

**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME VE DEĞERLENDİRME ANA BİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI**

**ÇOKLU PUANLANAN ÖLÇEKLERDE ANGOFF, ROC VE SINIR GRUP
YÖNTEMLERİ İLE KESME PUANININ BELİRLENMESİ**

DOKTORA TEZİ

DÖNDÜ BAHAR ŞAHİN SARKIN

Ankara, Mayıs, 2017



**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME VE DEĞERLENDİRME ANA BİLİM DALI
ÖLÇME VE DEĞERLENDİRME PROGRAMI**

**ÇOKLU PUANLANAN ÖLÇEKLERDE ANGOFF, ROC VE SINIR GRUP
YÖNTEMLERİ İLE KESME PUANININ BELİRLENMESİ**

DOKTORA TEZİ

DÖNDÜ BAHAR ŞAHİN SARKIN

DANIŞMAN: Yrd. Doç. Dr. HAMİDE DENİZ GÜLLEROĞLU

Ankara, Mayıs, 2017

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

D¼nd¼ Bahar Őahin Sarkın'ın hazırladıđı "Çoklu Puanlanan ¼lçeklerde Angoff, ROC ve Sınır Grup Y¼ntemleri ile Kesme Puanının Belirlenmesi" bařlıklı bu alıřma j¼rimiz tarafından Eđitimde ¼lme ve Deđerlendirme Anabilim Dalı / Programı'nda Doktora Tezi olarak kabul edilmiřtir.

İmza

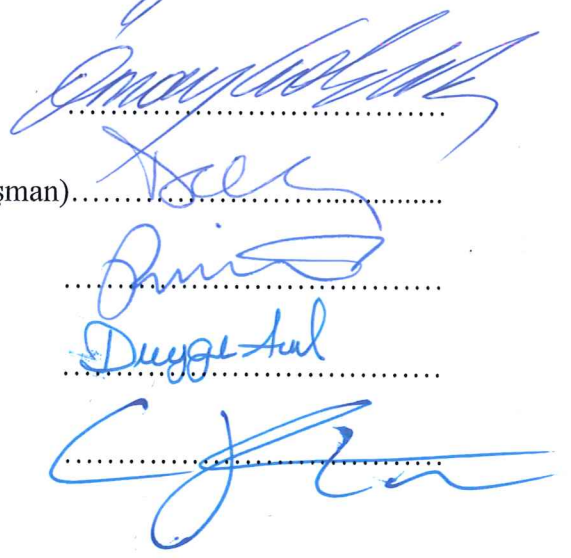
Bařkan: Do. Dr. ¼may okluk B¼KEOđLU

¼ye: Yrd. Do. Dr. H. Deniz G¼LLEROđLU (Danıřman)

¼ye: Yrd. Do. Dr. Fazilet TAŐDEMİR

¼ye: Do. Dr. Duygu ANIL

¼ye: Yrd. Do. Dr. C. Deha DOđAN



ONAY

Bu tez Ankara ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınay Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından/..../2017 tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca/..../2017 tarihinde kabul edilmiřtir.

Prof. Dr. İsmail G¼VEN

Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.



Döndü Bahar Şahin SARKIN

ÖZET

ÇOKLU PUANLANAN ÖLÇEKLERDE ANGOFF, ROC VE SINIR GRUP YÖNTEMLERİ İLE KESME PUANININ BELİRLENMESİ

Şahin Sarkın, Döndü Bahar

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı: Yrd. Doç. Dr. Hamide Deniz

Gülleroğlu Mayıs 2017, xii + 71 Sayfa

Bu araştırmada, üniversite öğrencilerinin kaygı düzeylerini belirleyen bir ölçeğin Angoff, ROC ve Sınır Grup yöntemleri kullanılarak kesme puanının belirlenmesi ve bu puana göre yüksek/düşük kaygı sınıflama uyumunun incelenmesiyle geçerlik çalışmalarına katkıda bulunmak amaçlanmıştır. Araştırma çoklu puanlanan, geçerlik ve güvenirlik çalışması yapılmış bir ölçeğe yeni bilgiler eklenmesi nedeniyle temel araştırma niteliğindedir. Araştırma grubunu bir devlet üniversitesinin eğitim fakültesinde öğrenim gören ve Kamu Personeli Seçme Sınavı'na katılacak olan 290 öğrenci oluşturmaktadır. Veri toplama amacı ile öğrencilere Sürekli Kaygı ve Sınav Kaygısı Ölçekleri uygulanmıştır. Kaygı durumu ile ilgili kesme puanını belirlemek için rehberlik ve psikolojik danışmanlık alanında görev yapan 10 yargıcının kararına başvurulmuştur. Bu araştırmadan elde edilen bulgular, Angoff ve ROC analizi ile hesaplanan kesme puanlarının (48 ve 48.5) birbirine çok yakın olduğunu göstermiştir. Angoff ve Sınır Grup yöntemleri ile belirlenen kesme puanlarına göre yüksek/düşük kaygılı kabul edilen öğrenci yüzdeleri arasında manidar bir fark ($\chi^2_{(sd=1)}=100.11, p<.05$) olduğu sonucuna ulaşılmıştır. Angoff ve Sınır Grup yöntemlerine göre belirlenen kesme puanlarının güvenirlik kanıtı olarak yargıcılar arası uyum katsayısı incelenmiştir. Sınır Grup yöntemine göre kesme puanı belirlendiğinde sınıf içi korelasyon katsayısının orta düzeyde ($r=.53$), Angoff yöntemine göre belirlendiğinde ise iyi düzeyde ($r=.65$) olduğu sonucuna ulaşılmıştır. ROC analizine göre hesaplanan kesme puanının, öğrencilerin kaygı düzeylerini sınıflamadaki doğruluk oranı araştırılmıştır. Araştırmada elde edilen AUC (eğri altında kalan alan) değerine göre Sürekli Kaygı Ölçeği'nin yüksek ve düşük kaygılı bireyleri %73 oranında (orta düzeyde) doğru sınıflandırabildiği sonucuna ulaşılmıştır. ROC analizi ile geçerlik kanıtı olarak pozitif ve negatif yordayıcı değerler

hesaplanmıştır. Ölçeğin tanı koymak yerine tarama amacıyla kullanıldığı göz önünde bulundurulduğundan kaygı düzeylerini yüksek olasılıkla belirleyebildiği görülmüştür. Sonuç olarak bu araştırmada çoklu puanlanan bir psikolojik ölçme aracının kesme puanının belirlenmesinde ROC analizi kullanıldığında diğer yöntemlerde olmayan duyarlılık, özgüllük ve pozitif/negatif yordayıcı değerlerinin bulunması daha ayrıntılı ve nesnel bilgiler sağlamıştır.

Anahtar kelimeler: Çoklu puanlama, Kesme puanı, Angoff, ROC, Sınır Grup.



SUMMARY

DETERMINING THE CUTOFF POINT OF MULTI- SCORING SCALES THROUGH ANGOFF, ROC AND BORDERLINE METHODS

Sahin Sarkin, Döndü Bahar

Doctor of Philosophy, Program for Measurement and Evaluation

Advisor: Asst. Prof. Dr. Hamide Deniz Gülleroglu

May 2017, xii + 71 Pages

In the present study, it is aimed to contribute to validity studies by determining the cutoff point of an inventory measuring university students' anxiety levels with Angoff, ROC, and Borderline methods and by examining high/low anxiety classification according to this point. The study is regarded as a basic research because new data have been added to a multi-scoring inventory whose validity and reliability studies were previously performed. The research group consists of 290 students who study at the education faculty of a state university and who will take the public personnel selection examination. Trait Anxiety and Test Anxiety Scales were applied to the students. 10 referees working in the field of guidance and psychological counseling were asked to determine the cut-off score related to anxiety. Findings obtained from this study show that the cutoff scores (48 and 48.5) calculated by Angoff and ROC analyses are very close to each other. It is concluded that there is a meaningful difference in the percentages of students who are considered to be high/low anxious according to the cutoff points determined by Angoff and Borderline methods ($\chi^2_{(sd=1)}=100.11, p<.05$). The consistency coefficient among the judges was examined to prove the reliability of the cut-off points determined by Angoff and Borderline methods. When the cut-off point was determined according to borderline method, intra-class correlation coefficient was found to be at moderate level ($r =.53$) and when it was determined according to Angoff method, it was found to be at good level ($r =.65$). The accuracy of the cutoff score calculated according to the ROC analysis in classifying anxiety levels of the students was investigated. According to the AUC (area under the curve) value obtained in the study, it is seen that the Trait Anxiety Inventory could correctly classify high and low anxious individuals at the rate of 73% (moderate level). Positive and negative

predictive values were calculated to prove the validity by ROC analysis. It is shown that the scale is quite likely to determine anxiety levels when it is considered that it is used for screening rather than diagnosis. To conclude, the study indicates that when ROC analysis was used to determine the cutoff score of a multi-scoring psychological measuring instrument, such characteristics as sensitivity, specificity, and positive/negative predictive values that other methods do not have provided more detailed and objective information.

Key words: Multi-scoring, Cutoff Score, Angoff, ROC, Borderline.



ÖN SÖZ

Türkiye’de uygulanan tipik tepki testleri bireyleri yönlendirme, seçme ve yerleştirmede belirleyici rol oynamaktadır. Psikolojik özelliklerin tanımı geçmişten günümüze önemli değişimlere uğramadığından duyuşsal özelliklerin ölçülmesinde yeni bir ölçme aracı geliştirmek yerine güvenilirliği ve geçerliği bilinen ölçme araçları yaygın olarak kullanılmaktadır. Bireylerin çoklu puanlanan ölçeğin tepki düzeylerine göre sınıflandırılması ve yapılan sınıflandırmaların doğruluğunun sınınanmaması, çoğu zaman tartışma konusu olmaktadır. Ayrıca ölçekten elde edilen puanların nasıl yorumlanacağına net olarak raporlanmaması, ölçek geliştirme veya uyarlama çalışmalarının en çok ihmal edilen bölümü olarak belirlenmiştir. Bireyler hakkında alınan kararların doğruluğu hem ölçmenin hem de ölçütün doğruluğuna bağlıdır. Çoklu puanlanan bir ölçekten elde edilen bireyin puanı, benzer bir grubun puanları ile karşılaştırılınca anlam kazanmakta aksi hâlde ham puanları tek başına yorumlamak ölçülen özelliğin miktarı ile ilgili bilgi sağlamamaktadır. Ölçekten elde edilen puana anlamlandıran ölçüt, belirli kuramsal beklentiden de oluşturulabilmektedir.

Bu araştırmada farklı yöntemlerle kesme puanı belirlemede sadece bir kesme puanı oluşturmaktan çok bireyleri sınıflandırma konusunda farklı bir bakış açısı sunulmaya çalışılmıştır. Bu bağlamda ölçekteki kesme puanlarının ve ölçekten alınan puanların karşılık geldiği tepki düzeyinin ne anlama geldiği tanımlanabilir ve ölçeğin puanlarına dayalı olarak yapılan sınıflamaların doğruluğu hakkında bilgi sunulabilir. Ayrıca ROC analizi yönteminin sınıflama doğruluğunun incelenmesiyle ölçek puanlarının geçerliğine katkı sağlanacağı düşünüldüğünden bu araştırmanın yapılmasına karar verilmiştir.

TEŞEKKÜR

Yüksek lisans sürecinde tanıdığım, doktora eğitimi ve tez süreciyle tamamladığım öğrencilik hayatımda değerli fikirleri ve hoşgörüsüyle yanımda olan, her türlü sıkıntıda ilgisini ve desteğini esirgemeyen, her zaman örnek aldığım ve alacağım değerli danışman hocam Yrd. Doç. Dr. Hamide Deniz Gülleroğlu'na,

Bu araştırmanın gerçekleştirilmesi süreci boyunca görüşleri ve önerileri ile destek veren, yanımda olan değerli hocalarım Doç. Dr. Ömay Çokluk Bökeoğlu ve Yrd. Doç. Dr. Fazilet Taşdemir'e,

Başta Prof. Dr. Nizamettin Koç olmak üzere Ankara Üniversitesi Ölçme ve Değerlendirme Bölümü'nde derslerini aldığım değerli hocalarıma,

Farklı şehirlerde olsak da karşılıksız sevgi ve ilgileriyle her daim yanımda hissettiğim anneme, babama, kardeşime ve desteği ile motivasyonumu hep yüksek tutan sevgili eşim Hakan'a

sonsuz teşekkürlerimi sunarım.

İÇİNDEKİLER

	Sayfa
TEZ BİLDİRİMİ.....	iii
ÖZET	iv
SUMMARY	vi
ÖN SÖZ.....	viii
ÇİZELGELER DİZİNİ	xii
ŞEKİLLER DİZİNİ	xii
BÖLÜM 1.....	1
GİRİŞ	1
1.1. Problem	1
1.1.1. Standart Belirleme Yöntemleri	6
1.1.2. Sınır Değer Saptamaya Dayalı Yöntemler	9
1.1.3. Eğri Altında Kalan Alan (Area Under Curve-AUC).....	13
1.2. Amaç	27
1.3. Önem.....	27
1.5. Sınırlılıklar	29
1.6. Tanımlar	30
1.7. Kısaltmalar	30
BÖLÜM 2.....	31
YÖNTEM.....	31
2.1. Araştırma Modeli.....	31
2.2. Araştırma Grubu	31
2.3. Veriler ve Toplanması	33
2.3.1. Veri Toplama Araçları.....	33
2.4. Verilerin Analizi.....	36
BÖLÜM 3.....	39
BULGULAR VE YORUMLAR	39
3.1. Angoff, ROC ve Sınır Grup yöntemleri ile belirlenen kesme puanı	39
3.2. Angoff ve Sınır Grup yöntemleri ile belirlenen kesme puanlarına göre düşük/yüksek kaygılı kabul edilen öğrencilerin yüzdeleri arasındaki fark.....	44
3.3. Kesme puanlarının belirlenmesinde yargıcılar arası uyum katsayısı	47

3.4. ROC analiziyle elde edilen eğri altında kalan alan değerinin öğrencileri ayırmadaki doğruluk oranı.....	50
3.5. Sürekli Kaygı Ölçeği'nin ROC analizi ile geçerlik düzeyi.....	51
BÖLÜM 4.....	55
SONUÇ VE ÖNERİLER.....	55
4.1. Sonuçlar	55
4.2. Öneriler	57
KAYNAKLAR.....	59
EKLER.....	64
EK1. Sınır Grup Yöntemine Ait Kararlar	64
EK2. Angoff Yöntemine Ait Kararlar	65
EK 3. Sınav Kaygısı Ölçeği	66
EK 4. Sürekli Kaygı Ölçeği	68
EK5. Sınır Grup Yöntemi için Yönerge	69
EK6. Angoff Yöntemi için Yönerge	70
ÖZ GEÇMİŞ.....	71

ÇİZELGELER DİZİNİ

Çizelge	Sayfa
1. ROC Analizi için Gerekli Olan Değerler	10
2. ROC Eğrisi Altında Kalan Alanın Değerleri ve Testin Yorumlanması	13
3. Araştırma Grubunu Oluşturan Öğrencilerin Lisans Programlarına Göre Dağılımı....	31
4. Araştırma Grubunu Oluşturan Yargıcıların Üniversitelere Göre Dağılımı	32
5. Angoff ve Sınır Grup Yöntemi ile Belirlenen Sürekli Kaygı Ölçeği Kesme Puanları	40
6. Araştırma Grubu Kaygı Düzeylerinin Sınav Kaygısı Ölçeği'ne Göre Sınıflandırılması	41
7. Sürekli Kaygı Ölçeğine İlişkin ROC Analizi ile Belirlenen Kesme Puanı	42
8. Farklı Yöntemlerle Belirlenen Kesme Puanları, Düşük/Yüksek Kaygılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri	44
9. Yöntemlere Göre Farklılaşan Düşük/Yüksek Kaygılı Öğrenci Oranlarının Manidarlığı	45
10. Yargıcıların Ölçek Puanlarına İlişkin Betimsel İstatistikler ve Normallik Testleri..	47
11. Yöntemlere Göre Kesme Puanı Belirlemede Yargıcılar Arası Uyum Katsayısı Sonuçları.....	48
12. ROC Eğrisi Altında Kalan Alan Değeri	51
13. Ölçeklere Göre Düşük/Yüksek Kaygılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri	52

ŞEKİLLER DİZİNİ

Şekil	Sayfa
1. Farklı ROC Eğrilerinin Sınıflama Durumları	12
2. Sürekli Kaygı Ölçeği'ne İlişkin ROC Eğrisi Altında Kalan Alan	50

BÖLÜM 1

GİRİŞ

Bu bölümde araştırmanın problemi, amacı, önemi, sayıtları, sınırlılıkları ve tanımları yer almaktadır.

1.1. Problem

Bilimin temel ölçütlerinden biri olan aktarabilirlik, üretilen bilginin doğruluğu ve yanlışlığının test edilebilirliği için gereklidir. Bilim dalları ilgilendikleri alanlardaki ilişkileri, özellikleri anlamak ve açıklamak için gözlem yapmakta ve değişkenler arasındaki ilişkileri nesnel, standart bir biçimde tanımlamak için matematik ve istatistiğin doğasına başvururlardır. Sosyal bilimlerde gözlenebilen değişkenler genelde soyut nitelikli olup soyut değişkenler doğası gereği dolaylı olarak gözlemlenmektedir (Baykul, 2000). Bu gözlenen ilişkilerin benzerlik ve farklılıklarını tanımlamak ise ölçme ile mümkündür.

Doğrudan gözlenemeyen ancak gözlenebilen davranışlar aracılığı ile yordandabilen psikolojik değişkenlerin öznelliği ölçme araçları ile nesnelliğe taşınmaktadır. Ölçen ile ölçülen arasındaki ilişkileri düzenleyen kurallar takımında ölçme aracı gözleyen, tepkiler ise gözlenendir. Fiziksel uyarıcının yerini psikolojik uyarıcıların aldığı psikolojide ölçme kuramlarının ve tekniklerinin önü açılmış ve psikometri bilimi gelişmiştir (Guilford, 1954). Psikometri, psikolojide bireylerin kişilik özellikleri, zekâ ve yeteneklerinin çeşitli testlerle belirlenmesi ve istatistiksel analizi üzerinde duran bir yaklaşım olarak tanımlanmaktadır (Nunnally, 1970). Psikometrik yaklaşımda bireyde aranan niteliklerin gözlenmesi, bir davranış örnekleminin nesnel ve standart ölçüsü olan psikolojik testler aracılığı ile o özelliğin büyüklüğünün ne kadar olduğuna karar vermekle mümkündür (Anastasi, 1988). Bu doğrultuda bireylerden bilgi edinme yöntemleri arasında psikolojik testlerin daha kolay uygulanabilmesi, objektif olarak puanlanabilmesi araştırmalarda bu ölçme araçlarının sıklıkla tercih edilmesinin nedenlerindedir (Cronbach, 1984). Eğitim ve psikolojide bireyin ölçülen özelliğinin büyüklüğünü belirlemede kullanılan testler, kişisel, sosyal ve politik açıdan öneme sahiptir. Testler, bireyin bilişsel, duyuşsal ve psikomotor özelliklerinin bütün olarak incelenmesinde kullanılan farklı yöntemlerle (görüşme, sosyogram, otobiyografi vb.)

toplanan bilgilerin birleştirilmesiyle birey hakkında alınan kararları sağlamlaştıran psikolojik ölçme araçlarıdır. Bireyleri bir üst eğitim kurumuna yerleştirme, personel seçme, işe alma, öğrencilerin öğrenme zorluklarını belirleme, davranış bozukluklarına tanı koyma vb. kararlar testlerden elde edilen puanlarla verilmektedir (Crocker ve Algina, 1986). Ancak unutulmamalıdır ki testler hiçbir zaman kendiliğinden kişiler hakkında kararlar veren mekanizmalar olmayıp bireyi anlamak için geliştirilen ve önemli bilgiler sağlayan araçlardır (Cronbach, 1984).

Psikolojik testler, çeşitli ölçütlere göre sınıflandırılmaktadır. Cronbach (1984) psikolojik testleri, maksimum performans ve tipik tepki olmak üzere iki başlıkta incelemiştir. Maksimum performans testleri, başarı ve yetenek testlerini; tipik tepki testleri ise kişilik, ilgi, tutum testlerini kapsamaktadır. Başarı ya da yetenek testlerinin çoğunun amacı, bireyleri yetenek düzeylerine göre sınıflamak ve sıralamaktır. İlgi, tutum, kişilik gibi psikolojik ölçmelerin amacı ise bireyler arasındaki farklılıkları açığa çıkarmaktır (Anastasi, 1988). Maksimum performans testleri (1) ölçülen niteliklerin açık seçik olması, (2) testle ilgili olarak geliştirilen doğru cevap anahtarının test henüz uygulanmadan önce dış ölçütlere göre saptanabilmesi, (3) bu tür testlerde test alan kişinin testten alınabilecek maksimum puana güdülenmiş olması gibi temel özellikleri yönünden tipik tepki testlerinden ayrılmaktadır (Özguven, 2007). Maksimum performans testlerinde birey öğrendiklerini belirli koşullarda sergilemektedir. Uygulanan sınavdan bireyin, yaşlarına göre daha yüksek puan alması, kendisiyle veya başkalarıyla yarış hâlinde olduğu yüz metre yüzme durumu gibi yani bir güçlük karşılığında bireyin birim zamanda yapabildiğinin en fazlasını ve iyisini yapabilme durumu maksimum performansı ifade eder. Tipik tepkiler ise birey, diğer grup üyelerine göre değerlendirildiğinde daha iyi ya da daha kötü yargılarına varılamayan davranışların oluşturduğu kümedir (Cronbach, 1984). Tipik tepki testlerinde; dış ölçütlere göre maddenin doğru ya da yanlış cevabının bulunmadığı, bireyin maddeleri incelediğinde hangi duyuşsal özelliğinin ölçüldüğünü belirleyemediği, bireyin maksimum puanı alma gibi bir durum için güdülenmediği görülür. Yani ölçülmek istenen özellik, herhangi bir zorlanma olmaksızın bireyin içinde bulunduğu koşullarda ne yaptığı ve nasıl davrandığı ile ilgilidir. Psikolojik testlerde bireylere belirli türden tepkileri ortaya çıkarması beklenen uyarıcılar sunulmakta ve bireylerin bu uyarıcılara gösterdikleri tepkiler gözlemlenerek bunların derecesi belirlenmeye çalışılmaktadır. Ölçme araçlarının temel işlevi, ölçekte yer alan maddelere (uyarıcılara) verilen cevaplara (gösterilen tepki) göre,

bireyin ölçülen özellik bakımından psikolojik boyut üzerindeki konumunu belirlemektir (Tezbaşaran, 2008).

Özellikle eğitim, psikoloji, sosyoloji ve psikiyatri gibi insan davranışlarını konu edinen birçok sosyal bilim alanında psikolojik testlerden yararlanılmaktadır. Türkiye’de bu alanlarda çalışan araştırmacılar testleri, eğitimde ve bilimsel araştırmalarda olduğu kadar akıl ve ruh sağlığı merkezlerinde, psikiyatri kliniklerinde, danışma merkezlerinde, okullarda, endüstri ve işletmelerde de kullanmaktadır (Öner, 1997). Eğitim alanındaki bilimsel çalışmalarda kullanılan psikolojik testler, genellikle özgün olarak geliştirilen veya uyarlama yoluyla kültürümüze kazandırılan ölçme araçlarıyla elde edilen verilere dayanmaktadır. Örneğin kaygı kavramı birçok sosyal bilim alanında çalışıldığı gibi bu araştırmada da çoklu puanlanan bir ölçme aracının ölçtüğü değişken olarak belirlenmiştir. Bu bağlamda araştırmanın asıl odak noktası çoklu puanlanan bir ölçme aracı olması nedeniyle kaygı kavramından genel olarak bahsedilmiştir.

Kaygıyı açıklamaya yönelik çalışmalar incelendiğinde kaygının eğitim alanında yapılan araştırmalarda sıklıkla kullanılan değişkenlerden biri olduğu belirlenmiştir. Bireyde uyarılmışlık durumu yaratan, nesnel olmayan bir tehlikeye karşı kişinin yaşadığı, bedensel, duygusal ve zihinsel değişimlerle kendini gösteren duygu durumu kaygı olarak tanımlanmaktadır (Öner, 1983). Evrensel bir olgu olan kaygı kültür, inanç, dil, ırk, din, cinsiyet ve yaş ayırmaksızın her bireyin yaşadığı mutsuz edici bir duygudur (Yiğit, Dilmaç, Deniz ve Hamarta, 2011). Bireylerin içinde buldukları durumlara göre kaygı kavramı ve süreci farklılaşabilmektedir. Örneğin sürekli kaygı, ortamda kaygılanmak için herhangi bir durum yokken uzun süreli ve şiddetli yaşanabilmektedir. Bireyin bulunduğu ortamda kaygı yaşatabilecek bir durum var olduğunda da bu durumla orantısız bir biçimde sürekli kaygı yaşamak yine mümkündür. Sürekli kaygı, bireyin kaygı yaşantısına olan yatkınlığı sonucu kişinin içinde bulunduğu durumları genellikle stresli olarak algılama ya da stresli olarak yorumlama eğilimidir (Ocaktan, Keklik ve Çöl, 2002). Sürekli kaygı, olumlu ya da olumsuz bir tepki durumu oluşturmayacak durumların birey tarafından tehlikeli ve zarar verici olarak algılanması sonucu oluşan hoşnutsuzluk ve mutsuzluk duygusudur (Öner, 1983). Sürekli olarak kaygısı yüksek olan birey, düşük olan bireye göre stres yaratan durumu daha çok tehdit edici ve tehlikeli olarak algılamaktadır. Bu algıdaki birey daha yüksek kaygı düzeyinde olduğundan diğerlerine göre daha şiddetli tepkide bulunma eğilimindedir. Eğitim öğretim sürecinde kaygının yüksek düzeyde yaşandığı anlar sınavlardır. Bu bağlamda

özel bir kaygı çeşidi olan sınav kaygısı, bireyin değerlendirilmesi durumunda ortaya çıkan korkuyla karışık hissedilen bir tedirginlik duygusudur. Güncel bir konu olan sınav kaygısı, birçok ülkede ilkokuldan sonraki eğitim öğretim sürecinde yoğun olarak yaşanmakta ve birçok olumsuz durumu beraberinde getirmektedir (Erözkan, 2004). İncelenen duyuşsal özellik ne olursa olsun psikolojik testlerden elde edilen bulguların genellenebilirliği ve işlevselliği, kullanılan ölçme aracının güvenilirliği ve geçerliğiyle paralellik göstermektedir. Bu nedenle araştırma sürecinin her aşaması önemli olmasına karşın, doğrudan gözlenemeyen değişkenlerin ölçülmesi ve değerlendirilmesi ölçme araçlarının niteliğine bağlıdır.

