

Eğitim Kurumlarında Kullanılan Psikolojik Testlerin Ölçme Standartlarına Göre İncelenmesi*

The Analysis of the Psychological Tests Using In Educational Institutions According To the Testing Standards*

Ezgi MOR DİRLİK**

Nizamettin KOÇ***

Öz

Bu araştırmada Rehberlik ve Araştırma Merkezleri ile okulların rehberlik servislerinde sık kullanılan psikolojik testlerden dördünün Amerikan Psikologlar Birliği (APA-American Psychological Association) tarafından yayınlanan Eğitim ve Psikolojide Ölçme Standartları ile Uluslararası Test Komisyonu (ITC-International Testing Commission) tarafından yayınlanan test uyarlama standartlarına uygunluğunu incelemesi amaçlanmıştır. Araştırmaya alınan psikolojik testler, amaçlı örneklem belirleme yöntemiyle belirlenmiştir ve Rehberlik ve Araştırma Merkezleri ile okulların rehberlik servislerinde sıklıkla kullanılan psikolojik testlerden seçilmiştir. Bu testler; Akademik Benlik Kavramı Ölçeği (ABKÖ), Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı (GEÇDA), Temel Kabiliyetler Testi 7-11 Yaş Formu (TKT 7-11) ve Wechsler Çocuklar İçin Zekâ Ölçeği Revize Edilmiş Formu (WISC-R)'dur. Araştırma kapsamında, APA tarafından 1999'da İngilizce olarak yayınlanmış Eğitim ve Psikolojide Ölçme Standartları kitabının geçerlik, güvenilirlik ve test geliştirme ve revizyon bölümlerindeki standartlar ile ITC tarafından geliştirilen test uyarlama standartları Türkçeye çevrilmiş ve kontrol listesi oluşturulmuştur. Söz konusu standartlar temel alınarak hazırlanan kontrol listesi kısa ve uzun olmak üzere iki form halinde oluşturulmuştur. Çalışmaya alınan psikolojik testler hazırlanan kontrol listesine göre araştırmacı tarafından incelenmiştir. İncelemelerin güvenilirliğini sağlama amacıyla her psikolojik test için yapılan incelemeler üç hafta arayla tekrarlanmıştır. İncelemeler sonucu elde edilen veriler SPSS 20.0 programında betimsel istatistikler kullanılarak incelenmiştir. Araştırma sonucunda incelenen psikolojik testlerin kontrol listesindeki standartları karşılama düzeyleri ve bu standartlar bakımından geliştirilmesi gereken yönleri belirlenmiştir. Elde edilen bulgulara göre, ABKÖ ve GEÇDA için incelemeye alınan standartların yeterli düzeyde karşılanmadığı belirlenmiş ve ABKÖ için yapılan güvenilirlik incelemelerinin geliştirilmesi gerektiği sonucuna ulaşılmıştır. Uyarlama çalışmaları olan WISC-R ve TKT 7-11 Yaş Formu'nun test uyarlama standartlarını karşılama düzeyleri incelendiğinde ise, WISC-R için standartların büyük çoğunluğunun karşılandığı, TKT 7-11 Yaş Formu için ise karşılanmayan birden çok standart olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Psikolojik testler, APA eğitimde ve psikolojide ölçme standartları, ITC test uyarlama standartları.

Abstract

The purpose of this research is to analyze four psychological tests which are frequently used in the Guidance and Research Centers and in the guidance services of the schools according to the standards for educational and psychological testing of APA (American Psychological Association) and test adaption standards of ITC (International Testing Commission). The tests were determined based on the goal-oriented sample selecting method and were selected from the most frequently used psychological tests in Guidance and Research Centers and school's guidance centers. These tests are: Scale of Academic Self-Concept (Akademik Benlik Kavramı Ölçeği-ABKÖ), Evaluation of Early Childhood Development Tool (Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı-GEÇDA), Primary Mental Abilities 7-11 (TKT 7-11), and Wechsler

*Bu çalışma Prof. Dr. Nizamettin Koç danışmanlığında ilk yazar tarafından 2013 yılında tamamlanan "Eğitim Kurumlarında Kullanılan Psikolojik Testlerin Ölçme Standartlarına Göre İncelenmesi" isimli yüksek lisans tez çalışmasından üretilmiştir.

**Yrd.Doç.Dr, Kastamonu Üniversitesi, Kastamonu-Türkiye, emor@kastamonu.edu.tr, ORCID ID: <https://orcid.org/0000-0003-0250-327X>

** Prof. Dr. Ankara Üniversitesi, Ankara-Türkiye, nkoc@ankara.edu.tr, ORCID ID: <https://orcid.org/0000-0002-3308-7849>

Intelligence Scale for Children Revised Form (WISC-R). In this research, the chapters related to the validity, reliability and test development and revision of “Standards For Educational And Psychological Testing” (APA, 1999) and the adaptation standards developed by ITC were translated into Turkish and a checklist was created by using these documents. The checklist has got two forms as short and long form. The tests were analyzed according to the short form of the checklist by researcher. In order to examine the reliability of these analyses, the analyses were repeated in three weeks’ time. Data of these analyses were exported to the Statistical Package for Social Sciences (SPSS) 20.0 and descriptive analysis was performed. As a result of this research, the meeting levels of the psychological tests to the test standards in the checklist and the features of the tests which should be improved according to the validity, reliability, test development and revision and test adaptation were determined. In conclusion, the standards analyzed have not been met satisfactorily by ABKÖ and GEÇDA, and according to the analyses of the reliability of ABKÖ, it has been found that more reliability analyses should be done for ABKÖ. As for the results that obtained for the adapted tests, it has been found that most of the adaptation standards have been met for WISC-R and there are many standards that have not been met by for Primary Mental Abilities 7-11.

Keywords: Psychological tests, APA standards for educational and psychological testing, ITC test adaptation standards.

GİRİŞ

Eğitimin evrensel amaçları arasında, bireylerin kendilerini tanımasını sağlayarak doğru karar verme güçlerini geliştirmek ve bu sayede kendilerini gerçekleştirmelerine yardımcı olmak bulunmaktadır (Koç, 1993). Rehberlik ve psikolojik danışma hizmetlerinin amacıyla da paralel olan bu amacı gerçekleştirmek için, bireyin kendisini tanımasına, kontrol etmesine, gelişimine uygun karar vermesine ve ilgi ve yeteneği doğrultusunda eğitim almasına yardımcı olma ve yönlendirme amaçlarıyla çeşitli bilgi edinme yollarına başvurulmaktadır. Önceden belirlenmiş amaçlar doğrultusunda, bireylerin psikolojik özelliklerinin ölçülmesi söz konusu olduğunda psikometrik ve izlenimci yaklaşım olmak üzere iki temel yaklaşım ortaya çıkmaktadır. Psikometrik yaklaşım izlenimci yaklaşıma göre daha fazla tercih edilmekte ve bu yaklaşım kapsamında, bireylere psikolojik bir test uygulanmakta ve izlenimci yaklaşıma göre daha kapsamlı ve derinlemesine bilgiler elde edilmektedir (Anastasi, 1988; Özgüven, 1994). Psikolojik testlerin diğer bilgi toplama yöntemlerine göre daha kolay uygulanabilmesi, objektif olarak puanlanabilmesi, geçerli ve güvenilir gözlemler yapmaya olanak sağlaması, bu ölçme araçlarının daha fazla tercih edilmesinin nedenlerini oluşturmaktadır (Cronbach, 1990).

Ölçekler, doğrudan gözlem yapmanın mümkün olmadığı ancak teorik olarak var olduğu kabul edilen değişkenlerin bireylerdeki düzeylerini ortaya çıkarmayı amaçlayan maddelerden oluşan ölçme araçlarıdır ve söz konusu soyut yapıların bireylerde ne düzeyde olduğunu belirlemek için geliştirilmekte ve bireylere uygulanmaktadır (DeVellis, 2012). Ölçek hazırlamak için iki temel yola başvurulabilir. Bunlardan ilki bireyin ait olduğu kültüre ilişkin özgün bir ölçek geliştirmek, bir diğer yol ise farklı bir kültür için geliştirilmiş, geçerlik ve güvenilirliğine ilişkin kanıtlar toplanmış bir ölçeği hedef kültüre uyarlamaktır (Erkuş, 2007). Gerek geliştirilmiş gerekse uyarlanmış ölçekler kullanılarak, bireyler hakkında bilgi toplanırken, geçerli ve güvenilir ölçmelere dayanarak değerlendirmeler yapmak hedeflenmekte, bundan dolayı da söz konusu ölçeklerin mevcut psikometrik niteliklerinin sorgulanması, teknik özelliklerinin belli standartlara uygunluğunun incelenmesi gerekliliği ortaya çıkmaktadır.

Son yıllarda Türkiye’de ulusal kongrelerde sunulan bildirimler ve dergilerdeki yayınlar incelendiğinde, geliştirilen ve uyarlanan ölçek sayısında büyük artış olduğu görülmektedir (Acar Güvendir ve Özer Özkan, 2015; Çüm ve Koç, 2013). Bu durum ölçme ve değerlendirme alanına ilişkin ilginin artmış olması nedeniyle sevindirici bir gelişme olarak nitelendirilebilir, ancak farklı alanlarda çalışan akademisyenler tarafından geliştirilen ve uyarlanan bu ölçeklerin birçoğu pek çok hata ve eksikliğe sahiptir (Erkuş, 2007). Ölçme ve değerlendirmeye ilişkin bilgi ve uygulama eksikliğinden kaynaklanan sorunlara ek olarak, geliştirilen ölçeklere yönelik gereken izleme çalışmalarının yeterince yapılmıyor olması da söz konusu ölçekler kullanılarak toplanan bilgilerin geçerlik ve

güvenirliğine gölge düşüren bir gelişmedir. Araştırmacılar tarafından çeşitli çalışmalar kapsamında ölçekler geliştirilmekte, deneme uygulamaları yapılmakta ve yürütülen çalışmanın amacına yönelik olarak kullanılmakta, fakat sonrasında testi geliştiriciler dahi testle ilgili gereken izleme çalışmalarını yapmakta yetersiz kalmaktadırlar. Geliştirilen psikolojik testlerin standardizasyon çalışmaları yapılmamakta, dolayısıyla testler ülke çapında yaygınlık kazanamamakta ve diğer araştırmacılar tarafından bilinmemektedir (Gülgöz, 1994; Öner, 1994). Söz konusu psikolojik testlerle geçerli ve güvenilir ölçmeler yapmak için gerek bu ölçme aracını geliştirenler tarafından gerekse farklı araştırmacılar tarafından izleme çalışmaları gerçekleştirilmeli ve bu testlerin teknik nitelikleri yeni çalışmalarla sorgulanmalıdır.

