

ADIM ADIM ÖLÇEK GELİŞTİRME: TEORİKTEN PRATİĞE

Yazarlar

Prof.Dr.İbrahim KORUK

Doç.Dr.Burcu BEYAZGÜL

Dr.Öğr.Üyesi Ufuk ACAR



© Copyright 2025

Bu kitabın, basım, yayın ve satış hakları Akademisyen Kitabevi A.Ş.'ne aittir. Anılan kuruluşun izni alınmadan kitabın tümü ya da bölümleri mekanik, elektronik, fotokopi, manyetik kâğıt ve/veya başka yöntemlerle çoğaltılamaz, basılamaz, dağıtılamaz. Tablo, şekil ve grafikler izin alınmadan, ticari amaçlı kullanılamaz. Bu kitap T.C. Kültür Bakanlığı bandrolü ile satılmaktadır.

Bu kitapta kullanılan SPSS ve AMOS yazılımlarına ait görsel arayüzler, IBM tarafından sunulan deneme (trial) sürümler aracılığıyla elde edilmiştir. Görseller yalnızca eğitim ve bilimsel amaçlı olarak kullanılmıştır. SPSS ve AMOS, IBM Corporation'a ait tescilli ticari markalardır. Bu yayının içeriği, IBM ile herhangi bir doğrudan ya da dolaylı iş birliği, sponsorluk veya onay ilişkisi içinde değildir. Kitapta yer alan tüm yazılım görselleri, yazar tarafından bilimsel anlatımı desteklemek amacıyla hazırlanmış ve sunulmuştur.

ISBN

978-625-375-618-5

Yayın Koordinatörü

Yasin DİLMEN

Kitap Adı

Adım Adım Ölçek Geliştirme:
Teorikten Pratiğe

Sayfa ve Kapak Tasarımı

Akademisyen Dizgi Ünitesi

Yazarlar

İbrahim KORUK

ORCID iD: 0000-0001-9564-2214

Burcu BEYAZGÜL

ORCID iD: 0000-0002-0417-3588

Ufuk ACAR

ORCID iD: 0000-0002-2663-8307

Yayıncı Sertifika No

47518

Baskı ve Cilt

Vadi Matbaacılık

Bisac Code

MED108000

DOI

10.37609/akya.3792

Kütüphane Kimlik Kartı

Koruk, İbrahim ...[ve başkaları].

Adım Adım Ölçek Geliştirme: Teorikten Pratiğe / İbrahim Koruk, Burcu Beyazgül, Ufuk Acar.

Ankara : Akademisyen Yayınevi Kitabevi, 2025.

226 s. : şekil, tablo. ; 160x235 mm.

Kaynakça var.

ISBN 9786253756185

GENEL DAĞITIM

Akademisyen Kitabevi A.Ş.

Halk Sokak 5 / A Yenışehir / Ankara

Tel: 0312 431 16 33

siparis@akademisyen.com

www.akademisyen.com

ÖNSÖZ

1400’lerde Galileo Galilei’nin “**Ölçebiliyorsan ölç, ölçemiyorsan ölçülecek hale getir**” sözleri, bilimsel gelişmede ölçmenin ne denli merkezi bir yere sahip olduğunu çarpıcı biçimde ortaya koymuştur. Bugün geldiğimiz noktada ise yalnızca ölçmeyi değil, **ölçemediğimizi nasıl ölçebiliriz** sorusunu sormaktayız.

Modern bilimde özellikle sağlık alanında bireylerin **algılarını, tutumlarını, duygularını ve kabullerini** anlamaya yönelik araştırmalar giderek artmaktadır. Sayamayacağımız, tartamayacağımız ya da doğrudan gözlemleyemeyeceğimiz bu soyut özelliklerin değerlendirilmesinde **ölçekler**, tüm sınırlılıklarına rağmen en güçlü araçlarımızdan biri haline gelmiştir.

Bu kitabın yazım süreci de tam bu noktada doğdu: Bizler, sağlık alanında çalışan araştırmacılar olarak ölçek geliştirme sürecinde çok sayıda **soru** ve **zorluk** ile karşılaştık. Var olan kapsamlı literatüre rağmen, özellikle uygulamaya dönük adımlarda cevap bulmakta ve süreci başlatmakta zaman zaman zorlandık. Bu deneyim, bizi, kendi yaşadığımız güçlükleri başkalarının daha kolay aşabilmesine katkı sunacak bir rehber hazırlamaya yöneltti.

Elinizdeki bu kitap, özellikle **ölçek geliştirmeye yeni başlayan araştırmacılar için** hem teorik bilgiyi hem de uygulama adımlarını bir araya getiren, kolay anlaşılır ve sistemli bir kaynak olarak tasarlandı. Kitapta; ölçek nedir sorusundan başlayarak, ihtiyaç belirleme, madde yazımı, pilot uygulama, açıklayıcı ve doğrulayıcı faktör analizleri, kesme noktalarının belirlenmesi gibi temel adımlar detaylı biçimde ele alınmaktadır.

Kitabın en dikkat çekici özelliği, **uygulama adımlarının bilgisayar programı ekran görüntüleriyle desteklenmiş olmasıdır**. Bu yönüyle kitap, yalnızca ne yapılması gerektiğini anlatmakla kalmaz, **nasil yapılacağını da gösterir**. Böylece okuyucular, öğrendikleri bilgileri doğrudan uygulamaya geçirebilirler.

Kitap içeriği, bir yandan bu alanda yazılmış pek çok kıymetli kaynaktan beslenirken, diğer yandan ilk kez ölçek geliştirme deneyimi yaşayacak okuyucular için süreci sadeleştiren örneklerle güçlendirilmiştir. Bu yönüyle kitabın, literatürde eksik kalan noktaları tamamlayıcı nitelikte bazı “ilkleri” de barındırdığını söyleyebiliriz.

Yazar olarak yer almaktan onur duyduğumuz bu kitabın, bilimsel araştırmalara katkı sunmasını ve yeni başlayanlara cesaret vermesini temenni ediyoruz.

Yazarlar

İÇİNDEKİLER

BÖLÜM 1	Ölçek Nedir? Neden Ölçmek İsteriz?.....	1
BÖLÜM 2	Ölçekte Bulunması Gereken Özellikler	11
BÖLÜM 3	İhtiyacın Belirlenmesi.....	25
BÖLÜM 4	Literatür Taraması ve Teorik Çerçevenin Oluşturulması.....	35
BÖLÜM 5	Madde Havuzunun Oluşturulması	45
BÖLÜM 6	Aday Ölçeğe İlk Şeklini Verme	59
BÖLÜM 7	Aday Ölçeğin Denenmesi	71
BÖLÜM 8	Açıklayıcı Faktör Analizi ve Boyutların Oluşturulması	105
BÖLÜM 9	Doğrulamalı Faktör Analizi ile Yapının Sınanması	137
BÖLÜM 10	Ölçek Puanını 'Değerleme'.....	205

BÖLÜM 1

ÖLÇEK NEDİR? NEDEN ÖLÇMEK İSTERİZ?

1. GİRİŞ

Ölçme, akademik araştırmaların temel taşlarından biridir. Bir olguyu bilimsel olarak inceleyebilmek ve onun hakkında geçerli ve güvenilir bilgilere ulaşabilmek büyük ölçüde ölçmenin başarısına bağlıdır (1). Ölçme, en yalın haliyle, belirli bir özelliğe sistematik ve nesnel kurallar çerçevesinde semboller veya sayılar atama işlemi olarak nitelendirilebilir (2). Bu tanım, sosyal bilimlerden sağlığa, mühendislikten eğitime kadar pek çok disiplin için geçerli bir çerçeve sunar.

Günümüzde, özellikle sağlık bilimleri ve sosyal bilimlerde araştırmalar hız kazandıkça, karmaşık kavramların (örneğin tutum, kaygı, yaşam kalitesi) ölçülmesi de ayrı bir önem kazanmıştır (3). Fiziksel ölçümlerde (boy, ağırlık vb.) kullanılan araçların kalibrasyon ve standardizasyonu daha kolay sağlanırken, psikolojik veya sosyal kavramların ölçümü belirli ilkeler çerçevesinde geliştirilen ölçeklerle yapılmaktadır (4). Ancak “ölçek” kavramı, yalnızca psikometrik araçları değil, her türlü sistematik ölçüm aracını ifade edebilir. Önemli olan, ölçülmek istenen özelliğin iyi tanımlanması ve uygun yöntemlerle sayısal olarak temsil edilmesidir.

Bu bölümde, “Ölçek nedir ve neden ölçmek isteriz?” soruları teorik çerçevede ele alınacaktır. Tüm disiplinler için ortak bir kavramsal temel sunmak amacıyla ölçek geliştirme yaklaşımlarına yer verilecektir. Bu şekilde araştırmacıların ölçek geliştirme süreçlerine başlamadan önce ihtiyaç duyacakları kuramsal altyapıya sahip olmaları hedeflenmektedir.

