

# Flynn Etkisi'nin Türk Örnekleminde Değerlendirilmesi: WÇZÖ-R ve WÇZÖ-IV Zeka Bölümü (ZB) Puanlarının Karşılaştırılması

Sait Uluç  
Hacettepe Üniversitesi

Burcu Korkmaz  
Hacettepe Üniversitesi

Özge Şahin  
Hacettepe Üniversitesi

## Özet

Flynn Etkisi psikoloji alan yazınındaki en tartışmalı konular arasında yer almaktadır. Araştırmacıların büyük çoğunluğunun bu etkinin varlığı konusunda hemfikir olmasına karşın etkinin kaynağı konusundaki tartışmalar süre gitmektedir. Günümüze kadar çok sayıda ülkede Flynn Etkisi'nin varlığının ve şiddetinin belirlenmesine yönelik çalışmalar yapılmıştır. Bu çalışmada Flynn Etkisi Türkiye örnekleminde ilk kez değerlendirilmektedir. Çalışma kapsamında Türkiye normları 1984 yılında oluşturulan WÇZÖ-R ile Türkiye normları 2012 yılında oluşturulan WÇZÖ-IV'ün dönüştürülmüş puanları karşılaştırılmıştır. Çalışmanın katılımcı örneklemini, yaşları 6-16 arasında değişen 87 çocuk ve ergenden oluşmaktadır. Bu çalışma sonunda, WÇZÖ-R puanlarının WÇZÖ-IV puanlarından anlamlı düzeyde yüksek olduğu ve fark puanının Flynn düzeltmesiyle yordanabilir nitelikte olduğu görülmüştür. Sonuçlar Flynn'in Amerika ve Avrupa örneklemelerinde elde ettiği bulgularla örtüşmektedir.

**Anahtar kelimeler:** Zeka, Flynn Etkisi, WÇZÖ-IV

## Abstract

The Flynn Effect is one of the most controversial issues in the psychology literature. Although the majority of researchers agree about the presence of this effect, the arguments about the source of the effect still continue. Many studies have been conducted to examine the presence and severity of this effect, in many countries. The Flynn Effect is evaluated for the first time in Turkey by the present study. For this aim, the scores of WISC-R for which Turkish norms had been established in 1984 were compared to composite scores of WISC-IV for which Turkish norms were established in 2012. The participants of the study were composed of 87 children and adolescents, whose ages varied between 6 and 16. Results of the study revealed significantly higher scores for WISC-R compared to WISC-IV, and this difference could be predicted with the Flynn correction. Results are consistent with the results which were obtained from American and European samples.

**Key words:** Intelligence, Flynn Effect, WISC-IV

Tüm dünyada olduğu gibi ülkemizde de her yıl eğitsel, klinik ya da hukuksal değerlendirme amacıyla yüzlerce zeka testi uygulanmaktadır. Bu testlerden elde edilen sonuçlar özel eğitim raporlarının hazırlanması, eğitim ya da tedavi planlarının oluşturulması, velayet durumlarının karara bağlanması, kişinin askerliğe uygunluğunun belirlenmesi ya da ceza sorumluluğunun bulunup bulunmadığının belirlenmesi gibi birçok kararda önemli bir rol oynamaktadır. Bu tür kararlar kişinin yaşamının geri kalanı üzerinde, sıklıkla kalıcı etkilere yol açmaktadır. Bu nedenle, gelişmiş ülkelerde bilişsel yetilerin değerlendirilmesine ilişkin çalışmalara yoğun kaynak ayrılmaktadır. Bu ülkelerde zeka başta olmak üzere, ilgili yetilerin uygun biçimde ölçülebilmesi için birbirinin seçeneği olabilecek ya da eş zamanlı kullanılacak çok sayıda araç geliştirilmiştir ve geliştirilmeye devam etmektedir (Flanagan, Ortiz ve Alfonso, 2007). Daha da önemlisi zeka ve zekanın ölçümüyle ilişkili tartışılmalı konular bilim insanları tarafından kuramsal ve görgül çalışmalar aracılığıyla ele alınmakta ve bu konulara açıklık kazandırılmaya çalışılmaktadır.

Zeka alan yazınındaki en tartışılmalı konulardan biri, Flynn Etkisi olarak da adlandırılan, zeka testi puanlarında yıllar içinde gözlenen düzenli artıştır. Son 10 yıl içerisinde gerek laboratuvarlarımızda yaptığımız çalışmalar, gerekse hastane ve kurumlarda yoğun olarak test uygulaması yapan meslektaşlarımızla kurulan kişisel iletişim, benzer bir eğilimin Türkiye’de yaygın kullanım alanı bulan Wechsler Çocuklar için Zeka Testi’nin yenilenmiş formu (WÇZÖ-R) için de geçerli olabileceğine işaret etmiştir. Ne yazık ki ülkemizde bu varsayımın sınanmasına olanak sağlayacak biçimde zeka puanlarındaki değişimin ya da tutarlığın belirlenmesine yönelik yapılmış düzenli izlem ölçümleri bulunmamaktadır. Bu çalışmada Türkiye’ye yeni uyarlanan Wechsler Çocuklar için Zeka Testi’nin dördüncü sürümünün (WÇZÖ-IV) puanları WÇZÖ-R puanlarıyla karşılaştırılarak bir değerlendirilme yapılması hedeflenmiştir.

Gelişen toplumların zeka düzeyindeki değişimler, hemen her zaman araştırmacıların ilgi duyduğu ancak yoğun fikir ayrılıkları yaşadıkları bir konu olmuştur. Bu konudaki ilk geniş kapsamlı ve sistematik çalışma 1984 yılında James R. Flynn tarafından yapılmıştır. Flynn (1984) Amerikalıların zeka düzeyindeki varsayılan gelişmeyi sınamak için Stanford- Binet Testi ve Wechsler testlerinin 1932 ile 1978 yılları arasında yapılmış norm çalışmalarını da kapsayan bir dizi karşılaştırma yapmıştır. Flynn’a göre “... *Eğer Amerikalıların IQ düzeylerinde zaman içinde bir artış söz konusuysa... katılımcılar önceki test normlarını izleyen normlara göre daha kolay bulmalıdır... Bir diğer deyişle, aynı katılımcılar ilk norm çalışması 1932’de diğer norm çalışması 1947’de yapılmış iki test aldıklarında, ilk norm çalışmasındaki puanları her zaman ikinci norm çalışmasındaki puanlarından*

*yüksek olmalıdır.*” Flynn varsayımlarını test etmek için yaş aralığı 2 ile 48 arasında değişen, toplam 7500 katılımcıyı içeren 73 çalışmayı bir arada değerlendirmiştir. Böylece, 18 test kombinasyonu elde etmiş ve bunların 17’sinde anlamlı bir değişim olduğunu saptamıştır. Elde edilen artışın aralığı yılda .25 puan ile .44 puan arasında değişmektedir. Bu bulguya dayanarak IQ puanlarında 46 yıllık bir süreçte 15 puan, yaklaşık 1 standart sapma artış olduğu gösterilmiştir. Bu ölçüde bir değişim kritik kararların alınmasında temel alınan ölçüm araçları için göz ardı edilemeyecek kadar önemli görünmektedir.

