

Wechsler Çocuklar İçin Zeka Ölçeği Geliştirilmiş Formunun Doğrulayıcı Faktör Analizi: Normal Zihinsel Gelişim Gösteren Çocukların Oluşturduğu Bir Örneklem

Cihat Çelik
Ankara Üniversitesi

İbrahim Yiğit
Ankara Üniversitesi

Gülşen Erden
Ankara Üniversitesi

Özet

Wechsler Çocuklar İçin Zeka Ölçeği Geliştirilmiş Formu (WÇZÖ-R), ülkemizde uzun yıllardır zihinsel değerlendirilmede yaygın olarak kullanılan ölçeklerden biridir. 1995 yılında norm çalışması yapıldıktan sonra, bu ölçeğin yapı geçerliliğini test etmeye yönelik ülkemizde herhangi bir çalışma yapılmamıştır. Bu çalışmanın amacı, normal zihinsel gelişim gösteren çocukların WÇZÖ-R puanları üzerinden Doğrulayıcı Faktör Analizi (DFA) yaparak ölçeğin yapı geçerliliğini incelemektir. Araştırmanın örneklemini, Genel Zeka Bölümü (ZB) 80-120 arası olan 932 çocuktan oluşmaktadır. Araştırmada, WÇZÖ-R'nin Zeka Bölümü ve 10 temel alt test puanları kullanılmıştır. Araştırmada, katılımcıların yaş ortalaması 10.97 ± 2.78 , Genel ZB ortalaması ise 101.93 ± 9.88 olarak bulunmuştur. Tek faktörlü model (g) ve iki faktörlü model (şözel ve performans) oluşturularak test edilmiş ve iki faktörlü modelin uyum indeksi değerlerinin [$\chi^2 (32, N = 932) = 132.05, \chi^2/sd = 4.12, GFI = .97, AGFI = .95, CFI = .88, RMSEA = .05$], tek faktörlü modelin uyum indeksi değerlerine [$\chi^2 (33, N = 932) = 141.10, \chi^2/sd = 4.27, GFI = .97, AGFI = .95, CFI = .86, RMSEA = .06$] göre daha iyi olduğu bulunmuştur. Araştırmanın bulguları, WÇZÖ-R yapısında var olan Wechsler'in bildirdiği iki faktörlü modelin bu çalışmanın örnekleminde de desteklendiğini göstermiştir. WÇZÖ-R'nin hem klinikte hem de eğitim alanlarında yaygın kullanımı göz önünde bulundurulduğunda, iki faktörlü modelden elde edilen puanların normal zihinsel gelişim gösteren bireylerde kullanımının daha yararlı bilgiler vereceği düşünülmektedir.

Anahtar kelimeler: Wechsler Çocuklar için Zeka Ölçeği Geliştirilmiş Formu (WÇZÖ-R), Doğrulayıcı Faktör Analizi (DFA), Normal zihinsel gelişim

Abstract

Wechsler Intelligence Scale for Children-Revised (WISC-R) is one of scales commonly used in intellectual assessment in our country for many years. There hasn't been any research aimed at testing construct validity of this scale after its norm study was conducted in 1995. The aim of this study is to test construct validity of the scale by performing confirmatory factor analysis for WISC-R scores of children who has intellectually normal development. The sample of the research consisted of 932 children whose General Intelligence Quantity (IQ) ranged between 80-120. In the research, WISC-R's IQ's and scores of 10 main subtests were used. In the research, mean age of the children is 10.97 ± 2.78 and General IQ is 101.93 ± 9.88 . One-factor model (g factor) and two-factor model (verbal and performance) was tested and it was found that fit index values of two-factor model [$\chi^2 (32, N = 932) = 132.05, \chi^2/sd = 4.12, GFI = .97, AGFI = .95, CFI = .88, RMSEA = .05$] was better than one-factor model [$\chi^2 (33, N = 932) = 141.10, \chi^2/sd = 4.27, GFI = .97, AGFI = .95, CFI = .86, RMSEA = .06$]. The research findings suggested that the two-factor model existed in structure of the WISC-R presented by Wechsler was supported in this sample. Given WISC-R is commonly used in both clinical and educational fields; it is considered that two-factor model can give more useful information about individuals with normal intellectual development than one-factor model.

Key words: Wechsler Intelligence Scale for Children-Revised (WISC-R), Confirmatory Factor Analyse (CFA), Normal intellectual development

Ülkemizde ve yurt dışında klinik uygulama ve araştırmalarda, bilişsel yapı içinde yer alan pek çok özelliği ölçmesi nedeniyle Wechsler Çocuklar İçin Zeka Ölçeği Geliştirilmiş Formu (WÇZÖ-R) (Wechsler Intelligence Scale for Children-Revised: WISC-R), zeka ölçekleri arasında en sık kullanılanlardan biri olmuştur (Erdoğan-Bakar, Soysal, Kiriş, Şahin ve Karakaş, 2005; Evinç ve Gençöz, 2007; Prifitera, Saklofske ve Weiss, 2005). WÇZÖ-R (Wechsler, 1974), 10 temel ve 2 yedek olmak üzere toplam 12 alt testten oluşan bir zeka ölçeğidir. Bu alt testlerden, Sözel Zeka Bölümü (ZB), Performans Zeka Bölümü (ZB) ve Genel Zeka Bölümü olmak üzere üç puan türü elde edilmektedir. Bu yönüyle, WÇZÖ-R'in orijinal halinin 2 faktörden (sözel ve performans) oluştuğu belirtilmektedir (Wechsler, 1974). Sözel zeka bölümünü içeren alt testler, daha çok kazanılmış/edinilmiş bilgi olarak belirtilen ve sözel kavramaya ilişkin becerileri ölçen bir faktördür (Savaşır ve Şahin, 1995). Horn (1989), kendi zeka kuramında bu alana kristalize zeka (Gc) adını vermektedir. Kristalize zeka alanı, kişilerin geçmiş yaşamından, eğitim deneyimlerinden ve kültürlerinden etkilenmektedir. Buna karşılık, akıcı zeka (Gf) olarak da kabul edilen performans zeka bölümü ise doğuştan getirilen zihinsel akıl yürütme becerisi olarak tanımlanmaktadır (Horn, 1989). Horn bu zeka alanını, yeni karşılaşılan bir durumda zihnin esnekliği ve baş etme becerisi olarak tanımlamakta, bunun yaşam deneyimleri ve kültürden bağımsız olduğunu belirtmektedir. Bu iki faktörün toplamından elde edilen Genel ZB puanının ise genel zekayı veya "g faktörünü" temsil ettiği kabul edilmektedir (Spearman, 1904). Ancak, Kaufman (1975) daha sonra WÇZÖ-R'in norm grubu üzerinde yaptığı faktör analizi çalışmasında, Sözel Kavrama, Algısal Organizasyon ve Dikkatin Dağılılılığı olmak üzere 3 faktörlü bir yapının ortaya çıktığını bulmuştur. Sözel Kavrama faktörü, Genel Bilgi, Sözcük Dağarcığı, Yargılama ve Benzerlikler alt testlerinden oluşurken; Algısal Organizasyon faktörü, Resim Tamamlama, Parça Birleştirme, Küplerle Desen, Resim Düzenleme ve Labirent alt testlerinden oluşmaktadır. Yeni faktör olan Dikkatin Dağılılılığının ise, Aritmetik, Sayı Dizisi ve Şifre alt testlerinden oluştuğu bildirilmiştir (Kaufman, 1975). Daha sonra hem normal hem de klinik örneklerde yapılan faktör analizi çalışmalarında da 3 faktörlü modelin ortaya çıktığı, fakat bu çalışmaların bazılarında üçüncü faktörün diğerlerine göre faktör yükünün daha düşük olduğu bildirilmiştir (Chan, 1984; Donders, 1993; Hill, Reddon ve Jackson, 1985; Hodges, 1982; Kaufman, 1979; McMahon ve Kuncce, 1981; Naglieri, 1981; Petersen ve Hart, 1979; Tingstrom ve Pfeiffer, 1988). Bu çalışmalarda çoğunlukla açınılayıcı (exploratory) faktör analizleri (AFA) yapılmıştır. Yapılan doğrulayıcı faktör analizi (DFA) çalışmalarında da WÇZÖ-R'da ikili ve üçlü faktör yapısının ortaya çıktığı görülmektedir (An-

