
Kuram ve Uygulamada SOSYAL BİLİMLER DERGİSİ

Social Sciences: Theory & Practice

ISSN: 2619-9408

Geliş/Received: 08.06.2024 *Kabul/Accepted:* 03.08.2024

Makale Türü: Derleme

Ölçek Geliştirme Süreçlerine İlişkin Bir Derleme Çalışması

*Fatma COŞKUN**

*Büşra EREN***

ÖZ

Sosyal bilimler alanında nicel paradigmaya dayalı araştırmaların yürütülmesinde, insanlara ait doğrudan gözlemlenemeyen örtük özellikleri ölçmek amacıyla kullanılan ölçeklerin rolü yadsınamaz bir gerçektir. Özellikle eğitim, psikoloji, sosyoloji gibi sosyal bilimlerin insan davranışına odaklanan daha özgül alt alanlarında üretilen bilgi, büyük oranda ölçekler aracılığıyla elde edilen verilere dayanmaktadır. Ölçeklerin kullanımı sadece bilimsel çalışmalar ile sınırlı değildir. Okul rehberlik servisleri veya psikoloji kliniklerinde kullanılan ölçekler, bireyler hakkında verilecek kararları büyük oranda etkilemektedir. Bu nedenle doğrudan veya dolaylı olarak insan yaşamı üzerinde etkili olan ölçeklerin geçerli ve güvenilir bir şekilde geliştirilmesi oldukça önemlidir. Ancak alan yazın incelendiğinde sosyal bilimler alanında ölçeklerin kullanımı bir asırdan daha eskiye dayanmasına rağmen ölçek geliştirme çalışmalarında pek çok kuramsal ve teknik hataların olduğu görülmektedir. Bu çalışmanın amacı, ölçek geliştirme süreçlerinde yaygın olarak görülen hataları betimleyerek ölçek geliştirme çalışmalarında uyulması gereken temel prosedürler ile birlikte ölçek geliştirme alanındaki yeni yaklaşımları okuyucuya tanıtmaktır. Bu açıdan, bu çalışmanın yeni ölçek geliştirme uygulamaları için bir kılavuz görevi üstlenerek geçerli ve güvenilir bilginin üretilmesine katkı sunması beklenmektedir.

Anahtar Kelimeler: Ölçek geliştirme prosedürleri, madde oluşturma, geçerlik ve güvenilirlik, Cronbach alpha, faktör analizi.

An Overview of Scale Development Processes

ABSTRACT

In the field of social sciences, the role of scales used to measure implicit characteristics of individuals that cannot be directly observed is undeniably significant in conducting research based on the quantitative paradigm. Particularly in more specific subfields of social sciences such as education, psychology, and sociology that focus on human behavior, the knowledge produced is largely based on data obtained through scales. The use of scales is not limited to scientific studies. Scales used in school guidance services or psychology clinics significantly influence decisions made about individuals. Therefore, it is crucial that scales, which have a direct or indirect impact on human life, are developed in a valid and reliable manner. However, a review of the literature reveals that, despite the use of scales in the social sciences dating back more than a century, there are numerous theoretical and technical errors in scale development studies. The aim of this study is to describe the common errors seen in scale development processes and to introduce the fundamental procedures that should be followed in scale development studies, along with new approaches in the field. In this regard, it is expected that this study will serve as a guide for new scale development practices and contribute to the production of valid and reliable information.

Atf Bilgisi: Coşkun, F., & Eren, B. (2024). Ölçek geliştirme süreçlerine ilişkin bir derleme çalışması. *Kuram ve Uygulamada Sosyal Bilimler Dergisi*, 8(2), 466-477. Doi: 10.48066/kusob.1498085

*Dr. Öğr. Üyesi, Kahramanmaraş Sütçü İmam Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, fatmacoskun@ksu.edu.tr, ORCID: orcid.org/0000-0002-6388-3504.

** Öğr. Gör. Dr., Milli Savunma Üniversitesi, Kara Astsubay MYO, Eğitim Bilimleri Bölümü, busra_karaduman@yahoo.com, ORCID: orcid.org/0000-0001-7565-1025

Keywords: Scale development procedures, item creation, validity and reliability, Cronbach's alpha, factor analysis.

Giriş

İçinde bulunduğumuz 21. yüzyılda dünyanın herhangi bir ülkesinde iş, eğitim, sağlık, askerlik gibi pek çok alanda bireylerin duyuşsal özelliklerinin belirlenmesi amacıyla ölçekler kullanılmaktadır. Ölçekler aracılığıyla bireylerin kaygı, motivasyon, algı, tutum, zekâ gibi doğrudan gözlemlenemeyen özellikleri değerlendirilmekte ve birtakım kararlar verilmektedir. Ölçekler aynı zamanda eğitim, psikoloji, sosyoloji gibi alanlarda bilimsel bilginin üretilmesi ve test edilmesinde kullanılmaktadır. Bu açıdan bilimin doğru ve tutarlı bilgiler üzerinden ilerleyebilmesi kullanılan ölçme araçlarının geçerlik ve güvenilirlik düzeyleriyle doğrudan ilişkilidir (Karakoç ve Dönmez, 2014).

Ölçekler aracılığıyla bir grup somut ifade kullanılarak doğrudan gözlemlenemeyen örtük özellikler ölçülmeye çalışılmaktadır (Carpenter, 2018). Başka bir deyişle ölçekler, genel olarak eğitim ve psikoloji gibi alanlarda bireylerin doğrudan gözlemlenemeyen özellikleri için bireylere örtük özelliklerle ilişkili birtakım maddeler sunularak bireylerin maddelere verdikleri tepkilerden hareketle örtük özelliğin yapısı ve düzeyi hakkında bilgi edinilmesi amacıyla kullanılmaktadır (Erkuş, 2012).

Ölçekler aracılığıyla doğrudan gözlemlenemeyen örtük özellikler ölçüldüğünden ölçeklerin niteliği oldukça önemlidir (Kaya Uyanık, vd., 2017). Bununla beraber ölçek geliştirme çalışmaları bittikten sonra da ölçeklerin kullanıldıkları yeni araştırmalar içerisinde takip edilmesi ve psikometrik özelliklerinin değerlendirilmesi gerekmektedir (Mor Dirlik, 2014). Ölçek geliştirme süreçleri, dikkat isteyen süreçler olmasına rağmen özellikle Türk kültüründe geliştirilen ölçeklerin sayısında büyük bir artış görülmektedir. Son yıllarda akademisyenler üzerindeki yayın baskısı nedeniyle doğrudan alanıyla ilişkili olan ve olmayan pek çok akademisyenin “en kolay” yayın türü algısı ile ölçek geliştirme çalışmalarına yöneldiği görülmektedir (Erkuş, 2007). Ulusal düzeyde geliştirilen ölçek sayısındaki artışın tersine ölçeklerin içeriğinde niteliksel açıdan pek çok yanlış veya kavram yanılgısı görülmeye devam etmektedir (Mor Dirlik, 2014). Ölçek geliştirme basit, kullanışlı ve analizlerinin kolay olduğu konusundaki yaygın görüş yanıltıcı olabilmektedir (Simms, 2008). Çünkü ölçek geliştirme süreci, sanıldığı gibi aksine yüksek düzeyde teorik, metodolojik ve istatistiksel yeterlikler gerektirmektedir (Carpenter, 2018).

