

ÇOCUKLAR İÇİN ANİMASYONLU
BİLİMSEL YARATICILIK
TESTİ (ÇABİYAT):
TASARIMI VE UYGULAMASI

DOÇ. DR. NAZMIYE NAZLI ATEŞGÖZ

EĞİTİM
yayınevi

**ÇOCUKLAR İÇİN ANİMASYONLU BİLİMSEL YARATICILIK TESTİ (ÇABİYAT):
TASARIMI VE UYGULAMASI**

Doç. Dr. Nazmiye Nazlı Ateşgöz

Genel Yayın Yönetmeni: Yusuf Ziya Aydoğan (yza@egitimyayinevi.com)

Genel Yayın Koordinatörü: Yusuf Yavuz (yusufyavuz@egitimyayinevi.com)

Sayfa Tasarımı: Kübra Konca Nam

Kapak Tasarımı: Eğitim Yayınevi Grafik Birimi

T.C. Kültür ve Turizm Bakanlığı

Yayıncı Sertifika No: 76780

E-ISBN: 978-625-5997-96-8

1. Baskı, Aralık 2024

**ÇOCUKLAR İÇİN ANİMASYONLU BİLİMSEL YARATICILIK TESTİ (ÇABİYAT):
TASARIMI VE UYGULAMASI**

Doç. Dr. Nazmiye Nazlı Ateşgöz

IV+216 s., 160x240 mm

Kaynakça var, dizin yok.

E-ISBN: 978-625-5997-96-8

Bu çalışma yazarın doktora tezinden üretilmiştir.

Copyright © Bu kitabın Türkiye'deki her türlü yayın hakkı Eğitim Yayınevi'ne aittir. Bütün hakları saklıdır. Kitabın tamamı veya bir kısmı 5846 sayılı yasanın hükümlerine göre kitabı yayımlayan firmanın ve yazarlarının önceden izni olmadan elektronik/mekanik yolla, fotokopi yoluyla ya da herhangi bir kayıt sistemi ile çoğaltılamaz, yayımlanamaz.

EĞİTİM
YAYINEVİ

Yayınevi Türkiye Ofis: İstanbul: Eğitim Yayınevi Tic. Ltd. Şti., Atakent mah.
Yasemen sok. No: 4/B, Ümraniye, İstanbul, Türkiye

Konya: Eğitim Yayınevi Tic. Ltd. Şti., Fevzi Çakmak Mah. 10721 Sok. B Blok,
No: 16/B, Safakent, Karatay, Konya, Türkiye
+90 332 351 92 85, +90 533 151 50 42, 0 332 502 50 42
bilgi@egitimyayinevi.com

Yayınevi Amerika Ofis: New York: Eğitim Publishing Group, Inc.
P.O. Box 768/Armonk, New York, 10504-0768, United States of America
americaoffice@egitimyayinevi.com

Lojistik ve Sevkiyat Merkezi: Kitapmatik Lojistik ve Sevkiyat Merkezi, Fevzi Çakmak Mah.
10721 Sok. B Blok, No: 16/B, Safakent, Karatay, Konya, Türkiye
sevkiyat@egitimyayinevi.com

Kitabevi Şubesi: Eğitim Kitabevi, Şükran mah. Rampalı 121, Meram, Konya, Türkiye
+90 332 499 90 00
bilgi@egitimkitabevi.com

İnternet Satış: www.kitapmatik.com.tr
+90 537 512 43 00
bilgi@kitapmatik.com.tr

 **kitapmatik**
projetinizi kitapçınıza

İÇİNDEKİLER

ÖNSÖZ.....	V
Giriş.....	1
BÖLÜM 1: YARATICILIK NEDİR?.....	7
GENEL VE ALANA ÖZGÜ YARATICILIK	8
Genel Yaratıcılık Yaklaşımı	9
Zihnin Yapı Kuramı	9
Alana Özgü Yaratıcılık Yaklaşımı.....	11
Üç Bileşenli Yaratıcılık Modeli	12
Sistemler Modeli	13
BÖLÜM 2: BİLİMSEL YARATICILIK NEDİR?.....	15
BİLİMSEL YARATICILIK KURAMLARI	16
Bilimsel Keşfin İkili Arama Modeli	16
Bilimsel Yaratıcılığın Yapı Modeli.....	19
Yaratıcı Bilimsel Çağrışımlar Modeli.....	20
BİLİMSEL YARATICILIK VE BİLGİ	20
BİLİMSEL YARATICILIK VE CİNSİYET FARKLILIKLARI	21
BİLİMSEL YARATICILIĞIN ÖLÇÜLMESİ	23
İlk Geliştirilen Bilimsel Yaratıcılık Testleri.....	23
Bilimsel Yaratıcılığı Kâğıt-Kalem Testleri ile Değerlendirme Yöntemi.....	25
Bilimsel Üretkenlik Testi (BÜT)	25
Bilimsel Yaratıcılığın Yapı Modeli'ne göre geliştirilen testler	29
Yaratıcı Bilimsel Çağrışımlar Testi	33
Çoğul düşünme becerilerine odaklanan diğer testler	34
Çoğul ve tekil becerilere yönelik bilimsel yaratıcılık testi	35
Öz Değerlendirme ve Başkaları Tarafından Değerlendirilme Yöntemi	36
Konsensüse Dayalı Değerlendirme Yöntemi.....	38
Performans Temelli Değerlendirme	40
BÖLÜM 3: Bilimsel Yaratıcılığı Değerlendiren Yeni Bir Araç: ÇABİYAT Testinin	
Tasarımı ve Uygulama Süreci.....	43
ÇABİYAT'ın içeriği	43
ÇABİYAT'ın uygulama yöntemi.....	45
ÇABİYAT'ın uygulamasında Kullanılan Araç Gereçler	45
ÇABİYAT'ın Uygulamasında Oturma Düzeni.....	46
ÇABİYAT'ın Uygulama Süresi	46
ÇABİYAT'ın Uygulama Adımları	47
ÇABİYAT'ın Puanlama Yöntemi.....	47
ÇABİYAT'ın Geliştirme Süreci	49
ÇABİYAT'ın Teorik Çerçevesi	50
Madde Havuzunun Oluşturulması	51
Madde Seçimi	54

Animasyonları Tasarlanma Süreci	56
Maddelerin animasyon hikâyelerine dönüştürülmesi	56
Hikâyelerin uzman görüşüne sunulması	56
Animasyon tasarlama sürecine yönelik hazırlıklar	58
Animasyonların tasarlanması	59
Animasyonların uzman görüşüne sunulması	59
Ön Deneme Uygulaması	59
Pilot Uygulama	60
Ana Uygulama	67
Ana Uygulama Çalışma Grubu	69
Ana Uygulama-Fen Bilimleri Başarı Testinin Uyarlanması	70
Testin Psikometrik Özelliklerinin Belirlenmesi	73
BÖLÜM 4: ÇABIYAT'ın PSİKOMETRİK ÖZELLİKLERİ	74
BETİMSEL analizler	74
Geçerlik analizleri	75
Yapı Geçerliği	75
Açımlayıcı faktör analizi	76
Doğrulayıcı faktör analizi	81
Yakınsak ve ıraksak geçerliği	84
Ölçüt Geçerliği	86
Gelişimsel geçerlik	86
Uyum geçerliği	90
Maddelerinin Yanlılık Analizi	91
Güvenirlilik analizleri	93
ÇABIYAT'ın İç Tutarlılığı	93
ÇABIYAT'ın Puanlayıcılar Arası Güvenirliliği	97
BÖLÜM 5: ÇABIYAT'IN PSİKOMETRİK ÖZELLİKLERİNE İLİŞKİN TARTIŞMA VE SONUÇ	98
ÇABIYAT'ın Geçerliğine Yönelik Tartışma ve Sonuç	98
ÇABIYAT'ın Yapı Geçerliğine Yönelik Tartışma ve Sonuç	99
ÇABIYAT'ın Ölçüt Geçerliğine Yönelik Tartışma ve Sonuç	101
ÇABIYAT'ın gelişimsel geçerliğine yönelik tartışma ve sonuç	101
ÇABIYAT'ın uyum geçerliğine yönelik tartışma ve sonuç	103
ÇABIYAT'ın Madde Yanlılığına Yönelik Tartışma ve Sonuç	105
ÇABIYAT'ın Güvenirliliğine Yönelik Tartışma ve Sonuç	106
İç Tutarlılığa Yönelik Tartışma ve Sonuç	106
Puanlayıcılar Arası Güvenirliliğe Yönelik Tartışma ve Sonuç	108
KAYNAKÇA	109

ÖNSÖZ

Bilim ve yaratıcılık, insanlık tarihinin her döneminde ilerlemenin temel itici güçleri olmuştur. Günümüzde bireylerin ve toplumların bilimsel düşünme yeteneklerini geliştirmesi, her zamankinden daha hayati bir önem taşımaktadır. Ancak, özellikle erken yaş gruplarında bilimsel yaratıcılığın nasıl ölçüleceği ve bu süreçte hangi yöntemlerin daha etkili olacağı hâlâ derinlemesine araştırılması gereken bir konudur.

Bu kitap, bilimsel yaratıcılığın ölçümüne yönelik çalışmalara yenilikçi bir katkı sunma amacıyla hazırlanmıştır. Anasınıfı, birinci ve ikinci sınıf düzeyindeki öğrencilerin bilimsel yaratıcılıklarını değerlendirmek için geliştirilen Çocuklar için Bilimsel Yaratıcılık Testi (ÇABİYAT), yalnızca bir ölçüm aracı olarak değil, aynı zamanda erken yaşlarda yaratıcılığı teşvik eden ve bu potansiyeli açığa çıkaran bir yaklaşım olarak öne çıkmaktadır.

Kitapta, bilimsel yaratıcılık testlerinin tarihsel gelişimi, mevcut testlerin avantaj ve eksiklikleri ayrıntılı bir şekilde ele alınmıştır. Bunun yanı sıra, erken çocukluk dönemine özgü ihtiyaçlar ve sınırlılıklar gözetilerek tasarlanan ÇABİYAT'ın teorik temelleri, tasarım süreci ve uygulama sonuçları detaylandırılmıştır. Özellikle animasyon temelli ölçüm yaklaşımı, bu alanda yeni bir açılım sunmaktadır.

Bu eserin, bilimsel yaratıcılık üzerine çalışan akademisyenlere, öğretmenlere ve eğitim politikalarına yön veren paydaşlara ilham vereceğini ümit ediyorum. Çocukların bilimsel düşünme becerilerinin geliştirilmesine yönelik her çaba, daha aydınlık bir geleceğin inşasına önemli bir katkı sağlayacaktır.

Son olarak, bu eserin doktora tezimden doğmuş, uzun ve titiz bir akademik yolculuğun meyvesi olduğunu ifade etmek isterim. Eğitim bilimleri alanında, bilimsel yaratıcılığın erken yaşlarda nasıl ölçülebileceğine dair yeni bir bakış açısı geliştirme hedefiyle yola çıkmış ve özgün bir değerlendirme aracı tasarlanmıştır. Bu çalışma, yalnızca bilimsel bir ürün değil, aynı zamanda bir öğrenme ve keşif serüveninin hikâyesini de yansıtmaktadır.

Doç. Dr. Nazmiye Nazlı ATEŞGÖZ

GİRİŞ

“Yaratıcılık bulaşıcıdır. Onu bulaştırın!”

Albert Einstein

Bilimsel yaratıcılık ölçümlerine yönelik çalışmalar 20. yüzyılın ortalarında başlamıştır. 1970’li ve 1980’li yıllarda ilk bilimsel yaratıcılık testleri geliştirilmiştir. Bu testler geçerlik ve güvenilirlik çalışmalarının yetersiz olması, kültürel öğeler içermesi, kuramsal temelinin yetersiz olması ve hedef alınan yaş gruplarına uygun olmaması nedeniyle eleştiriler almıştır (Hu ve Adey, 2002; Liang, 2002; Mohamed, 2005). Bilimsel yaratıcılığı ölçmek amacıyla kullanılan ilk testlerden biri Majumdar (1975) tarafından geliştirilmiştir. Majumdar ortaokul ve lise öğrencilerinin bilimsel yaratıcılığını ölçmek için farklı bilim alanlarında problem çözme becerilerine yönelik Hint kültürüne özgü bir test tasarlamıştır. Sinha ve Singh (1987) ortaokul öğrencilerinin bilimsel yaratıcılığını ölçmek için esneklik, özgünlük, gözlem yeteneği, hayal gücü, analiz etme becerisi ve dönüştürme yeteneklerine yönelik İngilizce ve Hinduca bir test geliştirmiştir. Test Hindistan’da yaygın olarak kullanılıyor olsa da teste yönelik herhangi bir geçerlilik ve güvenilirlik çalışması bulunmamaktadır (Sharma ve Mahrshi, 2017). Bir diğer bilimsel yaratıcılık testi Friendlar (1983) tarafından geliştirilmiştir. Friendlar lise öğrencilerinin bilimsel yaratıcılığını ölçmek amacıyla veri toplama, problem çözme, hipotez geliştirme ve deney tasarlama gibi becerileri değerlendirmeye yönelik açık uçlu maddeler tasarlamıştır. Araştırmacı testin test-tekrar test güvenilirliği düşük olarak bulmuştur. Frederiksen ve Ward (1978) bilim insanlarının keşif sürecinde sıklıkla kullandığı hipotez oluşturma, araştırma tasarlama ve araştırma sonuçlarını yorumlama olmak üzere üç beceriye odaklanarak lisans düzeyine yönelik bir bilimsel yaratıcılık testi geliştirmişlerdir. Testin güvenilirliğinin yeterli düzeyde olduğu ancak ölçüt geçerliliğinin düşük olduğu bulunmuştur. 1970’li ve 1980’li yıllarda geliştirilen testlerde bilimsel yaratıcılığın daha geniş bir çerçevede ele alındığı ve bu testlerin çeşitli sınırlılıklarının olduğu söylenebilir.

2000’li yıllarda kimi araştırmacılar söz konusu sınırlılıkları kısmen dikkate alarak çeşitli bilimsel yaratıcılık testleri geliştirmiştir. Ancak bilimsel

yaratıcılık testlerine yönelik halen çeşitli eksiklikler olduğu göze çarpmaktadır. Hu ve Adey (2002) genel yaratıcılıkkaraştırmalarını dikkate alarak Bilimsel Yaratıcılığın Yapı Modelini oluşturmuş ve bir bilimsel yaratıcılık testi geliştirmişlerdir. Modelin kuramsal alt yapısını Guilford'un Zihnin Yapı Kuramı ve yaratıcılığa yönelik araştırmalar oluşturmaktadır (Chin ve Siew, 2005). Modelde alana özgü bileşenlerin eksikliği dikkat çekmektedir. Sak ve Ayas (2008) bu eksikliği vurgulamışlar ve fen bilimlerine özgü beceriler ve genel yaratıcılık becerilerini sentezlemişlerdir. Bu doğrultuda çoğul düşünme becerilerine odaklanan çalışmaları ve Bilimsel Keşfin İkili Arama Modeli'ni temel alarak Bilimsel Üretkenlik Testini (BÜT) geliştirmişlerdir. BÜT'ün geçerlik ve güvenilirliği çeşitli çalışmalarla ortaya koyulmuştur (Ayas, 2010; Ayas ve Sak, 2014; Sak ve Ayas, 2013). Güçlü bir kuramsal yapıya sahip olan BÜT 6-8. sınıf öğrencilerine yöneliktir. Ayas, BÜT'ün kuramsal yapısını dikkate alarak 3-5. sınıf öğrencilerine yönelik formunu geliştirmiştir (Ayas, 2017). Testin daha küçük yaş gruplarına yönelik formu bulunmamaktadır. Bu bağlamda geçerliği ve güvenilirliği ortaya koyulmuş ve güçlü bir kuramsal yapısı olan testin anasınıfı, birinci ve ikinci sınıf öğrencilerine yönelik geliştirilen bir formunun olmaması alanyazın için önemli bir eksikliklerdir.

2000'li yıllarda geliştirilen testlere yönelik eksikliklerden biri de ileri erken çocukluk dönemindeki çocukların (5-8 yaş) bilimsel yaratıcılığını ölçmeye yönelik bilimsel yaratıcılık testlerinin sınırlı olmasıdır (Chin ve Siew, 2005). Bu yıllarda geliştirilen testlerin çoğunluğunun hedef kitesini ortaokul ve lise öğrencileri oluşturmaktadır (Ayas, 2017). Örneğin Hu ve Adey (2002) daha önce değinilen Bilimsel Yapı Modeli'ni temel alarak ortaokul ve lise (7-10. sınıf) öğrencilerine yönelik bir bilimsel yaratıcılık testi geliştirmişlerdir. Siew, Chang ve Chin (2014) ise Hu ve Adey'in (2002) geliştirmiş oldukları bilimsel yaratıcılık testinin beşinci sınıf öğrencilerine yönelik formunu geliştirmiştir. Mohamed (2006) beşinci sınıf öğrencilerine, Kanlı (2014a) ortaokul öğrencilerine (5-8. sınıf), Usta ve Akkanat (2015) 6-8. sınıf öğrencilerine, Yang, vd. (2019) ise 3-6. sınıf öğrencilerine yönelik bilimsel yaratıcılık testleri geliştirmişlerdir. Her bir testin geçerlik ve güvenilirlik çalışmaları bulunmaktadır.

Okul öncesi dönemdeki çocukların bilimsel yaratıcılığını ölçmek için Chin ve Siew (2005) Bilimsel Yaratıcılığın Yapı Modeli'ni temel alarak bir kâğıt kalem testi geliştirmişlerdir. Geliştirdikleri testte öğrencilerden sözel olarak yöneltile sorulara verdikleri yanıtları çizimleri, çizdikleri resimleri isimlendirmeleri ve belirledikleri isimleri yazmaları beklenilmektedir. Okul öncesi öğrencilerinin çizim ve yazma becerilerinin sınırlı olması nedeniyle bu ölçekle okul öncesi öğrencilerinin gerçek performansları ölçülemeyebilir. Birinci ve ikinci sınıf düzeyinde bilimsel yaratıcılığı ölçen bir diğer test Yaratıcı Potansiyeli Değerlendirme (Evaluation of Potential Creativity, EPoC) bataryasında yer alan

bir alt testtir (Lubart, Zenasni ve Barbot, 2013). Ölçeğin temelinde yaratıcılığın doğasında çoğul ve tekil düşünmenin beraber rol oynadığı fikri yer almaktadır. EPoC farklı disiplinlerde çocukların ve genç yetişkinlerin yaratıcılığını ölçmek amacıyla geliştirilmiş bir bataryadır. Ölçek 7 ile 11 yaşları arasında bireysel, 12 ile 16 yaşları arasında ise grup olarak uygulanmaktadır. Bireysel uygulamada öğrencilere sorular sözlü olarak yöneltilmekte ve öğrencinin sözel olarak ürettiği cevaplar kaydedilmektedir. Uygulamadan sonra ses kayıtlarına göre ölçeğin puanlaması yapılmaktadır (De Vries ve Lubart, 2017). Alanyazında var olan bilimsel yaratıcılık testleri incelendiğinde hedef kitlenin daha çok 3-10. sınıf öğrencileriyle sınırlı kaldığı ve anasınıfı, birinci ve ikinci sınıf öğrencilerine yönelik ölçümlerin yeterli sayıda olmadığı söylenebilir.

Erken çocukluk döneminde yaratıcılığın ölçümünde yaşanan sorunlar, küçük yaştaki çocuklarla yapılan herhangi bir ölçümde karşılaşılabilecek sorunları içermektedir. Bu sorunlar değerlendirmeyi yapan kişiden, uygulama ortamından, çocuktan ve ölçme yönteminden kaynaklanabilir (Starkweather, 1964). Değerlendirmeyi yapan kişiden kaynaklı sorunları önlemek için değerlendiriciye çocukların özellikleri ve ölçme yöntemine yönelik eğitimler verilmelidir. Uygulama ortamı dikkat dağıtıcı faktörlerden arındırılmalıdır (Ford ve Dahinten, 2005; Moran, vd., 1983). Çocuk kaynaklı faktörleri en aza indirebilmek için çocuğun isteğini, motivasyonunu ve merak düzeyini artırıcı materyallerin ölçüm yönteminde kullanılması önerilmektedir (Starkweather, 1964; Zachapoulou, Makri ve Pollatou, 2009). Ölçme yönteminden kaynaklı sorunları önlemek için ise çocukların gelişim özelliklerine en uygun yöntemin belirlenmesi gerekmektedir.

Genel yaratıcılık ölçümünde öz değerlendirme, kişilik envanterleri, testler gibi pek çok farklı yöntem kullanılmaktadır (Hocevar ve Bachelor, 1989; Piirto, 1999). Ancak erken çocukluk döneminde bu yöntemlerin çoğu yaş grubunun özelliklerine uygun olmadığı için kullanılmamaktadır. Örneğin bireyin kendi yaratıcılığını değerlendirdiği öz değerlendirme yöntemi bu yaş grubu için uygun bulunmamaktadır (Renzulli, vd., 1976). Erken çocukluk döneminde yaratıcılığın ölçümünde başkaları tarafından yapılan değerlendirme, performans temelli değerlendirme ve standart değerlendirme yöntemleri kullanılmaktadır (Fishkin ve Johnson, 1998; Şahin Zeteroğlu, 2010).

Başkaları tarafından yapılan değerlendirme yönteminde öğretmen, akran, ebeveyn gibi kişiler bireyin yaratıcılığını değerlendirmektedir (Kaufman, Plucker ve Baer, 2008). Örneğin Liang (2002) lise öğrencilerinin bilimsel yaratıcılığını değerlendirmek amacıyla akran değerlendirme yöntemini kullanmıştır. Ancak erken çocukluk döneminde daha çok öğretmen görüşleri dikkate alınmaktadır (örn. Renzulli ve Hartman, 1971). Öğretmenler çocuğun

soru sorma, benzersiz cevaplar verme, risk alma gibi davranışsal özelliklerini derecelendirmektedir (Houtz ve Krug, 1995). Ancak ebeveyn, öğretmen, akran gibi başkaları tarafından yapılan değerlendirme yönteminde geçerlik ve güvenilirlik bağlamında istenen düzeye ulaşamama, özneliliğin yüksek olması, maliyetli olma ve zaman alma gibi problemlerle karşılaşmaktadır (Alencar, Fleith ve Bruno-Faria, 2014; Hocevar, 1979; Kaufman, Plucker ve Baer, 2008). Ayrıca bu yöntemin inanç, önyargı gibi kişisel faktörlerden etkileneceği ve daha yüzeysel bilgiler elde edilebileceği ifade edilebilir. Bu bağlamda başkaları tarafından yapılan değerlendirme yönteminin erken çocukluk döneminde bilimsel yaratıcılığın ölçümünde tek başına kullanmaya uygun bir yöntem olmadığı söylenebilir.

Performans temelli değerlendirme yönteminde öğrenciye bir problem ya da problem durumu verilmekte, öğrenciden de problemin çözümüne yönelik bir ürün ortaya çıkarması beklenmektedir (Nitko, 2004). Bu yöntem ile öğrencilerin bilgiyi nasıl yapılandırdıklarına ve neleri yapabildiklerine dair bilgiler edinilir (Wortham ve Hardin, 2016). Son yıllarda yaratıcılığın değerlendirilmesinde kullanılmaktadır. Ancak performans temelli testlerin güvenilirlik ve geçerlik bağlamında sınırlılıklarının olduğu ve yaratıcılık gibi karmaşık ve üst düzey becerilerin ölçümüne uygun olmadığı vurgulanmaktadır (Plucker, Callahan ve Tomchin, 1996). Diğer taraftan bu yöntem uygulama bakımından zor ve zaman alıcıdır (Kaufman, Plucker ve Baer, 2008).

Standart değerlendirme yönteminde genellikle testler kullanılmaktadır. Yaratıcılık testlerinde bireylerden bir takım soru ya da görevlere yönelik yanıt üretmesi beklenmektedir. Yaratıcılık testlerinin büyük çoğunluğunu çoğul düşünme testleri oluşturmaktadır (Hong ve Milgram, 2010). Çoğul düşünme testlerinde bireylerden bir durum ya da soruna yönelik çok sayıda fikir üretmeleri beklenmektedir (Kaufman, Plucker ve Baer, 2008). Genel yaratıcılık ya da alana özgü çoğul düşünme testlerinin büyük çoğunluğunu grup olarak uygulanan kâğıt kalem testleri oluşturmaktadır. Ancak kâğıt kalem testleri küçük yaş grubundaki öğrencilerin gelişim özelliklerine uygun değildir. Ayrıca grup değerlendirmelerinde pek çok çocuk yönergeleri ve test maddelerini anlayamamaktadır. Grup uygulamalarında küçük yaşlardaki çocuklar uygulayıcıdan yardım alamadığı için kaygılanmaktadır. Bu nedenle erken çocukluk döneminde standart testlerin bireysel olarak uygulanmasının güvenilirliği artıracığı ön görülmektedir (Pfeiffer, Petscher ve Jarosewich, 2007; Ruf, 2005). Bireysel uygulanan standart testlerin de çeşitli sınırlılıkları bulunmaktadır. Erken çocukluk dönemindeki çocukların dikkat süresinin kısa olması ve test maddelerinin ilgi çekici olmaması standart testlerin güvenilirliğini düşürebilmektedir (Moran, vd. 1983). Bu nedenle erken çocukluk dönemindeki çocuklar için standart testlerin içeriğinde farklılaştırmalar yapılmalıdır (Tegano,

Moran ve Godwin, 1986; Zachopoulou, Makri ve Pollatou, 2009).

Başkaları tarafından yapılan değerlendirme ve performans değerlendirme yönteminin, erken çocukluk döneminde yaratıcılığın ölçümünde tek başına kullanılabilir bir yöntem olmadığı ifade edilebilir. Standart değerlendirme yöntemlerinin içeriğinde değişiklikler yapılması gerekmektedir. Örneğin standart testlerde maddelerin sunulmasında animasyonlardan yararlanılması, çocukların teste olan ilgilerini ve motivasyonlarını artırması, erken çocukluk döneminde yapılan değerlendirmelerin güvenilirliğini de artırmaktadır (Popp, Tuzinski ve Fetzer, 2016). Björk-Willén ve Aronsson (2014) erken çocukluk dönemindeki çocukların animasyonlu etkinliklere katılımlarını nitel bir çalışma ile incelemişlerdir. Araştırmacılar çocukların oyun saatlerinde oynayabilecekleri animasyonlu etkinlikler hazırlamışlar ve üç çocuğun animasyonlara verdiği tepkileri video ile kaydetmişlerdir. Araştırmacılar çocukların animasyonlardaki karakteri dikkatli bir şekilde dinlediğini ve karakterin sorduğu tüm soruları yanıtladıklarını gözlemlemişlerdir. Analizleri sonucunda çocukların animasyon karakterlerini, diyaloglarını ve animasyonda hikâyeleştirilen olayları gerçek olarak algıladıklarını ve animasyon karakterleri ile etkileşime geçtiklerini ortaya koymuşlardır. Puolakanaho vd. (2003) animasyonun görsel ve sözel faktörler içermesi nedeniyle erken çocukluk dönemindeki çocukların dikkatini çektiğini belirtmektedirler. Bu iddialarından hareketle çocukların fonolojik farkındalıklarını ölçmek için animasyonlu bir değerlendirme yapmışlar ve ölçümün geçerli ve güvenilir sonuçlar verdiğini ortaya koymuşlardır.

Animasyonlu değerlendirmelerin tıp alanındaki uygulamalarına da rastlanmaktadır. Erken çocukluk dönemindeki çocukların fiziksel ağırlarını belirlemede (Swanston, vd., 1993) ve klinik değerlendirmelerde (Morris, vd., 2013) animasyonlardan yararlanılmaktadır. Bu çalışmalarda animasyonlu değerlendirmeler ile çocukların sağlık durumlarına yönelik en doğru tanının verilebileceği ortaya koyulmuştur. Bu bağlamda animasyonun standart değerlendirmelere dâhil edilmesi ölçümlerin geçerliğini ve güvenilirliğini artıracığı ifade edilebilir. Eğitim alanında animasyonlar çocukların fonolojik farkındalık (Torgesen, vd., 2010), okuma yazma (Hitchcock ve Noonan, 2000; Travers, 2010) ve kavrama (Block, vd. 2002) becerilerini artırmak amacıyla kullanılmaktadır. Ayrıca animasyonlar kavramların öğretiminde kullanılmaktadır ve animasyonların çocukların öğrenme hızlarının üzerinde olumlu etkisinin olduğu pek çok çalışmada ortaya koyulmuştur. Ke vd (2006) yapmış oldukları meta analiz çalışmada 13515 katılımcının olduğu 35 araştırma bulgularını incelemiş ve animasyonların öğrenmede etkili bir araç olduğunu bulmuşlardır. Bu doğrultuda okuma yazma bilmeyen anasınıfı öğrencileri ile yeni okuma yazma öğrenmiş birinci ve ikinci sınıf öğrencilerinin

bilimsel yaratıcılığını ölçmek için kâğıt kalem testleri yerine özellikle bu sınıf düzeyindeki öğrencilerin dikkatini çeken animasyonla uygulanan testleri kullanmanın daha nitelikli bir değerlendirmeye fırsat vereceği düşünülebilir.

