayırt edilebilmesi için güçlük düzeyi yüksek olan maddelerin ölçeye eklenmesi gerektiği şeklinde bir yorum yapılabilir. Diğer yandan; ölçek maddelerinin büyük bir kısmının logit cetvelin üst kısmında bulunması ve logit cetvelin alt ucunda az sayıda maddenin yer alması, ölçeğin ayırt ediciliğinin artırılabilmesi için güçlük düzeyi düşük olan maddelerin ölçeye eklenmesi gerektiği anlamına gelebilir. Bu araştırmada, ODKÖ-ÖF'ye ait uyum istatistiklerinin istenilen aralıkta bulunduğu, madde ve birey yüzeylerine ilişkin yüksek güvenirlilik

katsayılarının elde edildiği ve çalışma grubundaki öğrencilerin logit cetveli boyunca geniş bir dağılıma sahip olduğu göz önüne alındığında ölçekte yer alan 16 madde dışında yeni maddelerin ölçme aracına eklenmesine ihtiyaç olmadığı sonucuna varılmıştır.

Rasch analizi, ölçme aracına eklendiğinde ölçeğin psikometrik özelliklerini iyileştirebilecek maddeler hakkında bilgi verebildiği gibi (Barker, Donovan, Schubert & Walker, 2016) ölçme aracından çıkarılması durumunda ölçeğin geçerlik ve güvenilirliğinde bir kayba yol açmayacak maddelerin belirlenmesine de olanak tanımaktadır (Zile-Tamsen, 2017). Brinthaup ve Kang (2014); aynı kapsamı ölçmeye yönelik olan ve logit cetvelinin aynı noktasında bulunan maddeleri, ölçeğin sağladığı içerikten anlamlı bir kayıp yaşanmadan ölçekten çıkarılabilecek maddeler olarak ifade etmektedir. Bu bağlamda, bir ölçekteki madde sayısının fazla olduğu düşünülüyor ve ölçeğin kullanılabilirliğinin artırılabilmesi için ölçme aracındaki madde sayısının azaltılması gerektiğine inanılıyorsa ilk olarak logit cetvelinin aynı noktasında bulunan maddelere bakılmalıdır. Logit cetveli üzerinde aynı noktada bulunan maddelerin içerik açısından da aynı özelliği ölçtüğünün belirlenmesi halinde, ilgili maddelerden yalnızca birinin ölçekte kalması yeterli olacak; diğerlerinin çıkarılması ölçeğin kapsam geçerliğinde ve ayırt ediciliğinde herhangi bir eksilmeye yol açmayacaktır (McCamey; 2014; Milliken vd., 2018). Bu çalışmada; ODKÖ-ÖF'deki madde sayısının ölçeğin kullanılabilirliğini olumsuz yönde etkileyebilecek bir büyüklükte olmadığı dikkate alındığında ölçekten madde çıkarılmasına gerek görülmemiştir.

İki yüzeyli Rasch analizinin ardından çok yüzeyli Rasch analizi uygulanmıştır. Birey, madde ve ABSE puanları şeklinde üç yüzey ile gerçekleştirilen çok yüzeyli Rasch analizi sonucunda, öğrencilerin olumsuz değerlendirilme korkularının ABSE puanlarına göre farklılaştığı tespit edilmiştir. Bu noktadan hareketle, olumsuz değerlendirilme korkusu ve akademik beklentilere ilişkin stres arasında anlamlı ilişki bulunduğu söylenebilir. Alanyazındaki çalışmalar (Kelecioğlu & Bilge, 2009) ile paralellik gösteren bu bulgu, ODKÖ-ÖF'nin ölçüt geçerliğine ilişkin bir kanıt sunmaktadır. Sonuç olarak, bu araştırma ile hem lise öğrencilerinin sınıf ortamında öğretmenleri ya da arkadaşları tarafından olumsuz değerlendirileceklerine dair yaşadıkları korkuyu ölçmek amacıyla kullanılabilir bir ölçek geliştirilmiş hem de Likert tipi ölçeklerin psikometrik özelliklerinin incelenmesinde ve bu ölçeklerden alınan puanlar üzerinden gerçekleştirilen farka dayalı istatistiklerde Rasch modelinin kullanıma örnek olabilecek bir kaynak alanyazına sunulmaya çalışılmıştır.

