

Likert Tipi Ölçekleri Kullanma, Modifiye Etme, Uyarlama ve Geliştirme Süreçleri

Processes of Using, Modifying, Adapting and Developing Likert Type Scales

Fulya GÖKDEMİR¹
Tülay YILMAZ²

¹Artvin Çoruh Üniversitesi, Sağlık Bilimleri Fakültesi, Hemşirelik Bölümü, Doğum Kadın Sağlığı ve Hastalıkları Hemşireliği Anabilim Dalı, Artvin, Türkiye

²İstanbul Üniversitesi-Cerrahpaşa, Sağlık Bilimleri Fakültesi, Ebelik Bölümü, İstanbul, Türkiye

ÖZ

Bilimsel araştırmalarda, etkili ve yenilikçi verilerin üretilmesi önemlidir. Bu amaçla gözlem, görüşme, deney, biyofiziksel ölçümler ve öz bildirime dayalı teknikler kullanılabilir. Ölçekler, öz bildirime dayalı veri toplama araçları olup araştırma verilerinin daha hızlı, kolay, ucuz ve standardize bir şekilde toplanmasını sağlar. Bu nedenle ölçek kullanma, modifiye etme, uyarlama ve geliştirme, insanı temel alan tüm bilim dallarında kritik bir rol oynar. Likert tipi ölçekler, bazı avantajları nedeniyle diğer ölçek türlerine kıyasla daha fazla tercih edilir. Konuyu bütüncül bir yaklaşımla ele alan kullanıcı dostu kaynakların oluşturulması ve araştırmacılar tarafından kullanımının yaygınlaştırılması ise bu konudaki farkındalığın artırılmasına, daha kaliteli ölçeklerin üretilmesine ve kullanılmasına katkıda bulunur. Bu nedenle, sunulan derlemede Likert türü ölçeklerin kullanılması, modifikasyonu, uyarlanması ve geliştirilmesi süreçleri bütüncül bir yaklaşımla ele alınmıştır. Makalenin bu konudaki farkındalığın artırılmasına, daha kaliteli ölçeklerin üretilmesine ve kullanılmasına katkıda bulunması amaçlanmıştır.

Anahtar Kelimeler: Geçerlik, güvenilirlik, ölçek

ABSTRACT

It is important to produce effective and innovative data in scientific research. For this aim, techniques based on observation, interview, experiment, biophysical measurements, and self-report can be used. The scales are self-report data collection tools and enable the faster, easier, cheaper, and standardized collection of research data. Therefore, using, modifying, adapting, and developing scales play a critical role in all human-based sciences. Likert-type scales are preferred more than other scale types due to some advantages. The creation of user-friendly resources that deal with the subject with a holistic approach and the dissemination of their use by researchers contribute to raising awareness on this issue and to the production and use of better quality scales. For this reason, the use, modification, adaptation, and development processes of Likert-type scales are discussed with a holistic approach in the present review. The aim of the article is to contribute to raising awareness on this issue, regarding the production and use of higher-quality scales.

Keywords: Reliability, scale, validity

Geliş Tarihi/Received: 25.10.2022

Kabul Tarihi/Accepted: 08.04.2023

Yayın Tarihi/Publication Date: 18.05.2023

Sorumlu Yazar/Corresponding Author:

Fulya GÖKDEMİR

E-mail: fulya08_tr@hotmail.com;

fulya08_tr@artvin.edu.tr

Cite this article as: Gökdemir F, Yılmaz T. Processes of using, modifying, adapting and developing likert type scales. *J Nursology* 2023;26(2):148-160.



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

GİRİŞ

Ölçekler; bilgi, duygu, ilgi, algı, tutum, inanç, eğilim, risk, yaşam kalitesi, davranış gibi özellikleri ölçebilen veri toplama araçlarıdır.^{1,2} Psikometrik açıdan kaliteli ölçekler ve etkili ölçümler elde edilen veriler, bilimsel araştırmaların yapı taşını oluşturur.³ Bu nedenle ölçek kullanma, modifiye etme, uyarlama ve geliştirme, insanı temel alan tüm bilim dallarında kritik bir önem taşır.²⁻⁵ İlk olarak, 1932 yılında Psikolog Rensis Likert (1903-1981) tarafından geliştirilen Likert tipi ölçekler, diğer ölçek türlerine (Bogardus, Thurstone, Gutmann, Osgood vb.) kıyasla daha sık kullanılır.⁶⁻⁸ Bu ölçeklerde, aynı yapıyı ölçtüğü düşünülen birbiriyle ilişkili en az üç soru bulunur.⁵ Ölçülen özelliğe ilişkin toplam bir puanın elde edilmesinin hedeflenmesi nedeniyle "Dereceleme toplamlarıyla ölçekleme" tekniği olarak da bilinir.^{8,9} Likert tipi ölçekleri, diğer ölçeklerden ayıran en önemli özellikler; toplam puanlarının olması, her bir alt boyutun (faktör) toplam puan içerisindeki büyüklüğünün hesaplanabilmesi, standardize ölçümler yapılabilmesi, tekrarlı ölçümlerde kullanılabilmesi, geçerlik ve güvenilirlik analizlerin yapılabilmesidir.^{1,8}

Likert tipi ölçeklerle veri toplama söz konusu olduğunda araştırmacılar; mevcut bir ölçeği kullanmayı, eski bir ölçeği modifiye etmeyi, farklı bir dil ve kültür için hazırlanmış ölçekleri kendi kültürlerine uyarlamayı ya da yeni bir ölçek geliştirmeyi tercih edebilir.^{1,10,11} Kaliteli veriler; ölçüm amacına uygun, geçerli, güvenilir, duyarlı, tutarlı, kararlı, hatalardan arınık, kullanışlı, objektif ve homojen ölçeklerden elde edilebilir.^{6,10} Bu nedenle ölçek kullanma, modifiye etme, uyarlama ve geliştirmeye yönelik çalışmaların titizlikle yürütülmesi gerekir.¹ Bu alanlardaki bilimsel çalışmaların sayısı hızla artmaktadır. Fakat kalitelerinin aynı düzeyde artmadığı ve bazı metodolojik hataların yapıldığı bilinmektedir.^{2,11,12} Kishore ve ark² bu durumun en önemli nedenleri olarak; kullanıcı dostu kaynaklara ve rehberlere erişilememesini, diğer araştırmaların planlanması, değerlendirilmesi, raporlanması için önemli kılavuzlar olmasına rağmen ölçek geliştirmeyi ele alan herhangi bir kılavuz ya da rehberin olmamasını, lisans ve lisansüstü eğitimlerde bu konulara yeterince yer verilmemesini göstermiştir. Güvendir ve Özkan¹³ ise Türkiye kaynaklı ölçek uyarlama ve geliştirme makalelerinde tam bir yöntem birliğinin olmadığını saptamış, ölçek geliştirme aşamalarını gösteren kılavuzlara, hakemlerin standardize değerlendirmeler yapabileceği kontrol listelerine, ölçeklerin koordinasyonu ve denetimini sağlayacak merkezlere ihtiyaç duyulduğunu bildirmiştir.

Ölçek kullanma, modifiye etme ve kısa form geliştirmeye yönelik literatür kaynaklarının sayısı sınırlıdır.^{1,10,14} Ölçek geliştirmeyi ele alan kaynaklar ise nispeten daha fazladır. Fakat bu çalışmaları standardize edecek rehberler yoktur.^{1-10,15-18} Ölçek uyarlama konusunda çeşitli kaynaklar ve Uluslararası Test Komisyonu (International Test Commission-ITC) gibi kuruluşlar tarafından hazırlanmış rehberler vardır.^{11,19-21} Fakat bu rehberlerin kullanım oranı düşüktür.²⁰ Bu durum, nitelikli bilimsel çalışmaların zayıf raporlanmasına ve kullanımlarının azalmasına neden olmaktadır.²² Ayrıca mevcut kaynaklarda ölçek kullanma, modifiye etme, uyarlama ve geliştirme süreçleri ayrı ayrı ele alınmış, ortak yönlerine ve farklılıklarına ise çok az değinilmiştir.^{1,10,15} Bu derlemenin amacı; likert tipi ölçeklerin kullanılması, modifiye edilmesi, farklı bir kültüre uyarlanması ve geliştirilmesi konularını bütüncül bir yaklaşımla ele almaktır. Böylece likert tipi ölçeklerle ilgili bilgi, farkındalık, eleştirel düşünme, psikometrik açıdan daha kaliteli ölçeklerin üretilmesi ve kullanılması gibi konulara olumlu katkı sağlayacağı düşünülmektedir.

Likert Tipi Ölçeklerin Kullanımı ve Modifikasyonu

Geçerli ve güvenilir olduğu daha önceden kanıtlanmış olan Likert tipi bir ölçeğin sorularında, hedef kitlesinde, dilinde ve dereceleme şeklinde herhangi bir değişiklik yapılmadan aynen uygulanmasına ölçek kullanma denir.^{1,10} Modifikasyon ise ulusal düzeyde kullanılan bir ölçeğin araştırılacak özellik ya da fenomeni ölçmede yetersiz kalması, ölçeğin kısa bir formunun geliştirilmek istenmesi, mevcut hedef kitlesinde (farklı bir yaş, cinsiyet, eğitim, meslek, hasta grubu vb.) ya da puanlamasında değişiklik gerekmesi gibi durumlarda uygulanan revizyon işlemidir.^{1,14} Modifikasyonda ifadeler aynı kalmakta fakat soru sayısında azaltma (kısa form geliştirme), ölçek maddelerinde nesne değişiklikleri, farklı hedef gruba uyarlama gibi işlemler yapılmaktadır.^{1,23} Kısa form geliştirme; kullanım kolaylığı sağlanması, zaman kazandırması, uygulayıcılar ve yanıtlayıcılar üzerindeki baskıları azaltması nedeniyle sağlık bilimleri alanında sıkça tercih edilmektedir.¹⁴ Bir ölçeğin ne kadar zaman sonra modifiye edilmesi gerektiği konusunda standardize süreler yoktur. Araştırmacılar, literatürdeki yenilik ve eleştirileri takip ederek ideal olan süreyi kendileri belirlemelidir.¹

Araştırmalarda kullanılacak ya da modifiye edilecek ölçeklere bilimsel dergiler, tez merkezi, internet siteleri ve ölçeği geliştirilenlerin kişisel blog sayfaları aracılığı ile ulaşılabilir.^{1,23} Psikometrik açıdan yetersiz ölçeklerin kullanılması, bireyin yaşamını olumsuz etkileyebilecek bazı kararların alınmasına neden olabilir.^{10,16,22} Bu nedenle herhangi bir ölçeği kullanmaya karar vermeden önce o ölçeğe ilişkin tüm araştırmaların incelenmesi gerekir.^{10,19} İnceleme sürecinde ölçeğin kullanım amacı, hedef kitlesi, güncelliği, psikometrik özellikleri, uygulama, puanlama, değerlendirme ilkeleri ve ölçek uygulayıcıların nitelikleri belirlenir.^{10,15,16} Literatürde aynı odak değişkeni ölçen birden fazla ölçek varsa araştırmacının amacına ve örneklem grubuna en uygun olan, psikometrik açıdan daha üstün olan, araştırma ekibi tarafından uygulanabilecek en güncel ölçeğin seçilmesi gerekir.^{1,10,16} Çocuklar için üçlü ve yetişkinler için beşli likert biçiminde derecelenmiş ölçekler ideal olarak kabul edilir.¹⁰ Dereceleme kısmında "kararsızım" ifadesi olan ölçeklerin kullanılması ise önerilmez. Çünkü bu ölçekler belirsizliği ve yanlılığı yansıtır, randomizasyonu ve normal dağılımı bozar.¹ Araştırmalar planlanırken, sonuçları raporlanırken ve ölçeklerle ilgili çalışmalara hakemlik yapılırken Sağlık Enstrümanlarının Seçimi için Fikir Birliğine Dayalı Standartların (COnsensus-based Standards for the selection of health Measurement INstrument-COSMIN) kullanılması ise daha kaliteli çalışmaların yapılmasına katkıda bulunur.²² Ölçeklerle ilgili çalışmaların metodolojik kalitesinin artırılması için aşağıdaki noktalara dikkat edilmesi gerekir.^{1,2,10,15,23}

