Türk Psikoloji Dergisi, Haziran 2017, *32*(79), 80–91

Üstün Yetenekli Çocuklarda WÇZÖ-R ve WÇZÖ-IV* Zekâ Puanlarının Karşılaştırılması

İbrahim Yiğit
Ankara Üniversitesi

Cihat Çelik
Ankara Üniversitesi

Gülsen Erden
Ankara Üniversitesi

Özet

Wechsler Çocuklar için Zekâ Ölçeği Geliştirilmiş Formu (WÇZÖ-R), üstün yetenekli çocukların belirlenmesinde ya da değerlendirilmesinde yaygın olarak kullanılmaktadır. Ülkemizde, WÇZÖ-R'den sonra Wechsler zekâ ölçeklerinin dördüncü sürümü olan WÇZÖ-IV yakın zamanda kullanılmaya başlanmıştır. Bu açıdan, hem üstün yetenekli çocukları belirlemede WÇZÖ-IV'ün etkililiğini değerlendirmek hem de kullanımda olan WÇZÖ-R ile WÇZÖ-IV arasında ortaya çıkabilecek puan farklarını belirlemek önem kazanmaktadır. Bu çalışmanın amacı, üstün yetenekli çocukların WÇZÖ-R ve WÇZÖ-IV üzerinde elde ettikleri puanları karşılaştırmaktır. Daha önce WÇZÖ-R ile üstün yetenekli olduğu belirlenen 50 çocuğa WÇZÖ-IV uygulanmıştır. Çalışmanın bulguları, WÇZÖ-R Toplam Zekâ Bölümü (ZB) ile WÇZÖ-IV Tüm Ölçek Zekâ Puanı (TÖZP) arasında ve WÇZÖ-R Performans ZB ile WÇZÖ-IV Algısal Akıl Yürütme Puanı (AAYDP) puanları arasında istatistiksel olarak anlamlı fark bulunduğunu göstermektedir. Buna göre, üstün yetenekli çocuklar söz konusu puanlarda, WÇZÖ-R üzerinde daha yüksek puanlar elde etmiştir. Ayrıca, karşılaştırılabilir alt test standart puanları açısından, çocukların WÇZÖ-R üzerinde, Sözcük Dağarcığı alt testi dışında, Benzerlikler, Küplerle Desen ve Şifre alt testlerinden daha yüksek puanlar aldıkları gösterilmiştir. Korelasyon analizleri, her iki sürümden elde edilen puanlar arasında istatistiksel olarak anlamlı ilişkilerin olduğunu ve bu iki sürümün benzer yapıları ölçtüğünü ortaya koymuştur. Değerlendiriciler, WÇZÖ-IV'ü üstün yetenekli çocukları belirlemede kullanırken, eski sürüme göre, bazı alt test standart puanlar ya da dönüştürülmüş puanlar arasında farklılıkların olabileceğini göz önünde bulundurmalıdır.

Anahtar kelimeler: Üstün yetenekli çocuklar, WÇZÖ-R, WÇZÖ-IV, Flynn etkisi

Abstract

Wechsler Intelligence Scales for Children-Revised (WISC-R) has been widely used in identifying or assessing gifted children. In our country, WISC-IV, which is the fourth version of Wechsler intelligence scales, will be in use soon. In this respect, it is important to assess both efficiency of WISC-IV in identifying gifted children and to determine the possible score differences between WISC-R and WISC-IV. The purpose of the present study is to compare WISC-R and WISC-IV scores of gifted children. WISC-IV was administered to 50 children who were identified as gifted by WISC-R. The findings of the study showed that there were statistically significant differences between WISC-R Full Scale Intelligence Quotient (FSIQ)/WISC-IV FSIQ and WISC-R Performance IQ/WISC-IV Perceptual Reasoning Index (PRI). Accordingly, gifted children had higher scores on WISC-R than WISC-IV. In terms of comparable subtest scores, they had also higher scores on WISC-R Similarities, Block Design and Coding subtests compared to WISC-IV's same subtests, except for the Vocabulary subtest. The correlation analysis revealed that there were significant relationships between WISC-R and WISC-IV composite and index scores and that both of these versions tested the same constructions. Administering WISC-IV for gifted identification, practitioners should consider some subtest or composite score differences compared to its previous version.

Key words: Gifted children, WISC-R, WISC-IV, Flynn effect

Yazışma Adresi: Arş. Gör. İbrahim Yiğit, Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi, Psikoloji Bölümü, 06100 Sıhhiye/Ankara
E-posta: psk.yigit@gmail.com

* "WÇZÖ-IV"ün standardizasyon ve norm çalışması, 109K533 ve 107K493 no.lu TUBİTAK-SOBAG projeleri kapmasında gerçekleştirilmiştir.

Üstün yetenekli çocuklar, akademik, zihinsel yetenek, yaratıcı ve üretici düşünme, sanat (örn., müzik, tiyatro), estetik, spor ve liderlik alanlarında yaşıtlarına göre daha yüksek performans sergileyen çocuklar olarak tanımlanmaktadır (Marland, 1972; Pfeiffer, 2002; Sattler, 2002). Bu çocukların, yetenekli oldukları alanda hızlı gelişme kaydettikleri ve eğitim açısından normal programların ötesinde, farklılaştırılmış ve onlara göre düzenlenmiş programlara ihtiyaç duydukları bilinmektedir (Sattler, 2002; Smutny ve Blocksom, 1990). Bu açıdan, üstün yetenekli çocukların yeteneklerine uygun okullarda eğitim görmesi ya da eğitilmesi, öncelikli olarak bu çocukların ya da ergenlerin doğru bir şekilde değerlendirilmesini ve belirlenmesini gerektirmektedir.

Üstün yetenekli çocukları değerlendirme ve belirleme amacıyla çeşitli değerlendirme araçları kullanılmaktadır. Kullanılan değerlendirme araçları, çocuğun yerleştirileceği programın içeriğine göre farklılık göstermekle birlikte, genellikle standart testler (örn., zekâ, başarı, yaratıcılık, sosyal olgunluk, kişilik testleri), anne-baba ve öğretmen raporları/bildirimleri, doğrudan çocuğun davranışını gözlemleme ve çocuğun yaratıcılığını değerlendirmeyi içermektedir (Alvino, McDonnel ve Richert, 1981; Sattler, 2002). Özellikle bireysel olarak uygulanan zekâ testlerinin en geçerli ve güvenilir araçlar olduğu belirtilmekte ve bu testler, üstün yetenekli çocukların bilişsel yeteneklerinin belirlenmesi ve gerekli eğitim programlarına yerleştirilmesi amacıyla sıklıkla kullanılmaktadır (Newman, Sparrow ve Pfeiffer, 2008; Pfeiffer, 2002; Sattler, 2002).

Alanda birçok zekâ testi olmasına rağmen, üstün yetenekli çocukları değerlendirme sürecinde Wechsler zekâ ölçeklerinin yaygın olarak kullanıldığı bildirilmektedir (Alvino ve ark., 1981; Roid, 1990; Stinnet, Harvey ve Oehler-Stinnet, 1994; Wilson ve Reschly, 1996). İlk olarak, 1939 yılında Wechsler-Bellevue Zekâ Ölçeği olarak geliştirilen ölçek, bu tarihten itibaren birçok revizyondan geçerek, sırasıyla Wechsler Çocuklar İçin Zekâ Ölçeği (WÇZÖ) (Wechsler, 1949), WÇZÖ-R (Wechsler, 1974), WÇZÖ-III (Wechsler, 1991), WÇZÖ-IV (Wechsler, 2003) ve son olarak WÇZÖ-V (Wechsler, 2014) adıyla günümüze kadar gelmiştir. Ölçek, 6-16 yaş arasındaki çocukların zekâ düzeylerini belirleme amacıyla kullanılmakta ve çeşitli alt test standart puanları ve dönüştürülmüş puanlar elde edilmesini sağlamaktadır (Wechsler, 2014). Üstün yetenekli çocukların belirlenmesinde temel ölçüt zekâ testi puanları olduğunda, genellikle Toplam ya da Genel Zeka Bölümü (ZB) puanı kullanılmaktadır. Buna göre, bu puan türünden 130 ve üstü ya da ortalamanın iki standart sapma üstü puan alan bir çocuk, üstün yetenekli olarak nitelendirilmektedir (Mark, Beal ve Dumont, 1998; Pfeiffer, 2012; Sattler, 2002). Wechsler zekâ ölçeklerinin, üstün yetenekli çocuklar için iyi bir değerlendirme bilgisi sunduğu ve geçerli sonuçlar ortaya koyduğu birçok çalışma tarafından da bildirilmiştir (Kaufman, 1992; Klausmeier, Mishra ve Maker, 1987; Silverman, Gilman ve Falk, 2004).

Belirli bir standardizasyon örneklemine ve normlarına sahip olan zekâ testlerinin ortaya çıkmasından bu yana dünya genelinde zekâ testlerinden elde edilen puanlarda sistematik ve yaygın bir yükselmenin olduğu gözlenmiştir (Kanaya, Scullin ve Ceci, 2003). Zekâ puanlarındaki bu yükselmenin, okullaşma, çevresel uyaranlardaki zenginleşme, testle olan tanışıklığın artması, beslenme ve sağlık alanlarındaki gelişmeler gibi birçok faktörden kaynaklandığı düşünülmektedir (Lynn, 2009). Flynn etkisi olarak bilinen bu olguya göre, zaman içinde genel zekâ (IQ) puanlarında meydana gelen bu yükselmeler, zekâ testi normlarının geçersiz hale gelmesine neden olmaktadır. Bir başka ifadeyle, zaman ilerledikçe ve zekâ testi normları eskidikçe, testin kullanımda kaldığı yıllar içinde değerlendirmeye alınan çocuklar, test üzerinde gittikçe daha iyi performans sergilemektedir. Böylece, zeka puanlarında 10 yılda bir ortalama üç puanlık bir artış söz konusu olmaktadır (Flynn, 1984; 1987; 1999; Steen, 2009). Kaufman (1990), 1930'dan bu yana ortalama zeka puanlarında sürekli bir artış olduğunu ifade etmiştir. Benzer şekilde, Wechsler (1991) de, ortalama zeka puanlarının aşamalı olarak arttığını ve zekâ testi normlarının düzenli aralıklarla güncelleştirilmesi gerektiğini bildirmiştir.