5. KAYNAKLAR

- Alkan, V. (2015). *Akademik ortamlarda olumsuz değerlendirilme korkusu ölçeğinin geliştirilmesi*. Yayınlanmamış yüksek lisans tezi, Ankara Üniversitesi, Ankara, Türkiye.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573. <http://dx.doi.org/10.1007/BF02293814>
- Anshel, M.H., Weatherby, N.L., Kang, M., & Watson, T. (2009). Rasch calibration of a unidimensional perfectionism inventory for sport. *Psychology of Sport and Exercise*, 10(1), 210-216. <http://dx.doi.org/10.1016/j.psychsport.2008.07.006>
- Aydın, S., Yavuz, F., & Yeşilyurt, S. (2006). Test anxiety in foreign language learning. *Balikesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(16), 146-160. <http://sbe.balikesir.edu.tr/dergi/edergi/c9s16/makale/c9s16m8.pdf> adresinden alınmıştır.
- Aypay, A. (2011). İlköğretim II. kademe öğrencileri için okul tükenmişliği ölçeği: Geçerlik ve güvenilirlik çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(2), 511-527. <http://toad.edam.com.tr/sites/default/files/pdf/ilkogretim-ii-kademe-ogrencileri-icin-okul-tukenmisligi-olcegi-toad.pdf> adresinden alınmıştır.
- Barker, B.A., Donovan, N.J., Schubert, A.D., & Walker, E.A. (2016). Using Rasch analysis to examine the item-level psychometrics of the infant-toddler meaningful auditory integration scales. *Speech, Language and Hearing*, 20(3), 130-143. <http://dx.doi.org/10.1080/2050571X.2016.1243747>
- Berber Çelik, Ç. (2014). *Akademik ertelemenin bazı psiko-sosyal değişkenlere göre açıklanması ve gerçeklik terapisine dayalı akademik erteleme ile başa çıkma eğitim programının etkililiğinin sınanması*. Yayınlanmamış Doktora Tezi, Karadeniz Teknik Üniversitesi, Trabzon, Türkiye.
- Beghetto, R.A. (2009). Correlates of intellectual risk taking in elementary school science. *Journal of Research in Science Teaching*, 46(2), 210-223. <https://doi.org/10.1002/tea.20270>
- Behizadeha, N., & Engelhard, G. (2014). Development and validation of a scale to measure perceived authenticity in writing. *Assessing Writing*, 21, 18-36. <http://dx.doi.org/10.1016/j.asw.2014.02.001>
- Bond, T.G., & Fox, C.M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Erlbaum.
- Brentari, E., & Golia, S. (2008). Measuring job satisfaction in the social services sector with the Rasch model. *Journal of Applied Measurement*, 9(1), 45-56. Retrieved from <http://www.unibs.it/sites/default/files/ricerca/allegati/10061.pdf>
- Brinthead, T.M., & Kang, M. (2014). Many-faceted rasch calibration: An example using the self-talk scale. *Assessment*, 21(2) 241-249. <http://dx.doi.org/10.1177/1073191112446653>
- Büyüköztürk, Ş. (2010). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi Yayınları.
- Castello, E. (2008). *Text complexity and reading comprehension tests*. Bern: Peter Lang
- Clifford, M.M. (1991). Risk taking: Theoretical, empirical and educational considerations. *Educational Psychologist*, 26(3), 263-297. <http://dx.doi.org/10.1080/00461520.1991.9653135>
- Cocoradă, E. (2011). Academic self-handicapping and their correlates in adolescence. *Bulletin of the Transilvania University of Brasov*, 4(53), 57-64. Retrieved from https://www.researchgate.net/publication/265796099_Academic_self-handicapping_and_their_correlates_in_adolescence
- Crawford, L., Leuzinger, J., Brannon, S., Hamner, J. (2015). Fear of negative evaluation: Differences amongst librarians. *Library Leadership & Management*, 29(3), 1-13. Retrieved from <https://journals.tdl.org/llm/index.php/llm/article/view/7096/6330>
- Çam, S., Sevimli, D., & Yerlikaya, E. (2010). Olumsuz değerlendirilmekten korkma ölçeği'ne (ODKÖ) ilişkin bir geçerlik ve güvenilirlik çalışması. *Ç.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 19(1), 132-140. <http://dergipark.ulakbim.gov.tr/cusosbil/article/viewFile/5000001396/5000002087> adresinden alınmıştır.
- Çekirdek, G. (2014). *Hazırlık sınıfı öğrencilerinin İngilizce başarılarını etkileyen bazı faktörlerin incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Eskişehir Osmangazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir, Türkiye.
- Çetin, B., İlhan, M., & Yılmaz, F. (2014). Olumsuz değerlendirilme korkusu ve akademik risk alma arasındaki ilişkinin kanonik korelasyonla incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 14(1), 135-158. <http://dx.doi.org/10.12738/estp.2014.1.1616>
- DeMars, C. (2010). *Item response theory*. Oxford, UK: Oxford University Press.
- Demirtaşlı, N., Yalçın, S., & Ayan, C. (2016). Ölçme ve değerlendirme dersine yönelik tutum ölçeğinin madde tepki kuramına dayalı olarak geliştirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(1), 133-144. <http://dx.doi.org/10.21031/epod.43804>