İzleyen çalışmalar bu ilk bulguların farklı zeka ölçekleri ve farklı ülkeler için de geçerli olduğuna işaret etmiştir (Flynn, 1987). Bu çalışmaların en kapsamlılarından biri yine Flynn (1987) tarafından gerçekleştirilen uluslararası bir meta analiz çalışması olmuştur. Bu çalışma kapsamında 14 ülkeden (Hollanda, Belçika, Fransa, Norveç, Yeni Zelanda, Kanada, Almanya, İngiltere, Avustralya, Japonya gibi) çok sayıda bilim insanı, araştırma enstitüsü ve resmi kurumun, ağırlıklı olarak 1950’den sonra yapmış olduğu bağımsız çalışmalar bir araya getirilmiş ve karşılaştırılmıştır. Karşılaştırmada Wechsler testleri, Stanford-Binet Testi, Otis Testi, Jenkins Sözel Olmayan Test, Iowa Temel Beceriler Testi, Raven Matrisleri gibi çok sayıda sözel ve sözel olmayan teste yer verilmiştir. Elde edilen sonuçların 1982 Amerika bulgularıyla oldukça tutarlı olduğu görülmüştür. Genel zeka puanlarında uluslararası düzeyde de ortalama 15 puanlık (5 ile 25 puan aralığında) bir artış olduğu bulunmuştur. Elde edilen bulgular, yaş gruplarına ya da ülkelere göre IQ puanlarının yükselme hızında önemli farklılıklar olduğunu gösterse de anlamlı düzeyde bir yükselmenin neredeyse evrensel bir durum olduğuna işaret etmiştir.

Flynn elde ettiği bulgulara dayanarak bazı önemli saptamalar ve çıkarımlarda bulunmuştur. Flynn’ın önemli saptamalarından biri Akıcı Zeka (Gf) ölçümlerindeki artışın Kristalize Zeka (Gc) ölçümlerindeki artışın iki katından daha çok olduğudur. Sözcük Dağarcığı ya da Aritmetik gibi öğrenmeden etkilenmeye açık, özgün beceri kazanımını ve sembollerle çalışmayı gerektiren Gc testlerinde yıllar içinde özgün eğitimin yaygınlaşmasına ve kalitesinin artışına eşgüdüllü bir yükselme gözlenmesi beklendiği bir durumdur. Flynn’ın bulguları bu tür yumuşak bir yükselmenin varlığına işaret etmektedir. Ancak, özgün bilgi ve beceri gerektirmeyen, kültürel etkilerden arınık, sözel olmayan Gf testlerinde gözlenen yüksek eğimli artışlar dikkat çekicidir. Flynn’ın bulguları özellikle Raven Matrisleri’ndeki artışın 30 yıl içerisinde neredeyse iki standart sapma kadar olduğunu göstermiştir. WÇZÖ-R söz konusu olduğunda performans alt ölçeklerindeki puan artışının sözel alt ölçeklerin iki katından fazla olduğu görülmektedir. Tüm bu bulgular çerçevesinde Flynn, zeka puanlarındaki artışın öğrenme-

den bağımsız, hatta öğrenme tarafından ketlenen bir durum olduğunu savunmuştur. Flynn'a (1982; 1987; 2006) göre zeka testi puanları genel zeka faktörü olan "g" ile zayıf ya da orta dereceli nedensel ilişkilere sahip olan *problem çözme becerilerini* ölçmektedir. Zeka puanlarındaki bu düzenli artış, günümüz toplumunun problem çözmeye dayalı yaşam tarzının çevresel kaynaklı ancak genlere de aktarılan bir yansıması olarak açıklanmıştır (Flynn, 2006).

Günümüzde Flynn Etkisi'nin varlığı ilgili bilim çevreleri tarafından büyük ölçüde kabul görse de, bu etkinin kaynağı hakkında önemli görüş ayrılıkları bulunmaktadır (McGrew, 2010). Bazı araştırmacılar bu etkinin çevresel kaynaklardaki iyileşmeyle yakından ilişkili olduğunu savunmaktadır. Örneğin, Lynn'a (2009) göre beslenme kalitesinde 1930'dan günümüze kadar gözlenen iyileşme Flynn Etkisi'ni ortaya çıkaran temel etkidir. Bazı araştırmacılar ise sağlık sistemindeki iyileşmenin (antibiyotik kullanımının yaygınlaşması ya da yeni tedavi tekniklerinin geliştirilmesi gibi) zeka puanları üzerinde önemli bir etkiye sahip olduğunu öne sürmektedir (Colom, Andres-Pueyo ve Juan-Espinoso, 1998). Zajonc ve Mullally'e (1997) göre ise çiftlerin giderek daha az sayıda çocuğa sahip olma eğilimi her bir çocuk için ayrılan kaynak miktarını arttırmakta ve ZB puanlarında yükselmeye yol açmaktadır. Bir diğer açıklama da, resmi ve resmi olmayan yollarla sağlanan eğitimin yaygınlığı ve kalitesindeki artışın temel etken olduğu görüşüdür (Daley, Whaley, Sigman, Espinosa ve Neumann, 2003; Eppig, Fincher ve Thornhill, 2010). Buna göre okul müfredatlarının iyileşmesinin yanı sıra bilgisayar oyunları ya da televizyonun yaygınlaşması da bilişsel olarak gelişmeye yol açmaktadır. Brand Hipotezi'ne göre ise yakın geçmişte çocukların riskten kaçınma eğilimleri fazladır. Günümüz çocukları ise daha fazla risk alarak daha güç soruları yanıtlama eğilimine sahiptir. Bu sosyo-kültürel değişim Flynn Etkisi'nin gerekçesi olarak sunulmaktadır. Pratik etkisi ise bir diğer açıklamadır. Buna göre, günümüz çocukları daha sık zeka testi benzeri testlerle karşılaşmakta ve bunun bir sonucu olarak hem soru tiplerine tanıdık hale gelmekte hem de bu tür bilişsel sorulara yönelik daha gelişmiş yaklaşım biçimleri geliştirmektedir. Genetik bakış açısını benimseyen araştırmacıların ise ZB artışına çok daha temkinli açıklamalarla yaklaştıkları gözlenmektedir. Örneğin, Jensen'e (1998) göre zekanın genotipinde gerçek bir artış söz konusu değildir. Geçmişteki olumsuz çevre koşulları kişilerin genetik potansiyellerinin çok altında bir performans göstermelerine yol açmıştır. Çevre koşullarının iyileşmesi zekayı arttırmamakta aslında var olan genetik potansiyelin gerçek sınırlarına ulaşmasına izin vermektedir. Ancak, bu modellerden hiçbiri ilgili değişimi açıklamak için tek başına yeterli olamamaktadır (Woodly, 2011).

Flynn Etkisi'ni açıklamaya yönelik tartışmaların düğümlendiği temel nokta *IQ İkilemi* (IQ Paradox) olarak adlandırılan olgudur. Geleneksel olarak, zeka puanlarının genetik art alanın ve çevresel koşulların birbirinden bağımsız (korelasyon göstermeyen) etkilerinin bir toplamı olduğu kabul edilmektedir ( $P_{(Fenotip)} = G_{(Genotip)} + E_{(Çevre)}$ ) (Kempthorne, 1957). Kalıtım ise genotipteki değişimin fenotipteki değişime oranıyla, 0 ile 1 arasında değişen bir etken olarak hesaplanmaktadır.