derson ve Dixon, 1995; Geary ve Whitworth, 1988; Lee ve Lam, 1988; O'Grady, 1989; Ramanaiah, O'Donnell ve Ribich, 1976; Silverstein, 1982; Witta ve Keith, 1994). Anderson ve Dixon (1995), normal ve psikiyatrik tanı alan örneklem grupları ile yaptıkları çalışmada, WÇZÖ-R'in kendi ikili faktör yapısının geçerli olduğunu bulmakla birlikte; WÇZÖ-R faktör yapısını en iyi açıklayan modelin, Kaufman'ın üç faktörlü modeli olduğunu ve diğer modellerin (tek faktörlü: g modeli; iki faktörlü: Horn modeli; üç faktörlü: Bannatyne ve Guilford modeli) Kaufman'ın modeline göre, daha zayıf uyum indeksi değerlerine sahip olduklarını bulmuşlardır. Benzer bir örüntüyü, WÇZÖ-R'in norm grubu üzerinde test eden Witta ve Keith (1994) yaptıkları analizde, 11 yaş grubu için kovaryans matrislerinin istatistiksel olarak farklı olmadığını bulmuştur. Yazarlar, üç faktörlü modelin iki faktörlü modele göre daha iyi uyum gösterdiğini ve 4 faktörlü modelden ise daha belirgin bir uyum farkı olduğunu bildirmişlerdir. Buna ek olarak, üç faktörlü modelin ki-kare analizi ile yapılan ölçümlerde ve 11 yaş grubunun değişmez hata matrislerinde en iyi uyumu sürdüren faktör olduğunu bildirmişlerdir (Witta ve Keith, 1994). WÇZÖ-R'in, ikili ve üçlü faktör yapısının doğrulayıcı faktör analizi ile test edildiği başka bir çalışmada, Wechsler'in iki faktörlü yapısı ile Kaufman'ın üç faktörlü yapısının desteklediği görülmüştür (Chan ve Lin, 1996).

Ülkemizde, Wechsler Çocuklar İçin Zeka Ölçeği Geliştirilmiş Formu (WÇZÖ-R), 1995 yılında Şavaşır ve Şahin tarafından uyarlanmış ve günümüzde kliniklerde, Milli Eğitim Bakanlığı'na bağlı Rehberlik Araştırma Merkezleri'nde (RAM) ve özel merkezlerde kullanımına devam edilmektedir. Uyarlandığı tarihten bu yana, söz konusu zeka ölçeği ile ilgili çeşitli çalışmalar yapılmasına rağmen, testin nasıl bir faktör yapısı ortaya çıkardığı DFA ile yeniden ele alınmamış ve incelenmemiştir. Yakın zamanda yapılan çalışmalarda ise, klinik veya normal örneklerde açınılayıcı faktör analizi yönteminin kullanıldığı görülmektedir (Erdoğan-Bakar ve ark., 2005; Erdoğan-Bakar, Soysal, Kiriş, Işık-Taner ve Karakaş, 2011; Kiriş ve Karakaş, 2004). Erdoğan-Bakar ve arkadaşları (2005) yaptıkları çalışmada, normal ve DEHB tanılı çocukların WÇZÖ-R puanları ve faktör yapılarını karşılaştırmış ve normal örnekte WÇZÖ-R'in Sözel, Performans ve Şifre'den oluşan 3 faktörlü bir yapı ortaya çıkardığını bulmuşlardır. Sözel ve performans puanları için alt test faktör yüklerinin .46 ile .89 arasında değiştiği, şifre için ise faktör yükünün .91 olduğu bildirilmiştir. DEHB tanılı grup için WÇZÖ-R'in ölçtüğü bilinen faktör yapısının ise bozulduğu belirtilmiştir (1. Faktör: Genel Bilgi, Benzerlikler, Resim Tamamlama ve Küplerle Desen, 2. Faktör: Yargılama, Resim Düzenleme ve Parça Birleştirme, 3. Faktör: Aritmetik, Sayı Dizileri. Şifre alt testi hem 1. hem de 2. faktörde

aynı yükte çıkmıştır). Erdoğan-Bakar ve arkadaşlarının (2011) yaptığı diğer bir çalışmada, önceki bulgulara yakın sonuçların elde edildiği söylenebilir. Benzer şekilde, normal ve DEHB tanılı grupların karşılaştırıldığı söz konusu çalışmada, normal örnekleme WÇZÖ-R'in sözel ve performans şeklinde düzenlenmiş yapısına ek olarak, üçüncü faktör olan dikkatin dağılılabirliği (sayı dizisi ve şifre alt testleri) faktörü bulunmuştur. Ancak, DEHB tanılı grup için elde edilen faktör örüntüsünün WÇZÖ-R'in yapısı ile tam uyuşmadığı bildirilmiştir. Kiriş ve Karakaş (2004), DEHB tanılı örneklem grubunun Raven Standart Progresif Matrisler (RSPM) Testi, WÇZÖ-R, Conner Anne-Baba/Öğretmen Derecelendirme Ölçekleri ve Bender Geşalt testlerinin puanlarını temel bileşenler analizi ile değerlendirmiş ve WÇZÖ-R'in 11 alt testinin tek bir faktör altında toplandığını ve sözel ve performans puanlarının DEHB'li olguları ayırtmadığını bildirmişlerdir. Sözü geçen bu çalışmada, WÇZÖ-R'in alt test faktör yüklerinin .57 ile .87 arasında değiştiği ve en düşük faktör yüküne sahip olan alt testlerin sırasıyla resim düzenleme (.57) ve yargılama (.58) olduğu bildirilmiştir.

Yukarıda da değinildiği gibi, ülkemizde WÇZÖ-R'in faktör yapısına yönelik yapılan çalışmalar, normal ve klinik örneklemlemlerle daha çok temel bileşenler analizi/açımlayıcı faktör analizi düzeyinde olmuştur. Yazında görüldüğü kadarıyla, ülkemizde henüz WÇZÖ-R'in Doğrulayıcı Faktör Analizi ile yapı geçerliliğinin test edildiği herhangi bir çalışmaya rastlanmamıştır. Yakın zamanda, Wechsler Çocuklar İçin Zeka Ölçeği-IV (WÇZÖ-IV)'ün norm çalışması bitirilmiş ancak henüz alanda kullanımına başlanmamıştır. Bu yönüyle, yeni sürümün kullanımının yaygınlaşması için belirli bir sürenin geçmesi gerektiği göz önünde bulundurulursa, WÇZÖ-R'in bir süre daha alanda kullanımına devam edileceği öngörülmektedir. Bu çalışmanın amacı, WÇZÖ-R'in norm çalışmasında ortaya çıkan faktörlerin, farklı bir normal örneklem grubunda ortaya çıkıp çıkmayacağını incelemektir. Bu amaç doğrultusunda, doğrulayıcı faktör analizinin uygun bir yöntem olacağına, yazında yapılan çalışmalar da göz önünde bulundurularak karar verilmiştir.

Yöntem

Örneklem

Araştırmanın örnekleme, 2005-2012 yılları arasında Türk Psikologlar Derneği ve Ankara Üniversitesi Psikoloji Bölümü'nde WÇZÖ-R Uygulama Eğitimi kapsamında WÇZÖ-R uygulanan 6-16 yaş arası ($Ort. = 10.97$, $S = 2.78$) 932 çocuktan oluşmaktadır. Bu çocukların test uygulamaları, test için gerekli eğitimi alıp, uygulama yeterliliğini (sertifika almaya) kazanmış olan psikologlar tarafından gerçekleştirilmiştir. Genel Zeka Bölümü, normal zeka düzeyi olarak kabul edilen 80-120 puan aralığı ($Ort. = 101.93$, $S = 9.88$) dışında kalan katılımcılar

araştırmaya dahil edilmemiştir. Katılımcıların, 502'si kız (%54), 430'u erkek (%46) çocuktan oluşmaktadır.

