

# Parametrik ve Parametrik Olmayan Madde Tepki Kuramında Farklı Örneklem Büyüklüklerine ve Boyutluluklarına Göre Parametre Değişmezliğinin İncelenmesi \*

Çiğdem REYHANLIOĞLU \*\*

Nuri DOĞAN \*\*\*

## Öz

Bu çalışmanın amacı farklı boyutluluk ve örneklem özelliklerinde Madde Tepki Kuramı (MTK) uygulamalarına göre kestirilen parametrelerin değişmezliğini incelemektir. Bu amaçla 2015 yılındaki Temel Eğitimden Ortaöğretime Geçiş (TEOG) Sistemi'nin birinci uygulamasındaki A kitapçığını alan öğrenci cevapları araştırma verisi olarak kullanılmıştır. Çalışma evreninin büyüklüğü 63,871'dir. Evrenden rastgele seçilmiş 50, 100, 200, 500, 1000 ve 5000 kişilik gruplar çalışmanın örneklemini oluşturmaktadır. MTK uygulamalarında tek boyutlu Matematik alt testinden ve yapay olarak oluşturulan iki boyutlu testin sonuçlarından yararlanılmıştır. Çalışmadan elde edilen bulgular sonucunda tek boyutlu test için Tek Boyutlu Parametrik Olmayan MTK'ye (TBPOMTK) göre 200 örneklem büyüklüğü itibarıyla madde parametresi değişmezliği sağlanmıştır. Aynı test Tek Boyutlu Parametrik MTK'ye (TBPMTK) göre analiz edildiğinde ise evren değere yakın madde parametresi kestirimleri için en az 1000 örneklem büyüklüğü ile çalışılması gerektiği sonucuna ulaşılmıştır. İki boyutlu testin TBPMTK ve TBPOMTK'ye ve Çok Boyutlu MTK'ye (ÇBMTK) göre analiz edilmesi ile elde edilen madde parametrelerinde değişmezliğin sağlanamadığı sonucuna ulaşılmıştır.

*Anahtar Kelimeler:* Madde Tepki Kuramı uygulamaları, örneklem büyüklüğü, parametre değişmezliği.

## GİRİŞ

Alanyazında bir testin çok boyutlu olduğu bir durumda bireylerin performansını belirlemede tüm teste ait puanının kullanılıp kullanılmayacağı veya nasıl kullanılacağı tartışılmaktadır. Dolayısıyla bir testin tek boyutlu olmadığı bir durumda tekboyutluluk varsayımı üzerine kurulan test teorileri çok boyutlu testlerden elde edilen verilerin analizinde yetersiz kalabilir. Bu durumda çok boyutlu bir testten elde edilen bireylere ait yetenek ve madde parametrelerinin kestirilmesi sürecinde kullanılan modeller önemlidir (Meara, Robin ve Sireci, 2000). Bu modeller tek boyutluluk varsayımı gerektirmeyen çok boyutlu verileri analiz edebilecek nitelikteki modeller olmalıdır. Diğer yandan ölçme kuramları incelendiği zaman bir çok varsayım gerektiren parametrik yöntemlerin kullanılması oldukça yaygındır. Ancak parametrik koşulların oluşmadığı ve tek boyutluluğun sağlanmadığı durumlara eğitim uygulamalarında sık sık rastlanmaktadır. Eğitimde ve psikolojide her zaman parametrik koşulların oluşmaması ve bireyler hakkında karar vermek için birden fazla alt boyutu olan testlerin kullanma zorunluluğu parametrik ve tek boyutlu modellerden farklı uygulamaların geliştirilmesini zorunlu kılmıştır. Bir başka ifade ile parametrik koşulların karşılanmadığı durumlarda parametrik olmayan modeller ve tek boyutluluğun sağlanamadığı durumlar için ise çok boyutlu modeller geliştirilmiştir. Ancak geliştirilen bu modeller kullanılmadan önce işlevliliğinin deneysel olarak ortaya konması gerekir. Bu amaçla kuramlar çerçevesinde geliştirilen bu modellerden elde edilen sonuçların, mevcut kuramlardan elde edilen sonuçlarla karşılaştırılması gerekir.

\* Bu çalışma, Parametrik ve Parametrik Olmayan Madde Tepki Kuramında Farklı Örneklem Büyüklüklerine ve Boyutluluklarına Göre Parametre Değişmezliğinin İncelenmesi isimli doktora tezinden üretilmiştir.

\*\* Dr., Gaziantep Koleji Vakfı Özel Okulları, Gaziantep-Türkiye, dr.cigdemreyhanlioglu@gmail.com: ve ORCID ID: 0000-0002-4685-0495

\*\*\* Prof.Dr., Hacettepe Üniversitesi, Ankara-Türkiye, e-posta: nuridogan2004@gmail.com ve ORCID ID: 0000-0001-6274-2016

Bu makaleye atıfta bulunmak için:

Reyhanlioğlu, Ç., & Doğan, N. (2020). Parametrik ve parametrik olmayan madde tepki kuramında farklı örneklem büyüklüklerine ve boyutluluklarına göre parametre değişmezliğinin incelenmesi. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 11(2), 98-112. doi: 10.21031/epod.584977

Geliş Tarihi: 01.07.2019  
Kabul Tarihi: 07.04.2020

Parametrik MTK modellerinin geliştirilmesi ve yaygınlaşması modern test teorisinin geliştirilmesinde ve kullanılmasında kuşkusuz önemli bir aşamadır. Bu modellerin kullanılması için büyük örneklemelere ihtiyaç duyulduğuna dair önemli bulgular elde edilmiştir. Bu durum parametrik modellerin okullarda ve diğer saha çalışmalarında uygulanabilirliği açısından önemli bir sınırlıktır. Bu sınırlıktan dolayı saha çalışmalarında ve özellikle okullarda madde parametreleri örneklem özelliklerine bağlı olarak, bireylerin başarı düzeyleri de madde parametrelerine bağlı olarak kestirilebilmektedir. Oysa MTK'nin en önemli özelliği olan madde ve yetenek parametrelerinin birbirinden bağımsız olarak kestirilmesini ifade eden değişmezlik özelliğinin sağlanmasının okullarda öğrenci yetenek düzeyleri ile test ve madde parametrelerine dayalı olarak alınan kararlar açısından önemli olduğu düşünülmektedir. MTK'nin değişmezlik özelliğinin sağlanması ile okullarda uygulanan testlere ait madde parametrelerinin öğrencilerin yetenek dağılımından bağımsız ve benzer şekilde yetenek dağılımlarının da test ve madde parametrelerinden bağımsız kestirilmesi mümkün olur (Price, 2017). Böylece değişmezlik özelliği sayesinde okullarda uygulanan sınavlara bağlı olarak alınan kararların daha isabetli olması beklenir. Bu nedenle küçük çalışma gruplarına uygulanabilmesi açısından özellikle okullarda parametrik olmayan MTK modellerinin kullanılması ve yaygınlaştırılması büyük önem arz etmektedir. Büyük örneklemelerin kullanılması zorunluluğuna bir çözüm sunması açısından parametrik olmayan MTK modelleri parametrik modellere göre önemli avantajlara sahiptir. Parametrik olmayan MTK modellerinin parametrik modellere göre bir diğer önemli avantajı da maddelere verilen tepkiler ile maddelerin ölçtüğü gizil değişken arasındaki ilişkinin daha az varsayıma sahip olmasıdır. Bunun nedeni parametrik olmayan MTK modellerine ait madde karakteristik eğrilerinin önceden tanımlanmış parametrik bir biçimlerinin olmamasıdır (Sodano & Tracey, 2011). Buna göre parametrik olmayan modellerin parametrik modellere göre daha kullanışlı olduğu ifade edilebilir. Bununla birlikte her ne kadar parametrik olmayan modeller önemli avantajlara sahip olsa da bu modelleri kullanabilmek için en az parametrik modeller kadar iyi çalıştığına ilişkin kanıt toplamaya ihtiyaç vardır ve bu çalışmadan elde edilen sonuçlar bu açıdan önemli kanıtlar ortaya koyacaktır. Bu çalışmada Türkiye'de kullanımı yeni yaygınlaşan bir kuram tek boyutlu parametrik olmayan MTK'den elde edilen sonuçlar ile bu kuramın parametrik karşılığı tek boyutlu MTK ve çok boyutlu testlerin analizinde kullanılan çok boyutlu MTK kapsamında elde edilen sonuçlar karşılaştırılmıştır.

### ***Araştırmanın Amacı***

Türkiye'deki alanyazın incelendiğinde, çok kategorili veriler üzerinde parametrik olmayan tek boyutlu ve çok boyutlu MTK uygulamalarının karşılaştırılması ile ilgili farklı çalışmalar olmasına rağmen (Koğar, 2018; Şengül-Avşar, 2018; Şengül-Avşar, 2017) iki kategorili veriler ile gerçekleştirilen sınırlı sayıda çalışmaya rastlanmıştır. Koğar (2014) ve Mor-Dirlik (2017) tarafından yapılan çalışmalar dışında araştırmaya ulaşılabilmiştir. Bu çalışma, iki kategorili gerçek veri üzerinden gerçekleştirilen Türkiye literatüründeki ilk çalışmalardan biri olması bakımından önemlidir. Bu bağlamda bu çalışmada iki kategorili tek boyutlu bir veri seti üzerinden parametrik ve parametrik olmayan MTK modellerinin işlevliliğinin test edilmesi amaçlanmıştır. Bununla birlikte çok boyutlu bir veri setinin tek boyutlu parametrik ve parametrik olmayan MTK modelleri ile çok boyutlu MTK modelinden elde edilen sonuçları karşılaştırılmıştır. Böylece tek boyutluluk varsayımının bozulduğu bir durumda tek boyutlu parametrik ve parametrik olmayan modellerin verinin analizinde yeterliliğinin ortaya konması amaçlanmıştır.