Psikolojik özelliklerin tanımı geçmişten günümüze önemli değişimlere uğramadığından duyuşsal özelliklerin ölçülmesinde yeni bir ölçme aracı geliştirmek yerine güvenilirliği ve geçerliği bilinen ölçme araçları yaygın olarak kullanılmaktadır. Test sonuçlarının anlamlı ve hatasız olması için testin standardizasyonunun yapılmış, normlarının belirlenmiş, geçerlik ve güvenilirlik kanıtlarının sunulmuş olması gerekir. Uygulayıcı tarafından bu dört özelliğin dikkatle incelenmesi ve testten elde edilen puanların nasıl yorumlanacağı ve ne anlama geleceğinin bilinmesi önemlidir. Test sonucu elde edilen bireyin puanı, benzer bir grubun puanları ile karşılaştırılınca anlam kazanmakta aksi hâlde ham puanları tek başına yorumlamak ölçülen özelliğin miktarı ile ilgili bilgi sağlamamaktadır. Ölçek geliştirme veya uyarlama çalışmalarının en çok ihmal edilen kısmı, ölçekten elde edilen puanların nasıl yorumlanacağına net olarak raporlanmamasıdır (Erkuş, 2007). Ölçek kullanıcısı, ilgili ölçek puanına sahip bireyi nasıl değerlendirmeli, toplam puan elde edilebilir mi yoksa alt ölçekler ayrı ayrı mı ele alınmalı, puanların artışı değişkendeki artışa ya da azalışa mı karşılık gelmekte, toplam puanlar hangi ölçüte göre yorumlanmalı gibi sorunlar yaşamaktadır. Bireye uygulanan psikolojik test sonrasında elde edilen veriler, bir ölçüte göre yorumlanmaktadır. Dolayısıyla bir ölçme aracının uygulandığı grubun bu test sonucunda aldığı puanları bilmek, bireyin diğer grup üyeleri arasındaki yerinin ve sırasının bilinmesi açısından önem taşımaktadır. Böylelikle aynı testi alan bireylerin puanları norm grubunun yüzdelik ve standart puanları ölçüt alınarak yorumlanmaktadır. Testin el kitabında yer alan normlar, olanı veya bekleneni göstermek yerine, bir ortalamayı ve testin geliştirilme aşamasında temsil ettiği grupların tipik tepkilerini yansıtmaktadır. Tipik tepki testlerinde normu oluşturan örneklemin ortalaması kesme puanı olarak kullanılmakta ve alınan puanlar örneklemin ortalamasına ve standart sapmasına göre

yorumlanmaktadır. Ancak normların önemli özelliklerinden biri zaman ve mekân açısından göreceli olması nedeniyle mutlak ve kalıcı değerler olarak görülmemesi gerektiğidir (Öner, 1997; Özgüven, 2007; Schultz ve Schultz, 1998). Çoklu puanlanan psikolojik ölçme araçlarında en çok düşünülen yanılgılardan biri de “normların mutlaka içinde bulunulan grubun ham puanlarının standartlaştırılması” bilgisidir. Oysaki norm değişkeni ilgili özellik açısından ölçme sonuçlarını farklılaştıran değişkenlere göre belirlenmektedir. Örneğin ölçülen değişken kadın ve erkeğe göre değişmiyorsa cinsiyet normu oluşturmanın anlamı yoktur. Ya da ölçülen özellik ile ilgili yaş grupları arasında manidar fark yoksa yaş sınıflamaları için norm tablosu oluşturmanın anlamı yoktur. Bireyler hakkında alınan kararların doğruluğu hem ölçmenin hem de ölçütün doğruluğuna bağlıdır. Ölçekten elde edilen puanı anlamlandıracak olan ölçüt, tipik tepkilerin ölçümünde o ölçeğe tepki veren diğer kişilerin puanlarından oluşturulduğu gibi bu ölçüt belirli kuramsal beklentiden de oluşturulabilir (Anastasi, 1998). Bu bağlamda farklı yöntemlerle oluşturulan kesme puanları kullanılarak ölçek puanlarının yorumlanmasına çözüm bulunulabileceği düşünülmektedir.

Türkiye’de uygulanan tipik tepki testleri bireyleri yönlendirme, seçme ve yerleştirmede belirleyici rol oynamaktadır. Testi alanların tepki düzeylerine göre sınıflandırılması ve yapılan sınıflandırmaların doğruluğunun sınanmaması, çoğu zaman tartışma konusu olmaktadır. Bireylerin sınıflandırılma durumları, yapılan ölçme ve değerlendirmenin türüne göre farklılaşabilmektedir. Değerlendirme yapılırken iki farklı ölçüt türü esas alınmaktadır. Bireyin test puanı, testi alan diğer bireyler içindeki yeri ile karşılaştırılıyorsa bu tür değerlendirmeler bağıl değerlendirme, bireyin test puanı önceden belirlenmiş bir ölçüt ile karşılaştırılıyorsa bu tür değerlendirmeler mutlak değerlendirme olarak adlandırılmaktadır (Tekin, 2000). Mutlak değerlendirmede bireyin gruptaki diğer bireylere göre nasıl bir tepki gösterdiği önemli değildir, önemli olan bireyin ne yapabildiğidir. Dolayısıyla ölçüt türünün farklılaşması alınan kararları da değiştirebilmektedir.

Eğitim alanında uygulanan psikolojik testlerin amacına hizmet edebilmesi için uygulayıcılardan, test hakkında yeterli bilgiye sahip olmaları ve testten elde edilen puanları doğru yorumlamaları beklenmektedir. Buna rağmen hem uygulayıcıların hem de test alan bireylerin sonuçlara şüphe ile baktıkları örnekler de çoktur. Birbirlerine göre değerlendirilen bireyler, test olurken birçok değişkenden etkilenebilmektedir. Dolayısıyla her bir bireyin puanı testin aritmetik ortalamasına yani ölçüt puanının

yüksek veya düşük olmasına yol açmaktadır. Tipik tepki testlerinden elde edilen puanlar, bireyin kuramsal tanımla ya da araştırmacının ortaya koyduğu ölçütlerle belirlenen özellik grubuna dâhil olup olmadığını ortaya çıkarmayı amaçlar (Cronbach, 1984). Bireyin ölçme sonunda yapılan değerlendirmede belirli bir gruba dâhil olduğu görülürse bu bireyin, o grubun genel özelliklerine sahip olduğu düşünülmektedir. Bağlı değerlendirmeyle bireyin ilgili testteki tepki düzeyini belirlemede yapılan hatalar, bireyi yeterli iken yetersiz olma durumu ile karşı karşıya getirebilmektedir. Hatalı ölçme sonuçları ve değerlendirmeler bireyi yanlış yönlendirebilmekte, bir sonraki aşamada alınacak kararları olumsuz etkileyebilmektedir. Bu bağlamda elde edilen puanları değerlendirebilmek için ölçütün nasıl oluşturulduğunun ve bireylerin tepki düzeylerinin neye göre farklılaştığının bilinmesi önemlidir. Bu sebeple bireyin çoklu puanlanan psikolojik testlerdeki tepkilerini değerlendirebilmek ve düzeylerini sınıflandırabilmek için kesme puanı belirleme yöntemlerinin incelenmesine gerek duyulmuştur.

Kesme puanı, bireyleri ölçülen özelliğin düzeylerine göre gruplara ayıran ve bu düzeylerin ne anlama geldiğini tanımlamada yararlanan eşik noktasıdır (Crocker ve Algina, 1986; Irwin, 2007). Kesme puanı belirleme yöntemleri günümüzde eğitim, tıp, sağlık, veterinerlik gibi alanlarda kullanılmakta ve farklı kavramlarla ifade edilmektedir. Kesme puanı, standart belirleme ve sınır değer saptama yöntemleriyle belirleme olmak üzere iki başlıkta incelenmiştir. Standart belirleme teorisyenleri, standart belirlemenin bilinebilir ya da kestirilebilir parametre keşfetme yöntemi olarak tanımlanmasını ve kavramsallaştırılmasını reddetmişlerdir (Kane, 1994). Çünkü standart belirleme, süregelen özelliği özgün kategorilere ayıran, önceden var olan kesme puanları bulmayı amaçlamaz. Her testin kendine özgü standartları bulunur ve bu standartlar zaman içerisinde değişikliğe uğrayabilir. Standart belirleme sürecinde yargıcılar, toplumun politik ve kültürel yapısı ve ülke kültüründeki ilgili psikolojik yapının varlığı gibi birçok değişken etkilidir. Dolayısıyla kesme puanı belirlenirken yargıcıların ve psikolojik yapının doğrudan ya da dolaylı gözlenebilen değişkenlerle ilişkisinin bilinmesi standart belirleme ve sınır değer saptama yöntemlerini birbirinden ayırmıştır.

1.1.1. Standart Belirleme Yöntemleri

Psikometride kullanılan standart belirleme kavramı, çeşitli uzmanlar tarafından tanımlanmıştır. Cizek (1993) tepkilerin iki ya da daha fazla derecesini ayırt eden puanı belirleyebilmek için uygun kurallar ve işlemleri takip etme süreci olarak tanımladığı

standart belirlemenin, psikometrinin istatistiksel, politik ve kültürel yönlerini bir araya getiren bir yöntem olduğunu vurgulamıştır. Hambleton (2001) standart belirlemeyi bireylerin test tepkilerini belirlenen kesme puanına göre belli düzeylere ayırma ve bu düzeylerin ne anlama geldiğini işevuruk hâle getirme olarak tanımlamaktadır.

Standart belirleme sürecinde kullanılacak birçok yöntem bulunmaktadır. Jeager (1989), temelinde test söz konusu olan yöntemleri “test merkezli”; öğrenci söz konusu olanları ise “öğrenci merkezli” olarak adlandırmıştır. İki sınıflama arasındaki ayrım, yargıcının (uzmanın) karar verirken neye odaklandığı ile ilgilidir. Yargıcılar ölçülen özellikle ilgili bilgiye sahip, alanında uzmanlaşmış, kesme puanının belirlenmesinde görev alan değerlendiricilerdir. Öğrenci merkezli yöntemde yargıcı karar alırken testi alan bireyleri ve bu bireylerin ölçülen özellikle ilgili minimum tepki düzeylerini göz önünde bulundurmaktadır. Test merkezli yöntemde ise yargıcı, testin içeriğine göre karar almaktadır. Bu araştırmanın amacı doğrultusunda çoklu puanlanan tipik tepki testlerine uygun standart belirleme yöntemlerinin birbirlerine göre avantajlı ve dezavantajlı yönleri tartışılmış ve hangilerinin kullanılacağına karar verilmiştir.

Test merkezli yöntemler yargıcıların test maddeleri ile ilgili kararlarına dayanmaktadır. Yargıcıların ölçülen duyuşsal özellikle ilgili bilgi sahibi olması, herhangi bir nesne, durum ya da olayın bireyde uyandırdığı ve bireyi kaçma veya kaçınma davranışına yönelten duyguların farklılığını bilmesi, farklı tepki düzeylerine ait grupları tanınması önemlidir.

Livingston ve Zieky (1982) test merkezli yöntemin beş temel adımdan oluştuğunu ifade etmektedir:

1. Yargıcıların belirlenmesi,
2. Her tepki düzeyi ile ilgili duyuşsal özelliğin derecesinin tanımlanması,
3. Yargıcıları seçilen standart belirleme yöntemi için eğitilmesi,
4. Verilerin toplanması,
5. Kesme puan elde edebilmek için verilerin analizinin yapılması.

Test merkezli yöntemlerde yargıcılar, testteki maddeleri gözden geçirerek belirli bir tepki düzeyi yani duyuşsal özelliğin sınıflama ayrımındaki minimum düzey için kesme puanı belirlemeye çalışmaktadır. Belirlenecek olan kesme puanı yargıcıların her bir madde için ölçülen özellikle ilgili sınır düzeydeki kişilerin tepki derecesinin yüzdesine bağlı olarak hesaplanmaktadır. Bu araştırmada test merkezli yöntemlerden Angoff yöntemi tercih edilmiştir.

1.1.1.1. Angoff Yöntemi

Yöntemin kolay uygulanabilir ve yargıcılarca anlaşılabilir olması kesme puanı belirlemede sık kullanılmaya nedenlerindedir (Berk, 1986; Mills ve Melican, 1988). Yargıcılar test maddelerine odaklanırken yüksek kaygı grubundaki minimum tepki düzeyine sahip kişiyi akıllarında tutmalıdırlar. Her bir yargıcı, ölçülen özellikle ilgili eşik düzeyindeki bireylerin maddelere ait hangi tepki dereceleri ile ayrılabileceğini düşünerek maddeleri cevaplamaktadır. Eşik düzeyindeki kişinin her bir maddeden aldığı puanların toplamı da kişinin testten alacağı puanı vermektedir (Cizek ve Bunch, 2007). Her yargıcının testte bulunan bütün maddelere ilişkin sonuçları toplanarak o yargıcıya ait minimum test puanı hesaplanmaktadır. Yargıcıların minimum test puanları ortalaması bu yöntem için nihai kesme puanını vermektedir.

Öğrenci merkezli yöntemlerde yargıcılar ele alınan konu ile ilgili olarak minimum düzeydeki öğrencileri temel alarak yani öğrenci gibi düşünerek değerlendirme yapmaktadır. Öğrenci merkezli değerlendirme yönteminde yargıcılar, testteki maddeler hakkında değil de testi alan bireyler hakkında bir yargıya varırlar. Bu sebeple yargıcıların testi alan bireylerin ölçülen duyuşsal özellikleri hakkında bilgi sahibi olmaları gereklidir (Livingston ve Zieky, 1982). Düşük ve yüksek kaygı düzeylerine sahip bireylerin tanınması, nasıl tepkiler verebileceklerinin bilinmesi yargıcıların cevaplarının güvenilirliği açısından önemlidir. Bu araştırmada öğrenci merkezli yöntemlerden Sınır Grup yöntemi tercih edilmiştir.

1.1.1.2. Sınır Grup Yöntemi

Yargıcılardan, bireyleri ölçülen özelliğe göre düşük ve yüksek tepki derecesine sahip şekilde tanımlamaları istenmektedir. İki sınıflandırmada arada kalan yani ne düşük ne de yüksek tepki derecesi verebilecek bireylerin yer aldığı yüksek kaygının minimum düzeyindeki bireyler, sınır grup olarak kabul edilmektedir. Yargıcıların ölçek puanlarının ortancası belirlenerek bu yöntem için kesme puanı elde edilir. Ortanca değer uç puan değerlerinden etkilenmediğinden Sınır Grup yöntemi için önemlidir. Çünkü sınır grup oluşturulurken ölçekten yüksek ve düşük puan alanlar bu grup içerisinde yer almamalıdır (Livingston ve Zieky, 1982).

Bu araştırmada ayrıca üçüncü bir yöntem olarak daha çok tıp, veterinerlik, sağlık bilimleri gibi alanlarda sınır değer oluşturulurken yararlanılan İşlem Karakteristik Eğrisi (Receiver Operating Characteristics Curve-ROC) tercih edilmiştir. Standart belirleme

yöntemlerinde kesme puanları, yargıcı puanlarının ortalaması ve ortancası hesaplanarak belirlenmektedir. Sınır değer saptamaya dayalı yöntemlerde ise düşük ve yüksek kaygılı bireylerden oluşan iki grup arasında ayırım yapmak için her olası puan kesme puanı olarak değerlendirilmektedir. ROC eğrisi ölçme araçlarının sınıflama doğruluklarını sayısal sonuçlarla açıklanması nedeniyle diğer yöntemlerden ayrılmaktadır. Son yıllarda eğitim alanında da kullanılmaya başlanan ve araştırmada yararlanılan sınır değer saptama yöntemlerinden ROC aşağıda açıklanmıştır.

1.1.2. Sınır Değer Saptamaya Dayalı Yöntemler

Sınır değer saptamaya dayalı yöntemlerde kesme puanı olarak hangi değer uygun olduğu, istatistiksel karar teorisi ile belirlenir. Ölçülen özellikle ilgili bireyleri tepki derecelerine göre sınıflandırmak için bazı istatistiksel ölçütler hesaplanır. Dirican (2001) test sonuçlarının değerlendirilmesinde genellikle kullanılan ölçütlerin; pozitif (doğru pozitif, yanlış pozitif) ve negatif (doğru negatif ve yanlış negatif) oran değerlerinin olduğunu belirtmektedir. Ölçülen özelliğin gerçekteki varlığı göz önünde bulundurulduğunda ölçüt testten ve kesme puanı belirlenen testten elde edilebilecek sonuçlar çapraz (dörtlü) tablolar oluşturularak dört grupta toplanmaktadır. Shapiro (1999) sınır değer saptama yöntemlerini işlem karakteristik eğrisi (ROC), aralık tahmini ($\pm 1,96s$) ve yüzde değerlerine göre standart belirleme olmak üzere üç gruba ayırmıştır. Bu araştırmada kullanılan sınır değer saptama yöntemlerinden İşlem Karakteristik Eğrisi (ROC) aşağıda açıklanmıştır.

1.1.2.1. İşlem Karakteristik Eğrisi (Receiver Operating Characteristics Curve - ROC)

ROC analizi; pozitif-negatif oran değerlerine göre hesaplanan duyarlılık ve özgüllük değerlerini kullanarak bir testin ayırt etme gücünün belirlenmesine, aynı özelliğe yönelik farklı testlerle teşhis gücünün kıyaslanmasına olanak sağlamaktadır (Dirican, 2001). ROC analizi, farklı klinik durumlarda uygulanan tanı testlerinin performansını belirlemek ve lojistik modeller, doğrusal sınıflama analizi gibi istatistiksel modellerin doğruluğunu değerlendirmek amacı ile kullanılan bir yöntemdir (Obuchowski, 2005). ROC eğrisi yardımıyla yapılan ROC analizinde, uygun kesme puanının belirlenmesi için bir referansa ihtiyaç duyulmaktadır. Örneğin sağlık alanında referans; gerçek durumu en iyi yansıtan, hasta ve sağlıklı durumunu birbirinden kesin olarak ayıran test olarak kabul edilir. Gerçekte hasta olan bir kişiye %100 doğrulukla hasta, gerçekte sağlıklı olan bir kişiye de %100 doğrulukla sağlıklı diyebilen bir testtir.

Bu testler genelde riskli, pahalı veya kullanılması zor yöntemlerdir. Örneğin otopsi sonucu, biyopsi sonucu vs. referans olarak kabul edilen yöntemlerdir (Karakaya, 2012). Ancak her alanda sağlık alanında olduğu gibi referanslar bulmak mümkün değildir. İlgili alanyazın incelendiğinde bu referansın bireylere önceden uygulanan aynı durumu ölçen bir ölçüt ölçek, test veya durum hakkındaki uzman görüşleri olabileceği görülmektedir (Koyuncu, 2015; Uluç, 2007; Taşdemir, 2013). Bu bağlamda araştırmada Spielberger (1980) tarafından geliştirilen, Öner ve Le Compte (1983) tarafından Türkçeye uyarlanan Sınav Kaygısı Ölçeği ölçüt olarak kabul edilmiştir. ROC analizi için gerekli olan duyarlılık, özgüllük, pozitif-negatif oran ve pozitif-negatif yordayıcı değerlerin hesaplamaları ve açıklamaları Çizelge 1’den yararlanılarak sunulmuştur (Akçay ve Demirel, 2011; Çamlıca ve Dişçi, 2008).

Çizelge 1.

ROC Analizi için Gerekli Olan Değerler

		Özelliğın Gerçekteki Varlığı (Ölçüt Test Sonucu)		
		Yüksek (Pozitif)	Düşük (Negatif)	Toplam
Test Sonucu	Yüksek (Pozitif)	A(DP)	B(YP)	A+B
	Düşük (Negatif)	C(YN)	D(DN)	C+D
		A+C	B+D	A+B+C+D

A (Doğru Pozitif, YP): Sürekli Kaygı Ölçeği puanlarına göre yüksek kaygılı olup gerçekte Sınav Kaygısı Ölçeği puanlarına göre de yüksek kaygılı olan öğrencilerin sayısı

B (Yanlış Pozitif, YP): Sürekli Kaygı Ölçeği puanlarına göre yüksek kaygılı olup gerçekte Sınav Kaygısı Ölçeği puanlarına göre düşük kaygılı olan öğrencilerin sayısı

C (Yanlış Negatif, YN): Sürekli Kaygı Ölçeği puanlarına göre düşük kaygılı olup gerçekte Sınav Kaygısı Ölçeği puanlarına göre yüksek kaygılı olan öğrencilerin sayısı

D (Doğru Negatif, DN): Sürekli Kaygı Ölçeği puanlarına göre düşük kaygılı olup gerçekte Sınav Kaygısı Ölçeği puanlarına göre de düşük kaygılı olan öğrencilerin sayısı

Çizelge 1’de sunulan değerler ROC analizi için gerekli olup yüksek ve düşük kaygılı bireylerin sınıflandırılması yapılırken karşılaşılabilecek durumları göstermektedir. Bireylerin kaygı durumları ölçüt ölçek ve kesme puanı belirlenen ölçeğe göre sınıflandırılmış, ölçeklere göre yüksek ve düşük kaygılı kabul edilen birey sayıları hesaplanmıştır. Bu değerler ayrıca işlem karakteristik eğrisi oluşturulurken

gerekli olan duyarlılık ve özgülük oranlarının belirlenmesinde kullanılan koordinatlardır.

Duyarlılık (Sensitivity)

Doğru Pozitif Orana eşittir. Gerçek durumun pozitif olduğu durumda test sonucunun pozitif olma olasılığıdır. Yani Sınav Kaygısı Ölçeği'nin yüksek kaygılı olarak belirlediği bireylerin Sürekli Kaygı Ölçeği ile hangi oranda saptanabildiğini belirten orandır. Sürekli Kaygı Ölçeği'nin yüksek kaygılı bireyleri belirlemede ne kadar duyarlı olduğunu gösterir. Duyarlılık değeri Eşitlik 1'deki gibi hesaplanır.

$$\text{Duyarlılık} = A / (A + C) = DP / (DP + YN)$$

(1)

Özgüllük (Specificity)

Doğru Negatif Orana eşittir. Gerçek durumun negatif olduğu durumda test sonucunun negatif olma olasılığıdır. Yani Sınav Kaygısı Ölçeği'nin düşük kaygılı olarak belirlediği bireylerin Sürekli Kaygı Ölçeği ile hangi oranda saptanabildiğini belirten orandır. Sürekli Kaygı Ölçeği'nin düşük kaygılı bireyleri ayırabilme gücünü belirtir. Özgüllük değeri Eşitlik 2'deki gibi hesaplanır.

$$\text{Özgüllük} = D / (B + D) = DN / (DN + YN)$$

(2)

Pozitif Yordayıcı Değer (PYD)

Yüksek kaygılı olup testten pozitif sonuç alan bireylerin oranıdır. Pozitif yordayıcı değer Eşitlik 3'teki gibi hesaplanır.

$$\text{PYD} = A / (A + B) = DP / (DP + YP)$$

(3)

Negatif Yordayıcı Değer (NYD)

Düşük kaygılı olup testten negatif sonuç alan bireylerin oranıdır. Negatif yordayıcı değer Eşitlik 4'teki gibi hesaplanır.

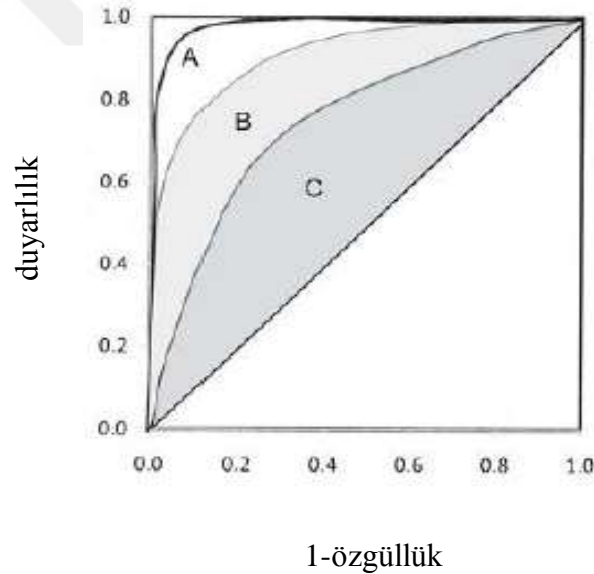
$$\text{NYD} = D / (C + D) = DN / (DN + YN)$$

(4)

Araştırılan testten elde edilen sonuçlar referans test sonuçlarıyla karşılaştırıldığında araştırılan testin pozitif ve negatif olarak sınıfladığı bireylerin sayısının gerçek durumlarının sayısı ile aynı olmadığı görülür. Bu durumda yanlış

negatiflik (1-duyarlılık) ve yanlış pozitiflik (1-özgüllük) değerlerinin hesaplanması söz konusudur.

ROC eğrisinin Y ekseninde duyarlılık ve X ekseninde 1-özgüllük değerleri yer almaktadır. Araştırılan testten ölçülen her bir değer kesim noktası olarak belirlenerek “duyarlılık”, “1-özgüllük” değerleri bu koordinat ekseninde gösterilir ve grafik üzerinde işaretlenir. Bu noktaların birleştirilmesiyle içbükey olan ve tüm eşik değerleri için araştırılan testin gerçek özelliklerini ortaya koyan ROC eğrisi elde edilir. Testin ayırt etme gücü ROC eğrisi üzerinden açıklanabilmektedir. Araştırılan testin ayırt etme gücü ne kadar iyi ise eğri o kadar yukarıya (yüksek duyarlılık bölgesi) doğru kayar. Eğri $y=x$ fonksiyonuna yaklaştıkça testin ayırt etme gücü azalır ve yanlış değerler oranının yüksek olduğu şeklinde yorumlanır (Eröz, 2010; Obuchowski, 2005; Sasse, 2002; Uluç, 2007). Üç farklı testten elde edilen ROC eğrileri Şekil 1’de gösterilmiştir.