- Cronbach alfa değeri 0,50'den az ve psikometrik açıdan yetersiz olan ölçekler, araştırmalarda kullanılmamalıdır. Bu ölçekler yeniden düzenlenmeli veya modifiye edilmelidir.
- Örneklem sayısı, güç analizi yapılarak belirlenmelidir.
- Tüm yasal izinler (ölçek kullanım izni, etik kurul izni, kurum çalışma izni, gönüllü onamı vb.) alınmadan, ölçekler uygulanmamalıdır.
- Katılımcılardan onam alınırken; ölçeğin kullanım amacı, araştırmaya katılım koşulları, gönüllülük ilkesi, çekilme hakkı, verilerin gizliliğinin nasıl korunacağı, çalışmanın yararları, riskleri, araştırmacının iletişim bilgileri hakkında bilgi verilmelidir.
- Ölçekler, benzer koşullar altında, önerilen sürelerde, bu konuda eğitim, sertifika, diploma gibi ön koşulları taşıyan kişiler tarafından uygulanmalı, puanlanmalı ve yorumlanmalıdır.
- Tartışmalı ve hassas konulara yönelik soruları içeren ölçekler uygulanacaksa, Etik Kurul izni alınırken bu durum belirtmeli ve veriler toplanırken kurum politikalarına uyulmalıdır.
- Elde edilen veriler; yasal düzenlemeler doğrultusunda saklanmalı, değerlendirme sonucu sadece bireyin kendisi, yasal vasisi ve yazılı izin verdiği diğer kişilerle paylaşılmalıdır.
- Ölçekten sağlanan bilgilerle bireyler bir mesleğe yönlendirilmemeli ve sadece ölçekten elde edilen puanlarla bireyler hakkında karar verilmemelidir. Diğer kaynaklardan (öykü, muayene, laboratuvar tetkikleri vb.) sağlanan bilgi ve bulgular da incelenmelidir.
- Verilerin eşleştirilmesi gerekiyorsa (test tekrar test, ön-test, son-test vb.) başlangıç aşamasında standart bir kodlama yapılarak olası karışıklıklar önlenmelidir.
- Geçerlik ve güvenilirlik analizlerinin üzerinden bir yıl geçmiş ölçekleri kullanan araştırmacılar, bu analizleri kendi örneklemelerinde yeniden gözden geçirmelidir.
- Ölçek, farklı bir gruba uygulanacak ya da modifikasyon yapılacaksa ölçek geliştiricilerden bu işlem için ayrıca izin alınmalı ve Tablo 1'de sıralanan hususlara dikkat edilmelidir.

Tablo 1. Likert Tipi Ölçeklerin Modifikasyonu ve Kısa Formlarının Geliştirilmesi

Modifikasyonda dikkat edilmesi gereken noktalar ¹	Kısa form geliştirilirken dikkat edilmesi gereken noktalar ¹⁴
<ul style="list-style-type: none"> Kapsamlı bir literatür taraması yapılarak orijinal ölçeğin psikometrik özellikleri, farklı gruplara uygulanma durumu, ulusal ve uluslararası araştırmalardaki kullanım sıklığı belirlenmelidir. Ölçeğin daha önce modifiye edilip edilmediği ve uzmanların modifikasyon konusundaki görüşleri incelenmelidir. Modifikasyon kararı verildiğinde ölçek kullanım izni ve diğer yasal izinler alınmalı, maddeler revize edilerek uzmanlardan görüş alınmalı ve ölçek geliştirme adımları uygulanmalıdır. Modifikasyon yapılırken, orijinal ölçek örneklem büyüklüğünden daha büyük olan bir örneklem seçilmelidir. Modifiye edilen ölçek için elde edilen geçerlik ve güvenilirlik bulguları, ölçeğin eski versiyonu ile kıyaslanmalıdır. Modifiye edilen ölçek, ölçeğin eski versiyonu ile yapısal açıdan benzer olduğunda ölçeğin eski versiyonunun kullanımına devam edilmelidir. Eski ve yeni versiyon arasında psikometrik açıdan farklılık varsa yeni bir örneklem (en az 150 kişi) veri toplanmalı, açıklayıcı ve doğrulayıcı faktör analizleri tekrarlanmalı, diğer güvenilirlik analizleri de yapılarak bulgular raporlanmalı ve çalışma sonuçları yayınlanmalıdır. Bu süreçte ölçeğin orijinal adı korunmalı, etik kurallara ve ölçeği geliştiren yazarların yazarlık haklarına saygı gösterilmelidir. Yeni versiyonu yapılan bir ölçeğin, eski versiyonunun/versiyonlarının kullanılmaması için tüm önlemler alınmalıdır. Modifiye edilen yeni ölçeğin kullanıldığı bilimsel yayınlar ve diğer araştırmacılar tarafından yapılan geri bildirimler takip edilmelidir. Modifikasyon işlemi yetersiz olduğunda yeni bir ölçeğin geliştirilmesi tercih edilmelidir. 	<ul style="list-style-type: none"> Ölçeğin uzun formu incelenmeli ve psikometrik açıdan kaliteli olmadığı belirlenirse, kısa form geliştirme işlemi yapılmamalıdır. Kısa form için seçilen maddelerin faktör yükü yüksek olmalıdır. Bu aşamada istatistiksel analizlerden ve uzman görüşlerinden de yararlanılmalıdır. Kısa forma alınan maddeler, uzun forma ait davranış evrenini yansıtabilmeli, ölçeğin geçerlik ve güvenilirliğini düşürmemelidir. Kısa formun psikometrik özellikleri, duyarlılık ve seçiciliği uzun formdan bağımsız olarak kanıtlanmalıdır. Aynı veri seti ile hem kısa hem de uzun form elde edilmemelidir. Kısa formun geçerliği belirlenirken yapı geçerliği sonuçları esas alınmalıdır. Kısa formun güvenilirliği, birden fazla yöntemle (iç tutarlılık analizleri, test tekrar test vb.) incelenmelidir. Uzun ve kısa formun psikometrik özellikleri karşılaştırılmalı, aynı ya da benzer ise yeni ölçek değiştirilebilir kısa form olarak kabul edilmeli, eşit alternatif form elde edilinceye kadar denemeye devam edilmelidir. Kısa formun, zaman tasarrufu sağladığı ve psikometrik özellikleri olumsuz etkilemediği ortaya konulmalı, ne kadar zaman tasarrufu sağladığı belirtilmelidir. Çok boyutlu ölçeklerin, her bir faktörünün psikometrik özellikleri ayrı ayrı incelenmeli, bir alt boyutta en az üç ve mümkünse dört madde bulunmasına dikkat edilmelidir. Analizler, herhangi bir alt boyutun tamamen atılması gerektiğini gösteriyorsa o faktöre ait kuramsal ve istatistiksel özellikler ayrıntılı olarak açıklanmalıdır. Kültürlerarası ölçek uyarlaması yapılacaksa öncelikle ölçeğin uzun formu, daha sonraki süreçte ise kısa formu uyarlanmalıdır.

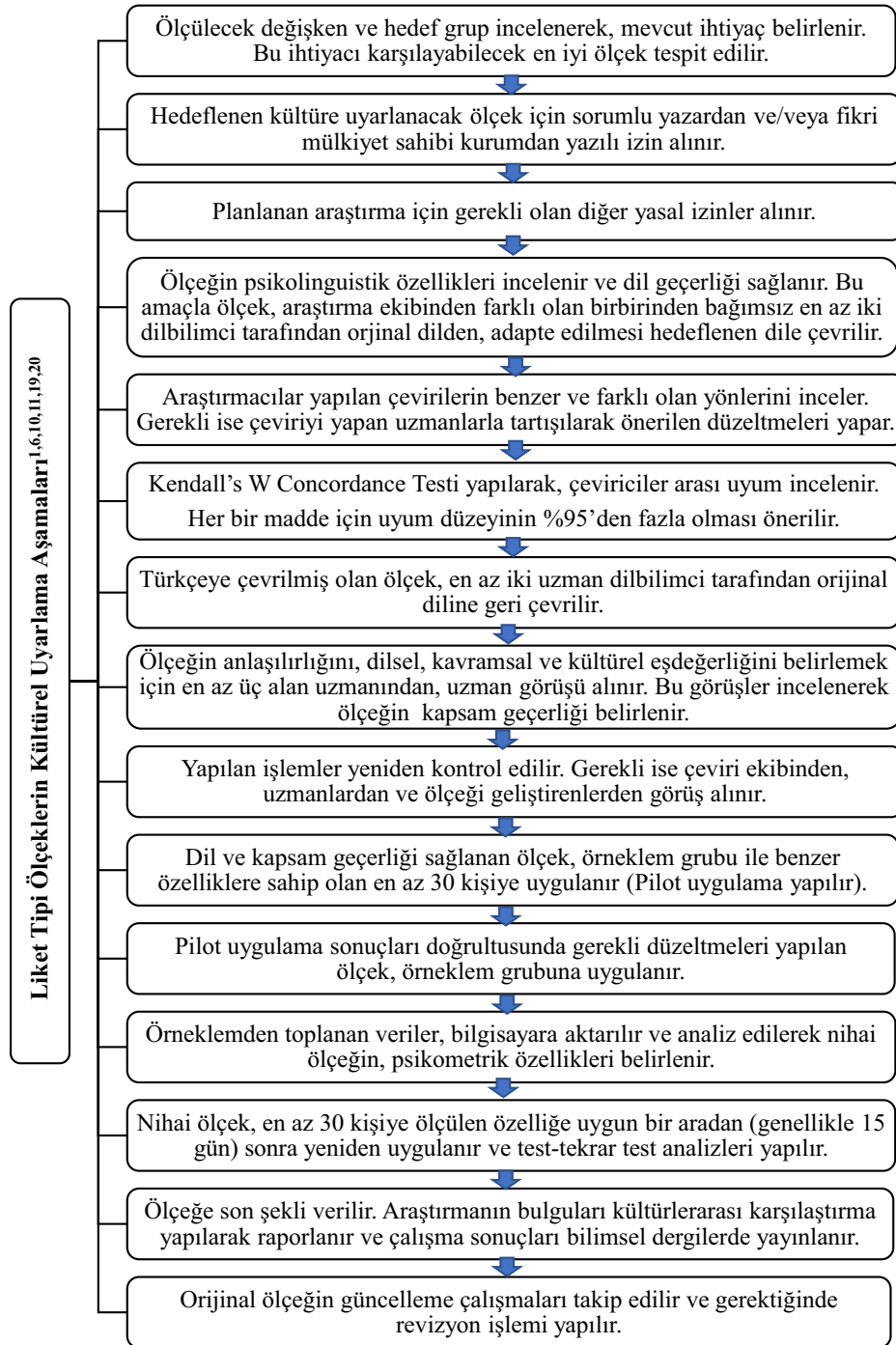
Likert Tipi Bir Ölçeğin Uyarlanması (Adaptasyonu)

Farklı bir kültür ve dilde geliştirilen bir ölçeğin, başka bir kültür ve dilde kullanılabilmesi için birtakım işlemlerden geçirilerek hazır hale getirilmesidir.¹¹ Bu çalışmaların temel amacı ölçeğin, özgün içeriğini ve kültürel eşdeğerliğini koruyarak hedef kültüre uyarlanmasını sağlamaktır.²⁰ Madde havuzu yazılmasına gerek olmaması, zaman tasarrufu sağlaması, ölçeğin psikometrik özelliklerinin daha önceden incelenmiş olması gibi avantajları nedeniyle sıkça kullanılmaktadır.^{16,20} Çok kültürlü ve çok merkezli çalışmaların yapılmasına, araştırmacılar arasındaki işbirliğinin artırılmasına, ölçeğin uygulandığı toplumlar arasındaki benzerliklerin ve farklılıkların ortaya çıkarılmasına katkıda bulunmaktadır.^{19,20} Bu nedenle kültürlerarası karşılaştırma yapan ölçekler elde etmek isteyen araştırmacıların, sık kullanılan ve bilim dünyası tarafından kabul gören ölçekleri kendi kültürüne uyarlaması, böyle bir amacı olmayanların ise dilsel ve kültürel problemleri en aza indirmek için yeni bir ölçek geliştirmesi tavsiye edilmektedir.¹⁶ Likert tipi ölçeklerin uyarlanması Şekil 1'deki ve geliştirilmesi ise Şekil 2'deki aşamalar takip edilerek gerçekleştirilmelidir.