$$(kalıtım) = \frac{G_{(değişim)}}{P_{(değişim)}}$$

Tek ve çift yumurta ikizlerine dayalı genetik çalışmalar aracılığıyla kalıtımın ( $h^2$ ) 1'e yaklaşan, .75'in üzerinde, yüksek bir orana sahip olduğu kestirilmiştir. Bu nedenle, zeka konusundaki kişilerarası farklılıkların tümüne yakınının kalıtım tarafından açıklanabileceği öngörülmüştür (Jensen, 1973; Heimstein ve Murray, 1994). Bu noktadaki ikilem, Flynn Etkisi'nde gözlenen hızlı ve büyük çaplı değişimin kalıtımın görece kararlı olma eğilimine ters düşmesinden kaynaklanmaktadır. IQ ikilemini çözmeye yönelik en kapsamlı açıklamalardan biri Flynn-Dikens Etkisi, diğer adıyla Katlamalı Sosyal Etki Varsayımı'dır (Dickens ve Flynn, 2001; 2002). Katlamalı Sosyal Etki Varsayımı'na göre zekanın genetik ve çevresel koşulların bağımsız etkilerinin toplamından oluştuğunu var saymak hatalı bir bakış açıdır. Her iki etkinin içinde de bir diğerine ait ortak bir varyans payı bulunmaktadır. Bir diğer ifadeyle, iyileşen çevresel koşullar belli bir oranda genetik alt yapıyı değiştirmektedir. Genetik alt yapının tüm toplumda aynı etki sonucu değişmesi kişinin içine katıldığı toplumsal yapıda bir iyileşmeye yol açmaktadır. Bu da çevresel koşulların daha radikal bir biçimde iyileşmesine yol açan bir döngü yaratmaktadır. Bu nedenle, çevrenin zeka üzerindeki etkisini doğrusal (lineer) değil, logaritmik (katlamalı) bir artış olarak ele almak gerekmektedir. İkilemi çözmek için önerilen genetik temelli bir diğer yaklaşım ise Mingroni Etkisi, diğer adıyla Heterosist Varsayımı'dır (Mingroni, 2004). Buna göre, kapalı toplumlarda homozigot oranı yükselmekte ve zeka gibi karakter özelliklerini baskılayan çekinik genlerin etkisi artmaktadır. Kapalı bir topluluk büyük çaplı bir toplumsal hareketlilik yaşadığında heterozigot oranı artmakta ve baskın genler çekinik genlerin baskılayıcı özelliklerini etkisiz hale getirmektedir. Mingroni'e göre (2007) 1950'den günümüze dünya çapındaki hızlı toplumsal hareketlilikler gen havuzlarındaki heterozigot oranını arttırmakta ve zeka gibi temel karakter özelliklerinde genetik temelli sıçramalara yol açmaktadır. Flynn Etkisi'nin kaynağı konusundaki görüş ayrılıkları günümüzde de sürmektedir. IQ ikilemi geleneksel çevre-kalıtım tartışmalarının güncel arenası olmaya devam etmektedir.

Bu çalışmada 1984 normlarıyla kullanılan WÇZÖ-R ve 2012 normları oluşturulmuş olan WÇZÖ-IV testleri kullanılarak Flynn Etkisi'nin Türk örneklemindeki geçerliliği üzerinde çalışılması amaçlanmıştır. İki araç önemli yapısal farklılıklar göstermesine karşın testin genelinde %70'e yakın bir ortaklık taşımaktadırlar. Ek olarak, standart puan ve birleşik puan hesaplama yaklaşımı açısından matematiksel bağlamda ortak bir temele sahiptir. Bu nedenle, bu tür bir karşılaştırma mümkün olabilmektedir. Bu çalışma Flynn Etkisi'nin Türkiye örnekleminde ilk değerlendirilişi olması nedeniyle oldukça özgün olabileceği düşünülmektedir. Ek olarak, yakın zamanda kullanıma girecek olan WÇZÖ-IV ile uzun süredir kullanımda olan WÇZÖ-R puanları arasındaki ilişkinin niteliği açısından uygulama alanına yönelik önemli bilgiler de sağlayabileceği düşünülmektedir.

### Yöntem

#### Örneklem

Çalışmanın örneklemini, Ankara ilinde yaşayan ve yaşları 6-16 arasında değişen 87 katılımcıdan oluşmaktadır. Katılımcılara ulaşmak için Hacettepe Üniversitesi'nde duyuru yapılmış ve ilgili yaş aralığında çocuğu ya da yakını olan çalışanlar çağrılmıştır. Gönüllüler değerlendirilirken çocuğun daha önce bir psikiyatrik tanı almış olması, önceden zeka testi uygulanmış olması ve bilişsel becerileri üzerinde etkili olabilecek süregelen bir hastalığının olması dışlama ölçütleri olarak kullanılmıştır. Katılımcılardan ve ailelerinden çalışma hakkında kısa bir bilgilendirme yapılmış ve yazılı onamları istenmiştir. Çalışmaya katılmayı kabul eden ve katılım koşullarını karşılayan 87 katılımcı araştırmaya alınmıştır. Bir katılımcı kişisel nedenlerden ötürü tekrar test uygulamasına katılamadığından analizler toplam 86 katılımcı üzerinden yürütülmüştür. Araştırmada yer alan katılımcıların 45'i kız (%51.7) iken 42'si (%48.3) erkek; ay olarak yaş ortalaması 128 ay ve standart sapması 30 ay (en düşük 72 ay en yüksek 193 ay) olarak belirlenmiştir. Katılımcıların babalarının yıl olarak toplam eğitim süreleri  $Ort. = 11.33$  ve  $S = 4.49$  (en düşük 5 yıl, en yüksek 23 yıl); annelerinin ise  $Ort. = 10.53$  ve  $S = 4.60$  (en düşük 5 yıl, en yüksek 18 yıl) olarak saptanmıştır.

#### Veri Toplama Araçları

**Wechsler Çocuklar için Zeka Ölçeği Yenilenmiş (WÇZÖ-R).** WÇZÖ-R, David Wechsler tarafından çocukların bilişsel yeteneklerini değerlendirmek için 1949 yılında ilk formu geliştirilen Wechsler Çocuklar için Zeka Ölçeği'nin, Wechsler tarafından 1974 yılında yeniden gözden geçirilmiş biçimidir (Wechsler, 1974). WÇZÖ-R, 6 yaş 0 ay ve 16 yaş 11 ay yaş aralığındaki çocuklara bireysel olarak uygulanmaktadır. Ölçek, Savaşır ve Şahin tarafından 1984 yılında Türkçeye uyar-

lanmıştır (Savaşır ve Şahin, 1984). WÇZÖ-R Testi'nin Türkçe biçimi 1995 yılında Türk Psikologlar Derneği tarafından kullanıma sunulmuştur. WÇZÖ-R, sözel ve performans olmak üzere 2 alt boyuttan oluşmakta ve 6'sı sözel, 6'sı performans olmak üzere toplam 12 alt testi içermektedir. Zeka Bölümü beş sözel ve 5 performans alt testin uygulanması ile hesaplanmaktadır. Ölçekten PPZB (Performans Puan Zeka Bölümü), SPZB (Sözel Puan Zeka Bölümü), TPZB (Tüm Puan Zeka Bölümü) olmak üzere üç tür dönüştürülmüş puan elde edilmektedir.

**Wechsler Çocuklar için Zeka Ölçeği-IV'üncü Sürüm (WÇZÖ-IV).** Wechsler Çocuklar için Zeka Ölçeği IV'üncü Sürüm (WÇZÖ-IV) 2003 yılında kapsamlı bir yapı değişikliği ve norm çalışmasının ardından alan kullanımına sunulmuştur (Wechsler, 2003). Ölçek Öktem ve arkadaşları tarafından 2011 yılında Türkçeye uyarlanmıştır (Öktem ve ark., 2011). WÇZÖ-IV 6 yaş 0 ay ile 16 yaş 11 ay yaş arasındaki çocukların bilişsel becerilerini ölçmek için kullanılan ve bireysel olarak uygulanan bir zeka testidir. Toplamda 10'u temel - Benzerlikler, Küplerle Desen, Sayı Dizisi, Şifre, Sözcük Dağarcığı, Resim Kavramları, Harf Rakam Dizisi, Mantık Yürütme Kareleri, Kavrama, Simge Arama - 5'i yedek -Resim Tamamlama, Çiz Çıkar, Genel Bilgi, Aritmetik, Sözcük Bulma- olmak üzere 15 alt test bulunmaktadır. WÇZÖ-IV'te yer alan bu alt testlerden 10 tanesi WÇZÖ-R ile ortak iken; Resim Kavramları, Harf Rakam Dizisi, Mantık Yürütme Kareleri, Kavrama, Çiz Çıkar, Sözcük Bulma, Simge Arama alt testleri WÇZÖ-IV'e özgü testlerdir. Genel Bilgi ve Aritmetik alt testleri WÇZÖ-R'de temel alt testler arasında yer alırken; okuldaki öğrenmelerden etkilenmeye açık olduğu ve bireyin genel bilişsel becerilerinden çok okul başarısıyla yüksek korelasyonu olduğu için WÇZÖ-IV'te yedek alt testler arasında yer almıştır (Uluç ve ark., 2011). WÇZÖ-IV'ten (1) Sözel Kavrama Dönüştürülmüş Puanı (SKDP), (2) Algısal Akıl Yürütme Dönüştürülmüş Puanı (AAYDP), (3): Çalışma Belleği Dönüştürülmüş Puanı (ÇBDP), (4) İşleme Hızı Dönüştürülmüş Puanı (İHDP) ve (5) WÇZÖ-IV Testi Tüm Ölçek Dönüştürülmüş Zeka Puanı olmak üzere 5 tür dönüştürülmüş puan elde edilmektedir. Alt testler ve Dönüştürülmüş puan gruplarına ilişkin kapsamlı bilgi Ek 1'de sunulmuştur.