Flynn etkisi göz önünde bulundurularak, zekâ testlerinin standardizasyon ve norm çalışmaları belirli aralıklarla yenilenmektedir. Bu yenilenmelerle ya da revizyonlarla birlikte, eski ve yeni sürüm arasında ortaya çıkan puan farkları gündeme gelmektedir. Zekânın zaman içinde sabit olduğu düşünülen psikolojik bir yapı olduğu ve bu nedenle bir zekâ testi yenilendiğinde yeni sürümün bir önceki sürümle benzer sonuçlar vermesinin önemli olduğu ileri sürülmektedir (Goh, Teslow ve Fuller, 1981; Hutton, Dubes ve Muir, 1992; Watkins, Campbell, Nieberding ve Hallmark, 1995). Bu açıdan, her iki sürümün eşitliğini ya da arasındaki ilişkiyi ve ortalama puanlar arasındaki sistematik farklılıkları belirlemek önemlidir (Vance, Maddux, Fuller ve Awadh, 1996). Bu durum, özellikle özel eğitim kurumlarına yerleştirme ya da uygunluğu tekrar değerlendirme amacıyla, düzenli aralıklarla zekâ testlerini uygulamak isteyen değerlendiriciler ve bu amaçlarla değerlendirilen çocuklar açısından daha da önem kazanmaktadır (Moffitt, Caspi, Harkness ve Silva, 1993). Yapılan çalışmalara bakıldığında, Wechsler zekâ ölçeklerinin eski ve yeni sürümünden elde edilen puanlar arasında anlamlı ortalama farklar gözlendiği bulunmuştur. Klinik ve normal örneklemlerle gerçekleştirilen bu çalışmalar, ortalama 3-6 arası bir puan farkı ile yenilenmiş testin daha düşük puanlar

ortaya koyduğunu göstermiştir: WÇZÖ/WÇZÖ-R (Doppelt ve Kaufman, 1977; Munford, 1978; Munford, Meyerowitz ve Munford, 1980; Munford ve Munoz, 1980; Reschly ve Davis, 1977; Solway, Fruge, Hays, Cody ve Gryll, 1976; Spitz, 1983); WÇZÖ-R/WÇZÖ-III (Bolen, Aichinger, Hall ve Webster, 1995; Doll ve Boren, 1993; Hishinuma ve Yamakawa, 1993; Lyon, 1995; Mahone ve ark., 2003, Newby, Recht, Caldwell ve Schaefer, 1993; Slate ve Saarnio, 1995; Vance ve ark., 1996, Wechsler, 1991). Alanyazında, WÇZÖ-III ile WÇZÖ-IV'ü karşılaştıran herhangi bir çalışmaya ulaşılamamıştır. Ülkemizde ise son yıllarda yapılan iki çalışmaya rastlanmaktadır. Bunlardan ilki, Dikkat Eksikliği ve Hiperaktivite Bozukluğu (DEHB) tanısı konulan çocukların WÇZÖ-R ile WÇZÖ-IV puanlarının karşılaştırıldığı çalışmadır (Çelik, Erden, Özmen ve Hesapçıoğlu, 2017). Bu çalışmada, DEHB'li çocukların WÇZÖ-R üzerinde, karşılaştırılabilir tüm puanlar açısından, WÇZÖ-IV'e göre daha yüksek puan aldıkları bulunmuştur. Bir diğer çalışma ise, zihinsel gelişimi normal olan çocukların WÇZÖ-R ZB puanlarının, WÇZÖ-IV'te karşılık gelen dönüştürülmüş puanlara göre, yüksek olduğunu göstermektedir (Uluç, Korkmaz ve Şahin, 2014). Bulunan puan farklılıklarına ek olarak, çalışmalarda testlerin eski ve yeni sürümlerinden elde edilen ZB puanları arasında anlamlı korelasyonlar olduğu ortaya koyulmuştur (Covin, 1977; Çelik ve ark., 2017; Munford, 1978; Sevier, Bain ve Hildman, 1994; Uluç ve ark., 2014, Vance ve ark., 1996; Wechsler, 1991; Wheaton ve Vandergriff, 1978).

Üstün yetenekli çocuklarda, iki test sürümünü karşılaştıran çalışmalar da, benzer puan farklılıklarının ortaya çıktığını göstermektedir. Üstün yetenekli çocukların WÇZÖ ve WÇZÖ-R ZB puanlarını karşılaştıran çalışmalar, bu çocukların WÇZÖ'den daha yüksek puanlar aldıklarını göstermektedir (Larrabee ve Holroyd, 1976; Solly, 1977; Wheaton, Vandergriff ve Nelson, 1980). Bu çalışmaların aksine, Wheaton ve Wandergriff (1978) daha önce WÇZÖ-R uygulanmış yaşları 10-12 arasında değişen 26 üstün yetenekli çocuğa, WÇZÖ uygulayarak iki sürümün puanlarını karşılaştırmış ve bu çocukların WÇZÖ-R Sözel ve Toplam ZB'den daha yüksek puanlar aldıklarını bulmuştur. WÇZÖ-III'ün standardizasyon çalışmasında (Wechsler, 1991), 23 üstün yetenekli çocuktan oluşan küçük bir gruba WÇZÖ-III uygulanarak, bu çocukların daha önce elde ettikleri WÇZÖ-R puanları ile karşılaştırma yapılmıştır. Bu çalışmanın sonuçları, Sözel, Performans ve Toplam ZB puanlarının, WÇZÖ-R'ye göre WÇZÖ-III'te daha düşük olduğunu göstermiştir. Buna göre, Sözel ZB ortalama puanında 5.8'lik (128.0'a 133.8), Performans ZB ortalama puanında 1.1'lik (124.6'ya 125.7), Toplam ZB ortalama puanında ise, 4.9'luk (128.7'ye 133.6) bir fark bulunmuştur. Diğer bir çalışmanın sonuçlarına göre ise, WÇZÖ-R ile WÇZÖ-III Toplam ZB puanları arasında 13.5'lik bir fark olduğu ortaya konmuştur (Bryant, 1992). Sevier ve arkadaşları (1994), 35 üstün yetenekli ilkokul öğrencisi ile gerçekleştirdiği çalışmalarında, WÇZÖ-III'ün WÇZÖ-R'ye göre Sözel ZB (14.57'lik fark), Performans ZB (7.60'lık fark) ve Toplam ZB (12.83'lük) puanlarında daha düşük puanlar ortaya koyduğunu bulmuştur. Ayrıca, alt test standart puanları açısından bakıldığında, resim düzenleme ve şifre alt testleri dışında, diğer tüm alt testlerin ortalama standart puanlarının WÇZÖ-R'de anlamlı olarak daha yüksek olduğu gözlenmiştir. Aynı çalışmada, değerlendirmeye alınan çocukların üstün yetenekliler programına uygunluğunun WÇZÖ-III kullanılarak değerlendirilmesi halinde, bu çocuklardan 14'ünün programa yerleştirilemeyeceği gözlenmiştir. Son olarak, yukarıdaki çalışmalardan farklı olarak, Sabatino, Spangler ve Vance (1995), 51 üstün yetenekli çocukla yürüttükleri çalışmada, WÇZÖ-R ile WÇZÖ-III puanlarının benzer sonuçlar ortaya koyduğunu bulmuştur. Buna göre, Sözel ve Performans ZB puanları arasında iki puanlık, Toplam ZB puanları arasında ise bir puanlık bir fark olduğu gözlenmiştir.

WÇZÖ-IV'ün üstün yetenekli çocukların değerlendirilmesinde geçerli ve güvenilir sonuçlar ortaya koyduğu bildirilmektedir (Silverman, 2005; Silverman ve ark., 2004; Yiğit, 2016). Üstün yetenekli çocukların, özellikle Sözel Kavrama Dönüştürülmüş Puanı (SKDP) ve Algısal Akıl Yürütme Dönüştürülmüş Puanı (AAYDP) üzerinde, Çalışma Belleği Dönüştürülmüş Puanı (ÇBDP) ve İşlemleme Hızı Dönüştürülmüş Puanına (İHDP) göre daha yüksek puanlar aldıkları ortaya koyulmuştur (Falk, Silverman ve Moran, 2004; Wechsler, 2003). Bir başka ifadeyle, üstün yetenekli çocuklar sözel soyut ve görsel akıl yürütmeyi değerlendiren alt testleri barındıran SKDP ve AAYDP'de iyi bir performans gösterirken, işlemleme becerilerini ölçen İHDP ve kısa süreli bellek işlevlerini ölçen ÇBDP'de daha düşük performans göstermektedirler (Flanagan ve Kaufman, 2004; Weiss, Saklofske, Prifitera ve Holdnack, 2006).

Wechsler zekâ ölçeklerinin ülkemizde kullanılmaya başlanması, WÇZÖ-R ile gerçekleşmiştir. Savaşır ve Şahin (1995) tarafından Türkçeye kazandırılan WÇZÖ-R, uzun yıllardır hastanelerde, kliniklerde, Rehberlik Araştırma Merkezleri'nde ve araştırma alanlarında yaygın olarak kullanılmaktadır. Özellikle, özel eğitim ihtiyacı olan çocukların değerlendirilmesinde temel değerlendirme aracı olarak rol oynamaktadır (Çelik, Yiğit ve Erden, 2015; Uluç, Öktem, Erden, Gençöz ve Sezgin, 2011). Son yıllarda, ülkemizde kurulan Bilim ve Sanat Merkezlerine (BİLSEM) üstün yetenekli çocukların alımında ya da kabulünde de WÇZÖ-R puanları kullanılmaktadır. WÇZÖ-R'nin yenilenen bir sonraki sürümü olan WÇZÖ-III'ün ise, Türkiye'de standardizasyon ve

norm çalışması yapılmamıştır. Yakın zamanda, Wechsler ölçeklerinin dördüncü sürümü olan WÇZÖ-IV'ün (Wechsler, 2003) standardizasyon ve norm çalışması tamamlanmış ve bu sürümün kullanılmasına geçilmiştir (Öktem, Gençöz, Erden, Sezgin ve Uluç, 2013). WÇZÖ-R ile karşılaştırıldığında, WÇZÖ-IV ile birlikte ölçeğin yapısı, içeriği ve yorumu açısından önemli yenilikler ve değişiklikler gerçekleştirilmiştir (Flanagan ve Kaufman, 2009; Wechsler, 2003). Buna göre, WÇZÖ-R'da var olan ve WÇZÖ-III'te de devam eden ikili yapı (Sözel ve Performans ZB) tamamen kaldırılmış ve bunun yerini, Sözel Kavrama, Algısal Akıl Yürütme, İşlemleme Hızı ve Çalışma Belleği şeklindeki dört yeni dönüştürülmüş (indeks) puan almıştır. Resim Düzenleme, Parça Birleştirme ve Labirentler alt testleri tamamen bataryadan çıkarılmış; Resim Kavramları, Harf-Rakam Dizileri, Mantık Yürütme Kareleri, Simge Arama, Çiz Çıkar ve Sözcük Bulma olmak üzere altı yeni alt test eklenmiştir. Bir başka değişiklik ise, daha önceki sürümlerde temel alt test olarak kullanılan Genel Bilgi ve Aritmetik alt testlerinin, yedek alt testler bölümüne taşınmasıdır. Ayrıca, güçlük düzeyi düşük ve yüksek maddeler eklenerek alt test madde sayıları artırılmıştır. Son olarak, test uygulama yönergelerine ve test materyallerine ilişkin çeşitli değişiklikler yapılmış ve uygulama kitapçığı daha kullanışlı hale getirilmiştir.