Şekil 1. Farklı ROC Eğrilerinin Sınıflama Durumları

Şekil 1’de (0, 0) noktasından (1, 1) noktasına doğru 45° açı yaparak çizilen köşegen, testin “şans çizgisi” olarak adlandırılır. Yanlış pozitif oran ile doğru pozitif oran değerleri birbirine eşit olan testin ROC eğrisi şans çizgisi şeklindedir. Bu durumda testin yüksek kaygılı bir kişiyi doğru bir şekilde pozitif olarak belirleme şansı yanlışlıkla negatif olarak belirleme şansına eşittir. İdeal bir ROC eğrisi ise (0, 0), (0, 1) ve (1, 1) noktalarını birleştirmektedir. Dolayısıyla ROC eğrisi, (0, 1) noktasına ne kadar yaklaşırsa testin ayırt edicilik gücü de o kadar yüksektir. Şekil 1’de gösterilen A eğrisi, B ve C’den üstündür. Çünkü A eğrisi, hem duyarlılığın hem de özgüllüğün %100

olduğu noktaya daha yakındır. Testler karşılaştırıldığında ise A eğrisinin elde edildiği testin, diğer testlere göre daha mükemmel sınıflama yapabildiği söylenir. ROC eğrisi üzerindeki her nokta, farklı bir duyarlılık ve 1-özgüllük değerlerinden oluşur. Araştırılan testin kesme puanı olarak hangi değer alınacağına karar verilirken duyarlılığın yüksek, 1-özgüllük değerinin düşük olduğu en iyi nokta belirlenir.

ROC eğrisi, araştırılan testin kendi doğruluğunu tanımlaması ve testler arasında güvenilir bir karşılaştırma yapmaya olanak sağlaması açısından sıklıkla kullanılmaktadır. ROC eğrisine ilişkin değerlendirmeler, eğri aracılığıyla hesaplanan AUC (Eğri Altında Kalan Alan), YI (Youden İndeksi) ve MDU (Maksimum Dikey Uzaklık) gibi değerler üzerinden yapılmaktadır (Köksal, 2011). Bu çalışmada ROC eğrisi değerlendirilirken incelenen, eğri altında kalan alan (AUC) aşağıda anlatılmıştır.

1.1.3. Eğri Altında Kalan Alan (Area Under Curve-AUC)

ROC eğrisi altında kalan alan ne kadar büyükse araştırılan test, ilgili özelliği belirlemede o kadar başarılıdır. Alan gerçekte ölçülen özellik bakımından pozitif olan gruptan seçkisiz olarak alınacak bir katılımcının test sonucuna bağlı olarak pozitif sınıflandırılma olasılığını göstermektedir (Uluç, 2007). AUC'nin olası değerleri 0.5 ile 1.0 arasındadır. Grove (2006), testin hiç ayırma gücü olmadığı durumlarda alanın beklenen değerinin 0.5 olduğunu, mükemmel bir testte ise sıfır yanlış pozitif ve sıfır yanlış negatif oranlarla alanın değerinin 1.0 olduğunu belirtmiştir. Hosmer, Lemeshow ve Sturdivant (2013), AUC değerinin ne kadar büyük olursa o kadar iyi bir ayırmadan bahsetmekle birlikte genel olarak Çizelge 2'de sunulan yorumların yapılabileceğini belirtmişlerdir.

Çizelge 2.

ROC Eğrisi Altında Kalan Alanın Değerleri ve Testin Yorumlanması

AUC Değerleri	Testle İlgili Yapılabilecek Yorumlar
AUC=0.5	Bir ayırım belirtmez, bu durum yazı tura atmaktan farksızdır.
0.5<AUC<0.7	“Zayıf” bir ayırım belirtir ancak yazı tura atmaktan daha iyidir.
0.7≤AUC<0.8	“Kabul edilebilir” bir ayırım belirtir.
0.8≤AUC<0.9	“Çok iyi” bir ayırım belirtir.
AUC≥0.9	“Mükemmel” bir ayırım belirtir.

Kaynak. Hosmer, Lemeshow ve Sturdivant, 2013.

Çizelge 2’de belirtildiği gibi eğri altında kalan alan değerine göre testle ilgili beş farklı düzeyde yorum yapılabilmektedir. Eğri altında kalan alan değerinin 1 olması, seçilen kesim noktası ile gerçek durumun %100 uyumlu olduğu anlamına gelmektedir. Bu durum testin kesme puanına göre olağanüstü bir ayrımın yapılabildiğini göstermektedir. ROC eğrisi altında kalan alanın hesaplanması için verilerin normal dağılım gösterip göstermediğine göre parametrik ve parametrik olmayan yöntemler kullanılmaktadır.

1.1.3.1. Parametrik Yöntem:

Gerçek durum ile test sonucu elde edilen duruma ait verilerin normal dağılımın karışımından oluştuğu varsayımına dayanır (Metz ve Pan, 1999). Parametrik yaklaşımda sürekli yapıdaki değişkenlerin iki değişkenli normallik varsayımı vardır. Binormal (iki biçimli) dağılımda puanlar pozitif ve negatif grupta normal dağılır. Parametrik yöneme göre ROC eğrisi Eşitlik 5’teki gibi hesaplanmaktadır (Krzanowski ve Hand, 2009):

$$ROC(t) = \Phi(a + b\Phi^{-1}(t))$$

t: yanlış pozitif oran

a: fonksiyon sabiti

b: eğrinin eğimi

Φ : kümülatif standart normal dağılım fonksiyonu

(5)

Eşitlik 5’teki parametrelere bağlı olarak her yanlış pozitif oran “t”ye karşılık, doğru pozitif oranı gösteren ROC(t) kestirilir. Test sonucu ve gerçekte olan durumun dağılımlarının ortalama ve standart sapmaları sırasıyla μ_1, μ_0 ve σ_1, σ_2 şeklinde gösterilir. Fonksiyonun sabitini ifade eden “a” değeri Eşitlik 6’daki gibi belirlenmektedir (Hanley ve McNeil, 1983).

$$a = \frac{\mu_1 - \mu_0}{\sigma_1}$$

(6)

ROC eğrisinin eğimini ifade eden “b” değeri ise Eşitlik 7 yardımıyla bulunmaktadır (Hanley ve McNeil, 1983).

$$b = \frac{\sigma_0}{\sigma_1}$$

(7)

ROC eğrisi altında kalan alan (AUC) formülü ise kümülatif standart normal dağılım fonksiyonu (Φ), fonksiyon sabiti (a) ve eğrinin eğiminden (b) yararlanılarak Eşitlik 8'deki gibi hesaplanmaktadır (Zhang, 2005).

$$AUC = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right) \quad (8)$$

ROC eğrisi Eşitlik 8 yardımıyla doğru pozitif orana karşılık gelen yanlış pozitif orana göre çizilebileceği gibi iki doğru sınıflama oranına göre de yani doğru pozitif orana karşılık doğru negatif oran değerlerine göre de çizilebilmektedir. Elde edilen eğriler aynı test için çizildiğinden simetrik olup eğri altında kalan alan değerinde bir farklılık olmamaktadır. Eğri altında kalan alan, düşük ve yüksek kaygılı gruptan seçkisiz olarak seçilen bireylerin ölçek sonuçlarının doğru sınıflanması olasılığını verir. Bu olasılık yüksek ve düşük kaygılı bireylerin ölçek sonuçlarının normal dağılım göstermesi koşuluyla geçerlidir.

1.1.3.2. Parametrik Olmayan Yöntem

Veri yapısı normal dağılım göstermediğinde ROC eğrisi altında kalan alan, eğriyi oluşturan noktaların birleştirilmesi sonucu oluşan geometrik şekillerin alanları toplanarak elde edilir (DeLong, DeLong ve Clarke-Pearson, 1988). ROC eğrisi altında kalan alan aslında bir olasılık ölçüsüdür. Bu istatistik doğru sıralama olasılığını verir. Bu alanın olasılıkla ifade edilen değeri “ θ ” simgesi ile gösterilir ve Eşitlik 9'daki gibi tanımlanır.

$$\theta = \frac{1}{n_{H+} \times n_{H-}} \sum_{i=1}^{n_{H+}} \sum_{j=1}^{n_{H-}} I(Y_{H+}, Y_{H-})$$

n_{H+} :yüksek kaygılı bireylerin sayısı

n_{H-} :düşük kaygılı bireylerin sayısı

(9)

Eşitlik 9'daki “I” tanımlanmış bir gösterge fonksiyonudur. Gösterge fonksiyonu; yüksek kaygılı bireylerin sayısı düşük kaygılı bireylerin sayısından fazla olduğunda 1, eşit olduğunda 0.5, düşük olduğunda ise 0 değerini almaktadır. Düşük ve yüksek kaygılı gruba ilişkin ölçek sonuçları bu kurala göre ($n_{H+} \times n_{H-}$) sayıdaki olası tüm karşılaştırmalar yapılarak gösterge fonksiyonu sonuçlarının ortalaması alınır.

Parametrik olmayan yöntemlerde Eşitlik 9 yardımıyla hesaplanan θ değeri, seçkisiz örnekleme yöntemi ile seçilen pozitif ve negatif özelliğe sahip katılımcı eşleştirildiğinde ($n_{H+} \times n_{H-}$) iki katılımcının doğru sınıflandırılmasının olasılığını vermektedir (Pepe, 2003). Yani ROC eğrisi altındaki gerçek alan q 'ya eşittir ve $q=P(XP>XN)$ şeklinde gösterilir. Rastgele seçilen pozitif ve negatif olan katılımcının doğru olarak sıralanma olasılığı Wilcoxon istatistiği ile belirlenmektedir (Hanley ve McNeil, 1983).

Eğitim bilimleri alanında kesme puanı belirleme yöntemleriyle ilgili yapılan araştırmalar incelendiğinde çoğunlukla başarı testlerinde birkaç yöntemin karşılaştırıldığı, avantajlı ve dezavantajlı yönlerinin belirlendiği görülmüştür (Çetin, 2011; Gündeğer, 2012; Impara ve Plake, 1998; Ömür ve Selvi, 2010; Stone, Koskey ve Sondergeld, 2011; Tanrıverdi, 2006; Taşdelen, 2009; Tülübaş, 2009). Ayrıca eğitim bilimleri alanında kullanılan testlerin kesme puanı belirlenirken ROC analizinin yapıldığı çalışmalar da mevcuttur (Koyuncu, 2015; Taşdemir, 2013; Vivo ve Franco, 2008). Alanyazın incelendiğinde kesme puanı belirleme yöntemlerinden hiçbiri için “en iyi yöntem, en doğru sınıflama yapabilen yöntem” şeklinde bir tanımlama yapılmamıştır. Irwin (2007), farklı kesme puanı belirleme yöntemleri aynı sınava aynı yargıcılar tarafından uygulandığında, kesme puanlarının değiştiğini belirtmektedir. Çetin ve Gelbal (2010) araştırmalarında standart belirleme yöntemlerinden öğrenci ve test merkezli yöntemlerden hangisinin kullanılacağına karar verilirken bireylerin, kullanılacak testin yapısının ve yargıcıların önemli olduğunu belirtmişlerdir. Cizek (2001), birden fazla kesme puanı belirleme yönteminin birlikte kullanılmasını önermekte ve sonuçların benzer çıkmasının elde edilen puanı güçlendireceğini vurgulamaktadır. Test merkezli standart belirleme yöntemlerinden Angoff ve öğrenci merkezli yöntemlerden Sınır Grup, yargıcılarca kolay anlaşılabilir olması, uygulamadaki kolaylığı da göz önünde bulundurulduğundan araştırmada tercih edilmiştir. ROC analizi ise kesme puanı belirlenirken duyarlılık ve özgüllük değerlerinin testin geçerliğine katkıda bulunması ve bireylerin sınıflama doğruluğuna ait değerler sunması nedeniyle tercih edilmiştir. Bu doğrultuda araştırmanın amacıyla ilgili Türkiye’de ve yurt dışında yapılmış standart belirleme ve sınır değer saptama yöntemleri kullanılarak yapılan kesme puanı belirleme araştırmaları aşağıda sunulmuştur.

Tanrıverdi (2006), Angoff, Angoff (1-0) ve Nedelsky standart belirleme yöntemleriyle oluşturulan ders geçme puanlarının sayısal ve sözel derslerde farklılık gösterip göstermediğini incelemiştir. Çoktan seçmeli testlerin kullanıldığı araştırmada, her ders için farklı standart belirleme yöntemleriyle oluşturulan geçme puanı üstünde puan alan öğrenci yüzdeleri ikiyeşerli olarak karşılaştırılmıştır. Türkçe dersi için Angoff ve Angof (1-0) yöntemlerine göre başarılı öğrenci yüzdeleri arasında .05 düzeyinde, diğer yöntemlere göre öğrenci yüzdeleri arasında .01 düzeyinde manidar fark olduğu sonucuna ulaşılmıştır. Matematik dersi için ikili karşılaştırmaların hepsinde .01 düzeyinde manidar fark olduğu sonucuna ulaşılmıştır. Ayrıca kesme puanı belirleme yöntemleri arasındaki uyum incelendiğinde Türkçe dersi için Angoff ve Angoff (1-0) yöntemleri arasında pozitif ve yüksek düzeyde uyumun olduğu sonucuna ulaşılmıştır. Matematik dersi için yine Angoff ve Angoff (1-0) yöntemleri arasında pozitif ve orta düzeyde uyum olduğu belirlenmiştir. Araştırmada, ölçülen özelliğin ve testin yapısına bağlı olarak öğrenciler hakkında verilecek geçti/kaldı kararları için standart belirleme yöntemlerinden hangisinin kullanılacağına değışebileceği sonucuna ulaşılmıştır.

Taşdelen (2009), Angoff ve Nedelsky standart belirleme yöntemleri ile oluşturulan kesme puanlarını Genellenabilirlik Kuramı'nı kullanarak karşılaştırmıştır. Çoktan seçmeli testlerin kullanıldığı araştırmada, iki yöntemle belirlenen kesme puanlarına göre başarılı sayılan öğrenci yüzdeleri arasında .05 manidarlık düzeyinde farklılık olduğu ve Nedelsky yöntemine göre başarılı kabul edilen öğrenci yüzdesinin Angoff yöntemine göre hesaplanan öğrenci yüzdesine oranla daha büyük olduğu sonucuna ulaşılmıştır. Ayrıca Genellenebilirlik Kuramı'na göre Nedelsky ve Angoff yöntemleri ile kesme puanı oluşturulurken 10 puanlayıcının yeterli olacağı belirlenmiştir. Nedelsky yöntemi ile yapılan puanlamada maddeler arası farklılığın ortaya çıkabildiği ancak kesme puanı belirlemede puanlayıcılar arasında tutarlılığın olmadığı görülmüştür. Angoff yöntemi ile yapılan puanlamada ise maddeler arası farklılığın ortaya çıkabildiği ve puanlayıcılar arasında tutarlılık olduğu görülmüştür.

Tülübaş (2009), çoklu puanlanan psikolojik testlerde kesme puanının belirlenmesinde Angoff ve Sınır Grup standart belirleme yöntemlerinin uygunluğunu incelemiştir. Yöntemlere göre elde edilen kesme puanları, Beck Depresyon ve Beck Umutsuzluk Ölçeklerinin normları ile karşılaştırılmıştır. Araştırmada 11 yargıcı ve 333 üniversite öğrencisi ile çalışılmıştır. Beck Depresyon Ölçeği için Angoff ve Sınır Grup yöntemi ile elde edilen kesme puanlarının normlardan .05 düzeyinde manidar olarak

farklı olduğu görülmüştür. Sınır Grup yöntemi ile elde edilen kesme puanının ölçeğin normlarından elde edilen kesme puanına daha yakın olduğu belirlenmiştir. Beck Umutsuzluk Ölçeği için ise hem Angoff hem de Sınır Grup yöntemiyle elde edilen kesme puanlarının aynı olduğu sonucuna ulaşılmıştır. Bu değer de ölçeğin normlarına yakın olduğu bulgularına ulaşılmıştır. Bu bağlamda standart belirleme yöntemlerinin norm oluşturma süreci içerisinde kullanılabilmesi, özellikle örneklem oluşturmada sıkıntı oluşturabilecek durumlarda bu yöntemlerden yararlanılabileceği sonucuna ulaşılmıştır.

Eröz (2010), testlerden elde edilen veri yapısına bağlı olarak ROC eğrisi altında kalan alana ilişkin istatistiksel yöntemleri karşılaştırmıştır. Üç farklı veri yapısı ve dağılımında, ROC eğrisi altında kalan alanı kestirmeye yönelik parametrik ve parametrik olmayan yöntemlerin hangilerinin daha iyi sonuç verdiğini incelemiştir. Bu amaçla tanı testinin verilerinin sürekli olması durumunda ya da süreksiz hâle dönüştürülmesi durumunda ROC eğrisi altında kalan alanlar (AUC) arasındaki farklılık incelenmiştir. Araştırma sonucunda, verilerin süreksiz hâle dönüştürülmesi durumunda test için belirlenen kesim noktası alındığında (AUC) en büyük alanlara ulaşılmıştır. Tanı testi sonuçlarının normal dağılım gösterdiği örneklerde parametrik yöntemin yansız olduğu sonucuna ulaşılmıştır.

Çetin ve Gelbal (2010), iki farklı ders için oluşturdukları çoktan seçmeli testlerle öğrencilerin geçti/kaldı durumlarının Angoff, Nedelsky ve Angoff (1-0) yöntemleri ile elde edilen kesme puanlarına göre farklılık gösterip göstermediğini incelemiştir. Araştırmada farklı yöntemlere göre belirlenen kesme puanlarına göre geçti/kaldı öğrenci yüzdeleri karşılaştırıldığında; Türkçe dersi için Nedelsky ve Angoff (1-0) yöntemleri arasında manidar bir fark olduğu, matematik dersi için ise Nedelsky ve Angoff (1-0) yöntemleri arasında manidar bir fark olmadığı bulgularına ulaşılmıştır. Öğrenciler hakkında verilecek kararlar için kullanılacak standart belirleme yöntemlerinden hangisinin seçileceğinin öğrenci ve yargıcı grubuna, derse ve kullanılacak testin yapısına bağlı olarak değişebileceği sonucuna ulaşılmıştır.

Ömür ve Selvi (2010) tarafından yapılan araştırmada, Angoff, Ebel ve Nedelsky yöntemleriyle ve farklı yargıcı sayıları (10, 20, 30) kullanılarak elde edilen kesme puanlarının sınıflama tutarlılığı indeksleri incelenmiştir. Çoktan seçmeli testlerin kullanıldığı araştırmanın sonunda, yargıcı sayısına bağlı olarak belirlenen kesme

puanlarında önemli bir deęişim olmamıştır. Üç yöntemde ve üç ayrı yargıcı sayısında en yüksek sınıflama tutarlılığı indeksinin yargıcı sayısının 10 olduğu Nedelsky yöntemiyle elde edildiğı sonucuna ulaşılmıştır.

Gündeđer (2012) çoktan seçmeli matematik testinden elde edilen verilere göre Angoff, Ebel, Angoff (1-0) standart belirleme yöntemleriyle elde edilen kesme puanları ve verilen kararları karşılaştırmıştır. Ayrıca bu üç farklı standart belirleme yöntemi arasındaki ilişki, belirlenen kesme puanlarının güvenilirliği ve yargıcıların standart belirlemedeki tutarlılığı genellenebilirlik kuramıyla incelenmiştir. Araştırmanın bulgularına göre, üç farklı yöntemle elde edilen kesme puanları farklı deđerdedir. En yüksek kesme puanının Angoff (1-0) yöntemi ile en düşük kesme puanının ise Ebel yöntemi ile elde edilmiştir. Yöntemlere göre başarılı kabul edilen öğrenci yüzdeleri arasında manidar bir farklılık olduğu belirlenmiştir. Öğrencilerin başarılı-başarısız olarak sınıflanmaları açısından Angoff– Angoff (1-0) ve Angoff–Ebel yöntemleri arasında yüksek derecede bir uyum varken Angoff (1-0)–Ebel yöntemleri arasında orta düzeyde uyumun olduğu sonucuna ulaşılmıştır. Belirlenen kesme puanlara göre Angoff ve Ebel yöntemleri arasında .01 düzeyinde manidar, Angoff ve Angoff (1-0) yöntemleri arasında .05 düzeyinde manidar bir farkın olduğu belirlenmiş, Ebel ve Angoff (1-0) yöntemleri arasındaki ilişkinin manidar olmadığı tespit edilmiştir. Ayrıca kesme puan belirlerken çalışılan en uygun yargıcı sayısının belirli bir düzeyden sonra artmasının güvenilirlik sonuçlarını deęiştirmeyeceğı sonucuna ulaşılmıştır.

Taşdemir (2013), Angoff ve Nedelsky standart belirleme yöntemleri ve sınır deđer saptama yöntemlerinden ROC ve Aralık Tahmini ile bir testin sınıflama doğruluklarını incelemiştir. Araştırmada, Kamu Personeli Seçme Sınavı'ndaki Türkçe ve matematik bölümlerinden tesadüfi olarak seçilen sorularla bir başarı testi oluşturulmuş ve ROC analizi için referans olarak öğrencilerin akademik başarı puanları kabul edilmiştir. Araştırmada standart belirleme yöntemleri için 10 farklı yargıcıdan görüş alınmıştır. Bu araştırmada, matematik ve Türkçe testleri için Angoff, Nedelsky, ROC ve Aralık Tahmini yöntemlerine göre farklı kesme puanları belirlenmiştir. Türkçe ve matematik testi için en yüksek kesme puanı Angoff yöntemi ile elde edilmiştir. Standart belirleme yöntemlerine göre iki ders için de başarılı sayılan öğrenci yüzdeleri arasında manidar farklılık olduğu sonucuna ulaşılmıştır. Angoff yönteminde Türkçe ve matematik maddelerine ait yargıcı kararları arasında uyumun olmadığı belirlenmiştir. İki ders için de Nedelsky yöntemine göre kesme puanı belirlemede yargıcı kararları

arasında uyumun olduğu sonucuna ulaşılmıştır. Sınır değer saptama yöntemlerinden ROC analizi ile belirlenen matematik testinin kesme puanı, öğrencileri başarılı/başarısız olarak doğru sınıflandırabilmiştir. Türkçe ve Matematik testinin aralık tahminine göre belirlenen kesme puanları, ROC analizinde belirlenen kesme puanlarıyla uyumlu bulunmuştur. Hem yargıcı kararı ile hem sınır değerlerle standart belirlemede, yöntemlerin birbirleri ile uyumlu olduğu görülmüştür. Kesme puanı belirleme ve öğrencileri sınıflamada doğru kararların verilmesinde yaygın olarak kullanılan Angoff ve Nedelsky yöntemleri gibi ROC analizinin de kullanılabileceği sonucuna ulaşılmıştır.

Koyuncu (2015) araştırmasında, sınır değer saptama yöntemlerinden ROC analizi ile çoklu puanlanan psikolojik bir testin kesme puanını belirleme çalışması yapmıştır. Araştırmada Sürekli Kaygı Ölçeği referans kabul edilerek KPSS Sürekli Kaygı Ölçeği'nin kesme puanı belirlenmiştir. Geliştirilen veya uyarlanmış çoklu puanlanan ölçeklerde ROC analizi ile kesme puanının nasıl belirlenebileceğini incelemek amacıyla yapılan çalışmada, cinsiyete ve örneklem büyüklüğüne göre de bulgular elde edilmiştir. Araştırma sonucunda, genel olarak KPSS Sürekli Kaygı Ölçeği'nin yorumlanmasında kullanılacak olan kesme puanı 28.5 olarak belirlenmiştir. Cinsiyete göre kesme puanı değerleri incelendiğinde, kız öğrenciler için kesme puanı 28.5, erkek öğrenciler için ise 26.5 olarak belirlenmiştir. KPSS Sürekli Kaygı Ölçeği'nin 0.54 olasılıkla gerçek durumda sürekli kaygısı yüksek olan kişileri doğru tespit edebildiği sonucuna ulaşılmıştır. Ölçeğin 0.82 olasılıkla da gerçek durumda sürekli kaygısı düşük olan kişileri doğru tespit edebildiği sonucuna ulaşılmıştır. Belirlenen kesim noktası ile KPSS Sürekli Kaygı Ölçeği'nin kabul edilebilir düzeyde düşük/yüksek kaygılı bireyleri ayırabildiği ve bu ayırımın gerçek durumla uyumunun %72 olduğu görülmüştür. Ayrıca tesadüfi olarak seçilen örneklem büyüklüklerine göre elde edilen kesme puanların aynı olmasına rağmen eğriyi oluşturan özgüllük, duyarlılık ve eğri altında kalan alan (AUC) değerlerinin değişkenlik gösterdiği belirlenmiştir. Çoklu puanlanan psikolojik testlerde kesme puanı belirleme yöntemi olan ROC analizinin kullanımının uygun olduğu sonucuna varılmıştır.

Buckhental, Plake, Impara ve Smith (2002) tarafından yapılan çalışmada Angoff yöntemi ile Angoff yönteminin eleştirilen yönlerini ortadan kaldırmak üzere geliştirilen İşaretleme (Bookmark) yöntemi karşılaştırılmıştır. İlk uygulamada yargıcılar Angoff ve İşaretleme yöntemlerine göre maddeler hakkında karar verirken ikinci

uygulamada yargıcılar öğrenci cevaplarını bilerek karar vermişlerdir. Angoff yönteminde verilerin yargıcılar tarafından bilinmesi durumunda kesme puanı değerinin düştüğü, İşaretleme yönteminde ise arttığı belirlenmiştir. Araştırmada Angoff ve İşaretleme yöntemleri ile belirlenen kesme puanlarına göre kalan öğrenci yüzdelerinde manidar bir farkın olmadığı görülmüştür. Ayrıca matematik testinin her iki yöneme göre de elde edilen kesme puanları arasında büyük farklılıklar gözlenmemiştir.