Standart testlerde maddelerin sunulmasında video ve animasyon gibi uygulamada kolaylık sağlayabilecek teknolojik uyarlamalardan ya da ürünlerden yararlanılması erken çocukluk döneminden sonraki dönemlerde yapılan değerlendirmelerin de güvenilirliğini artıracakı düşünülebilir. Kâğıt kalem testleri ile video, animasyon gibi teknolojik ürünlerin kullanıldığı değerlendirmelerin karşılaştırıldığı araştırma bulguları da bu fikri destekler niteliktedir. Örneğin Dancy ve Beichner (2006) 325 üniversite öğrencisine 30 maddelik çoktan seçmeli bir fizik testini kâğıt kalem yöntemi ile uygulamış ayrıca testin animasyonlu formunu aynı öğrencilere uygulamıştır. Ayrıca 30 öğrenci ile sesli düşünme yöntemi ile animasyonlu testi çözmelerini istemişlerdir. Çalışmalarının sonucunda öğrencilerin animasyonlu testten anlamlı bir şekilde daha yüksek puan aldıklarını ve animasyonlu testi daha iyi anladıklarını ortaya koymuşlardır. Animasyonun değerlendirme geçerliğini artırdığı sonucuna ulaşılmıştır. Bir başka çalışmada Chan ve Schmitt (1997) bir durumsal yargılama ölçeğinin (situational judgement test) kâğıt kalem formunu bir gruba, video temelli değerlendirme yöntemini ise bir başka gruba uygulayarak grupların performanslarını karşılaştırmışlardır. Araştırmanın sonucunda katılımcıların video temelli değerlendirme yönteminde kâğıt kalem yöntemine göre daha yüksek puan aldıklarını bulunmuştur. Ayrıca video temelli değerlendirme yönteminin uygulandığı katılımcı grubun değerlendirmeye yönelik motivasyonlarının daha yüksek olduğu ve bu yolla yapılan ölçümlerin daha geçerli ve güvenilir bulgular ortaya koyduğu görülmüştür. Bu bağlamda katılımcıların teste yönelik motivasyonlarının yüksek olduğu durumlarda daha güvenilir ve geçerli bulgulara ulaşılabileceği ifade edilebilir. Pek çok araştırma fen bilimleri alanında kullanılan animasyonların öğrencileri motive ettiğini ortaya koymuştur (Rosen, 2009).

Bu çalışma kapsamında anasınıfı, birinci ve ikinci sınıf öğrencilerine yönelik animasyonla uygulanan Çocukları İçin Bilimsel Yaratıcılık Testinin tasarım ve uygulama sürecine yer verilmiştir.

BÖLÜM 1: YARATICILIK NEDİR?

“Daha iyisini yapmak için bir yol vardır. Onu bulun!”

Thomas A. Edison

Yaratıcılık kavramının gelişimi üç dönemde incelenebilir (Dacey, 1999). İlkçağlardan orta çağa kadar olan birinci dönemde yaratıcılık daha çok ilahi güçlerle ilişkilendirilmekte ve gerçek üstü bir durum olarak ele alınmaktadır. Rönesans hareketleriyle beraber başlayan ikinci dönemde kalıtsal yaratıcılık konusu gündeme gelmiştir. Üçüncü dönemde ise yaratıcılığa yönelik bağlamsal ve psikolojik etkilerin varlığına vurgu yapılmış ve yaratıcılığa yönelik modern çalışmaların temeli oluşturulmuştur.

Modern yaratıcılık araştırmaları 1950 yılında Guilford’un Amerikan Psikoloji Derneği (American Psychological Association, APA) başkanı olarak yapmış olduğu bir konuşmasıyla birlikte başlamıştır (Sawyer, 2012). Bu konuşmadan sonra yaratıcılık farklı açılardan ele alınarak kavramsallaştırılmıştır (Plucker ve Runco, 1998). Kimi bakış açılarına göre kişilik ön plana çıkarılırken, kimi bakış açılarında ise bilişsel süreçlere vurgu yapılmıştır (Sawyer, 2012). Örneğin Csikszentmihalyi (1997) ortaya attığı yaratıcılık süreçlerinde şans faktörüne dikkat çekerken Amabile (1983) daha çok birey, çevre ve motivasyon unsurlarının önemini vurgulamıştır.

Yaratıcılığın çok boyutlu ve dinamik olması tanımlara da yansımıştır. Rhodes (1961) yaratıcılığa yönelik yapılan tanımları “yaratıcılığın dört P’si” olarak adlandırdığı birey (person), süreç (process), çevre (press) ve ürün (product) bileşenleri etrafında toplamıştır. Birey bileşeni alışkanlıklar, tutumlar, davranışlar, kişisel özellikler, değerler ve zekâ gibi faktörleri içermektedir. Süreç bileşeninde algı, motivasyon, düşünme, öğrenme ve iletişim gibi olgulara işaret edilmektedir. Çevre bileşeninde birey ile çevresi arasındaki ilişkiye vurgu yapılırken ürün bileşeni orijinal bir fikrin somutlaşarak fiziksel bir biçim alması sürecini kapsamaktadır.

Birey bileşenine odaklanan Dellas ve Gaier (1970) yaratıcı bireyin düşünsel yeteneklerinin ötesinde dürtüleri, ilgileri ve tutumlarıyla kendilerini ortaya koyduklarını iler sürmüşlerdir. Baer (1991) yaratıcılığı problem çözme

deneyimlerinde kullanılan düşünme alışkanlıkları, tutum, kişilik özellikleri ve beceriler olarak tanımlanmaktadır. Süreç bileşenine odaklanan Spearman ise (1930) yaratıcılığı işleyen bilinç ve bilinçaltı ile kurulan ilişkisel süreçler olarak tanımlamıştır. Torrance (1988) yaratıcılığı bir süreç olarak farklı aşamalarda ve farklı alanlarda uygulanabilir biçimde ele almıştır. Sternberg ve Lubart (1996) zihinsel yetenekler, bilgi, motivasyon, düşünce tarzı, kişilik ve çevre gibi altı önemli kaynağın yaratıcılığın ortaya çıkabilmesi için gerekli olduğunu ileri sürmüşlerdir. Buna göre, özellikle yaratıcılığı ödüllendiren ve teşvik eden bir çevre olmaksızın diğer kaynakların yaratıcılığın ortaya çıkması için yetersiz kalacağını vurgulamışlardır. Bu bağlamda yaratıcılığı çevreyle derinden ilişkilendirmişlerdir. Yine Amabile (1982) sosyal çevrenin etkisine vurgu yapmıştır. Ürün bileşenine odaklanan Czikszenmihalyi (2014) yaratıcılığın asla tek başına bireysel eylemlerin bir sonucu olarak ortaya çıkmadığını belirtmektedir. Ona göre yaratıcılık ürün odaklı bir eylemdir. Bu eylemin ortaya çıkmasını sağlayan güçler alan ya da sosyal kurumlar, kültürel alan ve bireydir. Yaratıcılığa yine ürün odaklı bir anlayışla yaklaşan Sak (2014), yaratıcılığın somut olarak ortaya koyulabilmesinde en etkin çıktı olarak ürünü işaret etmektedir. Bu görüşe göre yaratıcı olarak kabul edilebilecek bir ürünün yenilik ve uygunluk koşullarının yanı sıra yararlılık, kullanışlılık, orijinallik, ilham verme gibi niteliklere de sahip olması gerekmektedir çünkü ürünün niteliği ve önemi yaratıcılık düzeyini artıran ya da azaltan en önemli faktörlerin başında gelmektedir.

Yaratıcılık tanımları incelendiğinde ilk tanımların süreç bileşeni üzerine odaklandığı görülmektedir. Daha yakın tarihlerde yaratıcılık tanımlarının ürüne göre şekillendirildiği düşünülebilir. Öyle ki ürün tanımlamaları çok fazla kabul görmüş, hatta yaratıcı süreç ya da yaratıcı kişilik çalışmalarına ağırlık veren uzmanlar dahi tanımlamalarda yaratıcı ürün ölçütünden yararlanmaya başlamıştır (Amabile, 1983). Yaratıcılık tanımlarının odak noktalarının kişi bileşeninden ürün bileşenine evrildiği ifade edilebilir. Bach'tan Sheakespeare'e, Newton'dan Darwin'e ya da Chanel ve Saatchi'den Beatles'a kadar farklı kişilerden ve farklı alanlardaki yaratıcılıktan bahsedilebilir. Bununla birlikte yemek yapma, espri yapma, herhangi bir eşyayı tamir etme ya da benzeri bir günlük hayat pratiği içerisinde yer alan bir sorunu çözme de yaratıcılık olarak değerlendirilebilir (Boden, 2004). Bu durum yaratıcılığın "genel mi yoksa alana özgü mü olduğu" tartışmasını akıllara getirmektedir.

GENEL VE ALANA ÖZGÜ YARATICILIK

"Genel mi yoksa alana özgü yaratıcılık mı?" sorusu yıllardır tartışılmaktadır (Baer, 1994a, 2016; Kaufman vd., 2008). Kimi araştırmalarda yaratıcılığın alana özgü beceri ve özelliklerin ötesinde olduğu ve genel yaratıcılık performanslarının geçerli olduğu öne sürülmüştür. Kimi araştırmalarda ise,

yaratıcı performansın alana özgü olduğu çeşitli kanıtlarla ortaya koyulmaya çalışılmıştır (Baer, 1998, 2016).

Genel yaratıcılık (domain generality) yaklaşımına göre farklı yaratıcı davranışlar arasında güçlü ilişkiler ve ortak psikolojik durumlar vardır (Ivcevic, 2007). Bu yaklaşımda yaratıcılık, bireyin tüm bilgi alanlarındaki gösterdiği yaratıcı performansı etkileyen genel kapasite olarak kabul edilmektedir (Han, 2000). Bununla birlikte alan genel yaratıcılık yaklaşımına göre bir alan için doğru olduğu kabul edilen bir durum diğer tüm alanlar için de geçerli olmaktadır (Baer, 2012). Diğer taraftan yaratıcılığın alana özgü olduğu (domain specific) anlayışında farklı yaratıcı davranışlar arasındaki ilişki oldukça zayıf ve farklı psikolojik durumlar mevcuttur (Ivcevic, 2007). Bu nedenle alana özgü yaklaşımda yaratıcılık belli bir alanla sınırlı ve eşsiz bir kapasite olarak görülmektedir. Böylece bireyler diğer alanlardan bağımsız olarak tek bir alanda yüksek yaratıcılık seviyesine erişebilirler (Han, 2000).

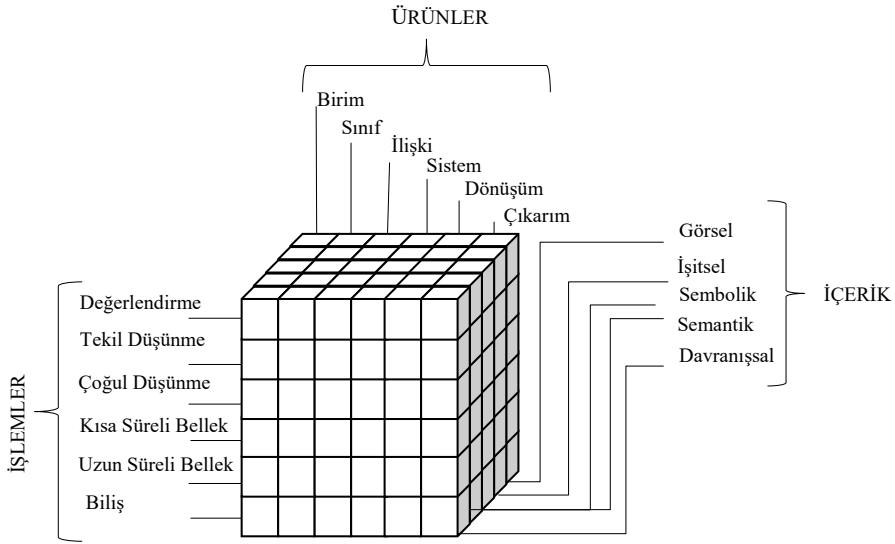
Genel Yaratıcılık Yaklaşımı

Yaratıcılığın genel bir yetenek olduğu düşüncesi uzun yıllar boyunca yaratıcılık alanında yapılan çalışmalara ve alan uzmanlarına ilham kaynağı olmuştur. Yaratıcılığın alanlar ve disiplinler genelinde uygulanabilirliği yaygın olarak kabul görmüştür (Hocevar, 1980; Milgram ve Milgram, 1976; Runco, 1992, 1999; Torrance, 1988). Alan uzmanları uzun yıllar boyunca yaratıcılığın genel oluşunu savunmuş ve alana özgü beceri ya da özellikleri aştığına dikkat çekmiştir (Baer, 1991, 2016). Belli bir dönemde, yaratıcılık türlerinin hemen hemen hepsinin açıklamasında genel yaratıcılık kuramlarından yararlanıldığı söylenebilir. Örneğin Hennessey ve Amabile (1999), Treffinger (1986) gibi araştırmacılar yaratıcılığı genel zihinsel özelliklerle ilişkilendirerek ele almıştır. Ancak genel yaratıcılık kuramları zaman içerisinde önemini kaybetmeye başlamıştır (Baer, 1998, 2016). Buna rağmen özellikle yaratıcılığa psikometrik açıdan yaklaşan araştırmacılar genel yaratıcılık kuramlarından beslenmektedir (Plucker, 1998). Baer (1991) en çok kabul gören genel yaratıcılık kuramının Guilford'un 1967 yılında ortaya koyduğu Zihnin Yapı Kuramı (Structure of Intellect Model) olduğunu öne sürmektedir.

Zihnin Yapı Kuramı

Zihnin Yapı Kuramı'nın temelleri 1955 ve öncesi yıllarda Guilford tarafından yapılan çalışmalara dayanmaktadır. Bu kuram Guilford tarafından geliştirilen Zekâ Teorisi'nin (Theory of Intellegience) dönüşümü niteliğindedir (Michael, 1999). Zihnin Yapı Kuramı'nda bilginin işlenmesi amacıyla yetenek ve fonksiyonlar bir araya getirilmektedir. Söz konusu yetenekler içerik, ürün ve işlem olmak üzere üç boyutta temsil edilmektedir (Woodman ve Schoenfeldt, 1990). Guilford bu üç boyutu ve boyutları temsil eden becerileri ortaya koymada

üç boyutlu bir küpten yararlanmaktadır. Başlangıçta modele göre bilginin işlenmesinde üç ana boyutun bileşenlerinin kombinasyonu ile 120 (4x5x6) yetenek türü tanımlanmıştır. 1985 yılında 150 (5x5x6), 1987 yılında ise 180 (5x6x6) yetenek türü oluşturulmuştur (Sternberg, 2003). Güncel modele Şekil 1.1'de yer verilmiştir.



Şekil 1.1. Zihnin Yapı Kuramı Modeli (Starko, 2017, s. 62)

Zihnin Yapı Kuramı'na göre zekâ içerik, işlem ve ürün olmak üzere üç ana bileşenden meydana gelmektedir. İçerik boyutu zihin tarafından tutulan ve kullanılan bilgi çeşitlerini kapsamaktadır. Ürün boyutu bilgi biçimini temsil etmektedir. İşlem boyutu ise zihin tarafından bilgi üzerinde gerçekleştirilen temel işlemleri içermektedir (Woodman ve Schoenfeldt, 1990). Şekil 2.1'de görüldüğü üzere girdileri temsil eden içerik bileşeni görsel, işitsel, semantik, sembolik ve davranışsal olmak üzere farklı beş alt bileşen daha yer almaktadır. Çıktıları temsil eden ürün bileşeni, sınıflar, ilişkiler, birimler, sistemler, dönüşümler ve çıkarımlarla beraber altı alt bileşen daha içermektedir. Zihinsel süreçleri temsil eden işlem bileşeni ise çoğul düşünme, tekil düşünme, kısa süreli bellek, uzun süreli bellek, değerlendirme ve biliş süreçleri gibi alt bileşenlerden oluşmaktadır (Sak, 2014). Guilford'un modelinde hem çoğul hem de tekil üretim yapılarının zihinsel işlem süreçlerinde beraber kullanılması alanda ortaya çıkan önemli farklılıklardan biri olmuştur (Maker, 1993).

Kurama göre çoğul düşünme bir soruna yönelik birden fazla olası yanıt üretebilme yetisi olarak tanımlanmaktadır (Starko, 2017). Tekil düşünme benzersiz tek bir fikir üretme ya da sonuç bulma yeteneği olarak

nitelendirilmektedir (Guilford, 1956). Guilford'a göre çoğul düşünme gerçek yaşamdaki yaratıcı davranışla ilgilidir (Hocevar, 1980). Ancak çoğul düşünme yaratıcı düşünme ile aynı anlama gelmemektedir. Gardner (2011) yapmış olduğu çalışmalarda, çoğul düşünme becerileri üzerine tasarlanmış testlerden elde edilen yüksek puanların yaratıcılığı yordamadığı sonucuna ulaşmıştır. Buna rağmen çoğul düşünme yaratıcılık sürecini temsil eden en önemli bileşenlerden biri olarak kabul edilmekte (Han, 2000) ve yaratıcılık testlerinde yoğun olarak kullanılmaktadır (Baer, 1999).

Çoğul düşünme becerilerinin yaratıcılık alanında kabul görmesini ve yaygınlaşmasını sağlayan neden Guilford'un çoğul düşünmeyi analiz, esneklik, akıcılık, orijinallik, nüfuz etme, tekrar tanımlama, sorunlara karşı duyarlılık, detaylandırma ve sentez olmak üzere sekiz boyutla ilişkilendirmesidir. Bu sekiz faktör arasından yaratıcılık alanında en yaygın kabul görenleri akıcılık, esneklik, orijinallik ve detaylandırma olmuştur (Michael, 1999). Akıcılık üretilen fikir sayısı, esneklik üretilen fikirlerin farklılığı, orijinallik üretilen fikirlerin özgünlüğü, detaylandırma ise üretilen fikirlerdeki ayrıntılar ile ilişkilidir (Sak, 2014).

Alana Özgü Yaratıcılık Yaklaşımı

Yaratıcılık performansının açıklanmasında tek faktörlü teorilerin işlevi zaman içerisinde sorgulanır hale gelmiştir. Yaratıcılığın tek yapıli bilişsel bir yapı olduğu ya da genel bir özellik olduğu yönündeki görüşler zayıflamıştır. Yaratıcılığa yönelik genel yönelimli yaklaşımların yerini daha çok alan yönelimli yaklaşımlar almaya başlamıştır (Baer, 1999). Çünkü yaratıcılığın alana özgü olduğu yaklaşımı hem teorik hem de deneysel çalışmalarla desteklenmiştir (Baer, 2016; Han, 2000).

Alana özgü bilgi ve becerilerin yaratıcılığa olan katkılarına yapılan vurgu zaman içerisinde artış göstermiştir (Amabile, 1983; Gruber ve Davis, 1988; Gardner, 2011; Csikszentmihalyi, 2014; Baer, 2016). Yaratıcılığın göreve özgü (task specific) olduğu görüşü araştırma sonuçları ile kanıtlanmıştır (Baer, 1994a, 1994b; Plucker, 1998). Gardner'ın Çoklu Zekâ Teorisi'nde önerdiği ilgili zekâ alanları, yaratıcılığın da alana özgü olabileceği fikrine katkıda bulunmuştur (Baer, 1999). Bununla birlikte, Runco (1987) üstün yetenekli çocukların sadece belli başlı alanlarda yaratıcı performanslar ortaya koyabildikleri sonucuna ulaşmış ve yaratıcılığın alana özgü olduğu yaklaşımına katkıda bulunmuştur.

Yaratıcılığın alana özgü yaklaşımında oldukça işlevsel roller üstlenen modeller ortaya konmuştur. En önemli modellerden biri Teresa Amabile tarafından 1983 yılında ortaya koyulan Üç Bileşenli Yaratıcılık Modeli'dir (Lubart, 1999). Yaratıcılığın alana özgü olduğunu savunan bir diğer önemli model ise Csikszentmihalyi tarafından geliştirilen Yaratıcılığın Sistemler Modeli'dir.

Üç Bileşenli Yaratıcılık Modeli

Üç Bileşenli Yaratıcılık Modeli iki önemli varsayım üzerine kurulmuştur. İlk varsayıma göre, günlük ve sıradan sayılabilecek yaratıcılık seviyelerinden tarihsel anlamdaki önemli icat, performans ya da keşifler gibi yüksek yaratıcılık seviyesinde kabul edilebilecek eylemlere doğru giden bir süreklilik vardır. İkinci varsayıma göre aynı alanda çalışan bireylerin yaratıcılık performansları farklılık gösterebilir.

Amabile modelinde yaratıcılığın alana özgü beceriler, görev motivasyonu ve yaratıcılıkla ilgili süreçler gibi bireysel boyuttaki bileşenler ve sosyal çevre gibi bireyin dışında var olan bileşenlerin bir araya gelmesiyle meydana geldiğini vurgulamaktadır. Buna göre yaratıcı bir çalışmanın ortaya çıkabilmesi için bu üç bileşenin her birinin mutlak varlığı gerekmektedir (Amabile, 1983, 2012; Amabile, vd.,1996).

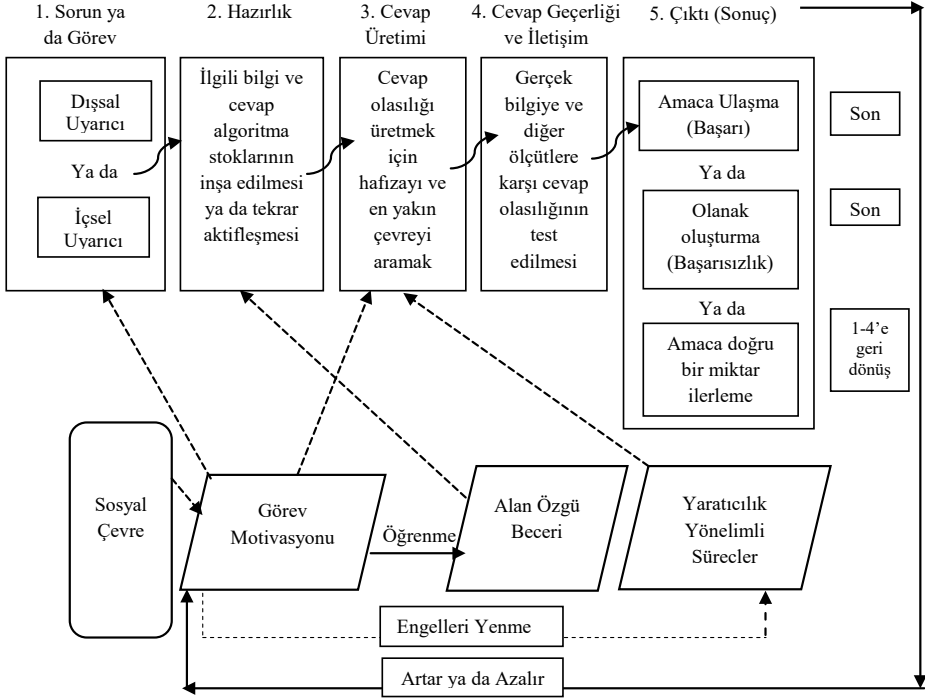
Alana özgü beceriler (domain-relevant skills) belirli bir alanda sahip olunan bilgi, teknik beceri, uzmanlık, yetenek ve zekâ gibi oluşumları temsil etmektedir (Amabile, 2012). Bu bileşene göre birey yaratıcı performans göstermek istediği alana ilişkin bilgi ve beceri sahibi olmalıdır (Starko, 2017).

Yaratıcılıkla ilgili süreçler (creativity-relevant processes) fikir üretme becerileri ve disiplinli çalışma biçimleriyle bağımsızlığa, risk almaya, sorunlara yönelik yeni bakış açıları geliştirmeye olanak sağlayan bireysel özellikleri ve bilişsel tarzları kapsamaktadır (Amabile, 2012). Yaratıcı çalışma biçimlerini de kapsayan bu bileşene göre bireylerin bir işe yönelik konsantrasyonları ve dikkatleri, odaklanmaları, belirsizliklere yönelik tahammülleri yaratıcılık için gerekli olan önemli özelliklerdir (Starko, 2017).

Görev motivasyonu (task motivation) en genel anlamda içsel bir motivasyon olarak tutku ile ilişkilendirilmektedir. Yaratıcılık dışsal kaynaklı motivasyonlar yerine tamamen bireysel ilgi ya da tatminlere dayanan motivasyonlarla gerçekleştirilen görevlerle ortaya çıkmaktadır (Amabile, 2012). Bireyler içsel motivasyonlarıyla bir işe yöneldiklerinde yaratıcılık yeteneklerinin en üst seviyesine çıkabilmektedir. Lubart (1999) görev motivasyonunun, bireyin söz konusu göreve dâhil olma ve görevi tamamlamaya yönelik tutumlarını içerdiğini iddia etmektedir. Görev motivasyonu bileşeni, Amabile'nin yaratıcılık alanına kazandırdığı en önemli katkı olarak değerlendirilmektedir (Starko, 2017).

Sosyal çevre (social environment) bireyin iş ve sosyal çevresi olarak tanımlanmaktadır. Söz konusu bileşen bireyin sahip olduğu içsel motivasyon ya da yaratıcılığa engel ya da uyarıcı olabilecek tüm dışsal motivasyonları kapsamaktadır. Bu anlamda birey yeni fikirlere karşı çıkılan, politik sorunların olduğu, baskı ve motivasyonu olumsuz etkileyen benzeri faktörlerin olduğu ortamda daha zayıf bir yaratıcı performans gösterir. İşbirliği, pozitif rekabet,

özgürlük ve yeni fikirlere saygının olduğu bir ortamda bireyin gösterebileceği yaratıcı performans daha yüksek olacaktır (Amabile, 2012). Modelin işleyişi Şekil 1.2’de yer almaktadır.



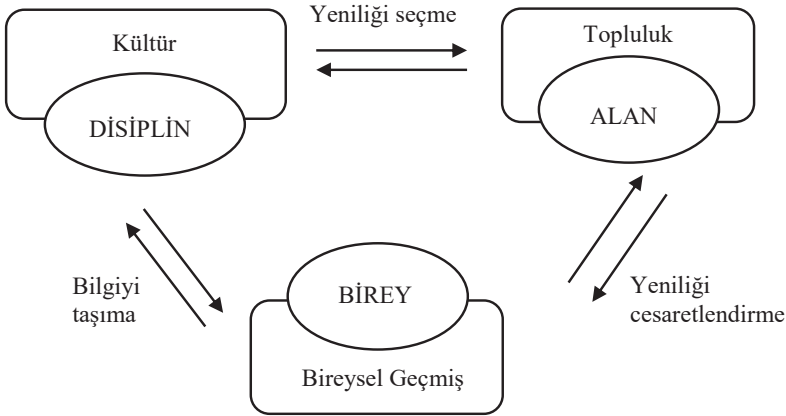
Şekil 1.2. Üç Bileşenli Yaratıcılık Modeli (Amabile, 1983, s.367)

Sistemler Modeli

Csikszentmihalyi (1999) yaratıcılığı daha çok kültürel ve sosyal süreçlerle ilişkilendirerek Sistemler Modeli'ni önermiştir. Model ile birlikte yaratıcılık çalışmalarının temel sorularından biri olan “Yaratıcılık nedir?” sorusu “Yaratıcılık nerededir?” olarak değişmiştir (Csikszentmihalyi, 2014). Çünkü modelde yaratıcılık hem bireysel hem de sosyal ve kültürel faktörlerden oluşan bir sistem kapsamında düşünülmektedir (Csikszentmihalyi ve Sawyer, 2014).

Modele göre disiplin, alan ve birey bileşenlerinin birbirleriyle karşılıklı etkileşimi sonucunda yaratıcılık ortaya çıkmaktadır. Disiplin bir takım sembolik kuralların ve süreçlerin var olduğu bileşendir. Alan, karar verici ya da otorite olarak bilinen bireyleri kapsamaktadır. Söz konusu uzmanlar fikirlerin ve çıktıların yaratıcı olup olmadığına karar vermektedir. Birey ise belirli bir disipline ait sembolleri kullanarak yeni fikirler üreten ve yeni oluşumları fark eden bileşeni temsil etmektedir. Bireylerin düşünceleri ya da eylemleri disiplinler içerisinde değişim yaratmaları ya da yeni disiplinlerin ortaya

çıkmasına neden olabilmeleri itibariyle önemlidir (Csikszentmihalyi, 1997). Sistemin her bir bileşeninin kendine özgü görevleri vardır ve her bir bileşen bir diğeri üzerinde etkiye sahiptir (Csikszentmihalyi ve Wolfe, 2014). Bileşenlerin birbirleri ile etkileşimi Şekil 1.3'te görselleştirilmiştir.