#### İşlem

Testler katılımcılara Hacettepe Üniversitesi'nin ya da Orta Doğu Teknik Üniversitesi'nin Klinik Psikoloji Yüksek Lisans Programlarından herhangi birine devam eden üçü erkek ve on sekizi kadın toplam 21 psikolog tarafından uygulanmıştır. Uygulamacılar iki aylık bir süre içinde beşer saatlik sekiz oturumdan oluşan bir eğitim paketi almışlardır. Eğitim paketi WÇZÖ-R ve WÇZÖ-IV testlerine ilişkin kuramsal bilgi aktarımını, gözetim

altında (aynalı gözlem odalarında) uygulama ve süpervizyon etkinliklerini içermiştir.

Uygulamalar yapılmadan önce, katılımcıların ailelerinden onay alınmıştır. Uygulamalar denetim ve süpervizyon altında gerçekleştirilmiştir. Tüm katılımcılara ilk olarak WÇZÖ-R bataryası bireysel olarak uygulanmıştır. WÇZÖ-IV uygulamaları ilk uygulamadan 2 ay sonra bireysel olarak gerçekleştirilmiştir. WÇZÖ-R uygulamaları ortalama 1.5 saat, WÇZÖ-IV uygulamaları ortalama 2 saat sürmüştür

### Bulgular

Bu çalışmanın temel amacını yöntem bölümünde tanımlanmış olan Türk çocuk-ergen örnekleminde üzerinden Flynn Etkisi'nin araştırılması oluşturmaktadır. Öncelikli olarak ölçeklerin dönüştürülmüş puanlarına ait psikometrik parametreler Tablo 1'de aktarılmıştır.

Ön psikometrik değerlendirmeler WÇZÖ-R puanları ve WÇZÖ-IV puanları arasında önemli farklar olduğuna işaret etmiştir. Çalışma örnekleminde WÇZÖ-R tüm test ZB puanlarına ait ortalamanın (112 puan) kendi norm grubunun birinci standart sapmasının pozitif ucuna (115 puan) yakın olduğu görülmektedir. Öte yandan, WÇZÖ-IV ait ortalama (102 puan) incelendiğinde kendi norm grubunun güven aralığı içinde kaldığı görülmektedir. Merkezi değerler incelendiğinde WÇZÖ-R için ortalama (112 puan), medyan (115 puan) ve mod (118 puan) değerlerinin giderek artan bir sıra içinde dizildiği göze çarpmaktadır. Buna dayanarak örneklem dağılımının pozitif yönde yatık olduğu ön görülebilir. WÇZÖ-IV'te ise merkezi değerlerin (102 puan, 104 puan, 102 puan) görece olarak eşit olduğu görülmektedir. Bu bulgu WÇZÖ-IV örnekleminin standart bir çan eğrisine yakın

olduğuna işaret etmektedir. Benzer örüntülerin diğer dönüştürülmüş puan grupları içinde geçerli olduğu görülmektedir.

### WÇZÖ-R ve WÇZÖ-IV Puanları Arasındaki İlişkiler

Bu bölümde WÇZÖ-R ve WÇZÖ-IV dönüştürülmüş puanları arasındaki ilişkiler Pearson Momentler Toplamı Katsayısı kullanılarak hesaplanmıştır. İlgili analizlerden elde edilen bulgular Tablo 2'de sunulmuştur. Bulgular her iki testten elde edilen Tüm Ölçek Dönüştürülmüş Zeka Puanları arasında oldukça yüksek bir ilişki ( $r = .81$ ) olduğuna işaret etmiştir. Tüm Ölçek Dönüştürülmüş Zeka Puanları arasında anlamlı bir fark bulunup bulunmadığı eşlenmiş gruplar için t-test kullanılarak hesaplanmıştır. Sonuçlar, WÇZÖ-R ( $Ort. = 112.4, S = 1.2$ ) puanlarının WÇZÖ-IV ( $Ort. = 102.4, S = 14.9$ ) puanlarından anlamlı derecede daha yüksek olduğuna işaret etmiştir ( $t = 10.12, S = .85, p < .000$ ). Buna dayanarak, her iki aracın ölçmeyi hedefledikleri genel zeka faktörü açısından pozitif yönde oldukça güçlü bir ilişkiye sahip olmasına karşın norm gruplarına dayalı sayısal göstergeler açısından anlamlı bir farklılık taşıdıkları söylenebilir.

WÇZÖ-R ile WÇZÖ-IV arasındaki farklar sadece norm grubunun yenilenmiş olmasıyla sınırlı değildir. WÇZÖ-IV önemli yapısal değişikliklerde taşımaktadır. Bu nedenle, Tüm Test Zeka Puanı dışındaki birleşik puanların doğrudan karşılaştırılması tümiyle güvenli bir değerlendirme yaklaşımı değildir. Ancak, alan yazın SZBP puanları ile SKDP puanları arasındaki ve PPZB puanları ile AAYDP puanları arasındaki ikili karşılaştırmaların mantıksal olarak kabul edilebilir olduğuna işaret etmektedir (Flanagan ve Kaufman, 2009). Buna dayanarak, Tablo 2 incelendiğinde SKDP ile SPZB aralarında ( $r = .78$ ) ve AAYDP ile PPZB arasında ( $r = .67$ ) güç-

**Tablo 1.** WÇZÖ-R ve WÇZÖ-IV Dönüştürülmüş Puanlarının Psikometrik Özellikleri

	SZBP	PPZB	SKDP	AAYDP	ÇBDP	İHDP	WÇZÖ-R	WÇZÖ-IV
Ort.	110	112	104	100	100	103	112	102
Medyan	114	113	103	100	100	106	115	104
Mod	114	123	94	106	106	109	118	102
S	16	16	17	13	15	16	15	15
Varyans	265	244	285	165	214	242	230	221
Ranj	85	75	82	66	67	76	77	80
Max.	147	147	143	129	132	141	140	134
Min.	62	72	61	63	65	65	63	54

Not. SKDP: Sözel Kavrama Dönüştürülmüş Puanı; AAYDP: Algısal Akıl Yürütme Dönüştürülmüş Puanı; ÇBDP: Çalışma Belleği Dönüştürülmüş Puanı; İHDP: İşleme Hızı Dönüştürülmüş Puanı; PPZB: WISC-R Performans Puan Zeka Bölümü; SZBP: WISC-R Sözel Puan Zeka Bölümü; WÇZÖ-R: WÇZÖ-R Tüm Puan Zeka Bölümü; WÇZÖ-IV: WÇZÖ-IV Testi Tüm Ölçek Dönüştürülmüş Zeka Puanı

