

# Kısa Form Geliştirme: Yöntemler, Öneriler ve İncelemeler

Hakan KOĞAR\*

## Öz

Bu derleme çalışmanın amacı, bir ölçme aracının kısa formu geliştirilirken kullanılabilecek yöntemleri açıklamak ve kısa form geliştirme çalışmalarına getirilen eleştirileri de dikkate alarak özellikle sağlık bilimleri alan yazınında yer alan bazı kısa form geliştirme çalışmalarını yöntemsel olarak incelemektir. Kısa form geliştirme çalışmalarının özellikle sağlık bilimlerine ait alanlarda yoğunlaştığı görülmektedir. Bu durumun temel nedeni olarak klinikte çalışan uzmanların üzerlerindeki baskıyı azaltacak hızlı ve güvenilir ölçme araçlarına ihtiyaç duymaları gösterilmiştir. Bu araştırma için seçilen 12 makalenin inceleme sonuçları göstermektedir ki kısa form geliştirme ile ilgili kurallara uyan çok az araştırma vardır. Bu durum, 20. yüzyılın ortalarından itibaren tartışılan kısa form geliştirme çalışmalarının halen tartışmaya açık olduğunu göstermektedir. Araştırmacılara ölçme aracının uzun formundan hareket ederek, bu çalışmada bahsedilen ölçütleri de dikkate alarak, ölçeğin kısa formunu geliştirmesi önerilmektedir. Etik kurallara uygun bir şekilde geliştirilen ve psikometrik özelliklere yeterli düzeyde sahip olan ölçme araçlarının seçilmesi önerilmektedir. Klinik araştırmacılar bilmelidir ki daha az madde içeren ölçme araçlarının daha az geçerli olduğu şeklindeki bir algı doğruyu göstermemektedir. Her ölçme aracı için aynı psikometrik standartlar aranmaktadır.

**Anahtar Kelimeler:** Kısa form, ölçek geliştirme, klasik test kuramı, madde tepki kuramı

## GİRİŞ

Bir ölçme aracının kısa formunu geliştirme çabaları, 20. yüzyılın başlarında Doll'un (1917) zekâ ölçmek için Binet-Simon zekâ testindeki bütün maddeleri kullanmanın gerekli olup olmadığını sorgulaması ile başlamıştır. 1950'li yıllarda hız kazanan kısa form geliştirme çalışmaları, zekâ ve yetenek testlerinde yer alan çok sayıda maddeye getirilen eleştiriler neticesinde (Levy, 1968) öncelikle klinik değerlendirmeler için kullanılan ölçme araçlarına yoğunlaşmıştır. O dönemdeki kısa form geliştirme çalışmalarını inceleyen Levy (1968) yaptığı çalışmada kısa form üretme çalışmalarının gerçek amacından uzaklaştığını ve tekdüze bir akademik aktivite haline geldiğini ifade ederek bu çalışmalara ilişkin eleştirisini belirtmiştir.

## Neden Kısa Form?

20. yüzyılın ortalarında kısa form geliştirme çalışmalarının temel amacı zamanı etkili kullanabilmektir (Levy, 1968). Zamanın ve emeğin ekonomik kullanımı ile de testlerin doğru tahminlerde bulunması arasında bir uzlaşma bulunması hedeflenmekteydi (Doppelt, 1956). Günümüzde kısa form üretme çabalarının arkasında ise çeşitli amaçların yattığı görülmektedir. Kısa form ile çok değişkenli ve çok kültürlü çalışmalarda kullanım kolaylığı sağlama, daha az davranış ölçerek zaman kazanma, çocuklar üzerinde daha kullanılabilir bir form oluşturma, seçme ve yerleştirme hedeflerine daha hızlı ulaşabilme, uzun form ile aynı geçerliğe sahip bir kısa form geliştirebilme gibi amaçlar bunlardan birkaçıdır. Kısa form geliştirme çalışmalarının özellikle sağlık bilimlerine ait alanlarda yoğunlaştığı görülmektedir. Bu durumun temel nedeni olarak klinikte çalışan uzmanların üzerlerindeki baskıyı azaltacak hızlı ve güvenilir ölçme araçlarına ihtiyaç duymaları gösterilmiştir (Smith, McCarthy ve Anderson, 2000).

\* Doç. Dr., Akdeniz Üniversitesi, Eğitim Fakültesi, Antalya-Türkiye, e-mail: hkogar@gmail.com, ORCID ID: 0000-0001-5749-9824

Bu makaleye atıfta bulunmak için:

Koğar, H. (2020). Kısa Form Geliştirme: Yöntemler, Öneriler ve İncelemeler. *Journal of Measurement and Evaluation in Education and Psychology*, 11(3), 302-310. doi: 10.21031/epod.739548.

Geliş Tarihi: 18.05.2020  
Kabul Tarihi: 08.09.2020

### ***Kısa Form Geliştirmede Kullanılan Psikometri Kuramları***

20. yüzyılın ortalarında kısa formların geliştirilmesinde, ölçme aracının uzun formu ile en yüksek korelasyonu veren madde setinin seçilmesi, yalnızca madde istatistiklerine göre bir madde örnekleme oluşturma ve geçerliği en yüksek faktör veya faktörleri seçme gibi çeşitli yolların izlenmiştir (Levy, 1968). Bu yollar içerisinde klasik madde istatistiklerine göre madde örnekleme belirleme yolunun sıklıkla kullanıldığı, bu istatistiklere ek olarak Guttman'ın skalogram analizleri gibi diğer bazı istatistiklere de başvurulduğu belirlenmiştir. Günümüzde kullanımı sınırlı olan bu yöntemlerle birlikte 1970 ve sonrasındaki teknolojik gelişmelerin de etkisi ile birlikte kullanılmaya başlanmış diğer yöntemler, psikometri kuramları bağlamında aşağıda ayrıntılı olarak açıklanmıştır.