Ölçek uyarlama sürecine başlamadan önce o ölçeğin hangi amaçla ve nasıl geliştirildiği, kullanım sıklığı, örnekleme, maddeleri, psikometrik özellikleri, ölçeğe yapılan eleştiriler, diğer ölçeklerden üstün ve zayıf olan yönleri incelemelidir.^{19,22} Ölçek uyarlamasının gerekliliği ve kullanılacak en iyi ölçek olup olmadığı belirlenmeli, uyarlamaya yönelik tüm işlemler (Şekil 1) titizlikle yürütülmelidir.^{16,19} Cinsel tercihler, dini ve siyasi yönelimler, toplumsal yapı ve kültürel tabu gibi özellikleri ölçek ölçeklerde çok daha hassas

davranılmalıdır.² Araştırmacının metodolojik kalitesini ve ölçek puanlarının kültürlerarası karşılaştırılabilirliğini arttırmak için ITC ve COSMIN rehberleri kullanılmalıdır.^{11,20,22}

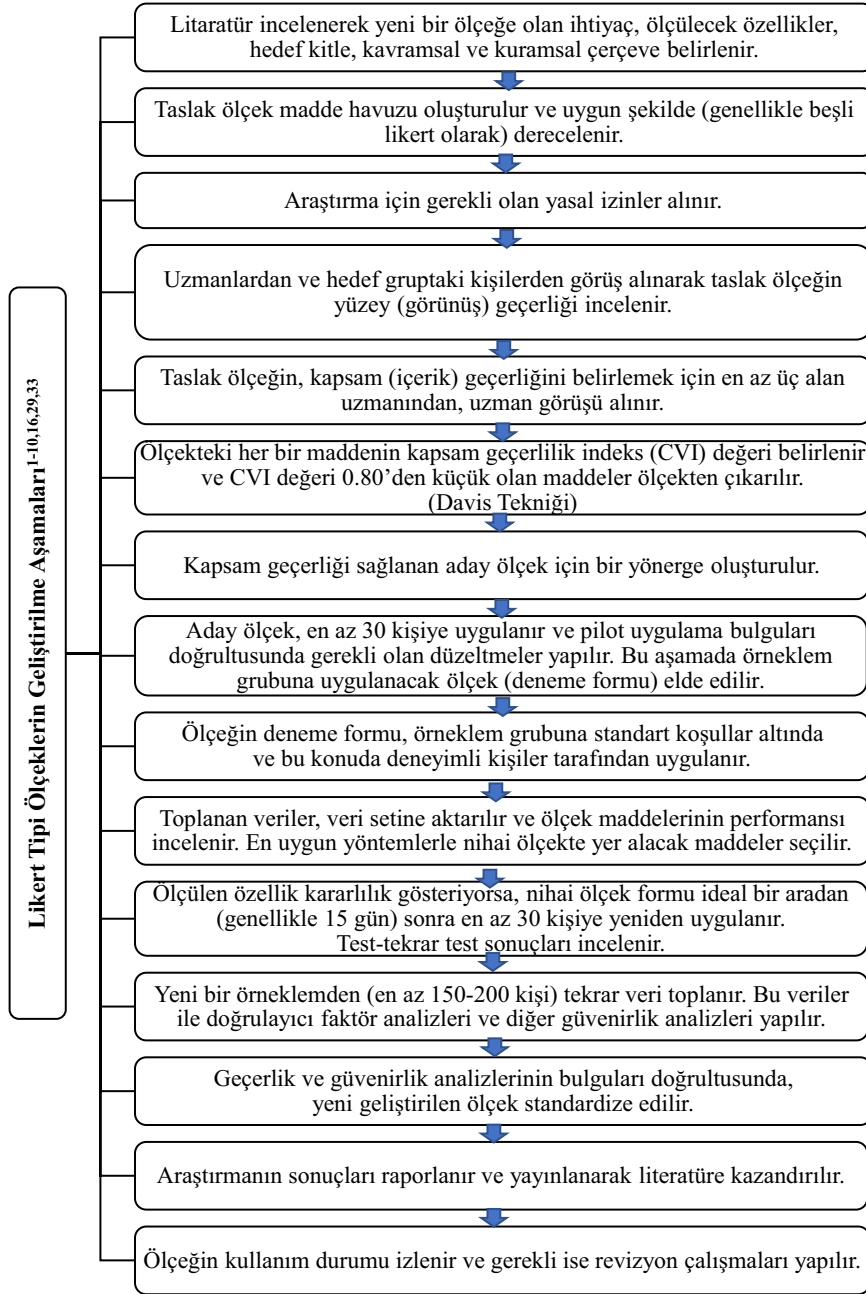
Ölçek uyarlaması yapılmadan önce sorumlu yazarla/fikri mülkiyet sahibi kurumla iletişime geçilerek yazılı izin alınmalıdır.^{6,11,19} İzin talep yazısında çalışma ekibi, ölçeğin kullanım amacı, araştırmacının örnekleme, ölçekte yapılabilecek küçük değişiklikler vb. belirtilmelidir.^{6,11} Orijinal ölçeğin uygulama, puanlama ve değerlendirme formları mutlaka istenmelidir.^{6,11} Etik sorunlar, zaman ve efor kayıpları yaşanmak için Türkiye'deki başka bir ekibe daha önce ölçek uyarlama izni verilip verilmediği sorulmalı, Türk kültürüne daha önce uyarlanmış ve uyarlama süreci devam eden ölçeklerin uyarlanması yapılmamalıdır.¹⁹ Araştırmacılar, uyarlama için izin istediklerinde orijinal ölçeğin kullanılmasını, atıfta bulunulmasını, sözleşme imzalanmasını, ücret ödenmesini, yetkinliğin belgelenmesini, çalışma ekibinde yer verilmesini, sonuçlarının paylaşılmasını isteme gibi taleplerle karşılaşabilir.^{6,11} Bu taleplerin nasıl karşılanacağını önceden planlanması ve diploma, sertifika gibi yetkinlik belgelerinin hazırlanması faydalı olabilir.¹⁵ Uyarlama ve ölçeğin klinikte kullanılması için ayrı ücretler talep edebilir. Bu nedenle alınan izinin, her iki işlemi de kapsamasına dikkat edilmesi gerekir.¹⁹ Araştırma sürecinde ölçeği geliştirenlerin isteklerine, sözleşme şartlarına, telif haklarına ve etik kurallara dikkat edilmelidir.¹⁹ Her iki tarafın açık kimlik bilgilerini içeren imzalı bir taahhütname alınması, tüm işlemlerin etik kurallar çerçevesinde yapılmasını sağlama açısından önemlidir.²⁰



Şekil 1. Likert Tipi Ölçeklerin Kültürel Uyarlama Aşamalarını Gösteren Akış Şeması.^{1,6,10,11,19,20}

Ölçek uyarlaması yapısal, dilsel, kültürel eşdeğerliği sağlamaya yönelik bir çok işlemin (Şekil 1) yapıldığı bir süreçtir ve bu süreçte çeşitli problemlerle karşılaşılabilir.^{10,11,16} En sık karşılaşılan problemler, ölçeğin geliştirildiği ve uyarlama işleminin yapılacağı grubun aynı olmaması, kapsam geçerliğinin, kültürel eşdeğerliğin, model uyumunun, orjinal ölçekteki psikometrik özelliklerin sağlanamamasıdır.^{3,19} Araştırmacılar, çevirilerden ve uygun olmayan bir ölçeği veya örnekleme seçmeden kaynaklanan hataları genellikle veri toplama ya da analiz aşamasında fark ederler.¹⁹ Fakat ölçek seçimi ve çevirisine yönelik işlemler, uyarlama işleminin kalitesini, ölçeğin hedef toplumdaki geçerliğini, kullanılabilirliğini, puanların

kültürlerarası karşılaştırılabilirliğini etkiler.^{11,20} Bu nedenle çeviriler titizlikle yapılmalı, anlamsal, kavramsal ve kültürel uyuma yönelik tüm önlemler alınmalıdır.^{6,19,20} Karaçam¹¹ orijinal ölçekteki ifadelerin, gerektiğinde kavramda eşdeğerliği sağlanmak amacıyla birden fazla madde ile hedef dile aktarılabilirliğini bildirmektedir. Fakat bu işlemin, ölçek geliştiriciler tarafından kabul edilmesi ve puanlamada değişikliğe neden olduğunun bilinmesi gerekir.^{6,19} İyi bir araştırma planının yapılması, çalışma ekibi, çevirmenler ve katılımcılar için yazılı talimatların oluşturulması uyarlamanın kalitesini artırır.^{11,20} Psikolinguistik (dil geçerliği) özellikler incelenirken ana dilleri, ölçeğin uyarlanacağı hedef kültürde olan fakat



Şekil 2. Likert Tipi Ölçeklerin Geliştirilme Aşamalarını Gösteren Akış Şeması.^{1-10,16,29,33}

her iki dili ve o alana yönelik bilimsel terimleri iyi bilen çevirmenler tercih edilmelidir.¹⁹ Geri çeviri yapanlar, ilk çeviri ekibinden farklı kişiler olmalı ve orijinal ölçeği önceden görmemelidir.^{11,20} Çeviriler tamamlandığında tereddüte düşülen noktalar için çevirmenlerden, uzmanlardan ve ölçeği geliştirenlerden görüş alınmalıdır.¹¹ Ölçeğin çeviri ve dil geçerliği sağlandıktan sonra, kapsam geçerliğinin incelenmesi için en az üç ve en fazla 20 alan uzmanından görüş alınmalıdır.¹⁹

Uzman görüşleri doğrultusunda kapsam geçerliği sağlanan ölçekler, 30-100 kişiden oluşan, örnekleme benzer bir grupta test edilmelidir.² Araştırmanın tüm verileri, yazılı gönüllü onamı alındıktan sonra ve benzer koşullar altında bu konuda yetkin kişiler tarafından toplanmalıdır.^{6,20} Özbildirim ölçekleri, aksi bir durum belirtilmediği sürece ölçek yönergesi doğrultusunda ve

katılımcının kendisi tarafından doldurulmalıdır.^{11,20} Pilot uygulama aşamasında katılımcıların maddeler hakkında ne düşündüğü, cevaplarını nasıl seçtiği, formları doldurken hangi sorularda zorlandıkları ve ne kadar zaman harcadıkları belirlenmelidir.²¹⁰ Bazı araştırmacılar pilot uygulama verileri ile ölçeğin psikometrik özelliklerinin incelenmesini önermektedir.^{1,20} Araştırmacılar bu öneri doğrultusunda analiz yapmayı planlıyorsa en az 200-300 kişi ile pilot uygulama yapmalıdır.^{1,19} Bu aşamada tespit edilen sorunlar giderilmeli, daha anlaşılır ve kullanışlı bir form oluşturularak asıl uygulama aşamasına geçilmelidir.²⁷

Araştırma sürecinde kaç kişiden veri toplanacağı ölçekte bulunan madde sayısı, araştırmacıların tercihleri, katılımcılara erişim koşulları, yapılması planlanan geçerlik ve güvenilirlik analizleri gibi etmenlerle ilişkilidir.²⁴⁻²⁶ Genellikle ölçekteki madde sayısının en

