

Eğitimde ve Psikolojide Kullanılan Testlerin Kültürlerarası Değerlendirme Amacıyla Uyarlanması

Editörler:

Ronald K. Hambleton • Peter F. Merenda
Charles D. Spielberger

Çevirenler:

Prof. Dr. Nizamettin Koç
Dr. Ahmet Yıldırım



Editörler: Ronald K. Hambleton, Peter F. Merenda ve Charles D. Spielberger
Çevirenler: Prof. Dr. Nizamettin Koç ve Dr. Ahmet Yıldırım

EĞİTİMDE VE PSİKOLOJİDE KULLANILAN TESTLERİN KÜLTÜRLERARASI DEĞERLENDİRME AMACIYLA UYARLANMASI

ISBN 978-605-318-844-5

DOI 10.14527/9786053188445

Kitabın orijinal adı: Adapting Educational and Psychological Tests for Cross-Cultural Assessment

Baskı sayısı: 1. Baskı

ISBN: 978-0805830255

Yayınevi: Psychology Press

Kitap içeriğinin tüm sorumluluğu yazarlarına aittir.

© 2017, PEGEM AKADEMİ

Bu kitabın basım, yayım ve satış hakları Pegem Akademi Yay. Eğt. Dan. Hizm. Tic. Ltd. Şti.'ye aittir. Anılan kuruluşun izni alınmadan kitabın tümü ya da bölümleri, kapak tasarımı; mekanik, elektronik, fotokopi, manyetik, kayıt ya da başka yöntemlerle çoğaltılamaz, basılamaz, dağıtılamaz. Bu kitap T.C. Kültür Bakanlığı bandrolü ile satılmaktadır. Okuyucularımızın bandrolü olmayan kitaplar hakkında yayınevimize bilgi vermesini ve bandrolsüz yayınları satın almamasını diliyoruz.

Pegem Akademi Yayıncılık, 1998 yılından bugüne uluslararası düzeyde düzenli faaliyet yürüten **uluslararası akademik bir yayınevidir**. Yayımladığı kitaplar; Yükseköğretim Kurulu'nca tanınan yükseköğretim kurumlarının kataloglarında yer almaktadır. Dünyadaki en büyük çevrimiçi kamu erişim kataloğu olan **WorldCat** ve ayrıca Türkiye'de kurulan **Turcademy.com** ve **Pegemindeks.net** tarafından yayınları taranmaktadır, indekslenmektedir. Aynı alanda farklı yazarlara ait 1000'in üzerinde yayını bulunmaktadır. Pegem Akademi Yayınları ile ilgili detaylı bilgilere <http://pegem.net> adresinden ulaşılabilir.

1. Baskı: Ağustos 2017, Ankara

Yayın-Proje: Özlem Sağlam
Dizgi-Grafik Tasarım: Tuğba Kuşcuoğlu
Kapak Tasarım: Pegem Akademi

Baskı: Ay-bay Kırtasiye İnşaat Gıda Pazarlama ve Ticaret Limited Şirketi
Çetinemeç Bulvarı 1314.Cadde No:37A-B
0312 472 58 55

Yayıncı Sertifika No: 14749
Matbaa Sertifika No: 33365

İletişim

Karanfil 2 Sokak No: 45 Kızılay / ANKARA
Yayınevi: 0312 430 67 50 - 430 67 51
Yayınevi Belgeç: 0312 435 44 60
Dağıtım: 0312 434 54 24 - 434 54 08
Dağıtım Belgeç: 0312 431 37 38
Hazırlık Kursları: 0312 419 05 60
İnternet: www.pegem.net
E-ileti: pegem@pegem.net

ÖN SÖZ

1989 yılında beş farklı ülkedeki öğrencilerin matematik başarı düzeylerini ele alan karşılaştırmalı bir rapor okudum. Bu raporun sonuçları beni şaşırttı. İşte bu yüzden sonuçları etkilemiş olabileceğini düşündüğüm bir dizi yöntemsel faktörün sonuçlar üzerindeki etkisini merak etmeye başladım. Bu faktörler; katılımcı her bir ülkede yapılan örnekleme çalışmasının niteliği, testin kapsamı ve formatı idi. Ancak bu faktörlerden daha ziyade testin İngilizceden başka dillere çevrilme şeklini merak etmiştim. Yöntemsel faktörler sonuçların geçerliğine gölge düşürmediği sürece, eğitimde uluslararası başarı çalışmaları, politikacılar ve eğitimciler için çok değerli katkılar sağlayabilir. Bu şaşırtıcı sonuçların, çevirmenlerin testleri farkında olmadan daha kolay ya da zor çevirmesinden kaynaklanmış olabileceği aklıma geldi. Çevirmenlerin özellikleri neydi? İşleri tamamlamaları için ne kadar süre verildi? Testin farklı dillerdeki formlarının denkliğini destekleyecek hangi deneysel kanıtlar elde edildi? Test çeviri yöntemleri hakkında görüş alışverişinde bulunmak için yukarıda belirtilen çalışmayı yürüten sorumlu test bürosunu aradım. Ne yazık ki testin gerçekte nasıl çevrildiği; testin farklı dil ve kültürlerdeki formlarının dilbilimsel, psikolojik ve istatistiksel denkliğinin nasıl sağlandığı ile ilgili verilen bilgiler beni çok fazla tatmin etmedi.

