

Yeni Geliştirilen Hacettepe-Bilgisayar Temelli Zihin Kuramı (Hacettepe-BTZK) Bataryasının Güvenirlik ve Geçerlik Çalışmaları



©Müge KADEMLİ¹, ©Banu CANGÖZ TAVAT²

ÖZET

Amaç: Başkalarının inanç, arzu ve niyetlerinin kendimizinkilerden farklı olabileceğini anlama yeteneği olarak tanımlanan zihin kuramının (ZK) 90'lı yıllarda popüler bir konu haline gelmesiyle ZK'yi ölçen çok sayıda görev geliştirilmiştir. Bununla birlikte aynı yapıyı ölçen farklı görevlerin kullanılması, standart bir görev bataryasının bulunmaması ve kullanılan ZK görevlerinin geçerlik/güvenirlik değerlerinin yetersiz olması araştırma sonuçlarında tutarsızlıklara yol açmıştır. Bu çalışmayla üç boyutlu renkli animasyonların kullanıldığı, sosyal etkileşime dayalı ve bilişsel/duygusal boyutları olan HACETTEPE-Bilgisayar Temelli Zihin Kuramı (HACETTEPE-BTZK) Bataryası'nın geliştirilmesi ve kapsamlı geçerlik ve güvenilirlik çalışmalarının yapılması amaçlanmıştır.

Yöntem: Dördü ikinci derece yanlış-inanç (iki bilişsel/iki duygusal) ve dördü ironi (iki bilişsel/iki duygusal) olmak üzere toplam sekiz senaryodan oluşan bataryanın geçerlik ve güvenilirlik çalışmaları yaşları 18-36 arasında değişen 214 sağlıklı yetişkinle yürütülmüştür.

Bulgular: Yapı geçerliği doğrulayıcı faktör analiziyle değerlendirilmiş, elde edilen uyum indeksleri modelin çok iyi bir uyuma sahip olduğunu göstermiştir [$\chi^2(19, N=214)=26,14$, $p>0,05$, $\chi^2/sd=1,38$, RMSEA=0,042, SRMR=0,05, GFI=0,97, AGFI=0,95, CFI=0,98, TLI(NNFI)=0,97]. Ölçüt geçerliği için kullanılan Dokuz Eylül Zihin Teorisi Ölçeği'nden alınan puanlarla HACETTEPE-BTZK Bataryası puanları arasında pozitif ve anlamlı ilişki ($r=0,32$, $p<0,05$) bulunmuştur. Yargıclararası güvenilirlik ve iç tutarlılık katsayıları sırasıyla $r=0,94$ ve $r=0,72$ olarak bulunmuştur.

Sonuç: HACETTEPE-BTZK Bataryası'nın yapıyı detaylı olarak ölçen, ekolojik geçerliği yüksek, kültürümüze uygun ve psikometrik açıdan güçlü bir değerlendirme aracı olduğu gösterilmiştir.

Anahtar Sözcükler: Bilişsel bilimler, güvenilirlik ve geçerlik, nöropsikolojik test, zihin kuramı

ABSTRACT

Reliability and Validity of the Newly Developed Hacettepe-Computer Based Theory of Mind (Hacettepe-CBTOM) Battery

Objective: With the rise in popularity of the theory of mind (ToM), defined as the ability to understand that others' beliefs, desires, and intentions may differ from one's own, numerous tools have been developed since the 1990s. However, the use of disparate tasks to measure the same construct, the lack of a standardized task battery, and the inadequate validity/reliability of existing TOM measures have contributed to inconsistent research findings. This study developed the HACETTEPE-Computer Based Theory of Mind Battery (HACETTEPE CBTOM), which utilizes three-dimensional colored animations, focuses on social interactions, and integrates cognitive/affective dimensions. Comprehensive validity and reliability studies were conducted.

Method: The validity and reliability studies of the battery, which consists of eight scenarios (four second-order false belief tasks: two cognitive/two affective, and four irony tasks: two cognitive/two affective), were carried out with 214 healthy adults aged 18-36.

Results: Construct validity was evaluated through confirmatory factor analysis, and the fit indices indicated an excellent model fit [$\chi^2(19, N=214) = 26,14$, $p>0,05$, $\chi^2/df=1,38$, RMSEA=0,042, SRMR=0,05, GFI=0,97, AGFI=0,95, CFI=0,98, TLI(NNFI)=0,97]. For criterion validity, a positive and significant correlation was found between the scores of the HACETTEPE-CBTOM Battery and the Dokuz Eylül Theory of Mind Scale ($r=0,32$, $p<0,05$). The inter-rater reliability and internal consistency coefficients were $r=0,94$ and $r=0,72$, respectively.

Conclusion: The HACETTEPE-CBTOM Battery is a culturally appropriate, ecologically valid, and psychometrically robust tool for detailed assessment of ToM.

Keywords: Cognitive sciences, neuropsychologic test, reliability and validity, theory of mind

Atf için: Kademli M, Cangöz Tavat B. (2025) Yeni Geliştirilen Hacettepe-Bilgisayar Temelli Zihin Kuramı (Hacettepe-BTZK) Bataryasının Güvenirlik ve Geçerlik Çalışmaları. *Türk Psikiyatri Derg* 36:353–363. <https://doi.org/10.5080/u27454>

Geliş Tarihi: 07.01.2024, **Kabul Tarihi:** 01.11.2024, **Çevrim İçi Yayın Tarihi:** 24.06.2025

¹Dr. Öğr. Üyesi, Süleyman Demirel Üniv., Psikoloji Bl., Isparta; ²Prof., Hacettepe Üniv., Psikoloji Bl., Ankara, Türkiye

Müge Kademli, e-posta: mugeaslankara@sdu.edu.tr

GİRİŞ

Zihin kuramı (ZK), bireyin diğer insanların sahip olduğu bilgi, duygu, arzu, niyet ve isteklerin, kendi sahip olduklarından farklı olabileceğini anlaması olarak tanımlanmakta olup, sosyal etkileşiminin özünde yer alan temel bilişsel becerilerden biridir (Wellman 1990). Frith'e (1992) göre ise ZK, kendimizin ve başkalarının, içsel ya da zihinsel üst temsilidir. Bu bağlamda, genel olarak başkalarının zihinsel durumları ile ilgili çıkarımlarda bulunma becerisi olarak tanımlanan ZK, kişinin kendi psikiyatrik sorunları da dahil olmak üzere, zihnine ilişkin içgörü oluşturmada sağlayan temel mekanizmadır (Wiffen ve ark. 2013). Bu nedenle, Ruhsal Bozuklukların Tanısal ve Sayımsal El Kitabı Beşinci Baskısı'nda (DSM-5) (Amerikan Psikiyatri Birliği 2013) nörobilişsel bozukluktan etkilenebilecek temel bilişsel işlevlerden biri olarak tanımlanmıştır.

Bilişsel ve duygusal olarak sınıflandırılan ZK'yi ölçmek için birinci derece doğru ve yanlış-inanç, ikinci derece doğru ve yanlış-inanç, metafor kavrama, ironi kavrama, gaf kavrama, beyaz yalan kavrama, blöf kavrama gibi çeşitli görevler kullanılmaktadır (Bora 2009, Corcoran ve ark. 1995, Değirmencioglu 2008, German ve Hehman 2006, Gregory ve ark. 2002, Happé 1994, Happé ve ark. 1998, Rowe ve ark. 2001, Sprong ve ark. 2007, Stone ve ark. 1998).

Bilişsel ZK, başkalarının düşünce ve inançlarına, duygusal ZK ise başkalarının duygu ve hislerine ilişkin çıkarımda bulunma becerisi olarak tanımlanmaktadır. Bilişsel ZK ölçümü için tercih edilen en geçerli yöntem doğru ve yanlış-inanç görevleridir (Wimmer ve Perner 1983). Birinci derece doğru-inanç görevleri hem kahramanın hem de katılımcının gerçeklikle ilgili aynı doğru bilgiyi bildikleri durumda katılımcının kahramanın ne bildiğini anlamasını gerektiren durumları içerir. Buna karşılık, birinci derece yanlış-inanç görevleri, kahraman ile katılımcının gerçeklikle ilgili bilgilerinin örtüşmediği durumları içerir ve senaryodaki kişinin yanlış inancına atıfta bulunmayı gerektirir. Bu durumda, diğerlerinin yanlış inanca sahip olabileceğini anlayabilen bir katılımcı, o kişinin ne bildiğini buna göre öngörmektedir (Stone ve ark. 1998). İkinci derece doğru ve yanlış-inanç görevlerinde ise ikinci kişinin üçüncü bir kişinin inancıyla ilgili olarak ne bildiğine atıfta bulunulur. Kısaca birinci ve ikinci derece yanlış-inanç görevleri arasındaki tek fark senaryoya üçüncü bir kişinin katılımı ile görevin bilişsel olarak daha zor hale gelmesidir. Yanlış-inanç görevleri, kişinin kendi bilgisini görmezden gelerek, diğer kişinin kendisinininkinden farklı bir bilgiye sahip olabileceğini anlama becerisini değerlendirdiği için ZK açısından önemlidir.