Veri Toplama Araçları

Wechsler Çocuklar İçin Zeka Ölçeği-Geliştirilmiş Formu (Wechsler Intelligence Scale for Children-WISC-R). Wechsler Çocuklar İçin Zeka Ölçeği (Wechsler Intelligence Scale for Children, WISC) 1949 yılında Wechsler tarafından 5 ile 15 yaşları arasındaki çocukların zeka düzeylerini ölçmek amacı ile geliştirilmiş olup; 1974 yılında yeniden düzenlenmiş (WÇZÖ-R) ve bu hali ile ölçeğin uygulanabildiği yaş grubu 6 ile 16 arasında çıkarılmıştır (Wechsler, 1974). WÇZÖ-R, ülkemizde Savaşır ve Şahin (1995) tarafından, 11 kent merkezinden seçilen, toplam 1639 çocuktan oluşan bir örneklem üzerinde standardizasyon çalışması yapılmış ve Türk kültürüne uyarlanmıştır. WÇZÖ-R toplam 10 temel ve 2 yedek alt test olmak üzere 6 sözel (Genel Bilgi, Benzerlikler, Aritmetik, Yargılama, Sözcük Dağarcığı, Sayı Dizisi) ve 6 performans (Resim Tamamlama, Resim Düzenleme, Küplerle Desen, Parça Birleştirme, Şifre, Labirent) alt testinden oluşmaktadır. Bu alt testlere yönelik standart puanların yanı sıra, Sözel Zeka Bölümü, Performans Zeka Bölümü ve Genel Zeka Bölümü katsayıları da elde edilmektedir. Bu Zeka Bölümü katsayıları için ortalama değer 100, standart sapma ise 15'tir. Her bir alt test için elde edilen standart puanların ise ortalaması 10, standart sapması 3'tür.

İşlem

Örnekleme oluşturan çocuklar uygulamaya dahil edilmeden önce, ailelerine uygulanacak test hakkında gerekli bilgiler aktarılmış; sonrasında ise gönüllü olarak katılmak isteyenlerden bilgilendirilmiş onam formu ile izin alınmıştır. Uygulamalar 10 temel alt test üzerinden gerçekleştirilmiş; yedek alt testler olarak bilinen Labirentler uygulanmamış ancak Sayı Dizisi alt testi süre ve durum uygun olduğunda uygulanabildiğinden tüm çocuklara uygulanamamıştır. Testin uygulandığı tarihten en az 6 ay önce bu testin uygulanmamış olması, çocuğun ruhsal duygusal ve nörolojik yakınmaları ve bu yakınmalar nedeniyle kliniğe başvurmamış olması, araştırma-ya dahil edilme ölçütleri olarak belirlenmiştir.

İstatistiksel Analiz

Bu çalışmanın amacı doğrultusunda, WÇZÖ-R'in faktör yapısının geçerliliğini test etmek için AMOS 21 paket programı kullanılarak Doğrulayıcı Faktör Analizi (Confirmatory Factor Analysis-CFA) yöntemi kullanılmıştır. Yazında yer alan kuramsal ve uygulamaya yönelik çalışmalara dayanarak, tek faktörlü model (g faktörü) (Spearman, 1904) ve iki faktörlü model (Sözel Zeka ve Performans Zeka) (Wechsler, 1974) olmak üzere iki model test edilmiştir. Sayı dizisi alt testi, ölçeğin, uygulama

süresinin uzunluğu, çocukların okula yetişmeleri, velisinin durumu ya da çocukların devam etmek istememeleri nedeniyle örneklem grubundaki çocukların tümüne uygulanamamıştır. Dolayısıyla, sadece toplam 123 çocuğa sayı dizisi alt testi uygulanmış olup, Kaufman'ın sayı dizisi alt testini içeren üç faktörlü modeli için yeterli sayıya ulaşılmadığından bu model test edilememiştir.

Doğrulayıcı faktör analizi sonucunda, önerilen modellerin veriye ne oranda uyum gösterdiğini belirlemek amacıyla çeşitli uyum ya da uygunluk indeksleri kullanılmaktadır. Uyum indeksleri, elde edilen verinin hipotetik modele uygunluğunu test edebilmesi açısından önemlidir (Meydan ve Şeşen, 2011). Bu çalışmada, yaygın olarak kullanılan χ^2 (Ki-kare), χ^2/sd , GFI (Goodness of Fit Index- Uyum İyiliği İndeksi), AGFI (Adjustment Goodness of Fit Index- Düzeltilmiş Uyum İyiliği İndeksi), CFI (Comparative Fit Index- Karşılaştırmalı Uyum İndeksi) ve RMSEA (Root Mean Square Error of Approximation- Yaklaşık Hataların Ortalama Karekökü) uyum indeksleri kullanılmıştır. χ^2 testi, geliştirilen model ile gözlem değişkenlerine ait kovaryans yapısında ortaya çıkan modelin farklı olup olmadığı hipotezini test etmektedir. Bu açıdan, χ^2 değerleri geliştirilen model ile elde edilen verinin uyuşmadığını gösterir ve küçük ya da istatistiksel olarak anlamsız değerler iyi uyuma işaret eder (Şimşek, 2007).

χ^2 değerinin örneklem sayısının artışıyla etkileneceği ve örneklem büyüdükçe istatistiksel anlamlılık kazanması nedeniyle diğer uyum indekslerinin kullanılması önerilmektedir (Byrne, 2001). Bunlardan ki-kare değerinin serbestlik derecesine oranı (χ^2/sd) uyum ölçütü olarak kullanılabilir. Bu oranın 5'ten küçük olması iyi uyum göstergesi olarak kabul edilmektedir (Kelloway, 1998). Diğer uyum indekslerinden RMSEA değerinin 0.05'e eşit veya küçük olması, model ile veri arasında mükemmel uyum; 0.08'e kadar olan değerlerin ise (0.08 de yeterli bir uyum değeri olarak kabul

edilmektedir) kabul edilebilir bir uyum olduğunu göstermektedir (Hu ve Bentler, 1999; Sümer, 2000). GFI, AGFI ve CFI uyum indeksi değerlerinin artması uyum iyiliğinin göstergesidir. GFI, AGFI ve CFI değerlerinin 0.90 ve üzerinde olması iyi uyum göstergesi olarak kabul edilmektedir (Byrne, 1998; Kline, 2005).

Bulgular

Her iki model için doğrulayıcı faktör analizi sonuçları χ^2 , serbestlik derecesi (sd), χ^2/sd , GFI, AGFI, CFI, RMSEA değerleriyle birlikte Tablo 1'de verilmiştir. Tablo 1 incelendiğinde, tek faktörlü model için χ^2 (35, $N = 932$) = 266.33, $p < .05$; iki faktörlü model için χ^2 (34, $N = 932$) = 223.95, $p < .05$ olarak bulunmuştur. Ki-kare istatistik değerleri tek başına değerlendirildiğinde, her iki modelde de bu değerlerin büyük ve istatistiksel olarak anlamlı çıktığı ve modellerin veri ile uyumlu çıkmadığı görülmektedir. Ancak, daha önce de ifade edildiği üzere, χ^2 değeri örneklem büyüklüğünden etkilendiğinden, örneklem sayısı arttıkça herhangi bir modeli reddetme olasılığı da artmaktadır. Bu açıdan değerlendirildiğinde, χ^2 değeri yerine diğer uyum indeksi değerlerinin kullanılmasının daha uygun olduğu bildirilmektedir (Bentler ve Bonett, 1980).

Diğer uyum indeksleri incelendiğinde, tek faktörlü model için GFI ve AGFI değerleri (sırasıyla, .94, .91) iyi uyuma işaret ederken; RMSEA değeri (.08) kabul edilebilir uyuma işaret etmektedir. χ^2/sd oranı 7.60 olarak bulunmuştur ve bu oran, modelin veriye uyum sağlamadığını göstermektedir. CFI değeri (.72) ise önerilen modelin veriye uyumlu olmadığını göstermektedir. Tek faktörlü modelin alt test faktör yüklerinin .17 ile .57 aralığında değiştiği (sırasıyla şifre, genel bilgi) ve faktöre yüklenen tüm alt testlerin istatistiksel olarak anlamlı olduğu görülmektedir. Doğrulayıcı faktör analizi sonuçlarına ilişkin değerler Tablo 2'den incelenebilir.