Şekil 1.3. Sistemler Modeli (Csikszentmihalyi, 1999, s. 315)

Modele göre birey, içerisinde yaşadığı kültürde elde ettiği bilgi birikimi üzerinde farklılıklar yaratabilir. Bireyin gelişmesinde ve şekillenmesinde çevre oldukça etkilidir. Çünkü bireyler, belli bir alan içerisinde ve belli bir düzeyde sahip olunan alan bilgisiyle yaratıcı performans gösterebilirler. Model, yaratıcılığın alana özgü bilgi birikimi gerektirdiğini savunmaktadır (Starko, 2017).

Modelin toplumsal ve bilimsel anlamda önemli sonuçları bulunmaktadır. Model, bireyin yaratıcı olarak kabul edilip edilmemesi durumunun bireysel özelliklerden ziyade disiplin içerisinde onaylanıp onaylanmamasına bağlı olduğuna dikkat çekmektedir. Bir buluşun yaratıcı olarak kabul edilebilmesi için buluşun kendisi yeterli değildir. Disiplin, alan ve birey arasında bir etkileşim olması koşuluyla yaratıcılık ortaya çıkabilir. Yaratıcılık ölçümü sosyo-kültürel faktörler tarafından belirlenebilmektedir (Sak, 2014).

Alana özgü yaratıcılığın temel alındığı modellerde yaratıcılığın disiplinlere göre şekilleneceği sonucu ortaya çıkmaktadır. Her disiplinin kendine özgü bileşenleri bulunmaktadır. Söz konusu bileşenlerle yaratıcılığın sınırları belirginleştirilebilir.

BÖLÜM 2: BİLİMSEL YARATICILIK NEDİR?

Yaratıcılığın alana özgü oluşumu fikri (Alexander, 1992; Amabile, 1996; Kaufman ve Baer, 2005, 2009) matematiksel yaratıcılık, sanatsal yaratıcılık, bilimsel yaratıcılık gibi farklı disiplinlere özgü yaratıcılığı gündeme getirmiştir. Ancak bilimsel yaratıcılığa yönelik yapılan çalışmaların sayısı ve kapsamı diğerlerine göre sınırlı kalmıştır. Bilimsel yaratıcılığa yönelik yapılan az sayıdaki çalışma arasından çoğu yaratıcı ürün, düşünce ve süreçler üzerine odaklanmıştır (Liang, 2002). Bilimin birikimli ilerleyen bir olgu olduğu dikkate alındığında, söz konusu birikimin en önemli kaynağının yeni icatlar ve yaratıcılık örnekleri olduğu söylenebilir. Yaratıcılık, bilimsel bilgi üretim süreçlerinde oldukça önemli bir rol oynamaktadır (Özdemir ve Dikici, 2017). Bilim insanları araştırmalarında bilimsel süreç becerileri, hayal gücü ve yaratıcılık gibi unsurlardan yararlanmaktadır (Hadzigeorgiou, vd., 2012). Problem çözme, hipotez geliştirme, deney tasarlama ve teknik yenilik gibi adımların hepsi bilime özgü olan belirli bir yaratıcılık biçimi gerektirmektedir (Lin, vd., 2003).

Bilimsel yaratıcılık hem genel hem de alana özgü bilgi ve beceri gibi bileşenleri kapsayan (Klahr, 2000), bir sorun çözme biçimi (Simon, 1977) ya da bir etkileşim olarak tanımlanmaktadır (Heller, 2007; Hu ve Adey, 2002). Bilimsel yaratıcılığın tanımlanmasında bilimsel süreç becerilerine odaklanan araştırmacılar da bulunmaktadır. Örneğin Klahr, Fay ve Dunbar (1993) bilimsel yaratıcılığın önemli gereklilikleri olarak bir takım karmaşık bilişsel beceriler ve bu becerilerin iç içe geçmesi sürecine vurgu yapmışlardır. Bilimsel yaratıcılık süreci hipotez oluşturma, deney tasarlama, deneysel sonuçları yorumlama ve hipotezlerin tekrar gözden geçirilerek düzenlenmesini kapsamaktadır.

Simonton (2004) bilimsel yaratıcılığı diğer yaratıcılık biçimlerinden daha üstün tutan ve bilimsel yaratıcılığa daha çok saygı duyan ortak bir inanışın varlığından bahsetmektedir. Bu inanışa göre Newton'un Matematiğin İlkeleri adlı çalışmasıyla Sheakespeare'in Hamlet, Plato'nun Devlet, Beethoven'ın Beşinci Senfoni ya da Da Vinci'nin Son Akşam Yemeği adlı çalışmaları aynı kabul edilemez. Bunun en önemli nedenlerinden biri Newton'un eserinin anlaşılabilmesi için bazen belli bir bilgi birikimi bile yetersiz kalırken, diğer eserlerin okuma-yazma bilen herhangi bir birey tarafından alınıp tüketilebilecek

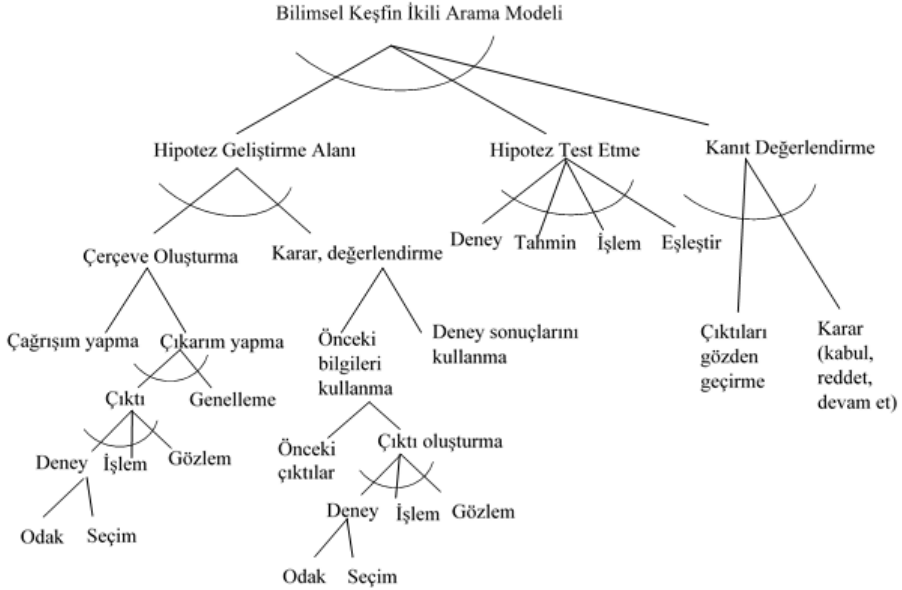
ve anlaşılabilir düzeyde olmasıdır. Bu doğrultuda Simonton bilimsel yaratıcılığın mantık, deha, şans ve zamanın ruhu (zeitgeist) olmak üzere dört temel bakış açısıyla ele alınabileceği fikrini ortaya koymuştur. Simonton'a göre bilim felsefesinin bilimsel yaratıcılığa yönelik mantıksal açıklama getirme eğilimleri yeni bir durum değildir. Örneğin Bacon'ın tümevarımsal akıl yürütme, Descartes'in ise tümdengelsel akıl yürütme yaklaşımlarıyla bilimsel yaratıcılığı ele aldıkları görülmüştür. Bununla birlikte Karl Popper, Albert Einstein gibi ünlü bilim insanları mantığın sınırlarını bilimsel yaratıcılığı açıklamak için yeterli görmemektedirler. Onlara göre yaratıcılık için canlı ve gösterişli bir hayal dünyasına, güçlü sezgilere ihtiyaç vardır. Bu görüş dehanın bilimsel yaratıcılıktaki önemini vurgulamaktadır. Kimi bilim insanları da bilimsel yaratıcılık süreçlerini düzensiz, tahmin edilemez ve kaotik olarak ele alarak yaratıcılıktaki şans faktörünü vurgulamaktadırlar. Yaratıcı ürünlerin sosyo-kültürel çevrenin bir ürünü olduğunu ileri süren ve bilimsel yaratıcılığı zamanın ruhu ile ilişkilendirerek ele alan bilim insanlarının da sayısı oldukça fazladır.

BİLİMSEL YARATICILIK KURAMLARI

Bilimsel yaratıcılığın doğasını anlamada araştırmacılar farklı yaklaşımları benimsemişlerdir. Kimi araştırmacılar (Rocke, 2010; Rothenberg, 1995) bilim insanlarının yaşamlarını incelemiş, kimi araştırmacılar (Jang ve Ko, 2017) bilim insanlarını laboratuvarında gözlemlemiş, kimi araştırmacılar (Klahr, 2000) psikoloji laboratuvarlarında öğrencilerin bilimsel keşif süreçlerini değerlendirmiş, kimi araştırmacılar (Hu ve Adey, 2002) ise bilimsel yaratıcılığı genel yaratıcılığa yönelik kuramlarla açıklamışlardır. Bu nedenle bilimsel yaratıcılığa yönelik farklı kuramlar geliştirilmiştir. Bu bölümde bilimsel yaratıcılığın doğasını anlamak için alanyazında ön plana çıkan ve bilimsel yaratıcılığın ölçümüne dayanak olan bilimsel yaratıcılık kuramlarına yer verilmiştir.

Bilimsel Keşfin İkili Arama Modeli

Klahr ve Dunbar (1988) bilimsel yaratıcılık sürecindeki yapıyı ortaya koymak amacıyla Bilimsel Keşfin İkili Arama Modeli'ni önermişlerdir. Modele göre bilimsel yaratıcılık deney ve hipotez alanlarındaki arayışlarla ortaya çıkmaktadır. Hipotez geliştirme alanındaki arayışlar önceki bilgiler ve deney sonuçlarından, deney alanındaki arayışlar ise geliştirilen hipotezlerden etkilenmektedir (Klahr, 2000). Klahr bilimsel yaratıcılığın ikili arama sürecinin hipotez geliştirme, deney tasarlama ve kanıt değerlendirme olmak üzere üç bileşenden oluştuğunu belirtmektedir. Modelin bileşenleri ve genel yapısı Şekil 2.1'de görselleştirilmiştir.



Şekil 2.1. Bilimsel Keşfin İkili Arama Modeli Bileşenleri (Klahr, 2000, s. 37)

Şekil 2.1'de görüldüğü üzere hipotez geliştirme alanı çerçeve oluşturma ve karar-değerlendirme olmak üzere iki bileşenden oluşmaktadır. İlk bileşen yeni bir hipotezin kapsamını oluşturmaktadır. İkinci bileşen ise yeni hipotezin iyileştirme sürecini temsil etmektedir. Çerçeve oluşturma, bir konunun ana hatlarını belirleme amacıyla var olan bilgilerin değerlendirildiği zihinsel bir süreçtir. Bu aşamada daha önceki bilgiler önemli bir rol oynar. Çünkü daha önceki bilgilerle çağrışım yapılarak yeni bir hipotez geliştirilir. Hipotez geliştirme yani çerçeve oluşturma süreci sadece çağrışımlar sonucunda gerçekleşmemektedir. Çıkarım yaparak da hipotez geliştirilebilir. Çıkarım yapma boyutu çıktı oluşturma ve genellemeden meydana gelmektedir. Çıktı oluşturma deney, işlem ya da gözlem sonuçlarının değerlendirilmesi sürecidir. Deney, odak noktaların belirlenmesi ve odak noktalara göre seçim yapılması sürecini kapsamaktadır. İşlem deneyin gerçekleştirilmesi, gözlem ise deney sonuçlarının değerlendirilmesini temsil etmektedir. Çerçeve oluşturmaya meydana getiren genelleme alt boyutu, oluşturulan çıktıların genel yargılara dönüştürülmesidir. Bu iki alt boyutun birbirini takip etmesinden sonra hipotez üretilmektedir.

Hipotez geliştirme alanının ikinci bileşeni olan karar değerlendirme süreci önceki bilgilerin ya da deney çıktılarının kullanılması ile gerçekleşmektedir. Bu bileşen sayesinde hipotezler geliştirilmektedir. Çerçeve oluşturma bileşeni gibi iki boyuttan oluşmaktadır. Çerçeve oluşturmada önceki bilgilerin kullanımı önemli bir rol oynarken, karar değerlendirme sürecinde deney çıktıları daha

etkilidir. Karar değerlendirme bileşeninde deney çıktıları incelenirken daha önce yapılmış deneylerin sonuçları ve gerçekleştirilen deney sonuçları bir arada değerlendirilerek üretilen hipotezde düzenleme yapılmaktadır (Klahr, 2000).

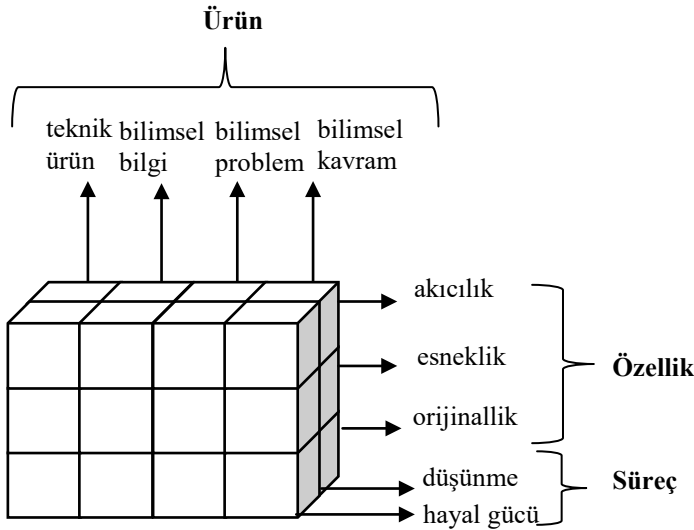
Rachelson (1977) hipotez geliştirmeyi bir problemin çözümüne yönelik açıklamaların üretildiği zihinsel bir süreç olarak tanımlamaktadır. Bilişsel psikoloji alanında çalışan uzmanlar analogi haritaları oluşturma (Gentner, 1983; Gick ve Holyoak, 1993), buluşsal arama (Kaplan ve Simon, 1990; Klahr ve Dunbar, 1988), hazırlama (Schunn ve Dunbar, 1996), hatırlama (Ross, 1984) ve kavramsal kombinasyonlar (Shrager, 1987) gibi yollarla hipotez üretildiğini iddia etmektedirler. Hipotez geliştirme bilimsel bir problemin çözümü ya da bilimsel buluşların ortaya çıkma sürecinin ilk adımı olması nedeniyle bilimsel yaratıcılığın önemli bileşenlerinden biri olarak kabul edilmektedir (Dunbar, 1997; Heller, 2007).

Hipotez test etme alanı deney, tahmin, işlem ve eşleştirme olmak üzere dört bileşenden oluşmaktadır. İlk bileşen deney tasarlamadır. Üretilen hipotez ve bu hipoteze yönelik tasarlanan deneyin sonuçları ikinci bileşenle tahmin edilmektedir. Üçüncü bileşen olan işlem, tasarlanan deneyin yapılmasıdır. Eşleştirme ise tahminlerle deney sonuçlarının karşılaştırıldığı son bileşendir (Klahr, 2000). Simon'a göre (1989) deney tasarlamasının amacı geliştirilen hipotezlerin doğruluğunu sınamaktır. Deney tasarlama sürecinde hipotezi test etmeye yönelik kanıtlar gözlenmektedir. Bu süreçte önceki deneyimlerin etkisinin en düşük düzeyde olması beklenmektedir (Zimmerman, 2000). Beş yaşından itibaren çocuklarda deney tasarlama becerilerinin olduğu görülmektedir. Bu yaş grubunun deney tasarlama becerilerinde önceki deneyimlerinin etkisi oldukça düşüktür (Sodian, Zaitchik ve Carey, 2000).

Kanıt değerlendirme, deney çıktılarını değerlendirme ve karar olmak üzere iki bileşenden oluşmaktadır. Deney sonuçları ve bu süreçte toplanan veriler değerlendirilerek hipotez ya kabul edilir ya da reddedilir. Değerlendirmelerde bir tutarsızlık olduğu görülürse deney süreci tekrar edilir. Bu karar sürecinde akla uygunluk, işlevsellik ve sadelik gibi ölçütler göz önünde bulundurulmaktadır. Bu ölçütlerin kanıt değerlendirmede etkili olduğu bilinse de karar sürecinin nasıl gerçekleştiği tam olarak bilinmemektedir (Klahr, 2000). Klahr ve Dunbar (1988), kanıt değerlendirme becerisinin hipotez ve deney alanları arasındaki bağlantıyı sağladığını ileri sürmektedirler. Araştırmacılara göre bilimsel yaratıcılığı yüksek düzeyde olan bilim insanları tüm kanıtları değerlendirir ve basit gerçekleri ortaya çıkarır. Modele göre bilimsel yaratıcılık deney ve hipotez alanlarındaki aramaların sonucunda ortaya çıkmaktadır. Bu iki alandaki aramaların sonuçlanabilmesi için hipotez geliştirme, deney tasarlama ve kanıt değerlendirme bileşenlerinin bir uyum içerisinde olması gerekmektedir.

Bilimsel Yaratıcılığın Yapı Modeli

Hu ve Adey (2002), alanyazında yaratıcılığa yönelik araştırmaları incelemişler ve bilimsel yaratıcılığın ölçülmesinde bir temel oluşturmak amacıyla Bilimsel Yaratıcılığın Yapı Modeli'ni önermişlerdir. Model kapsamında bilimsel yaratıcılık zihinsel bir özellik, orijinal ve toplumsal değeri olan bir ürün üretme yeteneği olarak ele alınmaktadır. Modelin temelini Guilford'un Zihnin Yapı Kuramı ve yaratıcılığa yönelik araştırmalar oluşturmaktadır (Chin ve Siew, 2005). Hu ve Adey (2002, s. 392) önerdikleri model kapsamında bilimsel yaratıcılığın yapısını açıklarken çeşitli varsayımlarda bulunmuşlardır. Bunların birincisi bilimsel yaratıcılığın diğer yaratıcılık türlerinden farklı olmasıdır. Bu varsayıma neden olarak bilimsel yaratıcılığın deney, problem oluşturma, problem çözme ve yaratıcı bilimsel etkinlikleri kapsamı gösterilmiştir. İkincisi bilimsel yaratıcılığın bir çeşit yetenek olarak kabul edilmesidir. Bilimsel yaratıcılığın yapısında bilişsel olmayan faktörler bulunmamaktadır ancak bilişsel olmayan faktörler bilimsel yaratıcılığı etkilemektedir. Üçüncüsü bilimsel yaratıcılığın bilgi ve yeteneğe bağlı olmasıdır. Dördüncüsü bilimsel yaratıcılığın statik ve dinamik bir yapıya sahip olmasıdır. Sonuncusu ise yaratıcılığın ve analitik zekânın bilişsel yetenek olduğu ancak ikisinin birbirinden farklı olduğudur. Araştırmacılar bu varsayımlarından yola çıkarak bilimsel yaratıcılığın yapısını ürün, özellik ve süreç olmak üzere üç boyut ile açıklamışlardır. Modeli resmetmek için Guilford'un Zihnin Yapı Kuramı'ndakine benzer bir küp şeklinden yararlanılmışlardır (Şekil 2.2).



Şekil 2.2. Bilimsel Yaratıcılığın Yapı Modeli (Hu ve Adey, 2002, s.391)

Şekil 2.2’de yer alan küpün her bir ekseninde ürün çeşitleri (teknik ürün, bilimsel bilgi, bilimsel kavram ve bilimsel problem), özellikler (akıcılık, esneklik ve orijinallik) ve süreç çeşitleri (düşünme ve hayal gücü) yer almaktadır. Küp, hemen hemen her biri birbirinden farklı bir yeteneği ya da yapıyı ifade eden 24 (4 x3x2) hücreden meydana gelmektedir. Bilimsel yaratıcılığın yapısında yer alan ürün çeşitleri bilimsel bir probleme ya da duruma yönelik çözümleri içermektedir (Chin ve Siew, 2005). Bilimsel yaratıcılığın yapısında yer alan özellikler boyutunda genel yaratıcılığın bileşenleri, süreç boyutunda ise düşünme ve hayal gücü yer almaktadır (Hu ve Adey, 2002). Düşünme yeni ürün ya da fikir üretme sürecini kapsamaktadır. Hayal gücü ise yeni fikir üretme yeteneğini temsil etmektedir (Chin ve Siew, 2005).

Yaratıcı Bilimsel Çağrışımlar Modeli

Kanlı (2014a, 2014b) bilimsel yaratıcılığın düşünme sürecini ortaya koymak amacıyla Yaratıcı Bilimsel Çağrışımlar Modeli’ni önermiştir. Modelin temelini, Mednick’in Çağrışımsal Düşünme Teorisi ve Sternberg ve Dunbar’ın analogi üzerine gerçekleştirdikleri araştırmalar oluşturmaktadır (Kanlı, 2017). Model kapsamında bilimsel yaratıcılık, alan bilgisinin çağrışımsal ve analogik düşünme sürecinden geçerek yeni düşünceye, ürüne ya da çözüme dönüşmesi olarak tanımlanmaktadır (Kanlı, 2014a, s.73).

Modele göre alan bilgisi, bilimsel yaratıcılığın ön koşuludur. Bu nedenle bilimsel yaratıcılık süreci bilginin üzerine inşa edilmiştir. Modele göre bilimsel yaratıcı düşünme sürecinde bir probleme yönelik çağrışım ve analogi kurularak aracılık ve benzerlik yoluyla çözümler oluşturulmaktadır (Kanlı, 2014a, 2014b, 2017). Model bilimsel yaratıcı düşünme sürecini ortaya koyan özgün bir modeldir. Modelin temelinin bilimsel araştırma bulgularına dayanması modeli güçlü kılmaktadır.

BİLİMSEL YARATICILIK VE BİLGİ

Alanyazında zengin bilgi birikiminin bilimsel yaratıcılığı artırıcı bir etkiye sahip olabileceği vurgulanmıştır (Klahr, 2000; Liang, 2002). Bilimsel yaratıcılık ve bilgi arasındaki ilişkiyi araştıran sınırlı sayıda çalışma bulunmaktadır (Kang, Park ve Hong, 2015). Diğer taraftan pek çok araştırmacı yaratıcılık ve bilgi arasındaki ilişkiyi sorgulamıştır (Weisberg, 2009). Yaratıcılık ve bilgi arasındaki ilişkiyi araştıran çalışmaları incelemek, bilimsel yaratıcılık ve bilgi arasındaki ilişkinin anlaşılmasına yönelik bir zemin oluşturacaktır.

Yaratıcılık ve bilgi arasındaki ilişkinin araştırılmasının nedenlerinden biri okullardaki çocukların sahip oldukları yaratıcılık potansiyellerinin geliştirilmesine yönelik olarak uygulanan programların sayısının artmasıdır. Eğitim programlarında yaratıcılık kavramına yer verilmesi bilginin yaratıcılıkla

olan ilişkisinin sorgulanmasının önünü açmıştır çünkü bilgi anlamının yapı taşlarından biri olarak görülmektedir (Craft, 2005).

Kimi araştırmacılar yaratıcı düşünme ve bilgi arasındaki ilişkiyi açıklamak için çeşitli varsayımlarda bulunmuşlardır. Önemli varsayımlardan biri bilgi ve yaratıcılık arasında açık ya da örtük bir gerilim olduğudur. Alana ilişkin bilgi birikiminin yaratıcılık için bir aracı olduğu fikri yaygın olarak kabul görmektedir. Bazen çok fazla bilgi birikimi, yaratıcılığı engelleyebilir. Bu ilişki ters-U grafiğiyle gösterilmektedir (Weisberg, 2009). Thomas Kuhn (1970) birikimli bilim insanlarının, gençlere ve birikimi daha az olan meslektaşlarına göre devrimsel nitelikte keşifler yapabileme şanslarının daha düşük olduğunu ortaya koymuştur (akt. Frensch ve Sternberg, 1989).

Yaratıcılık ve bilgi arasındaki gerilim varsayımına karşı olarak bilgi birikiminin yaratıcılığı pozitif yönlü olarak etkilediğini savunan bir görüş daha ortaya çıkmıştır. Bu görüş temel görüş (foundation view) olarak da adlandırılmaktadır. Temel görüşe göre, yaratıcılık ve bilgi arasındaki ilişki oldukça açıktır. Sternberg (2010) bir alanda ilerleyebilmek ya da alanı ileri götürebilmek için öncelikle yeterli bilgi birikimine sahip olunması gerektiğini öne sürmektedir. Bu görüşe göre etkili ve ses getirecek buluşların ortaya çıkabilmesi için belli bir bilgi birikiminin olması zorunludur (Weisberg, 2009). Polaroid fotoğraf makinasının mucidi olan Land “şans ya da ilham ile değil bilgi birikimiyle çalışmalarının başarıya ulaştığını” ifade ederek temel görüşü desteklemiştir (Copley, 1999, s. 512).

Pek çok araştırmacı temel görüşü desteklemektedir. Bu araştırmacılar yaratıcılık sürecinde bilgi ediniminin önemine vurgu yapmışlar ve bireylerin bilgilerinin yetersiz olduğu alanlarda yaratıcılık gösteremeyeceklerini belirtmişlerdir (örn. Baer, 2012; Boden, 2001; Craft, 2005; Csikszentmihalyi, 1996; Feldhusen, 1995; Gero ve Maher, 2013; Glaser, 1984; Hayes, 1989; Ivcevic, 2007; Kulkarni ve Simon, 1988). Bu bağlamda, alana özgü bilgi birikimi yaratıcı çalışmaların ön koşullarından biri olarak kabul edilmektedir (Weisberg, 2009). Amabile’ın Üç Bileşenli Yaratıcılık Modeli’nde ve Csikszentmihalyi’nin Sistemler Modeli’nde yaratıcı performansın ortaya çıkabilmesi için en önemli gerekliliklerden birinin alana özgü bilgi birikiminin olduğu savunulmaktadır (Cohen ve Ambrose, 1999; Starko, 2017). Sonuç olarak yaratıcılık ve bilgi arasındaki ilişkiye yönelik farklı fikirler ortaya atılmıştır. Bu görüşler doğrultusunda yaratıcı performansın sergilenebilmesi için asgari düzeyde bilgi birikimine gereksinim duyulduğu ifade edilebilir.

BİLİMSEL YARATICILIK VE CİNSİYET FARKLILIKLARI

Kadınların erkeklere kıyasla sorun çözme alanlarında daha yaratıcı olarak kabul edilmesi, yaratıcılık alanında cinsiyet farklılıkları üzerine odaklanan çok

fazla sayıda çalışma yapılmasının temel nedenlerinden biri olmuştur (Maccoby, 1996). Ancak yapılan araştırmaların önemli bir kısmında cinsiyet değişkeninin yaratıcılık alanında belirleyici bir faktör olmadığı sonucu elde edilmiştir. Örneğin Baer ve Kaufman (2008) tarafından yapılan çalışmada yaratıcılık alanındaki cinsiyet farklılıklarını inceleyen araştırmalar üzerinde durulmuştur. Toplam 78 araştırmanın ele alındığı çalışmanın sonucunda 35 araştırmada cinsiyet farklılığını destekler verilerin ortaya çıkmadığı bulunmuştur. Pek çok araştırmada (örn. Ayyıldız Potur ve Barkul, 2009; Chan, 2005; Kaufman, Baer ve Gentile, 2004; Lee, 2002) akıcılık, orijinallik, esneklik, başlıkların soyutluğu, zenginleştirme ve erken kapanmaya direnç gibi yaratıcılık bileşenlerine yönelik cinsiyet farklılıkları üzerinde durulmuş ve söz konusu yaratıcılık bileşenlerinin hiçbirinde cinsiyet bağlamında istatistiksel olarak anlamlı bir farklılık görülmediği ortaya koyulmuştur. Diğer taraftan, Matud, Rodriguez ve Grande (2007) tarafından yapılan bir araştırmada kızların akıcılık bileşeni anlamında erkeklerden daha iyi olduğu sonucu elde edilmiştir. Stoltzfus, vd. (2011) çalışmalarında ise akıcılık ve esneklik gibi yaratıcılık bileşenleri anlamında cinsiyet farklılıkları olduğu sonucu elde edilmiştir. Çalışmada kızlara göre erkeklerin akıcılık ve esneklikte daha iyi oldukları sonucu raporlanmıştır. Alanda yapılan çalışmalardan elde edilen sonuçlar arasındaki farklılık, yaratıcılık alanında cinsiyetin farklılık yaratan bir faktör olup olmadığını sorgulanır hale getirmektedir (Özdemir ve Sak, 2013, s. 56-57).