Duygusal ZK ise sıklıkla karmaşık duygusal durumları temsil eden resimli (görsel) uyarıcıların ya da kahramanın duygularını betimleyen hikayelerin (sözel) kullanıldığı görevlerle değerlendirilmektedir (Baron-Cohen ve ark. 2001, Shamay-Tsoory ve ark. 2006, Shamay-Tsoory ve ark. 2007, Bottiroli ve ark. 2016). Öte yandan, sosyal gaf, ironi, metafor, blöf

ve beyaz yalan içeren dolaylı dilsel anlatım görevlerinde kişilerin kelimelerin altında yatan gerçek mesajı anlama becerisi değerlendirilir (Happé 1994, Brüne ve Brüne-Cohrs 2006, Sprong ve ark. 2007). Youmans'a (2004, aktaran Değirmencioglu 2008) göre bu görevler, düz anlamlı olmayan konuşmanın (nonliteral speech) yorumlanmasını gerektiren daha karmaşık ve incelikli ZK görevleridir. Harrington ve arkadaşlarına (2005) göre ise bu görevler, konuşmanın pragmatik anlayışını yansıtırken, konuşmacının neyi bildiğini, neye inandığını ya da neye niyetlendiğini anlamayı gerektirir (Baron-Cohen ve ark. 1999).

Bilişsel ve duygusal türleri olabilen dolaylı dilsel anlatım görevlerinden ironi, ikinci derece yanlış-inanç görevi ile ilişkilendirilmektedir (Herold ve ark. 2002, Shamay-Tsoory ve ark. 2005, Shamay-Tsoory ve Aharon-Peretz 2007). Genellikle bir kişi ya da durum hakkında örtük veya üstü kapalı bir biçimde eleştiride bulunmak veya olumsuzluk ifade etmek üzere yapılan ironiye çoğu kez küçümseme, hor görme ve kınama da eşlik eder. Ironi yapan kişi, bağlam dışında bir durum yaratır ve ironiye muhatap olan kişinin bu durumun farkına vararak ironinin altında yatan asıl anlamı (yani söylenenin tam tersinin kastedildiğini) kavradığını varsayar. Ironinin muhatabı, söyleneni doğru olarak kavradıysa ironiyi bir hata ya da yalan şeklinde yorumlamaz.

Literatürde kullanılmakta olan ZK görevlerinin psikometrik (özellikle yapı ve ölçüt geçerliği) açıdan yeterince kapsamlı ve yöntemsel açıdan yeterince kontrollü olarak incelenmemiş olması (Harrington ve ark. 2005, Sprong ve ark. 2007, Bora 2009, Bora ve ark. 2009), görünüş ve ekolojik (dış) geçerliklerine ilişkin kısıtlı sayıda araştırma bulunması (Corcoran ve Frith 2003) ve farklı ZK görevlerinden elde edilen bulguların birbirleriyle karşılaştırılabilir olmaması çelişkili bulgulara sebep olmakta ve ölçü araçlarının gücünü tartışılır hale getirmektedir. Bu yüzden literatürde üzerinde uzlaşma sağlanmış, objektif ve güvenilir bir ZK ölçümüne duyulan gereksinim sıklıkla dile getirilmektedir (Harrington ve ark. 2005).

Başlangıçta çocuklar için geliştirilen, sonrasında yapısında ve içeriğinde hiçbir değişiklik yapılmaksızın yetişkinler için de kullanılan ZK görevleri "kolay" olmaları sebebiyle tavan etkisine yol açmakta (Saltzman ve ark. 2000, Henry ve ark. 2013) ve bu sorunun daha zor ZK görevleri geliştirilerek çözülebileceği öngörülmektedir. Türkiye'de de daha çok çocuk örnekleme (Sarı 2011, Karakelle ve Ertuğrul 2012, Kaysılı 2013, Şahin ve ark. 2020, Tülü ve Ergül 2022) incelenmiş olmakla beraber ZK'nin yetişkin örnekleme (Küçük 2018, Ertuğrul Yaşar 2022) de incelenmesi gerektiği belirtilmektedir. Son yıllarda hem klinik hem de sağlıklı yetişkin örnekleme ZK ölçümüne olan ihtiyacın artmasına bağlı olarak Batı kültürü için geliştirilmiş ölçeklerin Türkçeye uyarlanmasında bir artış gözlenmektedir (Yıldırım ve ark. 2011, Tanrıverdi 2022, Törenli Kaya ve ark. 2023). Öte yandan kullanılmakta olan

ZK görevlerinin el çizimi resimlerden oluşan çizgi hikayeler olması ekolojik geçerliği düşürürken, sonuçların gerçek hayat ortamlarına genellenabilirliğini de güçleştirmektedir (Henry ve ark. 2013). Sözü edilen dış geçerlik sorununu çözdüğünü iddia eden Sosyal Biliş Film Testi (SBFT) (Movie for the Assessment of Social Cognition: MASC) (Dziobek ve ark. 2006) gibi uyarıcı olarak videoların kullanıldığı testlerde ise psikometrik değerlendirme, görev detayları ve çeşitliliği ile videoların sayısı açısından sınırlılıklar bulunmaktadır. Sosyal etkileşim içeren animasyonların olduğu 11 adet videodan oluşan Edinburgh Sosyal Biliş Testi'nde (Edinburgh Social Cognition Test: ESCoT) (Baksh ve ark. 2018) ise bilişsel ve duygusal ZK ayrımı yapılmış olsa da videolardaki bazı görsellerin Türk kültürüne uygun olmaması (direksiyonun sağda olması ve trafiğin soldan akması gibi) ve duyguların puanlanmasında değerlendirmenin testöre bırakılmış olmasının standart uygulamayı zorlaştırması eleştirilmektedir (Tanrıverdi 2022). Ayrıca örneklemin sadece 37 kişiden oluşması ve yapı geçerliğinin incelenmemiş olması ESCoT'un Türk örneklemdaki psikometrik değerlendirilmesinde gözlenen sınırlılıklarıdır.

Bu araştırmanın amacı: a) ZK görevlerindeki sorunların giderildiği, dış geçerliği yüksek, tavan etkisi göstermeyen, farklı ZK türlerini içeren, psikometrik ve yönetsel açıdan güçlü HACETTEPE-Bilgisayar Temelli Zihin Kuramı (HACETTEPE-BTZK) Bataryası'nı geliştirmek ve b) ZK'nin bilişsel ve duygusal boyutlarını deneysel olarak ölçmektir.

YÖNTEM

Uygulamalar için gerekli etik kurul onayı, Hacettepe Üniversitesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu'ndan 13.06.2017 tarihli ve 16969557-868 sayılı karar ile alınmıştır.

HACETTEPE-BTZK Bataryası Geliştirme Sürecinin Aşamaları

I. Aşama: İkinci Derece Yanlış-İnanç ve İroni Kavrama Senaryolarının Oluşturulması

Batarya kapsamında 14 bilişsel ikinci derece yanlış-inanç (İYİ_B), 14 duygusal ikinci derece yanlış-inanç (İYİ_D), 14 bilişsel ironi kavrama (İ_B) ve 14 duygusal ironi kavrama (İ_D) görevi olmak üzere 28 adet ikinci derece yanlış-inanç (İYİ) ve 28 adet ironi kavrama (İ) görevi geliştirilmiştir. Görev senaryoları, yedi kişilik uzman grubu (akademisyen psikologlar) tarafından oluşturulmuştur. Bu bağlamda, uzman grup görevleri 28'i bilişsel (İYİ ve İ senaryosu) ve 28'i duygusal (İYİ ve İ senaryosu) olacak şekilde ayırıştırarak toplam 56 senaryo üretmiştir. Ardından bu senaryolar, senaryolardaki kişi sayısı, kişilerin gerçekliği bilip bilmeme durumu açısından eşdeğerliğin sağlanması amacıyla bazı düzenlemelere tabi tutulmuştur (ayrıntılı bilgi için bkz. Aslankara 2019). Düzenlemeler sonrasında, 28 İYİ ve

28 İ senaryosunun ZK'yi ve türlerini (bilişsel/duygusal) temsil etme gücünü belirlemek için 5'li Likert tipi ölçek (Hiç Temsil Etmiyor, Temsil Etmiyor, Kararsızım, Temsil Ediyor, Tamamen Temsil Ediyor) kullanılarak hazırlanan değerlendirme formları toplam 126 üniversite öğrencisine ($\bar{x}=20,19$, $SS=3,73$) uygulanmıştır. Bilişsel ve duygusal senaryolar, aynı türden senaryo üç kere art arda gelmeyecek şekilde (bilişsel/bilişsel/bilişsel gibi) seçkisiz sırayla sunulmuştur. Ayrıca sistematik hatayı azaltmak amacıyla değerlendirme formlarında senaryoların farklı sıralarda sunulduğu dört farklı versiyon (A, B, C ve D formları) kullanılmıştır.