Tablo 1. Uyum İndeksi Değerleri

Model	χ^2	sd	χ^2/sd	GFI	AGFI	CFI	RMSEA
Tek faktörlü model	266.33	35	7.60	.94	.91	.72	.08
Tek faktörlü model (1 hata ilişkilendirmesi)	185.50	34	5.45	.96	.94	.81	.07
Tek faktörlü model (2 hata ilişkilendirmesi)	141.10	33	4.27	.97	.95	.86	.06
İki faktörlü model	223.95	34	6.58	.95	.92	.78	.07
İki faktörlü model (1 hata ilişkilendirmesi)	171.66	33	5.20	.96	.94	.84	.06
İki faktörlü model (2 hata ilişkilendirmesi)	132.05	32	4.12	.97	.95	.88	.05

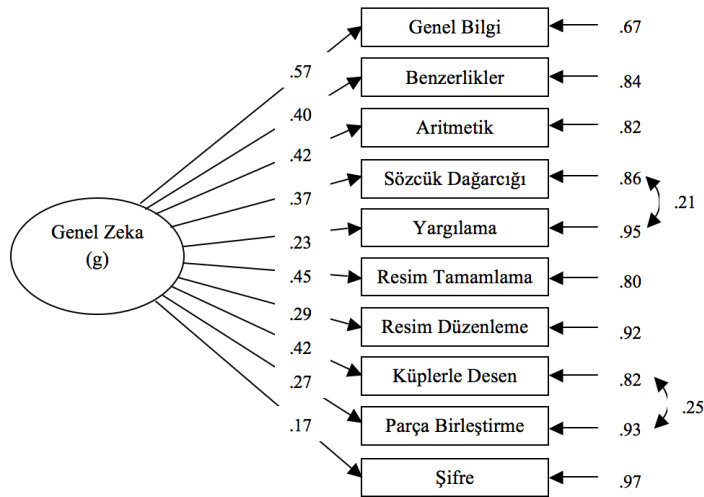
Tablo 2. WÇZÖ-R'in Doğrulayıcı Faktör Analizi Sonuçları

		β	t	SH	
Tek Faktörlü Model	Genel Zeka	Genel Bilgi	.57	14.13**	1.49
		Benzerlikler	.40	9.85**	1.13
		Aritmetik	.42	10.42**	1.15
		Sözcük Dağarcığı	.37	9.10**	.90
		Yargılama	.23	5.40**	.60
		Resim Tamamlama	.45	11.06**	1.20
		Resim Düzenleme	.29	7.03**	.81
		Küplerle Desen	.42	10.46**	1.18
		Parça Birleştirme	.27	6.64**	.75
		Şifre	.17	4.19**	.52
İki Faktörlü Model	Sözel Zeka	Genel Bilgi	.61	14.23**	1.60
		Benzerlikler	.41	9.81**	1.15
		Aritmetik	.42	10.23**	1.16
		Sözcük Dağarcığı	.39	9.39**	.95
		Yargılama	.22	5.07**	.58
	Performans Zeka	Resim Tamamlama	.52	11.53**	1.41
		Resim Düzenleme	.30	6.95**	.86
		Küplerle Desen	.46	10.36**	1.29
		Parça Birleştirme	.33	7.28**	.93
		Şifre	.16	3.73**	.49

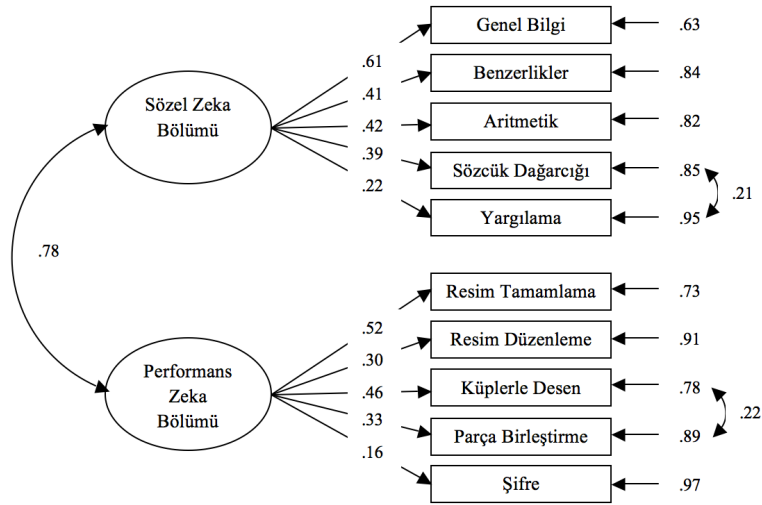
** $p < .01$

Önerilen modifikasyon indeksleri doğrultusunda ilk olarak, tek faktörlü modelde sözcük dağarcığı ve yargılama alt testlerinin hataları ilişkilendirilmiştir. Bu hata ilişkilendirmesi sonrasında, iki model ki-kare fark testi (χ^2 difference test) ile karşılaştırılmıştır (Tabachnick ve Fidell, 2001). Test sonucu, bu hata ilişkilendirmesinin modeli daha uyumlu hale getirdiğini göstermektedir [$\chi^2_{\text{fark}}(1, N = 932) = 80.83, p < .05$]. Bu hata ilişkilendirmesi gerçekleştirildikten sonra model yeniden test edilmiştir. Benzer şekilde, modifikasyon indekslerinde önerilen hata ilişkilendirmesi doğrultusunda, parça birleştirme ve küplerle desen alt testlerinin de hataları ilişkilendirilmiştir. Test sonucunda, bu hata ilişkilendirmesinin modeli daha uyumlu hale getirdiği görülmektedir [$\chi^2_{\text{fark}}(1, N = 932) = 44.40, p < .05$]. İki hatanın ilişkilendirildiği en son modelin, diğer modellerden anlamlı olarak daha iyi uyuma sahip olduğu ve model uyum indekslerinin daha iyi sonuçlar ortaya koyduğu görülmektedir [$\chi^2(33, N = 932) = 141.10, \chi^2/sd = 4.27, GFI = .97, AGFI = .95, CFI = .86, RMSEA = .06$].

İki faktörlü model için uyum indeksleri incelendiğinde, GFI ve AGFI değerlerinin (sırasıyla, .95, .92) iyi uyuma; RMSEA değerinin (.07) ise, kabul edilebilir uyuma işaret ettiği görülmektedir. χ^2/sd oranı 6.58 olarak bulunmuştur ve bu oran, modelin veriye uyum sağlamadığını göstermektedir. Tek faktörlü modelde olduğu gibi, iki faktörlü modelde de CFI değeri (.77) modelin veriye uyumlu olmadığını göstermektedir (Tablo 1).



Şekil 1. Normal Örnekleme ($N = 932$) WÇZÖ-R'in Tek Faktörlü Yapısı Not. Şekilde gösterilen değerler (sağdan sola doğru): WÇZÖ-R alt testleri için hata varyansları (unique/error variances), standardize edilmiş faktör yükleri.



Şekil 2. Normal Örneklemde ($N = 932$) WÇZÖ-R'nin İki Faktörlü Yapısı
 Not. Şekilde gösterilen değerler (sağdan sola doğru): WÇZÖ-R alt testleri için hata varyansları (unique/error variances), standardize edilmiş faktör yükleri, iki faktör arasındaki korelasyon katsayısı.

İki faktörlü modelde alt test faktör yükleri sözel zeka bölümü için .22 ile .61 aralığında (sırasıyla yargılama, genel bilgi), performans bölümü için ise .16 ile .52 aralığında (sırasıyla şifre, resim tamamlama) değiştiği ve her iki faktöre de yüklenen tüm alt testlerin istatistiksel olarak anlamlı olduğu görülmektedir (sonuçlar için bkz. Tablo 2).