### ***Klasik Test Kuramı (KTK)***

#### *Klasik madde istatistikleri*

Sosyal bilimlerde ölçek geliştirme çalışmalarına olan yoğun ilginin başladığı dönemlere kadar uzanan bu klasikleşmiş istatistiklerden en önemlileri madde güçlük indeksi, madde ayırt edicilik indeksi ve madde toplam korelasyon katsayısıdır. Madde güçlük indeksi, maddenin gruptaki bireylerin yetenek düzeylerine göre kolay veya zor olduğuna işaret eder. Henning'e (1987) göre bir maddenin çok kolay ya da çok zor olması puan dağılımının çarpık olmasına, dolayısıyla grubun yetenek düzeyine uygun bir madde hazırlanmadığına işaret edebilir. Toplam puana göre yüksek puana sahip grup ile düşük puana sahip grubu ayırt etmeyi hedefleyen madde ayırt edicilik indeksi, maddenin ölçekteki yerine karar veren önemli bir indeks değeridir. Madde toplam korelasyonu ise maddenin ölçtüğü özellik veya kapsam ile teste ait toplam puanın ölçtüğü özellik veya kapsamın ilişkisini verir. Her bir madde puanı toplam puanla ilişkili olmalıdır. Toplam puanla yüksek ilişki gösteren maddeler, faktör analizindeki faktör yükünde olduğu gibi, toplam puana ait varyansı yüksek düzeyde açıklayan maddelerdir. Diğer bir deyişle bu maddeler, yüksek geçerliğe sahiptir. Bu istatistiksel teknikler, hesaplamasının kolay olması nedeniyle 21. yüzyılda da sıklıkla kullanılmaktadır. Ancak özellikle madde toplam korelasyonunun Pearson momentler çarpımı korelasyon katsayısını temel alması nedeniyle yanıltıcı bulgulara neden olduğu bilinmektedir (Raykov ve Marcoulides, 2011).

#### *Biggers' (1976) Spearman-Brown tahmin yöntemi*

Klasik madde istatistiklerinin kullanımına eleştiri getiren Biggers (1976), uzun formun n tane paralel kısa formun bir bütünü, geliştirilen kısa formun ise bu paralel formlardan yalnızca biri olduğunu belirterek hangi kısa formun daha doğru bir seçim olduğunu belirlemenin mümkün olmadığını söylemektedir. Ayrıca madde atarak ya da seçerek kısa form geliştirmenin tersine çevrilemez deneysel bir yöntem olduğunu, yani önce kısa form geliştirip sonra madde eklenerek testin uzun formu elde edilmeye çalışmanın mümkün olamayacağını belirtmiştir. Bu amaçla Spearman-Brown tahmin yöntemini kısa form geliştirmenin bir alternatifi olarak sunmuştur. 40 maddeden oluşan dogmatizm ölçeğinin öncelikli olarak klasik madde analizleri yardımıyla kısa formunu geliştirmiştir. Sonrasında ölçeği tek-çift sayı olacak şekilde ikiye ayırmış, klasik madde analizleri ile elde edilen kısa formun toplam puanı ile ölçeğin uzun formunun toplam puanı arasındaki korelasyon katsayısını hesaplamıştır. Tek sayı ile ayrılmış ölçek yarısından elde edilen puanlar ile ölçeğin uzun formundan elde edilen puanlar arasındaki korelasyon .92, çift sayı ile ayrılmış yarıya ait puanlar ile ölçeğin uzun formundan elde edilen puanlar arasındaki korelasyon .93 olarak belirlenmiştir.

#### *Faktör analizi*

Faktör analizi, gözlenen değişkenler arasındaki ilişkilerden yola çıkarak bir modelin kurulduğu ve bu model aracılığıyla maddelerin bir veya birden fazla gizil değişkenle ilişkilendirildiği birçok değişkenli istatistiktir. Ölçek kısa formu geliştirme çalışmaları ile birlikte ölçek geliştirme ve uyarlama

çalışmalarında en sık kullanılan istatistiksel tekniktir. Ancak halen doğru uygulama örneklerine sıklıkla rastlandığını söylemek oldukça güçtür. Goretzko, Pahm ve Buhner'e (2019) göre faktör analizi kullanan çalışmalarda özellikle örneklem büyüklüğü belirleme ile doğru döndürme ve faktör açığa çıkarma tekniği seçme konularında sorunlar yaşanmaktadır. Fabrigar, Wegener, MacCallum ve Strahan (1999) inceledikleri araştırmalardan yola çıkarak faktör analizi kullanılacak çalışmalar ile ilgili bazı önerilerde bulunmuşlardır. Araştırmacılara göre bir faktörde yer alması gereken madde sayısı en az dört ve örneklem büyüklüğü ise en az 400 olmalıdır. Çok değişkenli normalliğin sağlandığı durumlarda en çok olabilirlik, diğer durumlarda ise veri dönüşümü ya da principal axis factoring gibi diğer teknikler kullanılmalıdır. Smith, McCarthy ve Anderson (2000), ölçek kısa formu geliştirme çalışmalarında faktör analizinin sıklıkla kullanıldığını belirtmiş, ölçeğin uzun formundan elde edilen veri setine uygulanan faktör analizi ile kısa formun geliştirilmesini ise eleştirmişlerdir. Bu şekildeki bir yaklaşım, ölçeğin uzun formu ile kısa formunun aynı yapıya sahip olduğunu varsaymak anlamına gelecektir. Ancak ölçeğin uzun formu ile kısa formunun aynı faktör yapısına sahip olduğu kesin değildir. Bu soruna çözüm olarak kısa formdaki maddeler üzerinden ayrı bir faktör analizi yapılmasını önermişlerdir. Bu bulgular ile uzun formdan elde edilen bulguların benzer olması, bu iki ölçek formunun birbirine alternatif olabileceğine işaret edecektir. Faktör yapılarındaki olası büyük farklılıklar ölçeğin kısa formu ile uzun formunun farklı özellikleri ölçtüğü anlamına gelebilmektedir.