II. Aşama: Madde Güçlük Analizleri

İYİ ve İ senaryolarına ilişkin madde güçlük analizi sonucuna göre İYİ senaryolarından biri (Senaryo 6) "zor ve ayırt edici değil" kategorisinde olduğu için; diğeri (Senaryo 4) "zor fakat ayırt edici" kategorisinde olmasına rağmen 5'li Likert tipi ölçek ile değerlendirilen ikinci derece yanlış-inanç görevini temsil etme gücü ortalaması 3'ün altında olduğu ($\bar{x}=2,98$, $SS=1,25$) için elenmiştir. Eleme sonrası, toplam 26 adet İYİ senaryosundan "zor fakat ayırt edici" olarak yorumlanan, temsil etme gücü en yüksek 4 bilişsel ve 4 duygusal olmak üzere toplam 8 İYİ senaryosu ile madde güçlük analizi sonucunda tamamı "zor fakat ayırt edici" olarak yorumlan İ senaryolarından temsil etme gücü en yüksek 4 bilişsel ve 4 duygusal toplam 8 İ senaryosu olmak üzere toplam 16 ZK senaryosu seçilmiştir.

III. Aşama: Uzman Değerlendirmesi

Seçilen 16 adet ZK senaryosu, II. Aşamada kullanılan hakem değerlendirme formuyla birlikte 23 kişilik uzman grubunun (2 uzman psikiyatrist, 21 psikolog) değerlendirmesine sunulmuştur. Uzman önerileri sonucunda, 16 adet senaryo arasından, 4 adet İYİ ve 4 adet İ senaryosu olmak üzere toplam 8 adet temsil etme gücü en yüksek senaryo (2 bilişsel ve 2 duygusal olmak üzere 4 adet İYİ ve 2 bilişsel ve 2 duygusal olmak üzere 4 adet İ senaryosu) seçilmiştir. Seçilen İ senaryoları karakterlerin cinsiyeti, İYİ senaryoları ise hem karakterlerin cinsiyeti hem de senaryodaki "son gerçekliği" bilen kişi sayısı açısından dengelenmiştir. Örneğin; bilişsel senaryoların birinde sadece 1 kişi "son gerçekliği" biliyor iken, diğer bilişsel senaryoda 2 kişi "son gerçekliği" bilmektedir. Seçilen 8 senaryonun bilişsel/duygusal ve İYİ/İ görevlerini temsil etme gücü ortalama ve standart sapmaları Tablo 1'de verilmiştir.

Yukarıda sıralanan üç aşama sonucunda, toplam 56 adet senaryodan HACETTEPE-BTZK Bataryası'nı oluşturacak 8 adet ZK senaryosu belirlenmiştir. Bu 8 senaryo animasyon tipi renkli videolara dönüştürülerek özgün bir ZK Bataryası yazılım programı oluşturulmuştur.

IV. Aşama: ZK Senaryolarına Uygun Soruların Geliştirilmesi

Belirlenen 8 ZK senaryosu için standart ZK soruları geliştirilirken senaryolardaki her bir karakterin bakış açısına göre

Tablo 1. Seçilen 8 ZK Senaryosunun 5'li Likert Tipi Ölçekle Değerlendirilen Bilişsel/Duygusal ve İkinci Derece Yanlış-İnanç/İroni Kavrama Görevini Temsil Etme Gücü Ortalamaları ve Standart Sapmaları

Senaryo No	Görevin Niteliği	$\bar{x}$	SS	ZK Görevi	$\bar{x}$	SS
12	Duygusal	4,67	0,48	İkinci Derece Yanlış-İnanç	4,43	0,6
22	Duygusal	3,76	1,64	İkinci Derece Yanlış-İnanç	4,24	1
9	Bilişsel	3,71	1,71	İkinci Derece Yanlış-İnanç	4,29	0,85
11	Bilişsel	4,62	0,5	İkinci Derece Yanlış-İnanç	4,71	0,46
22	Duygusal	4,1	1,14	İroni	4,71	0,56
26	Duygusal	4,14	0,79	İroni	4,67	0,48
5	Bilişsel	4,43	0,6	İroni	4,67	0,48
25	Bilişsel	4,57	0,51	İroni	4,62	0,5

ZK: Zihin kuramı, SS: Standart sapma

zamansal öncelik ve sonralık gözetilerek İYİ senaryoları için olası bütün soru kombinasyonlarını içeren toplam 53 soru oluşturulmuştur. Ancak bataryanın uygulanacağı potansiyel örneklem (sağlıklı yetişkinler ve psikiyatrik/nörolojik hasta grupları) açısından soru sayısının fazla olacağı öngörülerek soru sayısı 14'e düşürülmüştür. İ senaryoları için ise ironinin anlaşılıp anlaşılmadığını belirlemeye yönelik sorular ile mantık ve gerçeklik sorularını içeren toplam 10 soru oluşturulmuştur (İYİ ve İ senaryo sorularını belirlerken kullanılan ölçütler hakkında detaylı bilgi için bkz. Aslankara 2019).

V. Aşama: ZK Senaryolarına Yönelik Sorulara İlişkin Uzman Değerlendirmesi

ZK konusunda uzman psikiyatristlerin önerileri doğrultusunda güncellenen sorular, psikoloji alanında öğrenim gören 10 doktora öğrencisi tarafından cevaplanarak her bir ZK senaryosu için ayrı birer cevap anahtarı oluşturulmuştur.

VI. Aşama: ZK Görevlerine İlişkin Cevap Anahtarı Havuzunda Yapılan Düzenlemeler

Oluşturulan cevap anahtarının puanlayıcılar açısından daha ayrıntılı, net ve anlaşılır olmasını sağlamak için cevap anahtarı şekli (dil, anlatım ve Türkçe imla kuralları) ve içerik (cümle ve kelime kısıtlaması, uygulamacılar için her bir senaryonun cevap anahtarına eklenen özel bilgi notları gibi) açısından güncellenmiştir (detaylı bilgi için bkz. Aslankara 2019).

VII. Aşama: Geçerlik Çalışması

Kapsam ve Görünüm Geçerliği

Araştırmacılar hariç 23 uzmanın yorum ve düzeltmeleriyle ZK senaryolarının kapsam ve görünüm geçerliği (content and face validity) değerlendirilmiştir (detaylı bilgi için bkz. III. Aşama: Uzman Değerlendirmesi).

Yapı Geçerliği

Literatür incelendiğinde, kullanımda olan birçok ZK görevinin psikometrik özelliklerinin (özellikle yapı ve ölçüt geçerliği) incelenmemiş olduğu göze çarpmaktadır (Sprong ve ark. 2007). Benzer şekilde, bu görevlerin görünüm ve ekolojik geçerliğine ilişkin çalışmalar da kısıtlıdır (Concoran

ve Frith 2003). Bu çalışmada, kuramsal bir temele dayandırılarak oluşturulan örtük değişkenlerin (İYİ ve İ) yapısı ve birbirleriyle ilişkileri ile ilgili bir hipotez geliştirilmiş ve geçerlik, doğrulayıcı faktör analizi (DFA) (confirmatory factor analysis) ile incelenmiştir. DFA için yukarıda belirtilen aşamalara katılmamış, Ankara'daki üniversitelerin farklı bölümlerinde eğitim gören, yaş ortalaması 20,56 ($\pm 3,19$) olan ve herhangi bir psikiyatrik tanı bildirmemiş toplam 214 sağlıklı gönüllü üniversite öğrencisinden (%77,1'i kadın, %22,9'u erkek) veri toplanmıştır. Düzeltilmemiş işitme ve görme bozukluğu olan ya da bilgisayar uygulamasını yapmasına engel olacak bir fiziksel engeli bulunan, renk körlüğü olan, psikiyatrik, nörolojik veya psikolojik rahatsızlığı bulunan ve bu nedenlerle hastaneye yatış yapan, psikiyatrik ilaç kullanımı devam eden, ilaç kullanmayı yeni bırakan ya da bilişsel bir bozukluğa neden olabilecek herhangi başka bir ilaç kullananlar çalışmaya dahil edilmemiştir.

Geliştirilen ZK görevlerinden alınan puanlar sıralama düzeyinde (0, 0,5 ve 1 puan) olduğu için veri temizliği kapsamında uç değerler belirlenirken sürekli değişkenler için hesaplanan z puanı ($\pm 3,29$) (Field 2009) kullanılamamıştır. Normallik testleri açısından veri setinin çarpık ve basık dağılımının olmadığı ve normal dağılım gösterdiği görülmüştür.