Ülkemizde, üstün yetenekli çocukların değerlendirilmesinde ve ilgili eğitim kurumlarına (örn., BİLSEM) yerleştirilmesinde, ölçüt olarak zeka testi puanının kullanıldığı bilinmektedir. Bu durum göz önünde bulundurulduğunda, bu çocukların zihinsel/bilişsel yeteneklerinin doğru bir şekilde değerlendirilmesi önem kazanmaktadır. Yeni geliştirilen test sürümü nedeniyle, zekâ puanlarındaki düşüş, söz konusu çocukların üstün yetenekli olma özelliğini kaybetmelerine neden olabileceği ve yukarıda ifade edilen kurumlara yerleştirme oranlarını etkileyebileceği gibi, aynı zamanda özelleştirilmiş eğitim olanaklarından mahrum kalmalarına da neden olacaktır. Türkiye için WÇZÖ-IV'e geçiş aşamasının, aradaki test sürümü (WÇZÖ-III) atlanılarak gerçekleştiği düşünüldüğünde (Uluç ve ark., 2011), üstün yetenekli çocukları belirlemede hem WÇZÖ-IV'ün etkililiğini değerlendirmek hem de kullanımda olan WÇZÖ-R ile arasındaki puan farklarını belirlemek önem kazanmaktadır. Buradan hareketle, alanyazında Wechsler zekâ ölçekleri sürümleri arasındaki puan farkları olduğunu gösteren çalışmalar ışığında, üstün yetenekli çocukların oluşturduğu bir örneklemde, WÇZÖ-R ile WÇZÖ-IV'ten alınan puanların karşılaştırılması amaçlanmıştır. Bir başka ifadeyle çalışmanın amacı, Wechsler zekâ ölçeklerinin sürümleri arasında beklenen puan farklarının bu çalışmanın örnekleminde ortaya çıkıp çıkmayacağını incelemektir.

Yöntem

Örneklem

Araştırmanın örneklemini, ülkemizdeki üstün zekâ seviyesine sahip veya özel yetenekli çocuklara, ihtiyaç duydukları özel eğitimi veren özel bir kurumda ve Ankara ilindeki Bilim ve Sanat Merkezleri'nde (BİLSEM) eğitim gören toplam 50 üstün yetenekli öğrenci oluşturmuştur. Çalışmaya katılan 50 üstün yetenekli çocuğun %46'sı (n = 23) erkek, %54'ü (n = 27) kızdır. Çocukların WÇZÖ-R için yaş ortalaması, 153.98 ay (S = 28.30, ranj = 73-184); WÇZÖ-IV için yaş ortalaması ise 170.06 ay (S = 32.34, ranj = 79-196) olarak hesaplanmıştır.

Araştırmaya katılan çocukların annelerinin eğitim düzeylerine bakıldığında; dördünün (%8) ilkokul, sekizinin (%16) lise, 22'sinin (%44) yüksekokul, yedisinin (%14) üniversite ve ikisinin (%4) lisansüstü mezunu olduğu gözlenmiş ve yedi kişinin ise (%14) bilgisine ulaşılamamıştır. Babaların eğitim düzeylerine bakıldığında ise, birinin (%2) ilkokul, ikisinin (%4) ortaokul, ikisinin (%4) lise, 24'ünün (%48) yüksekokul, 12'sinin (%24) üniversite ve birinin (%2) lisansüstü mezunu olduğu gözlenmiş ve sekiz kişinin (%16) de bilgisine ulaşılamamıştır.

Veri Toplama Araçları

Wechsler Çocuklar İçin Zekâ Ölçeği-Geliştirilmiş Formu (WÇZÖ-R; Wechsler 1974). Wechsler tarafından 1949 yılında geliştirilen WÇZÖ'nün 1974 yılında yeniden düzenlenmiş hali olan WÇZÖ-R'nin uygulanabildiği yaş grubu 6-16 arasındadır. WÇZÖ-R'nin ülkemizdeki standardizasyon çalışması, Savaşır ve Şahin (1995) tarafından 11 kent merkezinden seçilen toplam 1639 çocuk üzerinde gerçekleştirilmiş ve ölçek, Türk kültürüne uyarlanmıştır. WÇZÖ-R'nin Türk kültürü üzerinde yapılan güvenirlik çalışmasında testin iki yarım güvenirliği, Sözel ZB için .97, Performans ZB için .93 olarak bulunmuştur. WÇZÖ-R toplam 10 adet temel alt test ve iki yedek test olmak üzere altı sözel (Genel Bilgi, Benzerlikler, Aritmetik, Yargılama, Sözcük Dağarcığı, Sayı Dizisi) ve altı performans (Resim Tamamlama, Resim Düzenleme, Küplerle Desen, Parça Birleştirme, Şifre, Labirentler) alt testinden oluşmaktadır. Ancak WÇZÖ-R'ın uygulama formunda değerlendirmenin sayı dizisi ve labirentler alt testi dışarıda bırakılarak 10 alt test üzerinden yapılması; bu iki alt testin diğer alt testlerin uygun olmadığı durumlarda verilmesi önerilmektedir (Savaşır ve Şahin, 1995). Bu alt testlere yönelik standart puanların yanı sıra, Sözel ZB, Performans ZB ve Toplam ZB katsayıları da elde edilmektedir. Bu ZB katsayıları için ortalama değer 100, standart sapma ise 15'tir. Her bir alt test için elde edilen standart puanların ise ortalaması 10, standart sapması üçtür.

Tablo 1. WÇZÖ-R ZB ve WÇZÖ-IV Dönüştürülmüş Puanlara İlişkin Betimleyici İstatistikler

WÇZÖ-R	*Ort.*	Medyan	Mod	*S*	Varyans	Ranj	Min.	Maks.
Sözel ZB	123.58	124	123	7.48	56	37	108	145
Performans ZB	130.28	130	130	8.88	78.81	36	110	146
Toplam ZB	129.38	129	133	5.94	35.30	27	120	147
WÇZÖ-IV	*Ort.*	Medyan	Mod	*S*	Varyans	Ranj	Min.	Maks.
SKDP	123.28	124	120	10.82	117.18	48	100	148
AAYDP	115	115	102	13.50	182.22	62	81	143
ÇBDP	114.56	112	109	14.82	219.80	62	85	147
İHDP	115.02	115	106	14.28	204.10	62	79	141
TÖZP	123.16	124	125	11.87	140.83	52	94	146

ZB: Zeka Bölümü; SKDP: Sözel Kavrama Dönüştürülmüş Puanı; AAYDP: Algısal Akıl Yürütme Dönüştürülmüş Puanı; ÇBDP: Çalışma Belleği Dönüştürülmüş Puanı; İHDP: İşlemleme Hızı Dönüştürülmüş Puanı; TÖZP: Tüm Ölçek Zeka Puanı.

Wechsler Çocuklar İçin Zekâ Ölçeği-IV (WÇZÖ-IV; Wechsler 2003). Wechsler Çocuklar İçin Zekâ Ölçeği-IV (WÇZÖ-IV) Wechsler ölçeklerinin son sürümü olup, 6-16 yaş aralığındaki çocukların bilişsel yeteneklerini ölçmek amacıyla geliştirilmiştir. WÇZÖ-IV, psikometrik ölçümlerle geçerliliği kanıtlanmış dört faktörü olup; toplam 10 temel ve beş yedek alt testten oluşmaktadır. WÇZÖ-IV'ten toplam beş dönüştürülmüş puan elde edilebilmektedir. Bunlar; Sözel Kavrama Dönüştürülmüş Puanı (SKDP): Benzerlikler, Sözcük Dağarcığı ve Kavrama (Yedek alt testler: Genel Bilgi, Sözcük Bulma); Algısal Akıl Yürütme Dönüştürülmüş Puanı (AAYDP): Küplerle Desen, Resim Kavramları ve Mantık Yürütme Kareleri (Yedek alt test: Resim Tamamlama); Çalışma Belleği Dönüştürülmüş Puanı (ÇBDP): Sayı Dizisi ve Harf-rakam Dizisi (Yedek alt test: Aritmetik); İşlemleme Hızı Dönüştürülmüş Puanı (İHDP): Şifre ve Simge Arama (Yedek alt test: Çiz Çıkar) şeklinde olmaktadır. 10 temel alt testin standart puanlarının toplamından da Tüm Ölçek Zekâ Puanı (TÖZP) elde edilmektedir. WÇZÖ-IV'ün, ülkemizdeki standardizasyon çalışması 2225 çocuk üzerinden yapılmıştır. Norm grubunda her bir yaş dilimi için yedi coğrafi bölge ülke nüfusuna oranlarıyla cinsiyet (kız-erkek) ve sosyoekonomik düzey (düşük-orta-yüksek) eşit olarak temsil edilmiştir (Öktem ve ark., 2013; Uluç ve ark., 2011). Bütün dönüştürülmüş puanlar için ortalama değer 100, standart sapma ise 15'tir. Her bir alt test için elde edilen standart puanların ortalaması 10, standart sapması üçtür.