Ölçek geliştirme çalışmalarının sağlam bir dayanağının olması gerekmektedir. Bu nedenle araştırmacılar ölçek geliştirme çalışmasına başlamadan önce, ölçülmek istenen yapıya ilişkin kullanılabilir bir ölçeğin var olup olmadığını incelemeli ve farklı bir kültürde istenen yapının ölçülmesi için alan yazında geçerli bir ölçek var ise uyarlama seçeneği de değerlendirmelidir. Ölçek geliştirme çalışmasında karar kılınmadan önce gerekli araştırmalar yapılarak bütün ihtimaller titizlikle gözden geçirilmelidir (Karakoç ve Dönmez, 2014). Nitekim Tavşancıl vd. (2014) yaptıkları araştırmada Türk kültüründe aynı örtük özelliğin ölçülmesi amacıyla farklı araştırmacılar tarafından farklı ölçeklerin geliştirildiğini belirlemişlerdir. Böyle bir durum, alan yazına yeni ve özgün bir katkı sağlamayıp araştırmacılar için zaman kaybına neden olabilmektedir. Bu nedenle araştırmacıların alan yazında ölçülmek istenen özelliğe dair geçerli bir ölçek olmadığına emin olduktan sonra ölçek geliştirme çalışmalarına başlaması daha doğru bir yaklaşım sergileyecektir (Kaya Uyanık, vd., 2017). Öte yandan ölçek geliştirme çalışmasını yürütecek araştırmacıların hem ölçülecek örtük özellik hem de ölçme ve değerlendirme adımları konusunda belli bir uzmanlık düzeyinde bilgi sahibi olması gerekmektedir. Aksi takdirde bu konuda uzman olmayan araştırmacılar tarafından yapılan ölçek geliştirme çalışmaları bilime en büyük zararı verecektir (Erkuş, 2014). Ölçek geliştirme prosedürleri dikkatli bir şekilde uygulanmadan geliştirilmiş ölçekler, kuramsal veya teknik açıdan hatalı olabilmektedir. Bu durum zaman ve emek kaybına neden olmaktadır. Ayrıca bu tür hatalı geliştirilmiş ölçeklerin akademik yayına

dönüştürülmesi durumunda, başka araştırmacılar tarafından kullanılarak yanlış bulguların elde edilmesine veya bireyler hakkında karar vermek amacıyla kullanıldığında ise yanlış kararlar alınmasına yol açmaktadır (Çüm ve Koç, 2013). Belirli alanlar dışında, doktora programlarında genel olarak ölçek geliştirme süreçleriyle ilgili herhangi bir ders okutulmamaktadır (Boateng vd., 2018). Bu nedenle pek çok araştırmacı alan yazında yer alan başka çalışmaları taklit ederek ölçek geliştirme süreçlerini uygulamaktadır (Carpenter, 2018).

Alan yazında, farklı araştırmacıların farklı dönemlerde yayımlanmış ölçek geliştirme çalışmalarını incelediği birtakım araştırmalar bulunmaktadır. Bu tür araştırmalardan birini gerçekleştiren Çüm ve Koç (2013) Türkiye’de 2005-2013 yılları arasında belirli psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarını, ölçek geliştirme ve uyarlama prosedürlerinin uygulanması açısından değerlendirmişlerdir. İncelenen ölçek geliştirme çalışmalarının %37’sinde (N=29) ölçek geliştirme adımlarına; %60’ında (N=29) ise ölçek geliştirme ilkelerine dair bilgilerin uygun olmadığı araştırmacılar tarafından raporlanmıştır. Tavşancıl, Güler ve Ayan (2014) Türkiye’de eğitim ve psikoloji alanlarında yayımlanan akademik dergilerdeki tutum ölçeği geliştirme çalışmalarını incelemişlerdir. Araştırmacılar, inceledikleri çalışmalardan hareketle çok sayıda yanlış ve eksik tutum ölçeği olduğunu belirlemişlerdir. Mor Dirlik (2014) Türkiye’de eğitim bilimleri alanında ölçek geliştirme konusunda yapılmış olan beş doktora tezini, ölçek geliştirme aşamaları açısından değerlendirmiştir. Araştırmacı, incelediği beş tezden ikisinin ölçek geliştirme prosedürlerini yüksek düzeyde doğru gerçekleştirdiğini, bir tezin orta düzeyde kaldığını ve iki tezin ise ölçek geliştirme prosedürlerini karşılama konusunda oldukça yetersiz kaldığını tespit etmiştir. Acar Güvendir ve Özer Özkan (2015) Türkiye’de 2006-2014 yılları arasında eğitim bilimleri alanında yayımlanan belirli dergilerdeki ölçek geliştirme ve ölçek uyarlama çalışmalarını incelemişlerdir. Araştırmacılar inceledikleri çalışmalarda birkaç adımın genel olarak benzer şekilde uygulandığını ancak süreçlere ilişkin üzerinde uzlaşmış prosedürler bütünü gözlenemediğini raporlamışlardır. Kaya Uyanık vd. (2017) Türkiye’de 2010-2015 yılları arasında eğitim bilimleri alanında yayımlanan belli akademik dergilerdeki ölçek geliştirme çalışmalarını, ölçek geliştirme ilkelerine uygunluğu açısından değerlendirmişlerdir. Araştırmacılar, inceledikleri elli yedi ölçek geliştirme makalesinden sadece bir tanesinde ölçek geliştirme ile ilgili tüm prosedürlerin uygulandığını; beş makalede prosedürlere ilişkin bilgi bulunduğu ancak prosedürlerin kısmen açıklandığı; kırk bir makalede prosedürlere ilişkin bilgilere yer verildiği ancak açıklama yapılmadığı; on makalede ise prosedürlere ilişkin hiçbir bilgi veya açıklamaya yer verilmediğini raporlamışlardır. Güler ve Ayan (2020) Türkiye’de 2002-2018 yılları arasında belirledikleri eğitim bilimleri dergilerinde yayımlanmış tutum ölçeği geliştirme çalışmalarını nitel araştırma yöntemiyle incelemişlerdir. Araştırmacılar, inceledikleri çalışmalarda var olan pek çok eksik ve yanlış ortaya çıkarmıştır. Öte yandan uluslararası yüksek indeksli dergilerde yayımlanan ölçek geliştirme çalışmalarının dahi ciddi kusurlar içerdiği görülmektedir (Carpenter, 2018). Sonuç olarak farklı dönemlere ait ölçek geliştirme çalışmalarını değerlendirmiş olan bütün bu çalışmalar göz önüne alındığında, ölçek geliştirme süreçleri ile ilgili ciddi düzeyde eksik veya yanlış bilginin olduğu savunulabilir. Bu çalışmanın amacı ölçek geliştirme süreçlerini olabildiğince basit, anlaşılır ve sade bir şekilde açıklayarak ölçek geliştirmek isteyen araştırmacılara kılavuz görevi üstlenecek bir yayın sunmaktır. Bu amaç doğrultusunda ulusal ve uluslararası düzeyde ölçek geliştirme süreçleriyle ilgili temel kaynaklar referans alınarak sürece ilişkin tüm adımlar olabildiğince basit bir şekilde açıklanmıştır.