Ölçek geliştirme çalışmalarındaki sıkıntıların benzerleri test uyarlama çalışmaları için de geçerlidir. Farklı kültürler için geliştirilmiş olan bir ölçeğin Türk kültürüne uyarlanması sırasında ortaya çıkan sorunlar birçok çalışmada incelenmiş ve genel olarak ölçülen psikolojik yapının kültürel eşdeğerliği incelenmeden testlerin kullanıma sunulduğu, çeviri işlemlerinin uyarlama olarak kabul gördüğü ve testlerin psikometrik nitelikleri bakımından takibinin ve güncelleştirilmesinin yapılmadığı sonuçlarına ulaşılmıştır. Ayrıca uyarlanan testlerden elde edilen puanların orijinal kültürden elde edilen normlar temel alınarak yorumlandığı da elde edilen sonuçlar arasındadır (Çıkrıkçı-Demirtaşlı, 2007; Çukur, 1999; Gülgöz, 1994; Savaşır, 1994). Bu durum da uyarlanmış psikolojik testler ile yapılan değerlendirmelere kuşku ile yaklaşılmasına neden olmaktadır. Uyarlama çalışmalarıyla ilgili yaşanan bu sorunlara çözüm üretmek, uyarlama çalışmalarının hedef kültüre uygunluğunu sağlamak, uyarlanan testlerden daha anlamlı ve isabetli sonuçlar elde etmek ve yeni testi hedef kültüre kazandırmak için uyarlama çalışmaları sırasında bir takım standartların gözetilmesini gerekmektedir (Cansever, 1982).

Gerek geliştirilmiş gerekse uyarlanmış psikolojik testlere ilişkin ülke çapında yapılan araştırmaların sayısının da ölçek sayısı ile paralel olarak son dönemlerde artış gösterdiği görülmektedir. Bu alandaki ilk çalışmalardan biri olan Özge (1981) tarafından yapılan araştırmada, o tarihlere ülke çapında kullanılmakta olan testler belirlenmiş ve mevcut özellikleri incelenmiştir. Bu çalışmaya göre söz konusu tarihte kullanımda olan 126 psikolojik testin yalnızca beşi Türkiye’de geliştirilmiş özgün psikolojik testler iken, 121 testin ise uyarlama ve çeviri olarak kullanılmakta olan testler olduğu belirlenmiştir. Ayrıca araştırmacı geçerlik, güvenilirlik ve norm çalışmalarının birlikte yapıldığı test sayısının beş olduğunu ve yalnızca altı psikolojik testin el kitabına sahip olduğunu belirlemiştir. Ramazan (1988) tarafından yapılan çalışmada ise psikolojik testlerle ilgili yayımlar, tez ve klinik çalışmalar ile araştırma merkezlerindeki kaynaklar taranmış ve Türkiye’de o tarihte kullanılmakta olan psikolojik test sayısı 164 olarak belirlenmiştir. Bu psikolojik testlerin 36’sının ölçek geliştirme, 60’ının uyarlama çalışması olduğu ve 68’inin ise yalnızca çeviri işlemine tabii tutulduğu saptanmıştır. Kağıtçıbaşı ve Savaşır (1988)’in Doğu Akdeniz ülkelerinde insan yeteneklerinin ölçülmesi ile ilgili yapılan çalışmaları inceledikleri çalışmanın Türkiye ile ilgili bölümünde ise ülke çapında yapılan araştırmaların çoğunun batı kültürlerinde geliştirilen ölçeklerin uyarlanması ile ilgili olduğu ve kültüre özgü yeni ölçeklerin geliştirildiği çalışmalara çok az rastlandığını belirtmiştir (akt. Savaşır, Sezgin ve Erol, 1992).

Psikolojik testlere ilişkin yapılan çalışmalar incelendiğinde, dikkat çeken diğer bir çalışmanın ise 1990 yılında Ölçme ve Değerlendirme Sistemi Özel İhtisas Komisyonu tarafından hazırlanan “Ölçme ve Değerlendirme Sistemi Geliştirme Çalışmaları 1” adlı rapor olduğu görülmektedir. Söz konusu rapor kapsamında, Rehberlik ve Araştırma Merkezleri’nin fiziki ortam ve koşullarının elverişsiz olduğu, psikolojik ölçme araçlarının yetersizliği, hem nicelik hem de nitelik olarak personelin ve bu uygulamalar için ayrılan bütçenin yetersiz olduğu belirtilmiştir. Ayrıca yönlendirme hizmetleri sunan elemanların mesleklerinin gerektirdiği psikolojik ölçme araçları konusunda merkezi ve ulusal düzeyde hizmet sunan bir kuruluşça desteklenmeleri, onların mesleklerini evrensel düzeydeki uygulamalarla bütünleşebilecek bir düzeyde yapabilmelerinin temel koşulu olarak kabul edilmiştir. Böyle bir kuruluşun oluşturulmasıyla hem yönlendirme hizmetlerinde çalışan bireylerin eğitim alabileceği, hem de psikolojik ölçme araçlarıyla ilgili yapılması gereken işlemlerin (uyarlama, geliştirme, dağıtım ve uygulama gibi) tek bir merkezde toplanması gerektiği belirtilmiştir. Bu rapora göre kurulması öngörülen Psikolojik Ölçme Araçları Geliştirme Merkezi’nde Psikolojik Ölçme

Araçları alanında formasyonu olan özellikle üniversitelerden olmak üzere yeterli sayıda uzmandan, yararlanılması gerektiği belirtilmiştir. Ayrıca ÖSYM ve TÜBİTAK gibi kurumlarla işbirliği yapılarak bilgisayar destekli programlardan yararlanılabileceği ve bir proje grubunun kurulmasıyla geliştirilmesi hedeflenen psikolojik testlerin eğitim-öğretim etkinlikleriyle bütünleşik bir şekilde deneme uygulamalarının okullarda yapılabileceği rapor kapsamında belirtilmiştir (MEB, 1990). Özetle söz konusu raporda psikolojik testlere ilişkin ülke çapında yaşanan tüm aksaklıklara dikkat çekilmiş ve çözüm önerisi olarak ise, kurumsallaşmaya gidilmesi gerekliliği belirtilmiştir.

Son yıllarda yapılan çalışmalarda ise genellikle ölçek geliştirme ve uyarlama çalışmalarını temel alan makaleler incelemeye alınmakta ve bu çalışmalar kapsamında geliştirilen ya da uyarlanan ölçeklerin ölçek geliştirme ve uyarlama ilkelerine uygunluğu incelenmektedir. Alanda önde gelen dergilerde yayınlanan makalelerin tarandığı geniş çaplı bir araştırma olan çalışmada, Doğan(2009) tarafından 958 makale incelemeye alınmış ve bu çalışmalarda 389 farklı psikolojik testin kullanıldığı belirlenmiştir. Bu testlerin %75'inin uyarlama, %13'ünün telif ve %12'sinin çeviri psikolojik testler olduğu ve uyarlama çalışmalarının büyük oranda dilsel veya kültürel denklik konusuyla ilgili ölçütleri sağlamadıkları belirlenmiştir. Ayrıca farklı çalışmalarda, ölçek geliştirme ve uyarlama makaleleri incelenmiş (Acar Güvendir ve Özer Özkan, 2015; Boztunç Öztürk, Eroğlu ve Kelecioğlu, 2014; Çüm ve Koç, 2013; Tavşancıl, Güler ve Ayan, 2014) ve geliştirilen ya da uyarlanan ölçeklerin geçerlik ve güvenilirlik gibi temel psikometrik nitelikleri belirlenmeye çalışılmıştır. Araştırmacı tarafından yapılan bir başka çalışmada ise temel amacı ölçek geliştirme olan doktora tez çalışmaları kapsamında geliştirilmiş olan ölçeklerin, test geliştirme ve revizyon standartlarına uyumu incelenmiştir. Beş ölçek geliştirme çalışmasının incelendiği bu çalışma sonucunda, incelemeye alınan tezlerden ikisinin standartları oldukça yüksek düzeyde, birinin orta düzeyde karşıladığı belirlenirken, diğer iki tez çalışması kapsamında geliştirilen ölçeklerin standartları karşılama düzeyinin oldukça düşük olduğu belirlenmiştir (Mor Dirlik, 2014).