Wang, Pan ve Austin (2003), farklı standart belirleme yöntemleri ve bu yöntemlerin geçme puanları üzerindeki etkilerini inceledikleri araştırmada standart belirleme yöntemlerini üç boyutlu bir taksonomiye yerleştirmişlerdir. Bu taksonomide kavramsal yaklaşımlar (çıktıya dayalı yaklaşım, gelişime dayalı yaklaşım), operasyonel yaklaşımlar (teoriye dayalı yaklaşım ve verilere dayalı yaklaşım) ve haritalama yöntemi (gözlenen puana dayalı ve örtük puana dayalı yaklaşım) boyutları yer almıştır. Araştırma kapsamında incelenen on iki standart belirleme yöntemine göre oluşturulan kesme puanları birbirinden farklı bulunmuştur. Aynı taksonomi boyutunda yer alan yöntemlerin bile kendi içlerinde oldukça fazla değiştiği, bu değişimin %29'dan %79'a kadar farklılaşabildiği sonucuna ulaşılmıştır.

Downing, Tekian ve Ydkowsky (2006) tarafından yapılan araştırmada Sınır Grup, Zıt Grup, Ebel ve Angoff, yöntemlerinin süreçlerini incelemişlerdir. Farklı standart belirleme yöntemlerinin farklı kesme puanları ürettiği ve kesme puanları belirlenirken yargıcıların kararlarının öznel kararlar olduğu sonucuna varılmıştır.

Takahashi, Uchiyama, Yanagisawa ve Kamae (2006) yaptıkları araştırmada ROC analizi ile diyabet tahmininde kullanılan hemoglobin A1c testinin kesme puanını hesaplamışlardır. Referans olarak kandaki glukoz değerleri alınmış ve diyabet hastalığının olup olmadığı durumu belirlenmiştir. Elde edilen verilerde ROC eğrisine dayalı olarak diyabet tahmininde kullanılan hemoglobin A1c testinin kesme puanı 5.3 (özellik % 92.1, duyarlılık % 84.2) olarak hesaplanmıştır.

Irwin (2007) araştırmasında Karşıt Gruplar 1 ve Karşıt Gruplar 2 adını verdiği yöntemlerle; Angoff, Sınır Gruplar ve Zıt Gruplar yöntemlerini karşılaştırmıştır. Karşıt Gruplar 1 ve Karşıt Gruplar 2 yöntemleri ile elde edilen kesme puanların, Angoff ve Zıt Grup yöntemleriyle elde edilen değerlere göre yüksek; Sınır Grup yöntemiyle elde edilen değere göre ise düşük olduğu sonucuna ulaşmıştır.

Vivo ve Franco (2008) akademik başarının yordayıcı değişkenlerinin değerlendirilmesinde ve üniversiteye giriş sınavının kesme puanını belirlemede ROC analizinin kullanılması üzerine bir araştırma yapmışlardır. Araştırmada, lisans öğrencilerinin akademik başarı sınıflandırmasında üniversiteye giriş değişkenlerinin geçerliğini değerlendirmek ve farklı düzeylerdeki öğrenci başarısını tanımlamak için kesme puanı belirlemek amaçlanmıştır. Üniversiteye girişte akademik başarıyı sınıflandırmak için ortaöğretim başarı puanı ortalaması, üniversite giriş genel sınav puanı, üniversite giriş özel kısım sınav puanı, ortaöğretim ve üniversite giriş sınavı ağırlıklandırılmış puanlar olmak üzere dört değişken incelenmiştir. Ortaöğretim başarı puanı ortalamasının, ortaöğretim ve üniversite giriş sınavı ağırlıklandırılmış puan değişkenlerine göre başarılı veya başarısız öğrenciyi sınıflandırmada daha iyi sonuç verdiği sonucuna ulaşılmıştır. Ayrıca diğer değişkenlere göre daha doğru sınıflandırma yapan ortaöğretim başarı ortalaması değişkeninin yordayıcılığının en yüksek olduğu görülmüştür.

Clauser ve diğerleri (2009) tarafından yapılan araştırmada, Angoff yöntemi için yargıcılarla belirlenen kesme puanlarının kararlılığının ortaya çıkarılmasında Genellenebilirlik Teorisi'nden yararlanılmıştır. Araştırmada, Angoff yöntemi için iki tür kanıt ortaya koyulmaya çalışılmıştır. İlki yargıcıların belirlediği kesme puanlarının kararlılığı, ikincisi ise yargıcıların yaptığı madde güçlüğü tahmini ile gerçek madde güçlüğü arasındaki ilişkidir. Ayrıca araştırmada uzmanların grup olarak yaptıkları tartışmaların bireysel tahminleri üzerindeki etkisine de bakılmıştır. Araştırmanın sonucunda, yargıcıların bir araya gelerek yaptıkları tartışmaların yargıcı kararları arasındaki farklılığı azalttığı belirlenmiştir. Tartışmaların, madde güçlük düzeylerine göre yapılan yargıcı kararlarında farklılığı artırdığı gözlenirken yargıcı tahminleri ile maddelerin gerçek güçlükleri arasındaki ilişkiyi güçlendiremediği görülmüştür. Ayrıca yargıcılar, zor maddeler için başarı oranını gerçekte olandan daha yüksek düzeyde tahmin ederken kolay maddeler için başarı oranını var olandan daha düşük düzeyde tahmin etmişlerdir.

Moraes, Freitas, Mondini ve Rosas (2009) Meksika'da çocuklarda ve ergenlerde doğum ağırlığını fazla kilolu ve normal olarak tanımlamak için ROC eğrisi yöntemiyle kesme puanı belirleme çalışması yapmışlardır. Araştırma grubu yaş ve cinsiyet değişkenlerine göre sınıflandırılmıştır. Çocukların doğum ağırlıkları değişkenine aşı

kartlarından ulaşılmıştır. Çocukların beslenme durumları yaş ve cinsiyet değişkenleri için özel vücut kitle indeksi kesme puanına göre tanımlanmıştır. Değişkenlere göre tahmin için lineer regresyon modelleri kullanılmıştır. Doğum ağırlık dilimleri ile vücut kitle indeksi ortalamaları doğrusal bir ilişki göstermiştir. Araştırmada sınıflama doğruluğunu incelemek için ROC eğrisi altında kalan hesaplanmıştır. Duyarlılık, özgüllük değerlerinin en uygun noktasına karşılık gelen doğum ağırlığı kesme puanı olarak belirlenmiş ve ROC eğrisi çizilmiştir. ROC eğrisi altında kalan alan yaşa ve cinsiyete göre her tabakada %78'den büyük olduğu ve ergenlerde cinsiyete göre farklılık gösterdiği belirlenmiştir. Oluşturulan ROC eğrisinde doğum ağırlığı kesme puanının kızlara göre erkeklerde biraz daha yüksek olduğu ve alanların 0.70'ten büyük olduğu hesaplanmıştır. Araştırma sonucunda fazla kilolu çocuk ve ergenlerin risk grubunda olup olmadığını belirlemek için doğum ağırlığı kesme puanının kullanılabilmesine ulaşılmıştır.

Stone, Koskey ve Sondergeld (2011) standart belirleme yöntemlerinden Angoff ve Angoff (1-0) modeli olmak üzere, kesme puanı belirleme çalışmalarının, ölçme süreçlerinin kavramsallaştırmaları ile ilişkili olan yapı kavramını göz önünde bulundurmadığını belirtmişlerdir. İki yöntemle yapılan beş yıllık araştırma sonuçları incelendiğinde Angoff yönteminin standart belirlemede kararlı bir yapı tanımlamadığı sonucuna ulaşılmıştır.

Tavakol ve Dennick (2012) çoktan seçmeli bir testin ROC analizi yöntemiyle kesme puanını belirleme üzerine bir araştırma gerçekleştirmişlerdir. Ölçüt test olarak yine aynı test kabul edilmiş ve standart belirleme yöntemleriyle yargıcıların belirlediği kesme puanı olan 11'e göre ROC analizi yapılmıştır. Öğrencilerin geçti/kaldı sınıflandırmaları, 11'den büyük puan alan öğrenciler geçti grubuna, 11 ve 11'den küçük puan alan öğrenciler kaldı grubuna dâhil edilerek yapılmıştır. ROC analizi yöntemiyle elde edilen kesme puanının, en yüksek Youden İndeks değeri olan 0.92 değerine karşılık gelen 10.5 noktası olduğu gözlemlenmiştir. 10.5 kesme puanına karşılık gelen duyarlılık değeri 1.0; özgüllük değeri 0.92 olarak hesaplanmıştır. AUC değeri ise 0.62 olarak hesaplanmıştır. ROC analizi ile belirlenen 10.5 kesme puanının standart belirleme yöntemleriyle yargıcılar tarafından belirlenen 11 kesme puanına çok yakın olduğu görülmüştür. ROC eğrisinin öğrencilerin geçti veya kaldı şeklinde sınıflandırılmasında uygun kesme puanı bilgisini sağladığı belirlenmiştir. Ayrıca ROC analizi yönteminin, standart belirleme yöntemlerinde olmayan duyarlılık ve özgüllük

değerlerini dikkate aldığı için daha objektif kesme puanı belirlediği sonucuna ulaşmıştır.

İlgili alanyazın incelendiğinde araştırmalar, ROC analizi ile hesaplanan duyarlılık, özgüllük, negatif ve pozitif yordama gücü ve eğri altında kalan alan değerlerinin ölçme aracının geçerliğine katkı sağladığını bir kez daha ortaya koymaktadır. Bu doğrultuda çoklu puanlanan ölçeklerin uyarlama ve geliştirilme çalışmalarında ROC analizi ile psikometrik niteliklerin incelendiği araştırmalar da aşağıda sunulmuştur.

Aydemir ve diğerleri (2006) tarafından yapılan araştırmada Sınır Kişilik Envanteri, Türk kültürüne uyarlanmış ve psikometrik nitelikleri ROC analizi ile incelenmiştir. Araştırma psikiyatri polikliniğine başvuran, Tanı Ölçütleri Başvuru El Kitabı'na (DSM-IV) göre tanı konulmuş 40 sınır kişilik bozukluğu, 35 klinik depresif bozukluğu ve 30 şizofreni hastası olmak üzere 105 hasta ve herhangi bir psikiyatrik bozukluğu olmayan 61 kişiden oluşan kontrol grubu ile yürütülmüştür. Güvenirlik analizinde tüm araştırma grubundan elde edilen verilerden hesaplanan Cronbach alfa iç tutarlılık katsayısı 0.92, test-tekrar test güvenirlik korelasyon katsayısı ise $r=0.67$ olarak bulunmuştur. Geçerlik analizinde ROC analizi ile belirlenen 15 kesme puanına göre özgüllük ve duyarlılık değerleri incelenmiştir. Belirlenen kesme puanının duyarlılık değeri %80, özgüllük değeri ise %79.3 olarak belirlenmiştir. Bu kesme puanının pozitif yordama değeri %55.1, negatif yordama değeri ise %92.5 olarak bulunmuştur. Sınır Kişilik Envanteri'nden elde edilen puanlar sırasıyla sınır kişilik bozukluğu bulunan grup ($n=40$) için 23.2 ± 8.3 , klinik depresif bozukluğu bulunan grup ($n=35$) için 12.3 ± 7.7 , şizofreni olan grup ($n=30$) için 11.0 ± 8.1 , sağlıklı olan grup ($n=61$) için ise 5.1 ± 3.7 olarak belirlenmiştir. Grupların puan ortalamaları arasındaki fark .01 düzeyinde manidar bulunmuştur. Sonuç olarak Sınır Kişilik Envanteri'nin, sınır kişilik bozukluğu grubunu diğer hasta gruplarından ve sağlıklı bireylerden iyi derecede ayırt edebildiği ve envanterin gündelik uygulamada güvenilir ve geçerli biçimde kullanılabilir olduğu sonucuna ulaşılmıştır.

Konuk ve diğerleri (2007) yaptıkları araştırmada Duygu Durum Bozuklukları Ölçeği'nin Türk kültürüne uyarlanmasını, geçerliğini ve ölçeğin kesme puanını ROC analizi ile belirlemişlerdir. Aynı hastalar Tanı Ölçütleri Başvuru El Kitabı'na (DSM-IV) göre değerlendirilmiş ve bu belirlemeler ROC analizi için referans olarak kabul

edilmiştir. Belirlenen tanılarına göre duygu durum bozukluk tanısı konulmayan hasta sayısı 7'dir (%2). Duyarlılık değerinin %64 ve özgüllük değerlerinin %77 olduğu ölçeğin kesme puanı 7 olarak belirlenmiştir. Bu kesme puanının pozitif yordama değeri %26.7, negatif yordama değeri %94.2 olarak bulunmuştur. Ölçek uygulanırken kesme puanının üzerinde alan bireylerin riskli olduğu, tanı koyabilmek için bu kişilerin daha detaylı incelenmesi gerektiği belirtilmiştir. Araştırmada ölçeğin iki grubu ayırma gücünün yüksek olduğu, tarama amaçlı duygu durum bozukluğunun belirlenmesinde kullanılabileceği ve diğer analizlerle geçerlik çalışmasının desteklenmesi gerektiği sonucuna varılmıştır.

Öner ve diğerleri (2009) tarafından Çocuklarda Uyku Ölçeği'nin (ÇUÖ) Türk kültürüne uyarlanması, geçerlik, güvenilirlik ve kesme puanı araştırması yapılmıştır. Çoklu puanlanan 22 maddelik ölçeğin, deney ve kontrol grupları oluşturularak psikometrik özellikleri belirlenmiştir. Kesme puanı, duyarlılık, özgüllük, negatif ve pozitif yordayıcı değerlerinin hesaplanması için İşlem Karakteristik Eğrileri kullanılmıştır. ROC eğrisinden elde edilen, kesim noktası olan 6 puan incelendiğinde duyarlılığın 0.94, özgüllüğün 0.92, pozitif yordayıcı değerin 0.93 ve negatif yordayıcı değerin 0.94 olduğu görülmektedir. Sözü edilen ölçeğin tanılı çocuklar ile sağlıklı çocukları (kontrol grubu) oldukça iyi bir şekilde ayırabildiği görülmüştür. Araştırmada olgunun genel değerlendirmesini kolaylaştıran ve psikometrik özellikleri yüksek düzeyde olan bir ölçeğin var olduğu sonucuna ulaşılmıştır.

Konkan, Şenormancı, Güçlü, Aydın ve Sungur (2013), Yaygın Anksiyete Bozukluğu Ölçeği'ni (YABÖ) Türk kültürüne uyarlayarak geçerlik, güvenilirlik ve ölçeğin kesme puanını belirleme çalışması yapmışlardır. Ölçek 4'lü likert tipi 7 maddeden oluşmaktadır. Araştırma psikiyatri uzmanlarınca yaygın anksiyete bozukluğu tanısı konulan deney grubu ve yaygın anksiyete bozukluğu tanısı olmayan kontrol grubu ile yürütülmüştür. ROC analizi sonuçlarına göre kesme puanı 8 olarak uygun bulunan ölçeğin pozitif yordayıcı değeri %79, negatif yordayıcı değeri %81 olarak belirlenmiştir. ROC analizine göre ölçeğin mükemmel ayırt ediciliğinin olduğu ve yaygın anksiyete bozukluğu için tarama testi olarak kullanılabileceği sonucuna ulaşılmıştır.

Türkiye'de ve yurt dışında kesme puanı belirleme ile ilgili yapılan araştırmalar incelendiğinde genellikle çoktan seçmeli başarı testleri ile çalışıldığı ve test merkezli standart belirleme yöntemlerinin kullanıldığı görülmüştür. Belirlenen kesme puanları

arasındaki farklılıklar, yöntemler arası kıyaslamalar, yargıcılar arası uyum indeksleri, yargıcı sayısı ve özellikleri genellikle odaklanılan noktalar olmuştur. Farklı yöntemlerin incelendiği ve yöntemlerin birbirleriyle karşılaştırıldığı birçok araştırmada özellikle Angoff ve Nedelsky yöntemleri kullanılmıştır. Angoff yöntemiyle belirlenen kesme puanının diğer yargıcı kararları ile belirlenen kesme puanlarından daha yüksek olduğu sonucuna ulaşılmıştır. Farklı standart belirleme yöntemleri ile belirlenen kesme puanlarına göre bireylerin ikili sınıflandırılma yüzdeleri incelendiğinde farkın genellikle anlamlı olduğu sonucuna ulaşılmıştır. Standart belirleme yöntemlerinde kaç yargıcı ile çalışmanın uygun olacağını inceleyen araştırmalarda on yargıcının yeterli olacağı, yargıcı sayısının artmasının belirlenen kesme puanlarını önemli derecede etkilemediği belirtilmiştir. Yargıcılar arası uyum indekslerinin incelendiği çalışmalarda, genellikle orta düzeyde uyumun olduğu sonucuna ulaşılmıştır. En yüksek yargıcılar arası tutarlılık indeksi Nedelsky, en düşük tutarlılık indeksi ise Ebel yöntemi ile elde edilmiştir. Sınır değer saptama yöntemlerinden ROC analizinin, kesme puanı belirlemede ve bireyleri doğru sınıflandırmada kullanılabileceği sonucuna ulaşılmıştır. ROC analizi ile belirlenen kesme puanının, standart belirleme yöntemlerinde olmayan duyarlılık ve özgüllük değerlerini dikkate aldığı için daha objektif kesme puanı belirlediği sonucuna ulaşılmıştır. Ayrıca ROC analiziyle elde edilen değerlerin, ölçek geliştirme ve uyarlama çalışmalarında geçerlik kanıtı olarak kullanıldığı görülmüştür. Ölçülen psikolojik yapıya bağlı olarak belirlenen kesme puanı üzerinde puan alan bireyler için bir ölçekle hemen tanı konamayacağı, bireylerin daha detaylı incelenmesi gerektiği belirtilmiştir.

Eğitim bilimleri alanında yapılan araştırmalar incelendiğinde çoklu puanlanan bir psikolojik test için kesme puanı belirlenirken test ve öğrenci merkezli yöntemlerle birlikte ROC analizinin kullanıldığı ve belirlenen kesme puanlarının uyumunun değerlendirildiği, ROC analizi ile geçerlik çalışmasının yapıldığı bir çalışmaya rastlanmamıştır. Ayrıca eğitim alanında kullanılan çoklu puanlanan psikolojik testin geliştirme veya uyarlama aşamasında kesme puanı belirleme çalışmalarının yapılmadığı, bu nedenle de test puanlarının yorumlanmasında aritmetik ortalama dışında bir ölçüt puanının kullanılmadığı belirlenmiştir. Bu araştırmada uyarlanan ve geliştirilen psikolojik testlerde ölçüt puanı oluşturulurken standart belirleme yöntemlerinin kullanılmasının alınan kararların doğruluğunu sağlayacağı düşünülmektedir. Dolayısıyla bu araştırma ile farklı kesme puanı yöntemlerinin karşılaştırılmasıyla test puanlarının yorumlanmasına ve testin geçerliğine katkı

sağlanarak test geliştiricilere ve uyarlama çalışması yapanlara farklı bir bakış açısı sunulması planlanmaktadır.

Bu doğrultuda çoklu puanlanan psikolojik testlerde kesme puanı belirleme yöntemlerinin kullanılmasına ve psikometrik niteliklerin incelenmesine duyulan gereklilik, bu araştırmanın problemini oluşturmaktadır.

1.2. Amaç

Bu çalışmada üniversite öğrencilerinin kaygı düzeylerini belirleyen bir ölçeğin Angoff, ROC ve Sınır Grup yöntemleri kullanılarak kesme puanının belirlenmesi amaçlanmıştır. Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt bulunmaya çalışılmıştır.

1. Üniversite öğrencilerinin Sürekli Kaygı Ölçeği'nden aldıkları puanların değerlendirilmesi için;
 - a. Angoff yöntemi ile belirlenen kesme puanı kaçtır?
 - b. Sınır Grup yöntemi ile belirlenen kesme puanı kaçtır?
 - c. ROC analizi ile belirlenen kesme puanı kaçtır?
2. Angoff ve Sınır Grup yöntemleri ile belirlenen kesme puanlarına göre düşük/yüksek kaygılı kabul edilen öğrenci yüzdeleri arasında manidar bir fark var mıdır?
3. Kesme puanlarının belirlenmesinde;
 - a. Angoff yöntemi kullanıldığında yargıcılar arası uyum katsayısı nedir?
 - b. Sınır Grup yöntemi kullanıldığında yargıcılar arası uyum katsayısı nedir?
4. ROC analiziyle Sürekli Kaygı Ölçeği'nden elde edilen eğri altında kalan alan değerinin düşük/yüksek kaygılı öğrencileri ayırmadaki doğruluk oranı nedir?
5. Sürekli Kaygı Ölçeği'nin ROC analizi ile geçerlik düzeyi nedir?

1.3. Önem

Psikolojik testler bireyler hakkında karar vermede, bilgi toplamada, tanı koymada, işe alımlarda, meslek seçiminde ve eğitim gibi pek çok farklı alanda kullanılmaktadır. “Eğitim ve psikolojide ölçme değerlendirme ile alınan kararlar, bireyleri, kurumları ve bir bütün hâlinde toplumu önemli ölçüde etkilemektedir. Etkilenen bireyler arasında, öğrenciler, ebeveynler, öğretmenler, eğitim yöneticileri, işverenler, işe başvuranlar, hastalar, danışmanlar, amirler ve alan uzmanları

bulunmaktadır. Etkilenen kuruluşlar arasında okullar, üniversiteler, iş yerleri, endüstri ve devlet kurumları bulunmaktadır. Testler amaçlara ulaşmaya yardımcı olduğu takdirde bireyler ve kurumlar, bundan yarar sağlayacaktır” (Amerikan Eğitim Araştırmaları Birliği, Amerikan Psikoloji Birliği ve Eğitim Ölçümleri Uluslararası Konseyi, 1997). Eğitim ve sağlık alanlarında ölçme araçlarından fazlaca yararlanılması, söz konusu ölçme araçlarının geliştirilme ve uyarlanma süreçlerinin araştırılması konusunda yapılan çalışmaların önemini artırmaktadır.

Uygulayıcılar, birçok psikolojik testi kullanırken test puanlarını neye göre, nasıl yorumlayacakları konusunda sorun yaşamakta ve aldıkları kararların doğruluğunu sorgulamaktadır. Günümüzde bir kültür için hazırlanmış olan psikolojik ölçme aracı farklı dil ve kültürlere çevrilerek de kullanılmaktadır. Eğitim alanında uyarlanan veya geliştirilen psikolojik testlerde değerlendirme yapılırken örneklemin ortalaması ölçüt olarak kabul edilmektedir. Ayrıca uyarlama çalışmaları incelendiğinde ölçek puanlarının yorumlanmasında kullanılan ölçütün uyarlanan ölçekten olduğu gibi alındığı, testin normu oluşturulurken kesme puanı belirleme yöntemlerinin uyarlama sürecine dâhil edilmediği belirlenmiştir (Erkuş, 2007). Uygulayıcıların kullanacakları ölçme araçlarına ait psikometrik niteliklerin neler olduğunu bilmeleri, bireylerin test puanlarının anlamlandırılabilmesi için önemlidir. Bahsedilen psikometrik özelliklerin yanında tipik tepki testlerinin maliyeti, uygulama ve puanlama kolaylığı gibi bazı pratik özelliklere sahip olması da bireylerin duyuşsal özelliklerini ortaya çıkarmada istenen niteliklerdendir.

Uygulayıcılar seçme, sınıflama, sıralama yaparken bireyin test puanını, testi alan diğer bireyler içindeki yeri ile karşılaştırıp bağıl değerlendirme yapmaktadır. Psikolojik testlerde bağıl değerlendirme yerine, testin uygulanma amacına göre seçilecek uzman niteliğindeki yargıcıların kararlarına dayalı olarak kesme puanı belirlenebilir. Mutlak ölçütlerin kullanılması, bireylerin ölçülen özellik bakımından düşük/yüksek gibi sınıflandırılmasında karşılaşılan sorunları giderebilir. Ayrıca testlere ilişkin geçerli ve güvenilir kesme puanlarının belirlenmesi ve puanlar arasında kalan düzeylerin hangi tepki düzeyi sınıflaması içinde kabul edildiği konusunda bireylerin bilgilendirilmesi de sorunların giderilmesine yardımcı olabilir. Çünkü yargıcılardan kendilerini bireyin yerine koyarak test maddelerini cevaplamaları ya da test maddelerinin ölçülen özelliklerle ilgili sınır düzeydeki bireyi ayırt edebilme durumlarını belirlemeleri beklenmektedir.

Yani testin ölçüt puanı yargıcı gruptan elde edilen puanlarla oluşturulmaktadır. Uyarlama aşamasında yapılan ROC analizi ile de testin elde edilen kesme puanına göre bireyleri ayırt etme gücü belirlenebilecektir. Dolayısıyla analizden elde edilen yordayıcı değere göre de testin psikometrik nitelikleri incelenebilir. Standart belirleme süreci sadece bir kesme puanı oluşturmaktan çok bir yol belirlemeyi amaçlamakta, bireyleri sınıflandırma konusunda farklı bir çerçeve sunmaktadır. Bu yolla testteki kesme puanlarının ve testten alınan puanların karşılık geldiği tepki düzeyinin ne anlama geldiği tanımlanabilir ve bu testin puanlarına dayalı olarak yapılan sınıflamaların doğruluğu hakkında bilgi sunulabilir. Ayrıca yöntemlerin sınıflama doğruluklarının incelenmesiyle test geliştirme ve uyarlama çalışmalarında kesme puanı belirleme yöntemlerinin kullanılmasının ölçek puanlarının geçerliğine katkı sağlayacağı düşünülmektedir.

Araştırma, çoklu puanlanan bir testin farklı yöntemlerle ölçüt puanının oluşturulması ve bu ölçüt puanlarına göre bireylerin sınıflandırılma durumlarının incelenmesiyle testin geçerliğine katkı, alana ve araştırmacılara yeni bir bilgi sunacağından önem taşımaktadır.

1.4. Sayıtlar

1. Cevaplayıcıların ölçek maddelerini yanıtlarken gerçek tepki düzeylerini yansıttıkları varsayılmıştır.
2. Araştırmaya katılan yargıcıların Angoff ve Sınır Grup standart belirleme yöntemlerine uygun biçimde puanlama yaptıkları varsayılmıştır.
3. Yargıcıların ölçek maddeleri ile ilgili kararlarını, yüksek kaygının minimum düzeyindeki öğrencilerin tepkilerini dikkate alarak verdikleri varsayılmıştır.