Alan yazında ölçek geliştirme süreçleriyle ilgili farklı araştırmacılar tarafından ele alınmış pek çok akademik yayın bulunmaktadır (AERA, APA ve NCME, 1999; DeVellis, 2012; Erkuş, 2012; Erkuş, 2014; Simms, 2008; Tavşancıl, 2018). Alan yazındaki ölçek geliştirme çalışmalarında ortak bir yaklaşımın benimsenememesinin bir nedeni olarak da ölçek geliştirme süreçleriyle ilgili birbirinden farklı pek çok kaynağın varlığı gösterilebilir. Ölçek geliştirme süreçleriyle ilgili yayınlar incelendiğinde, farklı yayınlarda ölçek geliştirme süreçleriyle ilgili ayrı sınıflandırmaların ve farklı vurguların yapılması

dikkat çekmektedir. Carpenter (2018) ölçek geliştirme ile ilgili alan yazından ve temel kuramlardan yola çıkarak ölçek geliştirme çalışmaları için araştırmacıların uymalarının gerekli olduğu on adım tanımlamıştır. Ancak Carpenter'ın gerçekleştirdiği bu çalışmadaki öneriler yalnızca faktör analizinin kullanıldığı araştırmalar ile sınırlandırılmıştır. Simms (2008) ölçek geliştirme için geleneksel ve modern yaklaşımları ele aldığı çalışmasında ölçek geliştirme sürecini maddi geçerlik, yapı geçerliği ve dış geçerlik maddeleri altında üç bölüme ayırarak açıklamıştır. DeVellis (2012) ölçek geliştirme süreçlerini birbirini izleyen sekiz adımda açıklamıştır. Bu adımlar sırasıyla; ölçülmesi amaçlanan yapının tanımlanması, madde havuzunun hazırlanması, ölçek tipinin belirlenmesi, uzman görüşünün alınması, belirlenen maddelerin uygulama formuna yazılması, ölçek geliştirme örnekleme ilgili formun uygulanması, maddelere dair değerlendirmelerin yapılması ve son olarak ölçek uzunluğuna karar verilerek nihai ölçek formunun raporlanması şeklindedir. Erkuş (2014) sürece ilişkin keskin ayrımlarından ziyade ölçek geliştirme sürecini ölçek geliştirmeye hazırlık, ölçek geliştirme süreci ve psikometrik analiz temaları altında ele almıştır. Morgado vd. (2017) ise ölçek geliştirme süreçlerini “madde oluşturma”, “teorik analiz” ve “psikometrik analiz” şeklinde üç tema altında açıklarken, Boateng vd. (2018) ölçek geliştirme süreçlerini dokuz adım içeren üç aşama şeklinde sınıflayıp açıklamıştır. Araştırmacılar söz konusu üç aşamayı; “ölçek maddelerinin yazılması”, “ölçeğin oluşturulması” ve “ölçeğin değerlendirilmesi” şeklinde isimlendirmişlerdir. Bu çalışma kapsamında ölçek geliştirme süreci, incelenen bütün alan yazını temsil edecek şekilde birbirini izleyen üç aşamada açıklanmıştır. Bu aşamalar “Taslak Ölçek Formunun Hazırlanması”, “İstatistiksel Analizlerin Uygulanması” ve “Uygulanan Ölçek Çalışmasının Raporlanması” şeklinde isimlendirilmiştir. Alan yazın taramasına dayanan bu çalışmada geleneksel derleme yöntemi kullanılmıştır (Li ve Wang, 2018). Bu araştırma kapsamında ölçek geliştirme sürecine ilişkin belirtilen üç aşama ve uygulama adımlarına dair bilgiler Tablo 1’de sunulmuştur.

Tablo 1. Ölçek geliştirme sürecinin aşamaları ve uygulama adımları

Taslak Ölçek Formunun Hazırlanması	İstatistiksel Analizlerin Uygulanması	Uygulanan Ölçek Çalışmasının Raporlanması
1-Ölçülmesi amaçlanan özelliğin teorik yapısına dair alan yazının incelemesi	1-Hangi istatistiksel yöntemlerin kullanılacağına karar verilmesi	1-“Nihai Ölçek Formu” nun hazırlanması
2-Hedef grubu temsil eden katılımcılara kompozisyon yazdırma ya da gözlem, görüşme vb. yollar ile nitel verilerin toplanması	2-İstatistiksel yöntem uygun örneklemin belirlenmesi	2-Ölçek geliştirme sürecinde uygulanan tüm adımların ve geçerlik-güvenirlik kanıtlarının şeffaf bir şekilde sunulması
3-Birinci ve ikinci adımdan hareketle, ölçülmesi amaçlanan yapının teorik ve işevuruk tanımının yapılması	3-Uygulama türünün belirlenmesi (çevrimiçi/yüz yüze)	3-Geliştirilen ölçeğin nasıl kullanılması gerektiğine dair yönerge ve bilgileri içeren kullanım kılavuzunun hazırlanması
4- Ölçek tipine karar verilerek maddelerin oluşturulması	4- Uygulama ortamından ve/veya uygulamayı gerçekleştiren kişiye ilişkin hata kaynaklarının değerlendirilmesi ve gerekli görülen hallerde eğitimlerin verilmesi	
5-Ölçek formunun hazırlanması ve uzman görüşüne sunulması	5-Verilerin Toplanması	
6-Pilot uygulama yapılarak katılımcı tepkilerinin gözlenmesi	6-Veri kalitesinin kontrol edilmesi ve varsayımların incelenmesi	
8-Pilot uygulama sonucunda son düzeltmelerin yapılması ve “Taslak Ölçek Formu” nun uygulama için hazır hale getirilmesi	7-Geçerlik ve güvenilirlik analizlerinin yapılması	

Kaynak: (AERA, APA ve NCME, 1999; DeVellis, 2012; Erkuş, 2012; Erkuş, 2014; Simms, 2008; Tavşancıl, 2018; Carpenter, 2018; Morgado vd., 2017; Boateng vd., 2018)

Bu çalışmada Tablo 1’de sunulan üç aşama ve uygulama adımları alışlagelmiş kavram yanılguları veya yanlış bilgilerin düzeltilmesine katkı sunması amacıyla, mevcut ölçek geliştirme çalışmalarında en çok yapılan hatalara atıflar yapılarak derleme çalışması olarak sunulmuştur. Bu kapsamda ayrıca ölçek geliştirme süreçleri ile ilgili yeni yaklaşımlar ve güncel gelişmeler hakkında da bilgiler sunularak çalışmanın amacı genişletilmiştir.