Güvenilir bilginin tekrarlanabilir bir yapıya sahip olması beklenir. Dolayısıyla aynı nedenlerin aynı koşullar altında aynı sonuçları vermesi beklenir. Yinelenmeyen, neden-sonuç ilişkisine dayanmayan bilgiler bilimsel sayılmayabilir. Bilimsel bilgi, incelenen alanların değişik ortamlarda tekrarlanmasıyla elde edilen sonuçlarından oluşur. Aynı zamanda bilimsel bilgi birikimlidir. Her bilim insanının yaptığı bir çalışmanın, kendinden önceki çalışmaları destekler nitelikte olması beklenir; desteklemediği durumda ise nedenlerinin açıkça ortaya konması beklenir. Bu bağlamda Türkiye ve dünya literatürü incelendiği zaman MTK uygulamalarından elde edilen sonuçların, ölçeklerdeki boyut sayısı (Smits,

Timmerman & Meijer 2012), yetenek düzeylerinin dağılımı (Syu, 2013) gibi farklı deđişkenler açısından karşılaştırıldıđı çok sayıda çalışmanın olduđu görülmüştür. Bu deđişkenlerden bir tanesi de örneklem büyüklüğüdür (Kođar, 2014; Köse, 2010; Sünbül, 2011). Örneklem büyüklüğü, aynı amaca hizmet eden parametrik ve parametrik olmayan modellerden hangisinin kullanılması gerektiđine karar verirken göz önünde bulundurulması gereken önemli bir faktördür. Bu çalışmada elde edilen sonuçların farklılaşp farklılaşmadıđı kuramsal uygulamaların yanı sıra örneklem büyüklüğüne göre de incelenmiştir. Bu bağlamda bu çalışma, MTK uygulamalarından elde edilen sonuçların örneklem büyüklüğüne göre karşılaştırıldıđı çalışmalardan biri olması açısından da önemlidir. Bu çalışma ile alanyazında parametrik ve parametrik olmayan MTK modellerinin kullanım tercihi için ölçüt olabilecek bir örneklem büyüklüğü belirlemek çalışmanın amaçlarından biridir. Bu amaç doğrultusunda TEOG uygulamasının farklı boyutlardaki alt testlerinden elde edilen yetenek düzeylerinin, örneklem büyüklükleri de göz önünde bulundurularak parametrik ve parametrik olmayan MTK ile çok boyutlu MTK kapsamında farklılaşp farklılaşmadıđını ortaya koymak amaçlanmıştır. Bu amaç çerçevesinde cevap aranan problem cümlesi “Farklı boyutluluk ve örneklem büyüklüğü deđişkenlerine göre parametrik ve parametrik olmayan MTK için parametrelerin deđişmezliđi ne düzeyde sağlanmaktadır?” şeklinde yapılandırılmıştır. Yapılandırılan problem cümlesi çerçevesinde cevap aranan alt problemler:

1. Tek boyutlu testlerde, parametrik ve parametrik olmayan MTK’ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduđuunda ve evrenden kestirildiđi durumda nasıldır?
2. Çok boyutlu testlerde, tek ve çok boyutlu MTK’ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduđuunda ve evrenden kestirildiđi durumda nasıldır?
3. Çok boyutlu testlerde, parametrik ve parametrik olmayan MTK’ye göre hesaplanan madde parametreleri için kestirilen standart hata ortalamaları, örneklem büyüklüğü 50, 100, 200, 500, 1000 ve 5000 olduđuunda ve evrenden kestirildiđi durumda nasıldır?

Alt problemlerin çözümünden elde edilen bulguları yorumlarken TBPOMTK, TBPMTK ve ÇBMTK için madde ayırtediciliđi ve madde güçlüđü için farklı göstergelerden yararlanılır. TBPOMTK için maddenin ayırtedicilik gücünü yorumlarken Hi parametresi kullanılmıştır ve Hi deđerinin 0.30’dan küçük olması Sijtsma ve Molenaar’a göre (2002) maddenin ayırtedicilik bakımından zayıf olduđunu gösterir. TBPOMTK için madde güçlüđünün göstergesi olarak klasik güçlük parametresi olan p deđerlerinden yararlanılmıştır. TBPMTK için ise ayırt ediciliđin göstergesi olarak a parametresi, madde güçlüđünün göstergesi olarak b parametresi kullanılır. Teorik olarak a ve b parametreleri (-∞, +∞) aralıđında deđerler alırlar. Son olarak ÇBMTK için, TBPMTK’de olduđu gibi ayırt ediciliđin göstergesi olarak a parametresi kullanılır. ÇBMTK’de testin her bir boyutu için ayrı bir ayırtedicilik parametresi kullanılmaktadır. Bu çalışmada kullanılan birleşik test iki boyutlu olduđu için iki tane ayırtedicilik parametresi kestirilmiştir. Bunlar a<sub>1</sub> ve a<sub>2</sub> parametreleridir. ÇBMTK için madde güçlüđünün göstergesi olarak d parametresi kullanılır. Benzer şekilde d parametresi TBPMTK’deki b parametresi gibi yorumlanmaktadır.

## YÖNTEM

Bu çalışma tek boyutlu parametrik ve parametrik olmayan MTK ile çok boyutlu MTK modellerine ilişkin betimleyici istatistikler elde etme, iki ya da daha fazla deđişken arasındaki ilişkinin varlıđını ve derecesini ortaya koyma açısından betimsel bir çalışmadır. Betimsel araştırmalar, olayların, objelerin, varlıkların, kurumların ve çeşitli alanların "ne" olduđunu açıklamaya çalışır (Kaptan 1977).

### *Evren ve Örnekleme*

Çalışmada kullanılan veriler, 2015 yılındaki TEOG’un birinci uygulamasında yer alan testlerin her biri için A kitapçıđını alan öđrenci cevaplarından elde edilmiştir. A kitapçıđında Türkçe, Matematik, Fen

Bilgisi, Din Kültürü ve Ahlak Bilgisi, İnkılap Tarihi, İngilizce, Almanca ve Fransızca alt testleri yer almaktadır. Bu çalışmada bütün alt testlerde A kitapçığında yer alan soruları cevaplayan öğrenciler çalışmanın evrenini oluşturmaktadır. Çalışmanın bundan sonraki kısmında çalışma evreni “evren” şeklinde ifade edilmiştir. Evren büyüklüğü 63.871’dir. Bu çalışmada elde edilen sonuçlar örneklem büyüklüklerine göre karşılaştırıldığı için, evrenden rastgele seçilmiş 50, 100, 200, 500, 1000 ve 5000 kişilik gruplar çalışmanın örneklemelerini oluşturmaktadır. Örneklem büyüklükleri belirlenirken özellikle küçük örneklem için saha uygulamaları göz önünde bulundurulmuştur. Örneklem büyüklüğünün alt sınırı belirlenirken 2017-2018 eğitim ve öğretim yılı içerisinde Gaziantep’teki özel kurumlarda okumakta olan ortalama 8. sınıf öğrenci sayısı göz önünde bulundurulmuştur. Gaziantep İl Milli Eğitim Müdürlüğü’ne bağlı olan Ar-Ge biriminden alınan bilgiye göre Gaziantep’te bulunan 24 özel ortaokulda okumakta olan toplam 1188 tane sekizinci sınıf öğrencisi bulunmaktadır. Dolayısıyla her bir okul başına düşen öğrenci sayısının ortalama değeri 49.5’tir. Bu nedenle örneklem büyüklüğünün alt sınırı 50 olarak belirlenmiştir. Diğer küçük örneklem büyüklükleri (100, 200 ve 500) de 50’nin çift katları olacak şekilde belirlenmiştir. Büyük örneklem büyüklüklerinin belirlenmesinde ise Hullin, Lissak ve Drasgow (1982), Goldman ve Raju (1986) ve Thissen ve Wainer (1982) tarafından yürütülen çalışma sonuçları göz önünde bulundurulmuştur.

Evrenden rastgele seçilen 50, 100, 200, 500, 1000 ve 5000 kişilik örneklem replikasyon yapılmaksızın sadece bir kez seçilmiştir. Örneklem seçiminde replikasyon yapılmaması çalışmanın sınırlılığı gibi görülmesine rağmen büyük örneklem için replikasyon yapılması durumunda karşılaşılabilecek problemlerin üstesinden gelmek için bu yola başvurulmuştur. Örnek olarak 5000 kişilik bir örneklem seçerken 50-100 replikasyon yapıldığında her bir örneklemde çok sayıda aynı birey bulunabilir ve bu bireyler parametre değişmezliğini şişirebilir. Bir başka ifadeyle değişmezliğin sağlanmasına otomatik neden olurlar ve/veya kestirimlerin yanlı olmasına neden olabilirler. Bu nedenle bu çalışmada replikasyon yapılmamıştır.

### **Veri Toplama Araçları**

Bu çalışmada 2015 yılı TEOG birinci sınavının A kitapçığında yer alan Türkçe, Matematik, Fen ve Teknoloji, T.C. İnkılap Tarihi, Yabancı Dil, Din Kültürü ve Ahlak Bilgisi alt testlerinden elde edilen verilerden yararlanılmıştır.