### ***Madde Tepki Kuramı (MTK)***

Madde tepki kuramı (MTK), KTK'nın çeşitli sınırlılıkları ve özellikle ölçeklerin psikometrik özelliklerinin belirlenmesindeki eksik yaklaşımlarına karşı geliştirilmiştir. Parametrik (Birnbaum, 1968; Rasch, 1960) ve parametrik olmayan (Mokken ve Lewis, 1982) olmak üzere iki farklı yaklaşımı içerir. Araştırmanın amacına ve varsayımların karşılanma düzeyine göre araştırmacılar bu iki temel yaklaşımdan birini seçmelidir. Gizil özellik ile madde cevapları arasında simetrik bir ilişkinin bulunması, tekboyutluluk ve yüksek örneklem büyüklüğü sağlandığında parametrik MTK modelleri kullanılabilir. Asimetrik bir dağılım ve düşük örneklem büyüklüğünün bulunduğu durumlarda parametrik olmayan MTK modelleri kullanılabilir. Parametrik ve parametrik olmayan MTK modellerinin tekboyutluluğun ihlal edildiği durumlara karşı dirençli olduğu da bilinmektedir (Embretson ve Reise, 2000; Sodano ve Tracey, 2011).

### ***Parametrik Madde Tepki Kuramı***

Faktör analizi gibi KTK'ye dayalı teknikler yalnızca birbirinden bağımsız maddeler arasındaki ilişkiler üzerinden bilgi elde edebilir. Ayrıca elde edilen bütün istatistiksel bulgular örneklem bağımlıdır. MTK'nın en büyük avantajı madde parametrelerinin değişmezliği iddiasında bulunarak örnekleme olan bağılılığa son vermesidir. MTK'de standart hatalar gizil özelliğe ait her bir düzey için tek tek hesaplanır. Bu sayede grubun tek bir hata değerine sabitlenmesinin önüne geçilir. Özellikle klinik ölçmelerde verilen kararlar açısından bu konu önemlidir. MTK yeteneği düşük ve yüksek olan grupları ayırabilen maddeler üzerinden bilgi elde eder. Ayrıca MTK'nın her bir yetenek düzeyinde ve her bir boyut için madde karakteristik eğrileri (MKE) vermesi sayesinde ölçeğin kısa formunu elde etmek için gerekli bilgi miktarı değerlendirilebilmektedir. Yüksek bilgi veren maddelerle ilgili yetenek düzeyi daha kesin olarak belirlenebilirken düşük bilgi veren maddeler ile yetenek düzeyini kestirmek daha az kesinlik ile mümkündür. Ölçülen özelliğin ranjına uygun bir şekilde en fazla bilgi içeren maddeler seçilebilmektedir. Daha iyi performans gösteren ve özelliğin farklı seviyeleri boyunca yeterli bilgiyi sağlayan maddeleri seçerek, yüksek psikometrik özelliklere sahip bir kısa form geliştirmek mümkün olmaktadır. MTK'de ayrıca, test bilgi fonksiyonları (TBF) sayesinde, Cronbach alfa gibi KTK'ye dayalı güvenilirlik belirleme teknikleri gibi tek bir katsayı vermek yerine, yapının ölçüldüğü her bir düzeyin kesinliği değerlendirilebilmektedir. TBF sayesinde yüksek bilgi elde edilen dolayısıyla düşük hata içeren yetenek düzeyleri belirlenebilmekte ve bu sayede yüksek yerel güvenilirlik ortaya koyulabilmektedir (Embretson ve Reise, 2000; Hambleton, Swaminathan, ve Rogers, 1991). TBF,

MKE'ler aracılığıyla geliştirilmektedir. Bu nedenle kısa form geliştirme çalışmalarında yüksek bilgi içeren maddeler seçilerek uzun formun sahip olduğu bilgi miktarına ulaşmak hedeflenmelidir.

### *Parametrik Olmayan Madde Tepki Kuramı ve Mokken Ölçekleme Analizleri*

Parametrik olmayan MTK, çok az sayıda varsayıma sahip olması nedeniyle 21. yüzyılın başlarından itibaren kullanımı yaygınlaşan bir MTK yaklaşımıdır. Uygulayıcılar açısından yorumlaması kolaydır. Özellikle keşfedici amaçlarla kullanımı yaygındır. Parametrik olmayan MTK uygulamalarında, parametrik MTK uygulamalarında olduğu gibi MKE'ler elde edilmektedir. Monoton olarak azalan, monoton olarak azalmayan, simetrik ve asimetrik olan her türlü dağılımda MKE elde edilebilmektedir (Meijer ve Baneke, 2004). Parametrik olmayan MTK modelleri, Mokken ölçekleme analizleri ve parametrik olmayan regresyon tahmin modelleri olmak üzere ikiye ayrılmaktadır. Mokken ölçekleme analizleri, Guttman ölçeklemesinin olasılığa dayalı olarak genişletilmiş halidir. Monoton Homojenlik Modeli (MHM) ve İkili Monotonluk Modeli (İMM) olmak üzere iki farklı yaklaşıma sahiptir. MHM, gizil özellik ile monoton bir ilişki gösteren madde cevap fonksiyonuna sahip ve homojen (tekboyutlu) olan madde gruplarına ait madde ve kişiler arasındaki ilişkileri tanımlamaktadır. İMM'nin basitleştirilmiş ve daha az varsayıma sahip halidir. Bu modellerdeki temel amaç bireyleri ve maddeleri sıralamaktır (Koğar, 2015). Parametrik olmayan ve parametrik MTK aynı madde seçim algoritmasını takip eder (Lei, Dunbar, ve Kolen, 2004).