2. ÖLÇME KAVRAMI

Ölçme, bir niteliğe veya özelliğe sistematik biçimde sayı ya da semboller atama süreci olarak tanımlanır (1). Bu tanım, ölçmeyi herhangi bir özelliğin miktarını,

ele alınır (1). Yapı geçerliği için faktör analizi (açıklayıcı ve doğrulayıcı) kritik önem taşır. Açıklayıcı faktör analizi, ölçeğin ölçtüğü boyutları veya faktörleri ortaya koyarken, doğrulayıcı faktör analizi kuramsal olarak öngörülen yapının verilerle ne ölçüde uyumlu olduğunu test eder (8). Tüm bu yöntemler, kuramsal olarak tanımlanan bir özelliğin gerçekten ölçülüp ölçülmediğini ortaya koymayı amaçlar. Bu da ölçme sürecini ampirik ve nesnel hale getirir (9).

10. BÖLÜM ÖZETİ

Ölçek, ölçülmek istenen bir özelliği sistematik ve nesnel biçimde sayısal verilere dönüştüren bir araçtır (2). Özellikle sağlık bilimleri ve sosyal bilimlerde, soyut kabul edilen pek çok kavram (tutum, kaygı, depresyon, yaşam kalitesi vb.) ancak ölçekler yoluyla nesnelleştirilerek araştırılabilir hale gelir (3).

Bütün disiplinlerde ölçmenin alt yapısı benzerdir: Geçerli ve güvenilir verilere ihtiyaç duyulur ve bu veriler, standart ölçüm araçları aracılığıyla elde edilir (7). Ancak psikometrik ölçümlerde “gerçek değerin” tam olarak bilinmemesi ve ölçümün dolaylı karakteri, ölçek geliştirme süreçlerinde faktör analizi, madde analizi, geçerlik ve güvenilirlik çalışmaları gibi aşamalara duyulan ihtiyacı artırır (9). Böylece sosyal ve sağlık bilimlerinde ölçme, kendine özgü yöntemsel bir çerçevede, ancak genel bilimsel ilkelerle uyumlu biçimde uygulanır.

KAYNAKLAR

1. Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Prentice Hall.
2. Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd ed.). Harper & Row.
3. DeVellis, R. F. (2016). *Scale Development: Theory and Applications* (4th ed.). SAGE.
4. Kline, T. (2005). *Psychological Testing: A Practical Approach*. SAGE.
5. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.
6. Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684), 677-680. <https://doi.org/10.1126/science.103.2684.677>
7. Tezbaşaran, A. A. (1997). Likert tipi ölçek geliştirme kılavuzu. *Türk Psikologlar Derneği*.
8. Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.
9. Büyüköztürk, Ş. (2007). *Sosyal Bilimler İçin Veri Analizi El Kitabı* (7. baskı). Pegem A Yayıncılık.
10. Karasar, N. (2018). *Bilimsel Araştırma Yöntemi* (32. baskı). Nobel Yayıncılık.
11. Binet, A., & Simon, T. (1905). *Methods Nouvelles Pour Le Diagnostic du Niveau Intellectuel Des Anormaux* [New Methods for the Diagnosis of the Intellectual Level of Subnormals]. *L'Année Psychologique*, 12(1), 191-244.
12. Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22(140), 1-55.

BÖLÜM 2

ÖLÇEKTE BULUNMASI GEREKEN ÖZELLİKLER

1. GİRİŞ

Tüm disiplinler göz önüne alındığında ölçeklerin özellikle psikolojik ve davranışsal değişkenlerin ölçümünde kritik rol oynadığı görülmüştür (1, 2). Ancak bir ölçeğin, araştırmalarda veya uygulamalarda bilimsel anlamda yararlı olabilecek sonuçlar üretmesi için belirli niteliksel ve teknik standartları karşılaması gerekir. Bu standartlar, ayrıca ölçeğin bilimsel kalitesini ve uygulamadaki geçerliliğini doğrudan belirleyen özelliklerdir.

Ölçekte bulunması gereken temel özellikler arasında başta geçerlik, güvenirlik, kullanılabilirlik, objektiflik ve standardizasyon gibi faktörler yer alır (3). Ayrıca, ölçeğin duyarlılığı, ayırt edicilik düzeyi, pratik uygulama kolaylığı, puanlama ve yorumlama yönergelerinin açıklığı gibi konular da ölçeğin kalitesini etkileyen önemli unsurlardır (4). Bu bölümde, ölçeğin niteliklerini sistematik biçimde ele alarak, araştırmacı ve uygulamacılara yol gösterici bir çerçeve sunulması amaçlanmaktadır.

Aşağıdaki yer alan alt başlıklarda, ölçek niteliğiyle ilgili temel yaklaşımlar, geçerlik ve güvenirlik kavramları derinlemesine incelenecek; ölçek standardizasyonu, duyarlılık, pratiklik ve diğer kritik unsurların kavramsal çerçevesi detaylı olarak ele alınacaktır. Burada sunulan çerçeve, ölçek geliştirme sürecini daha etkin kılmayı ve pratik uygulama bölümlerinin kolay anlaşılır olmasını hedeflemektedir (5).

2. GEÇERLİK (VALIDITY)

2.1. Geçerlik kavramı ve önemi

Geçerlik (validity), ölçeğin ölçmek istediği yapıyı ne derece doğru ölçtüğünü ifade eden en kritik özelliktir (6). Örneğin bir anksiyete ölçeği, gerçekten “anksiye-

Dolayısıyla, ölçek geliştirme ve uyarlama süreçlerinde bu özelliklerin bütüncül biçimde ele alınması şarttır. Birinde meydana gelecek eksiklik, diğerlerini de olumsuz etkileyecek ve sonuçta ölçeğin bilimsel değeri azalacaktır (9).

13. BÖLÜM ÖZETİ

Bu bölümde, ölçekte bulunması gereken özellikler geniş bir kuramsal çerçevede incelenmiştir. Bir ölçeğin bilimsel araştırmalarda veya klinik uygulamalarda etkili bir araç olabilmesi için geçerlik, güvenirlik, kullanılabilirlik, objektiflik, duyarlılık ve standardizasyon gibi temel nitelikleri karşılaması gerekir (5). Ayrıca, kültürel uyarlama, etik sorumluluklar, norm çalışmaları ve kesme noktalarının belirlenmesi gibi uygulamaya yönelik hususlar da ölçek niteliğine doğrudan etki eder (16).

Araştırmacılar, ölçek geliştirme sürecinde bu özellikleri göz önünde bulundurarak kapsamlı bir yöntem izlediklerinde, elde edilen ölçüm sonuçlarının bilimselliği, geçerliliği ve uygulanabilirliği artacaktır (1). Böylece, ele alınan konunun (Örneğin: kaygı, depresyon, yaşam kalitesi, tutumlar, öz-yeterlik gibi) daha iyi anlaşılmasına ve ihtiyaç duyulan müdahalelerin planlanmasına katkı sağlanacaktır (9).

Sonraki bölümlerde, ihtiyacın belirlenmesi, literatür taraması ve teorik çerçevenin oluşturulması, madde havuzunun oluşturulması, pilot uygulama ve madde seçimi, faktör analizleri (açıklayıcı ve doğrulayıcı) ve kesme noktası belirlenmesi gibi daha spesifik adımları detaylarıyla ele alacağız. Ölçekte bulunması gereken özellikler, söz konusu adımların her birinde dikkate alınarak ölçeğin kalitesi inşa edilecek ve son hale getirilecektir (3).

KAYNAKLAR

1. Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Prentice Hall.
2. DeVellis, R. F. (2016). *Scale Development: Theory and Applications* (4th ed.). SAGE.
3. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.
4. Kline, T. (2005). *Psychological Testing: A Practical Approach*. SAGE.
5. Cronbach, L. J. (1970). *Essentials of Psychological Testing* (3rd ed.). Harper & Row.
6. Stevens, S. S. (1946). On the Theory of Scales of Measurement. *Science*, 103(2684), 677-680. <https://doi.org/10.1126/science.103.2684.677>
7. Tezbaşaran, A. A. (1997). Likert tipi ölçek geliştirme kılavuzu. *Türk Psikologlar Derneği*.
8. Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.
9. Büyüköztürk, Ş. (2007). *Sosyal Bilimler İçin Veri Analizi El Kitabı* (7. baskı). Pegem A Yayıncılık.
10. Karasar, N. (2018). *Bilimsel Araştırma Yöntemi* (32. baskı). Nobel Yayıncılık.

11. Campbell, D. T., & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56(2), 81-105.
12. Likert, R. (1932). A Technique for the Measurement of Attitudes. *Archives of Psychology*, 22(140), 1-55.
13. American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. AERA.
14. Lord, F. M., & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*. Addison-Wesley.
15. Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological Testing and Assessment: An Introduction to Tests and Measurement* (9th ed.). McGraw-Hill.
16. Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.