### **İşlem**

Çalışmanın amacına uygun olarak TEOG’un birinci uygulamasında kullanılan A kitapçığında yer alan tek boyutlu ve iki boyutlu olan alt testlere ait sonuçların analiz sürecinde kullanılması amaçlanmıştır. Bu amaçla TEOG’da yer alan bütün alt testlerin KMO ve Bartlett Küresellik Testi sonuçlarına göre faktör analizine uygun olup olmadığı incelenmiştir. TEOG’da yer alan bütün alt testlere ait KMO değerleri 0.90’ın üzerinde çıkmıştır. Bartlett Küresellik Testi sonuçları ise bütün alt testler için istatistiksel olarak anlamlıdır. Bu durumda TEOG’da yer alan her bir alt test için elde edilen Bartlett testinin istatistiksel olarak anlamlı olması, verilerin çok değişkenli normal dağılımdan geldiğini ve dolayısıyla verilerin, faktör analizinin uygulanması için uygun bir yapıya sahip olduğunu gösterir (Çokluk, Şekercioğlu & Büyüköztürk, 2010). Bunun yanında örneklem büyüklüğünün faktör analizine uygunluğu açısından KMO değerinin 0.60’dan büyük olması istenir (Tabachnick & Fidell, 2001). Elde edilen sonuçlar doğrultusunda her bir alt test için polichoric korelasyon matrisinin kullanıldığı paralel analize dayalı boyutluluk sonuçları veren FACTOR 10.5.01 programından yararlanılarak faktör analizi gerçekleştirilmiştir. Yapılan boyutluluk analizi sonucunda TEOG’un bütün alt testlerinin baskın bir tek boyuta sahip olduğu belirlenmiştir. Her bir alt teste ait verinin, elde edilen tek faktörlü modelle olan uyumunu ortaya koymak için iki göstergeden yararlanılmıştır. Bunlar GFI ve RMSR’dır. GFI’nin 1’e yakınlığı ölçüsünde model ile veri uyumludur. RMSR ise Kelly’nin ölçüt değeri olan (0.0316)’dan küçük olursa model ile verinin iyi uyum sağladığı ifade edilebilir (Harman, 1962). Elde edilen GFI değerleri [0.998, 1] aralığında ve RMSR değerleri ise [0.001, 0.023] aralığında değer almıştır.



TEOG'da yer alan alt testlerin tamamının tek boyutlu bir yapıya sahip olduđu belirlenmesinin ardından ÇBMTK analizleri için kullanılmak üzere, iki tane alt testten madde seçerek iki boyutlu ve tek boyutlu MTK analizlerinde kullanılan testle aynı uzunluđa sahip (20 madde) yeni bir test oluşturulmuştur. ÇBMTK analizleri için gerekli olan testin oluşturulmasında yararlanılan alt testlerin seçiminde alt testlerin ikili kombinasyonlarından elde edilen faktör analizi sonuçları göz önünde bulundurulmuştur. Birleştirildiğinde iki boyutluluđu en iyi sađlayan alt testler Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi alt testleridir. Fen Bilgisi ile Din Kültürü ve Ahlak Bilgisi alt testlerinden kendi içlerinde korelasyonları yüksek ve diđer test maddeleriyle korelasyonları düşük 10'ar maddenin seçilmesiyle 2 boyutlu birleşik bir test oluşturulmuştur. Oluşturulan testin 2 boyutlu yapısına dair gerekli kanıtlar faktör analizi ile elde edilmiştir. Mevcut verinin, elde edilen iki faktörlü modelle olan uyumunu ortaya koymak için kullanılan iki gösterge arasından GFI deđeri 0.999 ve RMSR deđeri ise 0.01'dir. Sonuçta birleşik teste ait verilerin 2 faktörlü yapıya uyum sađladıđı sonucuna ulaşılmıştır.

### **Verilerin Analizi**

Verilerin analizi süreci MTK'nin gerekli varsayımlarının test edilmesiyle başlamıştır. TBPMTK için tek boyutluluk ÇBMTK için yerel bağımsızlık ve TBPOMTK için tek boyutluluk ve monotonluk varsayımları test edilmiştir. TBPMTK ve TBPOMTK için yerel bağımsızlık varsayımı ayrıca test edilmemiştir. Bunun nedeni tek boyutluluğun sađlanması için yeterli görülmektedir (Hambleton, Swaminathan ve Rogers, 1985). Bu nedenle yerel bağımsızlık varsayımı yalnızca ÇBMTK için test edilmiştir.

Yerel bağımsızlıđı test etmek için kullanılan yöntemlerden bir tanesi koşullu maddeler arası korelasyonların incelenmesidir (Ferrara, Huyny&Baghi, 1997; Akt: Bulut, 2015). Bu çalışmada yerel bağımsızlıđı test etmek için belli bir yetenek ranđı arasındaki (yüksek yetenek ve düşük yetenek grupları) maddelerarası korelasyonlardan yararlanılmıştır. Alt grup ile üst grupları belirlemek için ham puanların %20 ve %80'lik dilimleri kullanılır. Sınırlı yetenek düzeyindeki bireyler için elde edilen varyans ve kovaryans veya korelasyon matrislerinin köşegeninde yer alan elementlerin 0 veya 0'a çok yakın olması yerel bağımsızlık varsayımının karşılandıđını göstermektedir (Hambleton, 1991; McDonald, 1981; Akt: Bulut, 2015). Bu noktadan hareketle bu çalışmada düşük ve yüksek yetenek gruplarındaki bireylerin maddelere verdikleri cevaplar üzerinden elde edilen maddeler arası korelasyonlar elde edilmiştir. Elde edilen sonuçlara göre her iki yetenek grubunda da maddeler arası korelasyonlar çok düşük çıkmıştır. Böylece birleşik test için yerel bağımsızlık varsayımının karşılandıđı sonucuna ulaşılabilir.

Son olarak TBPOMTK için monotonluk varsayımı test edilmiştir. Bu çalışmada monotonluk varsayımının test edilmesi için R 3.0.2. yazılımı için Van Der Ark (2007) tarafından geliştirilen kullanılan Mokken paketinden yararlanılmıştır. Monotonluk varsayımının sonuçlarını yorumlamaya geçmeden önce önemli sembol ve kısaltmaları açıklamakta yarar vardır. (#AÇ) her bir madde için aktif çift sayısını, (#GMİ) gizil monotonluğun ihlalinin miktarı, (#GMİ/#AÇ) her madde çifti için ortalama olarak monotonluk ihlalinin miktarını, (maxGMİ) monotonluk ihlalinin miktarının en büyük deđerini, (TOP) toplam monotonluk ihlalinin miktarını, (TOP/#AÇ) her madde çifti için toplam monotonluk ihlalinin miktarını göstermektedir. Bütün bu deđerler 0'dan anlamlı bir şekilde büyük olduđu takdirde gizil monotonluk varsayımı ihlal edilmiş olur (Van der Ark, 2007). Yorumlama için önemli olan bir diđer gösterge ise madde ölçeklenebilirlik katsayısı olan  $H_j$ 'dir. Her bir maddeye ait olan  $H_j$ 'lerin 0,30'dan küçük olması maddenin ayırtedicilik bakımından zayıf olduđunu gösterir.  $H_j$  madde ayırtedicilik katsayısı olarak yorumlanır (Sijtsma ve Molenaar, 2002). Elde edilen sonuçlar dođrultusunda Din Kültürü ve Ahlak Bilgisi, İngilizce ve Fen bilgisi alt testleri için ise bazı maddeler tarafından varsayımın karşılanmadıđı görülmüştür. Türkçe ve Matematik alt testlerinde yer alan tüm maddelerin ise varsayımı karşılandıđı tespit edilmiştir. Özellikle Matematik testinde yer alan bütün maddeler için (#AÇ), (#GMİ), (#GMİ/#AÇ), (maxGMİ), (TOP) ve (TOP/#AÇ) deđerleri 0 çıkmıştır. Benzer şekilde matematik testinde yer alan bütün maddelere ait ölçeklenebilirlik katsayısı 0,30'un üzerindedir.

Elde edilen sonuçlardan hareketle TBPMTK ve TBPOMTK analizlerinde tek boyutluluk kanıtlarının daha güçlü olduğu belirlenen ve TBPOMTK için monotonluk varsayımını da karşılayan Matematik alt testinin sonuçlarından yararlanılmıştır. ÇBMTK için ise daha önce ifade edildiği gibi Din Kültürü ve Ahlak Bilgisi ve Fen Bilgisi testlerinden belli maddelerin seçilmesi ile oluşturulan iki faktörlü birleşik testin sonuçlarından yararlanılmıştır. Tek ve çok boyutlu veri analizi için kullanılan testlerin belirlenmesinden sonra, tek ve çok boyutlu MTK kapsamında hangi modellerin kullanılacağı belirlenmesi gerekir. TBPMTK için matematik testinin sonuçlarına 2 PLM ile 3 PLM uygulanmıştır. 2 PLM için kestirilen parametre sayısı 40 ve elde edilen -2 LL değeri 1419674.370'tir. 3 PLM için ise kestirilen parametre sayısı 60 ve elde edilen -2 LL değeri 1702461.230'dur. 20 serbestlik derecesinde elde edilen fark değeri ise 282786.86'dır. Elde edilen bu sonuç 20 (60-40) serbestlik derecesindeki  $\chi^2$  kritik değeri (31.410) ile kıyaslandığında, anlamlı çıkmıştır. Dolayısıyla, -2 LL değeri düşük olan 2PLM'nin 3PLM'ye göre anlamlı farklılık yarattığı; diğer bir ifadeyle 2PLM'nin veriye daha iyi uyum sağlayan model olduğu söylenebilir. Bununla birlikte Embretson ve Reise (2000), çoktan seçmeli test maddesi ile çalışıldığı durumlarda 3PLM'nin, kişilik verileri ile çalışıldığı durumlarda ise 1PLM veya 2PLM'den birinin kullanılmasını önerir. Ancak modellere ait -2LL değerleri de göz önünde bulundurularak TBPMTK modeli olarak 2PLM'nin uygulanmasına karar verilmiştir.