Daha sonra nitelikli test uyarlama uygulamalarıyla ilgili yaptığım incelemelerde; test geliştirme, test puanlarının eşitlenmesi, test normlarının belirlenmesi gibi ölçme alanındaki diğer önemli konularda yapılan çalışmaların derinliği ile test çevirme ve uyarlama alanındaki çalışmaların sahip olduğu yöntemsel derinliği karşılaştırdığımda hayal kırıklığına uğradım. Bu, benim kültürlerarası ölçme dünyasına ilk girişimdi. Yapılacak çok önemli işler olduğunu görebiliyordum.

1991 yılında test uyarlama metodolojisine olan ilgimi Uluslararası Test Komisyonu (International Test Commission – ITC) Kurulu'na taşıdım. Bugün, ITC, ulusal psikoloji kuruluşları, test merkezleri ve bireysel üyeleri bulunan ve dünya çapındaki ölçme uygulamalarını geliştirmeye kendini adanmış bir kurumdur. ITC Kurulu, test çevirme ve uyarlama ile ilgili bir kılavuz geliştirmek için bilim insanları ve uygulayıcılardan oluşan bir komite oluşturmaya karar verdi. Oluşturulan ve üyesi bulunduğum komitenin yürüteceği işler için Amerika Birleşik Devletleri'ndeki Yükseköğretim Kurulu (College Board) ve Ulusal Eğitim İstatistiği Merkezi'nden finansal destek elde ettiğimiz için şanslıydık. Birçok uluslararası kurumun ilgisini ve dikkatini komitenin yürüttüğü işe çekmeyi başardık. Bu kurumlar oluşturulan komitenin üye ihtiyacını karşıladılar. Bu kurumlar: Avrupa Psikolojik Değerlendirme Birliği (European Association of Psychological Assessment), Avrupa Test Yayıncıları Birliği (European Test Publishers Group), Ulusla-

rarası Kültürlerarası Psikoloji Birliği (International Association for Cross-Cultural Psychology), Uluslararası Eğitimde Başarıyı Değerlendirme Birliği (International Association for the Evaluation of Educational Achievement), Uluslararası Dil Testleri Birliği (International Language Testing Association) ve Uluslararası Psikoloji Bilimi Birliği (International Union of Psychological Science) idi.

Komite üyeleri üç yılı aşkın bir süre yoğun bir şekilde çalıştılar. Test çevirme ve uyarlama konuları ile ilgili yaşanan teknik gelişmeleri ortaya koymak için birkaç toplantı düzenlediler ve komite bu sürenin sonunda 22 maddeden oluşan “Test Uyarlama İçin Uluslararası Test Komisyonu Kılavuzu” adında nihai bir kılavuz ortaya koydu. Kılavuzda yer alan maddeler ve bu maddelerin her birinin gerekçesi Birinci Bölüm’de yer almaktadır.

Kılavuzun taslak formunun yorum ve katkı almak amacıyla üniversitelere gönderildiği zamanlarda, Florida Üniversitesi’nden Tom Oakland (ITC üyesi) ve ben kılavuzda yer alan maddeleri sunmak üzere bir konferans düzenlemeye karar verdik. ITC tarafından finanse edilen bu konferans 1999 ilkbaharında Amerika Birleşik Devletleri, Georgetown Üniversitesi’nde yapıldı. Konferansa katılım yüksekti ve bu katılım test çevirme ve uyarlama kılavuzunun ölçme alanı tarafından memnuniyetle karşılanacağı ve mevcut alan yazına önemli bir katkı sağlayacağını gösteriyordu.

Konferansla aşağı yukarı aynı zamanda, Prof. Dr. Charles Spielberger ve Prof. Dr. Peter Merenda (Prof. Dr. Spielberger kılavuzu geliştiren komitenin bir üyesiydi) test çevirme ve uyarlama alanındaki önemli teknik gelişmelere dikkat çekecek bir kitap hazırlamaya talip oldular. Prof. Dr. Spielberger kendi ölçme aracı olan Kendi Beyanına Dayalı Kaygı Envanteri’nin, 50’den fazla çeviri çalışmasında yer almıştı. Peter Merenda ise meslek yaşamının büyük bir bölümünde çeviri araştırmalarında aktif olarak yer almıştı. Üçümüz, konunun daha kapsamlı bir şekilde irdelenmesini sağlamak amacıyla Georgetown Üniversitesi’ndeki ITC Konferansı’ndaki bölümlerin ve konuşmaların çoğunun eklendiği bu kitabı oluşturmak için bir ekip oluşturduk.