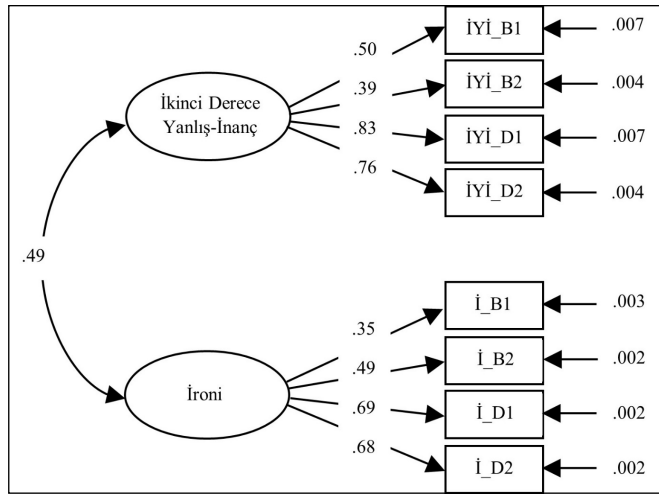
ZK bataryası görevleri, deneysel yöntemle geliştirilmiş senaryolar ile her bir senaryoya özgü birbirleri ile eşdeğer sorulardan oluşmaktadır. Örneğin, İYİ_B1 senaryosundaki 3. soru ile İYİ_B2 senaryosundaki 3. soru, senaryodaki olayın gerçekleştiği zaman, karakter sayısı dolayısıyla da zorluk düzeyi açısından eşdeğerdir. Aynı durum örtük değişkenler (İYİ ve İ) altında yer alan dörder adet senaryo için de geçerli olup senaryoların ölçülmesi hedeflenen örtük değişkenler altında yer alıp almadığının test edilmesi gerekmektedir. Bu amaçla 8 ZK senaryosundan elde edilen ortalama puanlar (İYİ için 10'ar, İ için 6'şar soru) üzerinden oluşturulan model DFA'ya tabi tutulmuştur.

DFA'da önerilen modellerin test edilmesi için öncelikle modelin tanımlanması gerekmektedir. Modeldeki her bir parametre için tek bir sayısal çözüm varsa ya da sayısal bir değer

verilebiliyorsa model tanımlanmış kabul edilir (Eroğlu 2003). Değişkenler arasındaki ilişkilere dayanarak üretilen hipotezlerin test edilmesi için modelin fazla tanımlanmış (overidentified) olması gerekir (Khine 2013). T-kuralına (Kenny 2011) dayanarak hesaplanan serbestlik derecesine göre önerilen modelin fazla tanımlanmış ve model-veri uyumunu sağlamış olduğu görülmüştür.

Faktör analizinin bir diğer varsayımı ise, verilerin aynı zamanda çok değişkenli normal dağılım göstermesidir. Çok değişkenli normal dağılım AMOS 24 paket programı kullanılarak hesaplanmış ve Mardia standardize çok değişkenli basıklık değeri 5,86 bulunmuştur. Yapılan karşılaştırma ($5,86 < 8$) sonucunda, mevcut veri setinin çok değişkenli normal dağılıma uyduğu görülmüştür.

Bu ön değerlendirmeler sonucunda, birinci düzey DFA modeli oluşturularak ZK bataryasının yapısındaki örtük faktörler ile bu faktörler arasındaki karşılıklı bağımlı etkiler AMOS 24 paket programı kullanılarak En Yüksek Olabilirlik (Maximum Likelihood: ML) kestirim yöntemine göre test edilmiştir. Örtük değişken olarak ele alınan İYİ ve İ elips şeklinde ifade edilmiştir (Şekil 1). Bu iki faktör birbiriyle ilişkili olup çift yönlü okla gösterilmiştir. Faktörleri temsil eden 8 adet gözlenen değişken ise dikdörtgen ile gösterilmiştir.



Şekil 1. Standardize Edilmiş Sonuçlar ile HACETTEPE-BTZK'nin Birinci Düzey DFA Sonuçları (iki faktörlü model)

Not: Şekilde gösterilen değerler (sağdan sola doğru): ZK alt testleri için hata varyansları (unique/error variances), standardize edilmiş faktör yükleri, iki faktör arasındaki korelasyon katsayısı.

İYİ_B1, İYİ_B2, İYİ_D1 ve İYİ_D2 İYİ faktörü; İ_B1, İ_B2, İ_D1 ve İ_D2 ise İ faktörü üzerine yüklenmiştir. DFA'ya ilişkin sonuçlar Şekil 1'de; belirlenen modele ilişkin uyum indeksi değerleri ise Tablo 2'de verilmiştir. Analizlerde, hangi uyum indekslerinin değerlendirileceğine ya da standart olarak kabul edileceğine dair kesinlik olmadığı için (aktaran Çapık 2014, Karagöz ve Ağbektas 2016) yaygın olarak kullanılan χ^2 , sd, χ^2 /sd, RMSEA, SRMR, GFI, AGFI, CFI ve TLI (NNFI) uyum indeksleri rapor edilmiştir. DFA sonuçlarına göre, elde edilen uyum indeksleri modelin çok iyi bir uyuma sahip olduğunu ortaya koymuştur [χ^2 (19, N=214)=26,14, $p > 0,05$, χ^2 /sd=1,38, RMSEA=0,042 (%90 güven aralığı değerleri: 0,000–0,078), SRMR=0,05, GFI=0,97, AGFI=0,95, CFI=0,98, TLI (NNFI)=0,97].

Modelin alt test faktör yüklerinin İYİ senaryoları için 0,39 ile 0,83; İ senaryoları için 0,35 ile 0,69 aralığında değiştiği ve örtük faktörlere yüklenen tüm alt testlerin istatistiksel olarak anlamlı olduğu bulunmuştur ($p < 0,001$). Ayrıca bütün faktör yükleri, Tabachnick ve Fidell'in (2014) belirttiği, kabul edilebilir faktör yükü değeri olan 0,32'nin üzerindedir. Bu değer, açımlayıcı faktör analizi (AFA) (exploratory factor analysis) için geçerli olsa da belirlenen bu ve diğer faktör yüklerinin (0,70 ve üzeri mükemmel, 0,63 çok iyi, 0,55 iyi, 0,45 orta ve 0,32 zayıf), ölçüm hatasından uzaklaştırılan DFA katsayıları için de rehber olabileceği düşünülmektedir (DiStefano ve Hess 2005). Son olarak, modelin açıklayıcı gücü, yapıların belirleyicilik katsayısı (R^2) hesaplanarak incelenmiştir. Falk ve Miller (1992), R^2 değerinin $\geq 0,10$ olması gerektiğini önermiştir. Bulgular, tüm R^2 değerlerinin 0,10'dan yüksek olduğunu göstermiştir. DFA sonuçları Tablo 3'te verilmiştir.

Senaryoların faktör puanı ağırlıkları Tablo 4'te sunulmuştur. Bu tablo incelendiğinde, geliştirilen senaryoların her birinin, beklendiği gibi kendi örtük değişkenleri altında yer aldığı görülmektedir. Ayrıca örtük değişkenler arasındaki standardize edilmiş korelasyon değeri 0,49 olup, bu değer istatistiksel olarak anlamlıdır ($p < 0,001$) (bkz. Tablo 5). Bu sonuca göre, İYİ ve İ'nin birbirlerinden farklı yapılar karşılık gelen ZK görevleri olduğu yorumu yapılabilir. DiStefano ve Hess'in (2005), örtük değişkenler arasındaki korelasyonun yapı geçerliğinin bir kanıtı olduğu görüşü doğrultusunda, istatistiksel olarak anlamlı bulunan korelasyon ($r=0,49$, $p < 0,05$), geliştirilen bataryanın yapı geçerliğine kanıt sağlamıştır.

Tablo 2. HACETTEPE-BTZK için DFA Uyum İndeksi Değerleri

Model	χ^2	sd	χ^2 /sd	RMSEA	SRMR	GFI	AGFI	CFI	TLI (NNFI)
DFA	26,14	19	1,38	0,042 (0,000 – 0,078)*	0,047	0,971	0,946	0,978	0,967

N=214; $p > 0,05$.

χ^2 = Ki-Kare, sd = Serbestlik Derecesi, RMSEA = Yaklaşık Hataların Ortalama Karekökü, SRMR = Standardize Edilmiş Artık Ortalamaların Karekökü, GFI = İyilik Uyum İndeksi, AGFI = Düzeltilmiş İyilik Uyum İndeksi, CFI = Karşılaştırmalı Uyum İndeksi, TLI (NNFI) = Tucker-Lewis İndeksi (Normlaştırılmamış Uyum İndeksi)

* Parantez içinde verilen değerler, %90 güven aralığı değerleridir.