ÇBMTK için birleşik testin sonuçlarına genişletilmiş M2PLM ile M3PLM uygulanmıştır. M2PLM için kestirilen parametre sayısı 40 ve elde edilen -2 LL değeri 1214446.23'dur. M3PLM için ise kestirilen parametre sayısı 60 ve elde edilen -2 LL değeri 1214490.81'dir. 20 serbestlik derecesinde elde edilen fark değeri ise 44.58'dir. Elde edilen bu sonuç 20 (60-40) serbestlik derecesindeki  $\chi^2$  kritik değeri (31.410) ile kıyaslandığında, anlamlı çıkmıştır. Bu durumda analizler için tercih edilen model -2 LL değeri daha küçük olan genişletilmiş M2PLM'dir.

Son olarak TBPOMTK için Andries van der Ark (2007) tarafından önerilen İkili Monotonluk Modeli (IMM) ve Monoton Homojenlik Modeli (MHM) arasından MHM'nin kullanılmasına karar verilmiştir. Bunun nedeni IMM tarafından açıklanabilen her veri setinin daha zayıf bir Modeli (MHM) tarafından açıklanabilmesidir (Andries van der Ark, 2007)

TBPOMTK analizlerinde R 3.0.2. yazılımı için Van Der Ark (2007) tarafından geliştirilen MOKKEN paketinden, tek ve çok boyutlu MTK analizinde ise Cai (2017) tarafından önerilen FlexMIRT 3.5 yazılımından yararlanılmıştır. TBPMTK parametre kestirimi 2 PLM'ye göre yapılmış, çok boyutlu MTK analizleri ise genişletilmiş 2 PLM'ye göre gerçekleştirilmiştir. Hem TBPMTK hem de ÇBMTK analizlerinde hata değerleri Cai'nin (2008) EM algoritmasının hata değerleri kestirilerek belirlenmiştir.

Bu çalışmada madde puanlarından değil test puanlarından yararlanılmıştır. Bu nedenle toplam test puanları için ortalama güçlük ve ayırtedicilik katsayıları kestirilmiştir. Her üç kurama ait modeller için de testin ortalama güçlüğü ve ayırtedicilik düzeyleri evrenden birer kez çekilen her bir örneklem için ayrı ayrı hesaplanmıştır. Her bir örneklem için kestirilen parametre ortalamalarının evren değerden ne kadar farklı olduğu incelenmiştir. Ancak bu fark istatistiksel olarak test edilmemiştir. Yapılan yorumlar sadece büyüklük küçüklük ilişkisi çerçevesinde yapılmıştır.

Çalışmanın odağında olan parametre değişmezliğinin incelenmesi amacıyla, parametre değişmezliğinin göstergesi olarak standart hata ortalamalarından (SHO) yararlanılmıştır (Koğar, 2014; Sünbül, 2011). Ancak bulgular yorumlanırken parametre ortalamalarının kestirimlerine ait SHO'ların yanı sıra parametre ortalamalarının da örneklem büyüklüğünden nasıl etkilendiği incelenmiştir. Bu amaçla her bir örneklem için kestirilen parametre ortalamalarına ait SHO'ların evren değerden ne kadar farklı olduğu incelenmiştir. Tıpkı parametre ortalamalarında olduğu gibi bu fark istatistiksel olarak test edilmemiştir. Araştırmanın amacı betimleme olduğundan elde edilen bulgular sadece büyüklük-küçüklük ilişkisi çerçevesinde yorumlanmıştır.

Çalışmaya ait bulguların elde edilmesi genel olarak iki başlık altında toplanabilir. Bunlardan ilki tek boyutlu olan Matematik testi tek boyutlu parametrik ve parametrik olmayan MTK altında modellenmiş, bu modellemeye göre madde parametreleri kestirilmiştir. İkinci boyutta ise iki boyutlu olan birleşik testin tek boyutlu olduğu kabulünden (iki boyutluluğun ihmalinden) yola çıkılarak tek boyutlu parametrik ve parametrik olmayan MTK'ye göre tek boyutlu bir model altında analiz

edilmiştir. Bunlara ek olarak iki boyutlu olan birleşik test doğasına uygun olarak ÇBMTK'ye göre modellenerek analiz edilmiş ve bu analizlerin sonucunda madde parametreleri kestirilmiştir.

### *Araştırmanın İç ve Dış Geçerliđi*

Araştırmanın iç geçerliđi bağımlı deđişkendeki deđişimlerin, bağımsız deđişkenlerle açıklanma derecesi ile ilgilidir. Bu çalışmada madde ve yetenek kestirimleri ile bu kestirime ait güvenilirlik düzeylerindeki deđişim, örneklem büyüklüğü ve MTK'nin farklı uygulamaları tarafından açıklanabildiđi için araştırmanın iç geçerliđi sağlanmıştır (Fraenkel & Wallen, 2006).

Araştırmanın dış geçerliđi bulguların genellenebilirlik derecesi ile ilgilidir. Bu çalışmadan elde edilen bulguların genellenebilirliđi kullanılan örneklem büyüklükleri, MTK uygulamaları ve kullanılan testlerin konu alanı ile sınırlı olduđu için araştırmanın dış geçerliđini bu çerçevede incelemek gerekir. Dolayısıyla kullanılan örneklem büyüklükleri, kullanılan MTK uygulamaları ve testlere ait konu alanı çerçevesinde araştırma sonuçlarının genellenebileceđi düşünülmektedir (Fraenkel ve Wallen, 2006).

## **BULGULAR**

### *Birinci Alt Problemin Çözümüne İlişkin Bulgular*

Birinci alt problemin çözümü için TBPMTK ve TBPOMTK'ye göre kestirilen matematik testi'ne ait madde parametreleri ve standart hata ortalamaları (SHO) Tablo 1'de görülmektedir.

Tablo 1. BPMTK ve TBPOMTK'ye Göre Kestirilen Matematik Testi'ne Ait Madde Parametreleri ve Standart Hata Ortalamaları

Örneklem Büyüklüğü	TBPOMTK				TBPMTK			
	H	H <sub>SHO</sub>	p	p <sub>SHO</sub>	a	a <sub>SHO</sub>	b	b <sub>SHO</sub>
50	0.32	0.09	0.48	0.109	1.07	0.32	0.09	0.37
100	0.39	0.07	0.45	0.109	1.31	0.28	0.24	0.23
200	0.33	0.05	0.43	0.107	1.32	0.22	0.34	0.18
500	0.33	0.03	0.43	0.107	1.41	0.15	0.34	0.10
1000	0.33	0.02	0.43	0.107	1.48	0.11	0.30	0.07
5000	0.33	0.01	0.43	0.107	1.50	0.05	0.29	0.03
Evren	0.33	0.00	0.43	0.107	1.52	0.01	0.30	0.01

TBPOMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin ilki ayırtediciliđin göstergesi olan ortalama H parametreleridir. Tabloda verilen H ortalamaları incelendiğinde evren deđerinin 0.33 olduđu görülmektedir. Evren deđere görel olarak en uzak H ortalaması 100 örneklem büyüklüğünden kestirilmiştir (H=0.39). Evren deđere en uzak H ortalaması bile evren deđerden çok farklı deđildir. Bu nedenle örneklemekten kestirilen H ortalamalarının büyüklük olarak evren deđere yakın olduđu sonucuna ulaşılabılır. Bununla birlikte 200 örneklem büyüklüğünden itibaren ise evren deđeri yansıtacak düzeyde kararlı bir yapıya sahip H'lerin kestirildiđi tabloda görülmektedir (H=0.33). H ortalamasına ait SHO'nun evren deđerini 0'a çok yakındır. Evren deđerini en az yansıtan örneklem büyüklüğünün en küçük örneklem olan 50 olduđu (H<sub>SHO</sub>=0.09) ve örneklem büyüklüğü arttıkça H parametresi ortalamasına ait SHO'ların evren deđerine yaklaştığı yine tabloda görülmektedir. Bunun yanı sıra örneklem büyüklüğü kaç olursa olsun SHO'lar büyüklük olarak evren deđere yakın büyüklükte kestirilmiştir. Elde edilen bu sonuçlara göre evren deđere yakın SHO'lar ile ortalama H parametresi kestirmek için büyük örneklemekten kullanılmasının zorunlu olmadığı sonucuna ulaşılabılır.

TBPOMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin bir diğeri ise güçlük göstergesi olan ortalama p parametreleridir. Evren değerine en uzak p ortalaması en küçük örnekleme kestirilmiştir ( $p=0.48$ ). Ancak Tablo 1'de görüldüğü gibi evren değere en uzak p ortalaması bile evren değerden çok farklı değildir.

Ortalama p parametresine ait SHO'nun evren değeri ise 0.107 olarak kestirilmiştir. Evren değere göreli olarak en uzak SHO ise 50 ve 100 örnekleme büyüklüklerinden kestirilmiştir ( $H_{SHO} = 0.109$ ). SHO 200 örnekleme büyüklüğünde ise en düşük değerini almıştır ( $p_{SHO} = 0.107$ ) ve bu değer evren değere eşit olduğu söylenebilir. Tıpkı H ortalamalarına ait SHO'larda olduğu gibi örneklemlerden kestirilen p ortalamalarına ait SHO'lar ile evren değeri arasında büyük farklılıkların olmadığı belirlenmiştir. Bununla birlikte 200 örnekleme büyüklüğü ile evren arasındaki örneklemlerde de SHO büyüklüklerinde bir değişim olmamıştır. Sonuç olarak TBPOMTK'ye göre kestirilen madde parametresi ortalamaları ve madde parametresi ortalamalarına ait SHO'ların 200 örnekleme büyüklüğünden itibaren evreni yansıtacak düzeyde kararlı bir yapıya sahip olduğu görülmüştür. Buna göre TBPOMTK için 200 örnekleme büyüklüğü itibariyle parametre değişmezliğinin sağlandığı sonucuna ulaşılmıştır.

TBPMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin ilki ayırtediciliğin göstergesi olan ortalama a parametreleridir. Örneklemlerden kestirilen a parametresi ortalamasının örnekleme büyüklüğüne bağlı olarak 1.07 ile 1.52 arasında değerler aldığı Tablo 1'de görülmektedir. Ortalama a parametresinin evren değeri 1.52 olarak kestirilmiştir. Ortalama a parametresi evren değerine en uzak değerini en küçük örnekleme büyüklüğü olan 50'de almıştır ve örnekleme büyüklüğü arttıkça a parametresi ortalamasına ait değerlerin de artma eğiliminde olduğu ve evren değerine yaklaştığı yine tabloda görülmektedir. 1000 örnekleme büyüklüğünden itibaren a parametresi ortalamaları evren değerine çok yakın büyüklükte kestirilmiştir.

TBPMTK için a parametresi ortalamasına ait SHO'nun evren değeri ise 0.01 olarak kestirilmiştir. Evren değerine en uzak SHO'nun 50 örnekleme büyüklüğünden kestirildiği tabloda görülmektedir ( $a_{SHO} = 0.32$ ). Örnekleme büyüklüğünün artmasıyla a parametresi ortalamasına ait SHO azalma eğilimi göstererek evren değere yaklaşmıştır. Tabloda görüldüğü gibi 500 örnekleme büyüklüğü itibariyle evren değere çok daha yakın bir SHO ile a parametresinin kestirildiği sonucuna ulaşılabilir.

TBPMTK'ye göre kestirilen tek boyutlu teste ait madde parametrelerinin bir diğeri ise güçlük göstergesi olan ortalama b parametresidir. Ortalama b parametresi ait evren değer 0.30 olarak kestirilmiştir. Evren değerine en uzak b parametresi ortalaması 50 örnekleme büyüklüğünden elde edilmiştir ( $b=0.09$ ). Evren değeri en iyi yansıtan örnekleme büyüklüğü ise 1000'dir. Elde edilen sonuçlara göre b parametresi ortalamasının, örnekleme büyüklüğünün değişmesi ile düzenli bir artma ya da azalma eğiliminde olduğunu söylemek mümkün değildir. Ancak 1000 örnekleme büyüklüğü itibariyle b parametresi ortalamasına ait değişimlerin daha az olduğu söylenebilir. Bu durumda 1000 örnekleme büyüklüğü itibari ile b parametresi ortalamalarının evrene yakın büyüklükte olduğu sonucuna ulaşılabilir.

TBPMTK için b parametresi ortalamasına ait SHO'nun evren değeri ise 0.01 olarak kestirilmiştir. Evren değerine en uzak SHO en küçük örnekleme büyüklüğü olan 50 üzerinden kestirilmiştir ( $b_{SHO} = 0.37$ ). Ortalama b parametresi ait SHO'nun örnekleme büyüklüğünün artmasına bağlı olarak istikrarlı bir şekilde evren değerine yaklaştığı Tablo 1'de görülmektedir. Tabloda görüldüğü gibi 500 örnekleme büyüklüğünden itibaren de evren değerine daha yakın büyüklükte bir SHO ile b parametresinin kestirildiği sonucuna ulaşılabilir.

TBPMTK üzerinden elde edilen madde parametresi kestirimlerinin incelenmesinin sonucunda, a ve b parametreleri ile a ve b parametrelerine ait SHO'lar sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanmadığı sonucuna ulaşılmıştır.



### İkinci Alt Problemin Çözümüne İlişkin Bulgular

İkinci alt problemin çözümü için birleşik testin sonuçlarından ÇBMTK ile TBPMTK'ye göre kestirilen birleşik test ait madde parametreleri ve standart hata ortalamaları (SHO) Tablo 2'de görölmektedir.

Tablo 2. Tek ve Çok Boyutlu MTK'ye Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları

Örneklem Büyüklüğü	ÇBMTK						TBPMTK			
	a <sub>1</sub>	a <sub>1</sub> SHO	a <sub>2</sub>	a <sub>2</sub> SHO	d	d SHO	a	a SHO	b	b SHO
50	0.98	1.47	0.51	0.75	1.12	0.99	1.23	0.54	-0.82	0.48
100	0.60	0.70	0.44	0.45	0.77	0.53	1.24	0.75	-0.74	0.47
200	0.68	0.57	0.42	0.30	1.02	0.43	1.48	0.52	-0.79	0.36
500	0.68	0.43	0.40	0.20	0.87	0.36	1.68	0.69	-0.55	0.19
1000	0.67	0.25	0.41	0.14	0.84	0.20	1.76	0.16	-0.48	0.09
5000	0.67	0.11	0.42	0.06	0.88	0.08	1.94	0.07	-0.49	0.03
Evren	0.41	0.01	0.43	0.01	-0.18	0.01	1.52	0.01	0.30	0.01

ÇBMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin ilki ayırtediciliđin göstergesi olan ortalama a<sub>1</sub> ve a<sub>2</sub> parametreleridir. ÇBMTK'den kestirilen a<sub>1</sub> ortalamasının evren deđeri 0.41 olarak kestirilmiştir. Ortalama a<sub>1</sub> için evren deđere en uzak kestirim en küçük örneklem büyüklüğü olan 50 üzerinden elde edilmiştir. a<sub>1</sub> ortalamasının evren deđerini iyi yansıtan bir örneklem büyüklüğü ise bulunmamaktadır. Ortalama a<sub>2</sub>'ye ait evren deđeri ise 0.43 olarak kestirilmiştir. Ortalama a<sub>2</sub>'nin evren deđerine en uzak kestirim en küçük örneklem büyüklüğünden elde edilmiş ve 100 örneklem büyüklüğünden itibaren ise evren deđerine yakın büyüklükte a<sub>2</sub> ortalaması kestirimleri elde edilmiştir. Elde edilen sonuçlara göre a<sub>1</sub> ve a<sub>2</sub> ortalamalarının, örneklem büyüklüğünün deđişimine bađlı olarak düzenli bir artma ya da azalma eğiliminin olduđunu söylemek mümkün deđildir.

ÇBMTK için ayırtedicilik parametrelerinin ortalamasına ait SHO'nun evren deđerinin ise 0,01 olarak kestirildiđi tabloda görölmektedir. Evren deđerine en uzak SHO kestirimi en küçük örneklem büyüklüğü üzerinden elde edilmiştir (a<sub>1</sub> SHO=1.47; a<sub>2</sub> SHO=0.75). Örneklem büyüklüğünün artmasıyla her iki ayırtedicilik parametresi ortalamasına ait SHO da azalma eğilimi göstermiştir ve evren deđere yaklaşmıştır a<sub>1</sub> ortalamasının 5000 örneklemden itibaren ve a<sub>2</sub> ortalamasının 1000 örneklemden itibaren evren deđere yakın büyüklükte bir SHO ile kestirildiđi Tablo 2'de görölmektedir. Bununla birlikte her iki ayırtedicilik parametresi ortalamasına ait SHO da örneklemlerden evrene dođru sürekli bir deđişim göstermektedir.

Sonuç olarak ÇBMTK'den kestirilen ayırtedicilik parametrelerine ait ortalama deđerleri de SHO'ları da sürekli bir deđişim gösterdikleri için parametre deđişmezliđinin sağlanamadıđı sonucuna ulaşılmıştır.

ÇBMTK'den kestirilen d parametresi ortalamasının evren deđeri (d=-0.18) olarak kestirilmiştir. Evren deđerine en uzak d parametresi ortalaması en küçük örneklem büyüklüğünden elde edilmiştir (d=1.12). Örneklemlerden elde edilen d parametresi ortalamaları incelendiđinde evren deđerden farklı oldukları görölmektedir. Elde edilen sonuçlara göre d parametresi ortalamasının örneklem büyüklüğünün deđişimine bađlı olarak düzenli bir artma ya da azalma eğiliminin olmadıđı da görölmektedir. Ancak 500 ve 5000 örneklem büyüklükleri arasındaki deđişimin diđer örneklem büyüklüklerine göre daha az olduđu Tablo 2'de görölmektedir.

ÇBMTK için d parametresinin ortalamasına ait SHO'nun evren deđeri ise 0.01 olarak kestirilmiştir. Evren deđerine en uzak SHO kestirimlerinin en küçük örneklem büyüklüğünden elde edildiđi ve örneklem büyüklüğü arttıkça SHO kestirimlerinin de evren deđerine yaklaştıđı tabloda görölmektedir.

d parametresi ortalamasına ait SHO'ların örneklem boyunca farklı değerler aldığı yine Tablo 2'de görülmektedir.

Tıpkı ayrıtedicilik parametrelerinin ortalamalarında olduğu gibi d parametresi ortalaması ile d parametresi ortalamasına ait SHO da örneklemelerden evrene doğru sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

TBPMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin ilki ayrıtediciliğin göstergesi olan ortalama a parametreleridir. Örneklemelerden kestirilen a parametresi ortalamasının örneklem büyüklüğüne bağlı olarak 1.23 ile 1.94 arasında değerler aldığı Tablo 2'de görülmektedir. a parametresi ortalamasının evren değeri 1.52 olup evren değerini en iyi yansıtan örneklem büyüklüğü 200'dür. Daha genel bir ifade ile bu çalışmada kullanılan orta büyüklükteki örneklemelerde evren değere daha yakın kestirimler elde edilirken, uçtaki (en büyük ve en küçük) örneklemelerde evren değere daha uzak kestirimler elde edilmiştir. Ayrıca a parametresi ortalamasının evren değerinin tek boyutlu testten elde edilen evren değer ile aynı olduğunu da belirtmek gerekir. Buradaki karşılaştırmayı anlamlı hale getiren nokta, karşılaştırılan değerlerin aynı evren üzerinden fakat iki farklı boyutluluğa sahip yapılardan elde edilmesidir. Bu noktadan hareketle ulaşılan sonuç tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen a parametresine ait evren değerlerin ortalamasının değişmediğidir.