Psikolojik testlere ilişkin yapılan inceleme çalışmalarının sayısındaki artış da ölçek geliştirme ve uyarlama konusunda sıkıntılar yaşandığının göstergelerinden biri olarak kabul edilmekte (Acar Güvendir ve Özer Özkan, 2015) ve ölçeklere ilişkin yapılacak izleme çalışmalarının geçerli ve güvenilir ölçümler elde edebilmek için büyük önem arz ettiğini göstermektedir. Psikolojik testlerin geniş kullanım alanları göz önünde bulundurulduğunda, eğitim kurumlarının da psikolojik testlerden sıkça faydalanılan kurumlardan biri olduğu ve bu kurumlarda bireyleri tanımak, teşhis etmek ve yönlendirme yapma gibi amaçlarla kullanılmakta olan psikolojik testlerin de niteliklerinin sorgulanması gerekli olduğu sonucuna ulaşılmaktadır. Söz konusu inceleme için uluslararası alanda kabul görmüş, geçerlik, güvenilirlik ve ölçek geliştirme ve uyarlama standartlarının bir kontrol listesi formunda hazırlanıp uygulayıcılara ve araştırmacılara sunulması, ölçek geliştirme ve uyarlama aşamasındaki sorunlara çözüm olabileceği gibi hali hazırda kullanımda olan ölçeklerin de psikometrik niteliklerinin belirlenmesini sağlayabileceği düşünülmektedir.

Araştırmanın Amacı

Ülke çapında bu alanda yapılan çalışmaların sınırlılığı gözetilerek, hem psikolojik testlerin temel psikometrik nitelikleri olan geçerlik ve güvenilirliğini sorgulamak hem de bir kontrol listesi oluşturarak alana katkıda bulunmak için, bu çalışmada Türkiye'de eğitim kurumlarından sıklıkla kullanılan bir grup psikolojik test belirlenerek, bu testlerin American Psychological Association(APA) ve International Testing Commission (ITC) tarafından belirlenmiş olan temel test standartlarına ne derece uygunluk gösterdiğinin araştırılması, mevcut durumlarının betimlenmesi ve varsa geliştirilmesi gereken niteliklerinin belirlenmesi amaçlanmıştır. Bu sayede dünya çapında geçerliliği kabul edilen standartların sıklıkla kullanılan psikolojik testlere uygulanması sağlanarak bu çalışmanın daha sonradan yapılacak çalışmalara örnek oluşturması ve psikolojik testlerin gelişimine katkıda bulunması beklenmektedir.

YÖNTEM

Bu araştırma, tarama modeli kapsamında yer alan betimsel bir araştırmadır. Betimsel araştırmalarda, var olan bir durumun, herhangi bir müdahale olmaksızın, olduğu gibi tanımlanması esastır ve bu tanımlamalar nitel ya da nicel yaklaşıma göre yapılmaktadır (Karasar, 2010). Bu çalışmada nitel yaklaşım tercih edilmiş, çalışmanın verileri doküman inceleme yöntemi ile elde edilmiştir. Elde edilen veriler ise betimsel ve içerik analizi yöntemlerinden yararlanılarak analiz edilmiştir (Yıldırım ve Şimşek, 2011). Bu bağlamda çalışma kapsamında incelemeye alınan psikolojik testlerin belirlenen standartlar gözetilerek derinlemesine incelenmesi ve mevcut durumlarının betimlenmesi amaçlanmıştır.

İncelemeye Alınan Psikolojik Testler

Araştırmanın amacı doğrultusunda standartları karşılama düzeyleri incelenecek psikolojik testlerin belirlenmesi için seçkisiz olmayan örnekleme yöntemlerinden amaçlı örnekleme yöntemi kullanılmış ve uzman görüşleri temel alınarak incelenecek psikolojik testler belirlenmiştir. Rehberlik ve Araştırma Merkezleri ile okulların rehberlik servislerinde sıklıkla kullanıldığı belirlenen testlerden oluşturulan bir anket Ankara ve Bursa'da devlet okullarında görev yapmakta olan psikolojik danışmanlardan oluşan 79 kişilik bir gruba uygulanmış ve katılımcılardan, ankette yer alan psikolojik testlerden geliştirilmesi gereken yönleri bulunanları belirtmeleri istenmiştir. Anket uygulaması sonrasında, uzmanların görüşlerine göre üzerinde çalışma yapılması gerekli görülen psikolojik testler belirlenmiştir. Bu psikolojik testler uzmanlar tarafından belirtilen sıklığa göre şu şekildedir; Wechsler Çocuklar İçin Zekâ Ölçeği Revize Edilmiş Form (WISC-R), Temel Kabiliyetler Testi 7-11 Yaş Formu (TKT 7-11), Akademik Benlik Kavramı Ölçeği (ABKÖ), Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı (GEÇDA) ve Stanford Binet Zekâ Testi'dir. Uzmanlar tarafından en çok belirtilen bu beş psikolojik testin, ikisi (Stanford Binet Zekâ Testi ve Wechsler Çocuklar İçin Zekâ Testi) aynı değişkeni, zekâyı, ölçtüğü için bu çalışma kapsamına yalnızca birinin alınması kararlaştırılmış ve WISC-R araştırma kapsamına alınmıştır. Böylece uzmanların görüşü ölçüt olarak alınarak araştırma çerçevesinde incelenecek olan dört psikolojik test belirlenmiştir. Bu çalışma kapsamında incelemeye alınan dört psikolojik testin temel özellikleri aşağıda verilmektedir.

Akademik benlik kavramı ölçeği (ABKÖ): Bireylerin dört yetenek ve 12 ilgi alanına ilişkin benlik kavramını ölçmeyi amaçlayan bir psikolojik testtir. Bu psikolojik test ile ölçülen yetenekler; sözel yetenek, sayısal yetenek, şekil uzay yeteneği ile göz el koordinasyonu iken ölçülen ilgi alanlar ise, fen bilimleri ilgisi, sosyal bilimler ilgisi, ziraat ilgisi, mekanik ilgi, ikna ilgisi, ticaret ilgisi, ayrıntı ilgisi, edebiyat ilgisi, güzel sanatlar ilgisi, müzik ilgisi ve sosyal yardım ilgisidir. Söz konusu psikolojik test dört seçenekli 170 maddeden oluşmaktadır. Testin uygulaması ve puanlaması tamamen testi alan birey tarafından yapılmaktadır. Ölçeğin el kitabında puanlama ve profil oluşturmanın testi alan birey tarafından yapılması gerektiği özellikle vurgulanmış ve bunun dışında yapılan herhangi bir uygulamada ölçekten beklenen faydanın sağlanamayacağı belirtilmiştir. Söz konusu psikolojik testin, testi alandan farklı bir birey tarafından puanlanması durumunda, testi alan bireyin kendini tanıma sorumluluğunun üzerinden alındığına, bu nedenle cevapların gelişigüzel verilebileceğine dair uyarılarda bulunulmuştur (Kuzgun, 2011).

Gazi erken çocukluk gelişimi değerlendirme aracı (GEÇDA): Bu psikolojik test, 0-72 ay çocuklarının gelişimlerinin ayrıntılı olarak değerlendirilmesini, eğitim yaşantılarının düzenlenmesini ve çocuklardaki olası gelişimsel geriliklerin erken tanınması amaçlanarak geliştirilmiş bir gelişim değerlendirme aracıdır. Psikomotor, bilişsel, dil gelişimi ve sosyal-duygusal gelişimi ölçmeyi amaçlayan dört alt test ve toplam 249 maddeden oluşan test, gelişimsel oyunlar sırasında çocuğun gözlenmesi yoluyla uygulanmakta ve standart bir materyal seti ve el kitabı ile kullanılmaktadır. Ayrıca testi uygulayacak bireylerin "kullanıcı sertifikası" na sahip olması gerekmektedir (Temel, Ersoy, Avcı ve Turla, 2005).

Temel kabiliyetler testi 7-11 yaş formu (TKT): Çok faktörlü zeka kuramının önde gelen savunucularından olan Thurstone tarafından geliştirilen bu test, çoklu faktör kuramına

dayanmaktadır. Thorndike çoklu faktör kuramı çerçevesinde, kelime anlamı, sayısal akıl yürütme, kavrama, ilişkileri görsel algılama gibi faktörlerden söz ederek, zekâyı soyut zekâ, sosyal zekâ ve mekanik zekâ olarak sınıflandırmıştır (Anastasi,1988). Thurstone tarafından ilk olarak 1941 yılında 11-17 yaşları arasındaki çocuklar için geliştirilen bu test daha sonra, 5-7, 7-11 ve 11-17 olarak üç ayrı form halinde düzenlenmiştir. Bu çalışma kapsamında incelenen TKT 7-11 formu, yedi alt teste sahip bir grup testidir. Bu alt testler; “kelimeler”, “yer kavramı”, “resimler”, “kelime gruplaması”, “şekil gruplaması”, “ayırt etme”,ve “hesap” testleridir. Türkçeye uyarlaması ilk olarak 1953 yılında Test Araştırma Bürosu tarafından yapılan test, 2001 yılında maddeler üzerinde çeşitli redaksiyonlar yapılarak tekrar güncelleştirilmiştir. Testin ikinci uyarlama çalışması, Özel Eğitim ve Rehberlik Genel Müdürlüğü tarafından 1998 yılında başlatılmıştır ve 2001 yılında son bulmuştur. Çalışma dâhilinde, maddeler üzerinde yapılan araştırmalarla, madde güçlük ve ayırt edicilik indeksleri hesaplanmış ve bu özellikler bakımında istenilen düzeyde olmayan maddeler ölçekten çıkarılmış, 256 maddeyle başlanan çalışma 181 madde ile sonlanmıştır. Testin tümü ve alt ölçekler için güvenilirlik analizleri yapılmıştır. Geçerlik çalışmalarında ise yordama geçerliği hesaplanmış ve ölçüt olarak ders ve sınıf geçme notları kullanılmıştır. Norm çalışmaları da yapıldıktan sonra uyarlanan testin beş özel bir genel yetenek puanı veren bir test olarak Rehberlik ve Araştırma Merkezleri ile okulların rehberlik servislerinde kullanılması uygun bulunmuştur (MEB, 2001a). Testin genel özellikleri incelendiğinde, testi uygulayacak bireylerin psikoloji, rehberlik ve psikolojik danışma veya eğitimde psikolojik hizmetler lisans programı mezunu olmaları ve bununla birlikte; TKT (7 – 11) ile ilgili uygulama sertifikasının verildiği eğitim yaşantılarına katılmış olmaları gerekmektedir. Uygulama yapılacak grup aynı yaştaki bireylerden oluşturulabileceği gibi, farklı yaşlardaki bireylerden de oluşturulabilir. Test uygulanan grup 7 – 11 yaş grubuna dahil ve 10 kişi olmalıdır. Test toplam 80 dakikada uygulanabilmektedir ve her alt test için farklı süreler ayrılmıştır(MEB, 2001b)