- Dinnel, D.L., Brittain, T., Thompson, T., Johnson, K., King, M., & Pust, K. (2002, August). *A structural model of self-worth protection and achievement: goals, evaluative anxiety, attributions, self-esteem and uncertainty*. Poster presentations in the Second Biennial Self-Concept Enhancement and Learning Facilitation (SELF) Research Centre International Conference, Sydney, Australia. Retrieved from http://www.ibrarian.net/navon/paper/A_Structural_Model_of_Self_Worth_Protection_and_A.pdf?paperid=1240182
- Eckes, T. (2009). Many-facet Rasch measurement. In S. Takala (Ed.), *Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (Section H)*. Strasbourg, France: Council of Europe/Language Policy Division.
- Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York, NY: Routledge.
- Güler, N., İlhan, M., Güneyli, A., & Demir, S. (2017). An evaluation of the psychometric properties of three different forms of daly and miller's writing apprehension test through Rasch analysis. *Educational Sciences: Theory & Practice*, 17(3), 721–744. <http://dx.doi.org/10.12738/estp.2017.3.0051>
- Haiyang, S. (2010). An application of classical test theory and many facet Rasch measurement in analyzing the reliability of an English test for non-English major graduates. *Chinese Journal of Applied Linguistics*, 33(2), 87-102. Retrieved from <http://www.celea.org.cn/teic/90/10060807.pdf>
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hasford, J., & Bradley, K.D. (2011). Validating measures of self control via rasch measurement. *Journal of Applied Business Research*, 27(6), 45-55. <http://dx.doi.org/10.19030/jabr.v27i6.6465>
- İlhan, M. (2016). Açık uçlu sorularla yapılan ölçmelerde klasik test kuramı ve çok yüzeyle Rasch modeline göre hesaplanan yetenek kestirimlerinin karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(2), 346-368. <http://dx.doi.org/10.16986/HUJE.2016015182>
- İlhan, M., & Çetin, B. (2013). Ortaokul öğrencilerinin matematik odaklı akademik risk alma davranışları: Bir ölçek geliştirme çalışması. *e-International Journal of Educational Research*, 4(2), 1-28. <http://dergipark.ulakbim.gov.tr/ijer/article/view/1073000246/1073000122> adresinden alınmıştır.
- Kelecioğlu, H., & Bilge, K. (2009). Akademik beklentilere ilişkin stres envanterinin uyarlanması: Geçerlik ve güvenirlik çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 36, 148-157. <http://dergipark.ulakbim.gov.tr/hunefd/article/view/5000048423/5000045743> adresinden alınmıştır.
- Leary, M.R. (1983). A brief version of the Fear of Negative Evaluation Scale. *Personality and Social Psychology Bulletin*, 9(3), 371-375. <http://dx.doi.org/10.1177/0146167283093007>
- Leary, M.R., & Kowalski, R.M. (1995). *Social anxiety*. New York: Guilford Publications Inc.
- Lee, M., Peterson, J. J., & Dixon, A. (2010). Rasch calibration of physical activity self-efficacy and social support scale for persons with intellectual disabilities. *Research in Developmental Disabilities*, 31(4), 903–913. <http://dxdoi.org/10.1016/j.ridd.2010.02.010>
- Linacre, J.M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, 3(1), 85-106. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.424.2811&rep=rep1&type=pdf>
- Linacre, J.M. (2014). *A user's guide to FACETS Rasch-model computer programs*. Retrieved from <http://www.winsteps.com/a/facets-manual.pdf>
- Lubetkin, B., & Oumano, E. (1991). *Bailing out: The sane way out of a doomed relationship & survive with hope and self-respect*. New York: Simon and Schuster.
- McCamey, R. (2014). A primer on the one-parameter Rasch model. *American Journal of Economics and Business Administration*, 6(4), 159-163. <http://dx.doi.org/10.3844/ajebasp.2014.159.163>
- Mckinney, A.P. (2003). *Goal orientation: A test of competing models* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3157771).
- Meads, D.M., & Bentall, R.P. (2008). Rasch analysis and item reduction of the hypomanic personality scale. *Personality and Individual Differences*, 44(8), 1772-1783. <http://dx.doi.org/10.1016/j.paid.2008.02.009>
- Midgley, C., & Urdan, T. (2001). Academic self-handicapping and achievement goals: A further examination. *Contemporary Educational Psychology*, 26(1), 61-75. <https://doi.org/10.1006/ceps.2000.1041>
- Milliken, A., Ludlow, L., DeSanto-Madeya, S., & Grace, P. (2018). The development and psychometric validation of the ethical awareness scale. *Research Methodology: Instrument Development*. Advance online publication. <https://doi.org/10.1111/jan.13688>