TBPMTK için a parametresi ortalamasına ait SHO'nun evren değeri 0.01 olarak kestirilmiştir. Evren değere büyüklük olarak en uzak SHO 100 örneklem büyüklüğünden elde edilmiştir ( $a_{SHO}=0.75$ ). a parametresi ortalamasına ait SHO'nun örneklem büyüklüğüne bağlı olarak düzenli bir artma ya da azalma eğilimi gösterdiği söylenemez. Ortalama a parametresine ait SHO ardışık örneklem büyüklüklerinin bazılarında evren değere daha uzak bir değer alırken bazılarında daha yakın bir değer almıştır. Çok boyutlu veri yapısının tek boyutlu bir model altında modellenmesi bu düzensizliğin başlıca sebebidir. Ortalama a parametresine ait SHO'ların 1000 örneklem büyüklüğüne kadar evreni yansıtmayan değerler aldığı yine tabloda görülmektedir. 500 örneklem büyüklüğünden itibaren ise istikrarlı bir şekilde azalarak evren değere yaklaşmıştır. Bunun yanı sıra a parametresi ortalamasına ait SHO'lar en küçük örneklem büyüklüğünden en büyük örneklem büyüklüğüne kadar sürekli bir değişim göstermektedir.

TBPMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin bir diğeri ise güçlüğün göstergesi olan ortalama b parametreleridir. b parametresi ortalamasının evren değeri 0.30 olarak kestirilmiştir. Evren değerine büyüklük olarak en uzak kestirim örneklem büyüklüğünün 50 olduğu durumda elde edilmiştir ( $b=-0.82$ ). Bununla birlikte b parametresi ortalamasının evren değerini iyi yansıtan bir örneklem büyüklüğü bulunmamaktadır. a parametresi için ifade edilen "tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen a parametresine ait evren değerlerin değişmediği" sonucu b parametresi için de geçerlidir. Bir başka ifadeyle kullanılan veri seti ister tek boyutlu olsun ister çok boyutlu olsun kestirilen b parametrelerine ait evren değerler birbiri ile aynıdır. Elde edilen sonuçlara göre b parametresi ortalamasının örneklem büyüklüğünün değişimine bağlı olarak net bir artma ya da azalma eğiliminin olmadığı görülmektedir.

TBPMTK için b parametresi ortalamasına ait SHO'nun evren değeri 0.01 olup, evren değerine büyüklük olarak en uzak b parametresi ortalamasına ait SHO en küçük örneklem büyüklüğü olan 50 üzerinden kestirilmiştir ( $b_{SHO}= 0.48$ ). Örneklem büyüklüğü arttıkça kestirilen SHO'lar da evren değerine yaklaşmıştır. 500 örneklem büyüklüğü itibariyle evren değerine çok yakın büyüklükte bir SHO ile b parametresinin kestirildiği de Tablo 2'de görülmektedir SHO ( $b_{SHO}=0.19$ ). Tıpkı a parametresi ortalamasında olduğu gibi b parametresi ortalamasına ait SHO'lar da farklı örneklem büyüklüklerinde farklı değerler almıştır.

Sonuç olarak, TBPMTK'den kestirilen madde parametrelerinin ve madde parametrelerine ait SHO'ların incelenmesinin sonucunda, a ve b parametreleri ile a ve b parametrelerine ait SHO'lar sürekli bir değişim gösterdiği için parametre değişmezliğinin sağlanamadığı sonucuna ulaşılmıştır.

### Üçüncü Alt Problemin Çözümüne İlişkin Bulgular

Üçüncü alt problemin çözümü için birleşik testin sonuçlarından TBPMTK ve TBPOMTK'ye göre kestirilen birleşik teste ait madde parametreleri ve standart hata ortalamaları (SHO) Tablo 3'te görölmektedir.

Tablo 3. TBPMTK ve TBPOMTK'ye Göre Kestirilen Birleşik Teste Ait Madde Parametreleri ve Standart Hata Ortalamaları

Örneklemler Büyüklüğü	TBPMTK				TBPOMTK			
	a	asho	b	bsho	H	Hsho	p	psho
50	1.23	0.54	-0.82	0.48	0.46	0.10	0.65	0.097
100	1.24	0.75	-0.74	0.47	0.38	0.07	0.65	0.096
200	1.48	0.52	-0.79	0.36	0.39	0.05	0.68	0.096
500	1.68	0.69	-0.55	0.19	0.38	0.03	0.65	0.096
1000	1.76	0.16	-0.48	0.09	0.40	0.02	0.65	0.096
5000	1.94	0.07	-0.49	0.03	0.40	0.01	0.66	0.095
Evren	1.52	0.01	0.30	0.01	0.33	0.00	0.43	0.107

Birleşik testin TBPMTK'ye göre analiz edilmesi ile elde edilen sonuçlar ikinci alt problemin çözümünde detaylı bir şekilde ifade edilmiştir. Birleşik testin TBPOMTK'ye göre analiz edilmesiyle elde edilen ilk parametre H parametresidir.

H ortalamasının evren değerinin 0.33 olduđu tabloda görölmektedir ( $H=0.33$ ). Evren değere en uzak  $H_i$  ortalaması 50 örneklem büyüklüğünden kestirilmiştir ( $H=0.46$ ). Bu durumda tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden kestirilen H parametresine ait evren değerlerin deđişmediđi sonucuna ulaşılabılır. H ortalaması açısından evreni en iyi yansıtan örneklem büyüklükleri 100 ve 500 örneklem büyüklükleridir.

H ortalamasına ait SHO için 0'a çok yakın büyüklükte bir evren değeri kestirilmiştir ( $H_{SHO}=0.00$ ). Evren değere büyüklük olarak en uzak SHO en küçük örneklem büyüklüğü olan 50 üzerinden kestirilmiştir ( $H_{SHO}=0.10$ ). H ortalamasına ait SHO'ların örneklem büyüklüğü arttıkça istikrarlı bir şekilde azalarak evren değerine yaklaştığı görölmektedir. Bunun yanı sıra örneklemlerden elde edilen SHO'ların evren değerinden büyük farklılıklar göstermediđi belirlenmiştir. Bu durumda çok boyutlu bir veri üzerinden küçük örneklem büyüklükleri ile çalışıldığı zaman, evren değere yakın SHO'lar ile ayırtedicilik parametresinin kestirilmesinin mümkün olduđu söylenebilir. Ayrıca tek ve çok boyutlu testler TBPOMTK'ye göre analiz edildiğinde elde edilen H parametrelerine ait SHO'ların birbirlerine çok yakın değerler aldıkları görölmüştür. Bu durumda tek boyutluluğun ihlal edildiđi durumda H ortalamalarının benzer SHO değerleri ile kestirildiđi sonucuna ulaşılabılır. Bununla birlikte SHO'lar örneklemler boyunca büyüklük olarak sürekli bir deđişim göstermektedir.

Tablo 3'te göröldüğü gibi H ortalamaları ve SHO'ları örneklemler boyunca birbirleri ile farklılık gösterdikleri için H ortalaması için de deđişmezliđin sağlanamadığı sonucuna ulaşılabılır. Tek boyutluluğun sağlandığı durumda ise küçük bir örneklem büyüklüğünden itibaren deđişmezliđin sağlandığı birinci alt problemin bulgularında ifade edilmiştir. Tek boyutlu modellerin kullanılması için sağlanması gereken en önemli varsayım olan tekboyutluluğun ihlali, TBPOMTK sonuçlarını deđişmezlik açısından etkilemiştir.

TBPOMTK'ye göre kestirilen iki boyutlu teste ait madde parametrelerinin bir diğeri ise güçlüğün göstergesi olan ortalama p parametresidir. p ortalamalarına göre testin örneklemlerde görelilik olarak daha kolay olduđu da Tablo 3'te görölmektedir. Bununla birlikte p ortalamalarının evren değeri ile büyüklük olarak evrene en uzak değeri arasında büyük farklılıkların olmadığı da görölmektedir. p ortalamasının evren değeri 0.43 olarak kestirilmiştir ( $p=0.43$ ). Büyüklük olarak evrene en uzak p ortalaması ise 200 örneklem büyüklüğünden kestirilmiştir. Tıpkı H ortalamasının evren değerinde

olduğu gibi olduğu gibi p ortalamasına ait evren değerinin de tek boyutlu testin TBPOMTK'ye göre analiz edildiğinde kestirilen evren değer ile aynı olduğu görülmüştür. Bu durumda yine tek boyutluluk varsayımının sağlandığı ve sağlanmadığı durumlarda bir başka ifadeyle tek ve çok boyutlu testlerden elde edilen ortalama p parametresine ait evren değerlerin değişmediği sonucuna ulaşılabilir.

Tablo 3'te görüldüğü gibi p ortalamalarına ait SHO'nun evren değeri 0.107 olarak elde edilmiştir ( $p_{SHO}=0.107$ ). Evren değere en uzak SHO değeri ise 5000 örneklem üzerinden kestirilmiştir ( $p_{SHO}=0.095$ ). Tek boyutlu test TBPOMTK'ye göre analiz edildiğinde evrenden elde edilen p ortalamalarının aynı SHO ile kestirilmesi dikkat çeken bir bulgudur. Buna göre tek boyutluluğun sağlandığı ve ihlal edildiği durumda p ortalamalarına ait SHO'ların aynı olduğu sonucuna ulaşılabilir. p ortalamalarına ait SHO'ların evren değeri ile evrene en uzak değeri arasında büyük farklılıkların olmadığı Tablo 3'te görülmektedir. Bu nedenle örneklem büyüklüğünün değişmesinin p ortalamalarının kestirildiği SHO'lar üzerinde büyük farklılıklara neden olmadığı söylenebilir. Bu durumda tek boyutluluğun sağlanmadığı ve büyük örneklemelere ulaşamadığı takdirde evren değere yakın büyüklükteki SHO'lar ile ortalama p parametresinin kestirilmesinin mümkün olduğu sonucuna ulaşılabilir. p ortalamasına ait SHO'lar en büyük ve en küçük örneklem büyüklüklerinde farklı değerler almış, diğer örneklem büyüklüklerinde ise değişmemiştir.