Tablo 3. HACETTEPE-BTZK'nin DFA Sonuçları

		β	t	SH	R ²
İYİ	İYİ_B1	0,50	7,09*	0,02	0,25
	İYİ_B2	0,39	5,37*	0,02	0,15
	İYİ_D1	0,83	12,29*	0,02	0,70
	İYİ_D2	0,76	11,06*	0,02	0,57
İ	İ_B1	0,35	4,44*	0,01	0,12
	İ_B2	0,49	6,34*	0,01	0,24
	İ_D1	0,69	8,84*	0,01	0,47
	İ_D2	0,68	8,77*	0,01	0,46

β = Faktör Yüğü, t = Test İstatistiğı, SH = Standart Hata, R² = Belirleyicilik Katsayısı

Tablo 4. HACETTEPE-BTZK Senaryolarının Faktör Puanı Ağırlıkları

	İYİ_B1	İYİ_B2	İYİ_D1	İYİ_D2	İ_B1	İ_B2	İ_D1	İ_D2
İYİ	0,074	0,070	0,296	0,250	0,013	0,027	0,056	0,047
İ	0,008	0,008	0,033	0,028	0,067	0,144	0,293	0,250

Tablo 5. HACETTEPE-BTZK Örtük Değışkenleri Arasındaki Standardize Edilmiş Korelasyon ve t Değıeri

			β	SH	t	p
İYİ	<-->	İ	0,49	0,08	6,29	0,000
p<0,001. β : Faktör Yüğü, SH: Standart Hata, t = Test İstatistiğı						

p<0,001. β : Faktör Yüğü, SH: Standart Hata, t = Test İstatistiğı

Özetle, bir ölçme aracının soyut olan yapıları ne derece doğru ölçtüğünü gösteren yapı geçerliğini (Çokluk ve ark. 2014) belirlemek için kullanılan DFA, önerilen kuramsal model ile gözlenen kovaryansın tutarlılığını ve benzer yapıları farklı örneklerde aynı doğruluk oranıyla ölçüp ölçmediğini (measurement invariance) değerlendirmektedir (Kline 1998). Bu kapsamda, DFA bulguları, önerilen kuramsal yapı olan ZK'nin başarıyla ölçüldüğünü ve yapı geçerliğinin sağlandığını göstermiştir.

DFA sonuçlarına bağlı olarak, önerilen model için ölçümlerin bir dış standart kullanılarak değil, birbirlerine karşı değerlendirilmesini içeren ve bir yapı geçerliği yöntemi olan yakınsak (convergent) ve ıraksak (discriminant) geçerlik de test edilmiştir (Kline 2011). Aynı yapı içindeki ölçümlerden hangilerinin birbirleriyle ilişkili olduğunu gösteren yakınsak geçerliğin değerlendirilmesinde standardize edilmiş faktör yükleri kullanılabilir. İyi bir yakınsak geçerlik için faktör yüklerinin $\geq 0,70$ olması gerekirken; $\geq 0,50$ ise kabul edilebilir olarak değerlendirilir (Hair ve ark. 2013, Ekinci Demirelli 2024, Yaşlıoğlu 2017). Bu bağlamda hem İYİ hem de İ için Bileşik/Yapı Güvenirlik (Composite/Construct Reliability: ComR/CR) değeri, Ortalama Açıklanan Varyans (Average Variance Extracted: AVE) değerinden ve 0,70'ten büyüktür. AVE değerleri ise hem İYİ hem de İ için 0,50'den büyük değildir, ancak İYİ için AVE değerinin 0,50'ye yakın olduğu görülmektedir. Ayrıca AVE değerinin 0,50'den küçük olduğu

durumlarda ComR değeri 0,70'ten büyükse örtük faktörlerin güvenilir olduğu varsayılmaktadır (Hair ve ark. 2021).

Hangi yapıların birbirinden farklı olduğunu, yani yapıların tekilliğini gösteren ıraksak geçerlik için ise iki değerin daha hesaplanması gerekir. Maksimum Paylaşılan Varyansın Karesi (Maximum Squared Variance: MSV), bir faktörün diğer faktörlerden herhangi biriyle paylaştığı en yüksek varyansın (iki yapı arasındaki korelasyonun) karesidir (Hair ve ark. 2013, Yaşlıoğlu 2017). Buna göre Paylaşılan Varyansın Karesinin Ortalaması (Average Shared Square Variance: ASV) değeri sadece iki örtük değişken olması nedeniyle hesaplanamamış; ıraksak geçerlik değerlendirmesi MSV ve AVE üzerinden yapılmıştır. MSV değeri (0,24), İYİ için 0,42 ve İ için 0,32 değerlerinden küçüktür. Ayrıca AVE'nin karekök değerleri (İYİ için 0,65 ve İ için 0,57) faktörlerarası korelasyon değerinden (0,49) büyüktür. Bu ölçümlerin dışında ıraksak geçerlik için Henseler ve arkadaşları tarafından 2015 yılında önerilen ve son yıllarda sıklıkla kullanılan Heterotrait-Monotrait korelasyon oranı (HTMT) da hesaplanmıştır. Bu hesaplama göre 0,90'nın altındaki HTMT korelasyon oranı değerleri ıraksak geçerliğin sağlandığı anlamına gelir (Henseler ve ark. 2015). Yapılan hesaplama sonucunda HTMT korelasyon oranı 0,53 olarak bulunmuş ve ıraksak geçerlik için bir kanıt daha elde edilmiştir.

Özet olarak, her iki örtük faktör için hem yakınsak hem de ıraksak geçerliğin sağlandığı kanıtlanmıştır. Örtük değişkenlerin β , ComR, AVE ve HTMT değerleri Tablo 6'da verilmiştir.

Tablo 6. HACETTEPE-BTZK Örtük Değişkenlerinin β , ComR, AVE ve HTMT Değerleri

Örtük Değişkenler	Gözlenen Değişkenler	β	ComR	AVE	HTMT
İYİ	← İYİ_B1	0,505	0,94	0,42	0,53
İYİ	← İYİ_B2	0,392			
İYİ	← İYİ_D1	0,835			
İYİ	← İYİ_D2	0,757			
İ	← İ_B1	0,352	0,90	0,32	
İ	← İ_B2	0,491			
İ	← İ_D1	0,687			
İ	← İ_D2	0,68			

β = Faktör Yüklü, ComR = Bileşik Güvenirlik, AVE = Ortalama Açıklanan Varyans, HTMT = Heterotrait-Monotrait Korelasyon Oranı

Ölçüt Geçerliği

Ölçüt geçerliği için geliştirilen ZK bataryası ile Dokuz Eylül Zihin Teorisi Ölçeği (DEZTÖ) (Değirmencioğlu 2008) arasındaki korelasyon incelenmiştir. DEZTÖ'nün şizofreni tanısı almış hastalarda kullanılabilecek geçerli ve güvenilir bir ölçek olduğu gösterilmekle birlikte sağlıklı yetişkin grupta geçerliğine ilişkin kanıt bulunamamıştır (ayrıntılı bilgi için bkz. Değirmencioğlu 2008). Bununla birlikte ölçek, Değirmencioğlu ve arkadaşları tarafından 2018 yılında aynı veri seti ile geçerlik ve güvenilirlik analizleri ikinci kez yapılarak DEZİKÖ adıyla güncellenmiştir (ayrıntılı bilgi için bkz. Değirmencioğlu ve ark. 2018).

Mevcut çalışmada ölçüt geçerliğini belirlemek için psikometrik özellikleri açısından kabul edilebilir nitelikte olan DEZTÖ/DEZİKÖ kullanılarak asıl çalışmaya katılmamış toplam 55 gönüllü üniversite öğrencisinden veri toplanmıştır. Pearson Momentler Çarpımı Korelasyon Analizi sonucunda, geliştirilen 8 ZK senaryosundan alınan puanlar ile DEZTÖ/DEZİKÖ puanları arasında pozitif yönde anlamlı, ancak zayıf bir korelasyon ($r=0,32$, $p<0,05$) bulunmuştur. Bu örüntünün, iki ölçeğin madde sayıları ve madde güçlük düzeyleri arasındaki farktan kaynaklı olabileceği düşünülmektedir.

VIII. Aşama: Güvenirlik Çalışması

Yargıçlararası güvenirlik için Spearman's rho değerleri ve iç tutarlılık güvenirlik analizi için Cronbach α katsayısı kullanılmıştır. Bataryanın puanlanması sıralama düzeyinde olduğu için α katsayısı, sıralamalı (ordinal) α olarak hesaplanmıştır (Gadermann ve ark. 2012, Zumbo ve ark. 2007). Bu amaçla, doğrusal ilişki için geçerli olan Product-Moment Korelasyon Matrisi, Polychoric Korelasyon (PK) matrisine çevrilmiştir. Çevirme işlemi PLOMAT-C programı ile yapılmış, ardından da PK matrisi üzerinden yazılan SPSS kodu ile sıralamalı α katsayısı hesaplanmıştır. Her bir senaryo için sıralamalı α katsayıları Tablo 7'de sunulmuştur. Ortalama puanlar üzerinden değerlendirilen batarya toplamının iç tutarlılık güvenirlik analizi için hesaplanan Cronbach α katsayısı 0,72'dir.

Tablo 7. HACETTEPE-BTZK'nin Sıralamalı α Katsayıları

Faktör	Madde Sayısı	α
İYİ_B1	10	0,931
İYİ_B2	10	0,827
İYİ_D1	10	0,925
İYİ_D2	10	0,918
İ_B1	6	0,496
İ_B2	6	0,417
İ_D1	6	0,488
İ_D2	6	0,478

α = Alfa Katsayısı

Ortalama puanlar üzerinden elde edilen senaryolar/maddeler arası korelasyon matrisi Tablo 8'de verilmiştir.