BÖLÜM 3

İHTİYACIN BELİRLENMESİ

1. GİRİŞ

Ölçek geliştirme sürecinin başlangıcında, hangi alanda ve hangi amaca yönelik bir ölçme aracına ihtiyaç duyulduğunu netleştirmek hayati önem taşır. “İhtiyacın Belirlenmesi” adı verilen bu aşama, hem ölçeğin amaç ve kapsamını şekillendirir hem de diğer aşamalarda izlenecek yöntem ve teknikleri belirler (1). Özellikle sağlık ve sosyal bilimler gibi geniş kapsamlı ve sürekli değişim hâlindeki alanlarda, ihtiyaç belirleme sürecinin sistematik ve çok boyutlu yaklaşımı zorunludur (2).

Bu bağlamda, ihtiyaç belirleme süreci yalnızca “ölçeğin hangi özelliği veya kavramı ölçeceğini” saptamakla sınırlı değildir. Aynı zamanda bu kavramın hangi popülasyonda, hangi kuramsal çerçevede ve hangi yöntemsel yaklaşımlarla ele alınacağına ilişkin stratejik kararlara da zemin hazırlar (3). Aşağıdaki bölümlerde, ihtiyaç belirleme kavramının tanımı, süreçte kullanılan yöntemler, kuramsal çerçeve oluşturmanın önemi, paydaş analizi, etik ve metodolojik ilkeler, saha uygulamaları ve nihayetinde sonuç ile genel değerlendirme başlıkları detaylı olarak ele alınacaktır.

2. İHTİYACIN BELİRLENMESİ KAVRAMI VE ÖNEMİ

2.1. Tanım ve kapsam

İhtiyacın belirlenmesi, temel olarak “hangi probleme ya da alana yönelik bir ölçme aracına gereksinim duyulduğunu” tanımlamayı içerir (4). Örneğin bir araştırmacı, hasta memnuniyetini ölçmek isterken başka bir araştırmacı, örgütsel bağlılık ya da sosyal kaygı düzeyini değerlendiriyor olabilir. Her iki durumda da

KAYNAKLAR

1. DeVellis, R. F. (2017). *Scale Development: Theory and Applications* (4th ed.). Thousand Oaks, CA: SAGE.
2. Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales in Health, Social and Behavioral Research: A Primer. *Frontiers in Public Health*, 6, 149.
3. Crocker, L., & Algina, J. (2006). *Introduction to Classical and Modern Test Theory*. Mason, OH: Cengage Learning.
4. Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: SAGE.
5. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.
6. Büyükköztürk, Ş. (2020). *Sosyal Bilimler İçin Veri Analizi El Kitabı* (27. baskı). Ankara: Pegem Akademi.
7. Fowler Jr, F. J. (2014). *Survey Research Methods* (5th ed.). Thousand Oaks, CA: SAGE.
8. Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York: The Guilford Press.
9. Polit, D. F., & Beck, C. T. (2017). *Nursing Research: Generating and Assessing Evidence for Nursing Practice* (10th ed.). Philadelphia, PA: Wolters Kluwer.
10. Maxwell, J. A. (2013). *Qualitative Research Design: An Interactive Approach* (3rd ed.). Thousand Oaks, CA: SAGE.
11. Karasar, N. (2021). *Bilimsel Araştırma Yöntemi* (35. baskı). Ankara: Nobel Yayıncılık.
12. Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales: A Practical Guide to Their Development and Use* (5th ed.). Oxford: Oxford University Press.
13. Sandelowski, M. (2000). Focus on Research Methods: Whatever Happened to Qualitative Description? *Research in Nursing & Health*, 23(4), 334–340.
14. Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford: Oxford University Press.
15. Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research* (3rd ed.). Thousand Oaks, CA: SAGE.
16. WHO (World Health Organization). (2011). *World Report on Disability*. Cenevre: WHO Press.
17. Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Boston: Pearson.
18. Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods* (4th ed.). Thousand Oaks, CA: SAGE.
19. Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE Handbook of Qualitative Research* (5th ed.). Thousand Oaks, CA: SAGE.
20. Griffiths, M. D., Kuss, D. J., & Pontes, H. M. (2016). A Brief Overview of Internet Gaming Disorder and its Treatment. *Australian Clinical Psychologist*, 2(1), 20108.
21. Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9(2), 27–40.
22. Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14–26.
23. Tashakkori, A., & Teddlie, C. (Eds.). (2010). *SAGE Handbook of Mixed Methods in Social & Behavioral Research* (2nd ed.). Thousand Oaks, CA: SAGE.
24. Bryson, J. M. (2011). *Strategic Planning for Public and Nonprofit Organizations* (4th ed.). San Francisco, CA: Jossey-Bass.
25. Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative Data Analysis: A Methods*

- Sourcebook (3rd ed.). Thousand Oaks, CA: SAGE.
26. Freeman, R. E. (2010). *Strategic Management: A Stakeholder Approach*. Cambridge: Cambridge University Press.
 27. Beck, C. T. (2013). *Routledge International Handbook of Qualitative Nursing Research*. New York: Routledge.
 28. Greenwood, M. (2007). Stakeholder Engagement: Beyond the Myth of Corporate Responsibility. *Journal of Business Ethics*, 74(4), 315–327.
 29. Orb, A., Eisenhauer, L., & Wynaden, D. (2000). Ethics in Qualitative Research. *Journal of Nursing Scholarship*, 33(1), 93–96.
 30. Emanuel, E. J., Wendler, D., & Grady, C. (2000). What Makes Clinical Research Ethical? *JAMA*, 283(20), 2701–2711.
 31. Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation* (4th ed.). San Francisco, CA: Jossey-Bass.
 32. Guba, E. G., & Lincoln, Y. S. (1994). Competing Paradigms in Qualitative Research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105–117). Thousand Oaks, CA: SAGE.
 33. Polit, D. F., & Beck, C. T. (2008). *Nursing Research: Generating and Assessing Evidence for Nursing Practice* (8th ed.). Philadelphia, PA: Wolters Kluwer.
 34. Teijlingen, E. R. V., & Hundley, V. (2001). The Importance of Pilot Studies. *Social Research Update*, (35). University of Surrey.
 35. Charness, N., & Boot, W. R. (2009). Aging and Information Technology Use: Potential and Barriers. *Current Directions in Psychological Science*, 18(5), 253–258.
 36. Salamon, L. M. (Ed.). (2002). *The Tools of Government: A Guide to the New Governance*. New York: Oxford University Press.
 37. Bryson, J. M., Quick, K. S., Slotterback, C. S., & Crosby, B. C. (2013). Designing Public Participation Processes. *Public Administration Review*, 73(1), 23–34.
 38. Nutbeam, D. (1998). Health Promotion Glossary. *Health Promotion International*, 13(4), 349–364.
 39. Stone, D. (2012). *Policy Paradox: The Art of Political Decision Making* (3rd ed.). New York: W.W. Norton & Company.
 40. Fawcett, J., & Garity, J. (2009). *Evaluating Research for Evidence-Based Nursing Practice*. Philadelphia, PA: F. A. Davis Company.
 41. Guthrie, G. (2010). *Basic Research Methods: An Entry to Social Science Research*. New Delhi: SAGE India.
 42. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, 25(24), 3186–3191.
 43. Baker, M. J. (2000). Writing a Research Proposal. *The Marketing Review*, 1(1), 61–75.
 44. Callahan, J. L., & Martin, D. (2007). The Spectrum of School-Community Partnerships. *Preventing School Failure: Alternative Education for Children and Youth*, 51(4), 5–10.

BÖLÜM 4

LİTERATÜR TARAMASI VE TEORİK ÇERÇEVENİN OLUŞTURULMASI

Ölçekler, merak ettiğimiz özelliğe göre bireylerin sınıflandırılmasında, toplum ya da bir grup birey içindeki yerinin belirlenmesinde ya da bireydeki özelliğin derecesinin belirlenmesinde rol alır. Bu anlamda ister sınıflandırmak, ister yerini belirlemek, isterse de derecesini belirlemek amacıyla yola çıkılsın, hepsinin özünde belirli kurallar ve standartlar olmak zorundadır (1). Bu nedenle, standartlaştırma öncesi bilmemiz gereken birçok nokta göze çarpar: Neyi, nasıl ve ne kadar ölçelim?

1. NEYİ ÖLÇÜYORUZ?

Ölçülmek istenen olay, durum, tutumun ne olduğunun net bir şekilde tanımlanması hem süreci kısaltması hem de teknik olarak ölçüme yaklaşabilmek adına önemlidir. Bu aşamada, merak ettiğimiz ve gerçekten ölçmek istediğimiz özelliğe her yönüyle hakim olmalıyız:

Ölçmek istediğimiz özelliğin tıbbi literatürde kabul edilen net bir tanımı var mı?