Ayrıca ortalama p değerleri ve SHO'ları en büyük ve en küçük örneklem büyüklüklerinde farklı değerler aldığı için p ortalaması için de değişmezliğin sağlanmadığı sonucuna ulaşılmıştır. Tek boyutluluğun sağlandığı durumda ise küçük bir örneklem büyüklüğünden itibaren değişmezliğin sağlandığı birinci alt problemin bulgularında ifade edilmiştir. H ortalaması için de benzer sonuçlar ortaya konmuştur.

TBPMTK ile TBPOMTK'den elde edilen sonuçlar karşılaştırılacak olursa, H ortalamalarının a parametresi ortalamalarına göre küçük örneklem büyüklüklerinde bile evren değerine yakın SHO'lar ile kestirildiği Tablo 3'te görülmektedir. Tablo 3'ten elde edilen sonuçlara göre, a parametresi ortalamasını evren değerine yakın bir SHO ile kestirebilmek için en az 1000 hatta 5000 örneklem büyüklükleri ile çalışmak gerekmektedir. Daha küçük bir örneklemde evren değerine yakın bir SHO ile ortalama ayırtedicilik parametreleri kestirilebildiği için TBPOMTK'nin TBPMTK'ye göre daha avantajlı olduğu sonucuna ulaşılabilir. Tek boyutlu test parametrik ve parametrik olmayan MTK'ye göre analiz edildiğinde de benzer bir sonuca ulaşılmıştır. O halde tek boyutluluk varsayımı karşılanırsa da karşılanmasa da evren değerine yakın büyüklükte bir SHO ile ortalama ayırtedicilik parametresi kestirmek için, TBPMTK için büyük örneklemelerin kullanılması gerekmektedir. TBPOMTK için ise böyle bir sınırlama yoktur. Benzer yorumlar parametrik ve parametrik olmayan MTK'den elde edilen ortalama güçlük parametreleri için de geçerlidir. Evren değerine yakın büyüklükte bir SHO ile ortalama b parametresinin kestirilmesi için en az 1000 örneklem ile çalışmak gerekmektedir. Ancak evren değerine yakın bir SHO ile ortalama p parametresinin kestirilmesi için daha küçük örneklem büyüklükleri ile de çalışılabileceği Tablo 3'te görülmektedir.

## SONUÇLAR ve TARTIŞMA

Birinci alt problemin bulgularından elde edilen sonuçlar doğrultusunda tek boyutlu testin TBPOMTK'ye göre analiz edilmesiyle 200 örneklem büyüklüğü itibariyle madde parametresi değişmezliğini sağlandığı görülmüştür. Bu durumda Büyük örneklem büyüklüklerine ulaşamadığı durumda ortalama madde parametresi kestirimi amacıyla küçük örneklemelerden elde edilen veriler için TBPOMTK uygulamasından yararlanılabildiği sonucuna ulaşılmıştır. Elde edilen bu sonuç MTK'nin özellikle okul uygulamalarında eğitimin farklı kademelerinde uygulanan sınavlara ait madde parametresi kestirimlerine imkân sağlaması açısından büyük bir önem taşımaktadır.

Tek boyutlu testin TBPMTK'ye göre analiz edilmesiyle TBPMTK için örneklem büyüklüğü arttıkça a parametresi ortalamasına ait SHO'nun azaldığı ve evren değerine yaklaştığı görülmüştür. Thissen ve Wainer (1982) parametre kestirimi için 10.000 ve daha fazla örnekleme ihtiyacı olduğunu, Goldman ve Raju (1986) ise a parametresinin doğru kestirimi için en az 1000 kişilik örnekleme ihtiyacı duyulduğunu ifade etmişlerdir. Örneklem büyüklüğü 5000'i geçtiği halde a parametresine ait ortalama



deđerin ve a parametresi ortalamasına ait standart hata ortalamasının deđişmeye devam etmesi, Thissen ve Wainer (1982) bulgularını destekler niteliktedir.

Tek boyutlu testin TBPMTK'ye göre analiz edilmesiyle elde edilen b parametresi ortalamasına ait SHO örneklemler büyüklüğü arttıkça azalma eğilimi göstermiş ve evren değerine yaklaşmıştır. b parametresi ortalaması ise örneklemler büyüklüğüne bađlı olarak düzenli bir artma ya da azalma eğilimi göstermemiştir. Sünbül (2011) bu durumu "örneklemler büyüklüğünün b parametresi üzerindeki önemsizliđi" şeklinde ifade etmiştir. Bununla birlikte Hulin, Lissak ve Drasgow (1982) 2 PLM için 1000'den daha büyük örneklemlerde kestirilen parametrelerin önemli deđişikliklerin olmadığını ortaya koymuştur. Bu nedenle elde edilen bulgular Hulin, Lissak ve Drasgow (1982)'nin bulgularıyla tutarlılık göstermektedir.

Sonuç olarak küçük örneklemler büyüklüklerinden elde edilen madde parametrelerinin evreni yansıtacak düzeyde kararlı olması, bir başka ifadeyle küçük bir örneklemler büyüklüğünden itibaren parametre deđişmezliđinin sağlanması TBPOMTK'nin TBPMTK'ye göre önemli avantajlara sahip olduğunu bir kanıttır. Bu nedenle özellikle okul uygulamalarında TBPOMTK uygulamaları TBPMTK'ye tercih edilebilir.

İkinci alt problemin bulgularından elde edilen sonuçlar doğrultusunda birleşik testin ÇBMTK'ye göre analiz edilmesiyle elde edilen  $a_1$  ve  $a_2$  parametreleri ortalamasının, örneklemler büyüklüğüne bađlı olarak düzenli bir artma ya da azalma eğilimi göstermediđi sonucuna ulaşılmıştır. Ackermann (2005) ise örneklemler büyüklüğü arttıkça maddelerin ayırtedicilik gücünün arttığı sonucuna ulaşılmıştır. Bu açıdan çalışmadan elde edilen sonuç, Ackermann (2005)'in elde ettiđi sonuç ile farklılık göstermektedir.

ÇBMTK analizleri sonuçlarına göre d parametresi ortalamasının örneklemler büyüklüğünün deđişimine bađlı olarak net bir artma ya da azalma eğiliminin olmadığı görülmektedir. d parametresi ortalaması ile d parametresi ortalamasına ait SHO da örneklemlerden evrene doğru sürekli bir deđişim gösterdiđi için parametre deđişmezliđinin sağlanamadığı sonucuna ulaşılmıştır.

Birleşik test ÇBMTK'ye göre analiz edildiđinde her bir boyut için elde edilen a parametrelerinin ortalamasının ikisi de (0,1) aralığında değerler almıştır. TBPMTK'ye göre kestirilen ortalama ayırtedicilik parametreleri ise 1'in üzerinde değerler almıştır. Elde edilen bu sonuç Ansley ve Forsyth (1985)'in, iki kuramdan birbirine yakın büyüklükte a parametresi kestirimini elde ettiđi çalışmasında ortaya koyduđu sonuçlardan farklılık göstermektedir. Bununla beraber çok boyutlu testin TBPMTK'ye göre analiz edilmesinden elde edilen sonuçlara göre a ve b parametresi ortalamaları ve a ve b ortalamalarına ait SHO'lar bütün örneklemlerde farklı değerler aldığı için parametre deđişmezliđinin sağlanamadığı sonucuna ulaşılmıştır.

Sonuç olarak çok boyutlu bir veri ister TBPMTK altında modellen sin ister ÇBMTK altında modellen sin evren değerine yakın bir büyüklükte SHO ile ortalama bir parametre kestirimi yapmak için en az 5000 örneklemler büyüklüğünün kullanılması gerekir. 5000 ya da daha büyük örneklemler kullanılması durumunda kestirilen SHO evren değerine yaklaşmış olur. Ancak her ne kadar parametre ortalamalarına ait SHO evren değerine yaklaşırsa da çok boyutlu bir verinin tek boyutlu modele göre analiz edilmesi ile elde edilen parametre ortalamaları ile çok boyutlu modele göre analiz edildiđinde elde edilen parametre ortalamaları örneklemler büyüklüğüne bađlı olmaksızın birbirinden farklı çıkmıştır. Bir başka ifade ile parametre deđişmezliđi sağlanmamıştır. Elde edilen bu sonuç tek boyutluluk varsayımı sağlanmadığı zaman parametre kestirimlerine ait sonuçların farklılaştığını ifade eden Drasgow ve Parsons (1983)'ün bulgularıyla tutarlılık göstermektedir. Burada deđişmezliđin sağlanamamasının ve parametre ortalamalarının örneklemler büyüklüğüne bađlı olarak net bir artma ya da azalma göstermemesinin nedeni olarak bu çalışmada örneklemler seçiminde replikasyon yapılmamış olması düşünülmektedir.

Üçüncü alt problemin bulgularından elde edilen sonuçlar doğrultusunda birleşik testin TBPOMTK'ye göre analiz edilmesiyle elde edilen H ortalamasına ait SHO'ların örneklemler büyüklüğü arttıkça istikrarlı bir şekilde azalarak evren değerine yaklaştığı sonucuna ulaşılmıştır. Kođar (2014) çalışmasında örneklemler büyüklüğü arttıkça H katsayılarına ait standart hatanın azaldığını ortaya koymuştur. Elde edilen bu sonuç Kođar (2014)'ün bulgularıyla tutarlılık göstermektedir. Bunun yanı

sıra örneklem büyüklüğü kaç olursa olsun SHO'lar büyüklük olarak evrene çok yakın değerler almıştır. Bu durumda tek boyutluluğun ihlal edilmesinin Hi ve p ortalamalarının kestirildiği SHO değerlerinin üzerinde önemli bir etkisinin olmadığı sonucuna ulaşılmıştır.