**Tablo 2.** WÇZÖ-R ve WÇZÖ-IV Dönüştürülmüş Puanları Arasındaki İlişkiler

	SKDP	AAYDP	ÇBDP	İHDP	PPZB	SZBP	WÇZÖ-R
AAYDP	.480**	-	-	-	-	-	-
ÇBDP	.476**	.553**	-	-	-	-	-
İHDP	.460**	.416**	.197	-	-	-	-
PPZB	.335**	.670**	.333**	.490**	-	-	-
SZBP	.783**	.515**	.526**	.554**	.629**	-	-
WISC-R	.651**	.682**	.500**	.604**	.887**	.905**	-
WÇZÖ-IV	.854**	.795**	.702**	.665**	.603**	.803**	.813**

\*\* $p < .00$

Not. SKDP: Sözel Kavrama Dönüştürülmüş Puanı; AAYDP: Algısal Akıl Yürütme Dönüştürülmüş Puanı; ÇBDP: Çalışma Belleği Dönüştürülmüş Puanı; İHDP: İşleme Hızı Dönüştürülmüş Puanı; PPZB: WISC-R Performans Puan Zeka Bölümü; SZBP: WISC-R Sözel Puan Zeka Bölümü; WÇZÖ-R: WÇZÖ-R Tüm Puan Zeka Bölümü; WÇZÖ-IV: WÇZÖ-IV Testi Tüm Ölçek Dönüştürülmüş Zeka Puanı

lü ilişkiler olduğu görülmektedir. İlgili dönüştürülmüş puan grupları arasındaki farklar yine eşlenmiş gruplar için t-test kullanılarak hesaplanmıştır. Analiz sonuçları SPZB'nin ( $Ort. = 110.4, S = 16.3$ ) SKD puanlarından ( $Ort. = 103.9, S = 16.9$ ) anlamlı düzeyde yüksek olduğuna işaret etmiştir ( $t = 5.44, S = .85, p < .000$ ). Benzer bir biçimde, PPZB'nin de ( $Ort. = 112.4, S = 15.6$ ) AAYD puanlarından ( $Ort. = 100.09; S = 12.8$ ) anlamlı düzeyde yüksek olduğu görülmektedir ( $t = 9.70, S = .85, p < .000$ ). Bu bulgular çerçevesinde diğer dönüştürülmüş puanlarda da Tüm Test Zeka Puanına benzer bir örtüntü ortaya çıktığı söylenebilir. Ek olarak, PPZB ile AAYDP arasındaki puan farkının (yaklaşık 12 puan) SPZB ile SKDP arasındaki puan farkının (yaklaşık 6 puan) iki katı olduğu görülmektedir.

### Flynn Etkisi'nin Değerlendirilmesi

Giriş bölümünde de aktarıldığı gibi Flynn'in çalışmalarını zeka puanlarının yılda yaklaşık .33 puanlık bir artış gösterdiğine işaret etmiştir. WÇZÖ-R Testi'nin Türkçe biçimi 1995 yılında Türk Psikologlar Derneği tarafından kullanıma sunulmuştur. Ancak, aracın uyarlama çalışmalarının 1978'e kadar uzandığı ve normlarının 1984 yılında toplandığı bilinmektedir. Bu çalışmada Flynn Etkisi'nin Türk örnekleminde de benzer biçimde ortaya çıkıp çıkmadığını belirleyebilmek amacıyla WÇZÖ-R'nin standardize edildiği 1984 yılından günümüze kadar geçen 29 yıllık süre .33'lük sabitle çarpılarak Flynn Etkisi'nin sayısal değeri 9.57 olarak hesaplanmıştır. Flynn düzeltmeli WÇZÖ-R Tüm Test Zeka Puan dağılımı, katılımcıların WÇZÖ-R puanlarından 9.57 puan çıkarılarak hesaplanmış ve grup ortalamaları Tablo 3'te sunulmuştur.

Tablo 3 incelendiğinde 1984 normlarına göre hesaplanan WÇZÖ-R puanlarına Flynn düzeltmesi yapıldığında elde edilen grup ortalamasının WÇZÖ-IV'ün güncel normlarına dayalı grup ortalamasına oldukça yakın olduğu görülmektedir. Analizler WÇZÖ-IV ve Flynn düzeltmeli WÇZÖ-R puanları arasında istatistiksel olarak da anlamlı bir fark bulunmadığını göstermiştir ( $t = 1.14, S = .85, p = .260$ ).

**Tablo 3.** Flynn Düzeltmesi

WÇZÖ-R 1984 normu	WÇZÖ-R 1984 normu (Flynn düzeltmeli)	WÇZÖ-IV 2011 normu
112.4	102.8	102.4

Not. WÇZÖ-R: WÇZÖ-R Tüm Puan Zeka Bölümü; WÇZÖ-IV: WÇZÖ-IV Testi Tüm Ölçek Dönüştürülmüş Zeka Puanı

### Tartışma

Araştırma sonucu elde edilen bulgular genel olarak değerlendirildiğinde, Flynn'in tüm test zeka puanları için diğer ülkelerde belirlemiş olduğuna benzer bir artış eğiliminin Türk örnekleminde de ortaya çıktığı görülmektedir. Çalışma örnekleminde bu artışın 29 yıllık zaman dilimi için yaklaşık 10 puan olduğu belirlenmiştir. Flynn'in Amerika örnekleminde için 46 yıllık zaman diliminde yaklaşık 15 puanlık bir artış belirlemiş olduğu

göz önünde bulundurulduğunda, bu çalışma kapsamında gözlenen artışın miktarının da alan yazınla tutarlı olduğu göze çarpmaktadır.

Bir diğer önemli bulgu da, kristalize zeka ve akıcı zeka puanlarındaki yükselme eğilimindeki farkın bu çalışmada da kendini göstermiş olmasıdır. Flynn (2006) yaptığı karşılaştırmalarda akıcı zekanın, beklenenin tersine, kristalize zekadan daha yüksek bir artış oranına sahip olduğu sonucuna ulaşmıştır. Wechsler testleri ile yaptığı karşılaştırmalar sonucu akıcı zekayı temsil eden PPZB'nin kristalize zekayı temsil eden SPZB'den iki kat daha fazla artış gösterdiği hesaplanmıştır. Bu çalışmada da hesaplamalar akıcı zeka puanlarındaki artışın kristalize zekanın tam iki katı olduğuna işaret etmektedir. Bu durum WÇZÖ-IV'ün yapısında ortaya çıkan geniş çaplı değişikliklerle açıklanabilir. Ancak, bu ölçüde benzer oranlara ulaşılmış olması, testler arasındaki yapısal farklılıklardan çok akıcı-kristalize zeka arasındaki artış oranlarındaki farklılığın evrensel bir nitelik taşıdığı biçiminde yorumlanmasına açıktır. Bu varsayımlardan hangisinin daha güvenilir olduğu WÇZÖ-IV'ün ilerleyen yıllarda yapılacak norm yenilemelerini izleyen karşılaştırmalarla açıklığa kavuşabilecektir.

Bu çalışma, bilindiği kadarıyla, Flynn Etkisi'nin Türk örnekleminde ilk ele alınışıdır. Elde edilen sonuçlar tartışmaya yer bırakmayacak biçimde Flynn'ın sonuçlarını destekler nitelikte olmuştur. Öte yandan, Flynn Etkisi'nin kaynağı konusunda süregiden tartışmalara bu çalışma çerçevesinde katkı sağlayabilmek oldukça güçtür. Yine de elde edilen sonuçlar sınırlı bir spekülasyona olanak tanıyabilmektedir. Flynn Etkisi'nin heterosist açıklaması Amerika ve Avrupa tarihiyle oldukça tutarlı görünmektedir. I. ve II. Dünya savaşları büyük çaplı hareketliliklere yol açmış, farklı uluslar birbirleriyle iç içe geçmiştir. Geriye dönük olarak bakıldığında bu süreç heterozigot oranını artırma potansiyeline sahiptir. Öte yandan, aynı tarihsel dilimde benzer bir nüfus hareketliliğinin Türkiye coğrafyasında yaşandığını söylemek oldukça güçtür. Türkiye daha çok iç göç ve hızlı kentleşme olgularıyla yüzleşmek durumunda kalmıştır. Pek çok açıdan dini ve etnik gruplar kapalılığını sürdürmüştür. Buna rağmen bu araştırmanın bulguları Avrupa ve Amerika'daki değişimin aynı yön ve şiddette Türkiye'de de ortaya çıktığını göstermektedir. Bu nedenle, heterosist açıklamasının farklı coğrafyalarda tutarlılığını koruyamayan retrospektif bir yaklaşım olduğu sonucuna varılması mümkün gözükmemektedir. Öte yandan, hızlı teknolojik gelişim, sağlık ve eğitim sistemlerindeki gelişmeler ve medyanın hem Avrupa ve Amerika hem de Türkiye için ortak etki kaynakları olduğu açıktır. Örneğin son 20 yıl içinde bilgisayar teknolojisi inanılmaz ölçüde gelişim göstermiştir. Üst düzey görsel-motor koordinasyon, gör-