Bilimin tarihine bakıldığında bilimsel ve teknik konularla daha çok erkeklerin ilgilendiği ya da önemli icatların erkekler tarafından yapıldığı görülmektedir. Bu durum bilimsel yaratıcılık anlamında da cinsiyet farklılıkları olabileceği düşüncesini ortaya çıkarmıştır. Ancak araştırma bulguları bu iddiayı desteklememektedir. Araştırmalarda erkek öğrencilerin akıcılık, esneklik, orijinallik puan ortalamalarının kız öğrencilerin puan ortalamalarından daha yüksek olduğunu ancak aradaki farkın anlamlı olmadığı ortaya konulmuştur (Ayverdi, vd., 2012; Baysal, Kara ve Üçüncü, 2013; Kanlı, 2017; Mohamed, 2006; Sansanwal ve Sharma, 1993; Shukla ve Sharma, 1986). Hu, vd. (2010) öğrencilerin açık ve kapalı uçlu bilimsel problem oluşturma becerisini incelemişler ve öğrencilerin oluşturdukları problemleri akıcılık, esneklik ve orijinallik ölçütleri doğrultusunda değerlendirilmişlerdir. Çalışmada bilimsel problem oluşturma becerisinde cinsiyet değişkeni bakımından bir farklılık oluşturmadığı raporlanmıştır. Alanyazında bilimsel yaratıcılıktaki cinsiyet farklılıkları bilimsel buluş ve yayınlarda cinsiyet faktörünü inceleyen çalışmalarda ve fen bilimleri derslerindeki akademik başarıda cinsiyet rolünü ele alan çalışmalarda da araştırılmaktadır (Özdemir, 2013).

Bilim insanların bilimsel buluş ve yayınlar anlamındaki üretkenliklerinin bilimin temelini oluşturması (Fox, 2005), bilim insanları tarafından yapılan

bilimsel buluş ve yayın sayısının önemini artırmaktadır. Bilimsel yayınlardaki üretkenlik anlamında da erkekler ve kadınlar arasında bir farklılık olmadığı sonucunu ortaya koyan birçok çalışma yapılmıştır. Örneğin Reskin (1978) kimya alanında doktorasını yapmış kadın ve erkeklerin yayın sayılarında istatistiksel olarak anlamlı farklılıklar olmadığını ortaya koymuştur. Sax, vd. (2002) yapmış oldukları araştırmada kadın ve erkekler arasında bilimsel yayınlar anlamında var olan farklılığın giderek eşit bir seviyeye geldiğini rapor etmişlerdir. Diğer taraftan, erkeklerin bilimsel yayınlarının anlamlı bir şekilde daha fazla olduğunu ortaya koyan çalışmalar da bulunmaktadır (Astin ve Bayer, 1979; Cole, 1979; Cole ve Cole, 1973; Cole ve Zuckerman, 1987; Long ve Fox, 1995).

Bilimsel yaratıcılık alanındaki cinsiyet farklılıklarını incelemek için kimi araştırmacılar fen bilimleri derslerine yönelmiştir (Özdemir, 2013). Kimi çalışmalarda (Erickson ve Erickson, 1984; Erickson ve Farkas, 1991; Johnson, 1987; Kelly, 1988) erkek öğrencilerin kız öğrencilere göre fen bilimlerinde daha başarılı olduğu bulunmuştur. Ancak kimi araştırmacılar bu iddiaya karşı çıkmış ve fen bilimleri derslerinde cinsiyete göre tutarlı bir başarı farklılığının görülmediğini ortaya koymuşlardır (Adamson, vd., 1998; Mattern ve Schau, 2002; Nowell ve Hedges, 1998; Shafiq, 2013).

BİLİMSEL YARATICILIĞIN ÖLÇÜLMESİ

Bilimsel yaratıcılığın ölçülmesine yönelik çalışmalar 1960'lı yıllarda başlamıştır. 1970'li ve 1980'li yıllarda ilk bilimsel yaratıcılık testleri geliştirilmiştir. Bu yıllarda bilimsel yaratıcılığın öz değerlendirme gibi farklı yöntemler ile değerlendirildiği de görülmektedir. 1990'lı ve 2000'li yıllarda ise bilimsel yaratıcılığın nasıl ölçüleceğine yönelik kuramsal tartışmalarla birlikte eğitim ve psikoloji alanında bilimsel yaratıcılığın önemine vurgu yapılmıştır. Bu süreçle beraber bilimsel yaratıcılığı ölçme çalışmaları hız kazanmıştır. Bilimsel yaratıcılığın ölçülmesine yönelik alanyazında var olan yöntemler kâğıt-kalem testleri ile değerlendirme, öz-değerlendirme ve başkaları tarafından değerlendirilme, konsensüse dayalı değerlendirme ve performans temelli değerlendirme olmak üzere dört grupta incelenebilir. Bu yöntemleri ele almadan önce ilk geliştirilen bilimsel yaratıcılık testlerine kısaca yer verilerek bilimsel yaratıcılığın ölçümüne yönelik genel bir bakış sunulmuştur.

İlk Geliştirilen Bilimsel Yaratıcılık Testleri

Bilimsel yaratıcılığın ölçülmesine yönelik ilk testlerden biri Majumdar (1975) tarafından geliştirilmiştir. Majumdar, Guilford'un Zihinsel Yapı Modeli'ni temel alarak fizik, kimya, biyoloji ve matematik alanlarında problem çözme süreçlerindeki bilimsel yaratıcılığı değerlendirmek için 29 maddelik Hint kültürüne özgü bir test geliştirmiştir. Testin hedef kitesini ortaokul ve

lise öğrencileri oluşturmaktadır. Sing 1981 yılında Majumdar'ın bilimsel yaratıcılık testinin 10. sınıflara özgü formunu tasarlamıştır. Misra ise 1986 yılında altıncı ve yedinci sınıf öğrencilerinin bilimsel yaratıcılığını ölçmek için bir çoğul düşünme testi geliştirmiştir (Mukhopadhyay, 2013). Sharma ve Shukla, 1986 yılında ortaokul öğrencilerinin bilimsel yaratıcılıklarını ölçmeye yönelik 12 maddeden oluşan sözel bir test geliştirmişlerdir. Test Hint kültürüne özgüdür (Pestonjee, 1997). Sinha ve Singh (1987) ortaokul öğrencilerinin bilimsel yaratıcılık düzeylerini ölçmek için bir yaratıcılık testi geliştirmişlerdir. Esneklik, özgünlük, gözlem yeteneği, hayal gücü, analiz etme becerisi ve dönüştürme yeteneklerini ölçmeye yönelik oluşturulan test, İngilizce ve Hinduca 84 maddeden oluşmaktadır. Teste yönelik herhangi bir geçerlik ve güvenilirlik çalışması yapılmamıştır. Gupta 1988'de lise öğrencilerine yönelik 30 maddelik bir çoğul düşünme testi geliştirmiştir (Mohamed, 2006; Sharma ve Mahrshi, 2017).

Hint kültürüne özgü testlerin dışında da testlere rastlanmaktadır. Örneğin Friendlar (1983) lise öğrencilerinin bilimsel yaratıcılıklarını ölçmek amacıyla bir bilimsel yaratıcılık testi geliştirmiştir. Friendlar veri toplama, problem çözüme, hipotez geliştirme ve deney tasarlama gibi becerileri değerlendirmeye yönelik açık uçlu maddeler tasarlamış ve tasarladığı maddeleri 143 lise öğrencisine uygulayarak bir araştırma gerçekleştirmiştir. Araştırma sonucunda, testin test-tekrar test güvenilirliği düşük bulunmuş ancak bilimsel beceriler arasında istatistiksel olarak anlamlı ilişkiler saptanmıştır (Liang, 2002). Frederiksen ve Ward (1978) bilim insanların keşif sürecinde sıklıkla kullandığı hipotez oluşturma, araştırma tasarlama ve araştırma sonuçlarını yorumlama olmak üzere üç beceriye odaklanarak lisans düzeyine yönelik bir bilimsel yaratıcılık testi geliştirmişlerdir. Becerileri ölçmek için psikoloji alanındaki kurgusal araştırma tasarımları ve bulguları kullanılmıştır. Örneğin hipotez geliştirme becerisini ölçmek için katılımcılara yaşları 8 ile 10 arasında değişen 822 çocuğun doğum kilosu ve IQ puanlarının ortalamaları verilmiş ve doğum kilosu fazla olan çocukların IQ puanlarının daha yüksek olduğu bulgusu paylaşılmıştır. Katılımcılardan bu bulguya yönelik hipotez üretmeleri ve en iyi ürettiği hipotezi seçmesi beklenmiştir. Uygulama süresi 45 dakika olan testin puanlamasında cevapların niteliği, sayısı ve sıradanlığı dikkate alınmaktadır. Ölçek 3500 adaya uygulanarak çeşitli analizler yapılmış ve ölçeğin yeterli düzeyde geçerli ve güvenilir olduğu ortaya koyulmuştur.

İlk bilimsel yaratıcılık testleri incelendiğinde testlerin ortaokul, lise ve lisans düzeyinde olduğu görülmektedir. Testlerin çoğunluğunun kültürel özellikler taşıdığı, geçerlik ve güvenilirlik çalışmalarının yeterli düzeyde yapılmadığı öne sürülmektedir (Hu ve Adey, 2002). Mohamed (2006) bu testlerin maddelerinin ilgili yaş grubuna uygun olmadığını belirtmektedir. İlk bilimsel yaratıcılık

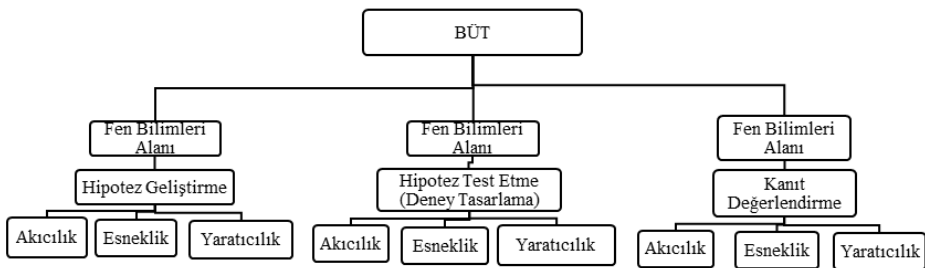
testlerinin sınırlılıklarının olduğu ifade edilebilir. Ancak bu testler güncel test geliştirme çalışmalarına bir zemin oluşturmaktadır.

Bilimsel Yaratıcılığı Kâğıt-Kalem Testleri ile Değerlendirme Yöntemi

Bilimsel yaratıcılığın ölçülmesinde yaygın olarak kullanılan yöntemlerden biri kâğıt kalem testleridir. Bu testlerde genellikle çoğul düşünme becerilerine odaklanmaktadır. Kimi araştırmacılar bilimsel yaratıcılığın ölçümünde hem tekil hem de çoğul düşünme becerilerinin göz önünde bulundurulması gerektiğini iddia etmektedir (Lubart, Zenasni ve Barbot, 2013; Yang, vd., 2016). Çoğul düşünme bir duruma ya da probleme yönelik olarak çok sayıda çözüm üretme sürecidir (Guilford, 1956, s. 274). Çoğul düşünme genellikle açık uçlu problemlerle değerlendirilmektedir. Sak ve Maker (2005) açık uçlu problemi birden fazla doğru yanıtı olan problem türü olarak tanımlamaktadır. Tekil düşünme ise bir duruma ya da probleme yönelik en iyi çözümü bulmadır (Guilford, 1956, s. 274). Bilimsel yaratıcılığın ölçümünde çoğul düşünme becerilerine odaklanan testlere örnek olarak Bilimsel Üretkenlik Testi (BÜT), Bilimsel Yaratıcılığın Yapı Modeli'ne göre geliştirilen testler ve Yaratıcı Bilimsel Çağrışımlar Testi verilebilir. Çoğul ve tekil düşünme becerilerine odaklanan testlere ise Çoğul-Tekil Bilimsel Yaratıcılık Testi örnek olarak verilebilir.

Bilimsel Üretkenlik Testi (BÜT)

BÜT altıncı, yedinci ve sekizinci sınıf öğrencilerinin fen bilimleri alanındaki yaratıcılıklarını ölçmek amacıyla geliştirilmiştir. BÜT grup olarak uygulanan bir kâğıt kalem testidir. Açık uçlu maddelerden oluşan testin uygulama süresi 40 dakikadır. Çoğul düşünme becerilerine odaklanan çalışmalar ve Bilimsel Keşfin İkili Arama Modeli temel alınarak geliştirilen BÜT üç boyuttan oluşmaktadır. İlk boyutta genel yaratıcılık becerileri, ikinci boyutta fen bilimlerine özgü beceriler, üçüncü boyutta ise fen bilimlerindeki alanlar yer almaktadır (Ayas, 2010; Ayas ve Sak, 2014; Sak ve Ayas, 2013). BÜT'ün tüm boyutları Şekil 2.3'te yer almaktadır.



Şekil 2.3. BÜT'ün Kuramsal Yapısı (Ayas ve Sak, 2014, s.198)

Şekil 2.3 incelendiğinde fen bilimleri alanlarının, fen bilimlerine özgü becerilerin ve genel yaratıcılık becerilerinin sentezlendiği görülmektedir. Fen bilimleri alanları fizik, kimya, biyoloji, ekoloji ve disiplinler arası alanları kapsamaktadır. Fen bilimlerine özgü beceriler hipotez geliştirme, deney tasarlama (hipotez oluşturma) ve kanıt değerlendirmedir. Genel yaratıcılık becerileri ise akıcılık, esneklik ve yaratıcılık şeklinde sınıflandırılmaktadır (Ayas ve Sak, 2014; Sak, 2010).

Bir çoğul düşünme testi olan BÜT beş alt testten oluşmaktadır. Birinci alt test “Sinek Deneyi”dir. Bu alt testte biyoloji alanında bir deney düzeneği görsel olarak sunulmaktadır. Öğrencilerden bu deney düzeneğine ilişkin üretebilecekleri kadar çok hipotez üretmeleri beklenmektedir. İkinci alt test olan “Etkileşim Grafiği”nde, öğrencilere disiplinler arası alanda bir grafik sunulmakta ve öğrencilerden bu grafiğe yönelik üretebilecekleri kadar çok hipotez üretmeleri beklenmektedir. Üçüncü alt test “Şeker Deneyi”dir. Bu alt testte öğrencilere kimya alanında bir deney düzeneği ve bir hipotez sunulmaktadır. Öğrencilerden hipotezi kanıtlayabilmek için deney düzeneğinde yapabilecekleri kadar çok değişiklik yapmaları beklenmektedir. “Yay Deneyi” dördüncü alt testtir. Bu alt testte fizik alanında bir deney düzeneği ve bu düzeneğe ilişkin bir amaç sunulmaktadır. Öğrencilerden amaca ulaşabilmek için deney düzeneğinde yapabilecekleri kadar çok değişiklik yapmaları beklenmektedir. “Besin Zinciri” son alt testtir. Bu alt testte öğrencilere ekoloji alanından bir besin ağı ve besin ağında yaşanan bir değişime yönelik bir grafik sunulmaktadır. Öğrencilerden bu değişime neden olabilecek kanıtlar bulmaları beklenmektedir (Sak ve Ayas, 2009, 2013). Hipotez geliştirme becerisini ölçmek için iki alt test, deney tasarlama becerisini ölçmek için iki alt test, kanıt değerlendirme becerisini ölçmek için ise bir alt test kullanılmaktadır.

BÜT’ün puanlaması akıcılık, esneklik ve yaratıcılık puanları ile yapılmaktadır. Akıcılık puanı üretilen doğru yanıtların sayısıdır. Esneklik puanı ise üretilen yanıtlardaki kategori sayısıdır. Yaratıcılık puanı olarak Snyder ve diğerlerinin (2004, s. 416) önerdiği formül (Creativity Quotient, CQ) kullanılmaktadır. Puan $\log_2 [(1+ u_1) \cdot (1+ u_2) \dots (1+ u_c)]$ formülü ile elde edilmektedir. Formülde yer alan 1,2... c kategorileri; “u” ise bir kategoride üretilen doğru cevap sayısını temsil etmektedir. Toplam akıcılık, tüm alt testlerden elde edilen akıcılık puanlarının toplanması ile elde edilmektedir. Toplam esneklik ve toplam yaratıcılık puanları da benzer şekilde hesaplanmaktadır (Ayas, 2010; Ayas ve Sak, 2014).

BÜT’ün geliştirilmesi ve psikometrik özelliklerinin incelenmesi kapsamında çeşitli araştırmalar gerçekleştirilmiştir. BÜT’ün geliştirilmesi kapsamında ilk olarak çoğul düşünme ve bilimsel düşünme becerilerine odaklanan çalışmalar incelenerek testin kuramsal yapısı ve kuramsal yapıya uygun madde

havuzu oluşturulmuştur. Fen bilimleri alanında çalışan uzmanların görüşleri doğrultusunda havuzdan beş madde seçilmiştir. Seçilen maddeler fizik, kimya ve biyoloji öğretmenlerinin görüşlerine sunulmuş ve gerekli düzenlemeler yapılarak test formu oluşturulmuştur (Ayas, 2010; Sak ve Ayas, 2013). Sonrasında üstün yetenekli olarak tanılanan 71 öğrenci (40 altıncı sınıf, 31 yedinci sınıf) ile ilk pilot çalışma gerçekleştirilmiştir (Ayas ve Sak, 2008). Bu çalışma kapsamında ölçeğin iç tutarlılığı için hesaplanan Cronbach Alfa değeri .76 olarak bulunmuştur. Alt test puanları ile toplam test puanları arasındaki korelasyon katsayılarının .50 ile .61 arasında değiştiği görülmüştür ($p < .01$). Puanlayıcılar arası güvenilirlik katsayılarının .91 ile .97 arasında değiştiği rapor edilmiştir. Yedinci sınıf öğrencilerinin toplam yaratıcılık puan ortalamalarının altıncı sınıfların puan ortalamalarına göre anlamlı bir şekilde yüksek olduğu sonucu bulunmuştur ($F = 3.75, p < .05, \eta^2 = .05$). Birinci pilot çalışmanın sonunda test ile ilgili revizyonlar yapılmış ve Üstün Yeteneklilerin Eğitimi Uygulama ve Araştırma Merkezi'ne (ÜYEP) başvuran 128 altıncı sınıf öğrencisi ile ikinci pilot çalışma gerçekleştirilmiştir (Sak ve Ayas, 2009). İkinci pilot çalışmada ölçeğin iç tutarlılığı için hesaplanan Cronbach Alfa değeri .81 olarak bulunmuştur. Birinci pilot çalışmaya göre bu değer yükseldiği görülmektedir. Alt test puanları ile toplam test puanları arasında korelasyon katsayılarının .56 ile .74 arasında değiştiği görülmüştür ($p < .01$). Toplam test için puanlayıcılar arası güvenilirlik katsayısı .95 olarak bulunmuştur. İkinci pilot çalışma kapsamında test tekrar test güvenilirliğini incelemek için 24 öğrenciye BÜT tekrar uygulanmış ve toplam test için test-tekrar güvenilirlik katsayısı .89 olarak bulunmuştur. İkinci pilot çalışma kapsamında BÜT'ün ölçüt geçerliğini test etme amacıyla öğrencilerin ÜYEP'e başvuru sürecinde uygulanan Matematiksel Yetenek Testi puanları (MYT) ve öğrencilerin fen bilimleri ve matematik derslerinin dönem sonu notları kullanılmıştır. Toplam test puanları (akıcılık, esneklik, yaratıcılık) ile MYT puanları, fen bilimleri ve matematik derslerinin dönem sonu ortalamaları arasında elde edilen korelasyon katsayılarının .35 ile .55 arasında değiştiği bulunmuştur ($p < .000$). BÜT'ün üstün yetenekli öğrencileri diğer öğrencilerden ayırt etme gücünü test etmek için üstün yetenekli öğrenciler ile diğer öğrencilerin bilimsel yaratıcılık puanları karşılaştırılmış ve katılımcıların %75.6'sının doğru bir şekilde tanılandığı sonucuna ulaşılmıştır. BÜT'ün pilot çalışmalarının ardından daha büyük örneklerle psikometrik özellikleri incelenmiş ve ölçeğin geçerli ve güvenilir bir test olduğu ortaya koyulmuştur (Ayas, 2010; Sak, 2010). BÜT'ün psikometrik özelliklerini incelemek amacıyla farklı çalışmalar da gerçekleştirilmiştir. Sak ve Ayas (2013) ÜYEP'e başvuran 288 altıncı sınıf öğrencisi ile ölçeğin ayırt edicilik geçerliğini, iç tutarlılık güvenilirliğini ve ölçüt geçerliğini incelenmiş ve ölçeğin yeterli düzeyde geçerli ve güvenilir olduğunu bulmuşlardır. Araştırmacılar, diğer çalışmalardan farklı olarak bu çalışma kapsamında Temel Bileşenler Analizi yapmış ve tek faktörlü

bir yapı elde etmişlerdir. Bu yapıda tek faktörün ölçeğin %34.45'ini açıkladığı bulunmuştur. Bu çalışmayı ÜYEP'e başvuran 693 altıncı sınıf öğrencisi ile başka bir çalışma takip etmiştir (Ayas ve Sak, 2014). Bu çalışmada araştırmacılar ölçeğin tek faktörlü yapısını test etmek için doğrulayıcı faktör analizi yaparak yapının doğrulandığı sonucuna ulaşmışlardır. BÜT'e yönelik gerçekleştirilen araştırmalar, testin geçerli ve güvenilir olduğunu ortaya koymuştur.

BÜT'ün geçerli ve güvenilir bir test olması alan yazında dikkat çekmektedir. Bu nedenle testin başka bir kültüre uyarlaması da yapılmıştır. Bermejo, vd. (2016) BÜT'ü İspanyol kültürüne uyarlamışlardır. Uyarlanan testin geçerliğini ve güvenilirliğini incelemek için 344 ortaokul öğrencisi ile bir çalışma gerçekleştirmişlerdir. Ölçeğin iç tutarlılığının ve puanlayıcılar arası güvenilirliğinin yeterli düzeyde olduğu ve ölçeğin tek faktörlü yapısının doğrulandığı sonucuna ulaşılmıştır. Araştırma bulguları incelendiğinde BÜT'ün altıncı, yedinci ve sekizinci sınıf öğrencilerinin bilimsel yaratıcılığını belirleme kullanılabilecek bir test olduğu ifade edilebilir. Güçlü bir kuramsal yapıya sahip geçerli ve güvenilir bir testin farklı sınıf düzeylerine göre uyarlanması alanyazın için önem arz etmektedir.

Ayas (2017) BÜT'ün üçüncü, dördüncü ve beşinci sınıf öğrencilerine yönelik formunu (BÜT 3-5) geliştirmiştir. BÜT'ün kuramsal yapısı sadeleştirilmiştir. Araştırmacı kanıt değerlendirme becerisinin hipotez geliştirme ve deney tasarlama becerileri arasında bir arabulucu beceri olmasından dolayı kanıt değerlendirme becerilerini ölçmeye yönelik maddelerin diğer iki beceri ile örtüştüğünü iddia etmektedir. İlgili araştırma bulgularına göre kanıt değerlendirme becerisinin 11 ya da 12 yaşından itibaren tutarlı bir biçimde ölçülebileceği ortaya koyulmuştur. Bu nedenlerden dolayı araştırmacı BÜT'ün kuramsal yapısından kanıt değerlendirme becerisini çıkarmış ve bilimsel yaratıcılığın hipotez geliştirme ve deney tasarlama olmak üzere iki bileşenden oluştuğunu ileri sürmüştür. Güncellenen kuramsal yapıya göre geliştirilen BÜT 3-5 fizik, kimya ve biyoloji alanlarına yönelik dört açık uçlu maddeden oluşmaktadır. BÜT 3-5, BÜT gibi grupla uygulanan fen bilimleri alanındaki çoğul düşünme becerilerini ölçen bir kâğıt kalem testidir. Testin psikometrik özellikleri 251 üçüncü sınıf, 220 dördüncü sınıf ve 166 beşinci sınıf olmak üzere 647 öğrencinin verileri ile incelenmiştir. Ölçeğin yapı geçerliğini incelemek için Temel Bileşenler Analizi ve doğrulayıcı faktör analizi yapılmış ve analiz bulguları iki faktörlü yapıyı doğrulamıştır. Ölçeğin ayırt edicilik geçerliğini incelemek için farklı sınıf düzeylerindeki öğrencilerin puanları karşılaştırılmış ve üst sınıfların lehine bir farklılık bulunmuştur. Bu kapsamda %27'lik üst grupla alt grup karşılaştırılmış ve üst grubun lehine farklılıklar bulunmuştur. Ölçeğin ölçüt geçerliğini test etmek için ölçeğin matematik ve fen bilimleri ders notları, fen bilimlerine yönelik tutum ve fen bilimlerindeki akademik benlik

algısına yönelik uyumu araştırılmış ve ölçeğin ölçüt geçerliğinin yeterli düzeyde olduğu sonucuna ulaşılmıştır. BÜT 3-5'in iç tutarlılığına ve puanlayıcılar arası güvenilirliğine yönelik yapılan analizler ölçeğin yeterli düzeyde güvenilir olduğunu ortaya koymaktadır. BÜT 3-5'in kuramsal yapısının hedef kitleye göre güncellenmesi ve geçerlik-güvenirlik çalışmaları ölçeğin iyi bir ölçüm aracı olduğunu ortaya koymaktadır. Testin madde sayısının az olması bir sınırlılık olarak düşünülebilir. Çünkü deney tasarlama ve hipotez geliştirme alt ölçeklerinde iki madde yer almaktadır. Ölçme değerlendirme uzmanları bir alt ölçekte en az üç madde olması gerektiğini belirtmektedirler (Kline, 2011).

BÜT'ün geliştirilme süreci ve teste yönelik araştırmalar incelendiğinde, testin çoğul düşünme becerileri ile bilimsel düşünme becerilerini sentezlemesi sağlam bir kuramsal yapı ortaya koymaktadır. Testin yeterli düzeyde geçerli ve güvenilir olması kullanımı açısından önemli göstergelerdir. Testin bir yaratıcılık testi olmasına rağmen orijinallik puanlarını içermemesi bir sınırlılık olarak düşünülebilir.

Bilimsel Yaratıcılığın Yapı Modeli'ne göre geliştirilen testler

Alanyazında kimi araştırmacılar Bilimsel Yaratıcılığın Yapı Modeli'ni temel alarak bilimsel yaratıcılık testleri geliştirmektedir. Hu ve Adey (2002) bu modeli kullanarak 6-8. sınıf öğrencilerine yönelik Bilimsel Yaratıcılığın Yapı Testi'ni geliştirmişlerdir. Chin ve Siew (2005), bu kuramı temel alarak anasınıfı öğrencilerinin bilimsel yaratıcılığını ölçmeye yönelik Görsel Bilimsel Yaratıcılık Testi'ni geliştirmişlerdir. Siew, Chong ve Chin (2014) de yine bu kuramı temel alarak beşinci sınıf öğrencilerin bilimsel yaratıcılıklarını ölçmeye yönelik bir test oluşturmuşlardır.

Bilimsel Yaratıcılığın Yapı Testi

Bilimsel Yaratıcılığın Yapı Testi altıncı, yedinci ve sekizinci sınıf öğrencilerinin fen bilimleri alanındaki yaratıcılığını ölçme amacıyla geliştirilmiştir (Hu ve Adey, 2002). Yedi açık uçlu maddeden oluşan ölçek grup olarak yaklaşık 60 dakikada uygulanmaktadır. Maddelerin puanlanması yaratıcılığın bileşenlerine (akıcılık, esneklik ve orijinallik) göre yapılmaktadır. Akıcılık, her bir doğru cevap sayısına, esneklik doğru cevapların yer aldığı kategori sayısına göre hesaplanmaktadır. Orijinallik puanı ise yanıtların sıklığıdır. Puanlar 0 (%10'un üzeri), 1 (%5 ile %10 arası) ve 2 (%5'in altı) aralığında hesaplanmaktadır. Ölçeğin "Bir elma toplama makinesi tasarlayınız. Tasarladığımız makinenin resmini çiziniz ve makinede yer alan her bir parçanın adını ve işlevini yazınız." maddesinin değerlendirilmesinde farklı bir yaklaşım benimsenmektedir. Öğrencinin tasarladığı üründe yer alan her bir fonksiyon için 3 puan verilerek esneklik puanı elde edilmektedir. Orijinallik puanı ise 1 ile 5 arasında puanlayıcının görüşüne göre yapılmaktadır.

Hu ve Adey (2002, s. 397-401) bilimsel yaratıcılık testini geliştirme sürecinde ilk olarak 48 madde geliştirmiş ve maddeleri uzman görüşüne sunmuşlardır. Çin Halk Cumhuriyeti'nde görev yapan 50 fen eğitimcisinin görüşlerine göre madde sayısı dokuza düşürülmüştür. Dokuz maddelik testi 60 öğrenciye uygulayarak pilot çalışma yapılmıştır. Çalışmanın sonunda testin yedi maddelik son formu oluşturulmuştur. Araştırmacılar testin son formunu İngiltere'de 160 öğrenciye uygulamışlar ve testin Cronbach Alfa iç tutarlılık katsayısını .89 olarak bulmuşlardır. Puanlayıcılar arasındaki korelasyon değerlerinin ise .793 ile .913 arasında değiştiği sonucuna ulaşılmıştır. Testin yapı geçerliğini incelemek için Temel Bilişenler Analizi yapılmış ve tek faktörlü bir yapı elde edilmiştir. Yedi maddenin %63 varyansla ölçeği açıkladığı rapor edilmiştir. Görünüş geçerliğini test edilmiş ve ölçeğin bilimsel yaratıcılığı ölçmeye uygun olduğu belirtilmiştir.