Benim yazdığım Birinci Bölüm, “Test Uyarlama İçin ITC Kılavuzu”nu tanıtmak amacıyla hazırlanmıştır. Buna ek olarak, test çevirme ve uyarlama alanında yayımlanan yayınların çoğu bu bölümde tanıtılmıştır. Test uyarlamada kavramsal ve yöntemsel konuları ele alan İkinci Bölüm, Hollanda Tilburg Üniversitesi’nden Prof. Dr. Fons van de Vijver ve Prof. Dr. Ype Poortinga tarafından hazırlanmıştır. İlk bölümde ITC kılavuzunu tanıtmak amacımız olmasaydı, kitapta ilk olarak bu bölüme yer verilecekti. Çünkü bu bölümün yazarları, bütün bölümlerle ilgili olan test çevirme ve uyarlama sürecinin anlaşılması için bir çerçeve sunmaktadırlar.

Üçüncü Bölüm Prof. Dr. Tom Oakland tarafından hazırlanmıştır ve yazar bu bölümde test uyarlama ile ilgili tüm önemli etik konuları ele almaktadır.

Dördüncü ve yedinci bölümler arasındaki bölümler, test çevirme ve uyarlama metodolojisindeki gelişmeleri sunmaktadır. Steve Sireci ve Liane Patsula (Amerika Birleşik Devletleri'nde Educational Testing Service'de çalışmaktadır) ve benim hazırladığım Dördüncü Bölüm test çevirme ve uyarlama sürecinde ortaya çıkan sorunlu test maddelerinin istatistiksel olarak belirlenmesine yönelik yaklaşımları ortaya koymaktadır. Prof. Dr. Sireci tarafından hazırlanan Beşinci Bölüm, bir testin farklı dillerdeki formlarının denkliğini sağlamak için iki dilli katılımcılardan yararlanılmasının güçlü ve zayıf yönleri üzerinde durmaktadır. Dr. Linda Cook, Dr. Alicia Schmitt-Cascallar ve Catharine Brown (sadece Yedinci Bölüm'de yer almıştır ve Educational Testing Service'de çalışmaktadır) tarafından hazırlanan altıncı ve yedinci bölümler birçok dilde formu olan testlerin istatistiksel olarak karşılaştırılması ile ilgili önemli yöntemleri ortaya koymakta ve test çevirme ve uyarlama sürecinde ortaya çıkan önemli sorunları incelemektedir. Alicia Schmitt-Cascallar'ın 2003 yılında beklenmedik bir şekilde vefat ettiğini üzülen ve belirtiyoruz. Alicia Schmitt-Cascallar, Georgetown Üniversitesi'ndeki ITC Konferansı'na davet edilmiş bir konuşmacıydı ve test çevirme ve uyarlama metodolojisi de olmak üzere ölçme alanındaki çalışmalara çok önemli katkılar sağlamıştır.

Kitabın sekizinci ve 14. bölümleri arasındaki bölümler okuyucuların dikkatini, temel olarak test uyarlama konuları ve yöntem bilimi ile ilgili konulardan test çevirme ve uyarlama uygulamalarının karmaşık dünyasına çekmek amacıyla hazırlanmıştır. Test çevirme ve uyarlama uygulamaları; yeterlik, sertifika, zekâ testleri, bilişsel testler, endüstride ve örgütlerde kullanılan testler, giriş testleri ve kişilik testlerini içine almaktadır. Dr. Cyndy Fitzgerald (önceden Microsoft'un, şimdilerde ise Caveon'ın danışmanlığını yapmaktadır) Sekizinci Bölüm'de Microsoft'un kendi yeterlik testlerini çevirmek ve uyarlamak için kullandığı süreci anlatmaktadır. Test çevirmenlerinin işini hızlandırmak amacıyla çevrim-içi (on-line) sistemlerin kullanımı bu bölümde örnek olarak görülmektedir. Dokuzuncu Bölüm'de, Dr. Carlos Maldonado ve Prof. Dr. Kurt Geisinger (Amerika Birleşik Devletleri, St. Thomas Üniversitesi'nden) dünyanın en yaygın kullanılan zekâ testlerinden birinin (Weschler Yetişkinler İçin Zeka Ölçeği) İngilizceden İspanyolcaya çevrilmesi ve uyarlanmasında karşılaşılan sorunları ortaya koymaktadır. Onuncu Bölüm'de, Prof. Dr. Norbert Tanzer (Amerika Birleşik Devletleri, Alliant Üniversitesi ve Avusturya, Graz Üniversitesi'nden) testlerin diller ve kültürlerarası uyarlanmasından ziyade bazı psikolojik testlerin eş zamanlı geliştirilmesi ile ilgili güçlü bir argüman ortaya atmaktadır. Prof. Dr. Fritz Drasgow ve Tahira Probst (Amerika Birleşik Devletleri, Illinois Üniversitesi'nden) 11. Bölüm'de temel olarak endüstride ve örgütlerde