sel bellek, problem çözme becerileri gerektiren bilgi işlemciler herkesin cebine girebilecek kadar küçülmüştür. Bilgi teknolojisindeki bu değişim dünyanın pek çok noktasında hemen hemen aynı hızda ortaya çıkmaktadır. Bu açıdan bakıldığında Katlamalı Sosyal Etki Varsayımı'nın zeka düzeyindeki artışa yönelik çok daha akla yakın bir açıklama sunduğu kabul edilebilir.

Kaynağı ne olursa olsun Flynn Etkisi'nin varlığını kabul etmek bizi kaçınılmaz olarak önemli bir sorunla yüzleşmek zorunda bırakmaktadır: *Geliştirdiğimiz ya da uyarladığımız ölçekler sonsuza kadar güvenli ve geçerli olarak kullanım potansiyeline sahip değildir.* Bu sorunu aşmanın yollarından biri Flynn düzeltmesine başvurulmasıdır. Amerika'da hukuk sistemi için, özellikle idam cezalarının infazı konusunda, Flynn düzeltmesinin kullanımı önemli tartışmalara yol açmaktadır (örn., Cunningham ve Tasse', 2010; Hagan, Drogin ve Guilmette, 2008; Duvall ve Morris, 2006). Bilim çevreleriyle paylaşılan sonuçlar, bazı davalarda maktulün ceza sorumluluğu konusunda karar verilirken Flynn düzeltmesinin temel alındığına işaret etmektedir. Ancak, birçok davada da Flynn düzeltmesinin göz ardı edildiği görülmektedir. Kişisel görüşümüzü mecbur kalmadığı sürece zeka testi sonuçlarının hesaplanması konusunda Flynn düzeltmesine başvurmanın uygun bir yol olmadığı yönündedir. Bunun en önemli gerekçesi, Flynn Etkisi'nin kaynağı konusunda tam bir açıklamanın bulunmamasıdır. Kaynağını bilmediğimiz bir etkinin ne zaman, hangi yönde ve ne hızla hareket edeceğini yordamak imkânsız görünmektedir. Bu nedenle, Flynn düzeltmesini bir sabite bağlamanın ancak bilimsel çalışmalar açısından güvenilir bir yöntem olduğu düşünülmektedir. Bir diğer çözüm de, belli aralıklarla zeka ölçeklerinin yeniden yapılandırılmasına ya da en azından kapsamlı bir norm yenilemesine gidilmesidir. On yılda bir yapılacak bu tür yenilemeler ilgili ölçeklerin geçerlik ve güvenilirliklerini garanti altına almanın iyi bir yolu olarak görülmektedir. Bu noktadaki en önemli sorun ise Türkiye gibi zeka testi üreticisi olmayan ülkelerin yaşadığı dış bağımlılıktır. Günümüzde Kaufman, Wechsler ya da Stanford Binet gibi önemli markaların ürettiği testler birkaç uluslararası firmanın kontrolü altındadır. Ne yazık ki, fiyat politikaları ve telif yaklaşımları şirket politikaları temel alınarak belirlenmektedir. Örneğin bir ölçüm aracının ilgili ülke için ne ölçüde gerekli olduğundan çok ülkedeki satış miktarının ne olacağı telif üzerinde daha etkili olabilmektedir. İlk uyarılma ve standardizasyon çalışmasını izleyen süreçte test normları üzerinde yenilemeye gitmek ya da ulusal ihtiyaçlar çerçevesinde yapı değişiklikleri yapmak yasal sorunlara yol açabilmektedir. Bu nedenle, köklü bir çözüm için ulusal ölçüm araçlarının geliştirilmesi zorunlu görünmektedir.

## Kaynaklar

- Colom, R., Andres-Pueyo, A. ve Juan-Espinosa, M. (1998). Generational IQ gains: Spanish data. *Personality and Individual Differences*, 25, 927-935. doi:10.1016/S0191-8869(98)00090-7.
- Cunningham, M. D. ve Tasse, M. J. (2010). Looking to science rather than convention in adjusting IQ scores when death is at issue. *Professional Psychology: Research and Practice*, 41(5), 413-419. doi: 10.1037/a0020226.
- Daley, T. C., Whaley, S. E., Sigman, M. D., Espinosa, M. P. ve Neumann, C. (2003). IQ on the rise: The Flynn effect in rural Kenyan children. *Psychological Science*, 14, 215-219. doi:10.1111/1467-9280.02434.
- Dickens, W. T. ve Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review*, 108, 346-369.
- Dickens, W. T. ve Flynn, J. R. (2002). The IQ paradox is still resolved: Reply to Loehlin (2002) and Rowe and Rodgers (2002). *Psychological Review*, 109, 764-771. doi: 10.1037//0033-295X.109.4.764.
- Duvall, J. C. ve Morris, R. J. (2006). Assessing mental retardation in death penalty cases: Critical issues for psychology and psychological practice. *Professional Psychology: Research and Practice*, 37, 658-665.
- Eppig, C., Fincher, C. L. ve Thornhill, R. (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society: Series B. Biological Sciences*, 277, 3801-3808.
- Flanagan D. P. ve Kaufman, A. S. (2009). *Essential of WISC-IV assessment (2. baskı)*. New York: Wiley.
- Flanagan, D. P., Ortiz, S. O. ve Alfonso, V. C. (2007). *Essentials of cross-battery assessment (2. baskı)*. New York: Wiley.
- Flynn, J.R. (1982) Lynn, the Japanese, and environmentalism. *Bulletin of the British Psychological Society* 35, 409-413
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29-51. doi:10.1037/0033-2909.95.1.29.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171-191. doi:10.1037/0033-2909.101.2.171.
- Flynn, J. R. (2006). IQ gains, WISC subtests, and fluid g: g-theory and the relevance of Spearman's hypothesis to race. G. R. Bock, J. A. Goode ve K. Webb, (Ed.), *The nature of intelligence* içinde (202-227). New York, NY: Wiley.
- Hagan, L. D., Drogin, E. Y. ve Guilmette, T. J. (2008). Adjusting IQ scores for the Flynn Effect: Consistent with the standard of practice? *Professional Psychology: Research and Practice*, 39(6), 619-625. doi: 10.1037/a0012693.
- Heimstein, R. J. ve Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Simon and Schuster.
- Jensen, A. R. (1973). *Educability and group differences*. New York: Harper and Row.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Lynn, R. (2009). What has caused the Flynn effect? Secular increases in the development quotients of infants. *Intelligence*, 37, 16-24. doi:10.1016/j.intell.2008.07.008.
- Kempthorne, O. (1957). *An introduction to genetic statistics*. Wiley, New York.
- Mingroni, M. A. (2004). The secular rise in IQ: Giving heterosis a closer look. *Intelligence*, 32, 65-83.
- Mingroni M. A. (2007). Resolving the IQ Paradox: Heterosis as a cause of the flynn effect and other trends. *Psychological Review*, 114(3), 806-829. doi: 10.1037/0033-295X.114.3.806.
- McGrew, K. S. (2010). The Flynn effect and its critics: Rusty linchpins and "Lookin' for g and Gf in some of the wrong places." *Journal of Psychoeducational Assessment*, 28, 448-468.
- Öktem, F., Gençöz, T., Erden, G., Sezgin, N. ve Uluç, S. (2011). *Wechsler Çocuklar İçin Zeka Ölçeği-IV Türkiye Norm Çalışması*, 109K533 nolu TÜBİTAK proje raporu. (Basılmamış rapor).
- Savaşır, I. ve Şahin, N. (1984). Wechsler çocuklar için zeka ölçeğinin Türk kültürüne uyarlanması ve standardizasyonu. *Doğa Bilim Dergisi*, 8(1), 91-97.
- 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, 49-57.
- Woodly, M. A. (2011). Heterosis doesn't cause the flynn effect: A critical examination of mingroni (2007). *Psychological Review*, 118(4), 689-693. doi: 10.1037/a0024759.
- Wechsler, D. (1949). *Wechsler Intelligence Scale for Children (WISC)*. New York: The Psychological Association.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised (WISC-R)*. New York: The Psychological Corporation.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth edition (WISC-IV)*. San Antonio, TX: Pearson.
- Zajonc, R. B. ve Mullally, P. R. (1997). Birth order: Reconciling conflicting effects. *American Psychologist*, 52, 685-699.