α Katsayısının 0,70 ve civarı olması "yeterli", 0,80 ve civarı olması "çok iyi", 0,90 ve üzeri olması ise "mükemmel" olarak değerlendirilir (Kline 2011). Aynı ölçütler, sıralamalı α katsayısının değerlendirilmesinde de kullanılmaktadır (Lorenzo-Seva ve Ferrando 2015). Buna göre elde edilen değerlerin, İYİ senaryoları için 0,83 ile 0,93; İ senaryoları için 0,42 ile 0,50 arasında değiştiği, batarya toplamı için ise 0,72 olduğu görülmektedir. Dolayısıyla İYİ senaryoları için iç tutarlılığın mükemmel olduğu, buna karşın İ senaryoları için zayıf olduğu söylenebilir. Cronbach α değerinin 0,50'den küçük olması, gözlenen puanlardaki varyansın çoğunun rastlantısal hatadan kaynaklı olduğunu gösterir. Bununla birlikte, örneklem yeterince büyükse, örtük değişken yöntemlerinin kullanıldığı durumlarda daha düşük güvenirlilik seviyelerinin de kabul edilebileceği söylenmektedir (Kline 2011). Ayrıca ComR değerlerinin hem İYİ hem de İ senaryoları için (ComR değerleri $>0,70$ ve AVE değerleri) güvenilir sonuçlar vermesi ve senaryolara ait soruların deneysel yöntemle geliştirilmiş olması gözetilerek bataryadan İ senaryolarına ait soru çıkartılmamıştır.

Tablo 8. HACETTEPE-BTZK Senaryo Maddelerine İlişkin Korelasyon Matrisi

	İYİ_B1	İYİ_B2	İYİ_D1	İYİ_D2	İ_B1	İ_B2	İ_D1	İ_D2
İYİ_B1	1							
İYİ_B2	0,30**	1						
İYİ_D1	0,41**	0,32**	1					
İYİ_D2	0,36**	0,26**	0,64**	1				
İ_B1	0,08	0,08	0,06	0,12	1			
İ_B2	0,25**	0,19**	0,20**	0,22**	0,24**	1		
İ_D1	0,25**	0,20**	0,30**	0,29**	0,24**	0,28**	1	
İ_D2	0,15*	0,08	0,25**	0,18**	0,23**	0,36**	0,48**	1

* p<0,05, ** p<0,01.

Koyu renkle gösterilen korelasyon değerleri, birbirleri ile ilişkili görevlerin korelasyonunu göstermektedir. Beklendiği üzere, aynı örtük değişken altındaki senaryolar daha yüksek korelasyon değerlerine sahipken, farklı örtük değişkenler altındaki senaryoların korelasyon değerleri düşüktür. Örtük değişkenler arasındaki korelasyon değeri 0,49 olup istatistiksel olarak anlamlıdır.

Yargıcılararası güvenirlik çalışması ise, ölçüt geçerliği çalışması için toplanan 55 gönüllü üniversite öğrencisinin verisi üzerinden yapılmıştır. Bataryada yer alan her bir senaryo için cevapların nasıl puanlanması gerektiği konusunda detaylı eğitim almış iki psikolog tarafından 55 katılımcının verisi bağımsız olarak değerlendirilmiştir. Yapılan analiz (Spearman's rho) sonucunda ZK'nin (8 senaryo toplamı için) yargıcılararası güvenirlik katsayısı $r_s = 0,94$ ($p < 0,001$) olarak bulunmuş; her bir senaryo için ise güvenirlik katsayılarının 0,57 (İ_B1) ile 0,997 (İYİ_B2) arasında olduğu gözlenmiştir. Her bir senaryo ve toplam ZK için elde edilen korelasyon matrisi p anlamlılık değerleriyle birlikte Tablo 9'da verilmiştir.

HACETTEPE-BTZK bataryasının geliştirilme aşamalarını içeren akış şeması Şekil 2'de özetlenmiştir.

TARTIŞMA

ZK ölçümünde kullanılan testlerle ilgili olarak giriş bölümünde açıklanan yöntemsel sorunları gideren, ZK'yi objektif ve güvenilir olarak ölçen ve psikometrik özellikleri açısından güçlü bir testin gerekliliği literatürde sıklıkla dile getirilmektedir (Harrington ve ark. 2005).

Yetişkinlerle yürütülen araştırmalarda ZK çoğunlukla tek bir görev ile incelendiği için yakınsak geçerlik değerlendirilememektedir. Örneğin, şizofreni tanısı almış yetişkinlerde ZK'nin birden fazla ZK görevi (birinci ve ikinci derece yanlış-inanç, ironi ve metafor görevleri) kullanılarak incelendiği 30 makale gözden geçirilmiş ve bu makalelerden sadece beş tanesinde yakınsak geçerliğin rapor edildiği görülmüştür (Harrington ve ark. 2005). Öte yandan ZK ölçümünde sıklıkla kullanılan birinci ve ikinci derece yanlış-inanç görevlerinin ıraksak geçerlikleriyle ilgili sonuçlar da çelişkilidir. Örneğin, hangi ZK görevlerinin zekadan bağımsız olduğunu ya da eğer klinik grupla çalışılıyorsa, hangi semptomların bozulan ZK performansı ile ilişkili olduğunu belirlemek için ıraksak geçerliği

Tablo 9. HACETTEPE-BTZK Senaryoları ve Toplam HACETTEPE-BTZK Puanı için Elde Edilen Yargıcılararası Güvenirlik Katsayıları

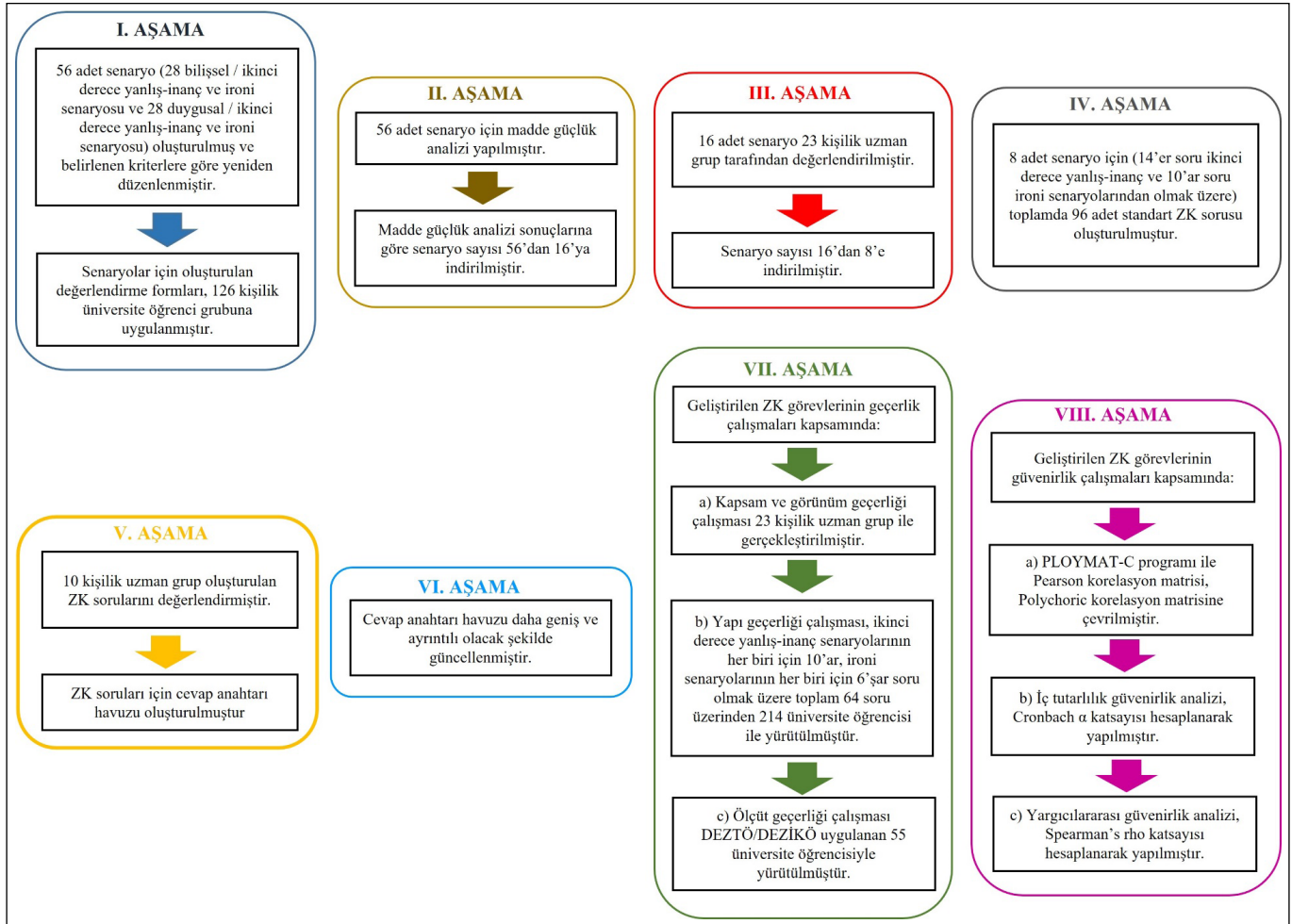
N = 55	Spearman's rho	p
İYİ_B1	0,907	0,000
İYİ_B2	0,997	0,000
İYİ_D1	0,931	0,000
İYİ_D2	0,808	0,000
İ_B1	0,573	0,000
İ_B2	0,870	0,000
İ_D1	0,778	0,000
İ_D2	0,750	0,000
ZK Toplam	0,935	0,000

Spearman's rho = Spearman'ın Sıralama Korelasyon Katsayısı

olan görevlerin kullanılması gereklidir (Harrington ve ark. 2005). Harrington ve arkadaşlarının (2005) da vurguladığı gibi, ZK görevlerinin geçerlik ve güvenirlik çalışmalarının yapılması acil bir ihtiyaçtır. Bu acil ihtiyacın bir sebebi de sayıları giderek artan ZK görevlerinden hangisi ve/veya hangilerinin ZK'nin en güvenilir ölçümü olduğu konusunda yaşanan karışıklıktır (Harrington ve ark. 2005).