Merak ettiğimiz özellik, daha önce tıbbi literatürde çalışılmış ve genel kabul gören bir tanıma sahip olabilir. Bu şekilde, ilgilendiğimiz özelliğin teorik çerçevesi halihazırda çalışılmış, tanı kriterleri konmuş, ayırıcı tanıya varan dışlama kriterleri oluşturulmuş olabilir. Böyle durumlarda güncel tıbbi bilgiye sadık olarak, bu kriterler üzerinden ölçüm aracı geliştirilmeye çalışılabilir.

Örneğin: Postpartum depresyon, doğumu takiben başlayan depresif ataklarla kendini gösteren bir hastalıktır. Bu durumdaki kadınlarda gerginlik, depresif ruh hali, uyku kalitesinin bozulması, yeme ve iştah bozukluğu, sosyal işlevlerde azalma görülebilir (2). Bu şekilde net bir tanımı olan postpartum depresyonu ölçmek adına, Cox ve Holden tarafından Edinburgh doğum sonrası depresyon

şüphesi olan çocuk olarak değerlendirilir (17). Bu ölçekte kullanılan bir madde aşağıda verilmiştir:

» *Çocuğunuz sembolik oyunlar (örneğin, oyuncak araba/bebekle oynarken sanki arabayı sürerek bir yerden başka bir yere götürüyormuş gibi/bebeğini uyutuyormuş gibi) oynar mı?* (17).

a) Evet

b) Hayır

KAYNAKLAR

1. Karakoç FY ve Dönmez L. (2014). Ölçek geliştirme çalışmalarında temel ilkeler. *Tıp Eğitimi Dünyası*, 13(40), 39-49.
2. American Psychiatric Association. <https://www.psychiatry.org/patients-families/Peripartum-Depression/What-is-Peripartum-Depression> 13.08.24.
3. Cox JL ve Holden JM (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale (EPDS). *Br J Psychiatry*, 150:782-786.
4. Morgado CMDc, Werneck GL, Hasselmann MH. (2013). Social network, social support and feeding habits of infants in their fourth month of life. *Ciênc Saúde Coletiva*. 18(2):367-376.
5. Maleki-Saghooni N, Barez MA, Karimi FZ (2020). Investigation of the relationship between social support and breastfeeding self-efficacy in primiparous breastfeeding mothers, *The Journal of Maternal-Fetal & Neonatal Medicine*, 33:18, 3097-3102.
6. Boccolini CS, Carvalho ML, Oliveira MI. (2015). Factors associated with exclusive breastfeeding in the first six months of life in Brazil: A systematic review. *Rev Saude Publica*. 49:91.
7. Beyazgul B, Koruk F, Koruk I. (2024). Emzirme Döneminde Baskılayıcı Davranışlar Yaşama Durumuna Yönelik Ölçek Geliştirme Çalışması: Emzirme Baskısı Ölçeği. 14. Uluslararası İstanbul Bilimsel Çalışmalar Kongresi, İstanbul, Türkiye.
8. DeVellis RF (2022). *Scale Development Theory and Applications*. 3.Edition: Çeviri: Totan T, Nobel Akademik.
9. Özdamar K. (2016). *Ölçek ve Test Geliştirme Yapısal Eşitlik Modellemesi*. Ankara: Nisan Kitabevi Yayıncılık.
10. Kartal M. ve Bardakçı S. (2018). *SPSS ve AMOS Uygulamalı Örneklerle Güvenirlik ve Geçerlik Analizleri*. Akademisyen Kitabevi. Ankara.
11. WHO. Abuse of older people. <https://www.who.int/news-room/fact-sheets/detail/abuse-of-older-people> 14.08.2024.
12. Özçakar N, Ergönen AT, Kartal M, Baydur H. (2017). Adaptation, reliability, and validity study of the Hwalek-SengstockElder Abuse Screening Test (HS/EAST): a Turkish version. *Turkish journal of medical sciences*, 47(6), 1894-1902.
13. Kartal M. ve Bardakçı S. (2019). *Tutum Ölçekleri*. Akademisyen Kitabevi. Ankara.
14. Çamlıca H ve Dişçi, R. (2008). Tanı testlerinde sınır değerlerinin belirlenmesi. *Türk Onkoloji Dergisi*, 23(1), 26-33.
15. Koruk F, Turan Z, Beyazgül B, Barlas F, Özgen HN, Koruk İ (2023). Breastfeeding during pregnancy and tandem breastfeeding attitude scale: Instrument development and psychometric properties. 2.Uluslararası Kadın Sağlığı Hemşireliği Kongresi, Ankara, İstanbul.
16. Mikhael EM, Hassali MA, Hussain SA, Shawky N. (2019). Development and validation of a comprehensive diabetes self-management scale. *Diabetes Metab Syndr* 13(3): 1717-1721.

17. Yıkgeç A. (2005). A validity study of the Modified Checklist for Autism in Toddlers (M-C-HAT) on a Turkish sample. Yayınlanmamış Yüksek Lisans Tezi, Boğaziçi Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
18. Karagöz Y, Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
19. Dişli B, ve Şahin NH. (2022). Menopoza özgü sıcak basması ölçeği: Geçerlik ve güvenirlik çalışması. *İstanbul Gelişim Üniversitesi Sağlık Bilimleri Dergisi*, (16), 48-62.
20. Çelik AS, Pasinlioğlu T. (2015). Sıcak Basması İnanç Ölçeği'nin geçerlilik ve güvenirlik çalışması. *Sağlık Bilimleri ve Meslekleri Dergisi*. 2(3):249-259.
21. Beyazgul B, Koruk F, Koruk I. (2024). Development of a scale for attitude toward sexually transmitted infections based on the Health Belief Model. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 298, 42-48.
22. Costa MF. (2020). Health belief model for coronavirus infection risk determinants. *Rev Saude Publica*. 54:47.
23. Skinner CS, Tiro J, Champion VL. (2015). Health Behavior and Health Education: Theory, Research and Practice. 5th ed. San Francisco: Jossey-Bass.
24. Jones K, Parker EJ, Steffens MA, Logan RM, Brennan D, Jamieson LM. (2016). Development and psychometric validation of social cognitive theory scales in an oral health context. *Aust N Z J Public Health*. 40(2):193-5.
25. Erol N, Sezgin N, Savaşır I. (1993).Gelişim Tarama Envanteri Geliştirilmesi ile İlgili Geçerlik Çalışmaları. *Türk Psikiyatri Dergisi*. 8(29): 16-22.

BÖLÜM 5

MADDE HAVUZUNUN OLUŞTURULMASI

Pek çok faktörden etkilenen durumlarda, tek bir soru ya da madde ile araştırılan özelliği ortaya koymak/ölçmek genellikle mümkün olmaz. Bu durumda ölçekler, tüm faktörleri de bünyesine alarak çözüm geliştiren araçlar olarak ortaya çıkar. Araştırılan özellik, ölçek geliştirme çalışmalarında örtük ya da gizli değişken olarak adlandırılır ve tek bir soruyla cevaplanamadığından bu gizli değişkeni saptayacak başka değişkenlere ihtiyaç vardır. Araştırılan özelliği ortaya çıkartmak için kullanılan değişkenler gözlenen değişken olarak adlandırılırlar. Gözlenen değişkenlerin hepsini içeren küme ise, araştırılan özelliği ortaya çıkartmak için tasarlanan madde havuzunu oluşturur. Madde havuzu, geliştirilmek istenen ölçeğin, ölçmeyi hedeflediği kavram ya da özellik (örneğin: stres, motivasyon, sağlık davranışı, öz-yeterlik gibi) ile ilgili potansiyel ölçek maddelerinin (soruların veya düz cümle şeklindeki önerme ifadelerinin) geniş bir listesidir. Bu havuz, olabildiğince kapsamlı olacak şekilde oluşturulur ve henüz ölçekte yer alacak kesin maddeler değildir. Sorular sorarak ya da önermelerde bulunarak gözlenen değişkenler bulunmaya çalışılır (1-3). **Yukarıda bahsedilen kavramlar örneklerle açıklanacak olursa:**

» *Son 1 ay içinde, evde yaptığınız işler dışında yaklaşık olarak ne sıklıkla yüzmeye, yürüyüş, bisiklete binme gibi fiziksel egzersiz yaptınız?* (4)
Bu, gözlenen değişkene bir örnektir.

» *Hastaların diyabet hastalığını yönetme becerisi* ise, örtük (gizli) değişkendir.

Bir başka örnek;

» *Herhangi biri sizin izniniz olmadan size ait olan şeyleri aldı mı?* (5) ifadesi gözlenen değişkeni sorgulamaktadır.

» *Yaşlı istismarı* ise, bu ölçek araştırma kurgusunda örtük değişkendir.