Tablo 2. Likert Tipi Ölçeklerde Sık Kullanılan Geçerlik Türleri

İşlemin Adı	Tanımı	Dikkat edilmesi Gereken Hususlar
Yüzey Geçerliği (Face Validity) ^{17,28}	Ölçeğin neyi ölçer görüldüğü ve bireyler tarafından nasıl algılandığının mantıksal bir yaklaşımla incelenmesidir.	Araştırmacılar, ölçeklerini öncelikle kendileri incelemeli ve daha sonra yakın çevresindeki arkadaşları, alan uzmanları, pilot uygulamaya katılan kişilerden görüşler almalıdır. En zayıf geçerlik türüdür ve sık kullanılmaz. Fakat kapsam geçerliği aşamasından önce yapıldığında, ölçekteki maddelerin düzeltilmesini ve gereksiz maddelerin elenmesini sağlar.
Kapsam Geçerliği (Content Validity) ^{2,6,10,17,33,37-43}	Ölçekteki maddelerin; ölçülmek istenen değişkenin tüm yönlerini temsil etme, yeterli nicelik ve nitelikte olma durumunun uzmanlar tarafından incelenmesi esasına dayanan geçerlik tekniğidir.	Oluşturulan maddeler, ölçülen değişkenin tüm temel yönlerini ele almalı ve konu ile alakasız soruları içermemelidir. Görüşü alınan uzmanlar, hem ölçülen konuda hem de ölçek geliştirme alanında çalışmaları olan bilim insanları arasından seçilmelidir. Araştırmacılar, en az üç uzmandan görüş almalı ve hangi tekniği kullanmak istiyorlarsa o yöntemin kriterlerine göre işlem yapmalıdır. Davis Tekniği'nde uzmanlar (5-20 kişi), her bir maddeye kapsamlı ilişkili olma durumuna göre bir ile dört arasında değişen puanlar vermelidir. Bir maddenin Kapsam Geçerlik İndeksi (KGİ, CVI) değeri bulunurken o maddeye üç ve dört puan veren uzmanların sayısı belirlenmeli ve toplam uzman sayısına bölünmelidir. Maddenin KGİ değeri 0,80'den küçük ise ölçekten çıkarılmalıdır. Polit-Back tekniği, Davis tekniğine benzemekle birlikte bu yöntemde hem maddelerin KGİ değeri (I-CVI) hem de bunların aritmetik ortalaması alınarak ölçek geneli için kapsam geçerlik değeri (S-CVI) bulunabilir. Şans faktörünü içerdikleri puanlayıcılar arası uyumun (Interrater agreement-IRA) kullanılması önerilmemektedir. Lawshe Tekniği kullanıldığında ise uzmanlar (3-40 kişi) her bir maddeyi gerekli olma durumuna göre bir ile üç puan arasında değerlendirmelidir. Bir maddenin Kapsam Geçerlik Oranı (KGO, CVR) bulunurken o maddeye gerekli diyen ve üç puan veren uzman sayısı, toplam uzman sayısının yarısının bölünmeli ve bir rakamı çıkarılmalıdır. Elde edilen KGO değerleri, uzman sayısına göre standardize edilmiş olan Minimum Kapsam Geçerlik Ölçütü (KGÖ) tablosu ile karşılaştırılmalıdır. Uzman sayısına göre önerilen KGÖ değerlerine ulaşılan maddeler ölçekte bırakılmalıdır. Aksi takdirde ilgili madde ölçekten çıkarılır. Uzmanlarla sürekli iletişimde kalınmalı ve pilot uygulamadan sonra gerekli görülürse tekrar uzman görüşü alınmalıdır.
Yapı Geçerliği (Construct Validity) ^{10,16,24-28,30-34,41-45,47}	Yapı geçerliliği, bir ölçeğin ölçülmesi amaçlanan teorik bir yapıyı ölçme derecesi olarak tanımlanır. Ölçeğin hangi alt boyutlardan (faktörler) oluştuğu, bu alt boyutlarda hangi maddelerin bulunduğu, ölçeğin ölçüm amacına hizmet edip etmediği ve alınan puanların ne anlama geldiği incelenir. Yapı geçerliği incelenirken nihai ölçeğin elde edilmesine yönelik bir çok analiz (faktör analizi, hipotez testi, bilinen grupların karşılaştırması vb.) yapılır. Yapı geçerliği sağlanan ölçeğin iç tutarlılık düzeyleri ve ölçüt bağıntılı geçerliliğinin de incelenmesi gerekir.	Ölçek uyarılama çalışmalarında verilerin analize uygunluğu incelenmeli, ölçekten alınan puanların karşılaştırılabilir olması için ölçeğin orijinal yapısı mümkün mertebe korunmalıdır. Ölçek geliştirme çalışmalarında ise veri setinin analizler için uygunluğu değerlendirilmeli ve ölçeğin faktör yapısının belirlenmesi için açıklayıcı faktör analizi (AFA) yapılmalıdır. Faktör analizi farklı yöntemlerle yapılabilir. Maksimum Olasılık Faktör Analizi (Maximum Likelihood) tekniği veriler normal dağılım gösterdiğinde kullanılır. Araştırmacılar, toplam varyans ile çalışmak istiyor ve veriler normal dağılıyorsa Temel Bileşenler Analizi (Principal Components Analysis) de yapabilir. Veriler normal dağılmıyor ve ortak varyans ile çalışılmak isteniyorsa ideal olarak Temel Faktörler Analizi (Principal Axis Factoring)'nin yapılması önerilir. Faktör analizi sonucunda, tek faktörlü ölçekler toplam varyansın, en az %30'unu, çok faktörlü ölçekler varyansın en az %40'unu açıklamalıdır. Ayrıca her bir faktör açıklanan varyans oranına en az %5 oranında katkıda bulunmalıdır. Faktör sayısı azaltılırken; açıklanan varyans oranları, Kaiser Kriterleri (Eigen value ≥ 1 olmalı), örneklem sayısının 200 kişiden fazla olması önkoşuluyla Scree Plot Grafiklerindeki keskin düşüş noktaları incelenmelidir. Ayrıca teorik yapıya uygunluk, rotasyon tekniği ve her bir faktörde en az üç madde bulunması gibi koşullar da göz önünde bulundurulmalıdır. Faktör puanları birbiri ile ilişkili ise eğik döndürme yöntemleri (Promax, Direct Oblimin) ve faktör puanları birbiri ile ilişkili değilse dik döndürme teknikleri (varimax, quartimax, equamax) kullanılmalıdır. Dik döndürme yapıldığında "Rotated Component Matrix" ve eğik döndürme yapıldığında ise "Patern Matrics" tablosundaki faktör yükleri yorumlanmalıdır. Analiz öncesinde yorumlamayı kolaylaştırmak için faktör yükü görüntüleme alt sınırı için 0,30 ve 0,32 gibi faktör yük değerleri tanımlanmalıdır. Daha güçlü bir faktör yapısı istiyorsa 0,40 veya 0,50 değerleri tercih edilmelidir. AFA yapılırken birden fazla faktöre çapraz yüklenen ve bu faktörlerde aldığı faktör yük değerleri arasındaki fark 0,10'dan (tercihen 0,15-0,20) daha küçük olan maddeler binişik olarak kabul edilmeli ve öncelikli olarak ölçetene çıkarılmalıdır. Daha sonra faktör yükü 0,30'dan ve düzeltilmiş madde-toplam puan korelasyon değeri 0,20'dan daha az olan maddeler ölçekten tek tek çıkarılmalıdır. Maddeler ölçekten atılırken, faktör yüklerinde ve korelasyon değerlerinde oluşan değişiklikler her aşamada yeniden değerlendirilmelidir. Faktör yükü 0,32-0,44 olan maddeler zayıf, 0,45-49 olan maddeler orta, 0,50-0,62 olan maddeler iyi, 0,63-0,70 olan maddeler çok iyi ve $\geq 0,71$ olan maddeler mükemmel olarak yorumlanmalıdır.

(Devamı)

Tablo 2. Likert Tipi Ölçeklerde Sık Kullanılan Geçerlik Türleri (Devamı)

İşlemin Adı	Tanımı	Dikkat edilmesi Gereken Hususlar
Yapı Geçerliği (Construct Validity) ^{1,10,24,25,28,32,41,45,48}		Faktör yükü 0,30-0,40 maddelerin ölçeğe alınabilmesi için örneklem sayısının en az 350 kişi, faktör yükü 0,40-0,50 arasında olan maddelerin ölçeğe alınabilmesi için örneklem sayısının en az 200 kişi olması gerekir. Faktör yükü 0,30'dan daha az olan maddeler ölçeğe alınacaksa, bu durumun kavramsal, kuramsal ve bilimsel nedenleri çok iyi bir şekilde açıklanmalıdır. Madde atma işlemi tamamlandığında ölçeğin her bir alt boyutunda en az üç ve mümkünse dört-beş madde kalmış olmalıdır. Bu maddelerin ortak varyansı açıklama oranları en az 0,10 ve mümkünse $\geq 0,20$ olmalıdır. AFA tamamlandığında en az 150 ve ideal olarak 250 kişilik yeni bir örneklemden veri toplanarak birinci ve ikinci düzey DFA yapılmalıdır. AFA ve DFA işlemleri aynı veri seti veya bu veri setinin iki eşit parçaya bölünmesi yolu ile de yapılabilir. Fakat bu durumun nedenlerinin belirtilmesi ve sınırlıklara eklenmesi gerekir. Birinci düzey DFA'da bulunan Path Diyagramı incelendiğinde, standardize çözüm değerleri (faktör yükleri) en az 0,30 ve mümkünse 0,50 olmalıdır. Maddelerin hata varyansları en fazla 0,90 olmalı ve tüm t değerleri anlamlı olmalıdır (kırmızı ok olan maddeler ölçekten atılmalıdır). Modelin verilerle uyum düzeyi belirlenirken, çoklu model uyum indeksleri incelenmelidir. Örneklem sayısı 200 kişiden fazla ise P değeri genellikle anlamlı olur ($P > ,05$). Bu durumda χ^2/df değerine bakılır. Eğer $\chi^2/df \leq 2$ ise model uyumu mükemmel ve $3 < \chi^2/df \leq 5$ arasında ise model uyumu kabul edilebilir düzeydedir. Bu değer 5'den büyükse model reddedilmelidir. Analizlerde RMSEA, SRMR ve RMR değeri $\leq 0,05$ ise model uyumu mükemmel, 0,05-0,08 ise kabul edilebilir, 0,08-0,10 ise kötü, $> 0,10$ ise kabul edilemez şekilde yorumlanmalıdır. NNFI (TLI) veya CFI değerleri 0,90'dan küçük ise oluşturulan ölçek modeli hatalıdır. AGFI değeri en az 85 olmalıdır. CFI, GFI, NNFI (TLI), NFI, IFI, RFI değerleri ≥ 90 ise kabul edilebilir uyum ve ≥ 95 ise mükemmel uyum vardır. CFI değeri 1,00'a, RMSEA ve SRMR sıfıra yaklaştıkça model uyumu da artmaktadır. Model uyumu sağlanmadığında, modifikasyon önerileri incelenmeli ve aynı alt boyuttaki maddelerin hata varyansları birleştirilmelidir. Modifikasyon sayısı üç ile sınırlı tutulmalıdır. İkinci düzey DFA sonuçları, birinci düzey DFA ile benzer ise birinci düzey DFA'daki uyum indeks değerleri yorumlanmalıdır. Analizi yapan kişiler, kullanılan teknikler ve istatistik programlar (LISREL, AMOS vb.) değiştirildiğinde aynı veri setinden farklı sonuçların elde edilebileceğini bilmelidir.
Yakınsak Geçerlik (Convergent Construct Validity) ^{7,17,23,47,49}	Kavramsal olarak aynı ya da benzer özellikleri ölçen iki farklı ölçeğin puanları arasındaki hipoteze dayalı ilişkilerin incelendiği yapı geçerliği türüdür. Birleşme ya da benzeşim geçerliği olarak da bilinmektedir.	Benzer özellikleri ölçen, ölçekler arasında pozitif yönde (yakınsak) ve orta düzeyde ($r = 0,30-0,50$, en fazla $r = 0,60$) anlamlı bir ilişki bulunur. Pearson korelasyon katsayısı en az 0,30 olmalı ve uygun bir ölçüt ölçek yoksa yeni ölçekten alınan puanlarla sosyodemografik özellikler (yaş eğitim, fiziksel aktivite vb.) karşılaştırılmalıdır.
İraksak Geçerlik (Divergent or Discriminant Construct Validity) ^{7,15,23,42,47}	Kavramsal olarak farklı özellikleri ölçen iki farklı ölçeğin puanları arasındaki hipoteze dayalı ilişkilerin incelendiği yapı geçerliği türüdür. Ayrışma geçerliği ya da ayırt edici geçerlik olarak da bilinmektedir.	Farklı yapıları ölçen ölçekler arasında negatif yönde (iraksak) veya çok zayıf düzeyde bir ilişki bulunması beklenir. Bu nedenle Pearson korelasyon katsayısının $r < 0,20$ olması gerekir.
Ölçüte Dayalı/Bağımlı/Bağıntılı Geçerlik (Criterion-Related Validity) ^{17,23,41,46,49}	Ölçme aracından elde edilen puanların, aynı yapıyı ölçen dış ölçütlerle (benzer ölçekler, altın standart testler, yaşam olayları vb.) ilişkisini inceleyen geçerlik tipidir. İki farklı şekilde yapılabilir: <i>Eşzamanlı Uyum Geçerliği (Concurrent Criterion Validity)</i> : Yeni geliştirilen ölçek ve geçerliği daha önceden kanıtlanmış başka bir ölçek ya da referans test bireylere, aynı anda uygulanır ve alınan puanlar arasındaki korelasyonlar incelenir. <i>Yordama (Tahmin) Geçerliği (Predictive Criterion Validity)</i> : Yeni ölçekten alınan puanlar ile gerçek hayatta beklenen yaşam olayları karşılaştırılır.	Pearson korelasyon katsayısı $r \geq 50$ ise mükemmel, $r \geq 0,40$ ise iyi ve $r \geq 0,30$ ise kabul edilebilir (makul) olarak yorumlanır. Ayrıca $r = 0,00-0,19$ ise ilişki yok, $r = 0,20-0,29$ zayıf ilişki, $r = 0,30-0,69$ orta düzeyde ilişki, $r = 0,70-0,79$ ise kuvvetli (yüksek güçlü) ilişki, $r = 80-0,90$ ise çok güçlü ilişki ve $r = 1,00$ ise mükemmel ilişki şeklinde de yorumlanabilir. Yordama geçerliğinde gerçek yaşam olayları kriter olarak alınmalı, kriter olarak kabul edilen ölçek ya da test, sınanan yeni ölçekten en az altı ay, en geç beş yıl sonra uygulanmalıdır. İki ölçek arasındaki korelasyon katsayısı $r \geq 0,30$ (en az 0,20) ise yeni geliştirilen ölçeğin yordama (tahmin) geçerliğinin olduğu kanısına varılmalıdır. Yeni geliştirilen ölçek ile ölçüt olarak kullanılan ölçek ya da test arasındaki ilişkiler belirtilmeli, yeni ölçeğin faydaları ve üstün yönleri vurgulanmalıdır.
Zıt-Bilinen Grup Karşılaştırması (Contrasted or Known Groups Validity) ^{6,40}	Yapı geçerliğine katkıda bulunan bir geçerlik türüdür. Bu yöntemde ölçek, ölçülen özellik açısından birbirine benzemeyen iki farklı gruba uygulanır ve sonuçlar karşılaştırılır.	Hastalar ve sağlıklılar, primiparlar ve multiparlar gibi farklı gruplar olduğunda uygulanır. Analiz sürecinde varyans analizi, ROC analizi, duyarlılık ve özgüllük analizi, bağımsız gruplarda t testi vb. kullanılabilir.