kullanılan testlerin diller ve kültürlerarası test denkliğini sağlama ile ilgili çalışmalarını sunmaktadırlar. 12. Bölüm'de, Drs. Michal Beller (Educational Testing Servicede çalışmaktadır) ve Naomi Gafni ile Phina Hanani (İsrail Ulusal Ölçme ve Değerlendirme Enstitüsü'nden) İsrail'de altı dilde kullanılmak üzere üniversiteye giriş sınavlarını hazırlamak için yaptıkları çalışmaları anlatmaktadırlar. Peter Merenda (Amerika Birleşik Devletleri, Rhode Islan Üniversitesi'nden) tarafından hazırlanan 13. Bölüm, yazarın meslek yaşamı süresince test çevirme ve uyarlama alanındaki gözlemlerinin ve bulgularının büyük bir kısmını ortaya koymaktadır. Bu alanda çok az araştırmacı onun kadar uzun ve başarılı çalışmalar yapmıştır. Son olarak 14. Bölüm'de Prof. Dr. Spielberger (Güney Florida Üniversitesi'nden) ve aynı üniversiteden iki meslektaşı, Manolete Moscoso ve Thomas Brunner kişilik testlerinin çevrilmesi ve uyarlanmasıyla ilgili sorunlar ve yöntemlerle ilgili oldukça fazla bilgi vermektedirler.

Kendim ve yardımcı editörler, Peter Merenda ve Charles Spielberger, adına 14 bölümden oluşan bu kitabın, önemi giderek artan test çevirme ve uyarlama konusu üzerine yapılan araştırmalara yön vermesini ve bu araştırmaları teşvik ederek Uluslararası Test Komisyonu'nun görevini daha ilerilere taşımasını umuyoruz. Bu alandaki gelişmeler, 1989 yılındaki ilk sorgulamamdan beri giderek artmıştır. Bugün, bu alan daha da ileriye götürülebilir. İyi uygulamalar için ITC tarafından kılavuz yayımlanmış, yöntembilim düzenlenmiş ve genişletilmiştir. Ayrıca uygulayıcıların takip edebileceği örnek niteliği taşıyan birçok çalışma mevcuttur. Aynı zamanda, yapılması gereken kayda değer birçok araştırma vardır. Bu kitabın, bu çalışmaları daha ileri taşıyacak başka araştırmacıları harekete geçirmesini umuyoruz.

Ronald K. Hambleton

ÇEVİRENLERİN ÖN SÖZÜ

Belirli bir dil ve kültür için geliştirilen testlerin ikinci bir dil ve kültürde kullanılmak üzere uyarlanmasına duyulan ilgi 20. yüzyılın ortalarından itibaren eğitim ve psikoloji alanında araştırma ve uygulamalar yapan kişiler arasında yaygınlık kazanmıştır. Bu ilgi günümüzde de tüm yoğunluğuyla artarak devam etmektedir. Bunun önemli nedenlerinden biri son yıllarda kültürlerarası farklılıkları ve karşılaştırmaları ele alan araştırmalarda sürekli bir artışın olmasıdır. Önümüzdeki yıllarda da uyarlanmış testlerin çeşitli alanlarda kullanımı ve bu testlere duyulan ihtiyaç konusunda daha büyük bir artışın yaşanacağı tahmin edilebilir.

Testlerin bir dil ve kültürden başka bir dil ve kültüre uyarlanmasını gerektiren pek çok neden vardır. Bu nedenlerden bazıları, test uyarlamanın; kültürler ve diller arası okul başarısı ile ilgili karşılaştırmalı çalışmaları kolaylaştırması, yeni testleri hazırlamak için gerekli para ve zamandan tasarruf sağlaması, değerlendirmede eşitliği sağlamasıdır. Ancak, test uyarlama ve puanların denkliliğini sağlama ile ilgili yöntemler ve kılavuzlar çok iyi bilinmemektedir. Bu nedenle, uyarlanan testler kullanılarak elde edilen verilere dayalı olan bulgular -uyarlama standartlarına uymayan, niteliksiz uyarlamalar nedeniyle- tartışmalı hale gelmektedir. Başka kültürlerde geliştirilmiş olup Türkiye’de kullanılagelen birçok ölçme aracının kullanıldığı araştırmalar ve karar alanlarında (seçme, sınıflama, yerleştirme, rehberlik, teşhis vb.) benzer sorunlar yoğun bir biçimde yaşanmaktadır. Nitekim ulusal ve uluslararası düzeylerde gerçekleştirilen birçok kongre, panel, konferans ve seminerde yetersiz/niteliksiz uyarlamaların yol açtığı sorunların yoğun bir biçimde dile getirildiği gözlenmektedir.