Ölçeğin güçlü yanı kuramsal bir yapıya göre tasarlanmasıdır. Ölçeğin maddeleri diğer genel yaratıcılık testlerine, puanlama yöntemi ise Torrance Yaratıcılık Testi'ne (Torrance, 1998) oldukça benzemektedir. Örneğin ölçeğin genel yaratıcılık testlerinde bir nesnenin farklı kullanım alanları sorulmaktadır (Kaufman, Plucker ve Baer, 2008). Bilimsel Yapı Testi'nin ilk maddesi de bu formatta hazırlanmıştır: "Bir cam parçasının düşünebildiğiniz kadar çok sayıda bilimsel kullanımlarını yazınız." Ölçeğin orijinallik puanlaması Torrance Yaratıcılık Testi'ne göre yapılmaktadır. Orijinallik puanlamasında dikkate alınan yüzdeler oranlarına yönelik herhangi bir bilimsel kanıt bulunmamaktadır. Ölçeğin yedinci maddesinin orijinallik puanı öznel bir değerlendirmeye göre elde edilmektedir. Görüldüğü üzere ölçeğin maddelerinin puanlanmasında standart bir yaklaşım benimsenmemiştir. Bu nedenle ölçekten toplam yaratıcılık puanı elde edilememektedir. Fakat yaratıcılığın değerlendirilmesinde toplam yaratıcılık puanının hesaplanması önerilmektedir (Kaufman, Plucker ve Baer, 2008). Ölçeğin geçerlik ve güvenilirlik çalışmaları kapsamında çeşitli sınırlılıkları bulunmaktadır. Örneğin ölçeğin madde seçiminde Çin Halk Cumhuriyeti'nde görev yapan 50 fen eğitimcisinin görüşlerine başvurulmuş ancak pilot çalışma ve ana çalışma İngiltere'de yapılmıştır. Ana çalışmanın örneklem sayısının görece küçük olması geçerlik-güvenirlik bulgularının etkisini düşürmektedir. Ölçeğin sınırlılıkları olmasına rağmen farklı kültürlere ve disiplinlere uyarlanması yapılmış ve kimi çalışmalarda bilimsel yaratıcılığı ölçmek amacıyla kullanılmıştır.

Hu ve Adey'in testi Estonya (Laius ve Rannikmäe, 2011) ve Endonezya kültürüne (Samsudin vd., 2018) uyarlanmıştır. Bu çalışmalarda testin geçerliği ve güvenilirliğine yönelik bir bulgu rapor edilmemiştir. Ölçeğin Türk kültürüne özgü araştırmalarda da kullanıldığı fakat yeterli geçerlik ve güvenilirlik çalışmalarının yapılmadığı görülmektedir (örn. Aktamış ve Ergin, 2008;

Demir ve Şahin, 2014). Bu durumu eleştiren Çeliker ve Balım (2012) geçerlik-güvenirlik çalışmaları yaparak Hu ve Adey'in testini Türkçeye uyarlamışlardır. Uyarlanan ölçek 6-8. sınıflarda öğrenim görmekte olan toplam 389 öğrenciye uygulanmıştır. Çalışmanın sonucunda, ölçeğin madde toplam korelasyonlarının .37 ile .74 arasında, Cronbach Alfa iç tutarlılık katsayısının ise .86 olduğu bulunmuştur. Bilimsel Yaratıcılığın Yapı Testi'nin araştırmacılar tarafından kullanılması ölçeğin alanyazındaki önemine dikkat çekmektedir. Ancak ölçeğin geçerlik ve güvenilirlik çalışmalarının genişletilmesi ve ölçeğin uygulamaya yönelik sınırlılıkları için çeşitli önlemlerin alınması gerekmektedir.

Görsel Bilimsel Yaratıcılık Testi

Chin ve Siew (2005) okul öncesi dönemdeki çocukların bilimsel yaratıcılığını ölçmek için Bilimsel Yaratıcılığın Yapı Modeli'ni temel alarak Görsel Bilimsel Yaratıcılık Testi'ni geliştirmişlerdir. Testin hedef kitlesinin özelliklerini ve anaokulu öğretim programını göz önünde bulundurarak Bilimsel Yaratıcılığın Yapı Modeli revize edilmiştir. Modele göre bilimsel yaratıcılık ürün (teknik ürün, bilimsel bilgi, bilimsel kavram ve bilimsel problem), özellik (akıcılık, esneklik ve orijinallik) ve süreç (düşünme ve hayal gücü) boyutlarından oluşmaktadır. Revize edilen modelde de üç boyut bulunmaktadır. Ürün boyutunda yer alan teknik ürün çıkarılmış ve özellik boyutunun kapsamı değiştirilmiştir. Revize edilen modelin özellik boyutu akıcılık, orijinallik, ayrıntı, soyutluk ve direnç türlerinden oluşmaktadır. Görsel Bilimsel Yaratıcılık Testi altı maddeden oluşmaktadır. Testin her bir maddesi sırasıyla 10 dakikalık bir süre içerisinde 5-10 kişilik gruplarla uygulanmaktadır. Uygulama sırasında öğrencilerden yanıtlarını çizmeleri ve çizdikleri resimleri isimlendirmeleri beklenmektedir. Öğrenci yanıtlarının puanlanmasında akıcılık, orijinallik, ayrıntı, soyutluk ve direnç özellikleri göz önünde bulundurulmaktadır. Ölçeğin dört maddesi için bu puan türlerinin tamamı, iki maddesi için ise direnç puanı hariç tüm puan türleri hesaplanmaktadır. Akıcılık ve orijinallik puanları Bilimsel Yaratıcılığın Yapı Testi'ndeki gibi hesaplanmaktadır. Ayrıntı puanı öğrencilerin çizimlerdeki her bir fonksiyon göz önünde bulundurularak elde edilmektedir. Soyutluk puanı öğrencilerin çizimlerine verdikleri isimlere göre 0 ile 3 arasında (0: Niteleyici bir özellik olmayan genel isimlendirme, 1: Somut özelliklerine yapılmış isimlendirme, 2: Hayali isimlendirme, 3: Soyut isimlendirme) değişmektedir. Direnç puanı ise öğrencilerin çizimlerinin özelliklerine göre 0 ile 2 arasında (0: Standart şekiller oluşturma, 1: Özellikli şekiller oluşturma, 2: Özgün şekiller oluşturma) farklılık göstermektedir.

Chin ve Siew, testin psikometrik özellikleri incelemek için 30 okul öncesi öğrencisi ile bir çalışma yapmışlardır. Cronbach Alfa iç tutarlılık katsayısının .81 olduğu, puanlayıcılar arasındaki korelasyon değerlerinin .54 ile .87 arasında değiştiği bulgusu rapor edilmiştir. Yapı geçerliğini incelemek için açılımlı

faktör analizi yapılmış ve tek faktörlü bir yapı elde edilmiştir. Altı maddenin açıkladığı toplam varyans %50.85 olarak bulunmuştur. Yapılan analizlerle ölçeğin yeterli düzeyde geçerli ve güvenilir olduğu sonucuna ulaşılmıştır.

Görsel Bilimsel Yaratıcılık Testi okulöncesi dönemdeki çocukların bilimsel yaratıcılık düzeyini ölçmeyi amaçlayan ilk çalışmalardan birisidir. Ölçeğin bir çoğul düşünme testi olmasına rağmen bir maddesinin açık uçlu olmadığı görülmektedir. Ölçeğin beşinci maddesinde öğrencilerden farklı yollarla kum ve tuz karışımını ayırmaları beklenmektedir. Bu problem durumunun tek yanıtı (çözünürlük farkı) bulunmaktadır. Ayrıca bu problem durumu farklı ülkelerde, örneğin ülkemizde, ortaokul öğretim programları kapsamında yer almaktadır. Okul öncesi öğrencileri için zor bir madde olarak değerlendirilebilir. Ölçeğin uygulanmasında okul öncesi öğrenciler yanıtlarını çizmekte ve çizdikleri resimleri isimlendirmektedirler. Bu düzeydeki öğrencilerin sınırlı çizim ve yazı becerileri, bilimsel yaratıcılığın ölçülmesinde kirlenici bir etki oluşturabilir. Ölçeğin yapı geçerliğinin test edilmesinde 30 katılımcı ile açılımlayıcı faktör analizi yapılmıştır. Kline (1999) faktör analizi yapılabilmesi için örneklem büyüklüğünün madde sayısına oranının en az 10 olması gerektiğini ifade etmektedir. Ölçeğin altı maddeden oluştuğu göz önünde bulundurulduğunda bu analizin yapılabilmesi için en az 60 katılımcı gerekmektedir. Bu bağlamda ölçeğin geçerlik ve güvenilirlik çalışmalarında sınırlılık olduğu ifade edilebilir. Test, okul öncesi öğrencilerinin bilimsel yaratıcılığını ölçmede kullanılabilecek bir taslak olarak değerlendirilebilir.

Beşinci sınıflara yönelik bilimsel yaratıcılık testi

Siew, Chong ve Chin (2014) beşinci sınıf öğrencilerinin bilimsel yaratıcılığını ölçme amacıyla Bilimsel Yaratıcılığın Yapı Modeli'ni temel alarak A ve B olmak üzere iki formu olan bir çoğul düşünme testi geliştirmişlerdir. Ölçeğin geliştirilme sürecinde Bilimsel Yaratıcılığın Yapı Modeli'ni temsil eden küpün 24 biriminin her biri için bir madde geliştirilmiştir. Pilot çalışma kapsamında maddeler uzman görüşüne sunulmuş ve ölçeğin dört maddelik son iki formu oluşturulmuştur. Testin A ve B formlarında yer alan maddeler birbirlerine paralel olarak hazırlanmıştır. Örneğin düşünme, hayal gücü ve bilimsel problem oluşturma becerisini ölçmek için öğrencilere A formunda "Bir kareyi dört eşit parçaya (aynı şekil) bölmek için mümkün olduğunca çok yöntem kullanınız. Cevaplarınızı kâğıda çiziniz" sorusu B formunda ise "Beş kibrit çöpü kullanarak mümkün olduğunca çok şekil oluşturunuz. Oluşturduğunuz şekilleri kâğıda çiziniz" sorusu sorulmuştur (Siew, Chong ve Chin, 2014, s. 114).

Ölçeğin psikometrik özelliklerini incelemek amacıyla testin her iki formu Malezya'da öğrenim gören 206 beşinci sınıf öğrencisine uygulanmıştır. Araştırmacılar A formunda yer alan dördüncü maddenin, B formunda yer alan ikinci maddenin %27'lik alt ve üst grubu ayırt edemediği, diğer maddelerin bu

grupları ayırt edebildiği sonucuna ulaşmışlardır. İç tutarlılığı değerlendirmek için hesaplanan Cronbach Alfa katsayısının A ve B formu için sırasıyla .77 ve .68 olduğu rapor edilmiştir. Puanlayıcılar arası güvenilirliğin A ve B formları için yeterli düzeyde olduğu belirtilmiştir. Yapı geçerliği için açımlayıcı faktör analizi yapılmış ve her iki form için de iki faktörlü bir yapı elde edilmiştir. Yapılan analizler sonucunda ölçeğin geçerli ve güvenilir olduğu iddia edilmiştir. Modelin üçlü yapısı yerine ikili yapısı doğrulanmıştır. Ölçek maddeleri değerlendirildiğinde birinci maddenin (“Bir kalemi daha ilginç ve sıra dışı bir forma dönüştürmek için ne tür bilimsel yöntemler kullanılabilir. Düşüncelerinizi çizerek açıklayınız”) çizim becerilerine, ikinci maddenin çağrışım kurma becerilerine (“Mıknatıs ile ilgili olabildiğince çok bilimsel kelime yazınız”), üçüncü maddenin (“Güneş ışınlarının yok olması durumunda neler olabileceğine yönelik bilimsel hikâyeler yazınız”) yazım becerilerine bağlı olduğu görülmektedir (Siew, Chong ve Chin, 2014, s. 113-114). Maddelerin ölçülmek istenilen yapıyı temsil etme gücünün düşük olduğu düşünülebilir.

Yaratıcı Bilimsel Çağrışımlar Testi

Kanlı (2014a), ortaokul öğrencilerinin (5-8 sınıf) bilimsel yaratıcılık düzeylerini ölçmek amacıyla Yaratıcı Bilimsel Çağrışımlar Modeli’ni önermiş ve bu model temelinde Yaratıcı Bilimsel Çağrışımlar Testi’ni (YBÇT) geliştirmiştir. Ölçeğin uygulanması yaklaşık bir saat sürmektedir. Ölçek çağrışımlar, analojik muhakeme ve analojik problem çözme olmak üzere üç alt testten oluşmaktadır. Ölçek toplamda 19 açık uçlu maddeden oluşmaktadır. Maddelerin puanlamasında ise üretilen doğru yanıt sayısı yani akıcılık puanları hesaplanmaktadır.

Çağrışımlar alt testinde öğrencilere fen bilimleri alanına özgü farklı kavramlar sunulmakta ve öğrencilerden bu kavramlar arasındaki ilişkileri bulmaları beklenmektedir. Örneğin öğrencilerden “ISI, IŞIK, ELEKTRİK” arasındaki ilişkiyi bulmaları istenmektedir (Kanlı, 2014a, s. 159). Analogik muhakeme alt testinde öğrencilere fen bilimleri alanına özgü analogi oluşturmaları beklenmektedir. Bu alt testte yer alan örnek bir soru şu şekildedir: “ÇAYDAKİ ŞEKER ile KARIŞTIRMAK arasındaki ilişki KAYA ile neler arasında vardır? (Kanlı, 2014a, s. 159). Analogik problem çözme alt testinde ise öğrencilere fen bilimlerine özgü bir problem sunulmakta ve öğrencilerden bu probleme yönelik çözüm üretmeleri beklenmektedir. Bu alt testte yer alan örnek bir soru şu şekildedir (Kanlı, 2014a, s. 159):

“Aslı ve Ercan bahçede ağaç dikmek için toprağı kazıyordu. Ercan toprağın üstünün kuru olmasına rağmen kazıldıkça alt tabakaların nemli olduğunu fark etti ve bunun neden olduğunu sordu. Aslı bir gösterimle açıklayacağını söyledi ve eve gittiklerinde bir kabın içini önce çakıl sonra kumla tamamen doldurdu ve dolu olan kabın içine su döktü. Ağzına kadar çakıl ve kumla dolu olan kap ancak

bir miktar suyla da dolduktan sonra taşmaya başladı. Sence Aslı bu gösterimle Ercan'a neyi anlatmak istedi? Toprağın alt tabakalarının nemli üstü tabakasının kuru olmasındaki bilimsel açıklamalar neler olabilir?"

Kanlı madde havuzunun oluşturulması ve iki pilot çalışma sonrasında 25 maddelik ölçeğin son halini oluşturmuştur. Ölçeğin bu formu 385'i normal, 293'ü üstün yetenekli olmak üzere 678 öğrenciye uygulanmıştır. Ölçeğin yapı geçerliğini incelemek için açımlayıcı ve doğrulayıcı faktör analizi yapılmıştır. Açımlayıcı faktör analizinde 6 madde ölçekten çıkarılmış ve üç faktörlü yapı doğrulanmıştır. Analiz sonucunda üç faktörün varyansa yaptığı katkı %47.4 olarak bulunmuştur. Aynı veri seti üzerinden yapılan doğrulayıcı faktör analizi ile de üç faktörlü yapı doğrulanmıştır. Analiz sonucunda uyum iyiliği değerlerinin iyi düzeyde olduğu rapor edilmiştir. Ölçeğin iç tutarlılığını incelemek için Cronbach Alfa katsayısı hesaplanmış ve bu değer toplam test için .89 olarak bulunmuştur. Puanlayıcılar arası güvenilirlik katsayısı toplam test için .984, test tekrar test güvenilirlik katsayısı toplam test için .935 olarak bulunmuştur. Üstün yetenekli öğrencilerle normal öğrencilerin puanları karşılaştırılmış ve katılımcıların %79.9'unun doğru bir şekilde tanılandığı sonucuna ulaşılmıştır.

Kuramsal bir yapıya göre tasarlanmış YBÇT geçerliği ve güvenilirliği ortaya koyulmuş bir bilimsel yaratıcılık testidir. Ancak ölçeğin yapı geçerliğine yönelik kanıtların sınırlı olduğu ifade edilebilir. Çünkü açımlayıcı ve doğrulayıcı faktör analizinin aynı veri seti üzerinden yapılması önerilmemektedir (Costello ve Osborne, 2005; Kline, 2013). YBÇT'nin bir yaratıcılık testi olmasına rağmen sadece akıcılık puanlarının hesaplanması ölçeğin diğer sınırlılığdır. Ölçüm puanlarına esneklik, orijinallik gibi diğer yaratıcılık puanları da eklenebilir. 19 maddelik ölçeğin yaklaşık bir saatte uygulanması, öğrencilerin sınırlı sayıda fikir üretmesine neden olabilir. Ölçeğin uygulama süresinin uzatılması gerekebilir.

Çoğul düşünme becerilerine odaklanan diğer testler

Alanyazında kuramsal bir yapıya göre oluşturulmayan ölçeklere de rastlanmaktadır. Örneğin Sharma ve Mahrshi (2017), 12-13 yaş aralığındaki çocukların bilimsel yaratıcılık düzeylerini ölçmek için alanyazında var olan testlerden yararlanarak bir bilimsel yaratıcılık testi geliştirmişlerdir. Sonuçlar, yeni ilişkiler, nedenler, nedenleri test etme, problem çözme ve bilimsel isimlendirme gibi becerileri ölçen test altı maddeden oluşmaktadır. Maddeler akıcılık, esneklik ve orijinallik ölçütlerine göre puanlanmaktadır. Ölçek 375 sekizinci sınıf öğrencisine uygulanarak, ölçeğin standardizasyonu yapılmıştır. Ölçeğin ölçüt geçerliği için Shukla ve Sharma'nın 1987 yılında geliştirmiş oldukları test uygulanmış ve ölçekler arasındaki uyumun .80 olduğu bulunmuştur. Ölçeğin test-tekrar test güvenilirlik katsayısı .90 olarak hesaplanmıştır. Araştırmacılar ölçeğin bir standardizasyon çalışması olduğunu

öne sürmektedirler. Ancak çalışma kapsamında sadece ölçüt geçerliği ve test-tekrar test güvenilirliği incelenmiştir. Çalışmanın örneklem sayısı da bir standardizasyon çalışması için yeterli değildir.

Kuramsal bir yapıya göre oluşturulmayan bir diğer ölçek Usta ve Akkanat (2015) tarafından geliştirilmiştir. Usta ve Akkanat bilimsel yaratıcılığın ölçülmesinde çoğul düşünme, tahmin yapabilme, problem oluşturma, problem çözme, hipotez geliştirme, sınıflama, döndürme ve yeniden yapılandırma gibi becerilerin değerlendirilmesinin gerekliliğine vurgu yapmaktadır. Bu doğrultuda 6-8. sınıf düzeyindeki öğrencilerin yaratıcılıklarını ölçmek için 10 tane açık uçlu madde geliştirilmiştir. Maddelerin değerlendirilmesinde akıcılık, esneklik ve orijinallik puanları elde edilmektedir. Araştırmacılar ölçeği 300 yedinci sınıf öğrencisine uygulamışlar ve ölçeğin iç tutarlılığı için Cronbach Alfa değerini .71 olarak bulmuşlardır. Çalışma kapsamında ölçeğin nasıl geliştirildiğine yönelik bilgi paylaşılmamış sadece ölçeğin bir çoğul düşünme testi olduğu belirtilmiştir.

Çoğul ve tekil becerilere yönelik bilimsel yaratıcılık testi

Yang, vd. (2016) bilimsel yaratıcılığın ölçülmesinde tekil ve çoğul becerilerin ölçülmesi gerektiğini iddia etmektedirler. Bu nedenle 3-6. sınıf öğrencilerin bilimsel yaratıcılığını ölçmek için tekil ve çoğul becerileri kapsayan bir bilimsel yaratıcılık testi geliştirmişlerdir. Dokuz maddeden oluşan ölçek 40 dakikada uygulanmaktadır. Ölçeğin yedi maddesi ile çoğul düşünme becerileri, iki maddesi ile de tekil düşünme becerileri ölçülmektedir. Çoğul düşünme becerilerini ölçen maddeler Hu ve Adey'in Bilimsel Yaratıcılık Testinden uyarlanmıştır. Örneğin geliştirilen ölçekte "Güneş olmasaydı, dünya nasıl bir yer olurdu?" gibi sorular uyarlanmıştır. Tekil düşünme becerileri ise Sternberg ve Lubart'ın (1995) geliştirdikleri bir ölçekten alınmıştır. Bu maddelerden birinde "Zemine sabitlenmiş bir yay ile tavana sabitlenmiş bir yayı bir noktada birleştirmek için bir strateji geliştiriniz. Bu yayların birbirinden el ile birleştiremeyecek kadar uzak olduğunu göz önünde bulundurunuz." sorusu sorulmuştur. Maddenin çözümünde öğrencilerden lastik bant, cam toplar, cam kavanoz ve pense materyallerini kullanmaları beklenilmektedir. Çoğul düşünme becerilerine yönelik maddeler akıcılık, esneklik ve orijinallik puanına göre değerlendirilmektedir. Tekil düşünme becerilerine yönelik maddeler 0-4 arasında puanlanmaktadır. Geliştirilen ölçeğin tamamı 158 ilkokul öğrencisine uygulanarak geçerlik-güvenirlik çalışmaları yapılmıştır. Araştırmacılar puanlayıcılar arası güvenilirliğinin .81 ile .98 arasında değiştiğini, madde ayırt edicilik katsayısının .28 ile .99 arasında olduğunu, madde güçlük değerlerinin .14 ile .50 arasında değiştiğini rapor etmişlerdir. Yang vd. (2019) başka bir çalışmada geliştirdikleri ölçeği revize etmişler ve geçerlik-güvenirlik çalışmalarını genişletmişlerdir. Yeni formu oluşturulan ölçek için iki pilot

çalışma yapılmış ve geçerlik ve güvenilirlik değerlerinin bir önceki versiyona göre daha iyi düzeyde olduğu bilgisi paylaşılmıştır. Ölçeğin iç tutarlılığını incelemek için hesaplanan Cronbach Alfa katsayısının çoğul düşünme becerisine yönelik maddeler için .60, tekil düşünme becerisine yönelik maddeler için .35, toplam ölçek için ise .61 olduğu sonucu paylaşılmıştır. Ölçeğin geçerlik-güvenirlik ve uygulama açısından çeşitli sınırlılıkları olduğu görülmektedir. Paylaşılan iç tutarlılık değerleri düşüktür. İlkokul öğrencilerinin dokuz maddelik bir ölçeği 40 dakika içerisinde tamamlamasının zor olduğu söylenebilir.

Öz Değerlendirme ve Başkaları Tarafından Değerlendirilme Yöntemi

Öz değerlendirme yönteminde birey kendi potansiyelini değerlendirmekte (Silvia, vd., 2012), başkaları tarafından değerlendirilme yönteminde ise öğretmen, akran, ebeveyn gibi kişiler bireyin potansiyelini değerlendirmektedir (Kaufman, Plucker ve Baer, 2008). Bu yöntemler yaratıcılık ölçümlerinde sıklıkla kullanılırken bilimsel yaratıcılık ölçümlerinde sınırlı kullanılmaktadır (Ayas, 2017).

Bilimsel yaratıcılığı öz değerlendirme yöntemi ile ölçen çalışmalara Taylor ve Ellison'un (1967) araştırması örnek olarak verilebilir. Taylor ve Ellison bilimsel yaratıcılığı en iyi yordayan değişkenlerin biyografik değişkenler olduğunu iddia etmektedirler. Bu nedenle bilimsel yaratıcılığı değerlendirmek için Alfa Biyografik Ölçeği'ni geliştirmişlerdir. Bu ölçek çoktan seçmeli formatta hazırlanmış 300 maddeden oluşmaktadır. Maddeler gelişim geçmişi, akademik geçmiş ve yetişkinlik dönemi, aile ve ilgi olmak üzere dört alana yönelik tasarlanmıştır. Ölçek maddelerinin bir kısmı ile bireyler kendilerini bilimsel alanlarda ne kadar yaratıcı olduklarını değerlendirmektedir. Ölçeğin geliştirilme sürecini dört aşamada 2000 bilim insanı ile gerçekleştirilmiştir. Bilim insanlarının büyük bir çoğunluğu Ulusal Havacılık ve Uzay Dairesi'nde (National Aeronautics and Space Administration, NASA) çalışmaktadır. Araştırmacılar ölçek geliştirme aşamalarında geçerliliği ve güvenilirliği düşük olan maddeleri belirlemiş ve ölçeğin 300 maddelik son formunu oluşturmuşlardır. Araştırma bulguları bilimsel yaratıcılığı yordayan değişkenlerin önem sırasının yetişkinlik dönemi ve ilgi, akademik geçmiş, gelişim geçmişi ve aile şeklinde olduğunu göstermektedir. Ölçeğin lise düzeyinde de uygulanabileceği iddia edilmektedir. Davis ve Belcher (1971) Alfa Biyografik Ölçeği'ni, Mednick'in Uzak Çağrışımsal Testi'ni ve Torrance Yaratıcı Düşünme Testi'ni 82 lise öğrencisine uygulamıştır. Davis araştırması kapsamında bu ölçeklerden elde edilen puanları karşılaştırmış ve Alfa Biyografik Ölçeği puanlarının Torrance Yaratıcı Düşünme Testi puanları ile arasında yüksek derece ve anlamlı ($r=.68$, $p<.001$) bir ilişki olduğunu bulmuştur. Alfa Biyografik Ölçeği'nin ölçüt geçerliğinin sağlandığı ifade edilebilir. Ancak ölçeğin madde sayısının çok fazla

olması uygulamayı zorlaştırabilir ve geçerlik-güvenirlilik çalışma bulgularını kirlitebilir.

Bilimsel yaratıcılığı öz değerlendirme yöntemi ile ölçen başka bir çalışma ise Zhu, vd. (2019) tarafından gerçekleştirilmiştir. Tekil ve çoğul düşünmenin bilimsel yaratıcılık ile olan ilişkisinin araştırıldığı çalışmada 588 lise öğrencisinin bilimsel yaratıcılığını ölçmek için Çinceye uyarlan bir anketin fen bilimlerine yönelik olan bölümü kullanılmıştır. Bu anketin orijinali Carson, Peterson ve Higgins (2005) tarafından geliştirilmiştir. Carson, Peterson ve Higgins'in anketinde bireyler kendi başarı düzeylerini sanat, bilim, edebiyat gibi 10 farklı yaratıcılık alanında değerlendirmektedir. Anketin bilimsel yaratıcılık ile ilgili olan bölümünde yedi madde bulunmaktadır. Katılımcılar bilim alanında yeteneğinin olmadığını düşünüyorsa ankette yer alan soruları yanıtlamamaktadır. Katılımcılar anket maddelerini evet ya da hayır olarak yanıtlamaktadır. Anket maddeleri bilimsel problem çözme, ulusal bir yarışta bilimsel başarı ödülü alma, bilim alanında bir burs kazanma, uluslararası bilimsel bir ödül kazanma, yayınlanmış bilimsel bir makaleye sahip olma, proje ödülü alma, atıf almış olma gibi durumlara yöneliktir. Anketin geçerlik ve güvenirlilik çalışmaları kapsamlı bir şekilde yapılmış ve geçerli ve güvenilir bir ölçüm aracı olduğu ortaya koyulmuştur.