- Moeeni, N., & Nejhad, M.N.M. (2014). The relationship between worry and fear of negative evaluation with social anxiety in students. *International Journal of Basic Sciences & Applied Research*, 3(1), 31-34. Retrieved from <http://isicenter.org/fulltext/paper-206.pdf>
- Özgüngör, S. (2006). Öz bilinç, olumsuz değerlendirilme korkusu, performans odaklı sınıf algısı ve not yönelimi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 19, 1-8. http://pauegitimdergi.pau.edu.tr/Makaleler/1357398936_9-%c3%96Z%20B%c4%b0L%c4%b0N%c3%87.pdf adresinden alınmıştır.
- Rameli, M.R.M., & Kosnin, A.M. (2017) Development and validation of math anxiety scale for secondary school students (MAS-SS): Application of Rasch analysis. *Man in India*, 97(19). 161-172. Retrieved from <http://eprints.utm.my/id/eprint/76856/>
- Rapee, R.M., & Heimberg, R. G. (1997). A cognitive behavioral model of anxiety in social phobia. *Behavioral Research and Therapy*, 35(8), 741-756. [https://doi.org/10.1016/S0005-7967\(97\)00022-3](https://doi.org/10.1016/S0005-7967(97)00022-3)
- Ricketts, S.N., Engelhard, G., & Chang, M.L. (2015). Development and validation of a scale to measure academic resilience in mathematics. *European Journal of Psychological Assessment*. <http://dx.doi.org/10.1027/1015-5759/a000274>
- Schaufeli, W. B., Martinez, I., Marques-Pinto, A., Salanova, M., ve Bakker, A. (2002). Burnout and engagement in university students: A cross-national study. *Journal of Cross-Cultural Studies*, 33(5), 464-481. <https://doi.org/10.1177/0022022102033005003>
- Shabani, M.B. (2012). Levels and sources of language anxiety and fear of negative evaluation among Iranian EFL learners. *Theory and Practice in Language Studies*, 2(11), 2378-2383. <https://doi.org/10.4304/tpls.2.11.2378-2383>
- Taşdelen Teker, G., Güler, N. & Kaya Uyanık, G. (2015). Comparing the effectiveness of SPSS and EduG using different designs for generalizability theory. *Educational Sciences: Theory & Practice*, 15(3), 635-645. <http://dx.doi.org/10.12738/estp.2015.3.2278>
- Tezbaşaran, A. (1997). *Likert tipi ölçek geliştirme kılavuzu*. Ankara: Türk Psikologlar Derneği.
- Utschig, A.C., Presnell, K., Madeley, M.C., Smits, J.A.J. (2010). An investigation of the relationship between fear of negative evaluation and bulimic psychopathology. *Eating Behaviors*, 11(4), 231-238. <https://doi.org/10.1016/j.eatbeh.2010.05.003>
- Uzun, G. (2017). *Akademik başarının okul, aile ve öğrenci özellikleri ile ilişkisinin çok düzeyli yapısal eşitlik modellenmesi ile incelenmesi*. Yayınlanmamış doktora tezi, Ankara Üniversitesi, Ankara, Türkiye.
- Ünal, E., Arık, S., & Uzun, B. (2016). Sınıf ve sosyal bilgiler eğitiminde öğrenim gören öğretmen adaylarının olumsuz değerlendirilme korkuları. *Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 22, 289-307. <http://www.mulkiyedergi.org/adyusbd/article/view/5000124109/5000164785> adresinden alınmıştır.
- Walker, E.R., Engelhard, G., & Thompson, N.J. (2012). Using Rasch measurement theory to assess three depression scales among adults with epilepsy. *Seizure*, 21(6), 437-443. <http://dx.doi.org/10.1016/j.seizure.2012.04.009>
- VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement*, 57(6), 995-1015. <https://doi.org/10.1177/0013164497057006009>
- Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, 33(4), 448-457. <http://dx.doi.org/10.1037/h0027806>
- Weeks W.J., Heimberg R.G., Rodebaugh T.L. (2008). The fear of positive evaluation scale: Assessing a proposed cognitive component of social anxiety. *Journal of Anxiety Disorders*, 22(1) 44-55. <https://doi.org/10.1016/j.janxdis.2007.08.002>
- Wright, B.D., & Linacre, J.M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370-371. Retrieved from <https://www.rasch.org/rmt/rmt83b.htm>
- Wright, B.D. & Masters, G. (1982). Rating scale analysis: Rasch measurement. Chicago: MESA Press.
- Zile-Tamsen, C.V. (2017). Using Rasch analysis to inform rating scale development. *Research in Higher Education*, 58(8), 922-933. <https://doi.org/10.1007/s11162-017-9448-0>