### *Karınca Kolonisi Optimizasyonu*

Bir psikometri kuramı olmamakla birlikte, kısa form geliştirme amacıyla geliştirilen en güncel ve etkili tekniklerden biri olan Karınca Kolonisi Optimizasyonu (Ant Colony Optimization - AKO) karıncaların yiyecek arama algoritmasını temel almaktadır (Dorigo ve Stützle, 2004). Yiyecek kaynağı ile koloni arasındaki en kısa yolun hesaplandığı bu algoritmanın kısa form geliştirme çalışmalarında kullanılabileceği düşünülmüştür. Yapısal Eşitlik Modeli (YEM) kullanılarak modellenmektedir. AKO algoritması doğru modele yakınsayarak en yüksek model uyumuna sahip modeli açığa çıkarmayı hedefler. Bunu tekrarlayan bir süreçle ortaya koyarak en iyi kısa formu elde etmeye çalışır.

### *Araştırmanın Amacı ve Önemi*

Bu çalışmada bir ölçme aracının kısa formu geliştirilirken sıklıkla kullanılan yöntemlerden bazılarının bahsedilmiştir. Kısa form geliştirme çalışmalarının sayısı oldukça fazla ve tarihi uzun yıllara dayalı olsa da, bu konu hakkındaki tartışmalar halen varlığını sürdürmektedir. Kısa form geliştirme çalışmalarına getirilen eleştiriler iki temel başlık altında incelenebilir. Birincisi, ölçme aracının geçerliğini diğer herhangi bir özelliğinden üstün tutmasıdır. Psikometri kuramlarına göre bir ölçme aracının geçerli olması bir zorunluluktur. Madde sayısını azaltma ya da zamanı daha etkili kullanma gibi daha çok kullanışlılığı ilgilendiren etmenler ikincil öneme sahiptir. Dolayısıyla bir ölçeğin kısa formunu geliştirirken öncelikli hedef, en az ölçeğin uzun formu kadar geçerli bir araç elde edebilmek olmalıdır. Ancak bu amaçtan sapan araştırmaların alanyazında yer aldığı bilinmektedir. İkincisi ise ölçme aracının kısa formu geliştirilirken yöntemsel hataların sıklıkla yapılması, dikkatsiz ve özensiz davranılması ve uzun form ile karşılaştırmasının yapılmamasıdır. Bu duruma neden olarak kısa form geliştirme ile ilgili yöntemsel bilginin alanyazında sınırlı kalması gösterilebilir (Smith, McCarthy ve Anderson, 2000). Bu araştırmanın amacı, bir ölçme aracının kısa formu geliştirilirken kullanılabilecek yöntemleri açıklamak ve kısa form geliştirme çalışmalarına getirilen eleştirileri de dikkate alarak eğitim ve sağlık bilimleri alanyazında yer alan bazı kısa form geliştirme çalışmalarını yöntemsel olarak incelemektir.

## ***Kısa Form Geliştirme Çalışmalarında Dikkat Edilmesi Gerekenler ve Bazı Çalışmaların İncelenmesi***

Çalışmanın bu bölümünde Levy (1968), Smith, McCarthy ve Anderson (2000), Hagtvet ve Sipos'un (2016) çalışmalarından yola çıkarak kısa form geliştirilirken dikkat edilmesi gereken konular maddeler halinde açıklanmıştır. Araştırma için 2011-2019 yılları arasında ERIC ve PUBMED indekslerinde taranan dergilerde yayınlanmış 12 kısa form geliştirme çalışması seçilmiştir. Bu çalışmaların tamamında yeni bir kısa form geliştirmek hedeflenmiş, aşağıda belirtilen kurallara uygun bir şekilde geliştirilip geliştirilmediği incelenmiştir. Bu çalışmalara ait künyeye Tablo 1'de yer verilmiştir.

**Tablo 1. İncelenen Çalışmalara ait Künye**

| Kaynak                                                                       | Ölçeğin Kısa Formu                                    |
|------------------------------------------------------------------------------|-------------------------------------------------------|
| Baiocco, Pallini ve Santamaria (2014)                                        | Adolescent Friendship Attachment Scale                |
| Woudstra, Meppelink, Maat, Oosterhaven, Fransen ve Dima (2019)               | Short Assessment of Health Literacy                   |
| Lim ve Chapman (2013)                                                        | Attitudes Toward Mathematics Inventory                |
| Jenkinson, Kelly, Dummett ve Morley (2019)                                   | The Oxford Participation and Activities Questionnaire |
| Rogers, M. E., Creed, P. A., Searle, J., ve Hartung, P. J. (2011)            | Physician Values in Practice Scale                    |
| Ferrario, Panzeri, Anselmi ve Vidotto (2019)                                 | Illness Denial Questionnaire                          |
| Morin, Valois, Crocker ve Robitaille (2019)                                  | Intellectual Disability Questionnaire                 |
| Nimon ve Zigarmi (2015)                                                      | Work Intention Inventory                              |
| Milavic, Padulo, Grgantov, Milic, Mannarini, Manzoni, Ardigo ve Rossi (2019) | The Psychology Skills Inventory For Sports            |
| Park ve Hill (2017)                                                          | Occupational Work Ethic Inventory                     |
| Siefert, Sexton, Meehan, Nelson, Haggerty, Dauphin ve Huprich (2019)         | DSM-5 Levels of Personality Functioning Questionnaire |
| Bohlmeijer, Klooster, Fledderus, Veehof ve Baer (2011)                       | Five Facet Mindfulness Questionnaire                  |