Wechsler çocuklar için zekâ ölçeği revize edilmiş form (WISC-R): WISC-R, 6-16 yaş arasındaki bireylere, bireysel olarak uygulanan bir zekâ testidir. Ölçek sözel ve performans bölümlerinden ve 12 alt testten oluşmaktadır. Orijinal formu İngilizce olan bu testin, Türkçeye uyarlama çalışmasında, testin 1949 yılında geliştirilmiş formunun Türkçe çevirisi ele alınmış, fakat yalnızca çevirisi yapılan ölçeğin uygulama için yeterli olmadığı ve bir takım değişikliklerin gerekli olduğu konusunda karara varılarak yeni bir standardizasyon çalışması başlatılmıştır. Bu çalışma kapsamında ilkökul kitapları taranarak , “Genel bilgi”, “Benzerlikler” ve “ Yargılama” alt testleri için sorular seçilmiş ve denenmiştir. Elde edilen verilere göre, bu alt testlerdeki soruların güçlük sıralaması yeniden yapılmıştır. “Sözcük dağarcığı” altesti için ilkökul kitapları taranmış 36.000 sözcük elde edilmiş ve bunların 36 tanesi seçilerek, bu alt test yeniden oluşturulmuştur. Performans bölümüne dahil olan alt testlerin üçü, resim tamamlama, resim düzenleme ve küplerle desen, 64 kişilik bir çocuk grubuna uygulanmış ve madde analizleri yapılmıştır. Her alt testteki maddeler için madde analizleri yapılmış, maddeler güçlük değerlerine göre yeniden sıralanmıştır. Tüm bu çalışmalardan sonra, bu psikolojik testin Türkiye normlarının elde edilmesi için daha geniş bir örneklem üzerinde uygulama yapılmış ve test için normlar oluşturulmuştur (Savaşır ve Şahin, 1995).

Veri Toplama Yöntemleri

Çalışma kapsamında incelemeye alınacak standart testlerin belirlenmesinden sonra, söz konusu psikolojik testlerin el kitaplarına ulaşılmış ve bu testler APA tarafından geliştirilen Eğitimde ve Psikolojide Ölçme Standartları ile ITC tarafından geliştirilen Test Uyarlama Standartları temel alınarak hazırlanan kontrol listesine göre incelenmiştir.

Kontrol listesinin hazırlanması aşamasında, APA tarafından 1999 yılında İngilizce olarak yayınlanan ”Standarts for Educational and Psychological Testing (Eğitim ve Psikolojide Ölçme Standartları)” kitabı incelenmiş ve bu kaynakta yer alan standartlardan geçerlik, güvenilirlik ve test geliştirme ve revizyon ile ilgili standartlar araştırmacı tarafından Türkçeye çevrilmiştir. Belirlenen standartlar yalnızca çeviri işleminden farklı olarak, dilsel anlamlılık da göz önünde bulundurularak yalınlaştırılmış ve kullanıcıların daha rahat anlayabileceği şekilde düzenlenmiştir. Bu standartlara ek olarak ITC tarafından belirlenen test uyarlama standartları üzerinde çalışılmış ve bu standartlar da

yalınlaştırılarak Türkçeye çevrilmiştir. Tüm bu standartlar kullanılarak bir kontrol listesi oluşturulmuştur. Uyarılama yaklaşımı ile Türkçeye çevrilen ve kontrol listesi haline getirilen bu form, beş ölçme değerlendirme uzmanı, iki türk dili uzmanı tarafından incelenmiş ve uzman görüşleri temel alınarak oluşturulan kontrol listesi üzerinde çeşitli revizyonlar yapılmıştır. İlgili tüm standartları içeren uzun form ve APA tarafından tüm psikolojik testler tarafından karşılanması gereken temel test standartları olarak belirtilen temel standartlardan oluşan kısa form olmak üzere kontrol listesinin iki ayrı formu oluşturulmuştur.

Araştırma kapsamına alınan psikolojik testler, el kitaplarında yer alan bilgiler doğrultusunda hazırlanan kısa form kontrol listesi çerçevesinde incelenmiştir. Araştırmaya dahil edilen dört psikolojik testten uyarlanmış olan TKT 7-11 Yaş Formu ve WISC-R sadece uyarılama standartlarına göre, diğer iki test, GEÇDA ve ABKÖ, ise geçerlik, güvenirlik ve test geliştirme ve revizyon standartlarına göre incelenmiştir. Psikolojik testlerin belirlenen standartları karşılama düzeyleri 1'den 3'e kadar puanlandırılarak belirlenmiş ve söz konusu standardın incelenen psikolojik test için hiç karşılanmadığı durumda 1 puan, kısmen karşılandığı durumda, 2 puan ve tamamen karşılandığı durumda ise 3 puan verilmiştir. Yapılan test incelemelerinin güvenirliğini saptamak amacıyla incelemeler araştırmacı tarafından üç hafta arayla tekrarlanmış ve iki inceleme arasındaki tutarlılık uyum yüzdesi indeksi kullanılarak hesaplanmıştır. İki inceleme arasındaki sürenin ne kadar olacağı konusunda alan yazında net bir bilgi bulunmazken, bu zaman aralığının çok uzun tutulması araştırmacının konuya tamamen yabancılaşmasına, çok kısa tutulması ise önceki incelemenin etkisi altında kalmasına neden olabileceği belirtilmektedir (Tavşancıl ve Arslan, 2001). Bu gibi olumsuz etkilerin araştırma sürecine dâhil olmasını önlemek için, iki inceleme arasındaki zaman üç hafta olarak belirlenmiştir.

Verilerin Analizi

Araştırmanın amaçları çerçevesinde elde edilen verilerin analiz edilmesi için içerik analizi tekniği kullanılmıştır. Elde edilen ham verilerin sayısallaştırılması için frekans dağılımından yararlanılmıştır. Hazırlanan kontrol listesindeki her bir psikolojik test standart için frekans tablosu oluşturulmuş ve böylece her psikolojik testin, kontrol listesindeki standartları karşılama durumu belirlenmiştir. Araştırmacı tarafından iki farklı zamanda yapılan incelemeler arasındaki tutarlılığın belirlenmesi için ise uyum yüzdesi indeksi değeri hesaplanmıştır. Uyum yüzdesi indeksi, aynı kodlamanın yapıldığı durumların mevcut tüm durumlara oranı hesaplanarak bulunan bir değerdir. Puanlayıcılar arası ya da puanlayıcı içi uyum yüzdesinin %70'den daha yüksek olması beklenmektedir (Tavşancıl ve Arslan, 2001).

BULGULAR

İncelemeye alınan psikolojik testlerin hazırlanan kontrol listesine uygunluğunun belirlenmesi için yapılan tekrarlı incelemelere ilişkin hesaplanan uyum yüzdesi indeksleri Tablo 1'de verilmiştir.

Tablo 1. İncelemeler Arası Uyum Yüzdesi İndeksi

Psikolojik Testler	Uyum Yüzdesi İndeksi
WISC-R	0.90
TKT 7-11	0.83
GEÇDA	0.78
ABKÖ	0.94

Tablo 1'de verilen değerler incelendiğinde, tüm testler için hesaplanan uyum yüzdesi indeksi değerlerinin 0.70'den yüksek olduğu görülmektedir. Uyum yüzdesi indeksinin 0.70 değeri ve

üstünde olması kabul edilebilir bir uyumu gösterdiğinden (Tavşancıl ve Arslan, 2001), her psikolojik test için yapılan incelemeler arasındaki uyumun kabul edilebilir olduğu sonucuna ulaşılmaktadır.

Geçerlik Standartlarının Karşılama Durumuna Yönelik Bulgular

Kontrol listesinin ilk bölümünde yer alan 10 geçerlik standardına göre incelenen GEÇDA ve ABKÖ'nin bu standartları karşılama durumu ve geçerlik standartları Tablo 2'de sunulmuştur.

Tablo 2. Geçerlik Standartlarının GEÇDA ve ABKÖ için Karşılama Durumuna İlişkin Özet Tablo

Geçerlik Standartları	GEÇDA	ABKÖ
Test puanlarının tüm yorumlarına ilişkin geçerlik kanıtları sunulmuştur.	Hiç karşılanmadı	Hiç karşılanmadı
Testin uygulanacağı grup/gruplar sınırlandırılmış ve testle ölçülecek yapı betimlenmiştir.	Tamamen karşılanadı.	Tamamen karşılanadı.
Testin amacı ve kullanımı ile tanımlanan evrenin bağlantısı kurulmuştur. Test kapsamı belirlenirken uygulanan işlemler açıklanmıştır.	Kısmen karşılanadı.	Kısmen karşılanadı.
Geçerlilik kanıtlarının elde edildiği grup ayrıntılı bir şekilde tanıtılmıştır.	Kısmen karşılanadı.	Kısmen karşılanadı.
Test maddelerinin kapsam geçerliği ile ilgili uzman yargısına başvurulmuşsa uzmanların eğitimi, deneyimi ve nitelikleri betimlenmiştir.	Kısmen karşılanadı.	Hiç karşılanmadı
Testin belli maddeleri ya da bazı alt testleri için özel yorumlar önerilmişse, bununla ilgili kanıtlar ve gerekçeler sunulmuştur.	Hiç karşılanmadı	Hiç karşılanmadı
Bazı test maddelerinin olası fakat kabul edilmeyecek yorumları varsa, puanlayıcı buna karşı uyarılmıştır.	Hiç karşılanmadı	Hiç karşılanmadı
Testin geçerliği bir ya da birden fazla ölçüte dayanarak belirlendiyse, bu ölçütlerin uygunluğu ve nitelikleri belirtilmiştir.	Kısmen karşılanadı.	Tamamen karşılanadı.
Teste dayalı olarak bileşik puanlar geliştirilmişse, alt puanların ağırlığına ilişkin temeller ve mantık açıklanmıştır.	Hiç karşılanmadı	Hiç karşılanmadı
Alt test puanlarının yorumlanmasında ve bu testler arasındaki puan farklarının yorumlanmasında kullanılacak bilgiler açıkça verilmiştir.	Hiç karşılanmadı	Tamamen karşılanadı.