Taslak Ölçek Formunun Hazırlanması

Yaygın kaniya göre ölçek geliştirme çalışmasının en önemli aşaması istatistiksel analizlerin uygulandığı aşama olarak görülmektedir. Ancak ölçek geliştirme çalışmalarının kalbi sayılabilecek asıl aşama taslak ölçek formunun geliştirildiği süreçtir. İstatistiksel analizlerin uygulandığı aşamada yalnızca geliştirilen taslak ölçek formunun geçerliğine ve güvenilirliğine dair deneysel kanıtlar elde edilmeye çalışılmaktadır. Öte yandan ölçek geliştirme çalışmalarında hangi kuram temel alınırsa alınsın taslak ölçek hazırlama aşaması hepsinde ortaktır. Bu nedenle ölçek geliştirme çalışmalarının ilk aşaması olan “Taslak Ölçek Formu Hazırlama” aşamasının oldukça özenli ve dikkatli bir şekilde yürütülmesi gerekmektedir.

Hinkin’e (1995) göre ölçek geliştirme çalışmalarında ölçekten kaynaklı hataları en aza indirmek için birtakım önlemlerin alınmasını gerektiren kısım, taslak ölçek formunun hazırlandığı ilk aşamadır. Bu aşamanın ilk adımı ölçülmesi amaçlanan özelliğe dair teorik yapının ortaya çıkarılması amacıyla alan yazının kapsamlı bir şekilde taranmasıdır. İkinci adım ise ölçülmesi amaçlanan özelliğin teorik yapıdan ziyade hedef grup ile gerçekleştirilen görüşmeler sırasında nitel yöntemler yoluyla belirlenmeye çalışılmasıdır. Ölçek geliştirme çalışmalarında ölçülmesi amaçlanan örtük özelliğin teorik ve kavramsal yapısı kapsamlı bir şekilde en ince ayrıntısına varıncaya kadar araştırılmalıdır. Ancak böyle bir inceleme yeterli olmayıp araştırmacıların ölçülmesi amaçlanan örtük özelliklerle ilgili hedef gruba kompozisyon yazdırma, hedef grubu gözleme, uzman görüşlerine başvurma, ilgili araştırmaların bulguları gibi farklı yöntemleri bir arada kullanarak örtük özelliğin yapısını oluşturması gerekmektedir (Erkuş, 2012). Ölçülecek özelliklerle ilgili kuramsal ve nitel araştırmaya dayalı incelemeler, soyut yapıyı temsil eden deneye dayalı niteliklerin tanımlanmasına rehberlik etmede en güçlü rolü oynamaktadır. Bu nedenle ölçek geliştirme çalışmalarında madde yazımına başlanmadan önce kuramsal ve nitel araştırmaya dayalı incelemeler aracılığıyla örtük özelliğin yapısı ve anlamı net bir şekilde ortaya konmalıdır (Carpenter, 2018). Ölçek geliştirme çalışmalarının bu aşamasında kullanılan nitel yöntemlerle ilgili daha ayrıntılı bilgiler için Rowan ve Wulff (2007)’un gerçekleştirdiği çalışmanın incelenmesi önerilebilir. Ölçek geliştirme çalışmalarında “Taslak Ölçek Formu Hazırlama” aşamasının üçüncü adımında araştırmacı ölçülmesi amaçlanan örtük özelliğin teoriye dayalı kuramsal tanımını ve işevuruk tanımını oluşturmalıdır. Güler ve Ayan (2020) bu adımla ilgili olarak inceledikleri ölçek geliştirme çalışmalarında, ölçülmesi amaçlanan özelliğin teorik temellerine dair bilgilere yer verildiğini fakat söz konusu özelliğin işevuruk tanımının genel olarak yapılmadığını veya eksik bırakıldığını tespit etmişlerdir. Erkuş (2012) ölçek geliştirme sürecini, ölçülmesi amaçlanan örtük özelliğin işevuruklaştırılması işlemi olarak tanımlamıştır. İşevuruk tanım, ölçülmesi amaçlanan özelliğin teorik tanımıyla ilişkili olmasının yanı sıra daha özel anlamıyla araştırma kapsamında kullanılan ölçek yapısı ile ilişkili tanımdır (Erkuş, 2012). Bu nedenle ölçek geliştirme çalışmalarında, araştırmacıların örtük özelliğe ait kuramsal ve işevuruk tanımını net olarak belirleyip ikisini birbirinden ayırt edebilmesi gerekmektedir.

“Taslak Ölçek Formu Hazırlama” aşamasının en zor sayılabilecek adımı, ölçek maddelerinin oluşturulmasıdır. Ölçek maddelerinin oluşturulması sürecinde mülakatlar, odak gruplar ve uzman geri bildirimleri ciddi bir role sahiptir (Carpenter, 2018). Ölçek maddelerinin her biri, örtük özelliklerle ilgili bir gözlemi temsil etmektedir. Bu nedenle ölçek maddelerinin yazımında uyulması gereken birtakım

kurallar vardır. Alan yazında farklı araştırmacılar tarafından yapılmış açıklamalardan hareketle ölçek maddelerini oluştururken uyulması gereken genel kurallar Tablo 2’de özetlenmiştir.

Tablo 2. Ölçek maddelerini oluştururken uyulması gereken kurallar

Madde Oluşturma Süreci
1-Oluşturulan her madde, ölçülmesi amaçlanan örtük özelliğin bir göstergesi olmak zorundadır.
2-Maddeler hedef grubun düzeyine (yaş, eğitim, zekâ, okuma vb.) uygun kelimeler içermelidir.
3-Maddeler tüm katılımcılar tarafından anlaşılacak biçimde sade, anlaşılır ve açık ifadelerden oluşmalıdır.
4-Herbir maddede yalnızca bir eylem veya durum yoklanmalıdır. Örneğin: “Ben sağlıklı beslenmeye ve spor yapmaya özen gösteririm” gibi bir maddede katılımcıların hangi eyleme tepki verdiği belirsizdir.
5-Ölçek maddelerinde günlük kullanıma dair argo veya popüler kültürün ürünü olan geçici kelimelerin kullanımından kaçınılmalıdır.
6-Maddelerde “tekrarlı” olumsuzluk oluşturacak kelimelerin kullanılmasına dikkat edilmelidir. Örneğin: “Spor yapmaktan keyif almam.” ifadesi yerine “Spor yapmaktan keyif alırım” ifadesi kullanılmalıdır.
7-Madde içeriğinde sıklık zarfları kullanılmamalıdır. Örneğin “Her zaman spor yaparım” ifadesi yerine “Spor yaparım” ifadesi kullanılmalıdır.
8-Olumlu ve olumsuz maddelerin sayısı birbiriyle dengeli olmalıdır.
9-Farklı alt gruplara taraflı davranacak maddeler yazmaktan kaçınılmalıdır. Örneğin: “Makyajımı temizlemek iyi hissettirir” gibi kadınlar için yanlı çalışma ihtimali yüksek olan bir madde yerine “Saçımı temizlemek iyi hissettirir” şeklinde kadınlar ve erkekler için genel olarak yazılmış bir ifade kullanılması daha uygun olacaktır.
10-Yanıt stilleri veya sosyal beğenirlik gibi hataya neden olabilecek durumları değerlendirebilmek amacıyla kontrol maddelerinin yazılması faydalı görülmektedir.