Kısaltmalar: Açıklayıcı Faktör Analizi (AFA), Doğrulayıcı Faktör Analizi (DFA), Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error of Approximation, RMSEA), Standardize Edilmiş Artık Ortalamaların Karekökü (Standardized Root Mean Square Residual, SRMR), Artık Ortalamaların Karekökü (Root Mean Square Residual, RMR), Normalleştirilmemiş Uyum İndeksi (Non-normed Fit Index, NNFI), Tucker Lewis İndeksi (Tucker Lewis Index, TLI), Karşılaştırılmalı Uyum İndeksi (Comparative Fit Index, CFI), Düzenlenmiş İyilik Uyum İndeksi (Adjusted Goodness of Fit Index, AGFI), Uyum İndeksi (Goodness of Fit Index, GFI), İyilik Normalleştirilmiş Uyum İndeksi (Normed Fit Index, NFI) Fazlalık Uyum İndeksi (Incremental Fit Index, IFI), Görelî Uyum İndeksi (Relative Fit Index, RFI).

Tablo 3. Likert Tipi Ölçeklerde Sık Kullanılan Güvenirlik Türleri

İşlemin Adı	Tanımı	Dikkat edilmesi Gereken Hususlar
İç Tutarlılık Güvenirliği (Internal Consistency Reliability) ^{1,4,16,28,30,34,43,50}	Ölçekteki tüm maddelerinin aynı yapıyı ya da değişkeni tutarlı ve homojen bir şekilde ölçüp ölçmediğinin incelenme işlemidir. İç tutarlılık analizleri hem yapı geçerliği hem de güvenirlilik için çeşitli kanıtlar sunmaktadır.	Likert türü ölçeklerin iç tutarlılığı en sık Cronbach's Alfa Katsayısı (α) ile belirlenir. Fakat McDonald'ın Omega (ω) yöntemi de kullanılabilir. Maddelerarası korelasyonlar, Madde-toplam puan korelasyon analizleri ve Testi Yarlama Tekniği (Split-Half Method) de ölçeğin iç tutarlılığı hakkında bilgi verir. Cronbach's Alfa katsayısının (α) ölçeğin geneli ve her bir yarısı için $\geq 0,70$ olması önerilir. Yeni bir ölçek geliştiriliyorsa bu değer $\alpha \geq 0,60$ olması yeterli kabul edilir. Cronbach's α değeri $\alpha \geq 0,90$ ise ölçeğin iç tutarlılığı mükemmel, $0,70 \leq \alpha < 0,90$ ise iyi, $0,60 \leq \alpha < 0,70$ kabul edilebilir, $0,50 \leq \alpha < 0,60$ ise zayıf, $\alpha < 0,50$ ise kabul edilemez düzeydedir. Cronbach's α değeri $< 0,50$ olan ölçeklerin kullanılmaması ve modifiye edilmesi önerilir. Çok faktörlü ölçeklerde, her bir alt boyutun α değerlerinin ayrı ayrı bulunması gerekir. İç tutarlılığı olumsuz etkileyen maddeler varsa, normallik koşulu sağlanamamışsa, maddeler arasındaki korelasyonlar düşükse ($< 0,20$), ölçekteki madde sayısı az (< 10) ise α değerleri düşük olur. Ölçekte birbirine benzeyen çok sayıda madde olduğunda ise Cronbach α değeri hatalı olarak yüksek çıkabilir.
Madde Toplam Puan Güvenirliği (Item-Total Correlation Reliability) ^{7,25,32,34}	Ölçekteki her bir maddenin ölçülen yapı ve ölçeğin toplam puanı ile ne düzeyde ilişkili olduğunu inceleyen güvenirlilik tekniğidir. Bir maddenin ölçeği temsil düzeyini belirlemek için genellikle düzeltilmiş madde toplam puan korelasyon (corrected item-total correlation reliability) analizleri kullanılır. Madde toplam-puan korelasyon değerleri (item-total correlation reliability) ve maddelerarası korelasyonlar değerleri (inter-item correlation reliability) de istendiğinde incelenebilir.	Olumlu ifadeler ölçeğin toplam puanı ile pozitif yönde, ters maddeler ise negatif yönde ve düşük düzeyde bir korelasyon gösterir. Örneklem homojense varyans oluşturmada yetersiz kalır ve korelasyonlar düşük çıkar. Pearson Momentler Çarpımı ile elde edilen madde toplam puan korelasyon değerleri istatistiksel açıdan anlamlı olmalıdır ($P < ,05$). Anlamlı değilse o madde ölçekten atılmalıdır ($P \geq ,05$). Likert tipi ölçeklerde yer alan maddelerin, düzeltilmiş madde-toplam puan korelasyon değerleri 0,30 ile 0,70 arasında olmalıdır. Bu değeri 0,20'den az veya 0,80'den fazla olan maddeler ölçekten çıkarılmalıdır. Madde toplam puan korelasyonunda ölçekteki tüm maddeler dikkate alınırken, düzeltilmiş madde-toplam puan korelasyonunda ise ilgili madde analizde dikkate alınmamaktadır.
Madde Analizine Dayalı Güvenirlilik Teknikleri (Güvenirlilik ve Ayırt Ediciliğe Yönelik Diğer Madde Analizleri) ^{17,33,34,46}	Tepki yanlılığı, taban ve tavan etkisi, maddelerin ayırt edicilik düzeyleri ve iç tutarlılığa olan katkıları, ölçek puanlarının toplanabilirliği gibi özellikleri incelemek amacıyla yapılan madde analizlerini içerir.	Hotelling T^2 testi istatistiksel açıdan anlamlı ($P < ,05$) ise tepki yanlılığı yoktur. Taban ve tavan etki düzeyleri %15-20'den daha fazla olmamalıdır. Tukey's Test for Nonadditivity testindeki P değeri $> ,05$ ise ölçek puanları toplanabilir. Ölçekten alınan toplam puanlar açısından bireyler büyükten küçüğe doğru sıralanıp %27'lik üst ve %27'lik alt gruptaki bireyler bağımsız örneklemde t testi ile karşılaştırılır. Eğer istatistiksel açıdan anlamlı farklılıklar ($P < ,05$) oluşuyorsa ölçekteki maddelerin üst ve alt gruptaki bireyleri ölçülen özellik açısından ayırt edebildiği söylenebilir.
Test-Tekrar Test Güvenirliği ^{7,9,32,41,47}	Ölçekten alınan puanların zaman karşısında ne düzeyde kararlı olduğunun incelendiği tekniktir. Bu yöntemde nihai ölçek formu, örneklemden seçilen bireylere farklı zamanlarda toplam iki kez uygulanır ve her iki ölçüm arasındaki ilişkiler incelenir.	Ölçülen özelliğin zaman karşısında stabil kaldığı düşünülüyorsa yapılmalıdır. Birinci ve ikinci uygulamanın koşulları birbiri ile aynı ya da benzer olmalıdır. Analiz için en az 30 çift veriye ihtiyaç vardır. Fakat denek kaybı olabileceği için yedek kişi alınmasında fayda vardır. Verilerin nasıl eşleştirileceği önceden planlanmalıdır. İki ölçüm arasındaki süre kısa ise korelasyon değeri hatalı olarak yüksek ve süre uzun ise düşük çıkabilir. Bu nedenle ölçülen özelliğe en uygun olan zaman aralığı seçilmelidir. Kâğıt kalem testleri için 15 günlük (iki-dört hafta) süre ideal olarak kabul edilir. Tutarlılık/kararlılık katsayısının; ölçüm aracından, ölçülen özelliğin (uyku, ruh hali, eklem ağrısı, bulantı vb.) değişken olmasından, katılımcılardan (kişinin yorgun olması, soruları yanlış okuması, sorular için ön hazırlık yapması, önceki yanıtlarını hatırlaması, sosyal beğenirlik kaygısı vb.), ölçüm koşullarından (ortamın özellikleri, uygulama yöntemi, ölçümün yapıldığı saat vb.) etkilenebileceği unutulmamalıdır. Veriler, normal dağılım gösteriyorsa bağımlı örneklemede eşleştirilmiş t testi (paired sample t-testi) ve aksi bir durumda ise Wilcoxon işaretli sıralar testi ile analiz yapılmalıdır. Ölçeğin zaman karşısında stabil olduğunun söylenebilmesi için subjektif testlerdeki kararlılık katsayısı en az 0,50 olmalıdır. Tutum ölçeklerinde ise bu değer en az $r = 0,70$ ve tercihen $r = 0,80$ olmalıdır.