### ***1. Öncelikli olarak ölçme aracının uzun formu yeterli geçerlik ve güvenirliğe sahip olmalıdır:***

Bir ölçeğin kısa formu geliştirilmek istendiğinde ilk yapılması gereken iş, bu ölçek formunun geçerlik ve güvenirlik değerlerini yargılamaktır. Geçerli ve güvenilir olmayan bir ölçme aracının, olası bir kısa formuna ait geçerlik ve güvenirlik değerleri de şüpheli olacaktır. İncelenen 12 araştırmadan ikisi ayrıntılı bir şekilde uzun formun psikometrik özelliklerine yer vermiştir. Diğer çalışmalar ya sadece güvenirlik katsayılarına yer vermekle yetinmiş ya da uzun formun geçerli ve güvenilir bir ölçme araç olduğunu belirtmekle yetinmiştir.

### ***2. Bir ölçme aracının kısa formu, uzun formu ile aynı psikometrik özelliklere sahip olmadığı sürece, o ölçeğin tek kısa formu değil alternatif kısa formlarından biridir:***

Ölçeğin uzun formunda yer alan maddelerden yapıyı en iyi açıklayan madde seti ya da yapıyı en iyi açıklayan maddeler arasından rastgele madde seçimleri yaparak kısa forma ait madde seti oluşturulmalıdır. Bu aşamanın ardından kısa formun “eşit” ya da “değiştirilebilir” form olduğu kararına geçilmelidir. “Eşit” kısa form, uzun formla aynı psikometrik özelliklere sahiptir, dolayısıyla uzun forma alternatif olarak kullanılabilir. “Değiştirilebilir” kısa form ise uzun form kadar yüksek psikometrik özelliklere sahip değildir. Dolayısıyla benzer bir yöntem ile tekrarlanabilecek başka bir çalışmada benzer formların elde edilebilmesi de olasıdır. “Değiştirilebilir” kısa formlar genellikle uzun formdan daha düşük geçerliğe sahiptir. Bu durumda bu forma alternatif olabilecek farklı formları, farklı faktör yapısını, farklı maddeleri ya da madde setlerini de araştırmacı açığa çıkarmalı ve tartışmalıdır. Aksi takdirde bu form, uzun formun bir alternatifi olamaz. Kısa formda daha az madde olması daha az geçerliğin yeterli olacağı anlamına gelmemektedir. Bu konu göz ardı edilemeyecek kadar önemlidir. İncelenen çalışmalardan ikisinin “eşit” form özelliği taşıdığı söylenebilmektedir. Bu

araştırmalarda geliştirilen kısa formlar en az aracın uzun formu kadar geçerli ve güvenilirlerdir. Ancak diğer on araştırmada uzun formun geçerlik ve güvenilirlik bilgilerine yeterince yer verilmediği için bu araştırmalar hakkında bir yoruma yer verilememiştir.

*3. Ölçme aracının ölçtüğü özelliğin doğasını yansıtacak şekilde bir davranış evreninden (uzun forma ait tüm maddeler), davranış örnekleme (kısa formda yer alacak maddeler) geçiş yapılmalıdır:*

Uzun formdaki madde havuzundan kısa forma madde seçerken dikkat edilmesi gereken unsurlardan biri de seçilen maddelerin oluşturacağı davranış örnekleminin, uzun forma ait davranış evreninin yansıtılabildiğini, yani genelleştirilebileceğini göstermektir. Bu konu psikometrik özellikler kadar önemli olup, kapsam geçerliği ile ilişkilidir. İyi açıklanmış ve tanımlanmış bir içerik, yapı geçerliği için de hayati öneme sahip öncelikli bir konudur. Bu içerik evrenini koruyabilmek adına, örnekleme alınacak maddelerin seçiminde yalnızca istatistiksel kanıtlar değil, alan uzmanlığı da önemlidir. İncelenen araştırmaların yalnızca biri uzun formun kapsamını ayrıntılı bir şekilde tartışmış ve kısa forma madde seçerken istatistiksel analizlerin dışında kapsamı da dikkate almıştır. Diğer araştırmaların tümünde yalnızca istatistiksel kanıtlar dikkate alınmıştır.

*4. “Ölçme aracının uzun formu geçerli ise kısa formu da geçerlidir” şeklindeki bir düşünce yanlıştır:*

Kısa form, her ne kadar uzun forma ait maddeleri de içerse bu durum kısa formun geçerli ve güvenilir olacağını garanti etmez. Kısa form daha az madde sayısı ve içeriğe sahiptir. Bu açılarından psikometrik olarak dezavantajlıdır. Bu nedenle, mutlaka psikometrik özelliklerin istatistiksel olarak kanıtlanması gerekir. İncelenen araştırmaların tümünde kısa formun geçerlik ve güvenilirliği için istatistiksel kanıtlar aranmıştır.