**EK1. Kelime İşleme Becerilerinin Değerlendirilmesinde Kullanılan Kelimeler**

<b>Küplerle Desen (Block Design)</b>	Temel Alt Test
<b>Küme</b>	Algısal Akıl Yürütme Kümesi
<b>Ölçüm Hedefi</b>	Soyut görsel uyarıcıyı çözümleyip birleştirebilme, sözel olmayan kavram oluşturma, görsel algılama ve örgütleme, görsel motor eşgüdüm, öğrenme ve görsel uyarıcıda şekil zemin ayırımını yapabilme
<b>Benzerlikler (Similarities)</b>	Temel Alt Test
<b>Küme</b>	Sözel Kavrama Kümesi
<b>Ölçüm Hedefi</b>	Sözel kavrama, işitsel kavrama, bellek, önemli ya da önemsiz görünüşler arasındaki farkı yakalayabilme ve sözel aktarım
<b>Sayı Dizisi (Digit Span)</b>	Temel Alt Test
<b>Küme</b>	Çalışma Belleği Kümesi
<b>Ölçüm Hedefi</b>	İşitsel kısa süreli bellek, sıralama becerisi, dikkati yoğunlaştırma
<b>Resim Kavramları (Picture Concepts)</b>	Temel Alt Test
<b>Küme</b>	Algısal Akıl Yürütme Kümesi
<b>Ölçüm Hedefi</b>	Soyut, kategorik yargılama becerisi
<b>Şifre (Coding)</b>	Temel Alt Test
<b>Küme</b>	İşlem Hızı Kümesi
<b>Ölçüm Hedefi</b>	İşlem hızı, kısa süreli bellek, öğrenme, görsel algı, görsel motor eşgüdüm, görsel tarama becerisi, bilişsel esnekliği, dikkat ve motivasyon
<b>Sözcük Dağarcığı (Vocabulary)</b>	Temel Alt Test
<b>Küme</b>	Sözel Kavrama Kümesi
<b>Ölçüm Hedefi</b>	Sözcük bilgisi, öğrenme, uzun süreli bellek, dil gelişiminin düzeyi, işitsel algılama, kavrama, sözel kavramsallaştırma, soyut düşünme ve sözel aktarım becerisi
<b>Harf-Rakam Dizisi (Letter-Number Sequencing)</b>	Temel Alt Test
<b>Küme</b>	Çalışma Belleği Kümesi
<b>Ölçüm Hedefi</b>	Sıralama, zihinsel düzenleme, dikkat, kısa süreli işitsel bellek, görsel uzamsal inceleme ve işlem hızı
<b>Mantık Yürütme Kareleri (Matrix Reasoning)</b>	Temel Alt Test
<b>Küme</b>	Algısal Akıl Yürütme Kümesi
<b>Ölçüm Hedefi</b>	Görsel bilgi işleme ve soyut yargılama becerileri
<b>Kavrama (Comprehension)</b>	Temel Alt Test
<b>Küme</b>	Sözel Kavrama Kümesi
<b>Ölçüm Hedefi</b>	Sözel yargılama, kavramsallaştırma, sözel kavrama ve aktarımı, geçmiş deneyimleri değerlendirme ve kullanma becerisi, pratik bilgiyi kullanma becerisi
<b>Simge Arama (Symbol Search)</b>	Temel Alt Test
<b>Küme</b>	İşlem Hızı Kümesi
<b>Ölçüm Hedefi</b>	İşlem hızı, kısa süreli görsel bellek, görsel-motor eşgüdüm, bilişsel esneklik, görsel ayırt etme ve dikkatini yoğunlaştırma, işitsel yargılama, algısal örgütleme ve planlama, öğrenme becerisi
<b>Resim Tamamlama (Picture Completion)</b>	Yedek Alt Test
<b>Küme</b>	Algısal Akıl Yürütme Kümesi
<b>Ölçüm Hedefi</b>	Görsel algılama, görsel örgütleme, dikkatini yoğunlaştırma, nesnelerin temel ayrıntılarının görsel tanınması
<b>Çiz Çıkar (Cancellation)</b>	Yedek Alt Test
<b>Küme</b>	İşlem Hızı Kümesi
<b>Ölçüm Hedefi</b>	İşleme hızını, görsel seçici algıyı, görsel uyanıklık

**EK1 (devam).** Kelime İşleme Becerilerinin Değerlendirilmesinde Kullanılan Kelimeler

<b>Genel Bilgi (Information)</b>	Yedek Alt Test
<b>Küme</b>	Sözel Kavrama Kümesi
<b>Ölçüm Hedefi</b>	Kristalize zeka, uzun süreli bellek, okul ve çevreden edindiği bilgileri aklına yerleştirip hatırlama
<b>Aritmetik (Arithmetic)</b>	Yedek Alt Test
<b>Küme</b>	Çalışma Belleği Kümesi
<b>Ölçüm Hedefi</b>	Zihinsel düzenleme, dikkatini yoğunlaştırma, dikkat, kısa ve uzun süreli bellek, sayıları değerlendirme becerisi ve zihinsel uyanklık
<b>Sözcük Bulma (Word Reasoning)</b>	Yedek Alt Test
<b>Küme</b>	Sözel Kavrama Kümesi
<b>Ölçüm Hedefi</b>	Sözel kavrama, analogik ve genel yargılama yeteneği, sözel soyutlama, bilgi kazanımı/ kullanımı, farklı tipteki bilgiyi birleştirip bütünleştirme becerisi

## Summary

# Reconsideration of Flynn Effect in a Turkish Sample: A Comparison between WISC-R and WISC-IV IQ Scores

Sait Uluç

Hacettepe University

Burcu Korkmaz

Hacettepe University

Özge Şahin

Hacettepe University

Hundreds of IQ tests concerning various fields are carried out in our country, as in many countries all over the world. Deductions and decisions based on the scores of these tests often have important effects on people's lives. In developed countries, a large number of tools have been and also continue being improved in order to appropriately measure the related competences, IQ having the priority (Flanagan, Ortiz, & Alfonso, 2007). One of the most controversial issues within the intelligence literature is the regular increase observed in the IQ test scores throughout the years, which is also called as Flynn Effect. There are many studies within the literature concerning the Flynn Effect. According to the results of Flynn's meta-analysis study (1987), which is one of the most inclusive studies, it is concluded that there is 15 scores increase on average in the overall IQ scores on an international level. Although the findings reached indicate that there are significant differences between the increasing speed of IQ scores according to the age groups or the countries, they point to the fact that there is a significant increase nearly on an international level.