KAYNAKLAR

- DeVellis RF. (2022). Scale Development Theory and Applications. 3.Edition: Çeviri: Totan T, Nobel Akademik.
- Özdamar K. (2016). Ölçek ve Test Geliştirme Yapısal Eşitlik Modellemesi. Ankara: Nisan Kitabevi Yayıncılık.
- Kartal M. ve Bardakçı S. (2018). SPSS ve AMOS Uygulamalı Örneklerle Güvenirlik ve Geçerlik Analizleri. Akademisyen Kitabevi. Ankara.
- Mikhael EM, Hassali MA, Hussain SA, Shawky N. (2019). Development and validation of a comprehensive diabetes self-management scale. *Diabetes Metab Syndr.* 13(3): 1717-1721.
- Özçakar N, Ergönen AT, Kartal M, Baydur H. (2017). Adaptation, reliability, and validity study of the Hwalek-SengstockElder Abuse Screening Test (HS/EAST): a Turkish version. *Turkish journal of medical sciences*, 47(6), 1894-1902.
- Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health.* 11;6:149.
- Ekici E, İkısık H, Ankaralı H, Manav G, Colak M, Kozan Cıkıncı EH. (2021). The validity and reliability of the Turkish version of the Family Nutrition and Physical Activity screening tool. *Marmara Medical Journal*, 34(3), 319-326.
- Karabulut Demir E. (2007). Ebeveyn tutum ölçeği (ETÖ) (Master's thesis, Sosyal Bilimler Enstitüsü).
- Işık A ve Şimşek GM. (2019). Ruh Sağlığı Çalışanlarının Madde Kullanım Bozukluğu Tedavisine Yönelik Tutumları Ölçeğinin Geliştirilmesi. *Kıbrıs Türk Psikiyatri ve Psikoloji Dergisi*, 1(Özel Sayı 1), 24-26.
- Koruk F, Turan Z, Beyazgül B, Barlas F, Özgen HN, Koruk İ (2023). Breastfeeding during pregnancy and tandem breastfeeding attitude scale: Instrument development and psychometric properties. 2.Uluslararası Kadın Sağlığı Hemşireliği Kongresi, Ankara, İstanbul.
- Beyazgul B, Koruk F, Koruk I. (2024). Development of a scale for attitude toward sexually transmitted infections based on the Health Belief Model. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 298, 42-48.
- Şahin DB ve Gülleroğlu HD. (2013). Likert tipi ölçeklere madde seçmede kullanılan farklı madde analizi teknikleri ile oluşturulan ölçeklerin psikometrik özelliklerinin incelenmesi. *Asya Öğretim Dergisi*, 1(2), 18-28.
- Ercan K, Akıllı M, Sezek F. (2009). Lise öğrencilerinin çevreye karşı tutumlarının cinsiyet açısından incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, (18), 43-54.
- Eskin M, Harlak H, Demirkıran F, Dereboy Ç. (2013). Algılanan Stres Ölçeğinin Türkçeye Uyarlanması: Güvenirlik ve Geçerlik Analizi. *Yeni Symposium Journal* Cilt 51, sayı 3.
- Yıkgeç A. (2005). A validity study of the Modified Checklist for Autism in Toddlers (M-C-HAT) on a Turkish sample. *Yayınlanmamış Yüksek Lisans Tezi*, Boğaziçi Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Eroğlu N ve Sabuncu N. (2019). Diyabet Öz Yönetim Skalası'nın (DÖYS) Türk toplumuna uyarlanması: geçerlik ve güvenilirlik çalışması. *Hemşirelik Bilimi Dergisi*, 1(3), 1-6.
- Kuzan R, Beyazgül B, Karadağ D. (2024). A Study on the Development of a Health Literacy Assessment Scale. *Harran Üniversitesi Tıp Fakültesi Dergisi*, 21(3), 393-401.
- Özdemir Z. (2018). Sağlık bilimlerinde likert tipi tutum ölçeği geliştirme. *Hacettepe Üniversitesi Hemşirelik Fakültesi Dergisi*, 5(1), 60-68.
- Karagöz Y, Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
- Uyumaz G ve Çokluk Ö. (2016). Likert Tipi Ölçeklerde Madde Düzeni ve Derecelendirme Farklılıklarının Psikometrik Özellikler ve Yanıtlayıcı Tutumları Açısından İncelenmesi. *Journal of Theoretical Educational Science/Kuramsal Eğitimbilim Dergisi*, 9(3).
- Sarı T ve Çakır G. (2016). Mutluluk korkusu ile öznel ve psikolojik iyi oluş arasındaki ilişkinin incelenmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 5(25), 222-229.
- Güvenir T, Özbek A, Baykara B, Arkar H, Şentürk B, İncekaş S. (2008). Güçler ve güçlükler anketi'nin (gga) Türkçe uyarlamasının psikometrik özellikleri. *Turkish Journal of Child and Adolescent Mental Health*, 15(2), 65-74.

BÖLÜM 6

ADAY ÖLÇEĞE İLK ŞEKLİNİ VERME

Ölçülmek istenen durumla ilgili madde havuzu oluşturulduktan sonra, yapılması gereken işlem aday ölçeğe ilk şeklini vermektir. Bunu yaparken; ilk teyit edilmesi gereken durum, geliştirmek istenilen ölçekte nihai olarak yer alması istenilen alt boyutları (faktörleri) temsil edecek yeterli sayıda maddenin havuzda olup olmadığıdır. Ölçek geliştirme çalışmalarında genel kurallardan birisi, her bir alt boyutu temsil eden en az 3 maddenin ölçekte yer almasıdır (1,2). Fakat, bu kural ölçeğin tamamlanmış ve kullanılabilecek duruma gelmiş son hali için geçerlidir. Bu durumda ölçekte yer alan her bir alt boyutu ölçecek en az 3 maddenin yer alabilmesi için, ölçek geliştirme aşamasında olası madde atılmalarını da öngörerek, daha fazla sayıda maddenin havuzda yer alması sağlanmalıdır. Bu yaklaşıma paralel olarak, olası madde düzeltmeleri ve madde atılmaları dikkate alınarak, sonuçta hedeflenen madde sayısının en az 4 katı madde ile yola çıkılması yaygın kabullerden biridir (1). Yine de örtük değişken ya da alt boyutlar için mümkün olabilecek tüm maddeleri eklediğini düşündüğü halde, yeterli madde sayısına ulaşamayan bir araştırmacının 2 katı ile yetinilebileceğini de belirten görüşler mevcuttur (3).

Hem uzman değerlendirmeleri sonrasında, hem 2.pilot uygulama sonrasında, hem de Açıklayıcı Faktör Analizi (AFA) aşamasında aday ölçekten madde atmak gerekebilir. Bu aşamalarda bazen o kadar çok madde atılmak zorunda kalınır ki, ilk planlanan alt boyutları temsil edecek sayıda madde kalmayabilir. Hatta bazen, tüm atılan maddeler belli alt boyutlarda yığılabilir ve ilgili alt boyutu ölçecek madde kalmayabilir. Bu gibi durumlarda araştırmacı üç alternatif yolu izleyebilir. Bunlar şöyle sıralanabilir:

- » Birincisi, maddesi eksilen alt boyutları tamamlamak için bir nevi başa dönüp ölçeği besleyebilir. Bu amaçla, yeniden literatür taraması yapabilir, uzmanlardan fikir alabilir ya da ölçeğin hedef grubu ile odak grup görüşmeleri yapabilir.

KAYNAKLAR

1. Özdamar K. (2016). Ölçek ve Test Geliştirme Yapısal Eşitlik Modellemesi. Ankara: Nisan Kitabevi Yayıncılık.
2. Kartal M. ve Bardakçı S. (2018). SPSS ve AMOS Uygulamalı Örneklerle Güvenirlik ve Geçerlik Analizleri. Akademisyen Kitabevi. Ankara.
3. DeVellis RF. (2022). Scale Development Theory and Applications. 3.Edition: Çeviri: Totan T., Nobel Akademik.
4. Özdemir Z. (2018). Sağlık bilimlerinde likert tipi tutum ölçeği geliştirme. Hacettepe Üniversitesi Hemşirelik Fakültesi Dergisi, 5(1), 60-68.
5. Jebb AT, Ng V, Tay L. (2021). A Review of Key Likert Scale Development Advances: 1995-2019. Front Psychol. 4; 12:637547.
6. Yeşilyurt S ve Çapraz C. (2018). Ölçek geliştirme çalışmalarında kullanılan kapsam geçerliği için bir yol haritası. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 20(1), 251-264.
7. Ayre C, Scally AJ. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and evaluation in counseling and development*. 47(1):79-86.
8. Lawshe CH. (1975). A quantitative approach to content validity. *Personnel psychology*. 28(4):563-75.
9. Kuzan R, Beyazgül B, Karadağ D. (2024). A Study on the Development of a Health Literacy Assessment Scale. *Harran Üniversitesi Tıp Fakültesi Dergisi*, 21(3), 393-401.
10. Akduman GG ve Cantürk G. (2010). Cinsel istismara uğrayan çocuklara karşı tutum ölçeği geçerlik ve güvenilirlik çalışması (Üniversite öğrencileri örnekleme). *Adli Tıp Dergisi*, 24(2), 22-29.
11. Davis LL. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*. 5(4), 194-197.
12. Erkuş A ve Selvi H. (2019). Ölçek uyarlama ve "norm" geliştirme. *Pegem Akademi*. Ankara.