Bir ölçme aracı geliştirmenin gerektirdiği uzmanlık hizmetlerindeki ve maddi kaynaklardaki sınırlılıklar nedeniyle az gelişmiş ve gelişmekte olan ülkelerde psikolog ve uzmanların yabancı bir kültürde geliştirilen ve standardize edilen ölçme araçlarını kendi kültürlerine uyarlayarak kullanma yoluna gittikleri gözlenmektedir. Ne yazık ki ölçme araçlarının bir kültürden başka bir kültüre uyarlanmasında sıklıkla izlenen yol, “uyarlama” adı altında ölçme araçlarının yalnızca dil bilimsel çevirisinin (bunun da ötesinde “niteliksiz çevirisinin”) yapılmasıdır. Bu hatalı uygulamalarla, Türkiye’de de yaygın bir biçimde karşılaşılmalıdır. Bunun temel nedeni, uyarlama çalışması yapan araştırmacılara rehberlik edecek Türkçe kaynakların sınırlı olmasıdır.

Dünya çapındaki ölçme uygulamalarını geliştirmeye kendini adanmış; ulusal psikoloji kuruluşları, test merkezleri ve bireysel üyeleri bulunan Uluslararası Test Komisyonu (International Test Commission – ITC) tarafından oluşturulan, hepsi psikolojik ölçme ve değerlendirme alanının yetkin uzmanları konumunda olan 21

kişilik bir grubun ortaya koyduğu bu eseri titiz bir çalışma çerçevesinde çevirerek alanımıza kazandırmış olmanın mutluluğunu yaşamaktayız.

Bu kitabın çevirisine karar verdiğimizde, gerekli izinlerin alınması çerçevesinde ilk yazıştığımız, kitabın editörlerinden Ronald K. Hambleton'a, bizi teşvik edici yazıları ve yönlendirici bilgileri için müteşekkirimiz. Ayrıca, kitabın Türkiye'de yayın hakkını alarak yayımlayan Pegem Akademi Müdürü Servet Sarıkaya'ya, kitabı okuyarak çok değerli katkılarda bulunan Doç. Dr. Kasım Kiroğlu'na ve Pegem Akademi çalışanlarına çok teşekkür ediyoruz.

Kitabın, başka kültürlerde ve dillerde geliştirilmiş ölçme araçlarını (testleri) Türk kültürüne uyarlayarak çeşitli amaçlarla kullanmak durumunda kalan araştırmacıların "ölçme aracı uyarlama ilke ve standartları"na uygun daha nitelikli uyarlama çalışmaları yapmalarına önemli katkı sağlaması umuduyla...

Prof. Dr. Nizamettin KOÇ

Dr. Ahmet YILDIRIM

*İleriki yaşamlarının önemli bir boyutunun
“Kültürlerarası Çalışmalar”a duyarlılık olması dileğiyle*

Özgün Bora KOÇ’a ve Zehra Beren YILDIRIM’a...

İÇİNDEKİLER

Ön Söz.....	iii
Çevirenlerin Ön Sözü	vii

1. KISIM

EĞİTİMDE VE PSİKOLOJİDE KULLANILAN TESTLERİN KÜLTÜRLERARASI UYARLANMASI: KURAMSAL VE YÖNTEMSSEL KONULAR

1. BÖLÜM: Testlerin Birden Çok Dil ve Kültüre Uyarlanmasıyla İlgili Konular, Uyarlama Desenleri ve Uyarlama İçin Teknik Yönergeler	1
<i>Ronald K. Hambleton</i>	
2. BÖLÜM: Test Uyarlamada Kavramsal ve Yöntemsel Konular	41
<i>Fons J. R. van de Vijver, Ype H. Poortinga</i>	
3. BÖLÜM: Test Uyarlama ile İlgili Seçilmiş Etik Konular	67
<i>Thomas Oakland</i>	
4. BÖLÜM: Test Uyarlama Sürecinde Ortaya Çıkan Hataların Belirlenmesinde Kullanılan İstatistiksel Yöntemler	95
<i>Stephen G. Sireci, Liane Patsula, Ronald K. Hambleton</i>	
5. BÖLÜM: Bir Testin Farklı Dillerdeki Formlarının Karşılaştırılabilirliğinin Değerlendirilmesi İçin İki Dilli Bireylerden Yararlanma	121
<i>Stephen G. Sireci</i>	
6. BÖLÜM: Bir Testin Farklı Dillerdeki Formlarından Alınan Puanların Karşılaştırılabilirliğinin Belirlenmesi	145
<i>Linda L. Cook, Alicia P. Schmitt-Cascallar</i>	
7. BÖLÜM: Başarı ve Yetenek Testlerinin Uyarlanması: Yöntemsel Konuların Gözden Geçirilmesi	177
<i>Linda L. Cook, Alicia P. Schmitt-Cascallar, Catherine Brown</i>	

2. KISIM

EđİTİMDE VE PSİKOLOJİDE KULLANILAN TESTLERİN KLTRLERARASI UYARLANMASI: BAđARI, YETENEK VE KİřİLİK TESTLERİNDEKİ UYGULAMALAR