Kaynak: (Hinkin, 1995; Erkuş, 2012; Carpenter, 2018; Rowan ve Wulff, 2007; Güler ve Ayan, 2020)

Tablo 2’de madde oluşturma sürecinde dikkat edilmesi gereken genel kurallar özetlenmiştir. Buna rağmen madde oluşturma sürecinin daha pek çok dikkat edilmesi gereken hususu olduğu unutulmamalıdır (DeVellis, 2012). Ölçek geliştirme çalışmalarının istatistiksel analiz aşamasında veya geliştirilen ölçeğin daha sonraki araştırmalarda kullanımları sırasında ortaya çıkan ölçüm hatalarının çoğu, ölçek maddelerinin kusurlu olmasından kaynaklanmaktadır (Hambleton, 2004; Hinkin, 1995). Güler ve Ayan (2020) inceledikleri çalışmaların bazılarında madde oluşturma süreçlerinde uyulması gereken ölçütlerin dikkate alınmadığını, olgusal söylemlerin tercih edildiğini, bazı maddelerin birden fazla farklı eylemi yokladığını, bazı maddelerin tutum objesi ile ilgili olmadığını ve madde içeriklerinde sıklık zarflarının kullanıldığını belirlemişlerdir. Şahin ve Boztunç Öztürk (2018) Türkiye’de 2010-2016 yılları arasında eğitim bilimleri alanında yayımlanan belirli akademik dergilerdeki ölçek geliştirme çalışmalarını, ölçek geliştirme süreçleri açısından incelemişlerdir. Araştırmacılar inceledikleri çalışmaların çok azında ters puanlanan maddelere dair bilgilerin olduğunu ve kontrol maddelerinin ise hiçbir makalede kullanılmadığını belirlemişlerdir. Ölçek maddelerinin oluşturulmasına ilişkin diğer önemli bir konu ise maddelerin kullanılacağı ölçek türü ve madde seçenek sayısının belirlenmesidir (Erkuş, 2012). Alan yazında uygulamadaki ve puanlamadaki pratikliği nedeniyle araştırmacılar tarafından en çok likert tipi ölçeklerin kullanıldığı görülmektedir. Ancak ölçek türünün belirlenmesinde ölçülmesi amaçlanan özelliğin yapısı doğrultusunda Thurstone, Guttman, Bogardus, Osgood gibi ölçeklerin de varlığı unutulmamalıdır (Tavşancıl, 2018). Ölçekte yer alan maddelerin seçenek sayısının belirlenmesinde hedef grubun özelliklerinin dikkate alınması gerekmektedir. Kesin kurallar olmamakla birlikte ilköğretim ikinci sınıf düzeyindeki bir grup için iki veya üç, üçüncü sınıf katılımcıları için 3,

dördüncü sınıf ve ortaokul düzeyindeki katılımcılar için dört, ortaöğretim ve yükseköğretim düzeyindeki katılımcı grubu için ise beş seçenekli maddelerin kullanılması önerilmektedir (Baykul, 2015).

“Taslak Ölçek Formu Hazırlama” aşamasında dikkat edilecek en önemli hususlardan biri de kapsam geçerliğinin sağlanmasıdır. Kapsam geçerliği, en düşük psikometrik geçerlik düzeyi olarak görülmektedir. Kapsam geçerliği, madde oluşturma sırasında ölçeğe dahil edilmektedir (Hinkin, 1995). Ölçek geliştirme çalışmalarında kapsam geçerliğinin belirlenmesi amacıyla genellikle uzman görüşüne başvurulmaktadır. Araştırmacılar uzman görüşü için ölçekte kullanılan dil konusunda, ölçek aracılığıyla ölçülmesi amaçlanan örtük özellik konusunda, ölçme ve değerlendirme alanında ve hedef kitlenin yapısı hakkında doktora derecesine sahip en az üçer bilim insanının görüşlerine başvurmalıdır. Uzman görüşleri doğrultusunda düzenlenen taslak ölçek formu, büyük gruba uygulanmadan önce pilot uygulama yapılmalıdır. Pilot uygulamada ölçek, fiili olarak uygulanmadan önce hedef kitleyi temsil eden küçük bir gruba uygulanarak yanlış anlaşılan ifadelerin düzeltilmesi ve ölçmeye karışabilecek hataların önceden belirlenip önüne geçilmesini mümkün kılmak amacıyla yapılmaktadır. Başka bir ifadeyle pilot uygulama uygun olarak ifade edilmeyen maddelerin elenmesini, tüm maddelerin en yüksek düzeyde anlaşılabilirliği için gözden geçirilmesini kolaylaştırmaktadır (Boateng vd., 2018). Pilot uygulamanın gerçekleştirildiği katılımcı grubu, hedef grubun çeşitliliğine bağlı olarak 5 ila 100 kişi arasında olmalıdır. Pilot uygulama grubu küçük ancak hedeflenen yanıtlayıcı grubuna mümkün olduğunca benzer olmalıdır (Carpenter, 2018). Pilot uygulama sonucunda gerekli tüm düzeltmeler yapıldıktan sonra “Taslak Ölçek Formu” büyük katılımcı grubuna uygulanmaya hazır hale gelmektedir.