İşlem

Test uygulamalarına başlamadan önce okul idaresinden, çocukların ailelerinden, araştırma hakkında bilgi veren 'bilgilendirilmiş onam formu' ile yazılı onam alınmıştır. Bu çocukların WZÇÖ-R testleri, Türk Psikologlar Derneği WÇZÖ-R uygulama sertifikasına sahip psikologlar tarafından uygulanmıştır. WÇZÖ-IV testleri de, daha önce WÇZÖ-R uygulama sertifikasına sahip olan, testin norm çalışmasını yapan proje ekibi tarafından eğitim verilen ve uygulama yeterliliği kazanan psikologlar tarafından uygulanmıştır. WÇZÖ-R Toplam ZB puanı 120 ve üstü olan çocuk ve ergenler araştırmaya dahil edilmiştir (Giriş bölümünde, üstün yetenekli çocuklar için Toplam ZB puanının, 130 ve üzeri olarak kabul edildiği ifade edilmiştir. Ancak ülkemizdeki uygulamalarda, 120 ve üstü Toplam ZB puanına sahip olanlar üstün yetenekli olarak değerlendirilmektedir). WÇZÖ-R uygulamasından en az altı ay önce bu testin uygulanmaması koşulu aranmıştır. İki test arasındaki uygulama süresi ise en az altı ay olarak belirlenmiştir. Her iki zekâ testi de kendi uygulama kitapçığında belirtildiği gibi standartlarına uygun bir şekilde uygulanmıştır. Değerlendirmeler temel alt testler ve Sayı Dizisi alt testi üzerinden yürütülmüştür.

Tablo 2. WÇZÖ-R ve WÇZÖ-IV Ortalama Puanlarının Karşılaştırılması

	WÇZÖ-R			WÇZÖ-IV				
	Ort.	*S*	*n*	*Ort.*	*S*	*n*	*Ort. fark*	*t*
Sözel ZB / SKDP	123.58	7.48	50	123.28	10.82	50	.30	.18
Performans ZB / AAYDP	130.28	8.88	50	115	13.50	50	15.34	**8.56***
Toplam ZB/TÖZP	129.38	5.94	50	123.16	11.87	50	6.22	**4.19***
ÇBDP	-	-	-	114.56	14.52	50		
İHDP	-	-	-	115.02	14.28	50		
Genel Bilgi	14.14	1.63	50	-	-	-		
Benzerlikler	13.78	1.83	50	13	2.34	50	0.78	**2.06****
Aritmetik	14.14	1.19	50	-	-	-		
Sözcük Dağarcığı	11.62	2.16	50	14.86	2.34	50	-3.24	**-7.57***
Yargılama/Kavrama	13.90	2.25	50	13.78	2.53	50	.12	.28
Sayı Dizisi	13.05	3.67	38	13.32	3.02	38	-.27	-.67
Resim Tamamlama	13.54	1.83	50	-	-	-		
Resim Düzenleme	11.52	2.73	50	-	-	-		
Küplerle Desen	15.74	2.10	50	13.86	3.16	50	1.88	**4.80***
Parça Birleştirme	14.50	2.18	50	-	-	-		
Şifre	15.86	2.86	50	13.28	3.17	50	2.58	**7.52***
Resim Kavramları	-	-	-	10.24	3.27	50		
Mantık Yürütme Kareleri	-	-	-	12.90	2.66	50		
Harf/Rakam Dizisi	-	-	-	12	2.34	50		
Simge Arama	-	-	-	11.82	2.57	50		

*$p < .001$;** $p < .05$. ZB: Zeka Bölümü; SKDP: Sözel Kavrama Dönüştürülmüş Puanı; AAYDP: Algısal Akıl Yürütme Dönüştürülmüş Puanı; ÇBDP: Çalışma Belleği Dönüştürülmüş Puanı; İHDP: İşlemleme Hızı Dönüştürülmüş Puanı; TÖZP: Tüm Ölçek Zeka Puanı.

Bulgular

Araştırma kapsamında değerlendirmeye alınan çocuk ve ergenlerin ZB ve dönüştürülmüş puanlara ilişkin betimleyici istatistikler Tablo 1'de sunulmuştur.

Araştırmanın hipotezlerinden biri olan WÇZÖ-R ZB ve alt test standart puanları ile WÇZÖ-IV alt test standart puanları ve dönüştürülmüş puanları arasında anlamlı farklılıkların olup olmadığını test etmek için eşleşmiş gruplar için t-testi analizi yapılmıştır.

Tablo 2'de görüldüğü gibi, WÇZÖ-R Toplam ZB, WÇZÖ-IV TÖZP'den ($t = 4.19$, $p < .001$); Performans ZB puanı, AAYDP'den ($t = 8.56$, $p < .001$) anlamlı düzeyde daha yüksek çıkmıştır. Benzer şekilde, WÇZÖ-R'nin Benzerlikler ($t = 2.06$, $p < .05$), Küplerle Desen ($t = 4.80$, $p < .001$) ve Şifre ($t = 7.52$, $p < .001$) alt testlerinin standart puanları, WÇZÖ-IV'ün benzer alt test standart puanlarına göre anlamlı düzeyde yüksek çıkmıştır. Buna karşılık, WÇZÖ-IV Sözcük Dağarcığı alt test standart puanı, WÇZÖ-R Sözcük Dağarcığı alt test standart puanından anlamlı düzeyde yüksek çıkmıştır ($t = -7.57$, $p < .001$).

Zekâ testinin iki sürümü arasındaki ilişkiyi değerlendirmek için, WÇZÖ-R Toplam, Sözel ve Performans ZB puanları ile WÇZÖ-IV'ün dönüştürülmüş puanları arasındaki Pearson korelasyon katsayıları hesaplanmıştır. Buna göre; benzer yapıyı ölçtüğü düşünülen Performans ZB ile AAYDP ($r = .42$, $p < .01$) ve Toplam ZB ile TÖZP ($r = .47$, $p < .01$) arasında pozitif yönde anlamlı ilişki olduğu ortaya çıkmıştır. Ancak, WÇZÖ-R Sözel ZB ile WÇZÖ-IV SKDP ($r = .23$, $p > .05$) arasındaki ilişki anlamlı çıkmamıştır (bkz. Tablo 3).

Tablo 3. WÇZÖ-R ZB Puanları ile WÇZÖ-IV Dönüştürülmüş Puanları Arasındaki İlişkiler

	WÇZÖ-R			WÇZÖ-IV				
N = 50	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1) Sözel ZB	-							
2) Performans ZB	-.16	-						
3) Toplam ZB	.61**	.68**	-					
4) SKDP	.23	.25	.39**	-				
5) AAYDP	-.01	.42**	.33*	.63**	-			
6) ÇBDP	.34*	.04	.30*	.31*	.13	-		
7) İHDP	-.18	.45**	.24	.11	.34*	.10	-	
8) TÖZP	.14	.43**	.47**	.78**	.81**	.55**	.55**	-

*$p < .05$; **$p < .01$. ZB: Zeka Bölümü; SKDP: Sözel Kavrama Dönüştürülmüş Puanı; AAYDP: Algısal Akıl Yürütme Dönüştürülmüş Puanı; ÇBDP: Çalışma Belleği Dönüştürülmüş Puanı; İHDP: İşlemleme Hızı Dönüştürülmüş Puanı; TÖZP: Tüm Ölçek Zeka Puanı.

Tartışma

Bu çalışmada, daha önce WÇZÖ-R ile üstün yetenekli olduğu belirlenen çocuklara WÇZÖ-IV uygulanarak, her iki test sürümünden elde edilen puanlar karşılaştırılmıştır. Bu doğrultuda, iki sürümün benzer şeyleri ölçtüğü düşünülen veya birbiriyle örtüşen dönüştürülmüş ve alt test standart puanları kullanılmıştır. Çalışma sonuçları, bu çocukların WÇZÖ-R Toplam ZB ile WÇZÖ-IV TÖZP ve WÇZÖ-R Performans ZB ile WÇZÖ-IV AAYDP puanları arasında anlamlı farklar olduğunu ancak WÇZÖ-R Sözel ZB ile WÇZÖ-IV SKDP puanları arasındaki farkın ise anlamlı olmadığını göstermiştir. Buna göre, söz konusu puanlar açısından WÇZÖ-R üzerinde daha yüksek puanlar elde edilmiştir. Üstün yetenekli çocuklarda, doğrudan WÇZÖ-R ve WÇZÖ-IV puanlarını karşılaştıran herhangi bir çalışma olmamasına rağmen, bu çocuklarda Wechsler zeka ölçeklerinin iki sürümünü karşılaştıran çalışmaların bulguları (Larrabee ve Holroyd, 1976; Sabatino ve ark., 1995; Sevier ve ark., 1994; Solly, 1977; Wheaton ve ark., 1980) söz konusu çalışmanın bulgularını desteklemektedir. Bu bulgular, aynı zamanda, zekâ testi normlarının her 10 yılda bir yenilenmesi gerektiğini ileri süren Flynn'ın (1984; 1987) çalışmalarını destekler niteliktedir. Buna göre, Flynn (1987) uluslararası bir meta-analiz çalışmasında Toplam ZB puanları arasında 5-25 arası puan farklıkların olabileceğini bildirmiştir. WÇZÖ-R'nin ülkemizde norm çalışması üzerinden bugüne kadar yaklaşık olarak 30 yıl geçtiği düşünüldüğünde, Flynn etkisine göre zekâ puanlarında yaklaşık 10 puanlık bir artışın olması beklenmektedir. Bu çalışmada, Toplam ZB ile TÖZP arasında 6.22 puanlık anlamlı bir fark olduğu ve bu farkın da Flynn'ın elde ettiği bulgularla tutarlı olduğu gözlenmiştir. Benzer şekilde, Performans ZB ile AAYDP arasında 15.34 puanlık anlamlı bir fark olduğu bulunmuştur. Flynn (1987) tarafından ortaya koyulan bir bulguya göre, WÇZÖ-R Performans ZB alt testlerinden elde edilen standart puanlar, Sözel ZB alt testlerinden elde edilen standart puanın iki katından daha fazla artış göstermektedir. Bu açıdan da, çalışmamızda elde edilen puan farkı alanyazınla tutarlıdır.