(Devamı)

Tablo 3. Likert Tipi Ölçeklerde Sık Kullanılan Güvenirlik Türleri (Devamı)

İşlemin Adı	Tanımı	Dikkat edilmesi Gereken Hususlar
İki Yarı Test Güvenirliği (Split Half Reliability) ^{6,7,9,32,34,41}	Toplam puan üzerinden değerlendirilen bir ölçeğin, iç tutarlılık ve kararlılık düzeyi hakkında bilgi edinmek için kullanılan geçerlik tekniğidir. Bu yöntemde yeni ölçek, katılımcılara tek bir oturumda uygulanır. Maddeler raslantısal olarak iki eşit parçaya bölünür ve her iki yarı arasındaki ilişkiler incelenir.	Madde sayısı fazla olan ölçeklerde kullanılır. İkiye bölme işlemi, ölçeğin birinci ve ikinci yarısı ya da tek ve çift sorular şeklinde yapılır. Ölçeğin birinci ve ikinci yarısı arasındaki korelasyon katsayısı en az 0,30 olmalıdır. Cronbach Alfa güvenirlilik katsayı değerleri ölçeğin hem birinci hem de ikinci yarısı için incelenir. Bu değer ideal olarak $\alpha \geq 0,70$ ve en az $\alpha \geq 0,60$ olmalıdır. Ayrıca Spearman-Brown ve Guttman katsayıları da raporlanmalıdır. Testin her iki yarısı arasındaki korelasyonlar yüksek olduğunda ölçeğin iç tutarlılığının iyi seviyede olduğu düşünülmelidir.
Paralel (Eşdeğer) Form Güvenirliği ^{6,7,32,41}	Aynı bireylerden oluşan gruba, bir ölçeğin farklı versiyonlarının uygulanması esasına dayanan güvenirlilik tekniğidir. Aynı özelliği ölçen paralel bir form varsa kararlılık katsayısının belirlenmesi için kullanılır. Fakat eşdeğer ölçekleri bulmak ve oluşturmak güç olduğu için sık kullanılmayan bir tekniktir.	Yeni ölçek, ilk oturumda ve paralel form ikinci oturumda uygulanmalıdır. Ölçeğin paralel formu yoksa araştırmacılar aynı soruları farklı kelimelerle ifade ederek ve uzman görüşü alarak, eşdeğer bir form oluşturulabilir. Ölçeklerin ölçtüğü nitelikler, soru sayıları, güçlük dereceleri, aritmetik ortalamaları ve standart sapma değerleri benzer ise yeni geliştirilen ölçek tutarlı ve güvenilirdir.

az beş katı büyüklükteki bir örnekleme çalışılması ve ortalama 300 kişiden veri toplanması önerilmektedir.^{4,25} Çapık ve ark¹⁹ örneklem sayısının 200 kişiden az olmaması gerektiğini bildirmektedir. Toplanan veriler, Tablo 2 ve 3'de gösterilen yöntemlerden uygun olanlarla incelenmeli, ölçeğin o kültür için geçerli ve güvenilir olduğu kanıtlanmalıdır.^{10,11} Ölçek uyarlama çalışmalarında genellikle doğrulayıcı faktör analizi (DFA) ve güvenirlilik analizlerinin yapılması yeterli görülmektedir.¹⁰ Orçan²⁷ sadece DFA yapıldığında en iyi uyum gösteren ölçek modelinin tespit edilemeyeceğini bildirmekte ve açılımlı faktör analizi (AFA)'nin de yapılması gerektiğini savunmaktadır. Gürbüz ve Şahin²⁸ kabul edilebilir sayıdaki modifikasyona rağmen ölçeğin, faktöriyel ve yapısal geçerliği sağlamadığında keşifsel nitelikteki AFA'nın yapılmasını önermektedir. Özdamar'a¹ göre pilot uygulama verileri ile AFA yapılmalı, daha sonra en az 150 kişilik farklı bir örneklemeden yeniden veri toplanmalı, bu verilerle AFA tekrarlanmalı, DFA ve güvenirlilik analizleri yapılmalıdır. AFA ve DFA'nın farklı örneklemeler ile yapılması idealdir.^{29,30} Fakat bu işlemlerin aynı veri seti veya bu veri setinin ikiye bölünmesi yoluyla yapılması da yaygın bir uygulama olup gerektiğinde kullanılabilir.³¹ Test tekrar test uygulaması, zorunlu olmayıp ölçülen özellik zaman karşısında stabil ise yapılmalıdır.^{7,9} Bu kapsamda ilk uygulamadan iki-dört hafta sonra en az 30 kişiden yeniden veri toplanması önerilir.^{9,10} Fakat test tekrar test için ne kadar zamanın geçmesi gerektiği tartışmalı bir konudur ve ölçülen özelliğe göre bir saatten bir yıla kadar değişen süreler belirlenebilir.³²

Araştırmacılar, literatürdeki farklı önerileri bilmeli ve kendi çalışmalarını için en uygun olan yöntemleri tercih etmelidir.^{7,24} Analizleri tamamladıklarında, mevcut bulguları orijinal ölçek ile karşılaştırılarak raporlamalı ve yayınlayarak literatüre kazandırmalıdır.^{1,10} Uyarlanacak ölçek, eşit derecede doğal ve kabul edilebilir olmalı, orijinal ölçekle aynı ya da benzer performans gösterebilmelidir.^{11,20} Ölçekten çok fazla sayıda madde çıkarılması ya da eklenmesi gerektiği belirlenirse yeni bir ölçek geliştirilmelidir.¹⁹

Likert Tipi Bir Ölçeğin Geliştirilmesi

Yeni bir kavram ortaya çıktığında, araştırılacak yapıyı ve değişkenleri ölçen ölçek olmadığında, literatürdeki ölçekler hedef kitle farklılığı, psikometrik özelliklerinde yetersizliği, uygulama, değerlendirme güçlüğü gibi nedenlerle ihtiyaca cevap vermediğinde yeni bir ölçek oluşturmak için yapılan işlemlere ölçek geliştirme denir.^{1,6,16,18,28} Bu işlemin birincil amacı, ilgilenilen odak değişkeni

ve yapıyı ölçen yeni bir ölçek yaratmaktır.³ Mevcut bir ölçeğin kullanılması, modifiye edilmesi ya da uyarlanması, ölçek geliştirmeye kıyasla çok daha kolay ve ekonomiktir.^{1,11,18} Fakat araştırmacılar, bazen çalışacakları alanda kullanabilecekleri, modifiye edebilecekleri ya da kendi kültürlerine adapte edebilecekleri nitelikli bir ölçek bulamayabilir.^{1,10,18} Ayrıca literatürde, konuyla ilgili ölçekler olsa bile farklı bir bakış açısıyla bakarak yeni bir ölçek geliştirmeye karar verebilir.¹

Likert tipi ölçek geliştirme, uzmanlık gerektiren zor bir işlemdir. Metodolojik yaklaşımın yansız bir-iki yıllık bir süre, önemli düzeyde teorik bilgi ve istatistiksel beceri gerektirir.^{2,33,34} Şekil 2'de gösterilen aşamalarının uygulanmaması; hatalı çalışma bulgularına, önyargılı ve güvenilir olmayan ölçeklerin geliştirilmesine neden olur.²⁷ Bu nedenle araştırmacıların; ön hazırlık aşamasında mevcut ihtiyacı, ölçmek istedikleri algı, tutum, davranış, inanç, niyet, risk gibi özellikleri, ölçeğin hedef kitesini, kimler tarafından kullanılacağını, madde sayısını, uygulama süresini, dereceleme, puanlama, yorumlama özelliklerini, maddelerin nasıl hazırlanacağını ve uzman görüşlerinin nasıl alınacağını çok iyi planlaması gerekir.^{7,10,16,33} Kapsamlı bir literatür taraması yapılması, uzman kişilerle görüşülerek ölçülecek fenomenin boyutlarının, yapısının, bileşenlerinin ve kavramsal çerçevesinin ortaya çıkarılması önemlidir.^{1,3} Ayrıca araştırma ekibi, uzman paneli, örneklem büyüklüğü, araştırmaya alınma ve dışlanma kriterleri, uygulama yönergesinin ve gönüllü onamının içeriği, hangi istatistiksel analizlerin yapılacağı, ölçek puanlarının nasıl yorumlanacağı, yasal izinlerin nasıl alınacağı, verilerin nasıl toplanacağı gibi birçok aşamanın planlanması gerekir.^{2,10,33}

İyi bir ölçeğin temelini, ölçülecek özelliklerle ilişkili nitelikli sorular oluşturur. Bu nedenle madde havuzu yazma, kaliteli bir ölçeğin temelini atmanın ilk ve en kritik koşulu olarak kabul edilir.² Ölçek maddeleri, tümdengelim (deductive) veya tümevarım (inductive) yaklaşımlarıyla ya da her iki yöntemin birlikte kullanılmasıyla yazılabilir.^{28,29} Tümdengelim yaklaşımı, kapsamlı bir literatür taramasını ve mevcut ölçeklere dayalı madde oluşturmayı içerir. Tümevarım yaklaşımında ise hedef kitleden çeşitli yöntemlerle (gözlem, odak grup görüşmesi, röportaj, kompozisyon yazdırma, uzman paneli, Delphi Tekniği vb.) toplanan nitel veriler esas alınır.²⁹ Nitel tekniklerin kullanılması, araştırmanın kalitesini ve madde havuzunun çeşitliliğini artırır. Fakat odak grup görüşmelerinde az sayıdaki kişiden veri toplanır ve katılımcıları bulma, görüşmeleri

yapma, verilerini sentezleme ile ilgili sorunlar yaşanabilir.³ Zhou³⁵ karma yöntemlerin kullanılmasının, ölçek geliştirmede en uygun araştırma tasarımı biçimi olduğunu bildirmektedir. McKim³ ise bu işlemin maliyeti ve zaman kayıplarını arttırdığını savunur. Literatür taraması tekniğinin; hakem denetiminden geçmiş yayınların seçilmesi, ekonomik olması, süreyi kısaltması, ölçülecek yapının netleştirilmesi, referans ölçeklerin belirlenmesi, nitel veri toplama kaynakları ve yanlılığın azaltılması gibi avantajları vardır. Bu nedenle daha sık tercih edilir.^{3,29} Ölçek geliştirme sürecine literatür taraması ile başlanması ve mümkünse nitel yaklaşımların da kullanılması tavsiye edilir.^{2,29} Maddeler yazılırken aşağıdaki noktalara dikkat edilmesi, ilgili alan ve ölçme değerlendirme uzmanlarından yardım alması ise daha kaliteli ölçeklerin geliştirilmesine katkıda bulunur.^{3,7,8,10,16,36}

- Geliştirilecek ölçeğin, kullanım amacı ve literatür taraması yapılırken kullanılacak yaklaşımlar (arama kriterleri, anahtar kelimeler, veri tabanları, taramanın yapılacağı yıllar, seçilecek makalelerin dili vb.) belirlenmelidir.
- Sorular hazırlanırken en az 30 ve mümkünse 90 makale incelenmeli, ideal makaleler seçilerek teorik açıdan analiz edilmeli ve bulgular bir tabloda özetlenmelidir.
- Yazılan maddeler; birinci tekil şahıs zamiriyle, şimdiki zaman ya da geniş zaman kipinde ve en fazla 20 kelime ile yazılmalıdır.
- Kısa, açık, anlaşılır, tarafsız, ölçülen konuya spesifik ve yanıt çeşitliliğini sağlayan nitelikte maddeler yazılmalıdır.
- Madde sayısı planlanırken, maddelerin %50'sinin analizlerde kaybedilme olasılığı ve katılımcıların sayısı dikkate alınmalı, nihai ölçekte kalması planlanan madde sayısının en az iki, mümkünse üç-dört katı madde yazılmalıdır.
- Taslak ölçeğin her bir alt boyutunda en az üç ya da dört, mümkünse 8-10 madde bulunmasına ve anket uzunluğunun 4-12 sayfa arasında olmasına dikkat edilmelidir.
- Maddeler, belirtke tablosunda özetlenmeli ve ifadeler incelemek geliştirilmelidir.
- Maddelerin içeriği ölçeğin amacına uygun olmalı, belirsiz, yoruma açık, birden fazla anlam içeren, yönlendirici ve kafa karıştırıcı ifadelerden kaçınılmalıdır.
- Tutum soruları yazılırken konuyla ilgili arzu edilen ve edilmeyen ifadeler yazılmalı, olgusal durumları ve gerçek olayları yorumlamaktan kaçınılmalıdır. Bilgi, duygu, davranış soruları eşit sayıda olmalıdır.
- Bireylerin farklı nedenlerle aynı yanıtlar vereceği sorulardan, herkes tarafından kabul veya reddedilme olasılığı yüksek olan maddelerden kaçınılmalıdır.
- Maddeler, ölçülmek istenen değişkenin tüm boyutlarını yansıtabilmeli, soruların sayısı ve ölçekteki sıralaması tarafsızlığı sağlar nitelikte olmalıdır.
- Maddeler, hedef kitlenin eğitim seviyesine ve kültürüne uygun olmalı, farklı geçmiş deneyimleri olan bireyler bu maddeleri benzer şekilde algılayabilmelidir.
- Aynı soru içerisinde birden fazla konu, durum, yargı olmamalıdır. Her bir maddenin, tek bir düşünceyi ya da niteliği ölçmesi sağlanmalıdır.
- Yönlendirici sorulardan, belli bir alana özgü teknik terimlerden, karmaşık ve bileşik cümlelerden, "sadece," "yalnızca," "bir tek," "eğer," "çünkü," "ve," "veya," "ya da" gibi sözcüklerden kaçınılmalıdır.
- Deneme formu, konu birliğini sağlayacak şekilde oluşturulmalı, ölçekteki maddeler kolaydan zora doğru sıralanmalıdır. Başlangıç soruları daha basit ve ilginç olmalıdır. Hassas, karmaşık ve sıkıcı sorular, en son bölüme bırakılmalıdır.