İstatistiksel Analizlerin Uygulanması

Ölçek geliştirme çalışmalarının “İstatistiksel Analizlerin Uygulanması” aşamasında, birinci aşama sonucunda oluşturulan “Taslak Ölçek Formu” nun geçerlik ve güvenilirlik düzeylerine ilişkin istatistiksel bulgular değerlendirilmektedir. Bu aşama hangi istatistiksel yöntemlerin kullanılacağına karar verilmesi, örneklemin belirlenmesi, uygulama türüne karar verilmesi, uygulama türünden kaynaklı hata türlerinin değerlendirilmesi, uygulama yapacak kişilerin bilgilendirilerek uygulayıcıdan kaynaklı hataların önüne geçilmesi, verilerin toplanması, veri kalitesi ve analiz yöntemlerine dair varsayımların incelenmesi ile geçerlik ve güvenilirlik analizi adımlarından oluşur. Bu aşamada elde edilen veriler üzerinden yapı geçerliğine ve ölçüt geçerliğine dayalı kanıtlar değerlendirilmektedir. Geleneksel yaklaşıma göre geçerlik incelemeleri kapsam geçerliği, görünüş geçerliği, ölçüt geçerliği ve yapı geçerliği şeklinde sınıflandırılmaktadır. Ancak özellikle öz değerlendirmeye dayanan ölçeklerde, yanıtlayıcıların sosyal beğenirlik etkisiyle ölçek maddelerini yanıtlamalarını önlemek amacıyla görünüş geçerliği aranmamaktadır (Tavşancıl, 2018). Kapsam geçerliği daha çok başarı testlerinde aranan bir geçerlik türü iken ölçeklerde ön plana çıkan asıl geçerlik türü yapı geçerliğidir (Baykul, 2015). Ancak bu durum ölçek geliştirme çalışmalarında kapsam geçerliğinin önemli olmadığı anlamı taşımamalıdır. Ölçek geliştirme çalışmalarında kapsam geçerliği “Taslak Ölçek Formu Oluşturma” aşamasında sağlanmaya çalışılmalıdır. Kaya Uyanık, vd. (2017) gerçekleştirdikleri araştırmada ölçek geliştirme çalışmalarında en çok göz ardı edilen veya uygulanmayan prosedürün ölçüt geçerliği olduğunu vurgulamışlardır. Araştırmacılar bu durumu alan yazında ölçüt olarak kullanılabilir başka bir ölçeğin olmaması, ölçek kapsamının farklı olması ya da zaman ve maliyet açısından sorun olabileceği gerekçesiyle uygulanmadığı şeklinde yorumlamışlardır. Öte yandan tarihsel süreç içerisinde ölçek geliştirme çalışmalarında geleneksel geçerlik kontrollerinin günümüz için yetersiz olduğu görülmüştür. Borsboom (2006) ölçeklerde yüzeysel yapı geçerliği kanıtlarının yeterli olmadığını ölçeklerin alt gruplar arasında ölçme değişmezliği gösterip göstermediğinin incelenmesi gerektiğini vurgulamıştır. Buna ek olarak ölçeklerden elde edilen puanların karşılaştırılabilir olması amacıyla ölçek geliştirme sürecinde alt gruplar arasında değişen madde fonksiyonu ve yanlılık analizlerinin gerçekleştirilmesinin de zaruri olduğu değerlendirilmektedir (Mellenbergh, 1989).

Yapı geçerliği, ölçekte yer alan maddelerin veya ölçekten alınan puanın altta yatan örtük özellik ile ilişkisini ortaya çıkarmaya yönelik birtakım kanıtların araştırılmasıdır. Araştırmacılar yapı geçerliği için yaygın olarak açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) yöntemlerini kullanmaktadır (Hinkin, 1995). Tarihsel olarak test ve ölçek geliştirme çalışmalarında çok uzun süredir klasik test kuramına dayalı (KTK) yöntemler kullanılmaya devam etmektedir. Ancak özellikle büyük gruplar için geliştirilen ölçeklerde madde tepki kuramı (MTK) giderek KTK'nın yerini almaya başlamıştır (Embretson, 1996). Ölçek geliştirme çalışmalarında yapı geçerliğinin belirlenmesinde kullanılacak pek çok yöntem ve teknik bulunmakla birlikte araştırmacılar hangi tekniğin kullanılması konusunda farklı görüşler ileri sürmüşlerdir. Erkuş (2012) ölçek geliştirme çalışmalarında kullanılan farklı yöntemler hakkında açıklamalarda bulunarak ölçek geliştirme süreçleri için faktör analitik tekniklerin vazgeçilmez olduğunu savunmuştur. Bekmezci ve Doğan (2021) ölçek geliştirme çalışmalarında yapı geçerliğini incelemek amacıyla kullanılan AFA, karar ağacı ve yapay sinir ağları yöntemlerini aynı veri seti üzerinde uygulayarak farklı ölçek modelleri oluşturmuşlardır. Araştırmacılar, devam niteliğinde yaptıkları DFA sonucunda AFA ve yapay sinir ağları ile oluşturulan modellerin uyum indekslerinin yüksek olduğunu belirlerken; karar ağacı ile oluşturulan modelin DFA ile doğrulanmadığını tespit etmişlerdir. Öte yandan Tóth-Király vd. (2017), ölçek geliştirme çalışmalarında en çok kullanılan istatistiksel analiz tekniklerinden AFA ve DFA' nın birtakım sınırlılıklar içerdiğini ileri sürerek ölçek geliştirme çalışmaları için yeni bir yaklaşım olan keşfedici yapısal eşitlik modellemesi (KYEM) yöntemini önermişlerdir. KYEM'in, AFA ve DFA' nın ayrı ayrı sağladığı avantajlarını birleştiren bir çözüm olarak geliştirildiğini savunan araştırmacılar yürüttükleri çalışmada, KYEM'in AFA ve DFA'ya göre üstün (uygulanabilir, esnek, güçlü bir alternatif) olduğunu belirlemişlerdir. Araştırmacılar, ayrıca DFA'da maddelerin diğer faktörlerden etkilenmeden tek bir faktör altında çalıştığı varsayımı nedeniyle faktörlerdeki çapraz yüklemelerin belirli bir düzeyde sistematik hataya ve yanlış yorumlamaya neden olduğunu belirtmişlerdir (Tóth-Király vd., 2017). Erkuş (2012) ise benzer sınırlılığı, sıfır olmayan çapraz yüklemeler olduğunda aynı anda hem mevcut hem de ifade edilmemiş sınırlayıcı kısıtlamaların (yani maddelerin yalnızca bir faktörle yüklenmesi varsayımı) faktörler arasındaki ilişkiyi yükseltmesi ile açıklamıştır. Şimşek (2006) ölçek geliştirme çalışmalarında AFA ve kümeleme analizinin birbiriyle benzer sonuçlara işaret ettiğini ancak kümeleme analizinin daha yüksek duyarlılığa sahip olduğunu belirlemiştir. Farooq (2016) ölçek geliştirme çalışmalarını istatistiksel yönden incelediği çalışmada yapısal eşitlik modeli yönteminin (YEM) ölçek geliştirme çalışmalarında yaygın olarak kullanılmadığını bunun yerine daha çok AFA, diskriminant analizi ve çoklu regresyon gibi yöntemlerin yaygın olduğunu belirlemiştir. Bekmezci ve Doğan (2021) yaptıkları çalışmada ölçek geliştirme çalışmalarında yapı geçerliğini belirlemek için kullanılan farklı yöntemlerin, geçerli madde ve boyut açısından farklı yapılar oluşturabileceğini ifade etmişler ve yapı geçerliğinin belirlenmesinde araştırmacıların tek yöntem yerine birkaç yöntemi birlikte kullanmasını önermişlerdir.