Çalışmanın bir diğer bulgusu, karşılaştırılabilen alt testler açısından, bu örneklemde WÇZÖ-R ile WÇZÖ-IV puanları arasında farkların olduğudur. Buna göre, Benzerlikler, Küplerle Desen ve Şifre alt test standart puanlarının WÇZÖ-R üzerinde daha yüksek olduğu görülmektedir. WÇZÖ-IV sürümü ile birlikte, üstün yetenekli öğrencileri daha iyi betimlemek için, alt testlerin en üstüne daha zor sorular yerleştirildiği bildirilmektedir (Newman ve ark., 2008). Bu üç alt test incelendiğinde, Küplerle Desen ve Benzerlikler alt testlerinde madde sayısında ve madde içeriklerinde önemli değişikler olduğu; Şifre alt testinde ise herhangi bir değişiklik olmadığı görülmektedir. Bununla birlikte, WÇZÖ-R'de bu üç alt testten aldığı standart puanı WÇZÖ-IV'te de almak isteyen bir çocuğun, bu alt testlerden daha fazla maddeyi doğru bir şekilde cevaplaması gerekmektedir. Ayrıca, bu örneklemde, üç alt test için WÇZÖ-R standart puanları 9-19 aralığındayken; WÇZÖ-IV'te 5-19 aralığında olduğu görülmüştür. Bu standart puanlar arasındaki farklılık, bu alt test puanları arasındaki farklılığı açıklamaktadır. Sözcük Dağarcığı alt test standart puanının WÇZÖ-IV'te yüksek çıkması beklenmedik bir bulgudur. WÇZÖ-R norm çalışması, yaklaşık 30 yıl önce gerçekleştirilmiştir. Bu sürümdeki Sözcük Dağar-

cığı alt testindeki sözcükler görece daha eskidir. Buna karşılık, WÇZÖ-IV Sözcük Dağarcığı alt testindeki kelimeler, daha günceldir ve mevcut müfredatlar taranarak hazırlanmıştır. Üstün yetenekli çocukların, hem bu kelimelere olan aşinalıklarından hem de sözel becerilerde yüksek performans göstermelerinden dolayı, WÇZÖ-IV Sözcük Dağarcığı alt testinden yüksek puanlar aldıkları düşünülmektedir.

WÇZÖ-R ZB ve WÇZÖ-IV dönüştürülmüş puanları arasındaki korelasyonlar incelenmiştir. Korelasyon analizi sonuçları, Toplam ZB ile TÖZP puanları ve AAYDP ile Performans ZB puanları arasında anlamlı ilişkiler olduğunu göstermektedir. Burada da en güçlü ilişkinin genel zekâ puanları arasında olduğu görülmektedir. Bu da iki sürümün genel zekâ düzeyini ölçmede birbirini desteklediğini ortaya koymaktadır. İki sürüm arasındaki ilişkiler genel olarak değerlendirildiğinde, genel zekâ puanları (TÖZP) arasında güçlü ve anlamlı ilişkilerin olması her iki sürümün devamlılığını göstermekte ve zekâyı değerlendirmede sonuçların birbirini desteklediği düşünülmektedir. Ülkemizde normal örneklemle (Uluç ve ark., 2014) ve DEHB tanısı konulan çocuklarla (Çelik ve ark., 2017) yapılan çalışmalarda, iki sürümden elde edilen puanlar arasındaki korelasyonlar açısından benzer bulgular elde edilmekle birlikte, bu çalışmalarda elde edilen korelasyon katsayılarının daha yüksek olduğu gözlenmiştir. Ek olarak, WÇZÖ-R Sözel ZB ile WÇZÖ-IV SKDP arasında anlamlı bir ilişki bulunamaması beklenmedik bir bulgu olarak karşımıza çıkmıştır. Bu bulgunun olası nedenlerinden biri, örneklem sayısınız az olması olabilir. Bir diğer neden olarak, WÇZÖ-IV SKDP'yi oluşturan alt test maddelerinde gerçekleştirilen önemli değişiklikler ve düzeltmeler gösterilebilir. Son olarak, üstün yetenekli çocukların oluşturduğu bu örneklemde, hem WÇZÖ-R'nin kendi içinde bazı ZB puanları (örn., Sözel ZB/Performans ZB) hem de WÇZÖ-IV'ün bazı dönüştürülmüş puanları (örn., İHDP/SKDP) arasında anlamlı ilişkiler bulunmamıştır. WÇZÖ-R'nin ZB puanları arasında anlamlı ilişkilerin olmamasının, değerlendirilen örneklem grubu açısından bu test sürümünün temsiliyet gücünü kaybetmiş olabileceğine ve bu nedenle, genel zekâ faktörünün bütünleşik ya da tutarlı yapısının bozulmuş olabileceğine işaret ettiği düşünülmektedir. Bu anlamda, WÇZÖ-IV'ün kendi içinde, genel zeka faktörünü destekleyen daha tutarlı bir yapıya sahip olduğu söylenebilir. Bununla birlikte, özellikle son dönemde yapılan bazı çalışmalarda, WÇZÖ-IV'ün dönüştürülmüş puanları arasında anlamlı ilişkilerin bulunmadığı görülmüştür (Hagmann-von Arx, Meyer ve Grob, 2008; Rowe, Kingsley ve Thompson, 2010). Beklenmedik görülen bu bulguların, örneklemin özelliklerinden kaynaklı olabileceği unutulmamalıdır. Bu açıdan, yukarıda aktarılan bulguların ileriki çalışmalarda açıklığa kavuşturulmasının, daha sağlıklı bilgiler sağlayacağı düşünülmektedir.

Bu çalışmada, WÇZÖ-R'ye göre daha önce üstün yetenekli olduğu belirlenen çocukların, WÇZÖ-IV ile değerlendirilmesi halinde, 17 çocuğun "üstün yeteneklilik" kategorisine dahil olamayacağı görülmüştür. Alanyazındaki çalışmalarda da benzer durumların ortaya çıktığı bildirilmiştir (örn., Sevier ve ark., 1994). Bu durum, WÇZÖ-IV'ün üstün yetenekli çocuklar için daha geçerli ve sağlıklı bir değerlendirme bilgisi ortaya koyduğunu gösterebilir. Aslında, ülkemiz açısından WÇZÖ-R'nin normları çok eski olduğundan, bu sürümle değerlendirilen normal zekâ aralığında olan çocukların üstün yetenekli olarak belirlenmesi de söz konusu olabilir. Dolayısıyla, WÇZÖ-R ile üstün yetenekli olarak değerlendirilen ancak WÇZÖ-IV ile üstün yetenekli olmadığı belirlenen çocukların, eski test normları nedeniyle üstün yeteneklilik kategorisine dâhil olabileceği söylenebilir. Bu anlamda, hâlihazırda üstün yetenekli olarak eğitim alan çocukların ve ailelerinin mağdur olmaması açısından, tekrar değerlendirme aşamasında WÇZÖ-R'nin kullanılmasının daha sağlıklı ve etik açıdan daha uygun olabileceği düşünülmektedir. Bununla birlikte, üstün yeteneklilik için ilk defa değerlendirilecek çocuklara, yeni norm ve standardizasyona sahip olan WÇZÖ-IV'ün uygulanması, değerlendirmeyi daha sağlıklı ve geçerli kılacak ve böylece çocuklar için daha doğru bir yönlendirme sağlanmış olacaktır. Her iki testin belirli aralıklarla (en az altı aylık bir süre) uygulanmış olması halinde, yenilenen test normları ile birlikte ortaya çıkan puan farkları göz önünde bulundurularak uygunluğa karar verilmesi ve bu süreçte her iki test sonuçlarının birlikte değerlendirilip karşılaştırmalı olarak ele alınması önem kazanmaktadır.

Üstün yetenekli çocukların WÇZÖ-IV SKDP ve AAYDP alanlarında daha iyi performans gösterdikleri bilinmektedir. Bu çalışmanın örneklemini oluşturan çocukların da benzer şekilde bu alanlardan yüksek puan aldıkları bulunmuştur. WÇZÖ-IV'ün sözel kavrama ve algısal akıl yürütme alanlarına ağırlık veren bir yapıda olması, bu özellikteki çocukların doğru bir şekilde değerlendirilmesine önemli bir katkı sağlamaktadır. Bununla birlikte, WÇZÖ-IV'te bazı alt testlerin (örn., Küplerle Desen) zaman ödülsüz puanlarının hesaplanabiliyor olması, işlemleme hızı görevlerinde, normal zeka aralığındaki çocuklara göre daha hızlı olmadıkları bilinen üstün yetenekli çocukların değerlendirmesini daha işlevsel hale getirebilir. Ek olarak, üstün yetenekli çocukları belirlerken Toplam ZB puanı ya da TÖZP yerine, diğer zekâ alanlarından en yüksek puanın kullanılmasını öneren çalışmalar söz konusudur (örn., Silverman, 2005). Bu öneri, değerlendirmede üstün yetenekli çocukların lehine sonuçlar doğursa da, WÇZÖ-IV'ün kuramsal altyapısında var olan zekânın bütüncül olması ("g" faktörü)

yaklaşımına uygun düşmemektedir. Nitekim Flanagan ve Kaufman (2004) WÇZÖ-IV puanlarını değerlendirme aşamasında, TÖZP'nin bütüncüllüğüne vurgu yapmıştır. Bununla birlikte, dönüştürülmüş puanlar arasında bir buçuk standart sapma (23 puan) fark olduğunda, TÖZP'nin bütüncül yapısından söz etmenin zor olduğunu belirtmişlerdir. Böyle bir durumda, TÖZP'den önemli çıkarımlarda/yorumlarda bulunmanın sağlıklı olmayacağını ileri sürmüşlerdir.

Üstün yetenekli çocukları değerlendirmede TÖZP'nin yanı sıra kullanılan bir diğer puan türü ise WÇZÖ-III'te oluşturulan (Prifitera, Weiss ve Saklofske, 1998) ve WÇZÖ-IV'te de devam eden (Prifitera, Saklofske ve Weiss, 2005) Genel Yetenek İndeksi (GYİ) (GAI-General Ability Index) puanıdır. GYİ, Sözel Kavrama (Benzerlikler, Sözcük Dağarcığı ve Sözel Kavrama) ve Algısal Akıl Yürütme (Küplerle Desen, Mantık Yürütme Kareleri ve Resim Kavramları) dönüştürülmüş puanlarının hesaplanmasında yer alan alt testlerin standart puanlarının toplamından elde edilir ve bu puanın belirlenmesine yönelik bir norm tablosu söz konusudur. Dönüştürülmüş puanlar arasında yüksek tutarsızlıklar olduğu zaman, bu puanın kullanılması önerilmektedir (Flanagan ve Kaufman, 2004). Üstün yetenekli çocukların daha iyi performans sergilediği alanlar olan sözel soyut ve görsel akıl yürütmeyi değerlendiren bu iki dönüştürülmüş puanın, Çalışma Belleği (manipüle edilen işitsel bellek) ve İşlemleme Hızı (kağıt-kalem görevlerinde hız) dönüştürülmüş puanlarına göre, üstün yeteneklileri belirlemede daha iyi göstergeler oldukları ortaya koyulmuştur (Falk ve ark., 2004; Flanagan ve Kaufman, 2004; Weiss ve ark., 2006). Bu açıdan bakıldığında, üstün yetenekli çocukları en iyi tanımladığı öne sürülen GYİ puanı için ülkemizde de norm çalışmasının yapılması, hem uygulama zamanının ve maliyetinin düşürülmesi açısından hem de üstün yetenekli çocukların zihinsel becerilerini daha kesin olarak ortaya koyması açısından yararlı olabilir.