Türkiye'de geliştirilen GEÇDA ve ABKÖ için incelemeye alınan geçerlik standartları ve bu standartların testlerce karşılama durumu Tablo 2'de görülmektedir. GEÇDA için yapılan incelemelerde 10 geçerlik standardının yalnızca birinin bu test için tamamen karşılanağı, dördünün kısmen karşılanağı ve beşinin hiç karşılanağı belirlenmiştir. İncelemeye alınan diğer psikolojik test olan ABKÖ için de benzer bir durum söz konusudur. Geçerlik standartlarından beşi bu test için hiç karşılanağken, ikisi kısmen, üçü ise tamamen karşılanağıdır.

Güvenirlik Standartlarının Karşılama Durumuna Yönelik Bulgular

Güvenirlik standartlarına uygunluğu incelenen ABKÖ ve GEÇDA için kontrol listesinin güvenilirlik standartları bölümünde yer alan 10 standardın tümü testlerin yapısı gereği incelemeye alınamamıştır. ABKÖ için sekiz standart incelemeye alınırken GEÇDA için dokuz güvenilirlik standardı incelenebilmiştir. İncelemelerde kullanılan standartlar ve testlerin bu standartları karşılama durumu Tablo 3'te verilmiştir.

Tablo 3. Güvenirlik Standartlarının GEÇDA ve ABKÖ için Karşılanma Durumuna İlişkin Özet Tablo

Güvenirlik Standartları	GEÇDA	ABKÖ
Yorumlanacak her toplam puan, alt test puanı ya da puanların kombinasyonu için güvenilirliğe ilişkin tahminler sunulmuştur.	Hiç karşılanmadı	Hiç karşılanmadı
Test puanlarının güvenilirliğinin hesaplanmasında kullanılan her yöntem açıkça tanımlanmış ve bu yöntemlerde kullanılan istatistiksel teknikler sunulmuştur.	Tamamen karşılanadı.	Hiç karşılanmadı.
Güvenirlik analizlerinde ve betimsel analizlerde kullanılan örneklem grubunun seçilme süreçleri rapor edilmiştir.	Tamamen karşılanadı.	Hiç karşılanmadı.
Testin uygulanma sürecinde önemli değişikliklere izin veriliyorsa, örneklem sayısı yeterli oldukça her uygulamada elde edilen puanlar için ayrı güvenirlik analizleri sunulmuştur.	Hiç karşılanmadı	Hiç karşılanmadı
Testin puanlanmasına öznel yargıların karışması olasılığı söz konusu ise, puanlayıcılar arası tutarlılık ve puanlayıcının birden fazla uygulanan ölçme uygulamalarındaki tutarlılığına dair kanıtlar sağlanmıştır.	Tamamen karşılanadı.	İncelenmeye alınmadı.
Testin birkaç farklı sınıf ya da yaş seviyesinde gruplarda uygulanması söz konusu ise ve her sınıf ya da yaş grubu için ayrı normlar sunulmuşsa, her grup ya da yaş seviyesi için ayrı güvenirlik bilgileri sunulmuştur.	Tamamen karşılanadı.	Hiç karşılanmadı
Farklı alt gruplarda beklenen ölçmenin standart hatasının ve güvenirlik katsayısının büyük oranda değişebileceğine dair genel olarak deneysel ve teorik sebepler varsa, her grupla ilgili güvenirlik bilgisi sağlanmıştır.	Hiç karşılanmadı	Hiç karşılanmadı
Test maddelerinin bir kısmı ya da alt testler kısmen farklı bir eğilimi ya da yeteneği ölçmek üzere hazırlanmışsa, güvenirlik tahmini için uygulanan işlemler testin çok faktörlü yapısına uygun olarak uygulanmıştır.	Hiç karşılanmadı.	Hiç karşılanmadı..
Ölçmenin standart hatası testin önerilen tüm yorumlamalarında kullanılacak puanlar için rapor edilmiştir.	Hiç karşılanmadı	Hiç karşılanmadı

Tablo 3 incelendiğinde, yapılan incelemeler sonucunda, ABKÖ için incelemeye alınan güvenirlik standartların hiçbirinin karşılanmadığı, GEÇDA için incelenen dokuz güvenirlik standardının beşinin hiç karşılanmadığı, dört standardın ise bu test için tamamen karşılanadığı belirlenmiştir. Bu durumda her iki test için yeni güvenirlik kanıtlarına ihtiyaç duyulduğu fakat özellikle ABKÖ için bu durumun daha gerekli olduğu sonucuna ulaşılmıştır.

Test Geliştirme ve Revizyon Standartlarının Karşılanma Durumuna Yönelik Bulgular

Çalışma kapsamında geliştirilen kontrol listesinin üçüncü bölümünde, toplam 15 standarttan oluşan “Psikolojik testlerin sahip olması gereken test geliştirme ve revizyon standartları” yer almaktadır ve Türkiye’de geliştirilmiş olan ve araştırmaya alınan GEÇDA ve ABKÖ için bu standartlar da incelemeye alınmıştır. Test geliştirme ve revizyon bölümünde yer alan 15 standarttan GEÇDA için 14’ü incelemeye alınırken, ABKÖ için 12 standart incelemeye alınmıştır. Her iki test için de bazı standartlar inceleme dışında tutulmuştur. Bunlardan GEÇDA için inceleme dışında tutulan standart “farklı ağırlıklandırılmış maddelerden oluşan testler için incelenmesi gereken bir standarttır” ve bu testin tamamı eşit ağırlıklandırılmış maddelerden oluştuğu için, bu standart GEÇDA için inceleme dışında bırakılmıştır.

İncelemeye alınan diğer psikolojik test olan ABKÖ için bu bölümde incelemeye alınmayan üç standart bulunmaktadır. Bunlardan ilki testte açık uçlu maddeler bulunması durumunda puanlayıcıya sunulması gereken puanlama kriterleri ile ilgili standarttır. Söz konusu testte açık uçlu bir madde bulunmadığı için bu standardın bu test için inceleme dışı tutulmuştur. Bu test için inceleme dışında tutulan diğer bir standart ise GEÇDA için de incelemeye alınmayan ve test maddelerinin farklı ağırlıklandırılmış olması durumunda işe koşulan standarttır. ABKÖ eşit ağırlıklandırılmış maddelerden oluştuğu için bu standart da incelenmeye alınmamıştır. Bu test için inceleme dışı tutulan son standart ise puanlayıcıların seçimi ve eğitimi ile ilgilidir ve ABKÖ bireyin kendi

puanlamasına dayandığı için tek puanlayıcı bireydir ve farklı puanlayıcılara ihtiyaç duyulmuyor olması bu standardın inceleme dışında tutulmasının gerektirmiştir. Her iki psikolojik test için de incelen test geliştirme ve revizyon standartlarını ve standartların bu testler için karşılanma durumu özet olarak Tablo 4’te sunulmuştur.

Tablo 4. Test Geliştirme ve Revizyon Standartlarının GEÇDA ve ABKÖ için Karşılanma Durumuna İlişkin Özet Tablo

Test Geliştirme ve Revizyon Standartları	GEÇDA	ABKÖ
Test geliştirme sürecine dayanan yeterli sayıda kanıt toplanmış ve rapor edilmiştir.	Hiç karşılanmadı	Hiç karşılanmadı.
Testin amacı / amaçları, test edilen özelliğin tanımı ve testin özellikleri ve maddeleri açıkça belirtilmiştir.	Kısmen karşılandı.	Tamamen karşılandı.
Test yönergeleri testin kapsamını, testte bulunan madde sayısını, madde formatını ve madde ve alt test düzenlemelerini tanımlamıştır.	Tamamen karşılandı.	Kısmen karşılandı.
Madde türleri, cevaplama biçimleri, puanlama ve test uygulama işlemleri, testin amacına, ölçülecek özelliğe ve testi alması hedeflenen grubunun özelliklerine dayanılarak seçilmiştir.	Tamamen karşılandı.	Tamamen karşılandı.
Test yönergeleri alanla ve ölçme ile ilişkili, test geliştirenler dışında kalan uzmanlar tarafından gözden geçirilmiştir.	Tamamen karşılandı.	Hiç karşılanmadı.
Maddelerin geliştirilmesi, gözden geçirilmesi, denenmesi ve madde havuzundan seçilmesinde uygulanan işlemler belirtilmiştir.	Kısmen karşılandı.	Hiç karşılanmadı.
Madde seçiminin hangi süreçle yapıldığı, madde seçiminde hangi verilerin kullanıldığı, madde güçlüğü, madde ayırt ediciliği gibi, rapor edilmiştir.	Kısmen karşılandı.	Kısmen karşılandı.
Testin gözden geçirilme süreci, deneysel analizleri ve maddelere ve cevaplama biçimlerine yönelik uzman yargılarına başvurulmuştur.	Kısmen karşılandı.	Hiç karşılanmadı.
Test maddeleri test yönergelerine göre farklı kategorilere ya da alt testlere ayrılmışsa sınıflama işlemleri ve bu sınıflamanın netliği ile uygunluğu belirtilmiştir.	Tamamen karşılandı.	Hiç karşılanmadı.
Testi cevaplayanların açık uçlu maddelerdeki performansını puanlarken kullanılacak kriterler belirtilmiştir.	Tamamen karşılandı.	İncelemeye alınmadı.
Uygulayıcı için hazırlanan yönergeler yeterince açıktır ve vurgulanarak sunulmuştur.	Tamamen karşılandı.	Tamamen karşılandı.
Testi cevaplayacaklar için hazırlanan yönergeler yeterince ayrıntılıdır. Uygun durumlarda materyal ve alıştırma örnekleri ya da soru örnekleri verilmiştir.	Tamamen karşılandı.	Tamamen karşılandı.
Test geliştirici testin uygulama koşullarının kabul edilebilir değişiklikler belirtilmiş ve bu farklı koşulların kabul edilebilir olma sebepleri belgelendirilmiştir.	Hiç karşılanmadı.	Tamamen karşılandı.
Puanlayıcıların seçimi, eğitimi ve puanları tanımlama/betimleme niteleme süreçleri test geliştirici tarafından rapor edilmiştir.	Kısmen karşılandı.	İncelemeye alınmadı.