BÖLÜM 7

ADAY ÖLÇEĞİN DENENMESİ

Hazırlık aşamasında, aday ölçekte önceden öngörülemeyen veya fark edilemeyen sorunlar ortaya çıkabilir. Maddelerde yazım ve noktalama hataları bulunması ya da maddelerin katılımcılar tarafından anlaşılmaması, sık karşılaşılan durumlardandır. Hatta, ölçekte yer alan yanıtlama aralığı ve seçenekler, hedef popülasyonun sosyo-kültürel seviyesine uygun olmayabilir. Bazen de, aday ölçekte yer alan bazı maddeler yanıtlayıcılar tarafından net bir şekilde anlaşıldığı halde, toplumun tutum ve davranışları bu maddenin yanıtlama seçeneklerinin heterojen dağılımına izin vermeyebilir. Örneğin, sosyal normların baskın bir şekilde kabule yönlendirdiği bazı tutumlara dair maddelerde cevaplar yoğun bir şekilde ‘Kesinlikle katılıyorum’ yanıtına yığılabilir. Böyle bir durumda ilgili maddenin ölçme ve belirleyici olma özelliği ortadan kalkmış olur. Tüm bu nedenlerle, aday ölçek ana uygulamadan önce pilot uygulamalarla sınanmalıdır.

Ölçek geliştirme çalışmalarında iki pilot uygulama önerilmektedir (1).

1. Pilot uygulama: Aday ölçeğin hedef popülasyonundan seçilen bir gruba maddelerin değerlendirilmesi amacıyla yapılır. Maddelerin hem yanıt heterojenliği hem de toplanabilirliği değerlendirilir. Net olmamakla birlikte yanıtlayan sayısının 20-50 kişi arasında olması kabul görmektedir (1). Değerlendirme sonrası uygun olmayan maddelerin revizyonu yapılır. Bu aşamada madde atılmaz.

2. Pilot uygulama: 1. pilot uygulama sonrası gerekli görülen madde revizyonları yapıldıktan sonra 2. pilot uygulama ile revize edilmiş maddeler denenir. Bu uygulama için aday ölçeğin hedef popülasyonundan 20-50 kişi arasında katılımcı alınır ve maddeler istatistiksel yöntemle değerlendirilir (1).

2. pilot uygulama sonrası yapılan analizlerle uygun olmadığına karar verilen maddeler atılabilir. Burada dikkat edilmesi gereken, aynı alt boyuttan çok fazla

Bu bölüm boyunca tanımlanan analizler sonrası maddelerin gerekli kriterleri sağlamadığı gözlenirse, hangi pilot uygulamanın yapıldığı yol haritasını belirlemede önemli olur. Eğer 1.pilot uygulama sonrası yapılan analizlerde bir sorun yaşanırsa ölçme uzmanı, alan uzmanı ve dil uzmanı tarafından ilgili maddeler yeniden değerlendirilip, maddeler yeniden düzenlenir. Daha sonra 2.pilot uygulamaya geçilir. Yapılan 2.pilot uygulamada da sorunların devam ettiği gözlenirse, madde atımına gidilmesi düşünülür. Aday ölçekte kalmasına karar verilen maddeler ile ‘yapı geçerliliği’nin incelenmesi için ana çalışmaya geçilir.

KAYNAKLAR

1. Kartal M. ve Bardakçı S. (2019). Tutum Ölçekleri. Akademisyen Kitabevi. Ankara.
2. IBM SPSS Statistics v.22.0 (IBM Corp.; Armonk, NY, USA).
3. Aksakoğlu G. (2013). Sağlıkta Araştırma ve Çözümleme. 3.Baskı, Meta Basım, İzmir.
4. Asuero AG, Sayago A, González AG. (2006). The correlation coefficient: An overview. *Critical reviews in analytical chemistry*, 36(1), 41-59.
5. Taylor R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35-39.
6. Emond EJ ve Mason DW. (2002). A new rank correlation coefficient with application to the consensus ranking problem. *Journal of Multi-Criteria Decision Analysis*, 11(1), 17-28.
7. Çelik Y. (2011). Nasıl? Biyoistatistik Bilimsel Araştırma SPSS.
8. Özdamar K. (2016). Ölçek ve Test Geliştirme Yapısal Eşitlik Modellemesi. Ankara: Nisan Kitabevi Yayıncılık.
9. Piedmont R ve Hyland M. (1993). Inter-Item Correlation Frequency Distribution Analysis: A Method for Evaluating Scale Dimensionality. *Educational and Psychological Measurement*. 53 (2); 369-378.
10. DeVellis RF. (2022). Scale Development Theory and Applications. 3.Edition: Çeviri: Totan T., Nobel Akademik.
11. Kartal M. ve Bardakçı S. (2018). SPSS ve AMOS Uygulamalı Örneklerle Güvenirlik ve Geçerlik Analizleri. Akademisyen Kitabevi. Ankara.
12. Karagöz Y ve Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
13. Tavakol M ve Dennick R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53.
14. Bujang MA, Omar ED, Baharum NA. (2018). A Review on Sample Size Determination for Cronbach's Alpha Test: A Simple Guide for Researchers. *Malays J Med Sci*. 25(6):85-99.
15. Chakrabarty SN. (2013). Best split-half and maximum reliability. *IOSR Journal of Research & Method in Education*, 3(1), 1-8.
16. Pronk T, Molenaar D, Wiers RW, Murre J. (2022). Methods to split cognitive task data for estimating split-half reliability: A comprehensive review and systematic assessment. *Psychonomic Bulletin & Review*, 29(1), 44-54.
17. Zimmerman DW. (1972). Test reliability and the Kuder-Richardson formulas: Derivation from probability theory. *Educational and Psychological Measurement*, 32(4), 939-954.
18. Saupe JL. (1961). Some useful estimates of the Kuder-Richardson formula number 20 reliability coefficient. *Educational and Psychological Measurement*, 21(1), 63-71.

19. Foster RC. (2021). Kr20 and kr21 for some nondichotomous data (it's not just cronbach's alpha). Educational and Psychological Measurement, 81(6), 1172-1202.
20. Erkuş A ve Selvi H. (2019). Ölçek uyarlama ve “norm” geliştirme. *Pegem Akademi. Ankara*.
21. Kuzan R, Beyazgül B, Karadağ D. (2024). A Study on the Development of a Health Literacy Assessment Scale. *Harran Üniversitesi Tıp Fakültesi Dergisi*, 21(3), 393-401.
22. Onbaşı Ö. (2022). e-Sağlıklı beslenme okuryazarlığı (e-SBO) ölçeğinin Türkçe'ye uyarlanması ve yetişkin bireylerde covid-19 pandemi öncesi ve sırasında e-SBO düzeyleri ile beslenmeye ilişkin parametreler arasındaki ilişkinin belirlenmesi.

BÖLÜM 8

AÇIKLAYICI FAKTÖR ANALİZİ VE BOYUTLARIN OLUŞTURULMASI

Açıklayıcı Faktör Analizi (AFA), ölçek geliştirme çalışmalarında madde indirgemede kullanılan bir analizdir. Örtük değişkeni en net şekilde tamamlayabilmek için yola çıkan madde havuzunda doğası gereği çok sayıda madde vardır. AFA'dan ilk beklenen bu madde sayısını azaltması yani madde indirgemesidir. Ölçek geliştirme çalışmaları içinde oluşturulan madde havuzunda yer alan maddelerin bazıları, diğer bazıları ile çok yakın ilişkili iken, diğer bazıları ile de daha az yakın ilişkilidir. Sonuç olarak oluşturulan madde havuzu ile örtük değişkeni tanımlamaya çalışmak, mümkün olduğunca maddeleri tanımladıkları bölümlere ya da özelliklere göre gruplamaktır. Bu gruplamalara alt boyut ya da faktör denilir. İşte AFA'dan beklenen ikinci şey de budur: maddeleri alt boyutlara göre gruplamak (1,2). Bu gruplamanın detayları aşağıda bahsedilmiştir.

Araştırılan örtük yapı, üzerinde iyi çalışılmış, tanımları ve alt boyutları da belli ise, o zaman hangi maddelerin hangileri ile daha yakından ilişkili olduğu, ölçeğin kabaca kaç alt boyuta yığılacağı kolaylıkla öngörülebilir. Böyle bir durumda madde havuzunu oluştururken, öngörülen bu alt boyutlara göre hazırlık yapmak daha isabetli olacaktır.