- Madde içerisinde birden fazla olumsuz ifade olmamalı, olumlu ve olumsuz maddeler aynı alanda toplanmamalı, rastgele dağıtılmalıdır.
- Olumsuz (ters) maddeler, nihai ölçeğin performansını düşürdüğü için maddelerin yarısının olumlu ve diğer yarısı olumsuz şekilde yazılmasından kaçınılmalıdır.
- Noktalama işaretlerine ve dilbilgisi yazım kurallarına dikkat edilmeli, Türk Dili uzmanlarından görüş alınarak ölçeğin anlaşılabilirliği artırılmalıdır.
- Ölçeğin amacını, madde sayısını, yanıtların nasıl işaretleneceğini, cevaplama süresini, isim yazma, test tekrar teste katılma gibi bilgileri açıklayan kısa, anlaşılır ve motivasyonu artırıcı özellikte bir ölçek yönergesi oluşturulmalıdır.
- Ölçek maddeleri, üç ile 11 arasında değişen kategorilere ayrılarak derecelenebilir. Fakat çocuklar için üçlü, yetişkinler için beşli ya da yedili derecelenmeler önerilir.
- Dereceleme yapılırken; katılımcıyı kararsızlığa ve belirsizliğe götüren kelimelerin (hepsi, daima, hiçbir, asla, bazen, ara sıra, her zaman, nadiren, çok az, kısmen, biraz, birkaç, tamamen katılıyorum, katılmıyorum, kararsızım vb.) birlikte kullanılmaması ve sosyal beğenilirden kaynaklanabilecek sorunların azaltılması önerilir.
- Kullanılan yazıların puntosu, örneklem grubuna uygun olmalı, sayfa düzenine ve likert derecelenimin her sayfanın en üst kısmında görünmesine dikkat edilmelidir.
- Ölçekteki likert derecelenmeler, maddelerin içerisinde olmamalıdır. Dereceleme ve puanlamalar, maddelerle uyumlu olmalı, olumlu maddelere yüksek ve olumsuzlara daha düşük puanlar verilmelidir.
- Örneklemden, veri toplanmadan önce benzer özellikteki küçük bir grupta pilot uygulama yapılmalı ve katılımcılardan gelen öneriler ölçeğe eklenmelidir.
- Geçerlik ve güvenilirlik analizlerinden sonra ölçekte, ölçülmek istenen yapıyı temsil edecek sayı ve nitelikte madde kalmadığı belirlenirse, araştırmacılar deneyimli kişilerden yardım almalı, mevcut maddeleri geliştirilmeli ya da yeni maddeler yazarak tekrar uzman görüşü alınmalıdır.

Uzman görüşü alma; ölçeğin kapsam geçerliğini belirleme, anlaşılabilirliğini artırma ve uygun olmayan maddeleri eleme amacıyla kullanılan en iyi nitel yöntemdir.²⁹ Bu amaçla kullanılan farklı teknikler bulunmakta ve seçilen teknik doğrultusunda üç ile 40 arasında değişen uzmana ihtiyaç duyulmaktadır.^{2,6,34,37} Uzmanlara değerlendirme için en az yedi günlük bir sürenin verilmesi ve hatırlatmalar (e-posta, telefon, yüzyüze görüşmeler vb.) yapılması önerilmektedir.² Uzman görüş formları standardize olmalı ve kullanılan tekniğe göre derecelendirilmelidir.^{2,37} Lawshe Tekniği'nde üçlü dereceleme (Gereksiz/Çıkarılmalı=1 puan, Yararlı ancak Yetersiz=2 puan, Gerekli/Uygun=3 puan) sistemi kullanılır.³⁸ Polit-Beck ve Davis Tekniği'nde ise dörtlü derecelenmeler (Uygun Değil/İlişkisiz=1 puan, Biraz Uygun/Majör Değişiklik Gerekli=2 puan, Uygun/Minör Değişiklik Gerekli=3 puan, Çok Uygun/İlişkili=4 puan) yapılır.^{2,6,37,39} Ayrıca madde hazırlama ve revize etme aşamasında sadece uzmanlardan değil, hedef gruptan ve uzman olmayan kişilerden de görüş alınması önerilir.^{2,28}

Kapsam geçerliği sağlanan ölçeklerin, pilot uygulamasının en az 30 kişilik bir grupta yapılması önerilir.² Bu aşamada taslak ölçeğin psikometrik özellikleri de incelenenirse en az 200 kişiden veri toplanmalıdır.¹⁴ Pilot uygulama standart koşullarda yapılmalı, veri toplama formlarının yanıtlama süresi ve anlaşılmayan maddeler belirlenmelidir. Katılımcıların önerileri ölçeğe eklenmeli ve gerekli ise yeniden uzman görüşü alınmalıdır.¹⁰ Sonraki aşamalarda ise

revize edilen ölçek ile hetorejen bir örneklemden veri toplanarak, nihai ölçeğin psikometrik özellikleri incelenmelidir.^{1,2,29} Örnekleme büyüklüğü, pilot uygulamadan sonra belirlenmeli ve pilot çalışmaya katılanlar örnekleme dahil edilmemelidir.^{34,40} Bu konuda farklı öneri ve yaklaşımlar olmakla birlikte genellikle ölçekteki madde sayısının en az beş (5 : 1), tercihen 10 (10 : 1) katı büyüklükteki bir örnekleme çalışılması ve en az 300 kişiden veri toplanması önerilmektedir.^{4,25} Bu oranın 15:1 veya 20:1 şeklinde olması ise sonuçların güvenilirliğini arttırmaktadır.^{4,25,26} Test tekrar test, yapılması planlanıyorsa nihai ölçek formu ile en az 30 kişiden aralıklı veya aralıksız yöntem kullanılarak iki kez veri toplanmalıdır.⁹ İki uygulama arasındaki süre ölçülecek özelliğe göre belirlenmelidir.³²

Veriler deneyimli kişiler tarafından toplanmalı ve Tablo 2 ve 3'deki yöntemlerden uygun olanları ile analiz edilerek, geliştirilen ölçeğin geçerlik ve güvenilirliğini gösteren kanıtlar elde edilmelidir.⁴¹⁻⁴⁵ Geçerlik ve güvenilirlik, tüm ölçme araçlarında bulunması gereken iki temel özelliktir.⁴¹ Ölçme aracının, kullanım amacına uygun olması, ölçülecek yapıyı tam olarak temsil etmesi ve farklı değişkenlerle karıştırmadan doğru şekilde öçebilmesine geçerlik denilmektedir.^{10,41,45} Yüzey geçerliği (face validity), kapsam geçerliği (content validity), yapı geçerliği (construct validity), yakınsak geçerlik (convergent construct validity), iraksak geçerlik (divergent or discriminant construct validity), ölçüt geçerliği (criterion-related validity) gibi türleri bulunmaktadır.^{10,41-44} Güvenirlik ise ölçme aracının hedeflenen yapıyı ölçmede duyarlı, hatalardan arınık ve homojen olması, benzer koşullarda tekrarlanan birbirinden bağımsız farklı ölçümlerde tutarlı ve kararlı sonuçlar üretebilmesidir.^{10,40,41,46} Likert tipi ölçeklerin güvenilirlikleri; iç tutarlılık analizleri (Cronbach alfa katsayısı vb.), madde analizleri (madde-toplam puan korelasyonu vb.), iki yarıya bölme tekniği (split half reliability), paralel (eşdeğer) formlar yöntemi, test-tekrar test işlemi, bilinen grupların karşılaştırılması, %27'lik alt ve üst grubun karşılaştırılması gibi yöntemlerle incelenmektedir.^{4,6,40,43,46} Yeni ölçeklerin faktör yapısı ise AFA ile belirlenmektedir.^{10,25} Faktör analizleri yapılmadan önce kayıp ve uç değerlerin kontrol edilmesi, normallik koşulunun sağlanması ve örneklem büyüklüğünün en az iki yöntemle incelenmesi gerekmektedir.²⁴ Örneklemin yeterli sayılabilmesi için her bir test için önerilen minimum katılımcı sayısı sağlanmalı, anti-amage korelasyon değerleri ve Kaiser-Meyer-Olkin Katsayısı (KMO) en az 0,50 olmalıdır.^{25,45} KMO değeri 0,50-0,59 ise örneklem büyüklüğü zayıf, 0,60-0,69 ise orta, 0,70-0,79 ise iyi, 0,80-0,90 ise çok iyi ve >0,90 ise mükemmel olarak yorumlanmalıdır.^{10,25,45} Normallik koşulunun sağlandığının söylenmesi için Bartlett Küresellik Testi anlamlı ($P < ,05$), maddelerin çarpıklık (skewness) değerleri <3 ve basıklık (kurtosis) değerleri <10 olmalıdır. Tablo 2'deki geçerlik ve Tablo 3'deki güvenilirlik kriterleri, hem mevcut ölçeklerin, hem de yeni geliştirilen ve uyarlaması yapılan ölçeklerin psikometrik özelliklerini değerlendirmek için kullanılabilir. Likert tipi ölçeklerin, geçerlik ve güvenilirlikleri, aşağıdaki yöntemler kullanılarak arttırılabilir.^{1,9,16,23,43}

- Örneklem seçimi dikkatli yapılmalı, her bireye eşit katılım hakkı tanınmalı, gönüllü onamı alınarak, yapılacak analizler için ideal sayıdaki örnekleme ulaşılmalıdır.
- Katılımcılar, gönüllü bireyler arasından önceden belirlenen kriterlere göre seçilmeli ve örneklemden çıkarılan kişilerin, hedef kitlenin tüm alt gruplarını kapsamaması sağlanmalıdır.
- Ölçeğin amacı, kaç madde olduğu, cevaplamanın nasıl yapılacağı, cevaplama süresi gibi bilgileri içeren anlaşılır bir ölçek yönergesi oluşturulmalıdır.

- Maddeler homojen ve anlaşılır olmalı, basitten karmaşığa doğru sıralanmalıdır. Aynı özelliği ölçen maddeler bir arada toplanmamalı, karıştırılarak sorulmalıdır.
- Pilot uygulama grubunun verileri analiz edilmeli ve saptanan sorunlar asıl uygulama aşamasından önce giderilmelidir.
- Korelasyon katsayılarının düşük çıkarmaması için homojen gruplardan değil, hetorejen gruplardan veri toplanmalıdır.
- Tüm veriler, benzer koşullar altında toplanmalı, öz bildirim ölçeklerini kişinin kendisi doldurmalı ve yapılan açıklamalar tüm katılımcılar için standart olmalıdır.
- Araştırmacıdan, ortamdan, deneklerden kaynaklanabilecek problemlere karşı dikkatli olunmalı ve gerekli önlemler alınmalıdır.
- Elektronik anketlere katılım daha fazladır fakat internete erişim sorunları, örneklem yanlılığı, analizlerin olumsuz etkilenmesi gibi risklerinin olduğu unutulmamalıdır.
- Analizlerden önce gerekli ön koşullar sağlanmalı ve ölçek maddeleri, doğru şekilde kodlanmış olmalıdır.
- Cronbach alfa katsayısı düşükse yapılan analizler ve ters maddelerin kodlanmasındaki hatalar kontrol edilmelidir.
- Ölçeğin güvenilirliğini ve iç tutarlılığını olumsuz etkilemesi nedeniyle Cronbach alfayı düşüren ve madde toplam puan korelasyon değeri 0,20'den az olan maddeler ölçekten çıkarılmalıdır.
- Ölçekteki derecelendirme kategorilerinin ve madde sayısının arttırılması, güvenilirliği artırır fakat cevaplayanları olumsuz etkileyebilir. Bu nedenle makul sayıda madde ve ideal olarak beşli ya da yedili likert derecelendirmeler kullanılmalıdır.
- Örneklem, büyüdükçe yapılacak analizlerin gücü artar ve hata olasılığı azalır. Bu nedenle olabildiğince büyük bir örnekleme (en az 200 kişi) çalışılmalıdır.
- Araştırmacılar, örneklem sayısının analizler ve model uyum indeksleri üzerine olan etkilerini bilmeli ve gerekli önlemleri almalıdır.