8. BLM: Geniř lekli Sertifika Programlarında Test Uyarlama	201
<i>Cyndy T. Fitzgerald</i>	
9. BLM: Wechsler Yetiřkinler İin Zekâ leęi'nin İřpanyolcaya Uyarlanması: nemli Sonuları Olan Bir Test Uyarlama alıřması	221
<i>Carlos Y. Maldonado, Kurt F. Geisinger</i>	
10. BLM: Testlerin Birden Fazla Dil ve Kltrde Kullanılması Amacıyla Geliřtirilmesi: Eř Zamanlı Test Geliřtirmeye Duyulan Talep	243
<i>Norbert K. Tanzer</i>	
11. BLM: Test Uyarlamada Psikometrik zellikler: Diller ve Kltrlerarası lme Deęiřmezlięinin Deęerlendirilmesi	271
<i>Fritz Drasgow, Tahira M. Probst</i>	
12. BLM: Birden Fazla Dilde Uygulanan Giriř (Seme) Sınavlarının Geliřtirilmesi, Uyarlanması ve Yrtlen Geerlik alıřmaları: İsrail rneęi	305
<i>Michal Beller, Naomi Gafni, Pnina Hanani</i>	
13. BLM: Eđitimde ve Psikolojide Kullanılan Testlerin Kltrlerarası Deęerlendirilmesi	329
<i>Peter F. Merenda</i>	
14. BLM: Duygu Durumlarının ve Kiřilik zelliklerinin Kltrlerarası Deęerlendirilmesi	351
<i>Charles D. Spielberger, Manolete S. Moscoso, Thomas M. Brunner</i>	
Yazar Dizini	377
Konu Dizini	385

I

Eğitimde ve Psikolojide
Kullanılan Testlerin
Kültürlerarası Uyarlanması:
Kuramsal ve Yöntemsel Konular

1. BÖLÜM

Testlerin Birden Çok Dil ve Kültüre Uyarlanmasıyla İlgili Konular, Uyarlama Desenleri ve Uyarlama İçin Teknik Yönergeler

Ronald K. Hambleton
Massachusetts Üniversitesi

Günümüzde, başarı, yetenek, kişilik testleri ve anketlerin çok dilli formlarına duyulan ihtiyacı ortaya koyan birçok bulgu mevcuttur (bkz. Ercikan, 2002; Hambleton, 2002; Hambleton ve de Jong, 2003; Harkness, 1998). Örneğin, Eğitimde Uluslararası Başarı Değerlendirme Birliği (International Association for the Evaluation of Educational Achievement - IEA) matematik ve fen testlerinin 30'dan fazla dilde hazırlanmasını gerektiren Üçüncü Uluslararası Matematik ve Fen Çalışması'nı (The Third International Mathematics and Science Study – TIMSS) 45'ten fazla ülkede uygulamıştır. Amerika Birleşik Devletleri'nde yeni test uyarlama projelerinin bilinen iyi örnekleri, Amerikan Yükseköğretim Kurulu (College Board) tarafından İspanyolca formu hazırlanan *Akademik Yetenek Testi (Scholastic Aptitude Test- SAT)*, Amerika Eğitim Konseyi tarafından hazırlanan *Genel Eğitim- sel Gelişim Testi (General Educational Development Test – GED)*, Amerika Birleşik Devletleri Eğitim Departmanı tarafından hazırlanan *Ulusal Eğitimde İlerleme Değerlendirmesi (National Assessment of Educational Progress (NAEP)* ve birkaç eyaletin eğitim birimleri tarafından hazırlanan başarı testleridir. Gelecekte daha fazla test uyarlama çalışmasının olması beklenmektedir. Çünkü a) son zamanlarda testlerin uluslararası değişimi daha yaygın hale gelmiştir, b) uluslararası sertifikalar vermek amacıyla daha fazla sınav kullanılmaktadır ve c) kültürlerarası araştırmalara olan ilgi artmaktadır.

Testlerin bir dil ve kültürden başka bir dil ve kültüre uyarlanmasının altında yatan birçok neden/gerekeçe – örneğin, kültürler ve dillerarası okul başarısı ile il-

gili karşılaştırmalı çalışmaları kolaylaştırması, yeni testleri hazırlamak için gerekli para ve zamandan tasarruf sağlaması, değerlendirmede eşitliği sağlama gibi- olmasına rağmen test uyarlama ve puanların denkliliğini sağlama ile ilgili yöntemler ve kılavuzlar çok iyi bilinmemektedir (Hambleton, 1993, 1994; Hui ve Triandis, 1985; van de Vijver ve Hambleton, 1996). Kültürlerarası araştırmalar yürüten birçok araştırmacı, nitelişiz uyarlanmış testler yüzünden alanlarındaki araştırmaların büyük bir kısmının geçersiz olacak kadar sorunlu olduğunu ifade etmişlerdir.