Her ne şekilde oluşturulursa oluşturulsun, AFA bu maddeleri bir dizi matematiksel işleme göre gruplayarak alt boyut ya da faktörleri belirlemeyi sağlar. Bu gruplama, grup içinde maksimum, gruplar arasında minimum korelasyon oluşturma esasına göre yapılır. (1,2). Burada yapılan alt boyut belirleme işleminin sadece istatistik temelli olduğu unutulmamalıdır. Yani, araştırmacı AFA'nın önerdiği alt boyutların çalışılan örtük değişkenin teorik çerçevesine uygun olup olmadığını kendisi değerlendirmelidir. Araştırmacı, AFA'nın önerdiğinden az ya da daha çok alt boyut beklediği durumlarda, istatistik sonuçlarla teorik çerçevenin birbiri ile uyumunu değerlendirmelidir. Bu nedenle analizi yapan ve ölçeği

toplam varyansın %14.67'sini açıklayan bir alt boyut ve özgün değeri 2.60 olan ve toplam varyansın %14.43'ünü açıklayan bir alt boyut olduğu görülmektedir. Toplamda, 3 alt boyuttan oluşan ölçeğin, örtük değişkenin %47.21'ini açıkladığı görülmektedir (11). Literatürde, ölçek maddelerince açıklanan varyans oranının tek boyutlu ölçeklerde toplam varyansın en az %30'unu, çok boyutlu ölçeklerde ise en az %40'ını açıklaması gerektiği belirtilmektedir (4,5). Bu nedenle, aday ölçeğin açıkladığı varyans oranının yeterli ve örtük değişkenin temsilinin iyi olduğu söylenebilir.

KAYNAKLAR

1. Karagöz Y ve Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
2. Kartal M. ve Bardakçı S. (2018). SPSS ve AMOS Uygulamalı Örneklerle Güvenirlilik ve Geçerlilik Analizleri. Akademisyen Kitabevi, Ankara.
3. Orçan F. (2018). Açımlayıcı ve doğrulayıcı faktör analizi: İlk hangisi kullanılmalı. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 9(4), 413-421.
4. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young S.L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer, *Frontiers in Public Health*. 11;6:149.
5. Carpenter S. (2018). Ten Steps in Scale Development and Reporting: A Guide for Researchers, *Communication Methods and Measures*.12(1):25-44.
6. Gaskin CJ, Happell B. (2014). On exploratory factor analysis: a review of recent evidence, an assessment of current practice, and recommendations for future use. *Int J Nurs Stud*. 51(3):511-21.
7. DeVellis RF. (2003). Scale development: theory and applications (applied social research methods). *Social Research*. 900:14929.
8. Şencan H. (2005). Sosyal ve davranışsal ölçünlerde güvenilirlik ve geçerlilik. Ankara: Seçkin Yayıncılık. p. 105-363.
9. Howard, M. C. (2016). A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *International Journal of Human-Computer Interaction*, 32(1), 51-62.
10. Nkansah, B. (2018). On the Kaiser-Meier-Olkin's Measure of Sampling Adequacy. *Mathematical theory and modeling*, 8, 52-76.
11. IBM SPSS Statistics v.22.0 (IBM Corp.; Armonk, NY, USA).
12. Dziuban C ve Shirkey E. (1974). On the Psychometric Assessment of Correlation Matrices. *American Educational Research Journal*, 11, 211 - 216.
13. Aslam M ve Arif O. (2020). Multivariate Analysis under Indeterminacy: An Application to Chemical Content Data. *Journal of Analytical Methods in Chemistry*, 2020.
14. Larsen R ve Warne R. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods*, 42, 871-876.
15. Lin-quan Z. (2015). Application of Eigenvalue and Principal Components Based on Factor. *Mathematics in Practice and Theory*.
16. Ledesma R, Valero-Mora P, Macbeth G. (2015). The Scree Test and the Number of Factors: a Dynamic Graphics Approach. *The Spanish Journal of Psychology*, 18.

17. Macrosson W. (1999). Scree Plots, Data Structure, and Random Variance. *Psychological Reports*, 84, 533 - 540.
18. Costello, A. B., & Osborn, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, 10(1), 7.
19. Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

BÖLÜM 9

DOĞRULAYICI FAKTÖR ANALİZİ İLE YAPININ SINANMASI

Doğrulayıcı Faktör Analizi (DFA), Yapısal Eşitlik Modellemesi (YEM)'nin ölçek geliştirme çalışmalarında kullanılan türüdür. YEM'de kullanılan tanımlamalar ve yapılan analizler ufak değişiklikler olmakla birlikte, DFA'da da kullanılır. Bu bağlamda, YEM'i anlamak, DFA için de faydalı olacaktır. DFA, AFA ile alt boyutları geliştirilmiş ve madde indirgemesi yapılmış aday ölçeğin, örtük değişkeni belirlemedeki yeteneğini uyum indeksleri üzerinden değerlendirerek sınavmasını yapar. YEM'de olduğu gibi, DFA'da da, araştırmacı tarafından geliştirilen teorik modelin uygunluğu ölçülür (1-3).

Klasik epidemiyolojik çalışmaların tasarımı bağımlı ve bağımsız değişken yaklaşımına göre kurgulanır. Genel kabul, bir sebep-sonuç ilişkisinde, çalışmanın sonuç değişkeni olan bağımlı değişken ve bu bağımlı değişkeni etkileyen bağımsız değişkenin olmasıdır. Fakat, YEM'de farklı olarak, değişkenler endojen ve eksojen değişken olarak yapılandırılır. Endojen değişken, diğer değişkenlerden etkilenen değişkendir. Eksojen değişken ise, etkileyen değişkendir. İlk bakışta endojen değişkenin bağımlı değişkene, eksojen değişkenin de bağımsız değişkene benzemesine rağmen, bu isimlendirmenin asıl çıkış noktası YEM analizlerinde birden çok ilişkinin test edilmesi ve bu nedenle bazı değişkenlerin kimi zaman bağımlı, kimi zaman da bağımsız değişken gibi davranabilmesidir (1).

Aşağıda Şeşen tarafından, çalışanlarda adalet algısının tükenmişliğe etkisini ve iş tatmininin aracı değişken rolünün YEM ile incelendiği bir çalışmanın modeli örnek olarak sunulmuştur. Burada tek yönlü giden oklar, bir değişkenin diğer bir değişken üzerindeki etkisini gösterir (4).

KAYNAKLAR

1. Meydan CH ve Şeşen H (2015). Yapısal eşitlik modellemesi amos uygulamaları, 2.Baskı, Detay Yayıncılık, Ankara.
2. Alavi M, Visentin D, Thapa D, Hunt G, Watson R, Cleary M. (2020). Chi-square for model fit in confirmatory factor analysis.. *Journal of advanced nursing*.
3. Ocak M. (2020). AMOS ile adım adım yapısal eşitlik modellemesi uygulamaları, Sonçağ Yayıncılık, Bursa.
4. Şeşen H. (2010). Adalet algısının tükenmişliğe etkisi: İş tatmininin aracı değişken rolünün yapısal eşitlik modeli ile testi. *Savunma Bilimleri Dergisi*, 9(2), 67-90.
5. Karagöz Y ve Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
6. Kartal M. ve Bardakçı S. (2018). SPSS ve AMOS Uygulamalı Örneklerle Güvenirlik ve Geçerlik Analizleri. Akademisyen Kitabevi. Ankara.
7. DeVellis RF. (2003). Scale development: theory and applications (applied social research methods). *Social Research*. 900:14929.
8. Beyazgül B, Karadağ D. (2024). Doğum Öncesi Bakım Hizmetlerinin Birinci Basamak Sağlık Hizmetleri İçerisinde Çok Boyutlu Değerlendirilmesine Yönelik Ölçek Geliştirilmesi. V. Uluslararası Farklı Boyutlarıyla Sağlık Konferansı, Çanakkale, Türkiye.
9. IBM Corp. (2021). IBM SPSS Amos for Windows, Version 26.0. Armonk, NY: IBM Corp.
10. Goretzko D, Siemund K, Sterner P. (2023). Evaluating Model Fit of Measurement Models in Confirmatory Factor Analysis. *Educational and Psychological Measurement*.
11. Jackson D, Gillaspay J, Purc-Stephenson R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations.. *Psychological methods*, 14 1, 6-23.
12. Sun J. (2005). Assessing Goodness of Fit in Confirmatory Factor Analysis. *Measurement and Evaluation in Counseling and Development*, 37, 240 - 256.
13. Erkorkmaz Ü, Etikan İ, Demir O, Özdamar K, Sanisoğlu SY. (2013). Doğrulayıcı faktör analizi ve uyum indeksleri. *Türkiye Klinikleri Journal of Medical Sciences*, 33(1), 210-223.
14. Aylar F ve Nagihan E. (2019). Derleme: Ölçek Geliştirme Çalışmalarında Doğrulayıcı Faktör Analizinin Kullanımı. *The Journal of Social Sciences*, 4(10), 389-412.
15. Nye C. (2022). Reviewer Resources: Confirmatory Factor Analysis. *Organizational Research Methods*, 26, 608- 628.
16. IBM SPSS Statistics v.22.0 (IBM Corp.; Armonk, NY, USA).