SONUÇ

Araştırmacıların; psikometrik açıdan kaliteli likert tipi ölçekleri seçebilme, modifiye etme, uyarlama ve geliştirme becerisine sahibi olması beklenmektedir. Fakat lisans ve lisansüstü eğitimlerde bu konulara yeterince yer verilmemektedir. Mevcut kaynaklarda bu konuların ayrı ayrı ele alınması, standardizasyonu sağlayacak rehberlerin, kullanıcı dostu kaynakların ve denetim mekanizmalarının yetersiz olması, bu alanlardaki çalışmaların metodolojik kalitesini düşürmektedir. Pek çok araştırmacı, yeni bir ölçek geliştirmek yerine literatürdeki ölçekleri kullanmayı ya da farklı bir kültür için geliştirilmiş ölçekleri kendi kültürlerine uyarlamayı tercih etmektedir. Bu işlemler daha ekonomiktir ve zaman kayıplarını azaltmaktadır. Fakat özensiz bir şekilde yapıldıklarında önemli sorunları da beraberinde getirmektedir. Araştırmaya, kültüre ya da hedef gruba uygun olmayan önyargılı, güvenilir ve yanlış veriler üreten ölçekler, araştırmacıları ve diğer bireyleri olumsuz etkilemektedir. Bu sorunların azaltılabilmesi için deneyimli kişilerden yardım alınması, kullanıcı dostu kaynakların ve denetim mekanizmalarının arttırılması gerekmektedir. Bu derlemede ölçek kullanma, modifiye etme, uyarlama ve geliştirme süreçleri bütüncül bir yaklaşımla ele alınmıştır. Makaledeki bilgi ve tablolar likert tipi ölçeklerle çalışan araştırmacıların daha doğru kararlar almasına, daha kaliteli ölçekleri kullanmasına ve üretmesine katkıda bulunabilir.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Fikir – T.Y., F.G.; Tasarım –F.G., T.Y.; Denetleme – T.Y.; Kaynaklar – F.G., T.Y.; Veri Toplanması ve/veya İşlenmesi – F.G., T.Y.; Analiz

ve/veya Yorum – F.G., T.Y.; Literatür Taraması – F.G.; Yazıyı Yazan – F.G.; Eleştirel İnceleme – T.Y., F.G.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemişlerdir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadıklarını beyan etmişlerdir.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – T.Y., F.G.; Design – F.G., T.Y.; Supervision – T.Y.; Resources – F.G., T.Y.; Materials – F.G.; Data Collection and/or Processing – F.G., T.Y.; Analysis and/or Interpretation – F.G., T.Y.; Literature Search – F.G.; Writing Manuscript – F.G.; Critical Review – T.Y., F.G.

Declaration of Interests: The authors declare that they have no competing interest.

Funding: The authors declared that this study has received no financial support.

KAYNAKLAR

- Özdamar K. *Eğitim, Sağlık ve Davranış Bilimlerinde Ölçek ve Test Geliştirme: Yapısal Eşitlik Modellemesi*. Eskişehir: Sözkese Matbaacılık, Nisan Kitabevi ; 2017.
- Kishore K, Jaswal V, Kulkarni V, De D. Practical guidelines to develop and evaluate a questionnaire. *Indian Dermatol Online J*. 2021;12(2):266-275. [CrossRef]
- McKim C. Using the literature to create a scale: an innovative qualitative methodological piece. *Int J Soc Res Methodol*. 2022;1-9. [CrossRef]
- Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Front Public Health*. 2018;6(149):149. [CrossRef]
- Carpenter S. Ten steps in scale development and reporting: A guide for researchers. *Commun Methods Meas*. 2018;12(1):25-44. [CrossRef]
- Esin MN. Veri toplama yöntem ve araçları & veri toplama araçlarının güvenilirlik ve geçerliliği. İçinde: Erdoğan S, Nahcivan N, Esin MN, Çev. 1. Baskı. *Hemşirelikte Araştırma Süreci, Uygulama ve Kritik* (ss. 193-231). İstanbul: Nobel Tıp Kitapevleri; 2014.
- DeVellis RF. *Scale Development: Theory and Applications*. 4th ed. London: SAGE Publications; 2017.
- Özdemir Z. Sağlık bilimlerinde Likert tipi tutum ölçeği geliştirme. *Hacettepe Univ Hemşirelik Fak Derg*. 2018;5(1):60-68. [CrossRef]
- Tavşancıl E. *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. 6. Baskı. Ankara: Nobel Akademik Yayıncılık; 2019.
- Seçer İ. *Psikolojik test geliştirme ve uyarlama süreci SPSS ve LISREL Uygulamaları*. 2. Baskı. Ankara: Anı Yayıncılık; 2018.
- Karaçam Z. Ölçme araçlarının Türkçeye uyarlanması. *J Midwif Health Sci*. 2019;2(1):28-37.
- Güler G, Ayan C. A Review of attitude scales developed in Turkey between 2002-2018 regarding the scale development process. *Ank Univ J Fac Educ Sci*. 2020;53(3):839-863. [CrossRef]
- Güvendir MA, Özkan YÖ. Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Electron J Soc Sci*. 2015;14(5):23-33.
- Koçar H. Kısa form geliştirme: yöntemler, öneriler ve incelemeler. *Eğit Psikol Ölçme Deger Derg*. 2020;113:302-310.
- Karakoç FY, Dönmez PDL. Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitimi Dünyası*. 2014;13(40):39-49. [CrossRef]
- Yurdabakan İ, Çüm S. Davranış bilimlerinde ölçek geliştirme (Açıklayıcı faktör analizine dayalı). *J Fam Med Prim Care*. 2017;11(2):108-126.
- Karagöz Y, Bardakçı S. *Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme*. Ankara: Nobel Akademik Yayıncılık; 2020.
- Erkuş A. *Psikolojide ölçme ve ölçek Geliştirme-I: temel Kavramlar ve İşlemler*. 5. Baskı. Ankara: Pegem Akademi Yayınevi; 2021.
- Çapık C, Gözüm S, Aksayan S. Kültürlerarası ölçek uyarlama aşamaları, dil ve kültür uyarlaması: güncellenmiş rehber. *Florence Nightingale J Nurs*. 2018;26:199-210.
- Hernández A, Hidalgo MD, Hambleton RK, Gómez-Benito J. International test commission guidelines for test adaptation: a criterion checklist. *Psicothema*. 2020;32(3):390-398. [CrossRef]
- Dirlik EM, Altıntaş Ö, Kartal SK. Uluslararası test standartlarının ve yönergelerinin sistematik ve karşılaştırmalı incelenmesi. *Opus-Uluslararası Toplum Araştırmaları Derg*. 2021;18(39):804-830.
- Gagnier JJ, Lai J, Mokkink LB, Terwee CB. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res*. 2021;30(8):2197-2218. [CrossRef]
- Robinson MA. Using multi-item psychometric scales for research and practice in human resource management. *Hum Resour Manage*. 2018;57(3):739-750. [CrossRef]
- Çokluk Ö, Şekercioğlu G, Büyüköztürk Ş. *Sosyal Bilimler için Çok Değişkenli İstatistik SPSS ve LISREL Uygulamaları*. 5. Baskı. Ankara: Pegem Akademi Yayıncılık; 2018.
- Field A. *Discovering Statistics Using IBM SPSS Statistics*. 5th ed. London: SAGE Publications; 2018.
- Kline RB. *Principles and Practice of Structural Equation Modeling*. 4th ed. New York: Guilford Publications; 2016.
- Orçan F. Açıklayıcı ve doğrulayıcı faktör analizi: İlk hangisi kullanılmalı? *Eğit Psikol Ölçme Deger Derg*. 2018;9(4):414-421.
- Gürbüz S, Şahin F. *Sosyal Bilimlerde Araştırma Yöntemleri Felsefe-Yöntem-Analiz*. 5. Baskı. Ankara: Seçkin Yayıncılık; 2018.
- Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. Scale development: ten main limitations and recommendations to improve future research practices. *Psicol Reflex Crit*. 2017;30(1):3. [CrossRef]
- Hair JF, Black WC, Babin BJ, Anderson RE. *Multivariate Data Analysis*. 9th ed. UK: Cengage Learning EMEA; 2018.
- Koyuncu İ, Kılıç AF. Açıklayıcı ve doğrulayıcı faktör analizlerinin kullanımı: bir doküman incelemesi. *Eğitim Bilim*. 2019;44(198):361-388.
- Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 5th ed. USA: Oxford University Press; 2015.
- Yusoff MSB, Arifin WN, Hadie SNH. ABC of questionnaire development and validation for survey research. *EIMJ*. 2021;13(1):97-108. [CrossRef]
- Gökdemir F, Yaşanan Uykü G. *Sorunlarını ve Uykü Sorunları ile Başa Çıkma Yöntemlerini Belirleme Ölçeklerinin Geliştirilmesi* (Yayımlanmamış Doktora Tezi) (Yılmaz T, tez Danışmanı). İstanbul: İstanb Univ-Cerrahpaşa Lisansüstü Eğitim Enstitüsü Ebelik Anabilim Dalı. 2021.
- Zhou Y. A mixed methods model of scale development and validation analysis. *Meas Interdiscip Res Perspect*. 2019;17(1):38-47. [CrossRef]
- Rosellini AJ, Brown TA. Developing and validating clinical questionnaires. *Annu Rev Clin Psychol*. 2021;17:55-81. [CrossRef]
- Yusoff MSB. ABC of content validation and content validity index calculation. *EIMJ*. 2019;11(2):49-54. [CrossRef]
- Lawshé CH. A quantitative approach to content validity. *Pers Psychol*. 1975;28(4):563-575. [CrossRef]
- Davis LL. Instrument review: getting the most from a panel of experts. *Appl Nurs Res*. 1992;5(4):194-197. [CrossRef]
- Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. *Niger Postgrad Med J*. 2015;22(4):195-201. [CrossRef]
- Mohajan HK. Two criteria for good measurements in research: validity and reliability. *ASHU- ES*. 2017;17(4):59-82. [CrossRef]
- Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiol Serv Saude*. 2017;26(3):649-659. [CrossRef]
- Sürücü L, Maslakçı A. Validity and reliability in quantitative research. *Bus Manag Stud An Int J*. 2020;8(3):2694-2726. [CrossRef]

44. Alpar R. *Uygulamalı İstatistik ve Geçerlik-Güvenirlilik: Spor, Sağlık ve Eğitim Bilimlerinden Örneklerle*. 5. Baskı. Ankara: Detay Yayıncılık; 2018.
45. Bursal M. *SPSS ile Temel Veri Analizleri*. 2. Baskı. Ankara: Anı Yayıncılık; 2019.
46. Büyüköztürk Ş. *Sosyal Bilimler İçin Veri Analizi El Kitabı. İstatistik, Araştırma Deseni SPSS Uygulamaları ve Yorum*. 27. Baskı. Ankara: Pegem Akademi Yayıncılık; 2020.
47. Hájjar ST. Statistical analysis: internal-consistency reliability and construct validity. *Int J Quant Qual Res Methods*. 2018;6(1):46-57.
48. İlhan M, Çetin B. LISREL ve AMOS programları kullanılarak gerçekleştirilen yapısal Eşitlik modeli (YEM) analizlerine ilişkin sonuçların karşılaştırılması. *Eğit Psikol Ölçme Deger Derg*. 2014;5(2):26-42.
49. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med*. 2018;18(3):91-93. [\[CrossRef\]](#)
50. Kılıç S. Cronbach'ın Alpha güvenirlilik katsayısı. *J Mood Disord*. 2016;6(1):47-48. [\[CrossRef\]](#)