1.5. Sınırlılıklar

1. Bu araştırma, bir devlet üniversitesinin eğitim fakültesinin son sınıflarında öğrenim gören öğrencilerle sınırlıdır. Bu öğrenciler mezuniyete yakın olmaları ve kamu personeli seçme sınavına (KPSS) hazırlanmaları sebebiyle kaygı düzeylerinin diğer sınıflara göre daha yüksek olacağı beklendiğinden araştırmaya dâhil edilmiştir.
2. Bu araştırmada kesme puanı belirleme işlemi standart belirleme yöntemlerinden Angoff ve Sınır Grup, sınır değer saptama yöntemlerinden ROC analizi ile sınırlıdır.

1.6. Tanımlar

Kesme Puanı: Öğrencilerin yüksek veya düşük kaygılı olarak sınıflandırılmasında kullanılacak, sınır kabul edilen puandır.

Kaygı: Bireyde uyarılmışlık durumu yaratan, nesnel olmayan bir tehlikeye karşı kişinin yaşadığı, bedensel ve zihinsel değişimlerle kendini gösteren duygu durumudur.

Sürekli Kaygı: Olumlu ya da olumsuz bir tepki durumu oluşturmayacak durumların birey tarafından tehlikeli ve zarar verici olarak algılanması sonucu oluşan hoşnutsuzluk ve mutsuzluk duygusudur.

Sınav Kaygısı: Bireyin değerlendirilmesi durumunda ortaya çıkan korkuyla karışık hissedilen bir tedirginlik duygusudur.

1.7. Kısaltmalar

ROC: Receiver Operating Characteristics Curve (İşlem Karakteristik Eğrisi)

AUC: Area Under Curve (ROC Eğrisi Altında Kalan Alan)

BÖLÜM 2

YÖNTEM

Bu bölümde araştırmanın modeli, araştırma grubu, verilerin toplanması ve verilerin analizine yer verilmiştir.

2.1. Araştırma Modeli

Sürekli Kaygı Ölçeği'nin Angoff, ROC ve Sınır Grup yöntemlerine dayalı olarak kesme puanının belirlenmesi amaçlandığından araştırmanın modeli bu yönüyle temel araştırmadır. Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel (2011) temel araştırmayı bilgi, kuram üretmeye yönelik çalışmalar olarak tanımlamışlardır. Karasar'a (2009) göre temel araştırmaların amacı, var olan bilgiye yenilerini katmaktır. Aynı zamanda araştırma, çoklu puanlanan bir ölçeğin kesme puanının farklı yöntemlerle değişiminin incelenmesi ve en uygun puana karar verilmesi yönüyle bilgi üretmeye yöneliktir.

2.2. Araştırma Grubu

Araştırmada yargıcılar ve öğrenciler olmak üzere iki farklı araştırma grubu yer almıştır. Araştırma grubunu, 2015-2016 Akademik Yılı Bahar Dönemi'nde Rize ili Recep Tayyip Erdoğan Üniversitesi Eğitim Fakültesi'nde öğrenim gören, farklı lisans programlarının son sınıflarındaki 290 öğrenciden oluşmaktadır. Belli niteliklere sahip, yeterli sayıdaki öğrenciler, amaçsal örnekleme yöntemlerinden ölçüt örnekleme göre seçilmiştir (Büyüköztürk ve diğ., 2011). Araştırma grubunu oluşturan öğrencilerin öğrenim gördükleri programlara göre dağılımları Çizelge 3'te sunulmuştur.

Çizelge 3.

Araştırma Grubunu Oluşturan Öğrencilerin Lisans Programlarına Göre Dağılımı

Bölüm	N	%
Fen Bilgisi Öğretmenliği	41	14
İlköğretim Matematik Öğretmenliği	89	31
Sınıf Öğretmenliği	71	24
Sosyal Bilgiler Öğretmenliği	32	11
Türkçe Öğretmenliği	57	20
Toplam	290	100

Çizelge 3'te görülebileceği gibi araştırma grubunda fen bilgisi öğretmenliği (41), ilköğretim matematik öğretmenliği (89), sınıf öğretmenliği (71), sosyal bilgiler öğretmenliği (32) ve Türkçe öğretmenliği (57) programlarında öğrenim gören ve 2016 Kamu Personeli Seçme Sınavı'na katılacak olan öğrenciler yer almıştır.

Kesme puanının belirlenmesi için değerlendirme yapmaları istenen yargıcılar ise düşük ve yüksek kaygılı öğrencilerin özelliklerini iyi bilen ve kaygı durumu ile ilgili çalışmaları olan rehberlik ve psikolojik danışmanlık alanında görev yapan öğretim elemanlarıdır. Alanyazında yargıcı grubun sayısına ilişkin farklı görüşler yer almaktadır. Jeager (1989) çalışmasında uzman sayısının fazla olmasının avantajlı ve dezavantajlı yönlerini açıklamakta ve 10 ile 15 arasında olan bir yargıcı grubun yeterli olduğunu belirtmektedir. Ayrıca Taşdelen (2009) genellenebilirlik kuramına göre standart belirleme ile ilgili araştırmasında 10 yargıcının yeterli olacağını belirtmiştir. Gündeğer (2012) Angoff yönteminde en uygun yargıcı sayısının belirlenmesine yönelik yaptığı çalışmada, yargıcı sayısının artırılması ile güvenilirliğin değişmediği sonucuna ulaşmıştır. Kesme puanı belirlenirken yargıcı sayısının kaç olması gerektiği ile ilgili olarak Hurtz ve Hertz (1999), Angoff yönteminde kaç yargıcı kullanılmalı gerektiğini araştırmışlar ve yaklaşık 10-15 yargıcının hatta bazı durumlarda daha az sayıda yargıcının bile yeterli olacağı sonucuna ulaşmışlardır. Bu bilgiler doğrultusunda araştırmada 10 yargıcı ile çalışılmıştır. Kesme puanı belirleme sürecine katılan yargıcıların çalıştıkları üniversitelere ilişkin bilgiler Çizelge 4'te sunulmuştur.

Çizelge 4.

Araştırma Grubunu Oluşturan Yargıcıların Üniversitelere Göre Dağılımı

Üniversite	Uzman Sayısı	%
Başkent Üniversitesi	2	20
Karadeniz Teknik Üniversitesi	4	40
Recep Tayyip Erdoğan Üniversitesi	4	40
Toplam	10	100

Çizelge 4'te görülebileceği gibi rehberlik ve psikolojik danışmanlık alanında görev yapan Başkent Üniversitesinden iki, Karadeniz Teknik Üniversitesinden dört ve Recep Tayyip Erdoğan Üniversitesinden dört yargıcı ile çalışılmıştır. Kaygı durumu ile ilgili değerlendirme yapmaları istenen yargıcı grup, doktorasını tamamlamış veya tamamlamak üzere olan öğretim elemanlarından oluşmaktadır.

2.3. Veriler ve Toplanması

Araştırmada verilerin toplandığı gruplar, öğrenciler ve yargıcılardır. Öğrenci verilerini Sürekli Kaygı ve Sınav Kaygısı Ölçeklerinden alınan puanlar oluşturmaktadır. Yargıcı verileri ise yargıcıların Sınav Kaygısı Ölçeği'nin maddeleri hakkında aldıkları kararlardır. Ölçekler aynı kavramla ilgili olması ve yanıtlama sürelerinin kısa olması nedeniyle öğrencilere tek seferde uygulanmıştır. Yargıcıların kesme puanı belirlerken verdikleri kararlar ise araştırmacı tarafından hazırlanan bir formula elde edilmiştir. Veriler toplanmadan önce yargıcıların yüksek kaygının minimum düzeyi ile ilgili ortak bir anlayış geliştirilmesi ve puanlamaları arasında tutarlılığın sağlanması amacıyla yöntemlerle ilgili ayrıntılı bilgi verilmiş ve kesme puanının ne anlama geldiği araştırmacı tarafından açıklanmıştır. Yargıcılardan, her bir ölçek maddesini Angoff ve Sınır Grup yöntemlerine göre cevaplamaları istenmiştir. İlk uygulamada araştırmacı tarafından yargıcılara, öğrenci merkezli olan Sınır Grup yöntemi ile ilgili bilgi verilmiştir. Yargıcılar, “yüksek kaygının minimum düzeyinde” olan bir öğrenciyi temel alarak yani onun tepkilerini göz önünde bulundurarak ölçek maddelerini cevaplamışlardır. İkinci uygulamada araştırmacı tarafından yargıcılara, test merkezli olan Angoff yöntemi için ayrıntılı bilgi sunulmuştur. Yargıcılar, her madde için hangi tepki derecesinin “yüksek kaygının minimum düzeyinde” olan öğrenciyi ayırt edeceğine karar vermişlerdir. Yani birinci uygulamada yargıcılara öğrenciyi göz önünde bulundurmaları, ikinci uygulamada ise ölçek maddelerini dikkate almaları istenmiştir. Her iki uygulamada da yargıcılardan “yüksek kaygının minimum düzeyindeki öğrenciyi” düşünüp cevap vermeleri istenmiştir. Ölçek, aynı yargıcılara onar gün arayla uygulanmıştır. Öğrenci ve test merkezli yöntemlere göre yargıcıların cevaplarından oluşan ölçek puanları, yargıcıların verilerini oluşturmaktadır.

2.3.1. Veri Toplama Araçları

Araştırmada farklı yöntemlerle kesme puanı belirlenen, çoklu puanlanan test Sürekli Kaygı Ölçeği'dir. İki faktör kaygı kuramına dayalı olarak Spielberger ve arkadaşları (1970) kaygının farklı iki durumunu ölçmek amacıyla Durumluk ve Sürekli Kaygı Envanteri'ni geliştirmişlerdir. Türkçeye uyarlanması Öner ve Le Compte tarafından 1983 yılında yapılmıştır. Envanter lise öğrencilerinin ve normal yetişkinlerin kaygı düzeylerini ölçmek amacıyla kullanılmakta ve toplam 40 maddeden oluşmaktadır. Envanterin her biri 20 maddelik ve dört dereceli Likert tipi iki ayrı ölçeği bulunmaktadır (Özgül, 2007). Bu envanterden Sürekli Kaygı Ölçeği'nin tercih

edilmesinin nedeni; bireyin içinde bulunduğu durum ve koşullardan bağımsız olarak kaygısını belirlemesi, rehberlik ve psikolojik danışmanlık alanında sıklıkla kullanılması ve araştırma grubunu oluşturan öğrencilerin yanıtlamaktan sıkılmayacakları uzunlukta olduğunun düşünülmesidir.

2.3.1.1. Sürekli Kaygı Ölçeği

Ölçek, bireyin içinde bulunduğu durum ve koşullardan bağımsız olarak genellikle kendini nasıl hissettiğini belirlemek amacıyla geliştirilmiştir. Bireyler, genel olarak hissettikleri duygulara göre ölçek maddelerini yanıtlamaktadırlar. Bireylerden, her maddeye yönergedeki kurallara uyarak; (1) pek az, (2) bazen, (3) çoğu zaman, (4) hemen her zaman gibi katılma derecelerini gösteren seçeneklerden birini işaretlemeleri beklenmektedir. Ölçeğin cevaplanması yaklaşık 10 dakika sürmektedir. Ölçekte olumlu ve olumsuz maddeler bulunmaktadır. Olumlu duyguları dile getiren maddeler puanlanırken 1 olarak puanlananlar 4'e, 4 olarak puanlananlar ise 1'e dönüştürülmektedir. Olumsuz duyguları dile getiren maddelerde 4 değerindeki yanıtlar kaygının yüksekliğini göstermektedir. Olumlu duygulara verilen 4 değerindeki yanıtlar düşük, 1 değerindeki yanıtlar yüksek kaygıyı göstermektedir. Sürekli Kaygı Ölçeği'nde olumlu duyguları belirten 7 madde (21, 26, 27, 30, 33, 36 ve 39. maddeler) ters puanlanmaktadır. Ölçeğin puanlanmasında olumlu ve olumsuz maddelerin her biri için iki ayrı anahtar hazırlanmaktadır. Olumsuz maddeler için elde edilen toplam ağırlıklı puandan olumlu maddelerin ters puanlanmış toplam ağırlıklı puanı çıkarılmaktadır. Bu sayıya önceden tespit edilmiş ve değişmeyen bir değer (35 puan) eklenmektedir. En son elde edilen değer bireyin sürekli kaygı puanıdır. Ölçekten elde edilen toplam puan değeri, 20 ile 80 arasında değişmektedir. Puanın yüksek olması kaygı düzeyinin yüksek olduğuna işaret etmektedir.

Öner tarafından 1977 yılında geçerlik ve güvenirlik çalışmaları yapılan Sürekli Kaygı Ölçeği'nin Cronbach alfa iç tutarlık katsayısı .85 olarak belirlenmiştir. Sürekli Kaygı Ölçeği'nin diğer kaygı ölçekleri ile psikiyatrik hasta ve hasta olmayan gruplar üzerinde yapılan benzer ölçekler geçerliği incelenmiştir. Ölçeğin yapı geçerliği normal ve hasta gruplar üzerinde "iki faktörlü kaygı kuramı" çerçevesinde deneysel olarak incelenmiştir. Önemli ve stres yaratıcı olaylarda önce ve sonra sürekli kaygı puanlarının değişiminin anlamlı düzeyde olmaması, kuramdan çıkarılan hipotezleri desteklemiş ve ölçeğin yapı geçerliğinin göstergesi olarak değerlendirilmiştir. Öner ve Le Compte'nin (1983) 10 gün ile 1 yıl arasında değişen zamanlarda yaptıkları yineleme uygulamaları

sonuçları; farklı koşullarda, kişilerin, sürekli kaygı puanlarında anlamlı değişimler olmadığını ortaya koymuştur. Ölçeğin nevrotik ve normal bireyleri birbirinden ayırt edebildiği ve Türk toplumunda geçerli olduğu, danışma psikiyatri ve ruh sağlığı merkezleri ile ilgili araştırmalarda kullanılabileceği sonucuna ulaşılmıştır (Öner, 1997).

ROC analiziyle kesme puanı belirlenirken yüksek/düşük kaygılı öğrencileri belirleyen başka bir ölçeğin kullanılması gerekmektedir. Araştırmada Sınav Kaygısı Ölçeği, ROC analizi için ölçüt ölçek olarak kullanılmıştır. Bu ölçeğin tercih edilmesinin nedeni; araştırma grubunun kamu personeli seçme sınavı döneminde olmaları nedeniyle kaygı durumlarının gerçeği yansıttığı varsayımı ve on dakikada cevaplanması nedeniyle öğrencilerin sıkılmayacakları uzunlukta olmasıdır. Ayrıca Öner (1997) Sürekli Kaygı Ölçeği puanları ile Sınav Kaygısı Ölçeği'nin alt test puanları arasındaki ilişkiyi hesaplamıştır. Duyuşsallık ve kuruntu alt test puanları ile Sürekli Kaygı Ölçeği puanları arasındaki ilişkinin pozitif ve orta düzeyde ($r=0.45-0.60$) olması ölçeklerin araştırmada kullanılma nedenlerindedir.

2.3.1.2. Sınav Kaygısı Ölçeği

Spielberger (1970) tarafından geliştirilen ölçek 1983 yılında Öner ve Le Compte tarafından Türkçeye uyarlanmıştır. Sınav ve sınanmayla ilgili olumsuz duygu ve düşünceleri belirleyen bu ölçek ilkokul 4. sınıftan itibaren tüm bireylere uygulanabilmektedir. 20 maddeden oluşan dört dereceli Likert tipi ölçek iki faktörlüdür. Ölçeğin birinci faktörünün 12 (1., 6., 7., 9., 10., 11., 13., 14., 15., 16., 18., 19.maddeler), ikinci faktörünün 8 (2., 3., 4., 5., 8., 12., 17., 20. maddeler) maddeden oluştuğu sonucuna ulaşılmıştır. İlk faktöre “duyuşsallık” ikinci faktöre “kuruntu” ismi verilmiştir. Bireyler ölçekteki maddelere genel olarak hissettiği duyguların sıklık derecesine göre yanıt vermektedirler. Bireyler maddeleri kaygı düzeyleri derecesine göre, (1) hiçbir zaman, (2) bazen, (3) sık sık, (4) her zaman gibi seçeneklerden birini işaretleyerek yaklaşık 10 dakikada cevaplandırmaktadırlar. Ölçeğin ilk maddesi ters puanlanmaktadır. Ölçekten kuruntu, duyuşsallık ve toplam test puanı olmak üzere üç farklı puan elde edilmektedir. Ölçekten alınabilecek toplam puan 20 ile 80 arasında değişmektedir.

Ölçekten elde edilen puanlar, ham puanları oluşturmaktadır. Bu puanların standart puan karşılıkları norm tablosuna bakılarak belirlenmektedir. Norm tablosuna göre T-puanı 50'nin altında olan bireyler, düşük kaygılı kabul edilmektedir. Bu

araştırmada ROC analizinin yapılabilmesi için ölçüt kabul edilen değişkenin iki kategorili sınıflama ölçeği şeklinde olması gerekmektedir. Araştırma grubu üniversite öğrencilerinden oluştuğundan norm tablosuna göre Sınav Kaygısı Ölçeği'nden elde edilen toplam puan, 44'e eşitse veya 44'ten küçükse düşük kaygı; 44'ten büyük ise yüksek kaygı durumunu göstermektedir.

Öner (1997), dört dereceli Likert tipi Sınav Kaygısı Ölçeği'nin Cronbach alfa iç tutarlılık katsayısının .93 olduğunu belirtmiştir. Elde edilen iç tutarlılık ölçüsü olan Cronbach alfa katsayısı oldukça yüksektir. Bu sonuç, ölçeğin güvenle kullanılabileceği anlamına gelmektedir. Sınav Kaygısı Ölçeği'nin faktör yapısını belirlemek, ayrıca İngilizce formuna benzeyip benzemediğini görmek amacıyla faktör analizi yapılmıştır. Ölçeğin faktör desenini ortaya koymak amacıyla faktörleştirme yöntemi olarak temel bileşenler analizi; döndürme yöntemi olarak da dik döndürme yöntemlerinden maksimum değişkenlik (varimax) yöntemi seçilmiştir. Türkçe ölçek İngilizce formunda olduğu gibi iki boyutludur. Ölçeğin Türkçe formunda da aynı on iki madde birinci faktörde, sekiz madde ikinci faktörde toplanmıştır. Ölçeğin alt boyutlarına ilişkin Cronbach alfa katsayısı 1. faktör için .90, 2. faktör için ise .70 olarak hesaplanmıştır. Bu bulgu, ölçeğin alt boyutlarında madde sayısı azaldıkça güvenilirlik katsayısının da azaldığı şeklinde yorumlanabilir.

2.4. Verilerin Analizi

Araştırmanın birinci alt amacı olan Sürekli Kaygı Ölçeği'nin Angoff ve Sınır Grup yöntemleriyle elde edilen kesme puanının belirlenmesi için yargıcıların birinci ve ikinci uygulamada ölçekten elde ettikleri puanlar kullanılmıştır. Öğrenci ve test merkezli bu yöntemlere göre kesme puanı değerine karar vermek için yargıcıların ölçek puanlarının ortalaması ve ortancası hesaplanmıştır. Ölçeğin ROC analizi ile elde edilen kesme puanının belirlenmesi için öğrencilerin, ölçüt kabul edilen Sınav Kaygısı Ölçeği'nden aldıkları puanlar kullanılmıştır. Araştırma grubundaki her bir öğrencinin puanı, ölçeğin el kitabında belirtilen norm tablosuna göre sınıflandırılmıştır. Puanlar, norm tablosundaki puan değerinden düşükse “düşük kaygılı”, yüksekse “yüksek kaygılı” olarak 1-0 şeklinde kodlanmıştır. ROC analizinde ölçeğin kesme puanı olarak hangi değer alınacağına karar verilirken duyarlılığın yüksek, 1-özgüllük değerinin düşük olduğu en iyi nokta belirlenmiştir. Bu noktanın gösterdiği değer, kaygı düzeyini en iyi ayıran kesme puanı olarak kabul edilmiştir.

Araştırmanın ikinci alt amacı, Angoff ve Sınır Grup yöntemleri ile belirlenen kesme puanlarına göre yüksek/düşük kaygılı kabul edilen öğrenci oranları arasında manidar bir farkın olup olmadığını belirlemektir. Sınıflama ölçeğinde olan iki değişken arasındaki farkın manidarlığı ki-kare bağımsızlık testi ile incelenmiştir. Bu test için; yöntemlere göre yüksek ve düşük kaygılı kabul edilen öğrenci sayılarına ve yüzdelere ait 2x2'lik kontincensi tablosu oluşturulmuştur. Bu tablo değişkenlerin (düşük/yüksek kaygı) yer aldığı satır ve sütunlardan oluşmaktadır. Araştırmada, değişkenlerin oluşturduğu satır sayısı "2" ve sütun sayısı "2" olduğundan 2x2'lik bir kontincensi tablosuna göre veriler analiz edilmiştir.

Araştırmanın üçüncü alt amacında, kesme puanları belirlenirken yargıcılar arası uyum katsayısının ne olduğuna yanıt aranmıştır. Yargıcılar arası uyum incelenirken yargıcıların yöntemlere göre her bir test maddesine verdikleri yanıtlar arasındaki sınıf içi korelasyon katsayısı (Intraclass Correlation Coefficient) hesaplanmıştır. Shrout ve Fleiss (1979) tarafından temelleri atılan sınıf içi korelasyon katsayısıyla ikiden fazla sayıdaki yargıcının yaptığı puanlamaların uyumu belirlenmektedir. Ateş, Öztuna ve Genç (2009), ölçme sonuçlarının sürekli olduğu ve yargıcılar arası uyumun değerlendirilmesinde en yaygın kullanılan güvenilirlik kanıtının, sınıf içi korelasyon katsayısının hesaplanmasıyla elde edileceğini belirtmişlerdir. Kesme puanlarının güvenilir olması, yargıcılar arasındaki tutarlılığa bağlıdır. Güvenirlik kanıtı olarak kabul edilen sınıf içi korelasyon katsayısı; gruplar arası varyansın, gruplar arası ve gruplar içi varyans toplamına oranıdır. Araştırmanın amacına ve ölçme sonuçlarının elde edilme biçimine bağlı olarak farklı sınıf içi korelasyon katsayısı türleri vardır. Shrout ve Fleiss (1979) sınıf içi korelasyon analizi için üç farklı model önermişlerdir. Araştırmacı güvenilirliği tek bir yargıcı için ya da yargıcıların ortalama puanlarına dayalı olarak hesaplar. Birden fazla yargıcı varsa veya birden fazla ölçme işlemi yapılmışsa bunların ortalamasının temel alınması önerilmektedir.

Araştırmanın dördüncü alt amacı olan Sürekli Kaygı Ölçeği'nden elde edilen puanlara göre yüksek ve düşük kaygılı olarak sınıflandırılan öğrencilerin gerçek durumlarını belirlerken Sınır Kaygısı Ölçeği'nden aldıkları puanlar ölçüt alınarak sınır değer belirleme yöntemlerinden ROC analizi yapılmıştır.

Araştırmanın beşinci alt amacı, ROC analizi ile Sürekli Kaygı Ölçeği'nin ölçüt bağıntılı geçerliğinin belirlenmesidir. Geçerlik kanıtı olarak pozitif ve negatif yordayıcı

değerler hesaplanmış ve yorumlanmıştır. Bu değerler için öncelikle öğrencilerin Sürekli Kaygı ve Sınav Kaygısı ölçeklerinden elde ettikleri puanlara göre 2x2'lik tabloda sınıflandırılma durumları belirlenmiştir. Ölçeğin norm tablosuna göre öğrencinin sınav kaygısı ölçek puanı 44'e eşitse veya 44'ten küçükse bireyler düşük sınav kaygısı gösterdiği için "negatif"; 44'ten büyükse yüksek kaygı gösterdiği için "pozitif" şeklinde kodlanmıştır. Diğer sınıflandırma ise Sürekli Kaygı Ölçeği'nin ROC analizi ile belirlenen kesme puanına göre yapılmıştır. 2x2'lik tablo, ölçeklerin kesme puanlarına göre yüksek ve düşük kaygılı kabul edilen öğrenci sayılarına ve yüzdelere ait bilgiler içermektedir. Tablodan elde edilen bilgilerden yararlanılarak geçerlik kanıtı olan pozitif ve negatif yordayıcı değerler eşitlikler yardımıyla hesaplanmıştır. Yordayıcı değerler, kesme puanı belirlenen ölçeğin, ölçüt ölçeğe göre öğrencilerin kaygı durumlarını hangi olasılıkla doğru belirlediğini göstermektedir. Pozitif yordayıcı değer; Sürekli Kaygı Ölçeği'nin hangi olasılık değeri ile gerçekte yüksek kaygılı öğrencileri doğru tespit ettiği sonucunu vermektedir. Negatif yordayıcı değer ise Sürekli Kaygı Ölçeği'nin hangi olasılık değeri ile gerçekte düşük kaygılı öğrencileri doğru tespit ettiğini belirlemektedir.

BÖLÜM 3

BULGULAR VE YORUMLAR

Bu bölümde araştırmanın amaçları doğrultusunda elde edilen bulgulara ve yorumlara yer verilmiştir. Bulguların sunulduğunda araştırmanın amacı doğrultusunda yanıt bulabilmek üzere ele alınan alt amaçların sırası izlenmiştir.

3.1. Angoff, ROC ve Sınır Grup Yöntemleri ile Kesme Puanı Belirleme

Araştırmanın birinci alt amacı olan farklı yöntemlere göre kesme puanının belirlenmesi için öncelikle test ve öğrenci merkezli yöntemler için yargıcıların kararlarına ihtiyaç duyulmuştur. Angoff ve Sınır Grup yöntemleri için 10 yargıcıdan, belirlenen yönergelerle Sürekli Kaygı Ölçeği maddeleri hakkında karar vermeleri istenmiştir. Yargıcılardan, her madde için yönergedeki kurallara uyarak katılma derecelerini gösteren seçeneklerden birini işaretlemeleri beklenmiştir. Yirmi maddelik ve dört dereceli Likert tipi ölçekten alınabilecek minimum puan 20 iken maksimum puan 80'dir. İlk uygulama, Sınır Grup yöntemi ile kesme puanı belirleyebilmek için yapılmıştır. Yargıcılar, "yüksek kaygının minimum düzeyinde" olan bir öğrenciyi temel alarak yani onun tepkilerini göz önünde bulundurarak ölçek maddelerini yanıtlamışlardır. Yargıcıların Sınır Grup yöntemine göre verdikleri kararlar EK 1'de sunulmuştur. İkinci uygulamada aynı ölçek farklı yönerge ile yargıcılara tekrar verilmiştir. Yargıcılar, her madde için hangi tepki derecesinin "yüksek kaygının minimum düzeyinde" olan öğrenciyi ayırt edeceğini belirlemişlerdir. İlk uygulamada yargıcılar öğrenciye odaklanırken ikinci uygulamada ölçek maddelerinin tepki derecelerine odaklanmışlardır. Yargıcıların Angoff yöntemine göre verdikleri kararlar EK 2'de sunulmuştur. Yargıcıların öğrenci ve test merkezli yöntemlere göre verdikleri kararlar doğrultusunda ölçek puanları oluşturulmuştur. Angoff ve Sınır Grup yöntemine göre kesme puanı belirlemek için yargıcıların ölçek puanlarının ortancası ve ortalaması hesaplanmıştır. Yargıcıların iki uygulamadan elde ettikleri ölçek puanları ve yöntemlere göre belirlenen kesme puanları Çizelge 5'te sunulmuştur.