BÖLÜM 10

ÖLÇEK PUANINI ‘DEĞERLEME’

Ölçek geliştirmede son adım, ölçekle elde edilen sayısal değere ne anlam verileceğinin belirlenmesidir. Herhangi bir alandaki bilgi ya da başarı düzeyini ölçmek amacıyla oluşturulmuş ölçeklerde yorumlama, kurgusal olarak maksimum performans testlerine benzer. Bilgi ya da başarı düzeyine göre katılımcı yanıtlarının sıralandığı ve puan arttıkça bilme ya da başarma durumunun pozitif yönde ilerlediği durumlarda, ölçekten alınan toplam puanın artması olumlu bir durum olarak değerlendirilebilir (1). Böyle durumlarda bir kesme noktası belirlemeye gerek olmayabilir. Ama çoğunlukla elde edilen puanla ne yapacağımıza karar vermemiz gereken bir durum oluşur.

Ölçek geliştirme süreci tamamlandıktan ve ölçeğin ölçülmek istenen örtük değişkeni ölçme bakımından geçerli ve güvenilir olduğu belirlendikten sonra, ölçekten alınan puana bir ‘değer’ ataması yapmak gerekir. Bu noktaya *değerleme* denir ve *değerlendirmeden* bazı yanları ile ayrılır. Değerleme ile yapılan şey, elde edilen ölçümle bir yargıda bulunma sürecidir. Psikoloji ve psikiyatri gibi alanlarda değerlendirme terimi ‘norm’ terimi ile eş kullanılır. Norm geliştirme, ölçüm değerine göre birey ya da grup verisi hakkında karar vermek için ona uygun ölçüt oluşturmaktır (2,3). Bu noktada değerlememizi ya da norm geliştirmemizi etkileyen birçok faktör devreye girer. Bu yeni geliştirilen ölçekten alınan toplam puan ne anlam ifade etmektedir? Ölçek puanının artması ya da azalması neyi göstermektedir? Ölçek puanının belli aralıklar arasında olması bir hastalığın varlığı için tarama testi mi olacaktır? Ölçek puanının referans grupla benzerlik ya da farklılığı mı bizim için önemlidir?

Değerlemenin, bir tek ölçümden yola çıkarak o ölçüm puanına sahip bireye dair bir karar verme ya da yargıda bulunma durumu olduğunu belirtmiştik. Bireyin aldığı puanın ya da ölçümünün değerlendirilmesinde kıyas yapıp yapılma-

lerinin dağılımları incelenir ve ortalama ya da ortancaları kesme noktası olarak belirlenir (3,22,23).

Bir akademik başarı testi geliştirilmekte ve bu testin belirli bir yeterlilik düzeyini ölçmesi amaçlanmakta olsun. Alan uzmanlarından, testi uygulayan 100 öğrenci içerisinde başarı düzeyine göre “yeterli” ve “yetersiz” olarak nitelendirilebilecek bireyleri sınıflandırmaları istenir. Uzmanlar, bu süreci yürütürken bazı öğrenciler hakkında kesin karar veremez; bu öğrencilerin başarı düzeylerinin sınırda olduğunu bildirirler. Bu kararsız gruptaki 20 öğrencinin test puanlarına bakıldığında puanlar 58 ile 64 arasında değişmektedir. Bu grubun ortalama puanı 61, ortanca puanı ise 60’tır. Bu nedenle 60 puan, sınır grup yöntemiyle belirlenen kesme puanı olarak tanımlanır. Artık bu testte 60 ve üzeri puan alan bireylerin “yeterli”, altında kalanların ise “yetersiz” kabul edileceği bilimsel bir ölçüt oluşturulmuştur.

KAYNAKLAR

1. Karagöz Y ve Bardakçı S. (2020). Bilimsel Araştırmalarda Kullanılan Ölçme Araçları ve Ölçek Geliştirme. Nobel Kitabevi, Ankara.
2. Anastasi A ve Urbina S. (1997). Psychological Testing (7th ed.). Upper Saddle River, NJ: Prentice Hall.
3. Erkuş A ve Selvi H. (2019). Ölçek uyarlama ve “norm” geliştirme. *Pegem Akademi. Ankara*.
4. Nartgün Z. (2007). Aynı Puanlar Üzerinden Yapılan Mutlak Ve Bağıl Değerlendirme Uygulamalarının Notlarda Farklılık Oluşturup Oluşturmadığına İlişkin Bir İnceleme. *Ege Eğitim Dergisi*, 8(1), 19-40.
5. Aksakoğlu G. (2013). Sağlıkta Araştırma ve Çözümleme. 3.Baskı, Meta Basım, İzmir.
6. Çelik Y. (2011). Nasıl? Biyoistatistik Bilimsel Araştırma SPSS.
7. McCluskey A ve Lalkhen AG. (2007). Statistics II: Central tendency and spread of data. *Continuing Education in Anaesthesia, Critical Care & Pain*, 7(4), 127-130.
8. IBM SPSS Statistics v.22.0 (IBM Corp.; Armonk, NY, USA).
9. Cizek GJ ve Bunch MB. (2007). Standard Setting: A Guide to Establishing and Evaluating Performance Standards on Tests. Sage Publications.
10. Kandane-Rathnayake RK, Enticott JC, Phillips LE. (2013). Data distribution: normal or abnormal? What to do about it. *Transfusion*. 53(4):701-2.
11. Erselcan T, Özen A, Yuksel D, Altun GD, Öztürk E, Balcı TA, Karayalcin B. (2009). Kemik mineral yoğunluğu ölçümü uygulama kılavuzu. *Türk J Nucl Med*, 18(1), 31-40.
12. WHO. (1994). Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO study group. Geneva, World Health Organization (WHO Technical Report Series, No. 843).
13. Erkuş A. (2012). Değerlendirme, değerlendirme, norm, norm geliştirme, normalleştirme nedir ne değildir? *İlköğretim Online*, 11(2), 1-6.
14. Zeyrek AŞ, Kurban NK, Arslan S. (2020). Bir başarı testi geliştirme çalışması: hemşirelik öğrencilerinin intramüsküler enjeksiyon becerilerini ölçme. *Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi*, 9(2), 133-141.

15. Kuzan R, Beyazgül B, Karadağ D. (2024). A Study on the Development of a Health Literacy Assessment Scale. *Harran Üniversitesi Tıp Fakültesi Dergisi*, 21(3), 393-401.
16. Kılıç S. (2013). Klinik karar vermede ROC analizi. *Journal of Mood Disorders*, 3(3), 135-40.
17. Cindoglu C, Beyazgul B, Tatligun M. (2024). Turkish Validity and Reliability of Comprehensive Diabetes Self-Management Scale. *Medical Bulletin of Sisli Etfal Hospital*, 58(1).
18. Demir O ve Köse İA. (2014). Angoff, Nedelsky ve Ebel standart belirleme yöntemleri ile belirlenen kesme puanlarının karşılaştırılması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*.
19. Ömür S ve Selvi H. (2010). Angoff, Ebel ve Nedelsky yöntemleriyle belirlenen kesme puanlarının sınıflama tutarlılıklarının karşılaştırılması. *Journal of Measurement and Evaluation in Education and Psychology*, 1(2), 109-113.
20. Taşdemir F. (2013). Angoff (1-0), Nedelsky ve sınır değerleri saptama yöntemleri ile bir testin sınıflama doğruluklarının incelenmesi. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi, Ankara.
21. Ozarkan HB. (2018). Açık uçlu testlerde standart belirleme sürecinde genişletilmiş Angoff ve karşıt gruplar yöntemlerinin karşılaştırılması (Master's thesis, Eğitim Bilimleri Enstitüsü).
22. Kılıç A ve Kilmen S. (2019). Angoff, Evet/Hayır ve Sınır Grup Yöntemlerine Göre Kesme Puanlarının Karşılaştırılması. *Boğaziçi Üniversitesi Eğitim Dergisi*, 36(2), 47-66.
23. Taşdemir F, Çokluk Ö. (2013). Angoff (1-0), Nedelsky ve Sınır Değerleri Saptama Yöntemleri ile Bir Testin Sınıflama Doğruluklarının İncelenmesi. *Mediterranean Journal of Humanities* mjh.akdeniz.edu.tr III/2, 241-261.