Üstün yetenekli çocukları belirlemede ya da değerlendirmede yalnızca zekâ testi kullanmak yerine başka ek değerlendirme araçlarının kullanılması da değerlendirmenin daha sağlıklı bir şekilde gerçekleştirilmesine katkı sağlayabilir. Nitekim ülkemizde ve yurt dışında yapılan çalışmalarda, üstün yetenekli çocukları belirlemede çok yönlü değerlendirmeye (örn., yetenek, başarı ve yaratıcılık testleri) vurgu yapılmıştır. Kaufman ve Harrison (1986), zekâ testlerinin, düzgün bir şekilde kullanılması halinde, üstün yetenekli çocukları değerlendirmede önemli bir klinik araç olduğunu bildirmekle birlikte, yüzde yüz kesin sonuçlar vermediği ve yaratıcılık, artistik gibi yetenekleri ölçmemesi noktasında kısıtlı olduğuna vurgu yapmıştır. Buradan yola çıkarak, üstün yetenekliler programına uygunluk için tek bir zekâ puanı yerine, çoklu değerlendirme ölçütlerinin kullanılmasının önemli olduğunu belirtmişlerdir. Sonuç olarak, üstün yetenekli çocuklar, müzikal, yaratıcı, atletik, estetik ve sanatsal yetenekler sergileyebildiklerinden (Pfeiffer, 2002; Sattler, 2002), çoklu değerlendirme yöntemlerinin kullanılması, çeşitlilik gösteren zekâ alanlarının birlikte değerlendirilmesine olanak sağlayabilir. Bu çalışma, iki sürüm arasında ortaya çıkan puan farkları olduğundan, en azından hem WÇZÖ-IV'e geçiş aşamasında hem de sonraki değerlendirmelerde, ek bir değerlendirme aracının (örn., başarı testi, yetenek testi) gerekliliğine işaret etmektedir.

Bu çalışma, ülkemizde kullanımda olan WÇZÖ-IV ile WÇZÖ-R puanlarını üstün yetenekli çocukların oluşturduğu bir örneklemde karşılaştırması açısından ilk olma özelliğini taşımaktadır. Bu anlamda çalışmanın, değerlendirilen bu grup için bazı çıkarımların elde edilmesine katkı sağlayacağı ve bu çocukların değerlendirmesinde, WÇZÖ-IV'e geçiş aşamasında, zekâ değerlendirmelerinde ortaya çıkabilecek sorunların önceden belirlenmesi ve bilinmesi anlamında önemli bilgiler vereceği düşünülmektedir.

Genel olarak, söz konusu çalışmada ortaya çıkan puan farklılıklarının, testin yapısında ve içeriğinde gerçekleştirilen önemli değişiklerden kaynaklı olduğu düşünülebilir. Buradan hareketle, her iki sürümün benzer yapıları ölçtüğünün gözlenmesinin yanında, birkaç noktanın göz ardı edilmemesi gerektiği düşünülmektedir. Bunlardan ilki, test sürümlerinde genel zekâ puanlarının hesaplanmasında farklı alt test standart puanlarının kullanılmasıdır. Üstün yetenekli çocukların WÇZÖ-R üzerinde yüksek puanlar aldıkları temel alt testlerden Genel Bilgi ve Aritmetik, WÇZÖ-IV'te yedek alt testler bölümüne taşınmıştır. WÇZÖ-IV ile yapılan bir çalışmada da Aritmetik ve Genel Bilgi alt testlerinin, üstün yetenekliliğin en güçlü ölçümleri arasında yer alması, bu alt testlerin önemini ortaya koymaktadır (Silverman ve ark., 2004). Buna ek olarak, Tablo 2'den de görüleceği üzere, bu çocukların bataryaya yeni eklenen Resim Kavramları, Mantık Yürütme Kareleri, Harf-Rakam Dizisi ve Simge Arama alt testlerinden ortalama standart puanlar aldıkları gözlenmiştir. Bu bulguların desteklenmesi ve puan farklılıklarının yeni sürümdeki birtakım değişikliğin sonucu olup olmadığını belirlenmesi için daha fazla çalışmaya ihtiyaç olduğu düşünülmektedir. İkincisi ise, WÇZÖ-R'nin norm yenilemesinin 30 yıl gibi bir süreden sonra yapılması böyle bir puan farkının ortaya çıkmasını açıklamaktadır. Uluç ve arkadaşlarının (2014) da belirttiği gibi, zekâ testi üreticisi olmayan bir ülke olarak dışa bağımlılığımız söz konusudur. Türkiye'nin WÇZÖ-R'nin norm grubunu yenilemede geç kaldığı da açıktır. Bu noktada, belirli aralıklarla test sürümlerinin yenilenmesi (10 yılda bir), testin geçerlik ve

güvenirlik açısından daha sağlam olmasına katkı sağlayacaktır. Bundan sonraki süreçte, özellikle üstün yetenekli çocukların değerlendirilmesinde, WÇZÖ-R yerine WÇZÖ-IV'ün kullanılmasının daha sağlıklı ve yararlı sonuçlar vereceği düşünülmektedir.

Kaynaklar

Alvino, J., McDonnel, R. C. ve Richert, S. (1981). National survey of identification practices in gifted and talented education. *Exceptional Children, 48*(2), 124–132.

Bolen, L. M., Aichinger, K. S., Hall, C. W. ve Webster, R. E. (1995). A comparison of the performance of cognitively disabled children on the WISC-R and WISC-III. *Journal of Clinical Psychology, 51*(1), 89–94.

Bryant, D. C. (1992). *A comparison of WISC-R and WISC-III for gifted children.* Unpublished master's thesis, West Virginia Graduate College.

Covin, T. M. (1977). Comparison of WISC and WISC-R Full ScaleIQs for a sample of children in special education. *Psychological Reports, 41,* 237–238.

Çelik, C., Erden, G., Özmen, S. ve Hesapçıoğlu, S. T. (2017). Dikkat eksikliği hiperaktivite bozukluğunda Wechsler Çocuklar İçin Zekâ Ölçeği'nin iki sürümünün karşılaştırılması ve okuma becerilerinin değerlendirilmesi. *Türk Psikiyatri Dergisi* (Baskıda).

Çelik, C., Yiğit, İ. ve Erden, G. (2015). Wechsler Çocuklar İçin Zekâ Ölçeği Geliştirilmiş Formunun doğrulayıcı faktör analizi: Normal zihinsel gelişim gösteren çocukların oluşturduğu bir örneklem. *Türk Psikoloji Yazıları, 18*(35), 21–29.

Doll, B. ve Boren, R. (1993). Performance of severely language-impaired students on the WISC-III, language scales, and academic achievement measures. *Journal of Psychoeducational Assessment, Advances in Psychological Assessment Monograph Series,* 77–86.

Doppelt, J. E. ve Kaufman, A. S. (1977). Estimation of the differences between WISC-R and WISC IQs. *Educational and Psychological Measurement, 37*(2), 417–424.

Falk, R. F., Silverman, L. K. ve Moran, D. (2004). Using two WISC-IV indices to identify the gifted. In 51st Annual Convention of the National Association for Gifted Children, Salt Lake City, UT.

Flanagan, D. P. ve Kaufman, A. S. (2009). *Essentials of WISC-IV assessment* (Vol. 56). John Wiley & Sons.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95*(1), 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, *101*(2), 171–191.

Flynn, J. R. (1999). Searching for justice: The discovery of IQ gains over time. *American Psychologist, 54*(1), 5–20.

Goh, D. S., Teslow, J. ve Fuller, G. B. (1981). The practice of psychological assessment among school psychologists. *Professional Psychology, 12*(6), 696.

Hagmann-von Arx P., Meyer, C. S. ve Grob, A. (2008). Assessing intellectual giftedness with the WISC-IV and the IDS. *Zeitschrift für Psychologie/Journal of Psychology, 216*(3), 172–179.

Hishinuma, E. S. ve Yamakawa, R. (1993). Construct and criterion-related validity of the WISC-III for exceptional students and those who are at risk. *Journal of Psychoeducational Assessment,* 94–104.

Hutton, J. B., Dubes, R. ve Muir, S. (1992). Assessment practices of school psychologists: Ten years later. *School Psychology Review*, *12*(6), 696–706.

Kanaya, T., Scullin, M. H. ve Ceci, S. J. (2003). The Flynn effect and US policies: the impact of rising IQ scores on American society via mental retardation diagnoses. *American Psychologist*, *58*(10), 778.

Kaufman, A. S. (1992). Test reviews: Evaluation of the WISC-III and WPPSI-R for gifted children. *Roeper Review, 14*(3), 154–158.

Kaufman, A. S. ve Harrison, P. L. (1986). Intelligence tests and gifted assessment: What are the positives? *Roeper Review*, *8*(3), 154–159.

Kaufman, A.S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn and Bacon.

Klausmeier, K., Mishra, S. P. ve Maker, C. J. (1987). Identification of gifted learners: A national survey of assessment practices and training needs of school psychologists. *Gifted Child Quarterly*, *31*(3), 135–137.

Larrabee, G. J. ve Holroyd, R. G. (1976). Comparison of WISC and WISC-R using a sample of highly intelligent children. *Psychological Reports, 38*(3c), 1071–1074.

Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence, 37,* 16–24.

Lyon, M. A. (1995). A comparison between WISC-III and WISC-R scores for learning disabilities reevaluations. *Journal of Learning Disabilities*, *28*(4), 253–255.

Mahone, E. M., Miller, T. L., Koth, C. W., Mostofsky, S. H., Goldberg, M. C. ve Denckla, M. B. (2003). Differences between WISC-R and WISC-III performance scale among children with ADHD. *Psychology in the Schools*, *40*(4), 331–340.

Mark, R., Beal, A. L. ve Dumont, R. (1998). Validation of a WISC-III Short-Form for the identification of Canadian gifted students. *Canadian Journal of School Psychology, 14*(1), 1–10.

Marland, S. P. (1972). Education of the gifted and talented: Report to the Congress of the United States by the U.S. Commissioner of Education and background papers submitted to the U.S. Office of Education, 2 vols. Washington, DC: U.S. Government Printing Office.

Moffitt, T. E., Caspi, A., Harkness, A. R. ve Silva, P. A. (1993). The natural history of change to intellectual performance: Who changes? How much? Is it meaningful? *Journal of Child Psychology and Psychiatry, 34*(4), 455–506.