Flynn made some notable observations and inferences based on the findings he had reached. One of Flynn's notable observations is that the increase in the Fluid Intelligence (Gf) measurements is more than twofold of the increase in the Crystallized Intelligence measurements. Within the framework of the findings of the research, Flynn argued that the increase in the intelligence scores is independent from the learning and that it is even inhibited by the learning. Consequently, that regular increase in the intelligence scores is explained as a reflection of the lifestyle of today's society based on problem solving, which is exogenous but also can be transferred to the genes (Flynn, 2006).

Today, although the Flynn Effect's presence is accepted by the related science community to a large extent, there is significant dissensus on the source of this effect (McGrew, 2010). Some researchers argue that this effect is closely related with the improvements in the

exogenous resources (Colom et al., 1998; Daley et al., 2003; Eppig, Fincher, & Thornhill, 2010; Lynn, 2009; Zajonc & Mullally, 1997). Others explain this effect on the IQ increase with a genetic point of view and handle the subject evenly (Jensen, 1998). However, none of those models can be sufficient on itself to explain the change in the question (Woodly, 2011).

The main point where the discussions trying to explain Flynn Effect knot is the fact called *IQ Paradox*. Conventionally, it is accepted that the intelligence scores are a total of effects of genetic background and exogenous conditions being independent from each other (Kempthorne, 1957). The paradox here arises from the fact that the rapid and large-scale change observed in the Flynn Effect contradicts heredity's tendency to be relatively stable. One of the most far-reaching explanations aimed at solving the IQ Paradox is the Flynn-Dickens Effect, Social Multiplier Effect Hypothesis in other words (Dickens & Flynn, 2001, 2002). According to the Social Multiplier Effect Assumption, it is a wrong point of view to assume that the intelligence is formed by total of the standalone effects of genetic and exogenous conditions. Another approach which is based on the genetic structure to solve the paradox is Mingroni Effect, Heterocyst Hypothesis in other words (Mingroni, 2004). According to this approach, homozygote rate increases in closed societies and the effect of recessive genes, which suppress the characteristics such as intelligence, grows high.

With this study, it is aimed to study on Flynn Effect's validity for Turkish sample benefiting from WISC-R and WISC-IV tests, the former being used with 1984 norms and the latter with the 2012 norms.

## Method

### Sample

**Practitioner Sample.** The practitioner sample of the study consists of 21 psychologists in total, three of which are men and eighteen of which are women who

continue their studies at Clinical Psychology Master Programs of Hacettepe University or Middle East Technical University.

**Participant Sample.** The participant sample of this study consists of 86 participants living in Ankara, whose ages range from 6 to 16. 45 (51.7%) of those participants are girls and 42 (48.3%) of them are boys. The average of age is 128 months on a monthly basis and standard deviation is 30 months. Total education period of their fathers on a yearly basis is detected as 11.33 on average and standard deviation is determined as 4.49; that of their mothers is assessed as 10.53 on average and standard deviation is detected as 4.60.

### Tools

**Wechsler Intelligence Scale for Children-Revised (WISC-R).** WISC-R is developed by David Wechsler to evaluate the cognitive skills of children whose ages range from 6 years and 0 months to 16 years and 11 months. This scale is adapted to Turkish by Savaşır and Şahin in 1984 (Savaşır & Şahin, 1984). WISC-R includes 12 subtests in total, six of which are verbal and six of which are performance-based. Three types of transformed scores are derived from this scale. Those types of transformed scores are Performance IQ Score (PIQS), Verbal IQ Score (VIQS), and Full Scale IQ Score (FSIQS).

**Wechsler Intelligence Scale for Children- Fourth Edition (WISC-IV).** Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV), was presented to the use of the field in 2003 and the scale was adapted to Turkish in 2011 by Öktem et al. WISC-IV is an intelligence test which is used to evaluate the cognitive skills of children whose ages range from 6 years and 0 months to 16 years and 11 months, consists of 15 subtests and is applied individually. 5 types of transformed scores are derived from WISC-IV. Those transformed scores are Verbal Comprehension Score (VCS), Perceptual Reasoning Score (PRS), Processing Speed Score (PSS), Working Memory Score (WMS) and Full Scale IQ Score (FSIQS).

### Process

Application of WISC-IV and WISC-R was carried out under control and supervision. Firstly, the WISC-R battery was applied to all the participants individually. WISC- IV applications were carried out 2 months later than the first application individually. Application of WISC-R lasted one and half an hour on average and application of WISC- IV lasted 2 hours on average.

### Results

Psychometric pre-assessments have pointed to the fact that there are significant differences between WISC-

R and WISC-IV scores. Within the study sample, it is observed that the average (112 scores) of all WISC-R test IQ scores is close to the positive pole (115) of the first standard deviation of its own norm group. On the other hand, when the average belonging to WISC-IV (102 scores) is examined, it is seen that the average stands within the confidence interval of its own norm group.

Findings indicate that there is a strong correlation ( $r = .81$ ) between the Full Scale Intelligence Scores derived from the WISC-R and WISC-IV. According to the t-test results, it is found out that WISC-R scores ( $M = 112.4$ ,  $SD = 15.2$ ) are significantly higher than those of WISC- IV ( $M = 102.4$ ,  $SD = 14.9$ ) ( $t = 10.12$ ,  $SD = .85$ ,  $p < .000$ ).

Looking into the results, we can reach the conclusion that there is a high correlation between VCS and VIQS ( $r = .78$ ); and PRS and PIQS ( $r = .67$ ). The results of t-test analysis pointed to the fact that VIQS ( $avg. = 110.4$ ,  $SD = 16.3$ ) are significantly higher than VCS ( $avg. = 103.9$ ,  $SD = 16.9$ ) ( $t = 5.44$ ,  $SD = .85$ ,  $p < .000$ ). Similarly, it is observed that PIQS ( $avg. = 112.4$ ,  $SD = 15.6$ ) are significantly higher than PRS ( $avg. = 100.09$ ,  $SD = 12.8$ ) ( $t = 9.70$ ,  $SD = .85$ ,  $p < .000$ ). In addition, it is concluded that the score gap (nearly 12 scores) between the PIQS and the PRS is twofold of the score gap (nearly 6 scores) between the VIQS and VCS.

Flynn's studies indicate that intelligence scores have increased nearly .33 annually. Within this study, the Flynn Effect's numerical value is calculated as 9.57 by multiplying the period of 29 years, which represent the period beginning with 1984, when the WISC-R was standardized, and continuing up to now, by the constant value of .33 in order to determine whether the Flynn Effect occurred similarly in the Turkish sample. All WISC-R test score distribution which is edited by Flynn is calculated by subtracting 9.57 from WISC-R scores of the participants. When Flynn's editing is applied to WISC-R scores which are calculated according to the 1984 norms, it is observed that the group average derived from this calculation is very close to the group average based on the WISC-IV's current norms. Analyses have indicated that there is not any significant difference between WISC IV and Flynn-edited WISC-R scores statistically ( $t = 1.14$ ,  $SD = .85$ ,  $p = .260$ ).

### Discussion

It is seen that a tendency for increase similar to those Flynn detected in the other countries for all IQ test scores show up in the Turkish sample as well. Within the study sample, this increase is identified as nearly 10 scores over a period of 29 years. It is evident that the amount of increase observed within this study and its proportion of  $\frac{1}{2}$  in fluid and crystallized IQ scores

are consistent with the literature. Although the Flynn Effect is consistent with the literature, it is agreed that as for the source of this effect, the Social Multiplier Effect Hypothesis is more appropriate than heterocyst approach for the Turkish sample.

Whatever the source of the Flynn Effect is, as the measurement tools employed will not be valid and reli-

able forever, we encounter a problem. One of the ways to overcome this problem is using Flynn's editing. However, as it is impossible to foresee in which way and how fast an effect whose source cannot be defined precisely, will come out, employing Flynn's editing is not considered as an appropriate way, either.