Çizelge 5.

Angoff ve Sınır Grup Yöntemi ile Belirlenen Sürekli Kaygı Ölçeği Kesme Puanları

Yargıcı	Angoff	Sınır Grup
Y-1	45	38
Y-2	47	36
Y-3	48	38
Y-4	50	44
Y-5	45	43
Y-6	49	45
Y-7	49	49
Y-8	51	50
Y-9	47	49
Y-10	49	49
Kesme puanı	$\bar{X} = 48$	Ort =44.50

Çizelge 5 incelendiğinde, Sınır Grup yöntemine göre yargıcıların ölçek maddeleri hakkındaki kararları doğrultusunda elde edilen toplam puan 36 ile 50 arasında değişmektedir. Angoff yöntemine göre ise yargıcıların ölçek maddeleri hakkındaki kararları doğrultusunda elde edilen toplam puan 45 ile 51 arasında farklılaşmaktadır. Yüksek kaygının minimum düzeyini belirleyen yirmi karar incelendiğinde en düşük puan Sınır Grup yönteminde, en yüksek puan ise Angoff yönteminde elde edilmiştir. Çizelge 5'te de sunulduğu üzere yargıcıların Sınır Grup ve Angoff yöntemlerine göre ölçek maddelerine verdikleri yanıtlardan elde edilen toplam puanların birbirine yakın olduğu görülmektedir. 7 yargıcının (Y-1, Y-2, Y-3, Y-4, Y-5, Y-6, Y-8) ölçek puanları, Angoff yönteminde daha yüksektir. 1 yargıcının (Y-9) ölçek puanının Sınır Grup yönteminde daha yüksek olduğu, diğer 2 yargıcının (Y-7, Y-10) ise iki yöntemde de eşit puanlara sahip oldukları görülmektedir. Angoff yöntemi için kesme puan belirlenirken yargıcıların ölçek puanlarının ortalaması alınmıştır. Sürekli Kaygı Ölçeği'nin Angoff yöntemine göre değerlendirilmesi sonucu kesme puanı 48 olarak elde edilmiştir. Sınır Grup yöntemi için kesme puanı belirlenirken yargıcıların ölçek puanlarının ortancası hesaplanmıştır. Aynı ölçeğin Sınır Grup yöntemine göre yargıcıların değerlendirmesi sonucu elde edilen kesme puanı değeri 44.5 olarak belirlenmiştir. Yargıcıların öğrenci ve test merkezli yöntemlere göre verdikleri kararların farklılığı, ölçekle belirlenen yüksek kaygının minimum düzeyi puanlarının da farklı olmasına neden olmuştur.

ROC analizine göre kesme puanı belirlenirken üniversite öğrencilerinin Sürekli Kaygı Ölçeği'nden aldıkları puanların değerlendirilmesi için ölçüt ölçek verilerine ihtiyaç duyulmuştur. Ölçüt ölçek puanları, Sürekli Kaygı Ölçeği'nin ROC analizi ile kesme puanının belirlenmesi için düşük ve yüksek kaygılı olarak sınıflandırılan öğrencilerin gerçek durumlarını belirleyen referans olarak kabul edilmiştir. Öğrencilerin Sınav Kaygısı Ölçeği'nden aldıkları puanlar gerçek kaygı durumlarını ifade etmiştir. ROC analizinin yapılabilmesi için ölçüt kabul edilen puanın iki kategorili sınıflama ölçeğinde olması gerekmektedir. Dolayısıyla öğrencilerin Sınav Kaygısı Ölçeği'nden aldıkları puanlar, ölçeğin norm tablosunda belirtilen puan dikkate alınarak sınıflama ölçeği düzeyine dönüştürülmüştür. Üniversite öğrencileri için Sınav Kaygısı Ölçeği'nin norm tablosuna göre ölçek puanı 44'e eşitse veya 44'ten küçükse bireyler düşük sınav kaygısı gösterdiği için "SK düşük=0", 44'ten büyükse yüksek kaygı gösterdiği için "SK yüksek=1" şeklinde kodlanmıştır. Araştırma grubunun Sınav Kaygısı Ölçeği'nden aldıkları puanların kaygı gruplarına göre sınıflandırılması ve dağılımı Çizelge 6'da sunulmuştur.

Çizelge 6.

<i>Araştırma Grubu Kaygı Düzeylerinin Sınav Kaygısı Ölçeği'ne Göre Sınıflandırılması</i>		
Araştırma Grubu	N	%
Pozitif Grup (SK yüksek=1)	91	31.40
Negatif Grup (SK düşük=0)	199	68.60
Toplam	290	100.00

Çizelge 6 incelendiğinde, araştırma grubu 290 pozitif ve negatif kaygılı öğrenciden oluşmaktadır. Bu öğrencilerin %31.4'ü (N=91) ölçüt kabul edilen Sınav Kaygısı Ölçeği'nden 44'ten yüksek puan alarak "yüksek kaygılı" grupta (SK yüksek=1), %68.6'sı (N=199) ise ölçekten 44'e eşit veya 44'ten küçük puanlar alarak "düşük kaygılı" grupta (SK düşük=0) yer almaktadır.

Bireylerin sınav kaygısı puanları ölçüt alınarak Sürekli Kaygı Ölçeği'nin kesme puanını belirlemek amacıyla işlem karakteristik eğrisi (ROC) oluşturulmuştur. ROC eğrisi oluşturulurken duyarlılık ve özgüllük değerleri kullanılmaktadır. Bu ölçek için kesme puanı olarak hangi değer alınmalı sorusunun yanıtını vermek amacıyla analiz sonucunda verilen her bir duyarlılık ve özgüllük değerleri incelenmiş ve en uygun nokta seçilmiştir. Yani duyarlılığın yüksek ve özgüllüğün düşük olduğu, iki değer arasındaki

farkın en yükseğe ulaştığı nokta belirlenmiştir. ROC analiziyle belirlenen kesme puanı için koordinatlar Çizelge 7’de sunulmuştur.

Çizelge 7.

Sürekli Kaygı Ölçeğine İlişkin ROC Analizi ile Belirlenen Kesme Puanı

Kesme Puan	Duyarlılık	Özgüllük	Kesme Puanı	Duyarlılık	Özgüllük
34.00	1.00	1.00	49.50	.593	.246
35.50	1.00	.995	50.50	.473	.176
37.00	1.00	.990	51.50	.407	.121
38.50	1.00	.925	52.50	.341	.090
39.5	.989	.925	53.50	.242	.060
40.50	.978	.859	54.50	.187	.045
41.50	.956	.819	55.50	.143	.035
42.50	.923	.769	56.50	.099	.015
43.50	.912	.744	58.00	.088	.010
44.50	.890	.658	59.50	.055	.005
45.50	.857	.573	60.50	.033	.005
46.50	.813	.472	62.00	.011	.000
47.50	.725	.402	64.00	.000	.000
48.50	.716	.366			

Çizelge 7’ye göre kesme puanı 48.5 olarak seçilirse bu puana karşılık gelen duyarlılık değerinin %72 ve özgüllük değerinin %37 olduğu görülmektedir. Bu kesme puanı, duyarlılık ve özgüllük değerleri arasındaki farkın (716-366=350) en yüksek olduğu noktaya karşılık gelmektedir. Bu değer kesme puanı olarak belirlenmesinin nedeni, ölçek için duyarlılığın ve özgüllüğün her ikisinde de aynı anda en iyi noktaya işaret etmesidir.

Üniversite öğrencilerinin Sürekli Kaygı Ölçeği’nden aldıkları puanların değerlendirilmesi için Angoff yöntemiyle 48, Sınır Grup yöntemi ile 44.5, ROC analizine göre ise 48.5 kesme puanı değerleri belirlenmiştir. Kesme puanları incelendiğinde sınır değer saptama yöntemlerinden ROC analizi ile en yüksek değer, standart belirleme yöntemlerinden Sınır Grup ile de en düşük değer elde edildiği sonucuna ulaşılmıştır. Kesme puanlarının farklı olması, Wang, Pan ve Austin’in (2003) aynı sınıflama grubunda yer alan yöntemlerin bile kendi içlerinde oldukça fazla farklılaştığını belirttikleri çalışmanın bulguları ile örtüşmektedir. Sürekli Kaygı Ölçeği için belirlenen kesme puanlarının farklı olmasının nedeni, üç yöntemde de verilen kararların farklı dayanaklarının olmasına bağlanabilir. Standart belirleme yöntemlerinde

yargıcıların ölçek maddelerini cevaplandırırken öğrenci ve madde odaklı düşünceleri istenmiştir. Yargıcılar Sınır Grup yönteminde kendilerini “yüksek kaygının minimum düzeyindeki” öğrencinin yerine koyarak ölçeği yanıtlamışlardır. Angoff yönteminde ise yargıcılar, her bir maddenin hangi tepki derecesinin yüksek kaygının minimum düzeyindeki öğrenciyi belirleyeceğine karar vermişlerdir. ROC analizi ile kesme puanı oluşturulurken ise öğrencilerin ölçüt ölçek puanları odak noktasıdır ve o ölçeğe göre kesme puanı belirlenmiştir. Standart belirleme yöntemlerinden Angoff ile sınır değer saptama yöntemlerinden ROC analizi sonucu elde edilen kesme puanları birbirine çok yakındır. Test merkezli yöntemler sınıfında bulunan Angoff yönteminde yargıcılar, madde odaklı kararlar verdiklerinden maddelerin belirleyiciliğini göz önünde bulundurmuşlardır. ROC analizinde ise ölçüt ölçek kullanılarak ölçekler arası hesaplanan belirli indekslerle asıl ölçeğin kesme puanı belirlenmektedir. Dolayısıyla kesme puanlarının Angoff ve ROC analizinde çok yakın çıkmasının nedeni yöntemlerin odak noktalarının benzer olmasına bağlanabilir. Livingston ve Zeiky (1989) standart belirleme yöntemlerine yönelik eleştirilerde bulunurken yüksek kaygının minimum düzeyindeki bireyin her bir yargıcının beyninde farklı şekilleneceğini iddia etmişlerdir. Uygulamalar öncesi araştırmacı tarafından yapılan toplantılarda yargıcılar arasında “yüksek kaygının minimum düzeyi” olan birey hakkında ortak bir bakış açısı oluşturulmaya çalışılmıştır. Ancak yargıcıların ölçeği cevaplandırmaları sırasında bunu dikkate almadıkları düşünülmektedir. Ayrıca yargıcıların kesme puanı belirleme yöntemleri ile ilgili bilgileri bu araştırma ile oluşmuş ve cevapları, beklendiği gibi ayırtırmada zorlanmışlardır. Tavakol ve Dennick (2012) yaptıkları araştırmada, çoktan seçmeli bir testin standart belirleme yöntemleri ile oluşturulmuş kesme puanını, ROC analizi ile tekrar hesaplamışlardır. Farklı kesme puanlarının hesaplandığı çalışmada, ROC analizinin standart belirleme yöntemlerinde olmayan duyarlılık ve özgüllük değerlerini dikkate aldığı için daha objektif kesme puanı belirlediği sonucuna ulaşılmıştır. Çoktan seçmeli testlerde cevaplar 1-0 şeklinde iken bu araştırmada kullanılan çoklu puanlanan ölçeğin cevapları derecelidir. Yargıcıların bilişsel alana yönelik testlerdeki cevapları daha kesin iken duyuşsal özellikleri ölçen ölçeklerde cevapları daha çok farklılaşmaktadır. Dolayısıyla çoklu puanlanan bir ölçme aracının kesme puanı belirlenirken hangi yöntemin kullanılacağına karar vermede birçok değişken göz önünde bulundurulmalıdır. Yöntemlerin öğrenci ve madde odaklı olması ve yargıcıların yöntemlerle ilk kez karşılaşılıyor olmaları kesme puanlarının farklı olmasına neden olmaktadır. Kesme puanı belirlenirken uygun ölçüt ölçeğin varlığı ve

ölçüt ölçeğin kesme puanının olması da ROC analizinin tercih edilmesi için gereklidir. ROC analizinin diğer yöntemlerde olmayan duyarlılık ve özgüllük değerleri araştırmacılara ölçek hakkında daha fazla bilgi sunmaktadır.

3.2. Angoff ve Sınır Grup Yöntemleri ile Belirlenen Kesme Puanlarına Göre Düşük/Yüksek Kaygılı Kabul Edilen Öğrencilerin Yüzdeleri Arasındaki Fark

Araştırmanın ikinci alt amacına yanıt bulabilmek için 10 yargıcının 20 ölçek maddesi hakkındaki kararları sonucunda kesme puanları oluşturulmuştur. Yargıcıların iki yöntem için maddelere verdikleri cevapların farklılığı kesme puanlarının da farklı olmasına neden olmuştur. Ayrıca yöntemlere göre yargıcıların puanlarının ortalaması ya da ortancası hesaplandığından kesme puan değerleri farklıdır. Angoff yöntemine göre her yargıcının testte bulunan bütün maddelere ilişkin sonuçları toplanarak o yargıcıya ait minimum test puanı hesaplanmıştır. Yargıcıların minimum test puanları ortalaması, bu yöntem için nihai kesme puanı olmuştur. Sınır Grup yönteminde ise aşırı yüksek ve düşük ölçek puanlarından etkilenmemek adına yargıcıların test puanlarının ortancası hesaplanmıştır. Sürekli Kaygı Ölçeği'nin kesme puanları Angoff yöntemiyle 48, Sınır Grup yöntemi ile 44.5 olarak hesaplanmıştır. Belirlenen kesme puanlarına göre öğrencilerin Sürekli Kaygı Ölçeği puanları sınıflama ölçeği düzeyine dönüştürülmüştür. Yöntemlere göre elde edilen kesme puanlarındaki farklılık, yüksek ve düşük kaygılı kabul edilen öğrenci sayılarını ve yüzdelerini de etkilemiştir. Angoff ve Sınır Grup yöntemleriyle elde edilen kesme puanı değerine göre düşük ve yüksek kaygılı kabul edilen öğrencilere ilişkin bilgiler Çizelge 8'de sunulmuştur.

Çizelge 8.

Farklı Yöntemlerle Belirlenen Kesme Puanları, Düşük/Yüksek Kaygılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri

Yöntem	Kesme Puanı	Yüksek Kaygılı Öğrenciler		Düşük Kaygılı Öğrenciler	
		N	%	N	%
Angoff	48.00	117	40.30	173	59.70
Sınır Grup	44.50	192	66.20	98	33.80

Çizelge 8'de görüldüğü gibi belirlenen kesme puanlarına göre araştırma grubunu oluşturan 290 öğrencinin sınıflandırılma oranları farklılık göstermektedir. Öğrencilerin %40.3'ü (N=117) Sürekli Kaygı Ölçeği'nin Angoff yöntemi ile belirlenen kesme puanından yüksek puan almıştır. Öğrencilerin %66.2'si (N=192) ise Sürekli Kaygı Ölçeği'nin Sınır Grup yöntemi ile belirlenen kesme puanından yüksek puan almıştır.

Düşük kesme puanının belirlendiği Sınır Grup yönteminde yüksek kaygılı kabul edilen öğrenci sayısının ve yüzdesinin fazla olduğu görülmektedir. Düşük kaygılı kabul edilen öğrenci sayısı ve yüzdesinin fazlalığı ise daha yüksek kesme puanının hesaplandığı Angoff yöntemine aittir. Yöntemlere ait yüksek kaygılı kabul edilen öğrenci yüzdeleri arasındaki bu farklılığın manidarlığı ki-kare bağımsızlık testi ile incelenmiştir. Bu test için kesme puanlarına göre yüksek ve düşük kaygılı kabul edilen öğrenci sayılarına ve yüzdelere ait 2x2'lik kontincensi tablosu oluşturulmuştur. Ki-kare bağımsızlık testi sonuçları Çizelge 9'da sunulmuştur.

Çizelge 9.

Yöntemlere Göre Farklılaşan Düşük/Yüksek Kaygılı Öğrenci Oranlarının Manidarlığı

Yöntem	Sınır Grup			χ^2	sd	p		
	Düşük	Yüksek	Toplam					
Angoff	Düşük	N	98	75	173	100.11	1	.00
		%	57	43	100			
	Yüksek	N	0	117	117			
		%	0	100	100			
	Toplam	N	98	192	290			
		%	34	66	100			

Çizelge 9'da yöntemlere ait kesme puanlarına göre düşük/yüksek kaygılı öğrenci yüzdeleri ikili olarak karşılaştırıldığında gözlenen bu farkın manidar olduğu sonucuna ulaşılmıştır ($\chi^2_{(sd=1)}=100.11, p<.05$). Düşük ve yüksek kaygılı öğrenci yüzdeleri arasındaki manidar fark, yöntemlere göre belirlenen kesme puanlarının farklı olmasından kaynaklanmaktadır. Test ve öğrenci merkezli yöntemlere göre yargıcıların ölçek maddelerine verdikleri cevaplar değişebilmekte, dolayısıyla yargıcılardan elde edilen ölçek puanları farklılaşabilmektedir. Yargıcı puanlarının ortalama ve ortanca değerlerine göre belirlenen kesme puanı, Angoff yönteminde daha yüksek; Sınır Grup yönteminde ise daha düşüktür. Dolayısıyla bu kesme puanlarına göre bireylerin düşük/yüksek kaygılı sınıflandırılma oranları farklılaşmıştır. Yüzdeler arasındaki bu fark da manidar bulunmuştur. Test merkezli yöntemlerden Angoff ile ölçeğin kesme puanına karar verildiğinde araştırma grubundan az sayıda birey yüksek kaygı düzeyindedir. Öğrenci merkezli yöntemlerden Sınır Grup ile ölçeğin kesme puanına karar verildiğinde ise araştırma grubundan çok sayıda birey yüksek kaygı düzeyindedir. Tülübaş'ın (2009) Beck Depresyon Ölçeği için yaptığı çalışmanın bulgularında da Sınır Grup yönteminin Angoff yöntemine göre daha fazla öğrenciyi depresyonlu olarak gruplandığı ve yüzdeler arasındaki farkın manidar olduğu sonucuna ulaşılmıştır.

Angoff yönteminde yargıcıların yüksek kaygının minimum düzeyindeki öğrencileri ve aynı zamanda maddelerin ayırt ediciliğini düşünmeleri istenmektedir. Sınır Grup yönteminde yargıcılar ölçeği daha genel, Angoff yönteminde ise madde bazında daha ayrıntılı incelemektedir. Dolayısıyla Impara ve Plake'nin (1998) de belirttiği gibi yargıcılar Angoff yöntemini anlamada ve bu yöneme göre karar vermede zorlanmış olabilirler. Test ve öğrenci merkezli yöntemlerle belirlenen kesme puanlarına göre sınıflandırılan öğrenci oranları arasındaki manidar farklılık, Tanrıverdi'nin (2006) başarı testleri ile yaptığı çalışmasının bulguları ile örtüşmektedir. Çalışmada aynı ders için farklı yöntemlerle belirlenen kesme puanları farklılaşmış ve başarılı öğrencilerin yüzdeleri arasındaki fark manidar bulunmuştur. Bu bulgular yorumlandığında öğrenci yüzdeleri arasındaki manidar farklılıkların nedeni her yöneme göre yargıcıların testlerdeki odak noktalarının değişmesi ve kesme puanlarının hesaplanmasının farklılaşması olarak görülebilir. Ayrıca Çetin ve Gelbal (2010), farklı yöntemlere göre belirlenen kesme puanlarına göre geçti/kaldı öğrenci yüzdelerini karşılaştırmışlardır. Başarı testlerinde derslerin sayısal ve sözel içerikli olması ile geçme puanlarının belirlenmesi arasında manidar fark bulunmuştur. Türkçe dersi için yöntemler arasındaki yüzdeler arasında .05 düzeyinde manidar bir fark olduğu, matematik dersi için yöntemler arasındaki yüzdeler arasında .01 düzeyinde manidar bir fark olmadığı sonucuna ulaşmışlardır. Derslerin sayısal ya da sözel içerikli olması, başarılı kabul edilen öğrenci yüzdeleri karşılaştırıldığında manidar farklılığın nedenleri olarak görülmüştür.

Yargıcıların yöntemlere göre verdikleri kararlar sayısal cevap içeren maddelerde daha az değişiklik göstermekte dolayısıyla kesme puanları birbirine yakın değerler almaktadır. Ancak sözel alanlarda yargıcıların maddeler hakkındaki kararları daha çok farklılaşmakta ve kesme puanı değerleri değişmektedir. Sözel alan çoktan seçmeli testlerde veya duyuşsal alan içerikli ölçeklerde öğrenci yüzdeleri arasındaki manidar farklılığın nedeni, yargıcıların yöntemlere göre kesme puanı değerlerinin daha çok farklılaşmasıdır. Bu farklılaşma kaygı gibi özelliklerin ölçüldüğü tipik tepki testlerinde doğru cevabın bulunmaması, test alan kişinin testten alınabilecek maksimum puana güdülenmemiş olmasından kaynaklanabilir. Yani kaygı ölçeği ile ilgili yapılan bu çalışmada öğrenci yüzdeleri arasındaki manidar farklılığın nedenleri ölçek içeriği ve yargıcı kaynaklı olabilir. Bu değişkenlik yargıcıların minimum düzeydeki bireyi düşünürken zorlanmış olmalarından, kaygı ile ilgili yaptıkları çalışmaların

farklılaşmasından, teorikte ve uygulamada kararlarının değişkenlik gösterebilmesinden kaynaklanabilir. Standart belirleme yöntemlerinden hangisinin seçileceği konusunun, öğrenci ve yargıcı grubuna ve ölçülen özelliğin yapısına bağlı olarak değişebileceği söylenebilir.

3.3. Kesme Puanlarının Belirlenmesinde Yargıcılar Arası Uyum Katsayısı

Araştırmanın üçüncü alt amacına yanıt bulabilmek için öncelikle yargıcıların yöntemlere ait Sürekli Kaygı Ölçeği puanlarının normal dağılım gösterip göstermediği incelenmiştir. Bu amaçla yargıcıların ölçek puanlarına ait betimsel istatistikler ve normallik testi sonuçları Çizelge 10’da sunulmuştur.

Çizelge 10.

Yargıcıların Ölçek Puanlarına İlişkin Betimsel İstatistikler ve Normallik Testleri

Betimsel İstatistikler	Angoff Yöntemi	Sınır Grup Yöntemi
N	10.00	10.00
Ortalama	48.00	44.10
Ortanca	49.00	44.50
Tepe Değer	49.00	49.00
Standart Sapma	1.99	5.26
Varyans	3.96	27.66
Çarpıklık Katsayısı	-0.66	-0.49
Basıklık Katsayısı	-0.36	-1.48
Shapiro Wilk’s İstatistiği	0.89	0.18
Shapiro Wilk’s Sig. (p)	0.88	0.12

Çizelge 10 incelendiğinde, yargıcıların Angoff ve Sınır Grup yöntemlerine göre Sürekli Kaygı Ölçeği puanlarına ilişkin betimsel istatistiklerde; tepe değer, ortanca ve ortalama değerlerinin birbirine yakın olduğu belirlenmiştir. Standart normal dağılımda basıklık ve çarpıklık katsayı değerlerinin ± 1 aralığında kalması, dağılımın normalden aşırı sapma göstermediğinin bir kanıtı olarak değerlendirilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Test ve öğrenci merkezli yöntemlere göre Çizelge 10 değerleri incelendiğinde çarpıklık katsayılarının (-0.49 ve -0.66) kabul edebilir ± 1 aralığında olduğu, bu nedenle dağılımların çarpık olmadığı söylenebilir. Yine Çizelge 10 incelendiğinde ölçek puanlarına ilişkin basıklık katsayılarının (-1.48 ve -0.36) ± 1 aralığından aşırı sapma göstermediği ve dağılımların basık olmadığı söylenebilir. Bu bulgulara göre puanların dağılımının normale yakın olduğu söylenebilir. Büyüköztürk (2012), Shapiro Wilk’s testinin, veri sayısının 50’den az olduğu durumlarda kullanılan

bir normallik testi olduğunu belirtmiştir. Shapiro Wilk's testi için hesaplanan p-değerinin $\alpha=.05$ 'ten büyük çıkması, bu anlamlılık düzeyinde puanların normal dağılımdan anlamlı (aşırı) sapma olmadığını göstermektedir. Çizelge 10'daki bulgulardan da görülebileceği gibi, yöntemlere ait ölçek puanlarının normal dağılım gösterdiği söylenebilir.

Yargıcıların ölçek puanlarına ilişkin betimsel istatistikler ve normallik testi sonuçları, kararlar arası uyumu belirlemek için parametrik yöntemlerin kullanılabilirliğini göstermektedir. Araştırmanın üçüncü alt amacı olan yöntemlere göre yargıcılar arası uyum katsayısını belirlemek için sınıf içi korelasyon katsayısı analizi yapılmıştır. Analiz sonuçları Çizelge 11'de sunulmuştur.

Çizelge 11.