*5. Çok boyutlu yapıya sahip ölçme araçlarında her bir boyut için içerik ve psikometrik özellikler incelenmelidir:*

Ölçeğe ait her bir maddenin, çok boyutlu yapılarda, ilgili boyutla ilişkilendirilmesi ile ölçeğin psikometrik özellikleri incelenmelidir. Bu durumda her bir boyutun geçerli ve güvenilir olduğuna dair kanıtlar sunulmalıdır. Örneğin, madde-toplam korelasyonlarına göre madde seçimi yapılacak ise toplam puan olarak ölçme aracının genel toplam puanı değil, faktör puanı seçilmelidir. Bir boyutta en az dört madde yer almasına dikkat edilmeli, bir boyutun ölçekten tamamen çıkarılması gerekli ise buna ait kuramsal ve istatistiksel alt yapı ayrıntılı bir şekilde verilmelidir. Boyuttaki madde sayısı azaldıkça kapsam geçerliğinin de azalacağı unutulmamalıdır. İncelenen araştırmalardan biri tek boyutlu bir yapıya sahip olduğu için inceleme dışı kalmıştır. Diğer araştırmalardan 10’u çok boyutlu yapıyı dikkate alarak istatistiksel analizleri gerçekleştirmiştir. Bir araştırma ise çok boyutlu bir yapıya rağmen gizil özelliğe dair kanıtları ölçeğin toplam puanı üzerinden elde etmiştir.

*6. Güvenirliliğe ait kanıtlar çeşitli güvenilirlik türleri çerçevesinde elde edilmelidir:*

Kısa forma ait geçerliliğin belirlenmesinde yapı geçerliği esastır. Ancak güvenirliliğin raporlaştırılmasında iç tutarlılık anlamında güvenirlilik, davranışa dayalı ölçmelerde puanlayıcılar arası tutarlılık ve kararlılık anlamında güvenirlilik gibi çeşitli güvenirlilik kanıtları elde edilmesi gerekir. Güvenirlilik hata ile ilişkili bir kavramdır ve bir ölçme işleminde tek bir hata türünden söz edebilmek mümkün değildir. Dolayısıyla hataya farklı açılardan bakan güvenirlilik katsayılarına yer verilmelidir. İncelenen on iki araştırmadan yalnızca bir tanesi hem iç tutarlılık hem de kararlılık anlamında güvenirlilik katsayıları elde etmiştir. Bu amaçla Cronbach alfa ve test-tekrar test güvenirlilik katsayıları kullanılmıştır. Araştırmalardan birinde güvenirlilik katsayısının raporlaştırılmadığı belirlenmiştir. Diğer araştırmaların tümü iç tutarlılık anlamında güvenirlilik katsayısı raporlaştırmıştır. Bu araştırmalardan sekizi Cronbach alfa güvenirlilik katsayısı, biri Raykov’un en büyük güvenirlilik katsayısını, biri ise kişi güvenirliliği ve kişi ayrıricılığı katsayılarını rapor etmiştir.

*7. Kısa formun psikometrik özellikleri uzun formdan bağımsız olarak da incelenmelidir:*

Ölçme aracının kısa formu, uzun form ile yüksek ilişki gösteren bir kopyasıdır. Yüksek olan bu ilişki, kısa formun geçerli ve güvenilir olduğunu kanıtlamaz. Geçerlik ve güvenirlilik kavramları birbirine aktarılabilir, geçişken değillerdir. Bu nedenle mutlaka uzun formdan bağımsız olarak da kısa formun

psikometrik özellikleri incelenmeli ve kanıtlar ortaya konulmalıdır. Bağımsız bir gruptan elde edilen bu kanıtlar, uzun forma ait geçerlik ve güvenirlik kanıtları ile karşılaştırılmalıdır. İncelenen araştırmaların yarısı uzun formdan bağımsız olarak da kısa forma ait geçerlik ve güvenirlik kanıtları elde ederken diğer yarısı yalnızca uzun formdan madde azaltmakla sınırla kalmıştır.

*8. Klinik ve davranışsal ölçme araçlarında kısa forma ait sınıflama doğrulukları da incelenmelidir:*

Bazı klinik ölçme araçları sınıflama yapma amaçlıdır. Olumsuz sınıflamayı (sendromu olan birine sendromu yok tanısı koyma) ve olumlu sınıflamayı (sendromu olmayan birine sendromu var tanısı koyma) yanlış yapmaktan kaçınmak hedeflenmelidir. Bu amaçla, uzun formdan bağımsız kanıtlar elde edilmelidir. Uzun formun doğru sınıflama ve tanı koyma işi yapması, kısa formun da bu amaçları yerine getirebildiğini garanti etmez. İncelenen araştırmalardan dördü klinik amaçlarla kullanılabilmektedir. Bu araştırmalardan hiçbirisi sınıflama doğruluğu ile ilgili bir kanıtı yer vermemiştir.

*9. Kısa form geliştirmenin kazandıracağı zamanın anlamlı ve önemli olduğu gösterilmelidir:*

Kısa form geliştirmenin en somut amaçlarından biri zamandan tasarruf etmektir. Ancak önceki bölümlerde de bahsedildiği gibi geçerlik ve güvenirlik zamandan daha önemlidir. Dolayısıyla araştırmacının kazandığı zamanın psikometrik özellikleri etkilemediğini göstermesi ve ne kadar zaman kazandığını açıklaması gerekir. Madde sayısı 80 olan bir uzun formu tamamlamak için ortalama olarak 40 dakika gerekir. Bu ölçme aracının kısa formunun 40 madde içerdiği varsayıldığında 20 dakikalık bir zaman tasarrufu sağlanacaktır. Ancak bilinmektedir ki, 40 maddelik bir azalışın geçerliğe ve güvenirliğe olumsuz etkileri olacaktır. Bu etkilerin düzeyi tartışılmalıdır. İncelenen araştırmalardan biri kısa form geliştirme sayesinde kazanılacak zamanı, ölçme aracının psikometrik özelliklerini de dikkate alarak tartışmıştır. Diğer araştırmalar ise sadece zamanın daha etkili kullanılacağını belirtmişlerdir.