İki faktörlü model için önerilen modifikasyon indeksleri doğrultusunda, sözcük dağarcığı ve yargılama alt testlerinin hataları ilişkilendirilmiştir. Bu ilişkilendirme sonrasında, iki model ki-kare fark testi ile karşılaştırılmıştır (Tabachnick ve Fidell, 2001). Test sonucu, bu hata ilişkilendirmesinin modeli daha uyumlu hale getirdiğini göstermektedir [$\chi^2_{\text{fark}}(1, N = 932) = 52.29, p < .05$]. Bu hata ilişkilendirmesi gerçekleştirildikten sonra model yeniden test edilmiştir. Benzer şekilde, modifikasyon indekslerinde önerilen hata ilişkilendirmesi doğrultusunda, parça birleştirme ve küplerle desen alt testlerinin de hataları ilişkilendirilmiştir. Test sonucunda, bu hata ilişkilendirmesinin modeli daha uyumlu hale getirdiği görülmektedir [$\chi^2_{\text{fark}}(1, N = 932) = 39.61, p < .05$]. İki hatanın ilişkilendirildiği en son modelin, diğer modellerden anlamlı olarak daha iyi uyuma sahip olduğu ve model uyum indekslerinin daha iyi sonuçlar ortaya koyduğu görülmektedir [$\chi^2(32, N = 932) = 132.05, \chi^2/\text{sd} = 4.12, \text{GFI} = .97, \text{AGFI} = .95, \text{CFI} = .88, \text{RMSEA} = .05$].

Tek faktörlü model ile iki faktörlü model, uyum istatistikleri ile tutarlı olarak, χ^2 değerlerinin farkları açısından karşılaştırıldığında, iki faktörlü modelin tek faktörlü modelden istatistiksel olarak anlamlı düzeyde daha iyi olduğu görülmektedir [$\chi^2_{\text{fark}}(1, N = 932) = 9.05, p < .05$].

Tartışma

Doğrulayıcı faktör analizi (DFA) ile elde edilen bulgular incelendiğinde, Wechsler'in kendi önerdiği 2 faktörlü modelin, bütün alt testlerin tek bir faktöre bağlandığı modele göre daha iyi uyum gösterdiği görülmektedir. Diğer bir deyişle, bu çalışmada WÇZÖ-R'nin test ettiği ileri sürülen sözel kavrama becerilerini ölçen Sözel ZB ile akıl yürütme ve muhakeme becerilerini ölçen Performans ZB faktörlerinin normal örneklem gruplarında doğru bir şekilde test edilebildiği/ölçülebildiği görülmektedir. Ayrıca, tek faktörlü modelin de bazı uyum indeksleri açısından kabul edilebilir düzeyde olduğu görülmektedir. Bu bulgular, WÇZÖ-R'nin sözel, performans ve bu iki puandan elde edilen genel ZB puanlarını değerlendirmede yeterli olabileceğini ortaya koymaktadır.

Bu çalışmanın bulguları, yurt dışında yapılan DFA analizi ile ilişkili çalışmaların sonuçları ile uyumlu çıkmıştır. DFA ile yapılan çalışmaların çoğunda, WÇZÖ-R'da ikili ve üçlü faktör yapıları test edilmiş ve ikili fak-

tör yapısının geçerli olduğu veya iyi uyum indekslerine sahip olduğu bulunmuştur (Anderson ve Dixon, 1995; Geary ve Whitworth, 1988; Lee ve Lam, 1988; O'Grady, 1989; Ramanaiah ve ark., 1976; Silverstein, 1982; Witta ve Keith, 1994). Ancak, Wechsler'in iki faktörlü modeli ile Kaufman'ın önerdiği üç faktörlü modelin karşılaştırıldığı çalışmalarda farklı sonuçlar çıkmıştır. Anderson ve Dixon (1995), normal ve psikiyatrik tanı alan örneklem grupları ile yaptıkları çalışmada, WÇZÖ-R kendi ikili faktör yapısının geçerli olduğunu bulmakla birlikte; WÇZÖ-R faktör yapısını en iyi açıklayan modelin, Kaufman'ın üç faktörlü modeli olduğunu öne sürmüştür. Üç faktörlü modelin iki faktörlü veya tek faktörlü modele göre daha iyi uyum gösterdiğini bildiren başka çalışmalar olmakla birlikte (Geary ve Whitworth, 1988; Lee ve Lam, 1988; Witta ve Keith 1994), Chan ve Lin (1996), DFA ile yaptıkları çalışmada, Wechsler'in iki faktörlü yapısının görece Kaufman'ın üç faktörlü yapısına göre daha iyi uyum indekslerine sahip olduğunu belirtmiştir. Bu çalışmamızda örneklem grubunun WÇZÖ-R sayı dizisi alt testi yeterli sayıya ulaşmadığından Kaufman'ın üç faktörlü modeli test edilememiştir.

Benzer şekilde, bu çalışmanın bulguları ülkemizde yapılan faktör analizi çalışmalarının bulguları ile tutarlı çıkmıştır (Erdoğan-Bakar ve ark., 2005; Erdoğan-Bakar ve ark., 2011; Kiriş ve Karakaş, 2004). Normal örneklem grupları ile DEHB tanı grupların karşılaştırıldığı çalışmalarda, normal örnekleme WÇZÖ-R'in sözel ve performans şeklinde düzenlenmiş yapısının ortaya çıktığı ancak DEHB tanı grup için elde edilen faktör örüntüsünün WÇZÖ-R'in yapısı ile tam uyum olmadığını bildirmiştir (Erdoğan-Bakar ve ark., 2005; Erdoğan-Bakar ve ark., 2011). Normal zihinsel gelişim gösteren çocukların oluşturduğu örneklemler ile açılımlı faktör analizi kullanılarak yapılan bu çalışmalarda ortaya çıkan WÇZÖ-R'in iki faktörlü yapısının, bu çalışmada da desteklenmediği görülmektedir.

Tek faktörlü modelde alt test faktör yükleri incelendiğinde, en düşük faktör yüküne sahip alt testlerin şifre, yargılama ve parça birleştirme alt testleri olduğu görülmektedir (bkz. Şekil 1). Bu bulgu, Kiriş ve Karakaş'ın (2004) WÇZÖ-R'in tek bir faktörde toplandığını bulduğu çalışmanın sonuçları ile uyumaktadır. İki faktörlü modelde ise, sözel kavramayı ölçen alt testlerden en yüksek faktör yüküne genel bilgi alt testinin, en düşük faktör yüküne ise yargılama alt testinin sahip olduğu görülmektedir (bkz. Şekil 2). Performans alt testlerinde ise küplerle desen ve resim tamamlama alt testlerinin en yüksek faktör yüküne; şifre ve resim düzenleme alt testlerinin ise en düşük faktör yüküne sahip olduğu görülmektedir. Bu bulgular, Erdoğan-Bakar ve arkadaşlarının yaptığı çalışmaların (2005;2011) sonuçları ile uyumaktadır. Bu çalışmaların birincisinde, sözel alt testlerden en düşük faktör yüküne yargılama alt testi; performans

alt testlerinden ise resim düzenleme alt testi olduğu görülmektedir. Sözü geçen bu çalışmada, sayı dizisi sözel becerileri değerlendirmeye yönelik yedek alt test olması rağmen, performans becerileri ölçen faktörün altında yüklendiği ve şifre alt testinin ise ayrı bir faktör olarak çıktığı bildirilmiştir (Erdoğan-Bakar ve ark., 2005). Erdoğan-Bakar ve arkadaşlarının (2011) yaptığı diğer çalışmada ise, sözel alt testlerde en yüksek faktör yükünün genel bilgi alt testinde olduğu, diğer sözel alt testlerin faktör yüklerinin birbirine yakın olduğu bildirilmiştir. Algısal organizasyon olarak bildirdikleri ikinci faktörde ise, en yüksek faktör yüküne resim tamamlama alt testinin; en düşük faktör yüküne ise resim düzenleme alt testinin sahip olduğu bulunmuştur. Ayrıca, şifre ve sayı dizisi alt testlerinin ayrı bir faktör olarak üçüncü bir faktörde toplandığı bildirilmiştir. Bu çalışmada ise, hem iki faktörlü modelde hem de tek faktörlü modelde şifre alt testinin faktör yükünün düşük olduğu gözlenmektedir. Bu bulgunun, bu alt testin sayı dizisi alt testi ile birlikte Kaufman'ın önerdiği modelde dikkatin dağılılırılığı olarak adlandırılan üçüncü faktörde olabileceği yönündeki bulguları desteklediği düşünülmektedir (Anderson ve Dixon, 1995; Erdoğan-Bakar ve ark., 2005; Erdoğan-Bakar ve ark., 2011; Geary ve Whitworth, 1988; Lee ve Lam, 1988; Witta ve Keith 1994). Ancak daha önce de ifade edildiği gibi, sayı dizisi alt testi uygulanan örneklem sayısı, analiz yapılacak yeterlikte olmadığından bu çalışmada üç faktörlü model test edilememiştir.