Munford, P. R. (1978). A comparison of the WISC and WISC-R on black child psychiatric outpatients. *Journal of Clinical Psychology, 34*(4), 938–943.

Munford, P. R., Meyerowitz, B. E. ve Munford, A. M. (1980). A comparison of black and white children's WISC/WISC-R differences. *Journal of Clinical Psychology*, *36*(2), 471–475.

Munford, P. R. ve Munoz, A. (1980). A comparison of the WISC and WISC-R on hispanic children. *Journal of Clinical Psychology*, *36*(2), 452–458.

Newby, R. F., Recht, D. R., Caldwell, J. ve Schaefer, J. (1993). Comparison of WISC-III and WISC-R IQ changes over a 2 year time span in a sample of children with dyslexia. *Journal of Psychoeducational Assessment Advances in Psychoeducational Assessment Monograph Series*, 87–93.

Newman, T. M., Sparrow, S. S. ve Pfeiffer, S. I. (2008). The use of the WISC-IV in assessment and intervention planning for children who are gifted. In Prifitera A, Saklofske DH, Weiss LG (Eds.), *WISC-IV Clinical Assessment and Intervention*. (pp. 217–242). San Diego, CA: Academic Press.

Öktem, F., Gençöz. T., Erden. G., Sezgin, N. ve Uluç S. (2013). *Wechsler Çocuklar İçin Zeka Ölçeği-IV (WÇZÖ-IV) Uygulama ve Puanlama El Kitabı Türkçe Sürümü.* Ankara: Türk Psikologlar Derneği Yayınları.

Pfeiffer, S. I. (2002). Identifying gifted and talented students: Recurring issues and promising solutions. *Journal of Applied School Psychology*, *19*(1), 31–50.

Pfeiffer, S. I. (2012). Current perspectives on the identification and assessment of gifted students. *Journal of Psychoeducational Assessment*, *30*(1), 3–9.

Prifitera, A., Saklofske, D. H. ve Weiss, L. G. (2005). *WISC–IV clinical use and interpretation: Scientist practitioner perspectives.* New York: Academic Press.

Prifitera, A., Weiss, L. G. ve Saklofske, D. H. (1998). The WISC-III in context. In A. Prifitera & D.H. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 1–38). New York: Academic Press.

Reschly, D. J. ve Davis, R. A. (1977). Comparability of WISC and WISC-R scores among borderline and mildly retarded children. *Journal of Clinical Psychology, 33*(4), 1045–1048.

Roid, G. H. (1990). *Historical continuity in intelligence assessment: Goals of the WISC-III standardization.* Washington, DC: U.S. Department of Education.

Rowe, E. W., Kingsley, J. M. ve Thompson, D. F. (2010). Predictive ability of the General Ability Index (GAI) versus the Full Scale IQ among gifted referrals. *School Psychology Quarterly, 25*(2), 119–128.

Sabatino, D. A., Spangler, R. S. ve Vance, H. B. (1995). The relationship between the Wechsler Intelligence Scale for Children-Revised and the Wechsler Intelligence Scale for Children-III scales and subtests with gifted children. *Psychology in the Schools, 32*(1), 18–23.

Sattler, J. M. (2002). *Assessment of children. Behavioral and clinical implications* (4th ed.). San Diego: Jerome M. Sattler, Publisher, Inc.

Savaşır, I. ve Şahin, N. (1995). *Wechsler çocuklar için zekâ ölçeği uygulama kitapçığı.* Ankara: Türk Psikologlar Derneği.

Sevier, R. C., Bain, S. K. ve Hildman, L. K. (1994). Comparison of WISC-R and WISC-III for gifted students. *Roeper Review, 17,* 39–42.

Silverman, L. K. (2005). *An overview of issues in assessing gifted children.* Gifted Development Center, Colorado.

Silverman, L. K., Gilman, B. J. ve Falk, R. F. (2004). Who are the gifted using the new WISC-IV. In 51st annual convention of the National Association for Gifted Children, Salt Lake City, UT.

Slate, J. R. ve Saarnio, D. A. (1995). Differences between WISC-III and WISC-R IQs: A preliminary investigation. *Journal of Psychoeducational Assessment, 13,* 340–346.

Smutny, J. F. ve Blocksom, R. H. (1990). *Education of the Gifted: Programs and Perspectives.* Bloomington: Phi Delta Kappa.

Solly, D. C. (1977). I. Comparison of WISC and WISC-R scores of mentally retarded and gifted children. *Journal of School Psychology*, *15*(3), 255–258.

Solway, K. S., Fruge, E., Hays, J. R., Cody, J. ve Gryll, S. (1976). A comparison of the WISC and WISC-R in a juvenile delinquent population. *The Journal of Psychology*, *94*(1), 101–106.

Sparrow, S. S., Pfeiffer, S. I. ve Newman, T. M. (2005). Assessment of Children Who are Gifted with the WISC-IV. In Prifitera A, Saklofske D H, Weiss LG (Eds.), *WISC–IV clinical use and interpretation: Scientist-practitioner perspectives.* (pp. 282–300). San Diego, CA: Academic Press.

Spitz, H. H. (1983). Intratest and intertest reliability and stability of the WISC, WISC-R, and WAIS Full Scale IQs in a mentally retarded population. *The Journal of Special Education*, *17*(1), 69–80.

Steen, R. G. (2009). *Are People Getting Smarter? In Human Intelligence and Medical Illness* (pp. 1–8). New York: Springer.

Stinnet, T. A., Harvey, J. M. ve Oehler-Stinnet, J. (1994). Current test usage by practicing school psychologists: A national survey. *Journal of PsychoeducationalAssessment, 12,* 331–350.

Uluç, S., Korkmaz, B. ve Şahin, Ö. (2014). Flynn etkisinin Türk örnekleminde değerlendirilmesi: WÇZÖ-R ve WÇZÖ-IV Zekâ Bölümü (ZB) puanlarının karşılaştırılması. *Türk Psikoloji Dergisi, 29*(73), 60–69.

Uluç, S., Öktem, F., Erden, G., Gençöz, T. ve Sezgin, N. (2011). Wechsler Çocuklar için Zekâ Ölçeği-IV: Klinik bağlamda zekânın değerlendirilmesinde Türkiye için yeni bir dönem. *Türk Psikoloji Yazıları*, *14*(28), 49–57.

Vance, H., Maddux, C. D., Fuller, G. B. ve Awadh, A. M. (1996). A longitudinal comparison of WISC-III and WISC-R scores of special education students. *Psychology in the Schools, 33,* 113–118.

Watkins, C. E., Campbell, V. L., Nieberding, R. ve Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*(1), 54.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. San Antonio: Psychological Corporation.

Wechsler, D. (1974). *Manual for the Wechsler intelligence scale for children, revised*. Psychological Corporation.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, Texas: The Psychological Corporation.

Wechsler, D. (2003). *WISC-IV: Administration and scoring manual*. Psychological Corporation.

Wechsler, D. (2014). *Wechsler intelligence scale for children* (5th ed.). Bloomington, MN: Pearson.

Weiss, L. G., Saklofske, D. H., Prifitera, A. ve Holdnack, J. A. (2006). *WISC-IV advanced clinical interpretation*. Academic Press.

Wheaton, P. J. ve Vandergriff, A. F. (1978). Comparison of WISC and WISC-R scores of highly gifted students in public schools. *Psychological Reports, 43,* 627–630.

Wheaton, P. J., Vandergriff, A. F. ve Nelson, W. H. (1980). Comparability of the WISC and WISC-R with bright elementary school students. *Journal of School Psychology, 18*(3), 271–275.

Wilson, M. S. ve Reschly, D. J. (1996). Assessment in school psychology training and practice. *School Psychology Review, 25*(1), 9–23.

Yiğit, İ. (2016). Üstün Yetenekli Çocukların Değerlendirilmesinde Wechsler Çocuklar İçin Zekâ Ölçeği-IV (WÇZÖ-IV) Formunun Kullanımı. *Türkiye Klinikleri Psikoloji Özel Dergisi, 1*(1), 36–43.

Turkish Journal of Psychology, June 2017, *32*(79), 92–94

Summary
Comparison of WISC-R and WISC-IV Intelligence Scores in Gifted Children

İbrahim Yiğit — Ankara University
Cihat Çelik — Ankara University
Gülsen Erden — Ankara University

Gifted children have been identified as having higher performance in the areas of academic, intellectual ability, creative or productive thinking, arts, esthetic, sports and leadership than their peers (Marland, 1972; Pfeiffer, 2002; Sattler, 2002). Gifted children make rapid progress in these areas; and they need differentiated and specific programs beyond the regular education programs (Sattler, 2002; Smutny & Blocksom, 1990). In this respect, gifted children should be assessed and identified properly for being educated in schools tailored for their abilities.

Various assessment tools are used to assess and identify gifted children. The assessment tools generally include standardized tests (intelligence, achievement, social maturity and personality tests), parent and teacher reports, observations on child's behaviors and assessment of child's creativity (Alvino, McDonnel, & Richert, 1981; Sattler, 2002). Especially, individually administered intelligence tests are defined as the most valid and reliable tools. These tests are often used in order to identify intellectual skills of gifted children and place them into necessary education programs (Newman, Sparrow, & Pfeiffer, 2008; Pfeiffer, 2002; Sattler, 2002). Although there are many intelligence tests in the assessment of gifted children, it is reported that Wechsler intelligence scales are used widely (Alvino et al. 1981; Avant, 1987; Roid, 1990; Stinnet, Harvey, & Oehler-Stinnet, 1994; Wilson & Reschly, 1996). Wechsler Intelligence Scale for Children (WISC) was developed to identify intelligence levels of children aged between 6-16 years and yielded various subtests and composite or index scores (Wechsler, 2003). The key criterion in determining gifted children is intelligence score -total or general intelligence quotient (IQ) score. Accordingly, a child obtaining a total IQ higher than or equal to 130 is identified as gifted (Sattler, 2002; Pfeiffer, 2012).