## SONUÇ

Bir ölçme aracının kısa formunu geliştirirken araştırmacıların en büyük yanılgılarından biri, orijinal ölçme aracı ile kısa formun geçerlik ve güvenirlik bilgilerinin aynı olduğu düşüncesidir. Bu durum bazı araştırmacıların geçerlik ve güvenirlik gibi psikometrik özellikleri yok saymasına, bazı araştırmacıların ise bu konuda gereken önemi göstermemesine neden olmaktadır. Kısa form geliştirilirken gözlenen madde sayısı azalır. Dolayısıyla içerik ve kapsam daralır. Bu nedenle bu iki test formunun birbirine alternatif olabilmesi güçleşmektedir.

Bu araştırma için seçilen 12 araştırma sosyal ve sağlık bilimler alanındaki iki önemli indekste taranmaktadır. İnceleme sonuçları göstermektedir ki kısa form geliştirme ile ilgili kurallara uyan çok az araştırma vardır. Bu durum, 20. yüzyılın ortalarından itibaren tartışılan kısa form geliştirme çalışmalarının halen tartışmaya açık olduğunu göstermektedir. Levy (1968); Smith, McCarthy ve Anderson'un (2000) çalışmaları ile karşılaştırıldığında, incelenen çalışmaların daha doğru hareket ettiği de bir gerçektir. Ancak incelenen çalışmalarda, ölçme aracının uzun formuna ait geçerlik ve güvenirlik bilgilerinin daha ayrıntılı verilmemesi, kısa formun en az uzun form kadar geçerli ve güvenilir olması gerektiği konusuna dikkat edilmemesi, ölçme aracının uzun formu kadar geçerli ve güvenilir bir kısa form ortaya koyulamadığı durumlarda "değiştirilebilir" kısa form kavramının ortaya çıkacağının farkında olunmaması, güvenirlik raporlaştırmasının yalnızca iç tutarlılık katsayıları ile sınırlı kalması kısa forma ait psikometrik özelliklerin uzun formdan bağımsız olarak elde edilmemesi, ölçme aracının uzun formunun kapsamını ayrıntılı bir şekilde açıklama ve kısa formun kapsamının uzun forma da genellenebileceğine dair kanıtlar elde etmeme gibi önemli sorunlar olduğu belirlenmiştir. Çalışmalarda, kısa form geliştirmenin temel amaçlarından biri olan zamandan ne kadar tasarruf edildiği ve bu tasarrufun psikometrik özellikleri nasıl etkilediğinin tespit edilmemesi gibi eksiklikler de bulunmaktadır. Ayrıca klinik ölçme araçlarında sınıflama doğruluğunun test edilmediği de saptanmıştır.

Araştırmacılara ölçme aracının uzun formundan hareket ederek, bu çalışmada bahsedilen ölçütleri de dikkate alarak, ölçeğin kısa formunu geliştirmesi önerilmektedir. Ölçme aracının kısa formunu uyarlama çalışmaları ise seçilen kısa formun psikometrik özelliklerinin şüpheli olabilmesi gibi bazı önemli konulardan dolayı tavsiye edilmemektedir. Bu nedenle kısa form uyarlama çalışmaları yerine öncelikle ölçme aracının uzun formunun ilgili kültüre uyarlanması, ardından uyarlanmış ölçme aracının kısa formunun geliştirilmesinin daha sağlıklı bir yaklaşım olacağı önerilmektedir.

Özellikle klinik araştırmacılar açısından bakıldığında kısa formu geliştirilmiş bir ölçme aracını kullanmadan önce hakemli bir dergide yayınlanmış olması ve daha fazla zaman tasarrufu sağlayacak olması, bu ölçme aracının seçimi için yeterli değildir. Etik kurallara uygun bir şekilde geliştirilen ve psikometrik özelliklere yeterli düzeyde sahip olan ölçme araçlarının seçilmesi önerilmektedir. Klinik araştırmacılar bilmelidir ki daha az madde içeren ölçme araçlarının daha az geçerli olduğu şeklindeki bir algı doğruyu göstermemektedir. Her ölçme aracı için aynı psikometrik standartlar aranmaktadır. Ayrıca seçilecek olan kısa formun geliştirilme aşamasında bu çalışmada önerilen basamaklara uygun hareket edilip edilmediği de eleştirel bir gözle incelenmeli, yeterli özelliklere sahip kısa form klinik veya diğer amaçlarla kullanılmalıdır.