Tablo 4’te GEÇDA ve ABKÖ için incelenen test geliştirme ve revizyon standartları yer almakta ve testlerin söz konusu standartları karşılama durumu görülmektedir. İncelemeler sonucu yapılan değerlendirmelere göre GEÇDA için incelenen iki standart hiç karşılanmazken, beş standart kısmen ve yedi standart ise tamamen karşılanmaktadır. ABKÖ için yapılan incelemelerde ise incelemeye alınan on iki standardın beşinin hiç karşılanmadığı, ikisinin kısmen ve beşinin ise tamamen karşılandığı belirlenmiştir.

Test Uyarlama Standartlarının Karşılanma Durumuna Yönelik Bulgular

Kontrol listesinin son bölümünde uyarlanmış olan testler için incelenmesi gereken standartlar yer almaktadır. Bu bölümde 15 standart bulunmakta ve araştırma kapsamında incelenen testlerden Türkçeye uyarlanmış olan psikolojik testler, Wechsler Çocuklar İçin Zeka Ölçeği Revize Edilmiş Form ve Temel Kabiliyetler Testi 7-11 Yaş Formu bu standartlara göre incelenmiştir. Her iki

psikolojik test için incelemeye alınan test uyarlama standartları ve tetlerin bu standartları karşılama durumu Tablo 5'te sunulmuştur.

Tablo 5. WISC-R ve TKT 7-11 Yaş Formu için Uyarlama Standartlarının Karşılama Durumuna İlişkin Özet Tablo

Test Uyarlama Standartları	WISC-R	TKT 7-11
Çalışmanın ana amaçları göz önünde bulundurularak testin orijinal formu ile hedef dildeki formu arasındaki kültürel farklılıkların etkileri olabildiğince azaltılmıştır.	Tamamen karşılandı.	Hiç karşılanmadı.
Çalışma yapılan evrendeki test veya ölçekler tarafından ölçülen özelliğin örtüşme düzeyi belirlenmiştir.	Tamamen karşılandı.	Hiç karşılanmadı.
Testi uyarlayanlar araştırma yapılan evrende bütün dilsel ve kültürel farklılıkları göz önünde bulundurularak, testin uygulanacağı evrenin kültürel ve dil yapısına uygun test yönergeleri, test puanları ve test maddeleri oluşturmuştur.	Tamamen karşılandı.	Hiç karşılanmadı.
Testi uyarlayanlar test tekniklerinin seçimi, madde formatı, test düzeni ve diğer yapılması gereken işlemleri araştırmanın yapılacağı evrene uygun olmasını sağlamıştır.	Tamamen karşılandı.	Hiç karşılanmadı.
Testi uyarlayanlar uyarlanan ölçeğin geçerliliğine dair bilgileri toplamıştır.	Tamamen karşılandı.	Tamamen karşılandı.
Testi uyarlayanlar, uyarlama sürecinin doğru işlemesi ve diller arası eşitlik sağlanması için hem dilsel hem de psikolojik açıdan sistematik bir yol izlemiştir.	Tamamen karşılandı.	Hiç karşılanmadı.
Uyarlama işlemleri sırasında, uyarlanan her maddenin orijinal testteki maddelere denk olması gözetilmiştir.	Tamamen karşılandı.	Hiç karşılanmadı.
Testin uygulanmasını etkileyebilecek çevresel faktörler hedef gruplarda orijinal gruplara olabildiğince denk olacak şekilde düzenlenmiştir.	Tamamen karşılandı.	Tamamen karşılandı.
Testi uyarlayanlar olası problemleri öngörüp bu problemleri çözebilmek için uygun materyal ve yönergeleri hazırlamışlardır.	Tamamen karşılandı.	Tamamen karşılandı.
Testin el kitabı testle ilgili bütün özellikleri ve testin yeni bir kültüre uyarlanması durumundaki tüm gereksinimleri içermektedir.	Tamamen karşılandı.	Tamamen karşılandı.
Testin uygulama eğitimleri hedef dilde gerçekleştirilmiştir.	Tamamen karşılandı.	Tamamen karşılandı.
Testin el kitabında testle ilgili uyulması gereken kesin kurallar bulunmaktadır.	Tamamen karşılandı.	Tamamen karşılandı.
Bir test başka bir kültüre uyarlandığı zaman, testin uyarlanmış hali ve orijinal halinin eşitliğini desteklemek açısından yapılan değişimler raporlaştırılmıştır.	Hiç karşılanmadı.	Hiç karşılanmadı.
Testi uyarlayanlar testin uygulanacağı hedef evrendeki bireylerin testteki performansları etkileyebilecek sosyo-kültürel ve ekolojik özellikler ile ilgili spesifik bilgilere sahiptir ve bu özelliklerin test sonuçlarının yorumlanması ile ilgili etkilerini belirlemişlerdir.	Tamamen karşılandı.	Tamamen karşılandı.
Testin uygulandığı evrenlerin örneklemeleri arasındaki puan farkları görünen değer olarak kabul edilmemiş, ampirik kanıtlarla değişkenlerin anlamlılığı kanıtlanmıştır.	Tamamen karşılandı.	Tamamen karşılandı.

Tablo 5 incelendiğinde çalışmaya alınan uyarlanmış testler olan WISC-R ve TKT 7-11 Yaş Formu için ITC tarafından geliştirilen test uyarlama standartlarının karşılama durumu görülmektedir. İncelenen testlerden WISC-R için bu standartların karşılama durumuna bakıldığında, yalnızca bir standardın hiç karşılanmadığı, kalan 14 standardın ise bu test için tamamen karşılandığı görülmektedir. Uyarlama çalışması olarak Türk kültürüne kazandırılan ve bu çalışma kapsamında incelenen bir diğer psikolojik test olan TKT 7-11 Yaş Formu için ise WISC-R'dan farklı bir durum söz konusudur. Bu test için incelemeye alınan yedi standart hiç karşılanmazken, sekiz standart tamamen karşılanmaktadır. Tablo 5'te yer alan bilgilere göre WISC-R'ın uyarlama çalışmasının oldukça nitelikli olduğu fakat aynı durumun TKT 7-11 Yaş Formu için geçerli olmadığı sonucuna ulaşılabılır.

İncelemeye Alınan Psikolojik Testlerin Geliştirilmesi Gereken Yönlerine İlişkin Bulgular

APA tarafından geliştirilen Psikolojide ve Eğitimde Ölçme Standartları ile ITC tarafından geliştirilen Test Uyarılma Standartları temel alınarak hazırlanan kontrol listesine göre dört psikolojik test incelenmiş ve bu testlerin söz konusu standartları karşılama durumları belirlenmiştir. Yapılan incelemeler sonucunda her psikolojik testin incelemeye alınan standartlara göre eksik yönleri tespit edilmiş ve yapılması gereken güncellemeler belirlenmiştir. İncelenen psikolojik testlerden, Türkiye’de geliştirilmiş olan Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı’nın geçerlik standartlarını karşılama düzeyinin düşük olduğunu ve GEÇDA’ nın geçerlik standartları temel alınarak tekrar incelenmesi gerekliliğini ortaya çıkarmıştır. GEÇDA’ nın psikolojik testlerin sahip olması gereken geçerlik standartlarını karşılama düzeyinin artırılması için yapılması gerekenler şunlardır:

- Testle ölçülen yapı detaylı olarak betimlenmelidir.
- Test puanlarının ilişkin yapılan yorumlamaların dayandığı kanıt ve teoriler sunulmalıdır.
- Testin amacı ile testle temsil edilen evrenin bağlantısı kurulmalıdır.
- Geçerlik kanıtlarının elde edildiği gruplar detaylı bir biçimde tanıtılmalıdır.
- Test maddelerini inceleyen uzmanların deneyimleri sunulmalıdır.
- Test yanıtlarını değerlendirirken puanlayıcıların yapabileceği olası fakat kabul edilmeyecek yorumlar için puanlayıcılar uyarılmalıdır.
- Geçerlik kanıtlarının toplandığı koşullar rapor edilmelidir.

GEÇDA için yapılan güvenirlik incelemesinde ise bu testin güvenirlik standartlarını karşılamada konusunda da yetersiz olduğu belirlenmiştir ve güvenirlik standartlarını karşılama durumunun iyileştirilmesi için yapılması gerekenler belirlenmiştir. Bunlar;

- Yorumlanan tüm puan türleri için güvenirlik bilgileri sağlanmalıdır.
- Farklı gruplar için ölçmenin standart hata değerleri belirlenmeli ve güvenirlik kestirim işlemleri yapılmalıdır.