<i>Yöntemlere Göre Kesme Puan Belirlemede Yargıcılar Arası Uyum Katsayısı Sonuçları</i>			
Yöntem	Yargıcı Sayısı	Sınıf içi Korelasyon Katsayısı	p
Sınır Grup	10	.53	.006
Angoff	10	.65	.000

Üç ya da daha fazla yargıcının aynı ölçme aracına yönelik yaptıkları puanlama uyumunun incelendiği bu yöntemde sınıf içi korelasyon katsayısının alabileceği en büyük değer 1'dir. Shrout ve Fleiss'e (1979) göre elde edilen korelasyon katsayısı; $r < .40$ ise zayıf, $r = .41-.59$ arası orta, $r = .60-.74$ iyi ve $r > .75$ ise mükemmel uyum olduğunu göstermektedir. Çizelge 11'e göre çoklu puanlanan bir ölçeğin sınır grup yöntemine göre kesme puanı belirlendiğinde sınıf içi korelasyon katsayısının ($r = .53$) orta, Angoff yöntemine göre ise ($r = .65$) iyi düzeyde olduğu sonucuna ulaşılmıştır. Bu durumda yargıcılar arasında iki yöntemde de kabul edilebilir bir uyum olduğu söylenebilir. Hesaplanan korelasyon katsayıları, Angoff yöntemindeki yargıcılar arası uyumun Sınır Grup yöntemindeki yargıcılar arası uyumdan daha yüksek olduğunu göstermektedir. Garson (2003), sınıf içi korelasyon katsayısı yüksekse yargıcı grubun içinde değişkenliğin olmadığı ancak grup ortalamaları arasında farkın olduğu şeklinde yorumlanabileceğini belirtmektedir. Dolayısıyla Sınır Grup yönteminde yargıcı grup içindeki değişkenlik, Angoff yöntemine göre daha fazladır. Sınır Grup yöntemine göre kesme puanı hesaplanırken yargıcı ölçek puanlarının ortanca değerinin alınması bu bulguyu doğrulamaktadır. Taşdemir'in (2013) çalışmasında da yargıcılar arası uyum katsayısı sınıf içi korelasyon katsayısı ile hesaplanmış, aynı yöntem için sözel ve sayısal

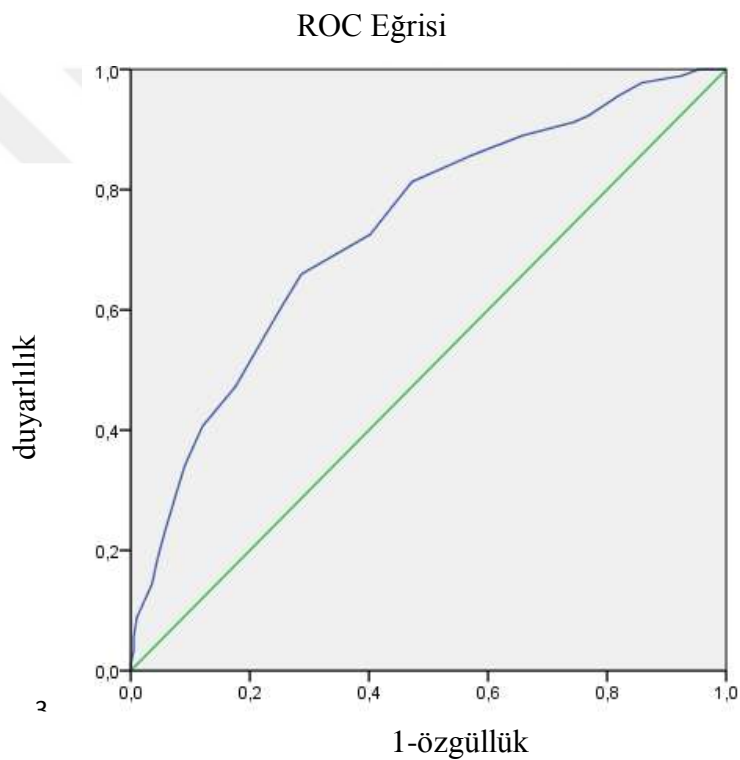
derslere göre farklı düzeylerde uyum olduğu sonucuna ulaşılmıştır. Sınıf içi korelasyon katsayısının orta ve iyi düzeyde bulunması, yargıcıların farklı düzeyde alan bilgisine sahip olmalarından, standart belirleme yöntemlerini tam olarak anlayamamalarından veya tepki düzeylerini net olarak ifade edememiş olmalarından kaynaklanabilir.

Yargıcılar arası uyum katsayısının farklı düzeylerde bulunması, yöntemlerin öğrenci ve madde odaklı olmaları, ölçülen değişkenin duyuşsal bir özellik olması ile ilişkilendirilebilir. Yargıcılar, Sınır Grup yönteminde “yüksek kaygının minimum düzeyinde” olan bir öğrenciyi temel alarak yani onun tepkilerini göz önünde bulundurarak ölçek maddelerini yanıtlamışlardır. Dolayısıyla yargıcıların kararlarında üniversite son sınıf öğrencileri ile empati kurabilme düzeyleri de etkili olmuştur. Kaygının tanımından yola çıkıldığında evrensel bir olgu olduğu ve kültür, cinsiyet, yaş ayrımı yapmaksızın her bireyin yaşadığı mutsuz edici bir duygu olduğu söylenebilir. Ancak kaygının düzeyleri içinde bulunulan ortama ve koşullara göre farklılaşmaktadır. Dayhoff (2000); kaygıların oluşmasında geçmişte tecrübe edilen olumsuz yaşantılar, bireyin biyolojik durumu ve içinde bulunulan sosyal yaşantılardan söz edilebileceğini belirtmiştir. Araştırmanın çalışma grubunu KPSS’ye hazırlanan üniversite son sınıf öğrencileri oluşturduğundan, farklı kaygı düzeylerine sahip bireylerin varlığı gözlenmiştir. Üniversite öğrencileri, özellikle geleceğe yönelik planların yoğunlaştığı dönem olan son sınıfta farklı kaygı düzeyleri yaşayabilmektedir. Eröz (2004) sınavlarda başarılı olma ile gelecekte olması istenenlerin gerçekleşmesi bağlamında öğrencilerin gerilim yaşadığını belirtmiş ve başarılı olma arzusu ile bu gerilimin kaygıya dönüştüğünü ifade etmiştir. Öğrencilerin tepkilerini göz önünde bulundurarak ölçek maddeleri hakkında karar veren yargıcılar ise iş kaygısı yaşamayan alanında uzman akademisyenlerdir. Yargıcıların ölçek maddeleri hakkındaki cevapları yüksek kaygının minimum düzeyindeki bireyle empati kurabildiği kadar oluşmuştur. Bireyin kaygı düzeyinin oluşumunda birçok karıştırıcı değişkenin varlığı söz konusu olduğundan yargıcıların kararlarının farklılaşması beklenen bir bulgudur. Bu bağlamda Sınır Grup yönteminde yargıcılar arası uyum katsayısı Angoff yöntemine göre daha düşüktür. Test merkezli yöntemde ise yargıcılar, her madde için hangi tepki derecesinin “yüksek kaygının minimum düzeyinde” olan öğrenciyi ayırt edeceğini belirlemişlerdir. Yargıcılar bu yöntemde öğrenci ile empati kurmadan, hangi tepki derecesinin ayırt edici olabileceğine karar vermişlerdir. Dolayısıyla yargıcıların yüksek kaygının minimum düzeyindeki bireyi, yüksek kaygılı bireylerden ayıran yanıtları dikkate alarak daha çok

uzmanlık bilgileri ile karar verdikleri söylenebilir. Yargıcıların yüksek kaygının minimum düzeyindeki birey hakkında aldıkları ortak kararların bu yöntemde uyum katsayısının daha yüksek olmasına neden olduğu söylenebilir.

3.4. ROC Analiziyle Sürekli Kaygı Ölçeği'nden Elde Edilen Eğri Altında Kalan Alan Değerinin Düşük/Yüksek Kaygılı Öğrencileri Ayırmadaki Doğruluk Oranı

Bireylerin sınav kaygısı puanları ölçüt alınarak Sürekli Kaygı Ölçeği'nin kesme puanını belirlemek için işlem karakteristik eğrisi (ROC) oluşturulmuştur. Duyarlılık ve özgüllük değerleri kullanılarak oluşturulan ROC eğrisi Şekil 2'de yer almaktadır.



Şekil 2. Sürekli Kaygı Ölçeği'ne İlişkin ROC Eğrisi Altında Kalan Alan

Bu ölçek için kesme puan olarak hangi değer alınmalı sorusunun yanıtını vermek için analiz sonucunda Şekil 2'de verilen her bir duyarlılık ve özgüllük değerleri incelenmiş ve optimum nokta (duyarlılık=.716, özgüllük=.366) seçilmiştir. Eğri altında kalan alan ise 0 ile 1 arasında değişen değerler almaktadır. Eğri, (0, 1) noktasına ne kadar yaklaşırsa alan değeri o kadar büyümektedir. Eröz (2010) yaptığı çalışmada, ölçüt testin verilerinin sürekli ya da süreksiz olması durumunda eğri altında kalan alanın farklılığını incelemiştir. Araştırma sonucunda, verilerin süreksiz hâle dönüştürülmesi durumunda test için belirlenen kesim noktası alındığında ROC eğrisi altında kalan

alanın en büyük değere ulaştığını belirtmiştir. Bu araştırmada da ölçüt ölçek olan Sınav Kaygısı Ölçek puanları süresiz veriye dönüştürülmüştür. ROC analizine göre bireyleri, yüksek ve düşük kaygılı olarak sınıflandırabilmek için Şekil 2'deki eğri altında kalan alan incelenmiştir. Çizelge 12'de ROC eğrisi altında kalan alana ait değerler sunulmuştur.

Çizelge 12.

ROC Eğrisi Altında Kalan Alan Değeri

AUC	s	p	%95 güven aralığı alt sınır	%95 güven aralığı üst sınır
.734	.032	.000	.672	.796

Çizelge 12'ye göre ROC analizi sonucunda, işlem karakteristik eğrisi altında kalan alan, (AUC)=0.734 ve manidar ($p<0.05$) bulunmuştur. Hosmer, Lemeshow ve Sturdivant (2013), AUC değerinin düzeyleri ile ilgili yaptıkları sınıflamada bu değer ne kadar 1'e yakın olursa belirlenen kesme puanına göre yapılan sınıflamanın o derece doğru olacağını belirtmişlerdir. Bu araştırmada işlem karakteristik eğrisi altında kalan alan değeri 0.734 olarak hesaplanmıştır. AUC değerinin düzeylerine göre yapılan sınıflama temel alındığında 0.734 değeri $0.7 \leq AUC < 0.8$ aralığında olduğundan, belirlenen kesme puanının "kabul edilebilir" düzeyde sınıflama yapabileceği sonucuna ulaşılmıştır. Bu değer ROC analizine göre hesaplanan kesme puanıyla, Sürekli Kaygı Ölçeği'nin yüksek ve düşük kaygılı bireyleri %73 oranında (orta düzeyde) doğru sınıflandırabildiğini göstermektedir. Koyuncu'nun (2015) çoklu puanlanan farklı bir kaygı ölçeği için yaptığı ROC analizinde AUC değerini kabul edilebilir düzeyde bulması, bu araştırmanın bulguları ile örtüşmektedir. ROC analizi ile kesme puanı belirlenirken aynı zamanda eğri altında kalan alanın yorumlanabilmesi, doğru sınıflandırma yapmada ayrıca kanıt sağlamaktadır.

3.5. Sürekli Kaygı Ölçeği'nin ROC Analizi ile Geçerlik Düzeyi

Geçerlik belirlemenin bir kanıt toplama süreci olduğu düşünülürse değişen koşullara göre ölçme araçları yeniden incelenmelidir. Korkmaz (2007), daha önce geliştirilmiş olan bir ölçme aracının farklı kültürlerde ölçmeyi amaçladığı psikolojik özelliği ölçüp ölçmediğinin önemli olduğunu, psikometrik niteliklerinin incelenmesi gerektiğini belirtmektedir. Araştırmanın beşinci alt amacına yanıt bulabilmek için

gerçek durumu yansıttığı varsayılan ölçüt ölçekle ve Sürekli Kaygı Ölçeği ile öğrencilerden toplanan verilerden yararlanılmıştır. ROC analizi ile geçerlik kanıtı sunabilmek için öğrenciler, Sürekli Kaygı ve Sınav Kaygısı Ölçekleri göz önünde bulundurulurken kaygı durumlarına göre sınıflandırılmışlardır. ROC analizi ile Sürekli Kaygı Ölçeği'nin geçerlik düzeyini belirlemek için iki kaygı ölçeği ile toplanan verilerden negatif ve pozitif yordayıcı değerler, duyarlılık ve özgüllük değerleri hesaplanmış ve yorumlanmıştır.

Üniversite öğrencileri için Sınav Kaygısı Ölçeği'nin norm tablosuna göre ölçek puanı, 44'e eşitse veya 44'ten küçükse bireyler düşük sınav kaygısı gösterdiği için "negatif"; 44'ten büyükse yüksek kaygı gösterdiği için "pozitif" şeklinde kodlanmıştır. Üniversite öğrencileri için Sürekli Kaygı Ölçeği'nin ROC analizi ile belirlenen ölçek puanı 48.5'ten küçükse bireyler düşük kaygı gösterdikleri için "negatif", 48.5'ten büyükse yüksek kaygı gösterdikleri için "pozitif" şeklinde kodlanmıştır. Ölçeklere göre düşük ve yüksek kaygılı kabul edilen öğrenci sayıları ve yüzdeleri Çizelge 13'te sunulmuştur.

Çizelge 13.

Ölçeklere Göre Düşük/Yüksek Kaygılı Kabul Edilen Öğrenci Sayıları ve Yüzdeleri

		Sürekli Kaygı Ölçeği			
			Düşük Kaygı (Negatif)	Yüksek Kaygı (Pozitif)	Toplam
Sınav Kaygısı Ölçeği	Düşük Kaygı (Negatif)	N	142	57	199
	Yüksek Kaygı (Pozitif)	N	31	60	91
		N	173	117	290

Geçerlik kanıtı olarak pozitif ve negatif yordayıcı değerler, formüller yardımıyla Çizelge 13'teki değerlerden hesaplanarak yorumlanmıştır. Pozitif yordayıcı değer, Sürekli Kaygı Ölçeği'nin hangi olasılık değeri ile sınav kaygısının yüksek olduğu durumu doğru tespit ettiği sonucunu vermektedir. Negatif yordayıcı değer ise Sürekli Kaygı Ölçeği'nin hangi olasılık değeri ile sınav kaygısının düşük olduğu durumu doğru tespit ettiğini belirlemektedir. Sürekli Kaygı Ölçeği'nin ROC analizine göre kesme puanı 48.5 olarak hesaplanmıştır. Çizelge 13 incelendiğinde araştırma grubunda bulunan 117 öğrenci, Sürekli Kaygı Ölçeği'nden aldıkları puanlar 48.5 kesme puanından yüksek olduğu için sürekli kaygının yüksek olduğu pozitif gruptadır. 173

öğrenci ise 48.5 kesme puanından düşük puanlar aldıklarından, sürekli kaygının düşük olduğu negatif grupta yer almaktadır. Pozitif grupta yer alan 117 öğrenciden 60'ı doğru pozitif, 57'si ise yanlış pozitif grupta yer alırken; negatif grupta yer alan 173 öğrencinin 142'si doğru negatif, 31'i ise yanlış negatif grupta yer almaktadır. Uluç (2007) ölçülen özelliğin gerçekte var olduğu durumu belirlemek için özgüllük ve duyarlılık değerlerinin yeterli olmadığını, yordayıcı değerlerle ulaşılan bilginin desteklenmesi gerektiğini belirtmektedir. Yordayıcı değerler, ölçülen özelliğin görülme sıklığına göre değişmektedir. Özelliğin görülme sıklığı, araştırma grubundaki yüksek kaygılı ve düşük kaygılı birey sayısına bağlı olarak hesaplanmıştır. Çizelge 13'teki iki ölçeğe göre de yüksek kaygı düzeyinde bulunan birey sayısının Sürekli Kaygı Ölçeği'ne göre yüksek kaygılı birey sayısına oranından hesaplanan pozitif yordayıcı değer ($PYD=60/117$) 0.52'dir. Bu sonuca göre Sürekli Kaygı Ölçeği'nin 0.52 olasılıkla sınav kaygısının yüksek olduğu durumu doğru tespit ettiği sonucuna ulaşılabilir. Yani bireyin sürekli kaygı düzeyi yüksek olarak belirlendiyse 0.52 olasılıkla gerçekten sınav kaygısı yüksektir. Çizelge 13'teki iki ölçeğe göre düşük kaygılı grupta bulunan birey sayısının Sürekli Kaygı Ölçeği'ne göre düşük kaygılı öğrenci sayısına oranından hesaplanan negatif yordayıcı değer ($NYD=142/173$) 0.82 olup bu sonuca göre Sürekli Kaygı Ölçeği 0.82 olasılıkla sınav kaygısının düşük olduğu durumu doğru tespit etmektedir. Bu değer, bireyin sürekli kaygı düzeyi düşük olarak belirlendiyse 0.82 olasılıkla gerçekten sınav kaygısı da düşüktür.

Sürekli Kaygı Ölçeği'nin negatif yordayıcı değeri, pozitif yordayıcı değerinden daha yüksektir. Bu bulgu, Konuk ve diğ. (2007) ile Konkan ve diğ.'nin (2013) ölçek uyarlama çalışmalarında hesaplanan negatif yordayıcı değerlerin, pozitif yordayıcı değerlerden yüksek bulunması bulguları ile örtüşmektedir. Bu çalışmalarda ölçülen özelliğin var ya da yüksek olduğu birey sayısının, özelliğin olmadığı ya da düşük olduğu birey sayısından daha az olduğu dikkat çekmektedir. Bu çalışmada da kaygı düzeyi düşük olan birey sayısı daha fazladır. Dolayısıyla sınıflandırma gruplarının sayılarının eşit olmaması negatif yordayıcı değerlerin yüksek olmasının nedeni olarak düşünülebilir. Uluç (2007), pozitif ya da negatif yordayıcı değerleri yüksek bir ölçme aracı kullanıldığı durumlarda, araştırma grubunda ölçülen özelliğin frekansının değişmesi nedeniyle bu değerlerden hangisine göre yorum yapılacağını değişeceğini belirtmektedir. Sürekli Kaygı Ölçeği'nin sınav kaygısının düşük olduğu durumu doğru tespit etme olasılığının, yüksek olduğu durumu doğru tespit etme olasılığından daha

fazla olduđu belirlenmiřtir. Arařtırmada, her iki ölçekte de düşük kaygılı sınıfında bulunan birey sayısı daha fazla olduđundan negatif yordayıcı değere göre ölçegin psikometrik özelliđi incelenmiřtir. Ölçegin ROC analiziyle hesaplanan 48.5 kesme puanına göre bireyler sınıflandırıldıđında, duyarlılık değerinin %72 ve özgüllük değerinin %37 olduđu belirlenmiřtir. Bu kesme puanına göre ölçegin yordayıcı değerinin yüksek olduđu söylenebilir. Bu ölçekle öğrencilerin ROC analiziyle hesaplanan kesme puanına göre yüksek veya düşük kaygı düzeyinde olduđu belirlenebilir. Ölçegin tanı koymak yerine tarama amacıyla kullanıldıđı göz önünde bulundurulduđundan sonuçlar, ROC analizi geçerlik kanıtı olarak kabul edilebilir. Bu kesme puanı dikkate alındıđında üniversite öğrencilerinde kaygı ile iliřkili olabilecek faktörlerin belirlenmesi ve öğrencilerde kaygıyı doğuran ya da besleyen kaynakların ele alınması için yönlendirmelerin kolaylıkla yapılabileceđi söylenebilir.

BÖLÜM 4

SONUÇ VE ÖNERİLER

Bu bölümde araştırmanın bulgularına dayalı olarak ulaşılan sonuçlara ve ileride yapılacak araştırmalara yönelik önerilere yer verilmiştir.

4.1. Sonuçlar

Araştırmada üniversite öğrencilerinin kaygı düzeylerini belirleyen bir ölçeğin Angoff, ROC ve Sınır Grup yöntemleri kullanılarak kesme puanı belirlenmiştir. Bu yöntemlerin yüksek/düşük kaygı sınıflama uyumu incelenerek ölçeğin geçerlik çalışmalarına katkı sağlanmıştır.

- Araştırmanın amacı çerçevesinde üniversite öğrencilerinin Sürekli Kaygı Ölçeği'nden aldıkları puanların değerlendirilmesi için kesme puanı değerleri Angoff yöntemiyle 48, Sınır Grup yöntemi ile 44.5, ROC analizine göre 48.5'tir. Yöntemler ölçek maddesi, öğrenci ve ölçüt ölçek odaklı olmasına rağmen Angoff ve ROC analizi ile hesaplanan kesme puanları birbirine çok yakındır. Ölçeğin tanı amacı koymak yerine tarama amacıyla kullanıldığı göz önünde bulundurulmalıdır. Belirlenen kesme puanlarına eşit ve bu puanların üzerinde puan alan bireylerin yüksek kaygılı olarak gruplandırılabilceği, bu bireylerle daha ayrıntılı görüşmeler yapılması gerektiği sonucuna ulaşılmıştır.
- Angoff ve Sınır Grup yöntemleri ile belirlenen kesme puanlarına göre yüksek/düşük kaygılı kabul edilen öğrencilerin yüzdeleri arasında manidar bir fark vardır. Yargıcı puanlarının ortalama ve ortanca değerlerine göre belirlenen kesme puanı; Angoff yönteminde daha yüksek, Sınır Grup yönteminde ise daha düşüktür. Kesme puanlarındaki bu farklılık, araştırma grubunun kaygı düzeylerine göre sınıflandırılma sayılarını ve yüzdelerini değiştirmiştir. Ayrıca kesme puanları oluşturulurken yargıcılardan, ölçülen özellikle ilgili minimum düzeydeki bireyi düşünmeleri istenmektedir. Yargıcıların cevapları oluştururken zorlanmaları kararlarını etkilemiştir. Standart belirleme yöntemleri ile kesme puanı oluşturulurken; ölçülen özelliğin yapısına, öğrenci ve yargıcı gruplarına bağlı olarak test veya öğrenci merkezli yöntemlere karar verilmesi gerektiği sonucuna ulaşılmıştır.

- Angoff ve Sınır Grup yöntemlerine göre belirlenen kesme puanlarının güvenilirlik kanıtı olarak yargıcılar arası uyum katsayısı incelenmiştir. Sınır Grup yöntemine göre kesme puanı belirlendiğinde sınıf içi korelasyon katsayısının orta düzeyde, Angoff yöntemine göre belirlendiğinde ise iyi düzeyde olduğu sonucuna ulaşılmıştır. Sınıf içi korelasyon katsayısının orta ve iyi düzeyde bulunmasının, yargıcıların farklı alan bilgisine sahip olmalarından, standart belirleme yöntemlerini tam olarak anlamamalarından ve tepki düzeylerini net olarak ifade edemediklerinden kaynaklandığı sonucuna ulaşılmıştır.
- ROC analizine göre hesaplanan kesme puanının, öğrencilerin kaygı düzeylerini sınıflamadaki doğruluk oranı araştırılmıştır. Bireylerin Sınav Kaygısı Ölçek puanları ölçüt alınarak Sürekli Kaygı Ölçeği'nin kesme puanını belirlemek için işlem karakteristik eğrisi (ROC) oluşturulmuştur. Araştırmada elde edilen AUC (eğri altında kalan alan) değerine göre Sürekli Kaygı Ölçeği'nin birey kaygı düzeyini belirlemede ayırım yapabildiği yani yüksek ve düşük kaygılı bireyleri %73 oranında (orta düzeyde) doğru sınıflandırabildiği sonucuna ulaşılmıştır.
- Sürekli Kaygı Ölçeği'nin ROC analizi ile geçerliği incelenmiştir. Sınav kaygısı ve sürekli kaygı düzeylerine göre sınıflandırılan düşük ve yüksek kaygılı öğrenci sayılarından pozitif ve negatif yordayıcı değerler hesaplanmıştır. Araştırmada, her iki ölçekte de düşük kaygılı sınıfta bulunan birey sayısı daha fazla olduğundan negatif yordayıcı değere göre ölçeğin psikometrik özelliği incelenmiştir. Ölçeğin ROC analizi ile hesaplanan duyarlılık, özgüllük ve negatif yordayıcı değerleri yüksektir. Ölçeğin tanı koymak yerine tarama amacıyla kullanıldığı göz önünde bulundurulduğundan kaygı düzeylerini yüksek olasılıkla belirleyebildiği sonucuna ulaşılmıştır.

Sonuç olarak bu araştırmada çoklu puanlanan bir psikolojik ölçme aracının standart belirleme ve sınır değer saptama yöntemlerine göre farklı kesme puanları hesaplanmıştır. Kesme puanının belirlenmesinde ROC analizi kullanıldığında diğer yöntemlerde olmayan duyarlılık, özgüllük ve pozitif/negatif yordayıcı değerlerinin bulunması daha ayrıntılı ve nesnel bilgiler sağlamıştır. ROC analizi ile hesaplanan kesme puanı kullanıldığında verilen kararların doğruluk oranı da hesaplanabilmektedir.

ROC analizi ile hesaplanan deęerlere gre leęin psikometrik zelliklerinin belirlenebileceęi gsterilmiřtir.

4.2. neriler

- Bu arařtırmada bireyler, dřk ve yksek kaygılı olmak zere iki kategoride sınıflandırılmıřtır. Bireylerin ikiden fazla kategoride sınıflandırılması iin oklu puanlanan aynı psikolojik testte kategoriler arası kesme puan belirleme alıřması yapılabilir.
- Bu arařtırmada oklu puanlanan psikolojik testlerde kesme puanı belirlenirken Angoff, Sınır Grup ve ROC olmak zere  yntem kullanılmıřtır. Angoff ve Sınır Grup yntemleri yargıcılarca kolay anlaşılabilir olmaları, uygulamadaki kolaylıkları gz nnde bulundurulduęundan arařtırmada tercih edilmiřtir. ROC analizi ise kesme puanı belirlenirken duyarlılık ve zgllk deęerlerinin testin geerlięine katkıda bulunması ve bireylerin sınıflama doęruluęuna ait deęerler sunması nedeniyle tercih edilmiřtir. Dięer kesme puanı belirleme yntemleriyle benzer arařtırmalar yapılabilir.
- Bu arařtırmada ęrenci grubuna lekler art arda uygulanmıřtır. Standart belirleme yntemlerine gre yargıcı kararları iin ise lekler birer hafta arayla uygulanmıřtır. Yapılacak arařtırmalarda bu sre farkı kesme puanı belirlemede bir deęiřken olarak belirlenip puanlar arasında farkın olup olmadıęı farklı yntemlerle incelenebilir.
- Bu arařtırmada yargıcı grup olarak kaygı konusunda alıřması olan uzman, rehberlik ve psikolojik danıřmanlarla alıřılmıřtır. Yargıcıların farklı niversitelerde bulunması ve yoęunlukları sebebiyle her bir yargıcı ile ayrı ayrı iletiřime geilerek kesme puanı ve yksek kaygının minimum dzeyindeki bireyin tanımı hakkında bilgilendirilmiřlerdir. Yargıcıların uygulamalardan nce oturumlarla bir araya getirilebileceęi farklı arařtırmalar yapılarak kesme puanı belirlemede bu durumun etkililięi belirlenebilir.
- Yargıcıların, minimum yksek kaygılı dzeydeki kiřiye iliřkin tanımlamaları birbirlerinden farklı olabilir. Bu nedenle yargıcılarla minimum yksek kaygı dzeyindeki kiřiyi tanımlamaya ynelik bir izleme alıřması yapılabilir.