Araştırmanın bulguları sonucunda her iki modelde alt testler anlamlı bir şekilde kendi faktörlerine yüklendiği görülmesine rağmen, bazı alt testlerin (şifre ve yargılama) faktör yüklerinin düşük olduğu görülmektedir (bkz. Tablo 2). Bunlardan biri olan şifre alt testi için yukarıda açıklama yapılmış olup, bunun ayrı bir faktör (dikkatin dağılılırılığı) olarak değerlendirilebileceği yazında belirtilmiştir. Testin kuramsal yapısı göz önünde bulundurularak her iki modelde yapılan hata ilişkilendirmeleri sonucunda, hem bazı genel uyum indekslerinin hem de alt test faktör yüklerinin sınırdaki olması, bu zeka ölçeğinin bazı alt testlerinin madde içeriklerinde bazı güçlükler olabileceğini ortaya koymaktadır. Ülkemizde, bu konuda yapılan bir çalışmada, sözel alt testlerin her biri için madde korelasyonları .68-.89 arasında değiştiği bulunmuştur (Çakır, Erden, Akgiray, Demir, 2006). Bir diğer çalışmada ise özellikle sözel alt testlerde madde güçlüklerinin değiştiği saptanmıştır (Demir, Erden, Adalı, 2006). Bu çalışma kapsamında madde analizi ve madde içeriklerine bakılmadığından, alt testlerin maddeleri bazında yorumlarda bulunmak güç olacaktır. Yine de çalışmanın bulguları göz önünde bulundurulduğunda, genel zihinsel değerlendirme dışında, alt test düzeyinde veya profil analizleri düzeyinde yapılan değerlendirmelerde; bazı alt testlerin yorumlanmasının dikkatli bir şekilde ele alınması gerektiği düşünülmektedir.

Sonuç olarak, bu çalışmanın bulguları, ülkemizde normal örneklem gruplarının zihinsel becerilerini değerlendirmede, WÇZÖ-R'in iki faktörlü yapısı ile elde edilen puanların kullanımının uygun olduğunu göstermektedir. WÇZÖ-IV'e geçiş döneminde hastane ve klinik uygulamalarda kullanımına devam edilecek olan WÇZÖ-R'in klinik örneklem gruplarında da doğrulayıcı faktör analizi ile test edilmesinin, bu örneklem gruplarında kullanımını daha güvenli hale getireceği öngörülmektedir. Ayrıca, Kaufman'ın ortaya koyduğu üç faktörlü modeli destekleyen çalışmalar göz önünde bulundurulduğunda, normal ve klinik örneklem gruplarında yapılacak olan çalışmalarda, sayı dizisi alt testinin de uygulanıp; bu üç faktörlü modelin test edilmesinin ilgili yazına önemli katkılar sağlayacağı düşünülmektedir.

Ülkemizde yakın zamanda norm çalışması bitirilen WÇZÖ-IV'ün (Uluç, Öktem, Erden, Gençöz ve Sezgin, 2011) kullanımına başlanması sürecinde, WÇZÖ-R'in da kullanımına bir süre daha devam edilmesi kaçınılmazdır. Bu çalışma bulguları, WÇZÖ-R'in klinik ve eğitim alanındaki uygulamalarda en azından ikili faktör yapısının korunması gerekliliği ortaya koymaktadır. Ülkemizde özellikle, zaman kısıtlılığı nedeniyle WÇZÖ-R uygulamalarında gelişmiş güzel temel alt test ya da testlerin çıkarılmasına ilişkin yaklaşımların olduğu bilinmektedir. Böylece, WÇZÖ-R'in bazı temel alt testlerinin uygulanmamasıyla yapılan değerlendirmelerin, test uygulanan çocuğun zihinsel gelişimi hakkında yeterli bilgi veremeyeceği düşünülmektedir. Klinik örneklemelerde yapılan değerlendirmelerde de iki faktörlü yapının önemi göz önünde bulundurulursa, bu yapının korunması gerektiği görülmektedir. Sonuç olarak, bu zeka ölçeğinin bütün temel alt testleri ile uygulanmasının ve çocuğun yakınmaları ya da tanısı kapsamında alt testler ve profil analizleri düzeyinde de değerlendirmelerin yapılmasının daha sağlıklı ve güvenilir sonuçlar vereceği düşünülmektedir.

Kaynaklar

- Anderson, T. ve Dixon E. W. Jr. (1995). Confirmatory factor analysis of the Wechsler Intelligence Scale for Children-Revised with normal and psychiatric adolescents. *Journal of Research on Adolescence*, 5(3), 319-332.
- Bentler, P. M. ve Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. Mahwah, NJ: Erlbaum.
- Chan, D. W. (1984). Factor analysis of the HK-WISC at 11 age levels between 5 and 15 years. *Journal of consulting and clinical psychology*, 52(3), 482.
- Chan, D. W. ve Lin, W. Y. (1996). The two-and three-dimensional models of the HK-WISC: A confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development*, 28(4), 191-99.
- Çakır, P., Erden, G., Akgiray, A. ve Demir, J. (2006). Wechsler intelligence scale for children – revised form of Turkish version: the current patterns of the responses. *International Testing Conference*, 6-8 Temmuz 2006, Brüksel.
- Demir, J., Erden, G. ve Adalı, N. (2006). Wechsler intelligence scale for children – revised form of Turkish version: the order of the item difficulty. *International Testing Conference*, 6-8 Temmuz 2006, Brüksel.
- Donders, J. (1993). Factor structure of the WISC-R in children with traumatic brain injury. *Journal of Clinical Psychology*, 49, 255-260.
- Erdoğan-Bakar, E., Soysal, A. Ş., Kiriş, N., Işık Taner, Y. ve Karakaş, S. (2011). Wechsler Çocuklar için Zeka Ölçeği Yeniden Gözden Geçirilmiş Formu'nun dikkat eksikliği hiperaktivite bozukluğunda ölçtüğü özellikler. *Çocuk ve Gençlik Ruh Sağlığı Dergisi*, 18(3), 155-174.
- Erdoğan-Bakar, E., Soysal, A. Ş., Kiriş, N., Şahin, A. ve Karakaş, S. (2005). Dikkat eksikliği hiperaktivite bozukluğunun değerlendirilmesinde Wechsler Çocuklar için Zeka Ölçeği Geliştirilmiş Formu'nun yeri. *Klinik Psikiyatri*, 8(1), 5-17.
- Evinç, G. Ş. ve Gençöz, T. (2007). Dikkat eksikliği hiperaktivite bozukluğu tanısı alan çocukların WISC-R profillerinin, farklı bir psikiyatrik tanı alan ve herhangi bir tanısı olmayan çocuklarla karşılaştırılması. *Türk Psikiyatri Dergisi*, 18(2), 109-117.
- Geary, D. C. ve Whitworth, R. H. (1988). Is the factor structure of the WISC-R different for Anglo-and Mexican-American children? *Journal of Psychoeducational Assessment*, 6(3), 253-260.
- Hill, T. D., Reddon, J. R. ve Jackson, D. N. (1985). The factor structure of the Wechsler scales: A brief review. *Clinical Psychology Review*, 5(4), 287-306.
- Hodges, K. (1982). Factor structure of the WISC-R for a psychiatric sample. *Journal of consulting and clinical psychology*, 50(1), 141.
- Horn, J. L. (1989). Measurement of intellectual capabilities: A review of theory. K. S. McGrew, J. K. Werder ve R. W. Woodcock, (Ed.), *WJ-R technical manual içinde* (197-245). Chicago, IL: Riverside.
- Hu, L. ve Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Kaufman, A. S. (1975). Factor analysis of the WISC-R at age levels between 6.5 and 16.5. *Journal of Consulting and Clinical Psychology*, 43, 135-147.
- Kaufman, A. S. (1979). *Intelligent testing: WISC-R*. New York: Wiley-Interscience.
- Kelloway E. K. (1998). Assessing model fit. *Using Lisrel for Structural Equation Modeling: A researcher's guide* (3. baskı) içinde (23-40). USA: Sage Publications.
- Kiriş, N. ve Karakaş, S. (2004). Dikkat eksikliği hiperaktivite bozukluğunun zeka testlerinden ve ilgili diğer nöropsikolojik araçlardan yordanabilirliği. *Klinik Psikiyatri Dergisi*, 7(3), 139-152.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling (2. baskı)*. New York: Guilford Press.
- Lee, L. M. P. ve Lam, Y. R. (1988). Confirmatory factor analyses of the Wechsler Intelligence Scale for Children-Revised and the Hong Kong-Wechsler Intelligence Scale for