Bu bölümün amaçları; a) test uyarlama ile ilgili hata kaynaklarını ve bu hataları azaltmanın yollarını ortaya koymak b) yedi büyük uluslararası kuruluşun desteğiyle Uluslararası Test Komisyonu (International Test Commission – ITC) tarafından hazırlanan test uyarlama kılavuzunu tanıtmaktır (bkz. Hambleton, 1994; van de Vijver ve Hambleton, 1996).

Daha fazla ilerlemeden, test uyarlama ve test çevirme ile ilgili bir ayrımın yapılması gerekmektedir. Bu bölümde **test uyarlama** kavramı daha yaygın ve sıklıkla kullanılan **test çevirme** kavramına tercih edilmektedir. Çünkü **test uyarlama** kavramı, başka bir dil ve kültüre uyarlanmak üzere geliştirilen bir testin uyarlama aşamasında yapılması gerekenleri daha çok kapsamakta ve yansıtmaktadır. Test uyarlama, bir testin farklı dil ve kültürlerde aynı yapıyı ölçebiliş ölçemeyeceğinden tutun da, çevirmenlerin seçimine, uyarlama sürecinde yapılması gereken düzenlemelere karar vermeye, testin uyarlanmasına ve uyarlanmış formunun testin orijinal formuyla denkliliğini test etmeye kadar birçok işlemi kapsamaktadır. Test çevirme ise, test uyarlama sürecinin tek bir adımıdır ve hatta bu adımda bile, izlenen süreç bakımından “test çevirme” yerine “test uyarlama” kavramının kullanılması daha uygundur. Çünkü test çevirmenleri testi çevirirken; testin uyarlandığı dil ve kültürde kültürel, psikolojik ve dilbilimsel olarak denk olan kavramları, kelimeleri ve ifadeleri bulmaya çalışmaktadırlar. Bu işlem, açık bir şekilde testin bire bir çevrilmesinin ötesine geçmektedir.

Bu bölüm boyunca, eğitimde ve psikolojide kullanılan bütün ölçme araçlarını, anketleri ve ölçekleri içine alacak şekilde “test” kavramını kullanıyoruz.

TEST UYARLAMADA HATA KAYNAKLARI VE GEÇERLİĞİ DÜŞÜREN FAKTÖRLER

Amerika Eğitimde Araştırma Birliği (American Educational Research Association - AERA), Amerikan Psikologlar Derneği (American Psychological Association - APA) ve Eğitimde Ulusal Ölçme Konseyi (National Council on Measurement in Education – NCME) gibi kuruluşların ortak çabasıyla geliştirilen **Eğitimde ve Psikolojik Ölçmelerde Test Standartları** (1985); eğitimde ve psiko-

lojide kullanılan testleri seçen, geliştiren, uygulayan ve kullanan eğitimde ölçme uzmanları ve psikologlar için özenle hazırlanmış yönergeleri içermektedir. Bu yayında yer alan standartların üçü özellikle test uyarlama ile ilişkilidir:

Standart 6.2. Bir test kullanıcısı, testin formatında, uygulanışında, yönergelerinde, dilinde ve kapsamında kapsamlı bir değişiklik yaparsa, testin değiştirilen bölümleri için tekrar geçerlik kanıtları elde etmelidir ya da değiştirilen testin geçerliğiyle ilgili kanıtlar elde etmenin gerekli ya da mümkün olmadığına dair mantıksal bir gerekçe sunmalıdır.

Standart 13.4. Bir test bir dilden ya da lehçeden başka bir dil ya da lehçeye çevrildiğinde, testin kullanılmasının hedeflendiği dil gruplarında testin güvenilirliği ve geçerliği saptanmalıdır.

Standart 13.6. İki-dilli testlerin iki formunun karşılaştırılabilir olması hedeflendiğinde, testlerin karşılaştırılabilir olduğuna dair kanıtlar elde edilmelidir.

Bu standartlar, bir testin bir dil ve kültürden başka bir dil ya da kültüre uyarlama çalışmalarında ortaya çıkabilecek hata kaynaklarının ya da geçerliği düşüren faktörlerin göz önünde bulundurulması ile ilgili bir çerçeve sunmaktadır. Test uyarlamada ortaya çıkan hata kaynakları ya da geçerliği düşüren faktörler üç geniş gruba ayrılabilir: a) kültürel/dilsel farklılıklar, b) teknik konular, desenler ve yöntemler c) sonuçların yorumlanması. Bu grupların her birinde yer alan hata kaynaklarına yeterince dikkat edilmemesi, uyarlanmış olan testin, orijinal formuna denk olmaması ile sonuçlanabilir. Denk olmayan testler de, denk varsayıldıklarında, ölçme işleminin yürütüldüğü gruplar hakkında yanlış yorumlarda ve çıkarımlarda bulunma gibi hatalarla sonuçlanabilir.