Uluslararası literatürde kullanılan görevlerden yararlanarak ülkemizde geliştirilen ve ZK'yi farklı yönleriyle ele alan DEZTÖ'nün şizofreni tanısı almış hastalarda kullanılabilecek geçerli ve güvenilir bir ölçek olduğu gösterilmiştir (Değirmencioğlu 2008). Daha sonra yapılan çalışma ile güncellenen adıyla DEZİKÖ'nün sağlıklı yetişkin grup için de geçerli bir ölçek olduğu bulunmuştur (Değirmencioğlu ve ark. 2018).

Bu çalışmada, DEZTÖ/DEZİKÖ ve literatürde yer alan diğer ZK görevlerinden farklı olarak zorluk düzeyi ve ekolojik geçerliği yüksek (gerçek yaşam olaylarını içeren), üç boyutlu renkli ve canlı animasyon türü bilgisayar temelli HACETTEPE-BTZK Bataryası görevleri sağlıklı yetişkin bireyler için geliştirilmiş ve kapsamlı bir dizi geçerlik ve



Şekil 2. HACETTEPE-BTZK Bataryası Görevlerinin Geliştirilme Sürecine İlişkin Aşamaları İçeren Akış Şeması

güvenirlik çalışması yapılmıştır. Ayrıca mevcut ZK testlerindeki sınıflamaya dayanan ikili puanlamanın (0-1 veya Evet-Hayır gibi) dezavantajlarını gidermek amacıyla sıralamaya dayanan bir puanlama sistemi (0, 0,5 ve 1 puan) tercih edilmiştir. Yapı geçerliği için kullanılan DFA sonucunda elde edilen uyum indeksleri modelin çok iyi bir uyuma sahip olduğunu göstermiştir [χ^2 (19, N=214)=26,14, $p>0,05$, $\chi^2/sd=1,38$, RMSEA=0,042, SRMR=0,05, GFI=0,97, AGFI=0,95, CFI=0,98, TLI (NNFI)=0,97]. Modelin alt test faktör yüklerinin İYİ senaryoları için 0,39 ile 0,83; I senaryoları için 0,35 ile 0,69 aralığında değiştiği ve örtük faktörlere yüklenen tüm alt testlerin istatistiksel olarak anlamlı olduğu gösterilmiştir ($p<0,001$). Ayrıca bütün faktör yükleri, kabul edilebilir faktör yükü değeri olan 0,32'nin üzerindedir. Son olarak, modelin açıklayıcı gücünü ifade eden tüm R^2 değerleri 0,10'dan yüksektir. DFA sonuçlarına bağlı olarak önerilen model için ölçümlerin bir dış standarda karşı değerlendirilmesi yerine birbirlerine karşı değerlendirilmesini içeren yapı geçerliği yöntemlerinden yakınsak ve ıraksak geçerlik test edilmiş ve her iki örtük faktör için hem yakınsak hem de ıraksak geçerliğin sağlandığı kanıtlanmıştır. Ölçüt geçerliği kapsamında kullanılan DEZTÖ/DEZİKÖ

ile HACETTEPE-BTZK Bataryası'ndan alınan puanlar arasında pozitif yönde anlamlı, ancak zayıf ($r=0,32$, $p<0,05$) bir ilişki bulunmuştur. Yargıçlararası güvenirlik ve iç tutarlılık (Cronbach α) katsayıları ise sırasıyla $r=0,94$ ve $r=0,72$ 'dir. Psikometrik parametreler bir bütün olarak değerlendirildiğinde, geliştirilen HACETTEPE-BTZK Bataryası'nın psikometrik açıdan güçlü bir ZK testi olduğu görülmektedir. Ayrıca ZK'yi ölçen diğer uluslararası ve ulusal testler/ölçeklerle karşılaştırıldığında alt test faktör yüklerini ve örtük faktörlerle alt testler arasındaki ilişkiyi faktör analizi yöntemini kullanarak belirleyen ender çalışmalardan biri olmasının yanı sıra sağlıklı yetişkin örneklem için geliştirilmiş, ekolojik geçerliği yüksek ve özgün nitelikteki senaryolardan oluşan iddialı bir ZK bataryası olma özelliğine sahiptir.

Sonuç olarak, psikometrik açıdan güçlü, farklı zorluk düzeylerinde maddeler içeren, modüler veya bir bütün olarak uygulanabilen, ekolojik geçerliği yüksek, sosyal etkileşim içeren, bilişsel ve duygusal boyutları olan bilgisayar temelli özgün HACETTEPE-BTZK Bataryası geliştirilmiştir. Bataryanın gelecekte yapılacak çalışmalarla çeşitli klinik örneklemelerde kullanılabilirliğinin incelenmesi önerilmektedir.

Hibe ve Destek Bilgisi: Animasyonların yaptırılması için özel bir yazılım şirketinden hizmet alımı yapılmıştır. Ücret, Öğretim Üyesi Yetiştirme Programı (ÖYP) bütçesinden karşılanmıştır. Ödeme, doğrudan söz konusu şirkete üniversite aracılığıyla yapılmıştır. Bunun dışında söz konusu çalışma için herhangi bir destek alınmamıştır.

Teşekkür: Pilot verilerin toplanmasına ve yargıcılararası güvenilirlik çalışmasına katkı sunan Psk. Özlem DURMAZ'a teşekkür ederiz.

*Bu makale ilk yazarın doktora tez çalışmasından üretilmiştir.