Tek boyutluluk varsayımının ihlal edilmesinin parametre kestirimleri üzerindeki en önemli etkisi, tek boyutluluk varsayımı karşılandığı durumda parametre değişmezliği sağlanırken, varsayım karşılanmadığında parametre değişmezliğinin de sağlanamaması olmuştur. Bu nedenle bu çalışmada Koğar (2014)'ün örneklem büyüklüğü arttıkça p değerleri için değişmezliğin sağlandığı bulgusuna ulaşamamıştır. Koğar (2014)'ün elde ettiği bu sonuç, 1. alt problemde elde edilen bulgularda ifade edildiği gibi tek boyutlu veriye ait p ortalamaları için elde edilmiştir. Bir başka ifadeyle tek boyutlu veriye ait p ortalamaları için örneklem büyüklüğü arttıkça parametre değişmezliği sağlanmıştır.

TBPMTK ile TBPOMTK'den elde edilen sonuçlar karşılaştırıldığında TBPOMTK'den kestirilen parametre ortalamalarının evren değerine yakın bir SHO ile kestirilebildiği sonucuna ulaşılmıştır. TBPMTK'den elde edilen parametre ortalamalarının evren değerine yakın bir SHO ile kestirilebilmesi için ise en az 1000 örneklem büyüklüğü ile çalışmak gerekmektedir. Daha küçük bir örneklemde evren değerine yakın bir SHO ile madde parametrelerinin kestirilebilmesi TBPOMTK'nin TBPMTK'ye göre avantajlı olduğunun bir kanıtıdır. Bununla birlikte tek boyutluluk varsayımının karşılanıp karşılanmamasının bu sonucu değiştirmedeği dikkat çeken bir bulgudur.

### Öneriler

Bu çalışmada örneklem büyüklüğünün etkisini incelemek için 50, 100, 200, 500, 1000 ve 5000 örneklem büyüklükleri ile çalışılmıştır. Benzer bir çalışma farklı örneklem büyüklükleri ile de gerçekleştirilebilir. Bu çalışmada TBPMTK için 2PLM, TBPOMTK için MHM ve ÇBMTK için M2PLM kullanılmıştır. Sonuçları incelenen kuramlara ait olan farklı modeller ile benzer bir çalışma yapılabilir.

Bu çalışmada kullanılan modeller üzerinden kestirilen parametre değerleri tüm teste ait ortalama parametre değerleridir. Benzer bir çalışma bir testte yer alan maddelere ait parametre kestirimleri için de gerçekleştirilebilir. Bu çalışmada örneklem büyüklüğünün etkisini araştırmak amacıyla parametre kestirimlerine ait standart hata ortalaması kullanılmıştır. Başka bir çalışmada örneklem büyüklüğünün etkisini araştırmak için farklı göstergelerden de yararlanılabilir.

Bu çalışmada şans başarısının etkisi göz önünde bulundurulmamıştır. Şans başarısından arındırılmış gerçek puanlar üzerinden de başka bir çalışma yapılabilir. Bu çalışmada örneklem büyüklüğünün parametre değişmezliğinin üzerindeki etkisi araştırılmıştır. Farklı bağımsız değişkenler ile de benzer bir çalışma yapılabilir. Bu çalışmada örneklem seçiminde her bir örneklem büyüklüğü yalnızca bir kez seçilmiştir. Başka bir çalışmada replikasyon yoluyla örneklem seçerek örneklem büyüklüğünün parametre değişmezliği üzerindeki etkisi araştırılabilir.

### KAYNAKÇA

- Ackerman, T. A. (2005). Multidimensional item response theory modeling. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Multivariate applications book series. Contemporary psychometrics: A festschrift for Roderick P. McDonald* (p. 3–25). Lawrence Erlbaum Associates Publishers.
- Andries van der Ark, L. (2007). Mokken Scale Analysis in R. *Journal of Statistical Software*, 20(11), 1-19.
- Ansley, T.N. and Forsyth, R.A. (1985). An Examination of the characteristics of Unidimensional IRT estimates derived from two dimensional data. *Applied Psychological Measurement*, 9(1), 37-48.
- Şengül Avşar A. (2018). Kategori sayısının psikometrik özellikler üzerine etkisinin mokken homojenlik modeli'ne göre incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 9(1), 49-63.
- Şengül Avşar A. , Tavşancıl E. (2017). Examination of polytomous items' psychometric properties according to nonparametric item response theory models in different test conditions, *Kuram ve Uygulamada Eğitim Bilimleri*, 17, 493-514.
- Cai L. (2017). *flexMIRT Version 3.51: Flexible multilevel multidimensional item analysis and test scoring* (Computer software). Chapel Hill, NC: Vector Psychometric Group.
- Cai, L. (2008). SEM of Another Flavour: Two new applications of the supplemented em algorithm. *British*

- Journal of Mathematical and Statistical Psychology*, 61, 309–329.
- Çokluk, Ö., Şekerciođlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok deđişkenli istatistik teknikleri*. Ankara: Pegem Akademi.
- Drasgow, F. & Parsons, K. (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7(2), 189-199.
- Embretson, S. E., & Reise, S. P. (2000). *Multivariate applications books series. Item response theory for psychologists*. Lawrence Erlbaum Associates Publishers.
- Fraenkel, J.R. & Wallen, N.E. (2006). *How to design and evaluate research in education* (Sixth edition). Boston: McGraw-Hill Pub.
- Goldman, S.H., & Raju, N. S. (1986). Recovery of one- and two-parameter logistic item parameters: An empirical study. *Educational And Psychological Measurement*, 46(1), 11-21.
- Hambleton, R. K. & Swaminathan, H. (1985). *Item Response Theory: Principles and applications*. Boston: Academic Publishers Group.
- Hulin, C. L., Lissak, R. I., & Drasgow, F. (1982). Recovery of two and three-parameter logistic item characteristic curves: A Monte Carlo Study. *Applied Psychological Measurement*, 6(3), 249-260.
- Harman, H. H. (1962). *Modern Factor Analysis* (2. edition). University of Chicago Press, Chicago.
- Kaptan, S. (1977). *Bilimsel Araştırma Teknikleri*, Ankara: Tekişik Matbaası ve Rehber Yayınevi.
- Kođar H. (2018). Examining invariant item ordering using mokken scale analysis for polytomously scored items. *Eđitimde ve Psikolojide Ölçme ve Deđerlendirme Dergisi*, 9(4), 312-325.
- Kođar, H. (2014). *Madde tepki kuramının farklı uygulamalarından elde edilen parametrelerin ve model uyumlarının örnekleme büyüklüđü ve test uzunluđu açısından karşılaştırılması*. (Doktora tezi, Hacettepe Üniversitesi, Eđitimde Ölçme ve Deđerlendirme Anabilim Dalı, Ankara). <http://tez2.yok.gov.tr/adresinden edinilmiştir>.
- Meara, K., Robin, F. & Sireci, S.G. (2000). Using multidimensional scaling to assess the dimensionality of dichotomous item data. *Multivariate Behavioral Research*, 35(2), 229–259.
- Mor-Dirlik, E. (2017). *Parametrik ve parametrik olmayan madde tepki kuramı modellerinden çeşitli faktörlere göre elde edilen madde ve yetenek kestirimlerinin karşılaştırılması*. (Doktora tezi, Ankara Üniversitesi, Eđitimde Ölçme ve Deđerlendirme Anabilim Dalı, Ankara). <http://tez2.yok.gov.tr/adresinden edinilmiştir>.
- Price, L. R. (2017). *Psychometric methods: Theory and practice*. New York, NY: The Guilford Press.
- Sijtsma, K., & Molenaar, I. W. (2002). *Introduction to Nonparametric Item Response Theory*. Thousand Oaks, CA: Sage.
- Smits, I. A. M., Timmerman, M. E., & Meijer, R. R. (2012). Exploratory mokken scale analysis as a dimensionality assessment tool: why scalability does not imply unidimensionality. *Applied Psychological Measurement*, 36(6), 516-539.
- Sodano, S. M., and Tracey, T. J. G. (2011). A brief inventory of interpersonal problems- circumplex using non-parametric item response theory: introducing the IIP-C- IRT. *Journal of Personality Assessment*, 93(1), 62-75. doi:10.1080/00223891.2010.528482
- Syu, J. J. (2013). *Applying person fit-in faking detection-the simulation and practice of non parametric item response theory*. (Doctoral Dissertation, National Chengchi University). Retrieved from <http://nccur.lib.nccu.edu.tw/bitstream/140.119/58646/1/251501.pdf>
- Sünbül, Ö. (2011). *Çeşitli boyutluluk özelliklerine sahip yapılarda, madde parametrelerinin deđişmezliđinin klasik test teorisi, tek boyutlu madde tepki kuramı ve çok boyutlu madde tepki kuramı çerçevesinde incelenmesi*. (Doktora tezi, Mersin Üniversitesi, Eđitimde Ölçme ve Deđerlendirme Anabilim Dalı, Ankara). <http://tez2.yok.gov.tr/adresinden edinilmiştir>.
- Tabachnick, G. B. ve Fidell, S. L. (2001). *Using multivariate statistics* (4<sup>th</sup> Edition), Boston MA: Allyn&Bacon.
- Thissen, D., & Wainer, H. (1982). Some standart errors in item response theory. *Psychometrika*, 47(4), 397-412.