- Children. *Educational and psychological measurement*, 48(4), 895-903.
- McMahon, R. C. ve Kuncze, J. T. (1981). A comparison of the factor structure of the WISC and WISC-R in normal and exceptional groups. *Journal of Clinical Psychology*, 37(2), 408-410.
- Meydan, C. H. ve Şeşen, H. (2011). *Yapısal Eşitlik Modellemesi, AMOS uygulamaları*. Ankara: Detay Yayıncılık.
- Naglieri, J. A. (1981). Factor structure of the WISC-R for children identified as learning disabled. *Psychological Reports*, 49(3), 891-895.
- O'Grady, K. E. (1989). Factor structure of the WISC-R. *Multivariate Behavioral Research*, 24(2), 177-193.
- Petersen, C. R. ve Hart, D. H. (1979). Factor structure of the WISC-R for a clinic-referred population and specific subgroups. *Journal of Consulting and Clinical Psychology*, 47(3), 643.
- Prifitera, A., Saklofske, D. H. ve Weiss, L. G. (Ed.). (2005). *WISC-IV clinical use and interpretation: Scientist practitioner perspectives*. New York: Academic Press.
- Ramanaiah, N. V., O'Donnell, J. P. ve Ribich, F. (1976). Multiple-group factor analysis of the Wechsler Intelligence Scale for Children. *Journal of Clinical Psychology*, 32(4), 829-831.
- Savaşır, I. ve Şahin, N. (1995). *Wechsler Çocuklar İçin Zeka Ölçeği uygulama kitapçığı*. (I. Savaşır ve N. Şahin, Çev.). Ankara: Türk Psikologlar Derneği.
- Silverstein, A. B. (1982). Alternative multiple-group solutions for the WISC and the WISC-R. *Journal of Clinical Psychology*, 38(1), 166-168.
- Spearman, C. (1904). "General Intelligence" objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292.
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49-74.
- Şimşek, Ö. F. (2007). *Yapısal Eşitlik Modellemesine giriş, temel ilkeler ve LISREL uygulamaları*. Ankara: Ekinoks Yayınları.
- Tabachnick, B. G. ve Fidell, L. S. (2001). *Using multivariate statistics (4. baskı)*. Needham Heights, MA: Allyn & Bacon.
- Tingstrom, D. H. ve Pfeiffer, S. I. (1988). WISC-R factor structure in a referred pediatric population. *Journal of Clinical Psychology*, 44(5), 799-802.
- Uluç, S., Öktem, F., Erden, G., Gençöz, T. ve Sezgin, N. (2011). Wechsler Çocuklar için Zeka Ölçeği-IV: Klinik bağlamda zekanın değerlendirilmesinde Türkiye için yeni bir dönem. *Türk Psikoloji Yazıları*, 14(28), 49-57.
- Wechsler, D. (1974). *WISC-R manual for the Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.
- Witta, E. L. ve Keith, T. Z. (1994). Multi-Sample hierarchical confirmatory factor analysis of the WISC-R: An old problem revisited. *The Annual Meeting of the Mid-South Educational Research Association*, 9-11 Kasım, Nashville, TN.

Summary

Confirmatory Factor Analysis of Wechsler Intelligence Scale for Children-Revised: A Sample of Children with Normal Intellectual Development

Cihat Çelik
Ankara University

İbrahim Yiğit
Ankara University

Gülşen Erden
Ankara University

Wechsler Intelligence Scale for Children-Revised (WISC-R) has become one of the most commonly used intelligence scales in our country and abroad for clinical practices and researches due to the fact that it assesses a great deal of characteristics involved in cognitive structure (Erdoğan-Bakar et al., 2005; Evinç & Gençöz, 2007; Prifitera, Saklofske, & Weiss, 2005). WISC-R provides scores for Verbal Intelligence Quotient (VIQ), Performance Intelligence Quotient (PIQ) and Full Scale Intelligence Quotient (FSIQ). From this aspect, the original structure of WISC-R consists of two factors (verbal and performance) (Wechsler, 1974). However, Kaufman (1975) then asserted that WISC-R had a three-factor structure in his factor analysis study with norm group of WISC-R. These factors are Verbal Comprehension, Perceptual Organization and Freedom From Distractibility. Perceptual Organization factor is defined by Picture Completion, Object Assembly, Block Design, Picture Arrangement and Mazes subtests, while Verbal Comprehension factor is defined by Information, Vocabulary, Comprehension and Similarities subtests. Freedom From Distractibility which is a new factor is defined by Arithmetic, Digit Span and Coding subtests (Kaufman, 1975).

Three-factor model was observed in factor analysis studies with both normal and clinical samples; however, in some of these exploratory factor analysis (EFA) studies, factor loading of third factor was found to be lower than the other factors (Chan, 1984; Donders, 1993; Hill, Reddon, & Jakson, 1985; Hodges, 1982; Kaufman, 1979; McMahon & Kunce, 1981; Nagleri, 1981; Petersen & Hart, 1979; Tingstrom & Pfeiffer, 1988). In addition, both two and three factor structure were observed in the confirmatory factor analysis (CFA) studies of WISC-R (Anderson & Dixon, 1995; Geary & Whitworth, 1988; Lee & Lam, 1988; O'Grady, 1989; Ramanaiah, O'Donnell, & Ribich, 1976; Silverstein, 1982; Witta & Keith, 1994). Anderson and Dixon (1995), discovered

that although two-factor structure of WISC-R was valid, the best model explaining WISC-R's factor structure was Kaufman's three-factor model and the other models (one-factor: g model; two-factor: Horn's model; three-factor: Bannatyne and Guildford models) had weaker fit indices than Kaufman's model.

Although various studies about WISC-R were conducted since it was adapted in 1995 in our country, factor structure of the scale has not been revised and reinvestigated yet. In addition, it was observed that EFA was performed in recent studies with clinical or normal samples (Erdoğan-Bakar et al., 2005; Erdoğan-Bakar et al., 2011; Kiriş & Karakaş, 2004). For instance, Erdoğan-Bakar and colleagues (2005; 2011) found that in intellectually-normal sample, WISC-R had a three-factor structure (Verbal, Performance and Coding); however for the ADHD-diagnosed group, factor structure was dispersed.

As mentioned above, CFA of WISC-R has not been carried out in any study in our country. The aim of this study is to examine whether the factors in WISC-R's norm study exist in a different sample of children who are in the range of normal intellectual level. Another aim of this study is to compare WISC-R's one and two factor structures.

Method

Sample

The sample of the research consisted of 932 (502 female, 430 male) children between the ages of 6-16 ($M = 10.97$, $SD = 2.78$) who were tested within the scope of WISC-R Administration Training in Psychology Department of Ankara University and Turkish Psychological Association. Participants whose General IQ are out of the range of 80-120 ($M = 101.93$, $SD = 9.88$), which is approved as a normal level of intelligence, were excluded from the study.