Liang (2002) bilimsel yaratıcılığın ölçümünde öz değerlendirme ve başkaları tarafından değerlendirilme yöntemini birlikte kullanmıştır. Liang, Tayvan'da yaşayan 11. sınıf öğrencilerinin bilimsel yaratıcılığını ölçme amacıyla bir çalışma gerçekleştirmiştir. Katılımcıların bilimsel yaratıcılık düzeylerini iki ölçüm aracının puanlarını birleştirerek belirlemiştir. Bu araçlar, Yaratıcılığı Derecelendirme Ölçeği (Creativity Rating Scale, CRS) ve Yaratıcı Etkinlikler ve Başarı Kontrol Listesi'dir (Creative Activities and Accomplishments Check Lists, CAACL). CRS, Renzulli'nin 1983 yılında geliştirdiği Üstün Yetenekli Öğrencilerin Davranış Özelliklerini Derecelendirme Ölçeği'nin (Scales for Ratings the Behavior Characteristics of Superior Students, SRBCSS) fen bilimlerine göre uyarlanmış halidir. SRBCSS 10 maddeden oluşmaktadır. Öğretmenler sınıflarındaki bir öğrenci için ölçekte yer alan her bir maddeye 0 ile 3 arasında (0: nadiren ya da hiçbir zaman, 1: ara sıra, 2: oldukça, 3: her zaman) bir puan vermektedir. CRS'de aynı puanlama yöntemi benimsenmektedir. Liang'ın araştırmasında puanlamayı öğretmenler yerine akranlar yapmaktadır. SRBCSS'de yer alan 10 madde, fen bilimlerine göre düzenlenmiştir. Örneğin SRBCSS'de yer alan "Meraklıdır, herhangi bir şey hakkında sürekli sorular sorar." ifadesi CRS'de "Fen bilimlerine meraklıdır, fen bilimleri ile ilgili herhangi bir şey hakkında sürekli sorular sorar." şeklinde sorulmaktadır. CRS'nin iç tutarlılık katsayısı .79, puanlayıcılar arası güvenirlilik katsayısı ise .91 olarak bulunmuştur. CAACL, Runco, Noble ve

Luptak tarafından 1990 yılında geçerlik güvenirlik çalışmaları yapılmış bir ölçektir. Ölçekte edebiyat, el sanatları, bilim ve sanat alanı olmak üzere dört alana yönelik 52 etkinlik yer almaktadır. Öğrencilerden her bir etkinliği ne sıklıkta gerçekleştirildiklerini (hiç, bir ya da iki kez, 3-5 kez, 5-10 kez, 10'dan fazla) belirtmeleri istenmektedir. Liang (2002) bilimsel yaratıcılığı ölçmek için fen bilimlerine yönelik 17 etkinliği kullanmıştır. Bu etkinlikler şunları kapsamaktadır: Bilimsel bir proje ya da makale için ödül kazanmak, makale yayınlamak, patentli icat geliştirmek, bilimsel derneklerden burs kazanmak, olimpiyatlara katılmak, deney tasarlamak, araç geliştirmek, bitki yetiştirmek, hayvan yetiştirmek, bilimsel yarışmaya katılmak, kamuya açık sergilenen bir projeye sahip olmak, bilimsel araştırma sürecine dâhil olmak, laboratuvar asistanlığı yapmak, fen bilimlerine yönelik yaz kursuna katılmak, bilimsel bir kulübe üye olmak, bilimsel öğretim materyali hazırlamak ve bilimsel kitap okumak. Liang bilimsel yaratıcılığı kişisel özellikler ve yaratıcı ürünlerin kesişimi olarak ele almaktadır. Bu nedenle öğrencilerin CRS ve CAACL ölçüm araçlarından aldıkları puanları çarparak bilimsel yaratıcılık puanını elde etmektedir. Liang'ın bilimsel yaratıcılığı ölçmede kullandığı yöntem incelendiğinde farklı bir kültüre uyarlanan ölçüm araçlarına yönelik yeterli düzeyde geçerlik ve güvenirlik çalışmalarının yapılmadığı görülmektedir.

Sonuç olarak öz değerlendirme ve başkaları tarafından değerlendirilme yönteminin daha çok lise öğrencileri ve yetişkinlerin bilimsel yaratıcılık düzeylerini ölçmede kullanıldığı görülmektedir. Bu yöntemin çeşitli sınırlılıkları bulunmaktadır. Kullanılan ölçeklerin kullanımında geçerlik ve güvenirlik bağlamında istenen düzeye ulaşamama, özelliğin yüksek olması, maliyetli olma ve zaman alma gibi problemlerle karşılaşmaktadır (Alencar, Fleith ve Bruno-Faria, 2014; Hocevar, 1979; Kaufman, Plucker ve Baer, 2008). Bu yöntemin inanç, önyargı gibi kişisel faktörlerden etkileneceği ve daha yüzeysel bilgiler elde edilebileceği ifade edilebilir. Yaratıcılığın kültürel unsurlardan etkilenmesi nedeniyle bu ölçüm yönteminin kültürlere göre farklılaşacağı düşünülebilir.

Konsensüse Dayalı Değerlendirme Yöntemi

Konsensüse dayalı değerlendirme yöntemi yaratıcılık ölçümlerinde sıklıkla kullanılan yöntemlerden biridir. Amabile (1982) tarafından geliştirilen yöntemde bir ürünün yaratıcılığının değerlendirilmesinde alan uzmanlarının görüşleri alınmaktadır. Bu yöntemin herhangi bir kuramsal yapısı bulunmamaktadır. Bu nedenle geçerlik çalışmaları yapılamamaktadır (Baer ve McKool, 2009). Yöntemin yaratıcılık ölçümleri için "altın standart" olarak değerlendirilmesi ve her alanda uygulanabilmesi nedeniyle yaratıcılık ölçümlerinde sıklıkla kullanılmaktadır (Baer, Kaufman, ve Gentile, 2004). Yöntem iki adımda oldukça basit bir şekilde uygulanmaktadır. İlk adımda katılımcılardan şiir

yazmaları, deney tasarımları, kompozisyon oluşturmaları, problem çözmeleri yani bir ürün oluşturmaları beklenmektedir. İkinci adımda ise alan uzmanları birbirinden bağımsız olarak ürünlerin yaratıcılık düzeyini puanlamaktadır. Sonrasında söz konusu puanların ortalaması alınarak ürünlerin yaratıcılık puanları elde edilmektedir (Baer ve McKool, 2009). Bilimsel yaratıcılığın ölçümünde bu yöntemin uygulamasına sınırlı sayıda da olsa rastlanmaktadır.

Atakaya (2018), konsensüye dayalı değerlendirme yöntemini kullanarak 3200 öğrencinin Bilimsel Üretkenlik Testi (BÜT) sorularına verdikleri cevapları alan uzmanlarına değerlendirmiştir. Araştırmacı öğrenci yanıtlarını değerlendirmeleri için fen bilimleri, fizik, kimya ve biyoloji alanlarında öğretmenlik yapan 26 kişilik bir uzman grubu oluşturmuştur. Uzman grup üyeleri birbirinden bağımsız olarak öğrenci yanıtlarının uygunluk ve yaratıcılık düzeyine 1 ile 6 arasında bir puan vermiştir. Puanlayıcılar arası güvenilirlik katsayısı “uygunluk endeksi” için .845, “yaratıcılık endeksi” için ise .739 olarak bulunmuştur. Uzmanlar arasında güvenilirliği değerlendirmek amacıyla hesaplanan Cronbach Alfa değeri uygunluk endeksi için .858, yaratıcılık endeksi için .779 olarak bulunduğu rapor edilmiştir.

Konsensüye dayalı değerlendirme yönteminin kullanıldığı bir diğer çalışma Mohamed (2006) tarafından gerçekleştirilmiştir. Mohamed alan yazında var olan araştırmaları inceleyerek bir model oluşturmuş ve bu modele göre beşinci sınıflara yönelik bir bilimsel yaratıcılık testi geliştirmiştir. Ölçek grup olarak uygulanmakta ve 75 dakika sürmektedir. Test Problemler ve Çözümler, Çiçekleri Gruplama ve Tasarım ve Deney olmak üzere üç alt ölçekten oluşmaktadır. Problemler ve Çözümler alt ölçekte öğrenciye çevresel problemleri yansıtan farklı fotoğraflar gösterilmektedir. Fotoğraflarda dolu çöp kovaları, gaz istasyonu, su kirliliği, hava kirliliği, Santa Cruz nehri etrafında yıkılmış ağaçlar gibi durumlar yer almaktadır. Bu alt testte öğrencilerden resimlerde gördükleri bilimsel problemleri listelemeleri, listeledikleri problemlerden birini seçmeleri ve bu probleme çözümler üretmeleri, sonrasında da ürettikleri çözümlerden birini seçerek bilimsel bir araç tasarımları beklenmektedir. Öğrencilerin araç tasarımında kullanmaları için elışı kâğıdı, strafor ve şönil gibi materyallerin olduğu bir paket verilmektedir. Öğrenci tasarımlarının fotoğrafı çekilmektedir. Fotoğraflar yani öğrenci ürünleri konsensüye dayalı değerlendirme yöntemi ile değerlendirilmektedir. Çiçekleri Gruplama alt ölçekte öğrencilere farklı bölgelerde çekilmiş çiçek resimleri sunulmaktadır. Öğrenciden çiçekleri Hilda Taba'nın öğretim stratejilerinden biri olan kavram gelişimi yöntemine göre gruplamaları istenmektedir. Kavram gelişimi yöntemi listeleme, gruplama, etiketleme, sınıflama ve yeniden sınıflama olmak üzere beş aşamadan oluşmaktadır. Bu alt ölçüğün sonunda öğrenci bir ilişki haritası çizer. Tasarım ve Deney alt ölçüğünde öğrencilerden çevreyle ilgili bir problem bulmaları ve

problemi çözmeleri beklenmektedir. Bu alt ölçekte öncelikle öğrencilerden belirledikleri probleme ilişkin hipotez üretmeleri beklenmektedir. Sonrasında öğrencilerden ürettikleri hipotezleri test edebilecekleri bir deney düzeneği çizmeleri beklenmektedir. Her bir alt ölçekte öğrenci ürünlerinin (araç tasarımı, ilişki haritası, deney tasarımı) bilimsel yaratıcılık düzeyi iki uzman tarafından puanlanmaktadır. Araştırmacı bu ölçęęi farklı okullarda öğrenim gören 138 öğrenciye uygulamış ve ölçęęin Cronbach Alfa değerini .90 olarak bulmuştur. Puanlayıcılar arası güvenilirliği incelemek için uzmanların ürünlere verdikleri puanların korelasyonları hesaplanmış ve yeterli düzeyde güvenilirliğe ulaşıldığı rapor edilmiştir. Bilimsel yaratıcılık testi değerlendirildiğinde ölçęęin hazırlanmasında kavramsal bir çerçeveden yararlandığı görülmektedir. Konsensüse dayalı değerlendirme yönteminde uzman sayısının en az 10 olması gerekmektedir (Kaufman, vd., 2008). Ancak bu çalışma kapsamında iki uzman bilimsel yaratıcılığı değerlendirmiştir. Bu bağlamda ölçęęin sınırlı olduğu söylenebilir.

Performans Temelli Deęerlendirme

Performans temelli değerlendirme eğitim alanında kullanılan ölçme değerlendirme yöntemlerinden biridir. Bu yöntemde öğrenciye bir problem ya da problem durumu verilmekte öğrenciden de bu duruma yönelik sözlü yanıt vermesi ya da bir ürün üretmesi beklenmektedir (Nitko, 2004). Performans temelli değerlendirme yöntemi ile öğrencilerin bilgiyi nasıl yapılandırdıkları ve neleri yapabildiklerine yönelik bilgiler edinilir. Öğrencilerin bilişsel düzeyine ilişkin anlamlı bilgilere ulaşılır (Wortham ve Hardin, 2016). Bu yöntemin çeşitli türleri bulunmaktadır. Bunlardan bazıları görev temelli değerlendirme, proje temelli değerlendirme ve görüşmedir (Gullo, 2005; Wortham, 2008). Bu türlerin bilimsel yaratıcılık ölçümlerinde kullanıldığı görülmektedir.

Görev temelli değerlendirmede bireye spesifik bir görev verilir ve bireyin görev sürecindeki performansı değerlendirilir (Wortham, 2008). Eichenberger (1978) bilimsel yaratıcılığı görev temelli bir yöntemle değerlendirmiştir. Üniversitede “Genel Fizik” dersi alan 42 öğrenciye iki görev vermiştir. Araştırma kapsamında bu öğrencilerden dokuzu görevi tamamlamadığı için 33 öğrencinin görev performansı değerlendirilmiştir. Öğrencilere her bir görev için bir hafta süre verilmiştir. Görevlerin nasıl hazırlanacağına yönelik bir sınırlama getirilmemiştir. Öğrenci ürünleri yazılı, sözlü, görsel gibi herhangi bir formatta olabilir. Öğrencilere verilen görevlerden birincisi şudur: “Newton’un üçüncü yasasını (her zaman uygulanan kuvvete karşı, eşit ve ters yönde bir tepki kuvveti vardır) test etmek için mümkün olduğunca çok deney düzeneği tasarlayınız.” (s. 422). İkincisi ise şu şekildedir: “Dünyanın yer çekimi kuvvetinden 10 kat daha fazla yer çekimi kuvveti olan bir gezegende yaşayan canlıların kullanması için olabildiğince çok makine tasarlayınız.” (s. 422). Öğrencilerin ürünleri akıcılık

ve orijinallik kapsamında değerlendirilmiştir. Akıcılık puanı 0 ile 5 (0: Ürün yoksa, 1: Bir ürün, 2: İki ürün, 3: Üç ya da dört ürün, 4: Beş ya da altı ürün, 5: Yedi ya da daha fazla ürün) arasında değişmektedir. Orijinallik puanı ürünün katılımcılar arasında kullanılan sıklığı göz önünde bulundurularak her bir ürün için 1 ile 5 arasında değerlendirilmektedir. Araştırmacı ölçüt geçerliğini test etme amacıyla Torrance Yaratıcılık Testi'ni katılımcılara uygulamış ve bilimsel yaratıcılık puanları ile uyumlu olduğunu rapor etmiştir.

Proje temelli değerlendirme kapsamında bireyler proje geliştirme sürecinde değerlendirilir (Wortham ve Hardin, 2016). Karademir (2016) çalışmasında üstün yetenekli öğrencilerin bilimsel yaratıcılığını proje temelli bir yaklaşımla ölçmeyi hedeflemiştir. Bu amaç doğrultusunda üçüncü ve beşinci sınıfta öğrenim gören 13 üstün yetenekli öğrenciye proje geliştirme adımları öğretilmiş ve projelerinin sonunda günlük hayatı kolaylaştıran bir ürün tasarlanmaları istenmiştir. Öğrencilerin bilimsel yaratıcılık düzeylerini ölçmek için projenin uygulama sürecinde öğrencilerle yapılan yarı yapılandırılmış görüşmeler, öğretmen gözlemleri ve proje uygulamasının sonunda elde edilen çizimler ve hazırlanan görsel materyaller değerlendirilmiştir. Yapılan görüşmelerde akıcılık, esneklik ve orijinallik puanlarının her biri için iki soru olmak üzere toplamda altı soru öğrencilere sorulmuştur. Örneğin akıcılığı ölçmek için “Projeyi hazırlarken hangi bilimsel yöntemleri kullandın?” sorusu, esnekliği ölçmek için “Projende değişiklikler yapman gerekseydi ne tür değişiklikler yapardın?” sorusu, orijinallığı ölçmek için ise “Senin projeyi arkadaşlarının projesinden ayıran özellikler nelerdir?” sorusu sorulmuştur. Öğretmen gözlemleri ürün tasarımı, günlük yaşam problemleri, bilimsel yöntem kullanımı, öz değerlendirme gibi belirlenen ölçütlere göre proje geliştirme sürecinde yapılmıştır. Öğrencilerle yapılan görüşmeler, öğretmen gözlemleri ve öğrenci ürünleri için içerik analizi yapılmış ve iki uzman bu verileri Bilimsel Yaratıcılığın Yapı Modeli'ne göre ürün, özellik ve süreç boyutunda ele almışlardır. Öğrencilerin bilimsel yaratıcılığını proje geliştirme sürecinde nitel bir yaklaşımla değerlendirmişlerdir.

Performans temelli değerlendirme yöntemlerinden biri olan görüşmenin yapılandırılmış görüşme, yapılandırmamış görüşme ve tanılayıcı görüşme olmak üzere üç farklı türü bulunmaktadır (Wortham, 2008). De Vries ve Lubart (2017) 7 ile 10 yaşları arasında 118 çocukla görüşme yaparak bilimsel yaratıcılık düzeylerini ölçmüşlerdir. Görüşmeler sırasında Yaratıcı Potansiyeli Değerlendirme (Evaluation of Potential Creativity, EPoC) ölçeği uygulanmıştır. EPoC 2011 yılında Lubart, Besançon ve Barbot tarafından geliştirilmiştir. Ölçeğin temelinde yaratıcılığın doğasında çoğul ve tekil düşünmenin beraber bir rol oynadığı fikri yer almaktadır. EPoC farklı disiplinlerde çocukların ve genç yetişkinlerin yaratıcılığını ölçmek amacıyla geliştirilmiş bir bataryadır.

Ölçek 7 ile 11 yaşları arasında bireysel, 12 ile 16 yaşları arasında ise grup olarak uygulanmaktadır. Ölçeğin puanlaması internet tabanlı bir programla yapılmaktadır (Lubart, Zenasni ve Barbot, 2013). EPoC bilimsel yaratıcılık ölçeğinde çoğul-keşfedici düşünme becerileri “Bazı kayaları kırmak neden daha zordur?” gibi sorularla ölçülmektedir. Bu sorular için akıcılık ve eşsizlik puanı elde edilmektedir. Tekil- bütüncül düşünme becerilerini ölçme sürecinde öğrencilere “yıldızlar-ışık-gece” gibi kavramlar verilmekte ve bu üç kavramı kullanarak orijinal bir problem oluşturmaları, oluşturdukları problemi çözmeleri ve çözümü kanıtlamaları beklenmektedir. Bu maddeler için orijinallik puanı elde edilmektedir (De Vries ve Lubart, 2017). De Vries ve Lubart bu testi uygularken iki oturumda yaptıkları görüşmeleri kaydetmişlerdir. Araştırma sonunda çoğul ve tekil düşünme becerileri arasında düşük bir ilişki olduğu sonucuna ulaşılmıştır. Araştırmacılar bu bulguyu, her iki becerinin bilimsel yaratıcılığın boyutları olduğu iddiasının bir kanıtı olarak değerlendirmişlerdir.

Performans temelli yaklaşımlarla bireylerin değerlendirilmesinde daha detaylı bilgilere ulaşıldığı görülmektedir. Ancak bu yöntemin uygulaması zor ve zaman alıcıdır. Bu nedenle yöntemin uygulanmasında iyi bir planlama yapılması uygulamanın etkililiğini artıracaktır.

BÖLÜM 3: BİLİMSEL YARATICILIĞI DEĞERLENDİREN YENİ BİR ARAÇ: ÇABİYAT TESTİNİN TASARIMI VE UYGULAMA SÜRECİ

Bilimsel Keşfin İkili Arama Modeli temel alınarak geliştirilen ÇABİYAT anasınıfı, birinci ve ikinci sınıf öğrencilerine yönelik bir testtir. Bu bölümde öncelikle testin içeriğine, uygulama yöntemine ve puanlamasına yer verilerek test tanıtılmış sonrasında ise testin geliştirilme süreci ele alınmıştır.

ÇABİYAT'ın içeriği

Anasınıfı, birinci ve ikinci sınıf öğrencilerinin bilimsel yaratıcılığını ölçmek amacıyla geliştirilen ÇABİYAT deney tasarlama ve hipotez geliştirme olmak üzere iki farklı alt ölçekten oluşmaktadır. Testten fen bilimleri alanında deney tasarlama ve hipotez geliştirme becerilerindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Deney tasarlama becerisi bir hipotezi kanıtlamak amacıyla deney düzeneği tasarlamayı ya da eksik olan deney düzeneğinde değişiklik yapmayı, hipotez geliştirme becerisi ise bir deney düzeneğine bağlı hipotez üretmeyi kapsamaktadır (Klahr, 2000). Hedef kitlenin bilgi düzeyinin sınırlı olması nedeniyle deney tasarlama ve hipotez geliştirme becerilerini ölçmek için kullanılan deney düzenekleri günlük yaşam içerisinde karşılaşılabilecek türden olayların ve durumların hikâyeleştirilmesi yoluyla öğrencilere sunulmaktadır.

Test, her bir alt ölçekte üçer madde olmak üzere toplamda altı maddeden oluşmaktadır. Deney tasarlama alt ölçüğünde Hamsterlar, Kum Havuzu ve Oyuncak Tünel isimli maddeler; hipotez geliştirme alt ölçüğünde ise Bataklıktaki Sinekler, Suluklar ve Oyuncak Gemi isimli maddeler bulunmaktadır. Maddelerin tamamı için animasyon senaryoları tasarlanmıştır. Animasyonlar öğrencilere çoklu ortam çeşitlerini saklamak için kullanılan kısa adıyla MP4 formatında sunulmaktadır. Maddelerin ait olduğu disiplinler ve ölçtüğü beceriler Tablo 3.1'de yer almaktadır.

Tablo 3.1. ÇABİYAT Maddelerine Yönelik Bilgiler

Madde	Disiplin	Beceri
Hamsterlar	Biyoloji	Deney Tasarlama
Kum Havuzu	Fizik	Deney Tasarlama
Oyuncak Tünel	Fizik	Deney Tasarlama
Bataklıktaki Sinekler	Biyoloji	Hipotez Geliştirme
Suluklar	Kimya	Hipotez Geliştirme
Oyuncak Gemi	Fizik/Kimya	Hipotez Geliştirme

Hamsterlar maddesi biyoloji alanında deney tasarlama becerisini ölçmektedir. Maddeden deney tasarlama becerisindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu maddede hamsterların yaşam alanı ile ilgili bir senaryo animasyon ile verilmektedir. Senaryoda bir çocuk ve babası hamsterlar için bir yaşam alanı hazırlamaktadır. Senaryoda baba hamsterlarla ilgili bir hipotez vermekte ve çocuktan bu hipotezi gerçekleştirmesi için hamsterlerin yaşam alanında değişiklikler yapmasını istemektedir. Öğrencilerden değişikliklere yönelik çok sayıda fikir üretmeleri beklenmektedir.

Kum Havuzu maddesi fizik alanında deney tasarlama becerisini ölçmektedir. Maddeden deney tasarlama becerisindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu maddede kum havuzunda topuyla oyun oynayan çocukla ilgili bir senaryo animasyon ile verilmektedir. Senaryoda çocuk parkta oynarken topun kum havuzunda oluşturduğu çukuru fark eder. Senaryoda anlatıcı çukurla ilgili bir hipotez vermekte ve çocuktan bu hipotezi gerçekleştirmesi için çeşitli değişiklikler yapmasını istemektedir. Öğrencilerden değişikliklere yönelik çok sayıda fikir üretmeleri beklenmektedir.

Oyuncak Tünel maddesi fizik alanında deney tasarlama becerisini ölçmektedir. Maddeden deney tasarlama becerisindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu maddede oyuncak tünel ile ilgili bir senaryo animasyon ile verilmektedir. Senaryoda bir çocuk ve halası oyuncak tünel ve oyuncak arabadan oluşan bir oyun düzeneği oluşturmaktadırlar. Senaryoda hazırladıkları düzenek ile amaçlarına ulaşamazlar. Amaçlarına ulaşabilmek için düzenekte çeşitli değişiklikler yapmak zorundadırlar. Senaryoda halası oyuncak tünel ve araba ile ilgili bir hipotez vermekte ve çocuktan bu hipotezi gerçekleştirmesi için oyun düzeneğinde değişiklikler yapmasını istemektedir. Öğrencilerden değişikliklere yönelik çok sayıda fikir üretmeleri beklenmektedir.

Bataklıktaki Sinekler maddesi biyoloji alanında hipotez geliştirme becerisini ölçmektedir. Maddeden hipotez geliştirme becerisindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu

maddede bataklik kenarina giden bir çocukla ilgili bir senaryo animasyon ile verilmektedir. Senaryoda çocuğun aklına batakliktaki sineklerin yaşamıyla ilgili bir soru gelmektedir. Öğrencilerden bu sorunun cevabına yönelik çok sayıda hipotezler üretmeleri istenmektedir.

Suluklar maddesi kimya alanında hipotez geliştirme becerisini ölçmektedir. Maddeden hipotez geliştirme becerisindeki akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu senaryoda bahçede oyun oynayan iki çocuğun yorulduktan sonra suluklarından su içmektedirler. Senaryoda çocuklar sularını içerken beklemedikleri bir durumla karşılaşmaktadırlar. Öğrencilerden ortaya çıkan bu durumun nedenlerine yönelik hipotezler üretmeleri istenmektedir.

Oyuncak Gemi maddesi fizik ve kimya alanında hipotez geliştirme becerisini ölçmektedir. Bu maddeden akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu maddede oyuncak gemi ile ilgili bir senaryo animasyon ile verilmektedir. Senaryoda bir anne, kızı için bir oyun alanı tasarlamaktadır. Oyunda anne, kızından oyuncak gemi ile ilgili bir eylemi gerçekleştirmesini istemektedir. Öğrencilerden söz konusu eylemin gerçekleşmesine yönelik hipotezler üretmeleri istenmektedir.

ÇABİYAT'ın uygulama yöntemi

ÇABİYAT anasınıfı, birinci ve ikinci sınıf düzeyinde eğitim gören, Türkçe konuşan ve Türkçe duyduğunu anlayan öğrencilere uygulanabilmektedir. Test bir uygulayıcı tarafından bireysel şekilde uygulanmaktadır. Testin uygulamasında öğrencilere maddeler animasyonla sunulmaktadır. Öğrencilerin yanıtları ise sözel olarak alınarak uygulayıcı tarafından ÇABİYAT kayıt formuna kaydedilmektedir. Test grup yaratıcılığının değerlendirilmesinin hedeflendiği durumlarda bir uygulayıcı tarafından grup şeklinde de uygulanabilmektedir. Araştırma kapsamında test bireysel şekilde uygulanmıştır. Bu nedenle testin bireysel olarak uygulama yöntemine yer verilmiştir. Testin uygulama adımlarına değinilmeden önce kullanılan araç gereçler, oturma düzeni ve testin uygulama süresi kısaca açıklanmıştır.

ÇABİYAT'ın Uygulamasında Kullanılan Araç Gereçler

Test araç gereçleri bir tablet, kâğıda basılı bir kayıt formu, kalem ve paravandan oluşmaktadır. Kayıt formunda uygulama ve puanlama yönergesi yer almaktadır. Formda test maddelerinin her biri için doğru cevap havuzu bulunmaktadır. Form testin uygulama kitapçığı olarak da işlev görmektedir. Uygulama sürecinde uygulayıcı, öğrencinin ürettiği tüm doğru cevapları test formuna kaydetmektedir. Ayrıca bu süreçte öğrencinin test formunu görmesini engelleyen paravan kullanılmaktadır.

ÇABİYAT'ın Uygulamasında Oturma Düzeni

Testin oturma düzeni Şekil 3.1'deki gibi görselleştirilmiştir. Şekil 3.1'de de görüldüğü üzere uygulama sırasında öğrenci uygulayıcının yanına oturmaktadır. Uygulayıcı test formunu kendi önüne, tableti ise öğrencinin önüne yerleştirmektedir. Aralarında ise öğrencinin formu göremeyeceği bir paravan düzeneği bulunmaktadır. Öğrenci tableti elinde tutabileceği gibi masanın üzerine de koyabilmektedir.



Şekil 3.1. Test Uygulamasında Oturma Düzeni

ÇABİYAT'ın Uygulama Süresi

Testin uygulama süresi, animasyonların ve öğrencilerin cevaplama süresinin toplamından oluşmaktadır. Deney tasarlama alt ölçeğinde yer alan Hamsterlar animasyonu 92 saniye, Kum Havuzu animasyonu 50 saniye, Oyuncak Tünel animasyonu ise 56 saniye sürmektedir. Hipotez geliştirme alt ölçeğinde yer alan Bataklıkta Sinekler animasyonu 55 saniye, Suluklar animasyonu 53 saniye, Oyuncak Gemi animasyonu ise 81 saniye sürmektedir. Test maddeleri için hazırlanan tüm animasyonların süresi 4 dakika 67 saniyedir. Öğrencilerin her bir madde için belirli bir süre sınırı olmamakla birlikte öğrencilerin test maddelerinin tamamını yanıtlama süreleri 10 ile 15 dakika arasında değişkenlik göstermektedir. Her bir soru için öğrencinin cevaplarının bitmesi beklenmektedir. Bu şekilde bir uygulama gerçekleştirilmenin nedeni ise zaman sınırlılığının öğrencinin performansı üzerinde olumsuz etki yaratacağı düşüncesidir. Sonuç olarak testin uygulama süresi 15 ile 20 dakika arasında değişkenlik göstermektedir.

ÇABİYAT'ın Uygulama Adımları

Testin uygulama süreci, uygulayıcının öğrenci ile tanışmasıyla birlikte başlamaktadır. Tanışma sonrasında uygulayıcı öğrenciye teste yönelik genel bir açıklama yapmaktadır: “Bugün seninle bazı çizgi filmler izleyeceğiz. İzlediğimiz her bir çizgi filmin sonunda bir soru sorulacak. Bu sorunun pek çok doğru cevabı var. Senden bu cevapları bulmanı istiyorum. Hazır mısın?” Öğrenci teste hazırlandıktan sonra ilk olarak deney tasarlama alt ölçeğindeki maddeler, sonrasında ise hipotez geliştirme alt ölçeğindeki maddeler sırasıyla uygulanmaktadır. Maddelerin her birinde aynı adımlar takip edilmektedir. Öğrenci animasyonu izledikten sonra uygulayıcı “Çizgi filmin sonunda sorulan soruyu anladın mı?” diye sormaktadır. Öğrenci anladığını söylerse “O zaman cevabın nedir?” diye yeni bir soru yöneltilmektedir. Öğrenci anlamadığını söylerse animasyon tekrar izletilmekte, gerekirse madde kısaca açıklanmaktadır. Öğrencinin yanıtlarına yönelik olarak doğru ya da yanlış gibi açıklama yapılmamaktadır. Fakat öğrencinin yanıtlarını tam olarak anlayabilmek için “Cevabınla ilgili biraz daha açıklama yapabilir misin?” sorusuna benzer sorular yöneltilebilmektedir. Öğrencinin daha fazla yanıt üretmesini sağlamak için yanıtından sonra “Çok güzel cevap verdin. Peki, başka ne olabilir?” diye başka bir soru sorulmaktadır. Öğrencinin her bir yanıtı test formuna kaydedilmektedir. Öğrencilerin sözlü olarak verdiği cevaplar test formunda yer alan, doğru cevap listesinden oluşan tablonun ilgili sütununa işaretlenmektedir. Öğrencinin verdiği doğru cevabın listede olmaması durumunda cevap tabloda boş bırakılan alana yazılmaktadır. Öğrenciden “bu soru için başka bir yanıtım yok” cevabı geldiğinde test maddesi sonlandırılmaktadır.