Ölçek geliştirme çalışmalarında, en az hangi istatistiksel yöntemin kullanılacağıın kararlaştırılması kadar önemli olan bir diğer konu ise söz konusu yöntemin nasıl uygulandığıdır. Farooq (2016) ölçek geliştirme çalışmalarında araştırmacıların genellikle çok değişkenli normallik ve çoklu bağlantı gibi varsayımların karşılanmasına dikkat etmediklerini ifade etmiştir. Güler ve Ayan (2020) ise inceledikleri çalışmaların büyük bir bölümünde AFA ve DFA' nın aynı veri seti üzerinden gerçekleştirildiğini belirlemişlerdir. Ancak alan yazında ölçülen yapının ortaya çıkarılmasına ilişkin önce AFA daha sonra ise belirlenen modelin uyumunu değerlendirmek için DFA yapılması gerektiği yer almaktadır (Çokluk, Şekercioğlu ve Büyüköztürk, 2012).

Alan yazında ölçek geliştirme çalışmalarında, iç tutarlılığın belirlenmesinde en çok kullanılan yöntemin Cronbach Alpha katsayısı olduğu görülmektedir. Şahin ve Boztunç Öztürk (2018) inceledikleri tüm ölçek geliştirme çalışmalarında, iç tutarlılığın yorumlanmasında Cronbach Alpha katsayısının kullanıldığını raporlamışlardır. Acar Güvendir ve Özer Özkan (2015) da farklı zaman

aralıklarındaki ölçek geliştirme çalışmalarını inceledikleri araştırmada, tüm araştırmacıların geliştirdikleri ölçeğin iç tutarlılığı belirlemek amacıyla Cronbach Alpha katsayısını hesaplandıklarını raporlamışlardır. Ayrıca araştırmacılar inceledikleri makalelerin çok azında eşdeğer yarılar güvenilirliğinin kullanıldığını belirtmişlerdir. Ölçek geliştirme çalışmalarında iç tutarlığın belirlenmesi amacıyla yaygın olarak Cronbach Alpha katsayısı tercih edilmesine rağmen yakın tarihte gerçekleştirilen bazı çalışmalar bu yöntemin birtakım sınırlılıklar içerdiğini ortaya çıkarmıştır. Bu nedenle alan yazında iç tutarlığın belirlenmesinde Cronbach Alpha katsayısının tek başına kullanılması yerine Omega katsayısı ile birlikte kullanılması önerilmektedir (Aydoğan, Kıbrıslıoğlu Uysal ve Doğan, 2019; Soysal, 2023).

Ölçek Çalışmasının Raporlanması

Ölçek geliştirme çalışmalarının son aşaması olan “Ölçek Çalışmasının Raporlanması” ında araştırmacılar “Nihai Ölçek Formu” nu hazırlayarak ölçek geliştirme sürecinde geliştirdikleri bütün işlemleri şeffaf bir şekilde raporlamalıdır. Ayrıca araştırmacılar, geliştirdikleri ölçeğin başka araştırmacılar tarafından nasıl kullanılacağına dair bir kılavuz kitapçık hazırlamalı veya geçerli bütün açıklamaları raporlamalıdır. Acar Güvendir ve Özer Özkan (2015) inceledikleri ölçek geliştirme çalışmalarının çok azında geliştirilen ölçeğin uygulama adımlarına ilişkin yönerge hazırlandığını raporlamışlardır. Mor Dirlik (2014) incelediği ölçek geliştirme konulu beş tezin hiçbirinde ölçek puanlarının nasıl yorumlanması gerektiği ile ilgili ayrıntılı bilgiler verilmediğini raporlamıştır. Ölçek geliştirme çalışmalarında analizler sırasında hangi maddelerin ölçekten çıkarıldığı, nihai ölçekte yer alan maddelerin hangi alt boyuta ait olduğu, hangi maddelerin ters puanlandığı, hangi maddelerin kontrol maddesi olarak kullanıldığı, ölçekten alınan en yüksek ve en düşük puanların ne olduğuna ilişkin bilgiler sonuç raporunda mutlaka sunulmalıdır. Ayrıca ölçek geliştirme çalışmalarında, ölçekten elde edilen puanların ne anlama geldiği ve kesme puanlarının nasıl belirlendiğine ilişkin açıklamaların yer alması gerekmektedir.

Araştırma ve Yayın Etiği

Bu çalışmada, Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesinde belirtilen tüm kurallara uyulmuştur. Yönergede *Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler* başlığı altında açıklanan eylemlerden hiçbiri gerçekleştirilmemiştir.

Etik Kurul İzni

Bu çalışma bir derleme çalışması olduğundan etik kurul izni gerektirmemektedir.

Yazarların Katkı Oranı

Bu çalışmada yazarlar aynı oranda katkı sağlamışlardır (1. yazar %50, 2. yazar %50).

Çıkar Çatışması

Çalışmanın yazarları arasında hiçbir çıkar çatışması yoktur.

Kaynakça

- Acar Güvendir, M. ve Özer Özkan, Y. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Electronic Journal of Social Sciences*, 14(52), 23-33. doi: 10.17755/esosder.54872
- AERA, APA, NCME (1999). *Standards for educational and psychological testing*. Washington DC: American Psychological Association.

- Aydoğan, S., Kıbrıslıoğlu Uysal, N. ve Doğan, N. (2019). Cronbach a ve Mcdonalds Ω değerlerinin gerçek verilerde karşılaştırılması. *The Journal of Academic Social Science*, 62(62), 460-468.
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması* (3. baskı). Pegem Akademi.
- Bekmezci, S. ve Doğan, N. (2021). Ölçek geliştirmede açmılayıcı faktör analizi, karar ağacı ve yapay sinir ağları sonuçlarının karşılaştırılması. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi*, 19(4), 135-154.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health*, 6, 149. doi: 10.3389/fpubh.2018.00149
- Borsboom, D. (2006). When does measurement invariance matter?. *Medical care*, 44(11), 176-181.
- Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication methods and measures*, 12(1), 25-44. doi: 10.1080/19312458.2017.1396583
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları* (2. Basım). Pegem Akademi.
- Çüm, S. ve Koç, N. (2013). Türkiye’de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Eğitim Bilimleri ve Uygulama*, 12(24), 115-135.
- DeVellis, R. F. (2012). *Scale development theory and applications* (Third Edition). Sage Publications.
- Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment*, 8(4), 341.
- Erkuş, A. (2007). Ölçek geliştirme ve uyarlama çalışmalarında karşılaşılan sorunlar. *Türk Psikoloji Bülteni*, 13(40), 17-25.
- Erkuş, A. (2012). Varolan ölçek geliştirme yöntemleri ve ölçme kuramları psikolojik ölçek geliştirmede ne kadar işlevsel: Yeni bir öneri. *Journal of Measurement and Evaluation in Education and Psychology*, 3(2), 279-290.
- Erkuş, A. (2014). *Psikolojide ölçme ve ölçek geliştirme-I temel kavramlar ve işlemler* (2. baskı). Pegem Akademi.
- Güler, G., & Ayan, C. (2020). Review of Attitude Scales Developed in Turkey Between 2002-2018 Regarding the Scale Development Process. *Journal of Faculty of Educational Sciences*, 53(3), 839-863. doi: 10.30964/auebfd.658488
- Hambleton, R. K. (2004). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K., Hambleton, P. F., Merenda & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment*, (pp. 15-50). Psychology Press.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of management*, 21(5), 967-988.
- Karakoç, F. Y. ve Dönmez, L. (2014). Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitimi Dünyası*, 13(40), 39-49. doi: 10.25282/ted.228738
- Kaya Uyanık, G. K., Güler, N., Taşdelen Teker, G., ve Demir, S. (2017). Türkiye’de eğitim alanında yayımlanan ölçek geliştirme çalışmalarının uygunluğunun çok yüzeyli Rasch modeli ile incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 8(2), 183-199.
- Li, S., & Wang, H. (2018). Traditional literature review and research synthesis. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave handbook of applied linguistics research methodology* (pp. 123-144). Palgrave Macmillan. https://doi.org/10.1057/978-1-137-59900-1_6
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International journal of educational research*, 13(2), 127-143.