Measure

Wechsler Intelligence Scale for Children - Revised [WISC-R (Wechsler, 1974)]. Wechsler Intelligence Scale for Children (WISC), developed by Wechsler in 1949 and revised in 1974 as WISC-R, is an individually administered intelligence test for children between the ages of 6-16. It was adapted to Turkish by Savaşır and Şahin (1995) and standardized with a sample of 1639 children. It 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) scales, providing scores for VIQ, PIQ and FSIQ, as well as standard scores for these subtests. The mean score for these intelligence quotients is 100 and standard deviation is 15.

Procedure

Before the children were included in the administration, permissions with informed consent were taken from families of the children. The administrations were carried out on the basis of 10 core subtests. In line with the criteria of research, the children, who had not been administered WISC-R for at least 6 months and who had not applied for a clinic or hospital due to psychological, emotional and neurological complaints were included in the study.

Statistical Analysis

For the purpose of this study, CFA was utilized to test the validity of WISC-R's factor structure by using AMOS 21 packaged software. Based on theoretical and practical studies in the literature, two model which are one-factor model (g factor) (Spearman, 1904) and two-factor model (VIQ and PIQ) (Wechsler, 1974) were tested. Kaufman's model could not be tested because of the insufficient number of children administered digit span subtest.

In CFA analyses, various fit indices are used to determine the eligibility of the obtained data to hypothetical model (Meydan & Şeşen, 2011). In this study, χ^2 (Chi-square), χ^2/df (chi-square/degrees of freedom ratio), GFI (Goodness of Fit Index), AGFI (Adjustment Goodness of Fit Index), CFI (Comparative Fit Index) and RMSEA (Root Mean Square Error of Approximation) fit indices were used. From this aspect, lower and statistically insignificant χ^2 values indicate good fit (Şimşek, 2007). However, χ^2 value is highly sensitive to sample size and when the sample size increases, χ^2 gets statistical significance (Byrne, 2001). Of the other fit indices, χ^2/df , lower than 5, is accepted as indicator of good fit (Kelloway, 1998). RMSEA value of .05 and below indicates the perfect fit between proposed model and data; the

values up to .08 indicates acceptable fit (Hu & Bentler, 1999; Sümer, 2000). For GFI, AGFI and CFI fit indices; the values of .90 and above are accepted as indicators of good fit (Byrne, 1998; Kline, 2005).

Results

The results of CFA for both models showed that chi-square value of each model [χ^2 (35, $N = 932$) = 266.33, $p < .05$ for one-factor model; χ^2 (34, $N = 932$) = 223.95, $p < .05$ for two-factor model] are high and significant and that the models don't fit in the data. For one-factor model, some of the other fit indices indicate no fit (CFI = .72), some indicate acceptable (RMSEA = .08) and good fit (GFI = .94, AGFI = .91). It was found that subtest factor loadings of one-factor model ranged between .17 and .57 (coding, information, respectively) and all subtests were statistically significant.

In accordance with proposed modification indices, firstly, errors of the subtests of vocabulary and comprehension were correlated. After this correlation, two models were compared with chi-square difference test (Tabachnick & Fidell, 2001). The result of the difference test indicated that the model fit got better [$\chi^2_{\text{difference}}$ (1, $N = 932$) = 80.83, $p < .05$]. Then, the model was retested. Similarly, the errors of subtests of object assembly and block design were also correlated and difference test indicated that model fit got better [$\chi^2_{\text{difference}}$ (1, $N = 932$) = 44.40, $p < .05$]. The last model which two errors were correlated was statistically better than the other models and model fit indices produced better results [χ^2 (33, $N = 932$) = 141.10, $\chi^2/df = 4.27$, GFI = .97, AGFI = .95, CFI = .86, RMSEA = .06].

For two-factor model, some of the other fit indices also indicate no fit (CFI = .77), some indicate acceptable (RMSEA = .07) and good fit (GFI = .95, AGFI = .92). In two-factor model, subtest factor loadings for VIQ ranged between .22 and .61 (comprehension, information, respectively); subtest factor loadings for PIQ ranged between .16 and .52 (coding, picture completion, respectively) and all subtests that loaded on both factors were statistically significant.

In accordance with modification indices proposed for two-factor model, errors of the subtests of vocabulary and comprehension were correlated. After this correlation, two models were compared with chi-square difference test (Tabachnick & Fidell, 2001). The result of the difference test showed that the model fit got better [$\chi^2_{\text{difference}}$ (1, $N = 932$) = 52.29, $p < .05$]. The model was retested after this error correlation. In addition, the errors of subtests of object assembly and block design were also correlated. Similarly, difference test showed that the model fit got better [$\chi^2_{\text{difference}}$ (1, $N = 932$) = 39.61, $p < .05$]. The last model in which two errors were correlated

was significantly better than the other models and model fit indices produced better results [$\chi^2(32, N = 932) = 132.05$, $\chi^2/df = 4.12$, GFI = .97, AGFI = .95, CFI = .88, RMSEA = .05].

As one-factor model compared with two-factor model in terms of the differences of χ^2 values, it was found that two-factor model was significantly better than one-factor model [$\chi^2_{\text{difference}}(1, N = 932) = 9.05$, $p < .05$].

Discussion

The results of CFA showed that fit indices of one-factor model was at an acceptable level; however, two-factor model had better fit indices than one-factor model. These findings suggest that WISC-R is adequate instrument for assessing the VIQ, PIQ and FSIQ.

The findings of this study are consistent with the results of other studies in the literature. In most of the studies using CFA, two-factor structure was found to be more valid or had better fit indices than one-factor model (Anderson & Dixon, 1995; Geary & Whitworth, 1988; Lee & Lam, 1988; O'Grady, 1989; Ramanaiah et al., 1976; Silverstein, 1982; Witta & Keith, 1994). In addition, the findings of this study are also consistent with the findings of other studies performed via EFA in our country, showing WISC-R's two-factor structure was found to be valid for the normal sample (Erdoğan-Bakar et al., 2005; Erdoğan-Bakar et al., 2011; Kiriş & Karakaş, 2004).

As the factor loadings of the subtests for one-factor model were investigated, the subtests with lowest factor loadings were coding, comprehension and object assembly, respectively. In addition, for two-factor model, it was found that information subtest had highest and comprehension subtest had lowest factor loadings within

subtests of VIQ and that picture completion subtest had highest and coding subtest had lowest factor loading within subtests of PIQ. These findings are consistent with the results of other studies in the literature (Kiriş & Karakaş, 2004; Erdoğan-Bakar et al., 2005; Erdoğan-Bakar et al., 2011). In this study, we found that coding subtest had lowest factor loading for both one-factor and two-factor model. This finding supports the hypothesis that coding subtest may be loaded on third factor named "freedom from distractibility" as proposed by Kaufman (Anderson & Dixon, 1995; Erdoğan-Bakar et al., 2005; Erdoğan-Bakar et al., 2011; Geary & Whitworth, 1988; Lee & Lam, 1988; Witta & Keith, 1994).

Although it was found that, the subtests significantly loaded on their own factors in both models, it was observed that some subtests (e.g., coding and comprehension) had lower factor loadings. Based on the findings of this study, it is suggested that these subtests should be interpreted carefully in intellectual assessment. Consequently, the findings of the study indicated that the use of scores provided by WISC-R's two-factor structure was more appropriate for assessing intellectual skills of normally-developed children in our country. The findings of this study report that WISC-R's two-factor structure should be conserved in administrations for clinical and educational fields. Also, it can be said that assessments without administering some core subtests of WISC-R won't give adequate information about intellectual development of children.

Recently, in our country, norm study of WISC-IV has just completed (Uluç, Öktem, Erden, Gençöz, & Sezgin, 2011); however, WISC-R will be in use for educational and clinical practices until WISC-IV is brought out. Consequently, it is suggested that the testing of WISC-R in clinical samples with CFA will make the use of the test in these samples more confident.