GEÇDA’ nın test geliştirme ve revizyon standartlarını karşılama durumunun geliştirilmesi için ise yapılması gerekenler aşağıda sıralanmıştır.

- Test geliştirme sürecine dayanan yeterli kanıt toplanmalıdır.
- Uygulama koşullarına ya da testi alan bireylere göre değişiklik gösterebilecek ancak kabul edilebilir olan koşullar belirtilmelidir.
- Testin amaçları, test edilen alanın özellikleri ve test maddeleri gereken açıklıkta belirtilmelidir.
- Maddelerin geliştirilmesi, gözden geçirilmesi ve seçilmesi sürecindeki işlemler daha detaylı olarak betimlenmelidir.
- Uzman yargılarına başvurulmalı ve puanlayıcıların seçimi, eğitimi ve puanları tanımlama düzeyleri detaylı olarak rapor edilmelidir.

Çalışma kapsamında incelemeye alınan bir diğer psikolojik test, Akademik Benlik Kavramı Ölçeği’dir. İncelenen standartlar göz önünde bulundurulduğunda, bu psikolojik teste ilişkin geliştirilmesi gereken nitelikler belirlenmiş ve bu testin özellikle güvenirlik standartları bakımından güncellenmesi ve güvenirliğine ilişkin yeni kanıtların toplanması gereklidir. Söz konusu testin güvenirlik standartlarını karşılama düzeyinin iyileştirilmesi için, yapılması gerekenler şunlardır:

- Yorumlanacak puan türlerine yönelik ölçmenin standart hatası hesaplanmalıdır.
- Güvenirlik analizlerinin yapıldığı yöntem ve kullanılan istatistikler açıklanmalıdır.
- Güvenirlik analizlerinin gerçekleştirildiği örneklem grubu ayrıntılı olarak betimlenmelidir.
- Farklı uygulamalar için farklı güvenirlik analizleri yapılmalıdır.
- Norm gruplarına ilişkin güvenirlik kestirimleri gerçekleştirilmelidir.
- Kullanılan güvenirlik belirleme yöntemleri testin çok faktörlü yapısına uygun olarak seçilmelidir.

İncelemelerden elde edilen sonuçlara göre, Akademik Benlik Kavramı Ölçeği'nin geçerlik standartları bakımından geliştirilmesi ve güncelleştirilmesi gereken noktalar şunlardır:

- Bu teste dair yorumlanacak puanların dayandığı teori ve kanıtlar sunulmalıdır.
- Kapsam geçerliğinin belirlenmesi için uzman yargılarına başvurulmalıdır.
- Testin alt testlerine dair yapılan yorumlarla ilgili kanıtlar sunulmalıdır.
- Madde yanıtlarının değerlendirilmesi sırasında puanlayıcıların yapabileceği olası fakat kabul edilmeyecek yorumlar için puanlayıcılar uyarılmalıdır.
- Geliştirilen bileşik puanlardaki alt puanların mantığı sunulmalıdır.

Test geliştirme ve revizyon standartları temel alındığında ise Akademik Benlik Kavramı Ölçeği'nin geliştirilmesi gereken yönleri şunlardır:

- Test geliştirme sürecine dayanan yeterli sayıda kanıt toplanmalıdır.
- Test yönergeleri daha detaylı olarak sunulmalıdır.
- Yönergeler için uzmanlardan görüş alınmalıdır.
- Maddelerin seçilme ve gözden geçirilme süreçleri betimlenmelidir.
- Testi oluşturan maddeler için uzman görüşüne başvurulmalıdır.
- Madde seçim sürecinde hangi verilerin kullanılmış olduğu da detaylı olarak sunulmalıdır.

Araştırma çerçevesinde Türkiye'de geliştirilen testlerin yanı sıra, uyarlama çabası olarak kullanılmakta olan Temel Kabiliyetler Testi 7-11 Formu ve Wechsler Çocuklar İçin Zeka Ölçeği Revize Edilmiş Form da yer almaktadır. Bu psikolojik testler uyarlanmış psikolojik testler oldukları için hazırlanan kontrol listesinde yer alan test uyarlama standartlarına göre incelenmiş ve bu standartlara göre geliştirilmesi ve güncelleştirilmesi gereken noktalar belirlenmiştir. Yapılan incelemelere göre, TKT 7-11 Yaş Formu için uyarlama standartları bakımından gözden geçirilmesi gereken noktalar şunlardır:

- Testin orijinal formu ile uyarlanmış formu arasındaki kültürel farkların etkileri olabildiğince azaltılmalıdır.
- Çalışma evrenindeki test ve ölçekler ile ölçülen özelliğin örtüşme düzeyi belirlenmelidir.
- Test yönergeleri ve maddeleri dilsel ve kültürel farklardan arındırılmalı ve madde formatı ve test düzeni hedef evrene uygun olarak seçilmelidir.
- Uyarlanan her maddenin orijinal testteki maddelere denk olup olmadığı araştırılmalıdır.
- Testin uyarlanmış hali ile orijinal halinin eşitliğini desteklemek için yapılan değişimler raporlaştırılmalıdır.

Bu güncelleştirme çalışmaları sonucunda söz konusu psikolojik test uyarlama standartlarının bütününe tamamen karşılayacak bir konuma kavuşabilecektir.

Araştırma kapsamında incelenen son psikolojik test ise bir diğer uyarlanmış test olan Wechsler Çocuklar İçin Zeka Ölçeği Revize Edilmiş Form'dur. Bu psikolojik test için yapılan incelemeler sonucunda, uyarlama standartlarının yüksek düzeyde karşılandığı belirlenmiştir. Söz konusu test için yalnızca bir uyarlama standardı hiç karşılanmamaktadır ve testin uyarlanmış hali ile orijinal halinin eşdeğerliğini desteklemek için yapılan değişimlerin raporlaştırılmasıyla bu test için belirlenen tüm uyarlama standartları karşılanabilecektir.

SONUÇLAR ve TARTIŞMA

Bu araştırmada kapsamında Rehberlik ve Araştırma Merkezleri ile okulların rehberlik servislerinde ortak olarak sıkça kullanılmakta olan dört test belirlenerek, bu testlerin eğitimde ve psikolojide ölçme standartları ile test uyarlama standartlarını karşılama durumu değerlendirilmiştir. Araştırma kapsamına alınan testlerden Türkiye'de geliştirilmiş olanlar, kontrol listesinde yer alan geçerlik, güvenilirlik ve test geliştirme ve revizyon standartlarına göre incelenmiş, Türkçeye uyarlanmış olanlarsa uyarlama standartları gözetilerek incelenmiştir.

Çalışma kapsamında elde edilen sonuçlar incelendiğinde, eğitim kurumlarında sıklıkla kullanılan psikolojik testlerden ABKÖ ve GEÇDA'nın özellikle geçerlik ve güvenilirlik gibi kritik psikometrik nitelikleri gereken düzeyde karşılamaktan uzak olduğu belirlenmiştir. Söz konusu ölçeklerin test geliştirme ve revizyon standartları bakımından durumları incelendiğinde ise, Acar Güvendir ve Özer Özkan (2015) tarafından yapılan çalışmada da belirtildiği üzere, ölçeklerin geliştirilme süreçleri arasında birliktelik bulunmadığı görülmektedir. GEÇDA ve ABKÖ için temel test geliştirme ve revizyon standartları tamamen karşılanmamakta ve farklı standartlar farklı düzeylerde karşılanmaktadır. Söz konusu standartlara uyumlu olarak sistematik bir yaklaşım izlenmeli ve tüm araştırmacıların standartları takip etmesi ile oluşabilecek hatalar en aza indirgenerek, tekrar eden hataların önüne geçilmelidir.

Araştırma çerçevesinde incelenen diğer iki psikolojik test yabancı kültürlerde geliştirilmiş ve Türkçeye uyarlanarak kazandırılmış olan testlerdir. Dolayısıyla bu testler uyarlama standartlarına göre incelenmiş ve mevcut durumları bu standartlar çerçevesinde betimlenmiştir. Uyarlanmış testlerden ilki Temel Kabiliyetler Testi 7-11 Yaş Formu'dur. Yapılan incelemeler sonucunda, 15 uyarlama standardının sekizinin bu test için tamamen karşılanmakta olduğu, yedisinin ise hiç karşılanmadığı belirlenmiştir. Çüm ve Koç (2013) tarafından yapılan çalışmada incelenen testlerde olduğu gibi, bu psikolojik test için de uyarlama çalışmasının gereklerinden olan eşdeğerlik incelemelerinin ihmal edildiği görülmektedir. Son olarak TKT 7-11 Yaş Formu için yapılan uyarlama çalışması sırasında çok sayıda maddenin test kapsamı dışında bırakılmış olduğu ve bu durumun da daha sonra gerçekleştirilecek uyarlama çalışmaları için uygunsuz bir örnek oluşturduğu sonucuna ulaşılmıştır.

Bu araştırma kapsamında test uyarlama standartlarına uygunluğu incelenen diğer psikolojik test olan WISC-R' in uyarlama standartlarını karşılama durumunun TKT 7-11 Yaş Formu'na göre daha yüksek olduğu belirlenmiştir. Bu test için incelenen 15 uyarlama standardının 14'ünün tamamen karşılanmakta olduğu, yalnızca bir standardın karşılanmadığı belirlenmiştir. Bu durumda söz konusu testin uyarlama standartlarına uygun bir şekilde Türk kültürüne uyarlandığı sonucuna varılmıştır.