Scores obtained from intelligence scales have been observed to having systematic and general increase around the world (Flynn, 1987; Kanaya, Scullin, & Ceci, 2003). This increase in the intelligence score may be resulting from many factors such as schooling, enrichment in environmental stimulus, increased familiarity with the test, and developments in the nutrition and health fields (Lynn, 2009). According to the phenomenon known as Flynn effect, increase in general IQ scores over time cause intelligence scale norms to be obsolete. In other words, as time elapses and intelligence scale norms lose their validity, assessed children shows increasingly higher performance on the scale. Thus, there exists a 3-point increase in IQ scores in every ten years (Flynn, 1984, 1987, 1999; Steen, 2009). Considering Flynn effect, standardization and norm studies of intelligence scales should be revised periodically. These revisions come up score differences between old and new versions. Studies comparing two versions of intelligence scales in gifted children revealed that gifted children obtained higher scores from the old versions of WISC than the newer versions (Bryant, 1992; Larrabee & Holroyd, 1976; Sevier et al., 1994; Wechsler, 1991; Wheaton & Wandergriff, 1978).

The use of WISC in our country began with the introduction of WISC-R (Wechsler Intelligence Scale for Children-Revised). Adapted into Turkish by Savaşır and Şahin (1995), WISC-R has been used widely for years in hospitals, clinics, Guidance Research Centers and research areas. Specifically, in the assessment of children who need tailored education, it plays an important role as an assessment tool (Çelik, Yiğit, & Erden, 2015; Uluç et. al., 2011). Standardization and norm study of WISC-IV that is the fourth version of Wechsler intelligence scales (Wechsler, 2003) has just been completed; and the process of its use has been started in our country (Uluç et. al. 2011).

Compared to WISC-R, WISC-IV has several important modifications in its construct, content and interpretation (Flanagan & Kaufman, 2009; Wechsler, 2003). According to these modifications, two-factor structure in both WISC-R and WISC-III was eliminated completely and replaced four new index scores (composite scores) as Verbal Comprehension Index (VCI), Perceptual

Address for Correspondence: Res. Asst. İbrahim Yiğit, Ankara University, Department of Psychology, 06100 Sıhhiye/Ankara
E-mail: psk.yigit@gmail.com

Reasoning Index (PRI), Processing Speed Index (PSI), Working Memory Index (WMI). Picture Arrangement, Picture Completion and Mazes subtests were eliminated from the scale while Picture Concepts, Letter-Number Sequencing, Matrix Reasoning, Symbol Search, Cancellation and Word Reasoning subtests were added. Another change was in the Information and Arithmetic subtests, which were the core subtests in WISC-R. They were transferred to the section of additional subtests in WISC-IV. Moreover, item numbers of the subtests were increased by adding items at the low and high levels of difficulty. Finally, several changes about the test materials and administration instructions were made; and the administration manual was made more functional.

Given that our country directly have begun to use WISC-IV, skipping WISC-III (Uluç et. al., 2011), it is important to examine the validity and effectiveness of WISC-IV in the identification of gifted children and to determine score differences between WISC-R and WISC-IV. Thus, in the light of the studies showing the score differences between two versions of WISC, the purpose of the present study is to compare WISC-R and WISC-IV scores in a sample of gifted children. In other words, it aims to determine whether expected score differences between these versions exist or not.

Method

The sample of the study consisted of 50 gifted students (ages 6-16) in the city of Ankara. WISC, developed by Wechsler in 1949, is an individually administered intelligence test for children between the ages of 5 and 15. The scale was revised in 1974 as WISC-R and the age range of the scale was changed to 6-16. In our country, the standardization study of WISC-R was performed with a sample of 1639 children who were chosen from 11 different city center and it was adapted to Turkish culture by Savaşır and Şahin (1995). WISC-R consists of 10 core and 2 supplemental subtests and organizes into Verbal (Information, Similarities, Arithmetic, Comprehension, Vocabulary, Digit Span) and Performance (Picture Completion, Picture Arrangement, Block Design, Object Assembly, Coding, Mazes) composite scores.

WISC-IV is an individually administered intelligence test for children between the ages of 6 and 16 (Wechsler, 2003). It consists of totally 10 core subtests and 5 additional subtests, including 4 factors. WISC-IV provides 4 index scores as VCI (Similarities, Vocabulary and Comprehension core subtests; and Information and Word Reasoning supplemental subtests), PRI (Block Design, Picture Concepts and Matrix Reasoning core subtests; and Picture Completion supplemental subtest), WMI (Digit Span and Letter-Number Sequencing core subtests; and Arithmetic supplemental subtest) and PSI (Coding and Symbol Search core subtests; and Cancellation supplemental subtest). FSIQ score is obtained by the sum of 10 core subtest scores.

Before the children were included in the study, necessary information about the test was conveyed to their parents and school board. Then, those wished to participate were given informed consent forms. The administrations were conducted by psychologists who have both WISC-R and WISC-IV Administration Certificate being given by Turkish Psychological Association. Children who obtained Full Scale IQ higher than or equal to 120 on WISC-R and had not been administered WISC-R for at least 6 months were included in the study.

Results

A paired sample t-test was conducted to test whether there were significant differences between WISC-R IQs and subtest scores; and WISC-IV index and subtest scores. Results indicated that WISC-R Full Scale IQ score was significantly higher than WISC-IV FSIQ ($t = 3.99$, $p < .001$); and WISC-R Performance IQ score was significantly higher than WISC-IV PRI ($t = 10.70$, $p < .001$). Similarly, WISC-R Similarities, Block Design and Coding scores were significantly higher than same scores on WISC-IV ($t = 2.06$, $p < .05$; $t = 4.80$, $p < .001$; $t = 7.52$, $p < .001$, respectively). In contrast, WISC-IV Vocabulary score was significantly higher than WISC-R Vocabulary score ($t = -7.57$, $p < .001$).

The Pearson correlation coefficients were calculated for examining relationships between WISC-R IQ scores and WISC-IV index scores. Accordingly, relationships between WISC-R Performance IQ and WISC-IV PRI; WISC-R Full Scale IQ and WISC-IV FSIQ were found to be significant ($r = .41$, $p < .01$; $r = .49$, $p < .01$, respectively). However, WISC-R Verbal IQ was not significantly related to WISC-IV VCI ($r = .27$, $p > .05$).

Discussion

In this study, children identified as gifted with WISC-R were administered WISC-IV and scores from both of the two versions were compared. The results of the study showed that these children scored higher on WISC-R Performance IQ and Full Scale IQ than WISC-IV PRI and FSIQ scores. Although there is no study comparing WISC-R and WISC-IV intelligence scores in gifted children, the findings of studies comparing other versions of WISC have supported the findings of the present study (e.g., Larrabee & Holroyd, 1976; Sabatino et al., 1995; Sevier et al., 1994), indicating that old version of the scale produced higher scores.

These results were also supported by Flynn's studies (1984; 1987) proposing that intelligence scale norms should be revised in every ten years. Flynn (1987) reported 5-25 score differences between Full Scale IQ scores. Given that WISC-R norm study for our country was 30 years ago, 10-point increase might be expected in the intelligence scores according to the Flynn effect. In this study, score difference of 5.88 was found and this difference was consistent with Flynn's findings. Additionally, there was a score difference of 15.18 between WISC-R Performance IQ and WISC-IV PRI. According to one of Flynn's (1987) findings, scores obtained from WISC-R subtests in Performance IQ increased more than twice of scores obtained from Verbal IQ subtests. In this sense, the score difference found in our study is consistent with the related literature.

The correlation analysis was conducted to determine relationships between WISC-R IQ scores and WISC-IV index scores. The results of the analysis showed that there were significant relationships among Full Scale IQ/FSIQ and Performance IQ/PRI. These strong and significant relationships between IQ and index scores were thought to address the continuity of both versions and support each other in assessment of the intelligence. In our country, studies with normal (Uluç et al., 2014) and clinical (ADHD-diagnosed children) (Çelik ve ark., 2017) samples produced similar correlation coefficients with the present study.

Another type of score used in the assessment of gifted children is General Ability Index (GAI) score. GAI score is obtained from the sum of the subtest scores in VCI and PRI and there is a norm table for calculation of GAI. When high inconsistencies between the index scores exist, GAI is suggested to be used (Flanagan & Kaufman, 2004). These two indexes which gifted children obtained higher scores and evaluate abstract and visual reasoning, have been revealed to be better indicators in identifying gifted children than both WMI (manipulated auditory memory tasks) and PSI (speed in paper and pencil tasks) indexes (Falk, Silverman, & Moran, 2004; Flanagan & Kaufman, 2004; Weiss, Saklofske, Prifitera, & Holdnack, 2006). In this regard, conducting a norm study for calculation of GAI scores in our country may be useful in terms of both reducing administration time and cost; and revealing intellectual abilities of gifted children more precisely.

The use of additional assessment tools in the identification of gifted children may contribute to the assessment to be able to make it in a healthier way. The studies in the literature emphasized multiple assessments in the identification of gifted children. In a study in our country, it was reported that the use of Raven Progressive Matrices Test as well as WISC-R presented better results for the identification of gifted children (Bildiren & Uzun, 2007). According to Kaufman and Harrison (1986), when intelligence scales are used properly, they are important clinical tools in the assessment of gifted children. However, these authors also reported that intelligence scales did not generate precise results and were limited to measure abilities such as creativity and artistic, suggesting that the use of multiple assessment criteria was important. Consequently, gifted children exhibits musical, creative, athletic, esthetic and artistic abilities (Pfeiffer, 2002; Sattler, 2002); and the use of multiple assessment tools may enable diverse intelligence areas to be assessed together. This study, as it revealed score differences between two versions, have been addressed the need for an additional assessment tool in the assessment of gifted children.

The intelligence score differences in the present study may often be thought as a result of the radical changes in the structure and content of new version (WISC-IV). Therefore, although both of the two versions measure same structures, several issues should be considered. The first issue is the use of different subtest scores in the calculation of Full Scale IQ/FSIQ score in the two test versions. The core subtests of Information and Arithmetic that gifted children had higher scores on WISC-R were transferred to the section of supplemental sub-tests in WISC-IV. In a study with WISC-IV, it was found that Arithmetic and Information subtests were amongst the strongest measures of giftedness (Silverman et al. 2004). Moreover, our study showed that these children obtained average scores from Picture Concepts, Matrix Reasoning, Letter-Number Sequencing and Symbol Search subtests, which were newly added to WISC-IV. In this regard, it is most likely to be score differences between the two versions. It is believed that more studies are necessary for supporting these findings and determining whether the score differences may be due to a number of changes in the new version. The second issue is that the revision of WISC-R for our country after a period of 30 years may cause such differences between the two versions. As reported by Uluç and colleagues (2013), our country has a foreign-source dependency since it has not yet produced its own intelligence scale. It is obvious for Turkey to be late for revising WISC-R's norm sample and standardization. In conclusion, revising test versions periodically (in every 10 years) would contribute to the test to be strong and valid assessment tool.