- Mor Dirlik, E. (2014). Ölçek geliştirme konulu doktora tezlerinin test ve ölçek geliştirme standartlarına uygunluğunun incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(2), 62-78. doi: 10.21031/epod.63138
- Morgado, F. F., Meireles, J. F., Neves, C. M., Amaral, A., & Ferreira, M. E. (2017). Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*, 30(3), 1-20. doi: 10.1186/s41155-016-0057-1
- Rowan, N., & Wulff, D. (2007). Using qualitative methods to inform scale development. *Qualitative Report*, 12(3), 450-466.
- Simms, L. J. (2008). Classical and modern methods of psychological scale construction. *Social and Personality Psychology Compass*, 2(1), 414-433. doi: 10.1111/j.1751-9004.2007.00044.x
- Soysal, S. (2023). Çok boyutlu test yapılarında alfa, tabakalı alfa ve omega güvenirlik katsayılarının karşılaştırılması. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, 5(1), 213-236.
- Şahin, M. G. ve Öztürk, N. B. (2018). Eğitim alanında ölçek geliştirme süreci: Bir içerik analizi çalışması. *Kastamonu Eğitim Dergisi*, 26(1), 191-199. doi: 10.24106/kefdergi.375863
- Şimşek, D. (2006). *Kümeleme analizi, çok boyutlu ölçekleme, doğrulayıcı ve açıklayıcı faktör analizi ile elde edilen yapı geçerliği kanutlarının karşılaştırılması*. (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Tavşancıl, E. (2018). *Tutumların ölçülmesi ve SPSS ile veri analizi* (6. baskı). Nobel Akademi Yayıncılık.
- Tavşancıl, E., Güler, G. ve Ayan, C. (2014). *2002-2012 yılları arasında Türkiye'de geliştirilen bazı tutum ölçeği geliştirme çalışmalarının ölçek geliştirme süreci açısından incelenmesi*. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi (Uluslararası Katılımlı) 9-13 Haziran, Hacettepe Üniversitesi, Ankara.

Extended Abstract

Introduction

In the 21st century, scales are used in various fields such as work, education, health, and military to determine individuals' affective characteristics. These scales assess unobservable traits such as anxiety, motivation, perception, attitude, and intelligence, aiding in decision-making processes. These tools are crucial for generating and testing scientific knowledge in fields like education, psychology, and sociology. The progression of science through accurate and reliable information is directly linked to the validity and reliability of the measurement tools used (Karakoç & Dönmez, 2014).

Scales aim to measure latent characteristics that cannot be directly observed through concrete statements (Carpenter, 2018). In fields like education and psychology, scales are used to present items related to latent traits, and the responses to these items provide insights into the structure and level of these traits (Erkuş, 2012). The quality of scales is critical since they measure latent characteristics that cannot be directly observed (Kaya Uyanık et al., 2017). Furthermore, after the development of scales, their use in new research must be monitored, and their psychometric properties must be evaluated (Mor Dirlik, 2014).

In recent years, there has been an increase in scale development studies in Turkish culture. Due to the pressure on academics to publish, many researchers, both related and unrelated to their fields, perceive scale development as the "easiest" type of publication (Erkuş, 2007). Despite the increase in the number of scales developed nationally, qualitative errors and misconceptions continue to be observed in the content of these scales (Mor Dirlik, 2014). The common belief that scale development is simple, practical, and easy to analyze can be misleading. The process of developing scales, contrary

to common belief, requires a high level of theoretical, methodological, and statistical competence (Simms, 2008; Carpenter, 2018).

Conclusion and Discussion

Scale development studies need to be based on a solid foundation. Researchers should examine whether there is an existing scale in the literature that measures the construct they intend to measure before starting the scale development process. If a valid scale exists in a different culture, the option of adapting that scale should also be considered (Karakoç & Dönmez, 2014). Tavşancıl and colleagues (2014) identified that different researchers have developed various scales to measure the same latent trait in Turkish culture. This situation does not contribute uniquely to the literature and can lead to a waste of time. Therefore, researchers should start scale development studies only after ensuring that there is no valid scale in the literature that measures the intended trait (Kaya Uyanık et al., 2017).

Researchers conducting scale development studies must have a high level of expertise in both the latent trait being measured and the steps of measurement and evaluation. Otherwise, scales developed by non-expert researchers can harm science (Erkuş, 2014). Scale development procedures that are not carefully applied can be theoretically or technically flawed, leading to incorrect findings or decision errors (Çüm & Koç, 2013). Generally, doctoral programs do not offer courses related to scale development processes (Boateng et al., 2018). Therefore, many researchers implement scale development processes by imitating other studies in the literature (Carpenter, 2018).

In the literature, there are studies that examine scale development studies published in different periods. Çüm and Koç (2013) evaluated scale development and adaptation studies published in certain psychology and educational sciences journals in Turkey between 2005-2013. They found that 37% of the scale development steps and 60% of the scale development principles were not properly followed. Tavşancıl and colleagues (2014) identified many incorrect and incomplete attitude scales in their review of attitude scale development studies published in educational and psychology journals in Turkey. Mor Dirlik (2014) evaluated five doctoral theses on scale development in the field of educational sciences and found that only two of the theses correctly applied the scale development procedures at a high level.

In this study, the scale development process is explained in three sequential stages based on the literature: "Preparation of the Draft Scale Form," "Application of Statistical Analyses," and "Reporting the Applied Scale Study." These stages aim to provide clear, understandable, and straightforward guidance to researchers. There are many academic publications in the literature that address scale development processes (AERA, APA & NCME, 1999; Erkuş, 2012; Erkuş, 2014; DeVellis, 2012; Simms, 2008; Tavşancıl, 2018). However, the existence of many different sources related to scale development processes prevents the adoption of a common approach. Carpenter (2018) defined ten steps for scale development, while Simms (2008) explained the process in three sections. DeVellis (2012) described the scale development processes in eight steps.