#### KAYNAKÇA

- Baiocco, R., Pallini, S., & Santamaria, F. (2014). The development and validation of an Italian short form of the adolescent friendship attachment scale. *Measurement and Evaluation in Counseling and Development*, 47(4), 247-255. <https://doi.org/10.1177/0748175614538060>
- Biggers, J.L. (1976). An a priori approach for developing short-forms of tests and inventories. *The Journal of Experimental Education*, 44(3), 8-10. <https://doi.org/10.1080/00220973.1976.11011528>
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M Lord, & R. Novick (Eds.), *Statistical Theories of Mental Test Scores* (pp. 397-479). Reading, MA: MIT Press.
- Bohlmeijer, E., Ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment*, 18(3), 308-320. <https://doi.org/10.1177/1073191111408231>
- Doll, E. A. (1917). A brief Binet-Simon scale. *Psychological Clinic*, 11, 197-211.
- Doppelt, J. E. (1956). Estimating the full scale score on the Wechsler Adult Intelligence Scale from scores on four subjects. *Journal of Consulting Psychology*, 20(1), 63. <https://doi.org/10.1037/h0044293>
- Dorigo, M., & Stützle, T. (2004). *Ant colony optimization*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/1290.001.0001>
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: L. Erlbaum Associates.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Ferrario, S. R., Panzeri, A., Anselmi, P., & Vidotto, G. (2019). Development and psychometric properties of a short form of the Illness Denial Questionnaire. *Psychology Research and Behavior Management*, 12, 727. <https://doi.org/10.2147/PRBM.S207622>
- Jenkinson, C., Kelly, L., Dummett, S., & Morley, D. (2019). The Oxford Participation and Activities Questionnaire (Ox-PAQ): development of a short form and index measure. *Patient Related Outcome Measures*, 10, 227-232. <https://doi.org/10.2147/PROM.S210416>
- Goretzko, D., Pham, T. T. H., & Bühner, M. (2019). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, (In Press), 1-12. <https://doi.org/10.1007/s12144-019-00300-2>
- Hagtevt, K. A., & Sipos, K. (2016). Creating short forms for construct measures: The role of exchangeable forms. *Pedagogika*, 66(6), 689-713.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Publications.
- Henning, G. (1987). *A guide to language testing- development, evaluation, Research*. London: Newbury House Publisher.
- Koğar, H. (2015). Madde tepki kuramına ait parametrelerin ve model uyumlarının karşılaştırılması: Bir Monte Carlo Çalışması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 142-157. <https://doi.org/10.21031/epod.02072>

- Lei, P. W., Dunbar, S. B., & Kolen, M. J. (2004). A comparison of parametric and nonparametric approaches to item analysis for multiple choice tests. *Educational and Psychological Measurement*, 64, 565–587. <https://doi.org/10.1177/0013164403261760>
- Leite, W. L., Huang, I. C., & Marcoulides, G. A. (2008). Item selection for the development of short forms of scales using an ant colony optimization algorithm. *Multivariate Behavioral Research*, 43(3), 411-431. <https://doi.org/10.1080/00273170802285743>
- Levy, P. (1968). Short-form tests: A methodological review. *Psychological Bulletin*, 69(6), 410. <https://doi.org/10.1037/h0025736>
- Lim, S. Y., & Chapman, E. (2013). Development of a short form of the attitudes toward mathematics inventory. *Educational Studies in Mathematics*, 82(1), 145-164. <https://doi.org/10.1007/s10649-012-9414-x>
- Meijer, R. R., & Baneke, J. J. (2004). Analyzing psychopathology items: A case seng for nonparametric item response theory modeling. *Psychological Methods*, 9, 354–368. <https://doi.org/10.1037/1082-989X.9.3.354>
- Milavic, B., Padulo, J., Grgantov, Z., Milić, M., Mannarini, S., Manzoni, G. M. ..., & Rossi, A. (2019). Development and factorial validity of the Psychological Skills Inventory for Sports, Youth, version–Short Form: Assessment of the psychometric properties. *PloS one*, 14(8), 1-17. <https://doi.org/10.1371/journal.pone.0220930>
- Mokken, R. J., & Lewis, C. (1982). A nonparametric approach to the analysis of dichotomous responses. *Applied Psychological Measurement*, 6, 417–430. <https://doi.org/10.1177/014662168200600404>
- Morin, D., Valois, P., Crocker, A. G., & Robitaille, C. (2019). Development and psychometric properties of the Attitudes Toward Intellectual Disability Questionnaire–Short Form. *Journal of Intellectual Disability Research*, 63(6), 539-547. <https://doi.org/10.1111/jir.12591>
- Nimon, K., & Zigarmi, D. (2015). Development of the work intention inventory short-form. *New Horizons in Adult Education and Human Resource Development*, 27(1), 15-28. <https://doi.org/10.1002/nha3.20090>
- Park, H., & Hill, R. B. (2018). Development and validation of a short form of the occupational work ethic inventory. *Journal of Career and Technical Education*, 32(1), 9-28. <https://doi.org/10.21061/jcte.v32i1.1588>
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Denmark's Paedagogiske Institut.
- Rogers, M. E., Creed, P. A., Searle, J., & Hartung, P. J. (2011). The physician values in practice scale-short form: Development and initial validation. *Journal of Career Development*, 38(2), 111-127. <https://doi.org/10.1177/0894845310363593>
- Sodano, S. M., & Tracey, T. J. (2011). A brief Inventory of Interpersonal Problems–Circumplex using nonparametric item response theory: Introducing the IIP–C–IRT. *Journal of Personality Assessment*, 93(1), 62-75. <https://doi.org/10.1080/00223891.2010.528482>
- Raykov, T., & Marcoulides, G. A. (2011). Classical item analysis using latent variable modeling: a note on a direct evaluation procedure. *Structural Equation Modeling*, 18(2), 315-324. <https://doi.org/10.1080/10705511.2011.557347>
- Siefert, C. J., Sexton, J., Meehan, K., Nelson, S., Haggerty, G., Dauphin, B., & Huprich, S. (2019). Development of a short form for the DSM–5 levels of personality functioning questionnaire. *Journal of Personality Assessment*, (In Press), 1-11. <https://doi.org/10.1080/00223891.2019.1594842>
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102-111. <https://doi.org/10.1037/1040-3590.12.1.102>
- Woudstra, A. J., Meppelink, C. S., Maat, H. P., Oosterha, ven, J., Fransen, M. P., & Dima, A. L. (2019). Validation of the short assessment of health literacy (SAHL-D) and short-form development: Rasch analysis. *BMC Medical Research Methodology*, 19(1), 122-131. <https://doi.org/10.1186/s12874-019-0762-4>