ÇABİYAT'ın Puanlama Yöntemi

ÇABİYAT bir çoğul düşünme testi olarak tasarlanmıştır. Çoğul düşünme testlerinde bireylerin yaratıcılığı açık uçlu sorularla ya da uyanlarla değerlendirilmekte ve bireylerden birden çok yanıt üretmeleri beklenmektedir (Kaufman, Plucker ve Baer, 2008). Genellikle bireylerin yanıtları akıcılık, esneklik ve orijinallik olmak üzere üç puan türünde incelenmektedir (Runco ve Mraz, 1992). Akıcılık puanı açık uçlu soru ya da uyan için üretilen doğru cevap sayısı, esneklik puanı bireylerin ürettiği cevaplardaki tema ya da kategori sayısı, orijinallik puanı ise üretilen cevapların özgünlüğü ya da seyrekliği ile ilgilidir (Guilford, 1968). Orijinallik puanı testin uygulandığı örneklem sayısı göz önünde bulundurularak elde edilmektedir (Hocevar, 1979).

ÇABİYAT puanlamasında her bir madde için akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları ve testin tamamından toplam akıcılık, toplam esneklik, toplam orijinallik ve toplam yaratıcılık bölümü puanları elde edilmektedir. Akıcılık puanı hesaplanması için üretilen doğru cevap sayısı,

esneklik puanı için cevaplardaki tema sayısı, orijinallik puanı için ise üretilen fikirlerin bulunma sıklığı hesaplanmaktadır. Yaratıcılık bölümü puanının elde edilmesinde $CQ = \log_2 (1+ u_1) + \log_2 (1+ u_2) + \dots \log_2 (1+ u_c)$ formülü kullanılmaktadır (Snyder vd., 2004, s. 416). Formülde yer alan 1, 2... c sayıları temaları; “u” ise bir temada üretilen doğru cevap sayısını temsil etmektedir. Bu değerın hesaplanmasına yönelik bir tablo kayıt formunun son sayfasında yer almaktadır. Toplam akıcılık puanı ÇABİYAT’ta yer alan altı maddenin her biri için hesaplanan akıcılık puanlarının toplanması ile elde edilmektedir. Benzer şekilde toplam esneklik, toplam orijinallik ve toplam yaratıcılık bölümü puanları hesaplanmaktadır. Akıcılık, esneklik ve yaratıcılık bölümü puanlarının nasıl elde edildiği bir örnek ile açıklanabilir. Örnekte yer alan açık uçlu bir maddeyi öğrenciye sorduğumuzu ve öğrencinin ürettiği cevapların örnekteki gibi olduğunu varsayalım.

Örnek madde: Evcilleştirilebilen hayvanlar nelerdir?

Öğrencinin cevapları: Kedi, köpek, at, bal arısı, Japon balığı, papağan, tavuk, eşek, güvercin, yılan, iguana, koyun, sarmaşık, maydanoz

Akıcılık puanı: 12 (Öğrencinin ürettiği doğru cevap sayısı)

Esneklik puanı: 5 [Tema sayısı: Balıklar (Japon balığı), kuşlar (güvercin, tavuk, papağan), memeliler (at, eşek, kedi, koyun, köpek), sürüngenler (iguana, yılan), ve böcekler (bal arısı)]

Yaratıcılık bölümü: Yaratıcılık bölümü puanının hesaplanmasında verilen formülün kullanılmaktadır. Formüle göre, her bir tema için öğrencinin ürettiği doğru cevap sayısına 1 eklenerek 2 tabanına göre logaritması alınır ve tüm temalar için elde edilen değerler toplanır.

Balıklar temasına üretilen doğru sayısı	Kuşlar temasına üretilen doğru sayısı	Memeliler temasına üretilen doğru sayısı	Sürüngenler temasına üretilen doğru sayısı	Böcekler temasına üretilen doğru sayısı
$CQ = \log_2 (1+1) + \log_2 (1+3) + \log_2 (1+5) + \log_2 (1+2) + \log_2 (1+1)$ $= 1 + 2 + 2.58 + 1.58 + 1 = 8.16$				

ÇABİYAT’ın akıcılık, esneklik ve yaratıcılık bölümü puanları örnekteki gibi yapılmaktadır. Orijinallik puanı, üretilen her bir doğru cevabın sıklık değerinin toplanması ile hesaplanmaktadır. Örneğin öğrenci *Oyuncak Tünel* maddesine iki doğru cevap vermişse, öncelikle her bir cevabın orijinallik değeri bulunmaktadır. Sonrasında bu değerler toplanarak orijinallik puanı elde edilmiş olmaktadır. Bir cevabın orijinallik değerinin hesaplanmasında kullanılan formül şu şekildedir:

$$\text{Orijinallik değeri} = 100 - \left(\frac{x}{N} \cdot 100\right)$$

Formülde yer alan x değeri katılımcı grubunda bu doğru cevabı üreten kişi sayısı, N değeri ise örneklem sayısını ifade etmektedir. Orijinallik değeri için öncelikle üretilen fikrin örnekleme bulunma yüzdesi hesaplanmakta sonrasında ise 100'den bu değer çıkarılmaktadır. Örneğin 800 kişilik bir örneklem grubunda *Oyuncak Gemi* maddesi için 32 kişinin A1 cevabını verdiğini varsayalım.

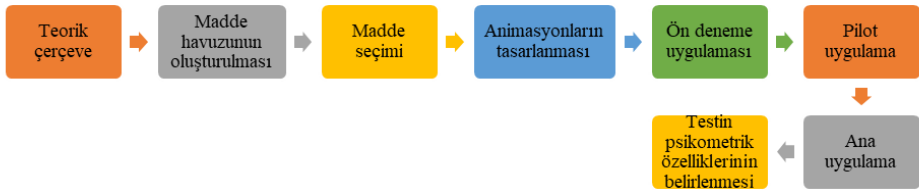
$$\text{Orijinallik değeri} = 100 - \left(\frac{32}{800} \cdot 100\right) = 96$$

Bu cevabın orijinallik değeri için 800 kişilik bir örneklem grubundaki herkesin bir test maddesi için doğru olan cevabı ürettiği varsayıldığında 0 (sıfır) olarak bulunmaktadır. 800 kişilik bir örneklem grubunda sadece 1 kişi, test maddesi için doğru olan cevabı üretmişse bu cevabın orijinallik değeri, yaklaşık 100 olarak bulunmaktadır. Örnekte de olduğu gibi, bir cevabın orijinallik değeri 0 ile 100 arasında değişmektedir.

ÇABİYAT puanlamasında test formunda yer alan doğru cevap listesi kullanılmaktadır. Bu liste testin geliştirilme sürecinde katılımcıların cevaplarına göre oluşturulmuştur. Liste temalara göre ayrılmış ve tablo haline dönüştürülmüştür. Formda yer alan tablolar tema, cevap kodu ve öğrenci yanıtı olmak üzere 3 sütundan oluşmaktadır. Yanıtların temalaştırılması yaratıcılık alanında çalışan bir uzman ve araştırmacı tarafından yapılmıştır. *Oyuncak Tünel* maddesi için 5, diğer maddelerin her biri için ise 4 tema oluşturulmuştur. Katılımcıların cevaplarına göre *Hamsterlar* maddesi için 28, *Kum Havuzu* maddesi için 27, *Oyuncak Tünel* maddesi için 31, *Bataklıkta Sinekler* maddesi için 32, *Suluklar* maddesi için 31, *Oyuncak Gemi* maddesi için ise 26 doğru cevap üretilmiştir.

ÇABİYAT'ın Geliştirme Süreci

ÇABİYAT'ın geliştirilmesinde izlenen aşamalar alan yazında önerilen farklı aşamaların sentezlenmesi ile belirlenmiştir. ÇABİYAT, teorik alt yapının belirlenmesi (Cohen ve Swerdlik, 2002), madde havuzunun oluşturulması (Downing, 2006), madde seçimi (DeVellis, 2017), animasyonların tasarlanması, ön deneme uygulaması (Cohen ve Swerdlik, 2002), pilot uygulama (Cresweel, 2014), ana uygulama (DeVellis, 2017) ve testin psikometrik özelliklerinin incelenmesi (Öner, 1997) olmak üzere sekiz aşamada geliştirilmiştir. Bu aşamalar Şekil 3.2'de görselleştirilmiştir.



Şekil 3.2. Test Geliştirme Aşamaları

ÇABİYAT'ın Teorik Çerçevesi

ÇABİYAT'ın teorik çerçevesi Bilimsel Keşfin İkili Arama Modeli (Klahr ve Dunbar, 1988) ve çoğul düşünme becerilerine odaklanan çalışmalara (Guilford, 1950, 1956, 1967; Torrance, 1972, 1988) dayanarak oluşturulmuştur. Bilimsel Keşfin İkili Arama Modeli'ni ortaya atan Klahr ve Dunbar (1988) bilimsel yaratıcılığın hipotez ve deney olmak üzere iki alanı içerdiğini öne sürmektedirler. Modele göre bilimsel yaratıcılık hipotez ve deney alanındaki arayışların sonucunda ortaya çıkmaktadır. Arayışlar ise iki şekilde gerçekleşmektedir. Bunlardan biri problem çözümünde daha önceki bilgilere göre hipotezler üretme sürecidir. Bu arayışta daha önceden öğrenilen bilgiler hafızadan çağrılarak hipotezler oluşturulmaktadır. Diğer arayış ise problem çözümünde yeni bir hipotez üretebilmek için deney tasarlamaktır. Klahr (2000) bu arayışlarda hipotez geliştirme, hipotez test etme (deney tasarlama) ve kanıt değerlendirme olmak üzere üç temel becerinin öncülük ettiğini ileri sürmektedir. Hipotez geliştirme bir çeşit problem çözme becerisidir ve bilimsel yaratıcılığın ilk adımı olarak kabul edilmektedir. Hipotez test etme ise ilk adımda üretilen hipotezlerin test edilmesi sürecini kapsamaktadır. Bu süreç deney, tahmin, işlem ve eşleştirme olmak üzere dört bileşenden oluşmaktadır. Kanıt değerlendirme deney çıktılarını değerlendirme ve çıktılara göre uygun kararı verebilme becerisidir. Kanıt değerlendirme becerisi hipotez ve deney alanlarında bir köprü niteliğindedir. Modele göre bilimsel yaratıcılık bu üç becerinin etkileşiminden ortaya çıkmaktadır.

ÇABİYAT'ın teorik çerçevesine göre bilimsel yaratıcılık genel yaratıcılık becerileri ile bilimsel becerilerin birleşimi sonucunda ortaya çıkmaktadır. Buna göre bilimsel yaratıcılık iki boyuttan oluşmaktadır. İlk boyutta deney tasarlama ve hipotez geliştirme becerileri, ikinci boyutta ise akıcılık, esneklik, orijinallik ve yaratıcılık bölümü becerileri yer almaktadır. ÇABİYAT'ın teorik çerçevesinin görseli Şekil 3.3'te yer almaktadır.



Şekil 3.3. ÇABİYAT'ın Teorik Çerçevesi

ÇABİYAT'ın teorik çerçevesine göre deney tasarlama bileşeni, bir hipotezi test etmek için hazırlanmış olan bir deney düzeneğinde hipotezi doğru test edebilmek için gerekli olan düzenlemelerin yapılmasını içermektedir. Bullock (1991), çocuklarda deney tasarlama becerisinin üç yolla ölçülebileceğini iddia etmektedir. İlk yol, bir hipotezin doğruluğunu test etmeye yönelik bir deney

düzeneği tasarlanmasıdır. İkinci yol doğruluğu test edilecek hipoteze yönelik tasarlanan deney düzeneğinde düzenlemelerin yapılmasıdır. Üçüncü yol ise doğruluğu test edilecek hipotez için sunulan deney düzeneklerinden uygun olanın seçilmesidir. Bu üç yol yapılandırılmamış, yarı yapılandırılmış ve yapılandırılmış şeklinde sınıflandırılmaktadır. ÇABİYAT'ın deney tasarlama bileşeni yarı yapılandırılmış bir yolla ölçülmektedir.

ÇABİYAT'ın teorik çerçevesine göre hipotez geliştirme bileşeni bir probleme yönelik çözüm üretme sürecidir. Klahr (2000) hipotez geliştirmeyi bir bilginin bir kısmını veya tamamını, önceki bilgilerden yararlanarak genel ya da kısa bir şekilde açıklamak olarak tanımlamaktadır.

Bilimsel Keşfin İkili Arama Modeli (Klahr ve Dunbar, 1988), Bilimsel Üretkenlik Testi'nde de (BÜT) (Ayas ve Sak, 2014; Sak ve Ayas, 2013) teorik çerçeve olarak kullanılmıştır. BÜT'ün teorik yapısına göre bilimsel yaratıcılık; hipotez geliştirme, hipotez test etme ve kanıt değerlendirme olmak üzere üç bileşenden oluşmaktadır. Bu üç beceri akıcılık, esneklik ve toplam yaratıcılık puanları ile ölçülmektedir (Ayas, 2010; Ayas ve Sak, 2014; Sak ve Ayas, 2013). ÇABİYAT'ın kavramsallaştırılması sürecinde alanyazında var olan araştırma bulguları ve hedef kitlenin bilgi düzeyi, bilişsel ve duyuşsal özellikleri göz önünde bulundurularak BÜT'ün teorik çerçevesinde çeşitli değişiklikler yapılmıştır. Bu değişikliklerden biri bilimsel yaratıcılığın bileşenleri arasından kanıt değerlendirme becerisinin çıkarılmasıdır. Araştırma bulguları erken çocukluk döneminde önceki bilgilerin kanıt değerlendirme becerisi üzerinde kirlenici bir etkiye sahip olduğunu ve bu becerinin ortaokul döneminden itibaren tutarlı bir şekilde ölçülebileceğini ortaya koymuştur (Amsel ve Brock, 1996; Bullock, 1991; Kuhn, vd., 1988; Schauble, 1996; Sodian, Zaitchik ve Carey, 1991). BÜT'ten farklı olarak ÇABİYAT'ın teorik çerçevesine orijinallik becerisi eklenmiştir. Bu değişikliğin nedeni ise yaratıcılık kavramında orijinalliğin çok önemli bir ölçüt olarak kabul edilmesidir. Plucker, Beghetto ve Dow (2004) yaratıcılığa yönelik 90 makaleyi incelemiş ve orijinalliğin yaratıcılık tanımlarında yer alan bir özellik olduğunu vurgulamışlardır. Runco ve Jaeger (2012) de yaratıcılık tanımlarında önemli ölçütlerinden birinin orijinallik olduğuna dikkat çekmektedir.

Madde Havuzunun Oluşturulması

ÇABİYAT'ın madde havuzunun oluşturulması sürecinde öncelikle havuzun büyüklüğü (Davey, Pitoniak ve Slater, 2015) ölçütü dikkate alınmıştır. Ölçek geliştirme sürecinin başında madde havuzunun büyüklüğünü belirlemek zordur. Düşük iç tutarlılığa karşı bir önlem oluşturmak için madde havuzunun oldukça geniş hazırlanması önerilmektedir (DeVellis, 2017). Alanyazında ölçeğin madde havuzunda yer alması gereken madde sayısına ilişkin bir fikir birliği

bulunmamaktadır. Buna karşın, yaygın görüşe göre ölçeğin son halinde yer alması planlanan madde sayısının en az 3-10 katı kadar maddenin geliştirilmesi önerilmektedir (Kaplan ve Saccuzo, 2012; Parshall, vd., 2002; Stocking, 1994). ÇABİYAT'ın son formunda deney tasarlama ve hipotez geliştirme alt ölçeklerinin her birinde üç madde olmak üzere toplamda altı maddenin yer alması planlanmıştır. Bu nedenle testin madde havuzunda en az 60 maddenin yer alması hedeflenmiştir.

Madde havuzunun oluşturulması önemli bir aşamadır çünkü maddeleri hazırlama süreci ölçeğin içerik geçerliği ile ilgilidir (Clark ve Watson, 1995; Seçer, 2015). İçerik geçerliği test maddelerinin ölçülmek istenen davranışı, tutumu ya da beceriyi ne ölçüde temsil ettiği ile ilgilidir (Anastasi ve Urbina, 1997; Urbina, 2014). İçerik geçerliğinin, teorik yapıya uygun madde geliştirilmesine bağlı olduğu ifade edilebilir. ÇABİYAT'ın içerik geçerliğini sağlamak için maddeler uzman bir ekiple testin teorik yapısına uygun olacak şekilde geliştirilmiştir.

Madde geliştirme ekibi, beşi erkek dördü kadın olmak üzere dokuz kişiden oluşmaktadır. Ekibin iki üyesi anasınıfı, iki üyesi sınıf öğretmeni, üç üyesi ise fen bilimleri öğretmeni olarak devlet okullarında görev yapmaktadır. Ekibin bir üyesi Bilim Sanat ve Eğitim Merkezi'nde (BİLSEM) fen bilimleri derslerini yürütmektedir. Ekibin son üyesi, lisansını fen bilgisi öğretmenliği bölümünde, lisansüstü eğitimine ise okul öncesi öğretmenliği bölümünde devam etmekte olan bir araştırma görevlisidir. Ekip üyelerinin mesleklerindeki deneyim yılları 8 ile 17 arasında değişmektedir ($\bar{X}= 12.6$; $ss= 3.2$). Ekip üyelerinden biri olan araştırma görevlisi iş yoğunluğu nedeniyle toplantıların yarısından azına katılabılmıştır. Diğer üyeler ise toplantıların tamamına katılmıştır. Ekip üyelerinin oluşturulması sürecinde katılımcılara uzman referansları ile ulaşılmıştır.

Madde havuzu 2017-2018 bahar döneminde sekiz haftalık bir süreç içerisinde geliştirilmiştir. Bu süreçte her hafta sonu araştırmacının moderatörlüğünde ortalama iki saat süren toplantı yapılmıştır. Toplantıların gerçekleştiği saat ekip üyeleri ile ortak olarak belirlenmiştir. Her bir toplantının sonunda ekip üyelerine Millî Eğitim Bakanlığı'nın belirlemiş olduğu ek ders ücreti temel alınarak beş saat üzerinden ücret ödenmiştir. Toplantılarda ekip üyelerinin izni ile ses kaydı alınmıştır. Toplantıların tarihi, süresi ve içeriğine yönelik bilgiler Tablo 3.2'de yer almaktadır.

Tablo 3.2. ÇABİYAT madde havuzunun oluşturulma süreci

	Tarih	Süre	İçerik
1. Toplantı	11.02.2017	2 saat	ÇABİYAT madde geliştirme sürecine yönelik eğitim
2. Toplantı	18.02.2017	1 saat 57 dakika	Hipotez geliştirme maddelerinin değerlendirilmesi
3. Toplantı	25.02.2017	1 saat 49 dakika	Hipotez geliştirme maddelerinin değerlendirilmesi
4. Toplantı	04.03.2017	1 saat 34 dakika	Hipotez geliştirme maddelerinin değerlendirilmesi
5. Toplantı	11.03.2017	2 saat 10 dakika	Deney tasarlama maddelerinin değerlendirilmesi
6. Toplantı	18.03.2017	1 saat 48 dakika	Deney tasarlama maddelerinin değerlendirilmesi
7. Toplantı	25.03.2017	2 saat 8 dakika	Deney tasarlama maddelerinin değerlendirilmesi
8. Toplantı	01.04.2017	2 saat 10 dakika	Herk iki beceri için geliştirilen maddeler

Tablo 3.2’de görüldüğü üzere ilk toplantıda araştırmacı tarafından madde geliştirme sürecine yönelik bir eğitim verilmiştir. Eğitimin başlangıcında ekip üyelerinden ÇABİYAT Geliştirme Grubu Bilgi Formu’nu doldurmaları istenmiştir. Formda madde geliştirme ekibine yönelik demografik sorular yer almaktadır. Bu form doldurulduktan sonra ekip üyeleri birbirleri ile tanışmıştır. Sonrasında eğitim sürecine başlanmıştır. İlk olarak alanyazında yer alan bilimsel yaratıcılığın tanımlarına, bileşenlerine (Chambers, 1964; Heller, 2007; Moravcsik, 1981 vb.) ve nasıl ölçüldüğüne yer verilmiştir. Alan yazında var olan ölçekler (Chin ve Siew, 2015; Hu ve Adey, 2002; Kanlı, 2014a; Sak ve Ayas, 2013; vb.) incelenmiştir. Ölçek incelemeleri yapıldıktan sonra ÇABİYAT’ın teorik yapısı, uygulama ve puanlama yöntemi tanıtılmıştır.

İlk toplantı dışındaki toplantılarda her bir ekip üyesinden ilgili beceri ile ilgili en az iki madde hazırlaması ve toplantı tarihinden iki gün önce araştırmacıya elektronik ortamda göndermesi talep edilmiştir. Madde havuzu oluşturma sürecinde ölçek maddelerinin somut materyallerle öğrenciye sunulması planlanmıştır. Bu nedenle madde havuzu oluşturma ekibinden somut materyallerin kullanabileceği maddelerin hazırlanması beklenmiştir. Maddelerin hazırlanmasında açık uçlu olma, hedeflenen beceriyi ölçme, öğrenme alanına uygun olma, anlaşılır bir dil kullanma ve hedef kitleye uygun olma ölçütleri dikkate alınmıştır.

Tablo 3.2’de görüldüğü üzere ekip üyelerinden ikinci, üçüncü ve dördüncü toplantıda hipotez geliştirme; beşinci, altıncı ve yedinci toplantıda deney tasarlama becerilerine yönelik iki madde geliştirmeleri istenmiştir. Son toplantıda ise hipotez geliştirme ve deney tasarlama becerilerine yönelik birer madde oluşturmaları beklenmiştir. Araştırmacı ekip üyelerinin geliştirdiği maddeleri değerlendirmiş ve çeşitli dönütler vermiştir. Ekip üyeleri dönütlere göre düzeltmeleri toplantıdan önce yaparak araştırmacıya göndermiştir. İlk toplantının sonunda madde oluşturma formu elektronik ortamda ve çıktı halinde ekip üyeleri ile paylaşılmıştır.

İkinci toplantıdan sekizinci toplantıya kadar gerçekleştirilen tüm oturumlarda araştırmacı kendisine gönderilen tüm maddeleri power-point sunumu olarak kaydetmiştir. Toplantılarda her ekip üyesi kendi hazırladığı maddeleri açıklamıştır. Maddelerin tanıtımından sonra tartışmalar yürütülmüş ve çeşitli düzenlemeler önerilmiştir. Araştırmacı toplantı ses kayıtlarına göre düzenlemeleri yapmıştır. Bu tartışmalar sırasında ölçmeyi hedeflediği becerilerden tamamen uzak olduğu konusunda hem fikir olunan maddeler, madde havuzundan çıkarılmıştır. Tartışmalar sırasında uygulanmasına karar verilen maddeleri anasınıfı ve sınıf öğretmenleri kendi sınıf ortamında uygulamış ve sonuçları grupta paylaşmışlardır. Her bir maddeye yönelik tartışmalardan sonra ekip üyeleri madde değerlendirme formuna göre ilgili maddeyi değerlendirmiştir. Madde değerlendirme formu her toplantı için araştırmacı tarafından oluşturulmuştur. Her toplantı başlangıcında da değerlendirme formu ekip üyelerine dağıtılmıştır. Madde değerlendirme formu ölçme ve değerlendirme uzmanlarının önerileri (Lane, vd., 2015; Seçer, 2015) ve ÇABİYAT'ın uygulama yapısına göre oluşturulmuştur. Madde değerlendirme formunda yer alan ölçütler şu şekildedir:

- *İlgili beceriyi ölçme:* Maddeler deney tasarlama veya hipotez geliştirme becerilerini ölçecek şekilde tasarlanmalıdır.
- *Açık uçlu olma:* Maddelere birden fazla doğru yanıt üretilebilmelidir.
- *Anlaşılır dil kullanma:* Maddelerde hedef kitleye uygun ve anlaşılır bir dil kullanılmalıdır. Bilimsel terimler doğru bir şekilde ifade edilmelidir. Kullanılan bilimsel kavramlar yalınlaştırılmalıdır.
- *Hedef seviyeye uygun olma:* Maddeler anasınıfı, birinci ve ikinci sınıf seviyesine uygun olmalıdır. Bununla birlikte, maddeler her öğrencinin en az bir yanıt üretebileceği zorlukta olmalıdır.
- *Hazırbulunuşluk düzeyine uygunluk:* Anasınıfı, birinci ve ikinci sınıf seviyesindeki öğrencilerin fen bilimlerine yönelik dersleri bulunmamaktadır. Bu nedenle maddeler öğrencilerin günlük yaşamlarında karşılaşılabileceği durumlar ve olaylar göz önünde bulundurularak tasarlanmalıdır.

Sonuç olarak toplantılarda hipotez geliştirme alt ölçeği için 65, deney tasarlama alt ölçeği için ise 51 madde geliştirilmiştir. Toplamda madde havuzu için 116 madde tasarlanmıştır. Son toplantıda ekip üyelerine katkılarından dolayı katılım ve teşekkür belgesi verilmiştir.

Madde Seçimi

Madde seçimi bir buçuk aylık bir süreç içerisinde gerçekleştirilmiştir. Madde seçim süreci 2016-2017 bahar dönemi Nisan ayının ilk haftası ile başlamış ve bu süreç Mayıs ayının ikinci haftasının sonunda tamamlanmıştır. Madde

havuzundan uygun maddelerin seçimi üç aşamada gerçekleştirilmiştir. İlk aşamada madde geliştirme ekibinde yer alan anasınıfı ve sınıf öğretmenlerinin öğrencileriyle yaptığı uygulamalar, toplantı ses kayıtları ve madde değerlendirme formu dikkate alınmıştır. İkinci aşamada bilimsel yaratıcılık alanında çalışan bir uzmanla görüşmeler gerçekleştirilmiştir. Üçüncü aşamada ise ikinci aşamada seçilen maddeler uzman görüşüne sunulmuştur.

İlk aşamada madde değerlendirme formundan elde edilen puanlar incelenmiştir. Ekip üyeleri değerlendirme formunda yer alan her ölçütü değerlendirmiş ve bu ölçütleri göz önünde bulundurarak ilgili maddeye 1 ile 5 arasında bir puan vermiştir. Ekip üyelerinin her birinin maddelere verdikleri puan ortalamaları alınmıştır. Ortalaması 3'ün altında olan 31 madde (20 hipotez geliştirme, 11 deney tasarlama), madde havuzundan çıkarılmıştır. Böylece madde havuzundaki madde sayısı 116'dan 85'e düşmüştür. 85 maddenin 45'i hipotez geliştirme maddesi, 40'ı ise deney tasarlama maddesidir. Sonrasında ise anasınıfı ve sınıf öğretmenlerinin öğrencileriyle yaptığı uygulama sonuçları ve toplantı kayıtları incelenmiştir. Ekip üyelerinin hepsinin çok iyi madde olarak değerlendirdiği 40 madde (20 hipotez geliştirme ve 20 deney tasarlama) belirlenmiştir. İlk aşamanın sonucunda madde havuzundaki madde sayısı 116'dan 40'a düşürülmüştür.

İkinci aşamada, 40 maddeyi bilimsel yaratıcılık alanında çalışan bir uzman ve araştırmacı birbirlerinden bağımsız olarak değerlendirmiştir. Uzman ve araştırmacı 40 maddeden 18 madde seçmişlerdir. Seçilen maddelerde farklılık olmasından dolayı bir toplantı yapılmıştır. Toplantıda uzmanların çatıştığı noktalar üzerinde tartışılmış ve toplantının sonucunda fikir birliği sağlanmıştır. Toplantı sonucuna göre 19 madde belirlenmiştir. Seçilen maddelerin ait olduğu disiplinler ve beceriler Tablo 3.3'te yer almaktadır. Tablo 3.3'te görüldüğü üzere maddelerin seçiminde maddelerin becerilere ve disiplinlere göre eşit bir şekilde dağılmasına dikkat edilmeye çalışılmıştır.