KAYNAKLAR

- Amerikan Psikiyatri Birliği (2013) Ruhsal Bozuklukların Tanısal ve Sayımsal El Kitabı, Beşinci Baskı (DSM-5) (Çev. ed.: E Köroğlu). Ankara, Hekimler Yayın Birliği, 2013.
- Aslankara M (2019) Zihin kuramının bilişsel psikoloji bağlamında incelenmesi: yönetici işlevlerden ayırıştırma ve bilgisayar temelli batarya geliştirme çalışması (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Baksh RA, Abrahams S, Auyeung B ve ark. (2018) The Edinburgh Social Cognition Test (ESCoT): examining the effects of age on a new measure of theory of mind and social norm understanding. *PloS one* 13: 1-16.
- Baron-Cohen S, O'Riordan M, Stone V ve ark. (1999) Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *J Autism Dev Disord* 29: 407-18.
- Baron-Cohen S, Wheelwright S, Hill J ve ark. (2001) The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry* 42: 241-51.
- Bora E (2009) Şizofreni spektrum bozukluklarında zihin kuramı. *Türk Psikiyatri Derg* 20: 269-81.
- Bora E, Yücel M, Pantelis C (2009) Theory of mind impairment in schizophrenia: meta-analysis. *Schizophr Res* 109: 1-9.
- Bottiroli S, Cavallini E, Ceccato I ve ark. (2016) Theory of Mind in aging: comparing cognitive and affective components in the faux pas test. *Arch Gerontology Geriatrics* 62: 152-62.
- Brüne M, Brüne-Cohrs U (2006) Theory of mind-evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci Biobehav Rev* 30: 437-55.
- Corcoran R, Frith CD (2003) Autobiographical memory and theory of mind: evidence of a relationship in schizophrenia. *Psychol Med* 33: 897-905.
- Corcoran R, Mercer G, Frith CD (1995) Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophr Res* 17: 5-13.
- Çapık C (2014) Geçerlik ve güvenilirlik çalışmalarında doğrulayıcı faktör analizinin kullanımı. *J Anatolia Nurs Health Sci* 17: 196-205.
- Çokluk Ö, Şekercioğlu G, Büyükköztürk Ş (2014) Sosyal Bilimler için Çok Değişkenli İstatistik: SPSS ve LISREL Uygulamaları. Ankara, Pegem Akademi.
- Değirmencioglu B (2008) İlk kez geliştirilecek olan Dokuz Eylül Zihin Teorisi Ölçeği'nin (DEZTÖ) geçerlik ve güvenilirlik çalışması (Yayımlanmamış doktora tezi). Dokuz Eylül Üniversitesi Sağlık Bilimleri Enstitüsü, İzmir.
- Değirmencioglu B, Alptekin K, Akdede BB ve ark. (2018) Şizofreni hastalarında Dokuz Eylül Zihin Kuramı Ölçeği'nin (DEZİKÖ) geçerlilik ve güvenilirlik çalışması. *Türk Psikiyatri Dergisi* 29: 193-201.
- DiStefano C, Hess B (2005) Using confirmatory factor analysis for construct validation: an empirical review. *J Psychoeduc Assess* 23: 225-41.
- Dziobek I, Fleck S, Kalbe E ve ark. (2006) Introducing MASC: a movie for the assessment of social cognition. *J Autism Dev Disord* 36: 623-36.
- Ekinci Demirelli A (2024) Kadın hekimlerde iş yaşam dengesi ölçeğinin yapı geçerliliğinin doğrulayıcı faktör analizi ile incelenmesi. *Uluslararası Yönetim İktisat ve İşletme* 20: 508-31.
- Eroğlu E (2003) Toplam kalite yönetimi uygulamalarının yapısal eşitlik modeli ile analizi (Yayımlanmamış doktora tezi). İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Ertuğrul Yaşar Z (2022) Zihin kuramının gözden geçirilmesi: farklı zihinsel durumlar ve yaşam boyu gelişim. *Gelişim ve Psikoloji Dergisi*, 3: 75-92.
- Falk RF, Miller NB (1992) A Primer for Soft Modeling. Ohio, University of Akron, s. 80.
- Field AP (2009) Discovering Statistics Using SPSS, 3. Baskı. London, Sage.
- Frith CD (1992) The Cognitive Neuropsychology of Schizophrenia. UK, Psychology.
- Gademann AM, Guhn M, Zumbo BD (2012) Estimating ordinal reliability for Likert-type and ordinal item response data: a conceptual, empirical, and practical guide. *Pract Assess Res Evaluation* 17: 1-13.
- German TP, Hehman JA (2006) Representational and executive selection resources in 'theory of mind': evidence from compromised belief-desire reasoning in old age. *Cogn* 101: 129-52.
- Gregory C, Lough S, Stone V ve ark. (2002) Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: theoretical and practical implications. *Brain* 125: 752-64.
- Hair JF, Black WC, Babin BJ ve ark. (2013) Multivariate Data Analysis: Pearson New International Edition, 7. Baskı. UK, Pearson, s. 686-90.
- Hair JF, Hult GTM, Ringle CM ve ark. (2021) A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM), 3. Baskı. New York, SAGE Publications.
- Happé FG (1994) An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J Autism Dev Disord* 24: 129-54.
- Happé FG, Winner E, Brownell H (1998) The getting of wisdom: theory of mind in old age. *Dev Psychol* 34: 358-62.
- Harrington L, Siegert R, McClure J (2005) Theory of mind in schizophrenia: a critical review. *Cogn Neuropsychiatry* 10: 249-86.
- Henry JD, Phillips LH, Ruffman T ve ark. (2013) A meta-analytic review of age differences in theory of mind. *Psychol Aging* 28: 826-39.
- Henseler J, Ringle CM, Sarstedt M (2015) A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Sci* 43: 115-35.
- Herold R, Tényi T, Lénárd K ve ark. (2002) Theory of mind deficit in people with schizophrenia during remission. *Psychol Med* 32: 1125-9.
- Karağöz Y, Ağbektas A (2016) Yapısal eşitlik modellemesi ile yaşam memnuniyeti ölçeğinin geliştirilmesi; Sivas ili örneği. *Bartın University Faculty Economics Administrative Sci* 13: 274-90.
- Karakelle S, Ertuğrul Z (2012) Zihin kuramı ile çalışma belleği, dil becerisi ve yönetici işlevler arasındaki bağlantılar küçük (36-48 ay) ve büyük (53-72 ay) çocuklarda farklılık gösterebilir mi? *Türk Psikoloji Dergisi* 27: 1-21.
- Kaysılı B (2013) Zihin kuramı: otizm spektrum bozukluğu olan ve normal gelişen çocukların performanslarının karşılaştırılması. *A.Ü. Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi* 14: 83-103.
- Kenny DA (2011) Terminology and basics of SEM. <http://davidakenny.net/cm/basics.htm> adresinden alınmıştır.
- Khine MS (Ed) (2013) Application of Structural Equation Modeling in Educational Research and Practice, 7. Baskı. NL, Sense.
- Kline RB (1998) Software review: software programs for structural equation modeling: Amos, EQS, and LISREL. *J Psychoeduc Assess* 16: 343-64.
- Kline RB (2011) Principles and Practice of Structural Equation Modeling. New York, The Guilford.
- Küçük Z (2018) Zihin kuramı ve gelişim süreçleri. *Uludağ Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi* 19: 475-503.
- Lorenzo-Seva U, Ferrando PJ (2015) POLYMAT-C: a comprehensive SPSS program for computing the polychoric correlation matrix. *Behav Res Methods* 47: 884-9.
- Rowe AD, Bullock PR, Polkey CE ve ark. (2001) 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain* 124: 600-16.
- Saltzman J, Strauss E, Hunter, M ve ark. (2000) Theory of mind and executive functions in normal human aging and Parkinson's disease. *J Int Neuropsychol Soc* 6: 781-8.
- Sarı OT (2011) Zihin Kuramı Hikâyeleri Testi'nin Türk çocuklarına uyarlanması ve okul öncesi dönemdeki normal gelişim gösteren, zihin engelli ve otizmlili çocukların zihin kuramı gelişimlerinin karşılaştırılması (Yayımlanmamış doktora tezi). Marmara Üniversitesi, İstanbul.
- Shamay-Tsoory SG, Aharon-Peretz J (2007) Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia* 45: 3054-67.

- Shamay-Tsoory SG, Shur S, Barcai-Goodman L ve ark. (2007) Dissociation of cognitive from affective components of theory of mind in schizophrenia. *Psychiatry Res* 149: 11-23.
- Shamay-Tsoory SG, Tibi-Elhanany Y, Aharon-Peretz J (2006) The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. *Soc Neurosci* 1: 149-66.
- Shamay-Tsoory SG, Tomer R, Aharon-Peretz J (2005) The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology* 19: 288-300.
- Sprong M, Schothorst P, Vos E ve ark. (2007) Theory of mind in schizophrenia. *Br J Psychiatry* 191: 5-13.
- Stone VE, Baron-Cohen S, Knight RT (1998) Frontal lobe contributions to theory of mind. *J Cogn Neurosci* 10: 640-56.
- Şahin B, Önal BS, Hoşoğlu E (2020) Gaf Tanıma Testi Çocuk Formunun Türkçeye uyarlanması ve psikometrik özelliklerinin incelenmesi. *Anadolu Psikiyatri Dergisi* 21: 54-63.
- Tabachnick BG, Fidell LS (2014) *Using Multivariate Statistics*: Pearson New International Edition, 6. Baskı. UK, Pearson, s. 699.
- Tanrıverdi O (2022) Edinburgh Sosyal Biliş Testi'nin (ESBİT) Türkçeye uyarlanması, geçerlilik ve güvenilirliğinin incelenmesi (Yayımlanmamış yüksek lisans tezi). İstanbul Medipol Üniversitesi, İstanbul.
- Törenli Kaya Z, Alpay EH, Türkkal Yenigüç Ş ve ark. (2023) Zihinselleştirme Ölçeği'nin Türkçe çevirisinin geçerlik ve güvenilirlik çalışması. *Türk Psikiyatri Derg* 34: 118-24.
- Tülü B, Ergül C (2022) 3-5 yaş grubu çocuklara yönelik zihin kuramı testi: geçerlik güvenilirlik çalışması. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi* 55: 31-61.
- Wellman HM (1990) *The Child's Theory of Mind*. Cambridge, The MIT Press.
- Wiffen BD, O'Connor JA, Gayer-Anderson C ve ark. (2013) "I am sane but he is mad": insight and illness attributions to self and others in psychosis. *Psychiatry Res* 207: 173-8.
- Wimmer H, Perner J (1983) Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cogn* 13: 103-28.
- Yaşlıoğlu MM (2017) Sosyal bilimlerde faktör analizi ve geçerlilik: keşfedici ve doğrulayıcı faktör analizlerinin kullanılması. *İstanbul Bus Res* 46: 74-85.
- Yıldırım EA, Kaşar M, Gündük M ve ark. (2011) Gözlerden Zihin Okuma Testi'nin Türkçe güvenilirlik çalışması. *Türk Psikiyatri Derg* 22: 177-86.
- Zumbo BD, Gadermann AM, Zeisser C (2007) Ordinal versions of coefficients alpha and theta for Likert rating scales. *J Mod Appl Stat Methods* 6: 21-9.