Niteliksiz test uyarlamadan kaynaklanan yanlış yorumda bulunma ile ilgili güzel bir örnek aşağıda verilmiştir. (Bu örnek, Columbia Teachers College'dan mesleki yaşamı boyunca uluslararası ölçme alanında bir öncü olan Richard Wolf tarafından aktarılmıştır.) Okuma becerisinin ölçüldüğü uluslararası karşılaştırmalı bir çalışmada (1990'lı yıllarda), Amerikalı öğrencilere kelime çiftleri verilmiş ve öğrencilerden anlam bakımından bu kelime çiftlerinin benzer ya da farklı olup olmadığına karar vermeleri istenmiştir: *Sanguine – pessimistic* kelime çifti, Amerikalı öğrencilerin şans başarısının sadece biraz üzerinde performans gösterdikleri (doğru bilme oranı: %54) kelime çiftlerinden biriydi. İngilizce konuşulmayan bir ülkede ise öğrencilerin yaklaşık %98'i bu soruyu doğru yanıtlayarak bu sorudaki performansları bakımından birinci olmuşlardır. Öğrencilerin performansları arasındaki bu büyük farklılığın neden kaynaklandığını daha iyi anlamak için yapılan değerlendirme sonucunda, öğrencileri birinci olan ülkenin dilinde *sanguine* kelimesinin karşılığının olmadığı anlaşılmıştır. Bu yüzden de *sanguine* kelimesinin

yerine *optimistic* kelimesinin karşılığı kullanılmıştır. Bu durum, soruyu kayda değer ölçüde kolay hale getirmiştir. Eğer bu soru, Amerikalı öğrencilere *optimistic* – *pessimistic* şeklinde sorulmuş olsaydı bu öğrenciler daha yüksek bir yüzdeyle bu soruyu doğru yanıtlarlardı. Bu örnek, test uyarlama sürecinin birbirine denk iki test ortaya koymakla sonuçlandığına dair güçlü kanıtlar olmaksızın uluslararası karşılaştırmalı sınavlarla ilgili çıkarımda bulunmanın tehlikesine işaret etmek amacıyla verilmiştir. 1990 yılından önce, uluslararası çalışmalar için birçok test uyarlama girişimi, birkaç nitelikli çevirmenden daha fazlasına sahip değildi. Bu durum, günümüzde hem TIMSS hem de Ekonomik İşbirliği ve Kalkınma Örgütü (Organization for Economic Cooperation and Development – OECD) tarafından geliştirilen Uluslararası Öğrenci Başarısı Değerlendirme Programı'nda (Programme for International Student Assessment - PISA) görülen test uyarlamadaki üst düzey gelişmişlik ile karşılaştırılabilir (OECD/PISA; bkz. Grisay, 2003; Hambleton, 2002).

Bu bölüm, uygulamada karşılaşılan birkaç yaygın hata kaynağı ve bu hataların üstesinden nasıl gelinebileceği tartışması ile devam etmektedir.

Puanları Etkileyen Kültürel ve Dilsel Farklılıklar

Kültürlerarası sonuçların değerlendirilmesi ve yorumlanması süreci, testlerin sadece çevrilmesi ya da uyarlanması gibi dar bir bağlamda düşünülmemelidir (van de Vijver ve Leung, 1997, 2000). Bu süreç daha ziyade, değerlendirme aşamasının bütün bölümleriyle –ölçülen yapının denkliği, testin uygulanması, kullanılan madde formatları ve bireylerin performansını etkileyebilecek hız– birlikte düşünülmelidir. Bu dört faktöre, bu bölümde kısaca değinilmektedir. Sonraki bölümlerde ise, bu faktörlere detaylı bir şekilde değinilmiştir.

Ölçülen Yapının Denkliği. Ölçülen yapının denkliği, testle ölçülen yapının her bir dil ve kültürdeki işe vuruk tanımlarının denkliğinin yanı sıra kavramsal ve işlevsel denkliğini de içine almaktadır (Harkness, 1998). Çalışmanın yürütüldüğü kültürlerde ölçülen yapının birbirine denk olduğunun belirlenmesi, uluslararası, kültürlerarası ya da dillerarası karşılaştırmalar yapmanın ön koşuludur. Birbirine denk olmayan yapıların kullanılması, dillerarası araştırmalarda yapılan en ciddi hatalardan biridir. Örneğin, testle ölçülen içerik bir ülkedeki öğretim programını yansıtırken diğer ülkedeki öğretim programını yansıtmıyorsa bu iki ülkedeki öğrencilerin matematik başarıları bakımından karşılaştırılması, hiçbir anlam taşımamaktadır. Bunun başka bir örneği, ölçülen yapı bakımından *yaşam kalitesi* olabilir. Bu yapı (yaşam kalitesi), bir ülkede arabalar, evler, televizyon gibi birçok eşyaya sahip olmaya karşılık gelirken, başka bir ülkede hayatta kalmak için biraz yiyeceğe sahip olmak ve kişinin gerektiğinde ulaşabileceği bir doktorun olmasına karşılık