Çalışma kapsamında elde edilen bulgulara göre, psikolojik testlere ilişkin gereken izleme çalışmalarının yapılması ve testlerin niteliklerinin yeniden sorgulanmasına ve testlerin psikometrik yönlerinin gelişimine katkıda bulunacağı sonucuna ulaşılmıştır. Tavşancıl, Güler ve Ayan (2014) tarafından da belirtildiği üzere, ülke çapında kullanılmakta olan psikolojik testlerin hem denetimini yapan hem de koordinasyon sağlayan bir test geliştirme merkezi kurularak, bu merkez tarafından psikolojik testler için gerekli olan izleme çalışmaları gerçekleştirilebilir. Kurulacak olan test merkezi ile psikolojik testlerin sistematik olarak bilimsel olarak incelenmesi, bu testlerin psikometrik niteliklerinin zamana göre değişiminin belirlenmesi sağlanarak, söz konusu testler kullanılarak verilen kararların isabetli olma olasılığı artırılabilecektir. Son olarak bu çalışma kapsamında geliştirilen kontrol listesi Acar Güvendir ve Özer Özkan (2015) tarafından belirtilen ölçek geliştirme ve uyarlama için bir el kitabı mahiyetinde kullanılabilir.

KAYNAKÇA

- Acar Güvendir, M. ve Özer Özkan Y. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerinin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 23-23.
- American Psychological Assosiation, American Educational Reserach Association & National Council on Measurement in Education. (1999). *Standarts for educational and psychological testing*. Washington: American National Research Education.
- Anastasi, A.(1988). *Psychological testing*.(6th Edition). New York: Macmillan).
- Boztunç Öztürk, N., Eroğlu, M. G. ve Kelecioğlu, H. (2014, Haziran). Eğitim bilimleri alanında yapılan ölçek uyarlama makalelerinin incelenmesi. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde sunulan bildiri, Hacettepe Üniversitesi, Ankara.
- Cansever, G. (1982). *Klinik psikolojide değerlendirme yöntemleri*. İstanbul: Boğaziçi Üniversitesi Yayınları Gözlem Matbaacılık Koll. Sti.
- Cronbach, L. J. (1990). *Essentials of psychological testing* (5th Edition). New York: Harper Collins.

- Çıkrıkçı-Demirtaşlı, R. N. (2017). Psikolojik ölçmelere ilişkin doğru bilinen yanlışlar. *Türk Psikoloji Bülteni*, 13(41), 65-68.
- Çukur, C. S. (1999). Kültürün psikoloji bilimindeki yeri üzerine görüşler. *Türk Psikoloji Yazıları*, 1(3), 1-16.
- Çüm, S. ve Koç, N. (2013). Türkiye’de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Eğitim Bilimleri ve Uygulama*, 12(24), 115-135.
- DeVellis, R. F. (2012). *Scale development theory and application* (3rd Editon). SAGE Publications, Inc.
- Erkuş, A. (2007). Ölçek geliştirme ve uyarlama çalışmalarında karşılaşılan sorunlar. *Türk Psikoloji Bülteni*, 13(40), 17-25.
- Gülgez, S. (1994). Test kullanımında temel konular. *Türk Psikoloji Dergisi*, 9(33), 1-8.
- Karasar, N. (2010). *Bilimsel araştırma yöntemi-kavramlar, ilkeler, teknikler*.(21. Baskı) Ankara: Nobel. Yayınevi.
- Koç, N. (1993). Eğitim sistemimizde ölçme ve değerlendirme alanındaki gelişmeler. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 25(2), 387-407.
- Kuzgun, Y. (2011). *Akademik Benlik Kavramı Ölçeği el kitabı*. Ankara: Nobel.
- Milli Eğitim Bakanlığı, Ölçme ve Değerlendirme Sistemi Komisyonu. (1990). *Ölçme ve değerlendirme sistemi özel ihtisas komisyonu raporu*. Ankara: Milli Eğitim Bakanlığı
- Milli Eğitim Bakanlığı. (2001a). *Temel kabiliyetler testi 7-11 uyarlama çalışması el kitabı*. Ankara: Milli Eğitim Basımevi.
- Milli Eğitim Bakanlığı. (2001b). *Temel kabiliyetler testi 7-11 yönergesi*. Ankara: Milli Eğitim Basımevi.
- Mor Dirlik, E. (2014). Ölçek geliştirme konulu doktora tezlerinin test ve ölçek geliştirme standartlarına uygunluğunun incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(2), 62-78.
- Öner, N. (1994). Güvenirliği ve/veya geçerliği sınanmış psikolojik testler. *Türk Psikoloji Dergisi*, 9(33), 9-18.
- Özge, S. (1981). *Psychological tests used in Turkey: A preliminary survey* (Yayınlanmamış yüksek lisans tezi). Bogaziçi Üniversitesi, İstanbul.
- Özgüven, İ. E. (1999). *Psikolojik testler*. Ankara: PDREM.
- Ramazan, O. (1988). *Survey for a source book on psychological tests used in Turkey* (Yayınlanmamış yüksek lisans tezi). Bogaziçi Üniversitesi, İstanbul.
- Savaşır, I. (1981). *Psikolojik testler* (Rapor No. 7). Ruh Sağlığı ve Hastalıkları, Türkiye Sınır ve Ruh Sağlığı Derneği Yayınları, Ankara.
- Savaşır, I. (1994). Ölçek Uyarlamasındaki Sorunlar ve Bazı Çözüm Yolları. *Türk Psikoloji Dergisi*, 9(33), 27-32.
- Savaşır, I., Sezgin, N. ve Erol, N. (1992). 0-6 yaş çocukları için gelişim tarama envanteri geliştirilmesi: Ön çalışmalar. *Türk Psikiyatri Dergisi*, 3(2), 33-42.
- Savaşır, I. ve Şahin, N. (1995). *Wechsler Çocuklar İçin Zeka Ölçeği*. Ankara: Türk Psikologlar Derneği.
- Tavşancıl, E. ve Arslan, E. (2001). *İçerik analizi ve uygulama örnekleri*. İstanbul: Epsilon.
- Tavşancıl, E., Güler G. ve Ayan C. (2014, Haziran). *2002-2012 yılları arasında Türkiye’de geliştirilen bazı tutum ölçeği geliştirme çalışmalarının ölçek geliştirme süreci açısından değerlendirilmesi*. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresinde sunulan bildiri, Hacettepe Üniversitesi, Ankara.
- Temel, F., Ersoy, Ö., Avcı, N. ve Turla, A. (2005). *Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı*. Ankara: Gazi Üniversitesi.
- Yıldırım, A., Şimşek, H. (2011). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri*. (8. Baskı). Ankara: Seçkin Yayıncılık.

EXTENDED ABSTRACT

Introduction

Psychological tests are the source of knowledge used in many fields such as education, personnel selection and journalism. Besides the wide range of usage, by using the results of the tests, there are many crucial decisions made for people. In order to make right decisions about people by using these tests, it is essential to analyze and investigate the quality of psychometric features of the tests. For this reason, the goal of this study is to analyze the psychometric qualities of the psychological tests which are frequently used in the Guidance and Research Centers and the guidance services of the public schools. Analyses of the tests used in this area is much more important because these tests are used to make decisions about students’ educational life and according to the results of the tests,

students prefer their occupations, or they are diagnosed about their pace of learning and something like that. Thanks to the analyzing of the tests by using the standards for educational and psychological testing of APA (American Psychological Association) and test adaption standards of ITC (International Testing Commission), it is possible to get more valid inferences about the people and the psychological traits.

Method

The four psychological tests were analyzed in this research and the tests were determined based on the goal-oriented sample selecting method. In order to decide which tests should be analyzed, a questionnaire was prepared which was composed of the list of tests that are in common usage in these institutions. The questionnaire was administered 79 counselors and their opinions were asked to state which tests require revisions. According to results of questionnaire, four tests were determined as requiring revisions. These tests are: Akademik Benlik Kavramı Ölçeği, Gazi Erken Çocukluk Gelişimi Değerlendirme Aracı, Primary Mental Abilities 7-11, and Wechsler Intelligence Scale for Children Revised Form. After determining the tests that were analyzed, a checklist was prepared by researcher to investigate the psychometric properties of the tests. In order to prepare the checklist, the chapters related to the validity, reliability and test development and revision of “Standards For Educational and Psychological Testing” that was published by APA in 1999 were translated into Turkish. Actually, the process was not just a translation, it was considered to make more clear and easy to understand, so the standards were also simplified. In addition to the APA standards, the adaptation standards developed by ITC were translated into Turkish and the checklist was composed by using these standards. The checklist has got two forms as short and long form. While the long form of the checklist is composed of all of the standards included in the relevant chapters, the short form includes only the basic standards which should be investigated for all kinds of psychological tests. The tests that were chosen in this study were analyzed according to the short form of this checklist by researcher, because in the long form, there are so many standards which can be analyzed only the special kind of psychological test. In order to question the reliability of these analyses, the analyses were repeated in three weeks’ time by the researcher and the fit index between the analyses was computed. Data of these analyses were exported to the Statistical Package for Social Sciences (SPSS) 20.0 and descriptive analysis was performed.

Results and Discussion

As a result of this study, the standards that have been met and have not met satisfactorily by the tests were determined. By this way, the psychometric qualities of the tests, such as validity and reliability were defined. As for the GEÇDA, most of the validity, reliability and test revision standards were met partially. So it has been found that, for this test, it is a necessity to gather some new proves for validity, reliability and test development process. As for ABKÖ, none of the reliability standards were met, so this test should be revised especially for reliability. The adapted tests, WISC-R and TKT 7-11, were analyzed according the adaptation standards and it was found that while all of the standards except one were met for WISC-R, only the half of the adaptation standards were met for TKT 7-11. It is clear that adapted test have met more standards than the others, but especially for TKT 7-11, some revisions should be done in order to get more valid and reliable results. To conclude, all of the psychological test should be revised and their psychometric properties should be investigated to make valid and reliable decision for people’s lives.