Böylece yargıcıların düşük veya yüksek kaygı düzeyindeki bireyleri düşünerek tepki derecelerinin değişiminden kaynaklanabilecek farklılıklar ortadan kalkabilir.

- Eğitim alanında da tıp ve sağlık alanlarında olduğu gibi ölçek geliştirme ve uyarlama çalışmalarında ROC analizinden yararlanılarak geçerlik çalışmaları yapılabilir.
- Bu araştırmada belirlenen kesme puanlarıyla eğitim ve psikoloji alanlarında kullanılan ölçeklerin tarama amacıyla kullanılabilmesi vurgulanmıştır. Duyuşsal alana ilişkin testlerde başarı testlerinde olduğu gibi bir puan göz önünde bulundurularak bireyler hakkında kesin kararlar verilmemelidir. Kesme puanına göre sınıflandırılan birey, ölçülen özellikle ilişkili olabilecek değişkenlerin belirlenmesi için kolaylıkla yönlendirilebilecektir.

KAYNAKLAR

- Akçay, A., ve Demirel, A. (2011). Pyometralı köpeklerde kan parametrelerinin optimal pozitiflik eşiğinin özgün oranlar ve ROC eğrisi yöntemi ile belirlenmesi. *Erciyes Üniversitesi Veterinerlik Fakültesi Dergisi*, 8 (3), 153-163.
- Amerikan Eğitim Araştırmaları Birliği, Amerikan Psikoloji Birliği, Eğitim Ölçümleri Uluslararası Konseyi (1997). *Eğitimde ve psikolojide ölçme standartları*. (Çev. S. Hovardaoğlu ve N. Sezgin). Ankara: Türk Psikologlar Derneği Yayınları.
- Anastasi, A. (1988). *Psychological testing*. New York: MacMillan Pub. Co. Inc.
- Ateş, C., Öztuna, D., ve Genç, Y. (2009). Sağlık araştırmalarında sınıf içi korelasyon katsayısının kullanımı. *Türkiye Klinikleri J Biostat*, 1 (2), 59-64.
- Aydemir, Ö., Demet, M. M., Danacı, A. E., Deveci, A., Taşkın, E. O., ve Mızrak S., (2006). Borderline kişilik envanterinin Türkçeye uyarlanması, güvenirlik ve geçerliği. *Türkiye’de Psikiyatri Dergisi*, 8 (1), 6-10.
- Baykul, Y. (2010). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Berk, R. A. (1986). A consumer’s guide to setting performance standards on criterion-referenced tests. *Review of Educational Research*, 56 (1), 137-172.
- Buckhendal, W. C., Smith, W. R., Impara, C. J., and Plake, S. B. (2002). A comparison of Angoff and Bookmark standard setting methods. *Journal of Educational Measurement*, 39 (3), 253-263.
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi.
- Büyüköztürk, Ş., Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., ve Demirel, F. (2011). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi.
- Cizek, G. J. and Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Cizek, G. J. (2001). *Setting performance standards: Concepts, methods and perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cizek, G. J. (1993). Reconsidering standards and criteria. *Journal of Educational Measurement*, 30 (2), 93-106.
- Clauser, B. E., Harik, P., Margolis, M. J., McManus, I. C., Mollon, J., Chis, L. And Williams, S. (2009). An empirical examination of the impact of group discussion and examinee performance information on judgements made in the angoff standard setting procedure. *Applied Measurement in Education*, 22, 1-21.
- Crocker L., and Algina J. (1986). *Introduction to classical and modern test theory*. Orlando. Florida: Holt. Rinehart and Winston Inc.

- Cronbach, L. J. (1984). *Essentials of psychological testing*. New York: Harper Collins Publishers.
- Çamlıca, H., ve Dişçi, R. (2008). Tanı testlerinde sınır değerlerinin belirlenmesi. *Türk Onkoloji Dergisi*, 23 (1), 26-33.
- Çetin, S., ve Gelbal, S. (2010). Farklı standart belirleme yöntemlerinin geçme puanları üzerine etkisi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 43 (1), 79-95.
- Çetin, S. (2011). *İşaretleme ve Angoff standart belirleme yöntemlerinin karşılaştırılması* (Doktora tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Çokluk Ö., Büyüköztürk Ş., ve Şekercioğlu G. (2010). *Sosyal bilimler için çok değişkenli istatistik*. Ankara: Pegem Akademi.
- Dayhoff, S. A. (2000). *Working through social anxiety*. New Mexico: Effectiveness-Plus Publications.
- DeLong, R. E., DeLong, D. M., and Clarke-Pearson, D.L. (1988) Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*, 44 (3), 837-845.
- Dirican, A. 2001. Tanı testi performansının değerlendirilmesi ve kıyaslanması. *Cerrahpaşa Tıp Dergisi*, 32, 25-30.
- Downing, S. M., Tekian, A., and Yudkowsky, R. (2006). Research methodology: Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teaching and Learning in Medicine*, 18 (1), 50-57.
- Erkuş, A. (2007). Ölçek geliştirme ve uyarlama çalışmalarında karşılaşılan sorunlar. *Türk Psikoloji Bülteni*, 13(40), 17-25.
- Eröz, B. (2010). *Veri yapısına bağlı olarak ROC eğrisi altında kalan alana ilişkin istatistiksel yöntemlerin karşılaştırılması* (Yüksek lisans tezi). Hacettepe Üniversitesi, Ankara.
- Erözkan, A. (2004). Üniversite öğrencilerinin sınav kaygısı ve başa çıkma davranışları. *Muğla Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 12 (2), 13-38.
- Garson, G. D. (2003). *Reliability analysis*. Carolina State University. Retrieved from <http://www2.chass.ncsu.edu/garson/PA765/reliab.htm/intraclass>.
- Grove, W. M. (2006). *Mathematical aspects of diagnosis*. United States of America,
- Guilford, J. P. (1954). *Psychometrics methods*. New York: McGraw-Hill Book Co.
- Gündeğer, C. (2012). *Angoff, Yes/No ve Ebel standart belirleme yöntemlerinin karşılaştırılması* (Yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Hambleton, R. K. (2001). *Setting performance standards: concepts, methods and perspectives*. Mahwah, NJ: Lawrence Erlbaum Publishers.

- Hanley, J. A. and McNeil, B. J. (1983). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 29- 36.
- Hosmer, D. W., Lemeshow, S., and X. Sturdivant, R. (2013). *Applied logistic regression*. New York: John Wiley and Sons.
- Hurtz, G. M., and Hertz, N.R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A Generalizability Theory Study. *Educational and Psychological Measurement*, 59 (6), 885-897.
- Impara, J. C., ve Plake, B. S. (1998). Teachers' ability to estimate item difficulty: a test of the assumptions in the angoff standard setting method. *Journal of Educational Measurement*, 35 (1), 69-81.
- Irwin, P. (2007). An alternative examinee-centered standard setting strategy. Universty of Nebraska, USA.
- Jeager, R. M. (1989). *Certification of student competence*: In R.L. Linn (Ed) Educational Measurement.(pp. 485-514). New York: Macmillan.
- Kane, M. (1994). Validating the performance standards associated with passing scores. *Review of Educational Research*, 63(4), 425-461.
- Karakaya, J. (2012). *Üç yönlü Roc analizi ve ortak değişken düzeltilmesi* (Doktora tezi). Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü, Ankara.
- Karasar, N. (2009). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayınları.
- Konkan, R., Şenormancı, Ö., Güçlü, O., Aydın, E., ve Sungur, M. (2013). Yaygın anksiyete bozukluğu-7 (yab-7) testi türkçe uyarlaması, geçerlik ve güvenirliği, *Nöropsikiyatri Arşivi Dergisi*, 50, 53-58.
- Konuk, N., Kıran, S., Tamam, L., Karaahmet, E., Aydın, H., ve Atık, L. (2007). Duygu durum bozuklukları ölçeğinin Türkçe uyarlamasının bipolar bozukluk taramasında geçerliği. *Türk Psikiyatri Dergisi*, 18 (2),147-154.
- Korkmaz, M. (2007). Psikolojik ölçmenin yeni kuralları ve Türkiye'deki durumu. *Türk Psikoloji Bülteni*, 13 (40), 8-14.
- Koyuncu, S. (2015) *Psikolojik ölçeklerde ROC analizi yöntemiyle standart belirleme* (Yüksek lisans tezi). Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Köksal, B. (2011). *Regresyon analizinde ROC eğrisi kestirimi ile model seçimi* (Yüksek lisans tezi). Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Krzanowski, W. J. and Hand, D. J. (2009). *ROC curves for continuous data*. Boca Raton: Chapman and Hall/CRC Press.
- Livingston, S. A., and Zieky M. J. (1989). A comparative study of standard setting methods. *Applied Measurement in Education*, 2 (2),121-141.
- Metz, C.E., and Pan, X. (1999). "Proper" binormal ROC curves: Theory and maximum-likelihood estimation. *Journal of Mathematical Psychology*, 43 (1), 1-33.

- Mills, C. N., and Melican G. J. (1988). Estimating and adjusting cut of score: features of selected methods. *Applied Measurement In Education*, 1, 201-275.
- Moraes, S. A., Freitas, I. C. M., Mondini, L., and Rosas, J. B. (2009). Receiver operating characteristic (ROC) curves to identify birth weight cutoffs to predict overweight in Mexican school children. *Journal de Pediatria*, 85 (1), 42-48.
- Nunnally, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill Book Co.
- Obuchowski, N. A. (2005). ROC analysis. *American Journal of Roentgenology*, 184 (2), 364- 372.
- Ocaktan, M. E., Keklik, A., ve Çöl, M. (2002). Abidinpaşa sağlık ocağında çalışan sağlık personelinde spielberg durumluk ve sürekli kaygı düzeyleri. *Ankara Üniversitesi Tıp Fakültesi Dergisi*. 55 (1), 21-28.
- Ömür, S., ve Selvi, H. (2010). Angoff, Ebel ve Nedelsky yöntemleriyle belirlenen kesme puanlarının sınıflama tutarlılıklarının karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1 (2), 109-113.
- Öner, N. ve A. Le Compte. (1983). *Durumluk-sürekli kaygı envanteri el kitabı*. İstanbul: Boğaziçi Üniversitesi Yayınları
- Öner, P., Barut, Y., Öner, Ö., Üneri, Ö.Ş, Bodur, Ş., Turgut, S., ve Munir, K. M. (2009). Çocuklarda uyku ölçeğinin geçerlik ve güvenilirliği. *Klinik Psikofarmakoloji Bülteni*, 19 (4), 382-395.
- Öner, N. (1997). *Türkiye’de kullanılan psikolojik testler: bir başvuru kaynağı*. İstanbul: Boğaziçi Üniversitesi Yayınları.
- Özgüven, E. İ. (2007). *Psikolojik testler*. Ankara: PDREM Yayınları.
- Pepe, M. S. (2003). The statistical evaluation of medical tests for classification and prediction. New York United States.: Oxford Statistical Science Series.
- Sasse, E. A. (2002). Objective evaluation of data in screening for disease. *Clinica Chimica Acta*, 315 (1), 17-30.
- Schultz, D. P. ve Schultz, S.E. (1998). *Psychology and work today, an introduction to industrial and organizastional psychology*. New Jersey: Prentice-Hall Inc.
- Shapiro, D. E. (1999). The interpretation of diagnostic tests. *Statistical Methods in Medical Research*, 8 (2),113-34.
- Shrout, P. E., Fleiss, J. L. (1979). *Intraclass correlations: uses in assessing rater reliability*. *Psychol Bull*, 86 (2), 420-428.
- Spielberger, C. D., Gorsuch, R. C., and R. E. Luschene. (1970). *Manual for the state-trait anxiety invenntory*. California: Consulting Psychologists Press.
- Stone, G. E., Koskey, K. L. K., and Sondergeld, T. A. (2011). Comparing construct definition in the Angoff and objective standard setting models: playing in a house of cards without a full deck. *Educationaland Psychological Measurement*, 71(6), 942-962.

- Takahashi K., Uchiyama H., Yanagisawa S., and Kamae I. (2006). *The Logistic Regression and ROC Analysis of Group-Based Screening for Predicting Diabetes Incidence in Four Years Applied Medical Statistics and Decision Sciences*, TheKobe Journal of Medical Science, 52 (6), 171-80.
- Tanrıverdi, S. (2006). *Standart belirleme yöntemlerinin geçme puanları üzerindeki etkisi* (Yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Taşdelen, G. (2009). *Nedelsky ve Angoff standart belirleme yöntemlerinin Genellenabilirlik kuramı ile karşılaştırılmasına ilişkin bir araştırma* (Yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Taşdemir, F. (2013). *Angoff (1-0), Nedelsky ve sınır değerleri saptama yöntemleri ile bir testin sınıflama doğruluklarının incelenmesi* (Doktora tezi). Ankara Üniversitesi Eğitim Bilimler Enstitüsü, Ankara.
- Tavakol, M., and Dennick, R. (2012). Standard setting: The application of the receiver operating characteristic method. *International Journal of Medical Education*, 3, 198-200.
- Tekin, H. (2000). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınevi.
- Tezbaşaran, A. (2008). *Likert tipi ölçek hazırlama kılavuzu*. Ankara: Türk Psikologlar Derneği Yayınları.
- Tülübaş, G. (2009). *Psikolojik testlerde Angoff ve Sınır grup yöntemleri ile kesme puanlarının belirlenmesi* (Yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Uluç, S. (2007). ROC Analizi I: Özgüllük, duyarlılık, pozitif yordayıcı değer ve negatif yordayıcı değer hesaplamaları. *Türk Psikoloji Bülteni*, 13 (40), 26-31.
- Vivo, J. M., and Franco, M. (2008). How does one assess the accuracy academic success predictors? ROC analysis applied to university entrance factors. *International Journal of Mathematical Education in Science and Technology*, 39 (3), 325-340.
- Wang, L., Pan, W., and Austin, J. T. (2003). Standards-setting procedures in accountability research: Impacts of conceptual frameworks and mapping procedures on passing rates. *Paper presented at the annual meeting of the American Educational Research Association*, Chicago.
- Yiğit, R., Dilmaç, B., Deniz, M. E., ve Hamarta, E. (2011). Sürücülerin sürekli ve durumluk kaygılarının bazı değişkenler açısından incelenmesi, *Uluslararası Avrasya Sosyal Bilimler Dergisi*, 2 (4), 37-44.
- Zhang, B. (2005). A semiparametric hypothesis testing procedure for the ROC curve area under a density ratio model, *Computational Statistics and Data Analysis*, 50, 1855-1876

EKLER

EK1. SINIR GRUP YÖNTEMİNE AİT KARARLAR

Sınır Grup Yöntemi İçin Yargıcı Kararları										
Maddeler	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10
1	2	1	2	1	1	1	3	3	1	3
2	1	2	2	2	2	1	2	2	2	2
3	1	1	2	1	1	1	2	3	2	3
4	2	2	1	3	1	4	3	2	3	3
5	2	1	1	1	1	1	3	3	2	2
6	2	3	3	1	2	1	3	3	1	3
7	2	3	2	1	2	1	3	3	1	3
8	2	1	1	1	1	1	3	3	1	3
9	1	1	1	1	2	1	4	3	1	3
10	2	1	2	2	1	1	3	3	1	3
11	1	1	1	1	1	1	3	3	2	2
12	1	1	1	1	1	1	3	3	1	3
13	2	3	2	1	2	1	3	3	1	4
14	2	2	1	2	2	1	2	3	2	3
15	1	1	2	2	1	1	3	3	2	2
16	2	1	1	1	1	1	3	3	1	1
17	1	1	2	1	1	1	3	3	2	3
18	1	1	2	1	1	1	3	3	1	2
19	2	3	3	2	1	1	3	3	3	3
20	2	1	1	1	1	2	1	2	2	3
	38	36	38	44	43	45	49	50	49	49
Sınır Grup Yöntemi Kesme Puan: 44.5										

EK2. ANGOFF YÖNTEMİNE AİT KARARLAR

Angoff Yöntemi için Yargıcı Kararları										
Maddeler	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10
1	2	1	1	2	1	1	3	3	2	2
2	3	2	2	3	1	3	3	3	2	3
3	2	2	2	2	2	2	2	3	2	2
4	2	1	4	3	1	3	3	2	2	4
5	2	2	2	1	2	1	3	3	2	2
6	2	1	2	1	1	2	3	3	1	2
7	2	1	1	1	1	2	3	3	1	2
8	2	1	1	1	1	1	3	3	2	3
9	2	1	1	2	2	2	4	3	1	3
10	2	1	2	1	1	1	3	3	2	2
11	1	1	1	1	1	1	3	3	1	1
12	1	1	1	1	1	1	2	3	2	2
13	2	1	1	1	1	2	3	3	1	2
14	2	2	1	2	1	2	2	3	2	3
15	2	2	2	2	1	2	3	3	2	1
16	2	1	2	1	1	1	3	3	2	2
17	1	1	1	2	1	2	3	2	2	1
18	2	2	2	1	2	2	4	3	2	1
19	2	1	2	1	2	1	2	3	3	2
20	2	1	2	2	2	2	3	3	2	2
	45	47	48	50	45	49	49	51	47	49
Angoff Yöntemi kesme puan: 48										

EK 3. ÖLÇÜT ÖLÇEK-SINAV KAYGISI ÖLÇEĞİ**SINAV KAYGISI ÖLÇEĞİ****Açıklama :**

Bu araştırma, çoklu puanlanan ölçeklerin standart belirleme yöntemleri ve sınır değer saptama yöntemleri ile kesme puan belirleme ve psikometrik özelliklerini inceleme amacı ile yapılan bir doktora tez çalışmasıdır. Aşağıda bireylerin kendilerini tanımlamak için kullandıkları bir dizi ifade sıralanmıştır. Her birini okuyarak genel olarak nasıl hissettiğinizi anlatan ifadenin sağındaki boşluklardan uygun olanını işaretleyiniz. Doğru ya da yanlış yanıt yoktur. İfadelerin hiçbiri üzerinde fazla zaman harcamadan sınavlarda genel olarak nasıl hissettiğinizi gösteren yanıtı işaretleyin. Araştırma sonuçları başka bir amaçla kullanılmayacaktır. Katılımınızdan dolayı teşekkür ederim.

Öğr.Gör.D. Bahar ŞAHİN
R.T.E.Ü/Eğitim Bilimleri Bölümü

	hiçbir zaman	bazen	sık sık	her zaman
1.Sınav sırasında kendimi rahat ve güvenli hissedirim.				
2.O dersten alacağım notu düşünmek sınav sırasındaki başarıyı olumsuz yönde etkiler.				
3.Önemli sınavlarda donup kalırım.				
4.Sınavlar sırasında bir gün okulu bitirip bitiremeyeceğimi düşünmekten kendimi alamam.				
5.Bir sınav sırasında ne kadar çok uğraşırsam kafam o kadar çok karışır.				
6.Sınavlarda kendimi huzursuz ve rahatsız hissedirim.				
7.Önemli bir sınav sırasında kendimi sinirli hissedirim				
8.Başarısız olma düşünceleri,dikkatimi sınav üzerinde toplamama engel olur.				
9.Bir sınava çok iyi hazırlandığım zamanlar bile kendimi oldukça sinirli hissedirim.				
10.Önemli sınavlarda sinirlerim öylesine gerilir ki midem bulanır.				
11.Bir sınav kağıdını almadan hemen önce çok huzursuz olurum.				
12.Önemli sınavlarda kendimi adeta yenilgiye iterim.				
13.Sınavlar sırasında kendimi çok gergin hissedirim.				
14.Önemli bir sınav sırasında paniğe kapılırım.				
15.Sınavların beni bu kadar rahatsız etmemesini isterim.				
16.Önemli bir sınava girmeden çok endişelenirim.				
17.Sınavlar sırasında başarısız olmanın sonuçlarını düşünmekten kendimi alamam.				
18.Önemli sınavlarda kalbimin çok hızlı attığını hissedirim.				
19.Sınav sona erdikten sonra endişelenmemeye(kurmamaya) çalışır,fakat yapamam.				
20.Sınavlar sırasında öylesine sinirli olurum ki aslında bildiğim şeyleri bile unuturum.				

EK 4. KESME PUANI VE PSİKOMETRİK ÖZELLİKLERİ İNCELENEN SÜREKLİ KAYGI ÖLÇEĞİ

Sürekli Kaygı Ölçeği

Açıklama:

Aşağıda kişilerin kendilerine ait duygularını anlatmada kullandıkları birtakım ifadeler verilmiştir. Her ifadeyi okuyarak o anda nasıl hissettiğinizi ifadelerin sağ tarafında bulunan sıklık durumlarına göre işaretleyiniz. Doğru ya da yanlış cevap yoktur. Herhangi bir ifadenin üzerinde fazla zaman harcamayınız.

		hiçbir zaman	bazen	çok zaman	her zaman
21.	Genellikle keyfim yerindedir				
22	Genellikle çabuk yorulurum				
23	Genellikle kolay ağlarım				
24	Başkaları kadar mutlu olmak isterim				
25	Çabuk karar veremediğim için fırsatları kaçıırım				
26.	Kendimi dinlenmiş hissediyorum				
27.	Genellikle sakin, kendine hakim ve soğukkanlıyım				
28	Güçlüklerin yenemeyeceğim kadar biriktiğini hissedirim				
29	Önemsiz şeyler hakkında endişelenirim				
30.	Genellikle mutluyum				
31	Her şeyi ciddiye alır ve endişelenirim				
32	Genellikle kendime güvenim yoktur				
33.	Genellikle kendimi emniyette hissedirim				
34	Sıkıntılı ve güç durumlarla karşılaşmaktan kaçınırım				
35	Genellikle kendimi hüzünlü hissedirim				
36.	Genellikle hayatımdan memnunum				
37	Olur olmaz düşünceler beni rahatsız eder				
38	Hayal kırıklıklarını öylesine ciddiye alırım ki hiç unutamam				
39.	Aklı başında ve kararlı bir insanım				
40	Son zamanlarda kafama takılan konular beni tedirgin ediyor				

EK5. SINIR GRUP YÖNTEMİ İÇİN YÖNERGE**Değerli Öğretim Elemanı,**

Bu araştırma, üniversite son sınıf öğrencilerinin kaygı düzeylerinin standart belirleme yöntemleri ile incelenmesi amacıyla yapılacak bir doktora tez çalışmasıdır. Siz değerli alan uzmanlarından isteğim, “yüksek kaygının minimum düzeyinde” olan bir öğrenciyi temel alarak (öğrenci gibi düşünerek) ölçeği cevaplamanızdır. Katılımınızdan dolayı çok teşekkür ederim.

Yüksek Kaygının Minimum Düzeyi: Kaygılı olduğu düşünülen öğrencilerin ölçekten alması beklenen en düşük puan

NOT: Ölçekten elde edilen puanlar kuramsal olarak 20 ile 80 arasında değişir. Puanlar 80’e doğru yaklaştıkça yüksek kaygı seviyesini, 20’ye doğru yaklaştıkça düşük kaygı seviyesini ifade eder.

**Öğr. Gör. D. Bahar Şahin / R.T.E.Ü Eğitim Fakültesi
Ankara Üniversitesi Eğitim Bilimleri Enstitüsü
Ölçme ve Değerlendirme Bölümü Doktora Öğrencisi**

EK6. ANGOFF YÖNTEMİ İÇİN YÖNERGE**Değerli Öğretim Elemanı,**

Bu araştırma, üniversite son sınıf öğrencilerinin kaygı düzeylerinin sınır değer saptama yöntemleri ile incelenmesi amacıyla yapılacak olan çalışmadır. Siz değerli alan uzmanlarından isteğim, “yüksek kaygının minimum düzeyinde” olan öğrenciyi ölçek maddelerinin seçenekleri belirleyecek olsaydı hangi seçeneğin bu öğrenciyi ayırt edeceğini ölçek üzerinde işaretlemenizdir. Katılımınızdan dolayı çok teşekkür ederim.

Yüksek Kaygının Minimum Düzeyi: Kaygılı olduğu düşünülen öğrencilerin ölçekten alması beklenen en düşük puan

NOT: Ölçekten elde edilen puanlar kuramsal olarak 20 ile 80 arasında değişir. Puanlar 80’e doğru yaklaştıkça yüksek kaygı seviyesini, 20’ye doğru yaklaştıkça düşük kaygı seviyesini ifade eder.

**Öğr. Gör. D. Bahar Şahin / R.T.E.Ü Eğitim Fakültesi
Ankara Üniversitesi Eğitim Bilimleri Enstitüsü
Ölçme ve Değerlendirme Bölümü Doktora Öğrencisi**

ÖZ GEÇMİŞ

Adı-Soyadı: Döndü Bahar ŞAHİN SARKIN

Doğum Tarihi: 01.06.1986

İletişim Bilgileri: Recep Tayyip Erdoğan Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Ölçme ve Değerlendirme Ana Bilim Dalı

E-posta Adresi: bahar.sahin@erdogan.edu.tr

Öğrenim Durumu:

Derece	Bölüm/Program	Üniversite	Yıl
Lisans	Fizik Öğretmenliği	Selçuk Üniversitesi	2005-2010
Yüksek Lisans	Ölçme ve Değerlendirme	Ankara Üniversitesi	2010-2012
Doktora	Ölçme ve Değerlendirme	Ankara Üniversitesi	2012-2017

İş Deneyimi:

Unvan	Görev Yeri	Yıl
Öğretim Görevlisi	Recep Tayyip Erdoğan Üniversitesi	2012-Devam