Tablo 3.3. ÇABİYAT Madde Seçimindeki Maddelerin Dağılımı

	Fizik (Fiziksel Bilim)	Kimya (Maddenin Doğası)	Biyoloji (Canlı Bilimi)
Deney tasarlama	3 madde	3 madde	3 madde
Hipotez geliştirme	3 madde	3 madde	4 madde

Son aşamada 19 madde, 3 uzmana bir toplantıda sunulmuştur. Uzmanların ikisi özel eğitim alanında biri ise ölçme değerlendirme alanında araştırmalar yapmaktadır. Uzmanlardan biri doçent, ikisi ise profesörlük unvanına sahiptir. Uzmanlar, maddelerin animasyon tasarımına uygunluğuna ve animasyon kullanımının uygulamayı kolaylaştıracağına dikkat çekmiş ve maddelerin animasyonla tasarlanmasını önermiştir. Sonuç olarak maddelerin animasyona dönüştürülmesine karar verilmiştir.

Animasyonları Tasarlanma Süreci

Maddelerin animasyon olarak tasarlanması sürecinde beş aşama takip edilmiştir. İlk olarak maddeler hikâyelere dönüştürülmüştür. İkinci aşamada hikâyeler uzman görüşüne sunularak içerik geçerliği incelenmiştir. Üçüncü aşamada animasyonların tasarımına hazırlık yapılmıştır. Dördüncü aşamada animasyonlar tasarlanmıştır. Son aşamada ise animasyonlar uzman görüşüne sunularak değerlendirilmiştir.

Maddelerin animasyon hikâyelerine dönüştürülmesi

İlk aşamada yaklaşık iki aylık bir zaman diliminde madde seçim sürecinde belirlenen 19 maddenin her biri, yaratıcılık alanında çalışan bir uzman, bir okul öncesi öğretmeni ve araştırmacı tarafından hikâyelere dönüştürülmüştür. Anasınıfı, birinci ve ikinci sınıf öğrencilerinin öğretim programlarında fen bilimleri dersi yer almamasına rağmen fen bilimleri eğitimi çocukların doğal ortamında gerçekleşmektedir (Alisinanoğlu, Özbey ve Kahveci, 2011). Hatta fen bilimleri eğitiminin çocuk doğduğu andan itibaren başladığı söylenebilir. Buldukları dünyayı anlama dürtüsü ile doğan bebekler çevresiyle etkileşimde bulunarak yeni bilgilere ulaşır (Wortham, 2010). Beş duyu organlarıyla algıladığı canlı ve cansız varlıklar hakkında deneyim kazanır ve doğaya yönelik fikirler edinirler (Küçükturan, 2003). Bu döngü yaşamları boyunca devam eder. Bu nedenle hikâyeler çocukların gündelik yaşamlarında karşılaşılabilecekleri olay ya da durumları içerecek şekilde hazırlanmıştır. İlk aşamanın sonunda oluşturulan hikâyelerin dil yapısı, Türk Dili ve Edebiyatı alanında öğretim görevlisi olarak çalışan ve bu alanda dersler veren bir uzman tarafından değerlendirilmiştir. Uzman yaklaşık iki hafta sürecinde her bir hikâyeye için çeşitli düzenlemeler önermiştir. Uzmanın önerdiği düzenlemeler yapıldıktan sonra ikinci aşamaya geçilmiştir.

Hikâyelerin uzman görüşüne sunulması

Animasyon tasarlama sürecinin ikinci aşamasında ölçeğin içerik geçerliğini oluşturmak için 19 animasyon hikâyesi iki farklı uzman grubunun görüşüne sunulmuştur. İlk uzman grubu fen bilimleri öğretmenlerinden, ikinci uzman grubu ise okul öncesi öğretmenlerinden oluşmaktadır. İlk uzman grubundan görüş alındıktan sonra ikinci uzman grubuna ulaşılmıştır. İlk uzman grubunda üçü erkek biri kadın olmak üzere dört fen bilimleri öğretmeni ve bilimsel yaratıcılık çalışan iki erkek akademisyen yer almaktadır. İlk grupta yer alan öğretmenler madde havuzu geliştirme ekibinde yer alan fen bilimleri öğretmenleridir. Bu grupta yer alan uzmanların mesleklerindeki deneyim süresi 5 ile 17 yıl arasında değişmektedir. Uzmanların tamamına elektronik ortamda uzman görüş formu gönderilmiş ve yaklaşık bir haftalık zaman içerisinde görüşleri alınmıştır. Bu gruba hikâyelerde hangi becerinin ölçüldüğü

ve hikâyelerin anlaşılır olup olmadığı sorulmuştur. Bu aşamada maddelerin kuramsal yapıya uygun olup olmadığı, bir başka deyişle ölçeğin içerik geçerliği test edilmiştir. İçerik geçerliğinin test edildiği aşamada, maddelerin ölçekte kalabilmesi için uzmanların uyuşma oranının %90-100 arasında olması koşul olarak kabul edilmiştir (Büyüköztürk, 2014). Uzman dönütleri incelendiğinde 19 maddenin ölçtüğü beceri konusunda %90'nın üzerinde uyum yakalanmıştır. İçerik geçerliği test edildikten sonra hikâyelerin hedef kitlenin en küçük grubu olan anasınıfı öğrencilerine uygunluğunu değerlendirmek amacıyla okul öncesi öğretmenlerinin (ikinci uzman grubu) görüşüne başvurulmuştur. Bu grupta yer alan 15 okul öncesi öğretmenin ikisi erkek on üçü kadındır. Okul öncesi öğretmenlerinin mesleklerindeki deneyim süresi 3 ile 10 yıl arasında değişmektedir. Öğretmenler, Türkiye'nin farklı bölgelerinde Millî Eğitim Bakanlığı'na bağlı okullarda hizmet vermektedir. Öğretmenlere "Hikâye anlaşılır mı?", "Hikâye anasınıfı düzeyine (60-66 ay) uygun mu?" ve "Sınıfınızdaki (60-66 ay) en zayıf öğrenci (kaynaştırma öğrencisi hariç) maddeye kaç cevap üretebilir?" soruları sorulmuştur. Öğretmenlerin görüşleri Tablo 3.4'te sunulmuştur.

Tablo 3.4. Okul öncesi öğretmenlerinin animasyon hikâyelerine yönelik değerlendirmeleri

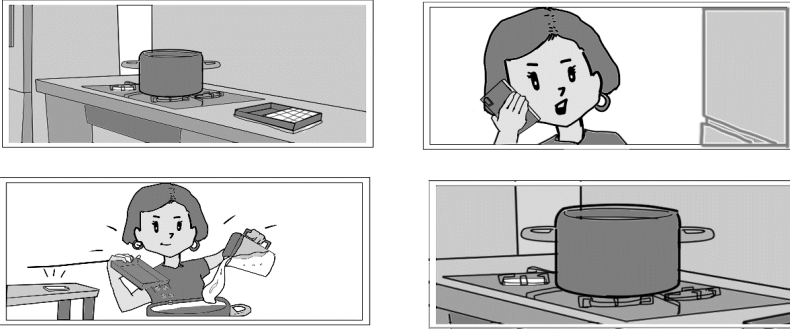
BECERİ: ANİMASYON	ANLAŞILIRLIK		UYGUNLUK		ÜRETİM (\bar{X}) "Sınıfınızdaki en zayıf öğrenci) maddeye kaç cevap üretebilir?"
	f Evet	f Hayır	f Evet	f Hayır	
DT: Buz Küpleri	14	1	14	1	1.5
DT: Hamsterlar	15	0	15	0	1.8
DT: Islak Tişört	15	0	15	0	2.5
DT: Işık Deneyi	11	4	5	10	0.7
DT: İlaç Deneyi	5	10	1	14	0.06
DT: Kum Havuzu	14	1	14	1	1.6
DT: Oyuncak Tünel	12	3	14	1	1.7
DT: Pembe ve Mor Çiçek	9	6	4	11	0.7
DT: Şerbet	14	1	9	6	1.3
HG: Balıklar	9	6	7	8	0.7
HG: Bataklıktaki Sinekler	15	0	12	3	1.2
HG: Boz Ayı ile Kutup Ayısı	15	0	15	0	1.7
HG: Gemi	8	7	12	3	0.9
HG: Kafesteki Kuşlar	5	10	2	13	0.5
HG: Köy	14	1	13	2	1.8
HG: Paraşüt	11	4	7	8	0.6
HG: Sabun	8	7	1	14	0.2
HG: Suluklar	15	0	14	1	1.7
HG: Terazi	15	0	12	3	1.3

DT: Deney Tasarlama, **HG:** Hipotez Geliştirme

Tablo 3.4 incelendiğinde okul öncesi öğretmenlerine göre sekiz maddenin (*Işık Deneyi, İlaç Deneyi, Pembe ver Mor Çiçek, Balıklar, Gemi, Kafesteki Kuşlar, Paraşüt ve Sabun*) zor olduğu görülmektedir. Öğretmenler, öğrencilerin bu maddelere yanıt üretemeyeceği konusunda hemfikirdir. Bu maddelerin büyük bir çoğunluğunda, anlaşılabilirlik ve uygunluk konusunda uyumsuzluklar olduğu görülmüştür. *Gemi* maddesinin üretim ortalaması 1'e oldukça yakındır. Bununla beraber bu maddenin ilgili yaş grubuna uygun olduğu ancak anlaşılmadığı görülmüştür. Bu nedenle *Gemi* maddesinin kurgusu değiştirilmiş ve *Oyuncak Gemi* olarak yeniden adlandırılmıştır. Maddenin senaryosu yeniden dil uzmanına gönderilmiş ve uzmanın önerileri dikkate alınmıştır. Bu madde dışındaki maddeler ölçekten çıkarılmıştır. Bu aşamanın sonunda toplam altı deney tasarlama ve altı hipotez geliştirme olmak üzere 12 maddenin animasyon olarak tasarlanmasına karar verilmiştir.

Animasyon tasarlama sürecine yönelik hazırlıklar

Üçüncü aşama animasyon tasarlama sürecine hazırlık niteliğindedir. Animasyonların tasarlanmasından önce animasyon senaryolarının görsel olarak hazırlanması ve seslendirmelerinin oluşturulması önerilmektedir (Bülbül, 2019; Popp, Tuzinski ve Fetzer, 2016). Bu doğrultuda öncelikle araştırma kapsamında tasarlaması planlanan 12 hikâyenin görsel senaryo taslağı grafik ve animasyon alanında uzman olan bir tasarımcı ile hazırlanmıştır. Örnek olarak *Şeker* maddesi için hazırlanan görsel senaryo taslağı Şekil 3.4'te sunulmuştur.



Şekil 3.4. Örnek Görsel Senaryo Taslağı

Görsel senaryo taslakları hazırlandıktan sonra hikâyelerin her birinde yer alan konuşma metinleri biri kadın biri erkek olmak üzere iki tiyatro oyuncusu tarafından seslendirilmiştir. Seslendirmeler için sanatçıların kelime başına belirledikleri ücret ödenmiştir. Seslendirmeler bir iletişim uzmanı, bir okul öncesi öğretmeni ve araştırmacı tarafından dinlenmiştir. Seslendirmeler teknik açıdan uygunluk, dikkat çekicilik ve vurgulamaların doğruluğu ölçütlerine göre incelenmiştir. Uzmanlardan gelen dönütler doğrultusunda gerekli düzenlemeler yapılmış ve bir sonraki aşamaya geçilmiştir.

Animasyonların tasarlanması

Animasyonlarda çeşitli programlar ile oluşturulan çizgi karakterler ya da gerçek nesnelere kullanılabilir (Popp, Tuzinski ve Fetzer, 2016). Araştırma kapsamında hedef grubun yaş grubunun küçük olması nedeniyle animasyonlarda gerçeğe yakın çizgi karakter kullanılmasına karar verilmiştir. Bu doğrultuda her bir animasyonun senaryosuna bağlı olarak farklı karakterler oluşturulmuştur. Görsel senaryo taslaklarının hazırlanmasına yardımcı olan tasarımcı araştırmacının yönlendirmeleri doğrultusunda animasyonları tasarlamıştır. Tasarımcı, bilgisayar uygulamaları ile 2 boyutlu 12 animasyon hazırlamıştır. Hazırlanan animasyonların süresi 41 saniye ile 117 saniye arasında değişmektedir. Tasarım çalışması belirli bir ücret karşılığı yapılmıştır.

Animasyonların uzman görüşüne sunulması

Geliştirilen animasyonlar üstün yeteneklilerin eğitimi alanında çalışan ve yaratıcılık konusuna hâkim beş uzmanın ve iki animasyon uzmanının görüşlerine sunulmuştur. Uzmanlara, “Animasyon hikâyesi anlaşılır mı?”, “Animasyonda yer alan görseller uygun mu?”, “Animasyon sesleri anlaşılır mı?” ve “Animasyon için diğer önerileriniz nelerdir?” soruları yöneltilmiştir. Uzmanlar her bir animasyon hikâyesinin anlaşılır olduğu konusunda hem fikir olmuşlar, seslendirmeler ve görsel öğelere yönelik çeşitli önerilerde bulunmuşlardır. Animasyonlarda yer alan arka fon müziği sesi yüksek bulunmuş, görsellerde çeşitli mantık hataları tespit edilmiştir. Uzmanlardan gelen dönütler doğrultusunda tasarımcı ile beraber animasyonlarda düzenlemeler yapılmış ve çeşitli sesli ve görüntülü efektler eklenmiştir.

Ön Deneme Uygulaması

Ön deneme, uygulamaya hazır hale getirilen ölçek maddelerinin anlaşılabilirliğini, maddelerde anlaşılmayan yerlerin saptanmasını ve ölçeğin uygulama süresinin belirlenmesini sağlayan ölçek geliştirme adımlarından biridir. Ön deneme uygulaması mutlaka araştırmacının denetiminde gerçekleştirilmeli ve katılımcılardan dönüt alınmalıdır. Ön deneme sürecinin gerçekleştirilmesi araştırmacıya zaman ve emek açısından olumlu katkı sağlamaktadır (Erkuş, 2016).

ÇABİYAT'ın ön deneme uygulaması 17 Nisan ile 20 Nisan 2018 tarihleri arasında Eskişehir Tepebaşı İlçe Milli Eğitim Müdürlüğü'ne bağlı bir ilkokulda 20 öğrenci ile gerçekleştirilmiştir. Öğrencilerin 11'i erkek, 9'u kızdır. Katılımcıların yedisi anasınıfı, yedisi birinci sınıf ve altısı ikinci sınıf öğrencisidir. Öğrenciler çalışmaya katılmaya gönüllü olan öğrenciler arasından rassal olarak seçilmiştir.

Ön deneme uygulaması kapsamında test araştırmacı tarafından bireysel olarak uygulanmıştır. Testin uygulanması iki oturum şeklinde

gerçekleştirilmiştir. İlk oturumda deney tasarlama maddesi olarak geliştirilen maddeler (*Buz Küpler, Hamsterlar, Islak Tişört, Kum Havuzu, Oyuncak Tünel ve Şerbet*), ikinci oturumda ise hipotez geliştirme maddesi olarak geliştirilen maddeler (*Bataklıktaki Sinekler, Boz Ayı ile Kutup Ayısı, Köy, Suluklar, Oyuncak Gemi ve Terazi*) uygulanmıştır. Her bir oturum yaklaşık yarım saat sürmüştür. Oturumlar sırasında her bir animasyon öğrenciye sırayla izletilmiş ve cevapları ÇABİYAT kayıt formuna yazılmıştır. Cevaplama sürecinin bitiminde, öğrenciye animasyonu anlayıp anlamadığı, eğlenip eğlenmediği sorulmuş ve öğrenciden animasyonun zorluk düzeyini değerlendirmeleri istenmiştir. Öğrenci görüşleri uygulama sırasında not alınmıştır.

Ön deneme uygulamasında öğrencilerin *Şerbet* maddesi haricindeki maddelere ürettikleri yanıtların ikiden fazla olduğu görülmüştür. Görüşmeler sırasında *Şerbet* maddesinin zor olduğu belirtilmiştir. Ön deneme uygulaması sürecinde maddelerin animasyon şeklinde olması nedeniyle testin öğrencilerin dikkatini çektiği gözlemlenmiştir. Görüşmeler sırasında öğrenciler test uygulanırken hem eğlendiklerini hem de düşündüklerini belirtmişlerdir. Bu süreçte anasınıfı öğrencilerinin diğer sınıftaki öğrencilere kıyasla daha çekingen olduğu gözlenmiştir.

Ön deneme uygulaması sırasında maddelerde öğrencilerin dikkatini dağıtan görsel ve işitsel öğeler tespit edilmiş ve animasyonlardan çıkarılmıştır. Uygulama sonunda *Şerbet* maddesinin testten çıkarılmasına karar verilmiştir. 11 animasyonda düzenlemeler yapılmış, yaratıcılık alanında çalışan üç uzmanın görüşüne sunulmuş ve uzmanların dönütlerine göre animasyonların son hali oluşturulmuştur. Öğrencilerin yanıtları temel alınarak da cevap havuzu oluşturulmuştur. Ön deneme uygulamasından sonra 11 madde ile pilot uygulamanın gerçekleştirilmesine ve doğru cevap havuzunun kullanılmasına karar verilmiştir.

Pilot Uygulama

Pilot çalışma 30 Nisan-18 Mayıs 2018 tarihleri arasında Eskişehir Tepebaşı İlçe Millî Eğitim Müdürlüğü'ne bağlı bir ilkokulda öğrenim gören birinci ve ikinci sınıf öğrencileriyle, 28 Mayıs-8 Haziran 2018 tarihleri arasında ise yine aynı ilçenin bir anaokulunda öğrenim gören anasınıfı (5 yaş grubu) öğrencileriyle gerçekleştirilmiştir. Pilot uygulamanın katılımcıları 41 anasınıfı, 40 birinci sınıf ve 40 ikinci sınıf öğrencisi olmak üzere toplam 121 öğrenciden oluşmaktadır. Öğrencilerin 67'si erkek, 54'ü kızdır.

Pilot uygulama araştırmacı tarafından 2 oturumda gerçekleştirilmiştir. İlk oturumda 6 animasyon; ikinci oturumda ise 5 animasyon çocuklara izletilmiş ve yanıtları uygulama formuna kaydedilmiştir. İlk oturum yaklaşık 20, ikinci oturum ise 15 dakika sürmüştür. Oturumlar sürecinde öğrencilerin animasyona

yönelik soruları, yorumları ve cevapları nitel veri olarak kullanılmak üzere kaydedilmiştir.

Araştırmacı pilot uygulama sürecinde öğrenci yanıtlarını ve gözlemlerini kaydetmiştir. Nitel bulgulara göre öğrencilerin *Islak Tişört, Boz Ayı ile Kutup Ayısı* ve *Terazi* maddeleri için yanıt üretmede zorlandıkları gözlemlenmiştir. Deneysel tasarlama maddelerine, hipotez geliştirme maddelerine kıyasla daha çok cevap üretildiği gözlemlenmiştir. Öğrencilerin ilk oturumdan sonra, ikinci oturuma çok daha istekli katıldıkları ve oturumlardan keyif aldıkları öğrenci yorumlarından anlaşılmıştır. Oturumlar sırasında bazı öğrencilerin animasyon sahnelerinde yer alan gereksiz detaylara takıldığı ve cevaplarını bu detaylara göre şekillendirdiği tespit edilmiştir. Gözlemlere dayanarak, hedef yaş grubu için ölçekte altı ile sekiz maddenin yer almasının süre açısından uygun olacağı öngörülmüştür.

Pilot uygulama sürecinde elde edilen veriler akıcılık (toplam doğru yanıt sayısı) puanlarına göre analiz edilmiştir. Veri analizi kapsamında madde ortalamalarını incelemek için betimsel istatistikten, maddeler arasındaki bağıntıları incelemek için ise maddeler arası korelasyon değerlerinden yararlanılmıştır. Madde analizi için madde güçlüğü ve madde ayırt edicilik endeksleri hesaplanmıştır. Ölçeğin iç tutarlılığını test etmek için Cronbach Alfa değeri ve doğrulanmış madde toplam korelasyonları incelenmiştir. Son olarak test maddelerinin teorik yapıyı ne seviyede yansıttığını belirlemek amacıyla da açımlayıcı faktör analizi yapılmıştır. ÇABİYAT'ın pilot çalışmasında uygulanan 11 maddeye ilişkin betimsel bulgulara Tablo 3.5'te yer verilmiştir.

Tablo 3.5. ÇABİYAT Pilot Çalışmasında Uygulanan Maddelere Yönelik Betimsel Bulgular

	N	En Düşük	En Düşük Sıklık	En Yüksek	En Yüksek Sıklık	\bar{X}	ss
DT1:Hamsterlar	121	0	1	15	1	5.13	1.928
DT2: Kum Havuzu	121	0	2	12	1	3.91	2.352
DT3: Oyuncak Tünel	121	1	2	13	1	5.12	2.147
DT4: Buzdan Küpler	121	0	1	13	1	4.76	2.236
DT5:Islak Tişört	121	0	1	11	2	3.37	1.954
HG1: Boz Ayı ile Kutup Ayısı	121	1	4	12	1	4.37	2.130
HG2: Köy	121	1	6	9	1	3.69	1.539
HG3:Oyuncak Gemi	121	1	6	13	1	4.90	2.378
HG4: Bataklıktaki Sinekler	121	1	7	15	1	5.09	2.757
HG5:Terazi	121	0	3	8	5	3.01	1.951
HG6: Suluklar	121	0	1	14	1	4.86	2.203
Toplam Akıcılık	121	15	1	114	1	48.21	17.409

DT: Deneysel Tasarlama. HG: Hipotez Geliştirme.

Tablo 3.5 incelendiğinde maddelere üretilen cevapların sayısının 0 ile 15 arasında, ortalamalarının ise 3.01 ile 5.13 arasında değiştiği görülmektedir. Madde ortalamaları değerlendirildiğinde *Kum Havuzu*, *Islak Tişört* ve *Terazi* maddelerinin diğerlerine oranla zor olduğu düşünülebilir. Maddelere en az cevap üretiminin 0 ve 1 olduğu görülmüştür. Ancak maddelere hiç yanıt üretemeyen ya da sadece bir yanıt üreten öğrencilerin sıklığı 1 ile 7 arasında değişmektedir. Bu bulgu öğrencilerin en az yüzde 95'inin maddelere yanıt ürettiğini göstermektedir. Maddelere üretilen maksimum cevap sayısı 8 ile 15 arasında değişmektedir. Bu değerler *Terazi* ve *Köy* maddelerinin diğerlerine kıyasla açık uçluluk bağlamında daha sınırlı olduğuna dair ipucu sunmaktadır. 11 maddelik ÇABİYAT testinden alınan toplam akıcılık puanı 15 ile 114 arasında değişmektedir. Öğrencilerin toplam akıcılıktan aldığı puan ortalaması ise 48.21'dir. ÇABİYAT'ın pilot formunda yer alan maddeler arasındaki ilişkiyi incelemek için Pearson momentler çarpımı korelasyon analizi yapılmıştır. Analiz sonunda elde edilen katsayılar Tablo 3.6'da yer almaktadır.

Tablo 3.6. ÇABİYAT Pilot Maddeleri Korelasyon Katsayıları

	DT2	DT3	DT4	DT5	HG1	HG2	HG3	HG4	HG5	HG6	TA
DT1	.550**	.425**	.415**	.334**	.441**	.461**	.352**	.357**	.175	.314**	.587**
DT2		.626**	.511**	.542**	.508**	.494**	.548**	.559**	.365**	.491**	.772**
DT3			.468**	.559**	.580**	.499**	.543**	.582**	.406**	.499**	.768**
DT4				.535**	.395**	.603**	.450**	.539**	.310**	.482**	.705**
DT5					.435**	.521**	.524**	.685**	.545**	.581**	.774**
HG1						.532**	.514**	.570**	.350**	.386**	.707**
HG2							.502**	.586**	.414**	.488**	.738**
HG3								.648**	.618**	.554**	.782**
HG4									.540**	.636**	.843**
HG5										.524**	.648**
HG6											.742**

DT: Deneysel Tasarlama. HG: Hipotez Geliştirme. DT1: Hamsterlar. DT2: Kum Havuzu. DT3: Oyuncak Tünel. DT4: Buzdan Küpler. DT5: Islak Tişört. HG1: Boz Ayı ile Kutup Ayısı. HG2: Köy. HG3: Oyuncak Gemi. HG4: Bataklıkta Sinekler. HG5: Terazi. HG6: Suluklar. TA: Toplam Akıcılık ** p<.01

Tablo 3.6 incelendiğinde *Hamsterlar* maddesi ile *Terazi* maddesi arasında pozitif ancak anlamsız bir ilişki olduğu görülmüştür ($r = .18$, $p > .05$). Bu maddeler arasındaki ilişki dışında tüm maddelerin aralarındaki ilişkinin pozitif ve anlamlı olduğu bulunmuştur. Maddeler arasındaki ilişki düzeyi orta ya da yüksektir. Ayrıca maddelerin tamamı ile toplam akıcılık puanı arasında pozitif anlamlı ve yüksek düzeyde bir ilişki vardır ($r > .50$, $p < .01$). Bu bulgular maddelerin tamamının ölçüğe hizmet ettiğini göstermektedir.

Pilot uygulamada madde analizi kapsamında madde güçlüğü ve madde ayırt edicilik indeksleri incelenmiştir. Madde ayırt ediciliği için alt ve üst %27'lik

grubun ($N_{alt} = N_{üst} = 33$) puanları bağımsız örneklem t testi ile karşılaştırılmıştır. Tablo 3.7’de madde güçlüğü ve madde ayırt edicilik değerleri yer almaktadır.

Tablo 3.7. ÇABİYAT pilot madde analizine yönelik bulguları

	Madde	Varyansların		Varyansların Eşleşliği durumuna		
	Güçlüğü	Eşleşliği için	Levene	göre t-Testi sonuçları		
	P_i	F	p	T	sd	p
DT1:Hamsterlar	.342	3.131	.082	5.805	63	.000
DT2: Kum Havuzu	.326	8.288	.005	9.946	51.572	.000
DT3: Oyuncak Tünel	.394	2.023	.160	9.000	63	.000
DT4: Buzdan Küpler	.366	11.246	.001	7.290	44.455	.000
DT5:Islak Tişört	.306	15.730	.000	9.214	38.872	.000
HG1: Boz Ayı ile Kutup Ayısı	.364	16.072	.000	7.437	43.397	.000
HG2: Köy	.410	7.152	.010	9.059	45.812	.000
HG3:Oyuncak Gemi	.377	12.895	.001	10.693	45.826	.000
HG4: Bataklıktaki Sinekler	.339	7.122	.010	10.854	47.560	.000
HG5:Terazi	.376	15.097	.000	8.132	45.125	.000
HG6: Suluklar	.347	7.358	.009	7.464	42.314	.000

DT: Deney Tasarlama. HG: Hipotez Geliştirme.

Tablo 3.7 incelendiğinde maddelerin güçlük indeksi değerlerinin .306 ile .410 arasında değiştiği görülmektedir. Bu değerlerin .30 ile .70 arasında olmasından dolayı madde güçlük düzeylerinin ortalama seviyede olduğunu belirtebiliriz. Ölçek için iyi bir maddenin madde güçlüğü değerinin .50’ye yakın olması beklenmektedir (Anastasi ve Urbina, 1997). Bu bağlamda testte yer alan maddelerin kabul edilebilir düzeye yakın olduğu ifade edilebilir. Diğer taraftan testin toplam akıcılık puanlarına göre belirlenen alt ve üst %27’lik grupların her bir madde için ortalama puanları arasındaki fark anlamlı bulunmuştur ($p < .001$). Bu bulgu tüm maddelerin alt ve üst grupları ayırt edebildiğini göstermektedir.

Alt grup üst grup madde ayırt ediciliğine ek olarak maddelerin sınıf düzeyine göre ayırt ediciliği tek yönlü varyans analizi (ANOVA) testi ile incelenmiştir. Her bir madde için bu testin tekrarlanması nedeniyle farkların karşılaştırmasında geleneksel anlamlılık düzeyi (.05) yerine, .004 değeri göz önünde bulundurulmuştur. Bu değer geleneksel anlamlılık düzeyinin yapılan analiz sayısına bölünmesi ile elde edilmiştir (.05/11). Tek yönlü ANOVA sonucuna göre *Oyuncak Tünel*, *Buzdan Küpler*, *Boz Ayı ile Kutup Ayısı*, *Bataklıktaki Sinekler* ve *Suluklar* maddelerinde sınıf düzeylerine göre madde ortalamalarında anlamlı bir farklılık bulunmamıştır ($p > .004$). Ancak *Hamsterlar*, *Kum Havuzu*, *Islak Tişört*, *Köy*, *Oyuncak Gemi* ve *Terazi* maddeleri için sınıf düzeylerine göre anlamlı bir farkın olduğu görülmüştür ($p < .004$). Gerçekleştirilen izleme testlerinde ikinci sınıfların, birinci sınıflardan ve anasınıflarından tüm maddelerde anlamlı bir şekilde daha yüksek puan aldığı ($p < .001$) ancak birinci sınıfların anasınıflarına göre anlamlı bir şekilde

yüksek puan almadığı görülmüştür ($p>.05$). Bu bulgunun pilot uygulama grubundaki birinci sınıf öğrencilerinin puanlarından kaynaklanmış olabileceği düşünülebilir.

Pilot uygulama sürecinde ölçeğin iç tutarlılığını incelemek için Cronbach Alfa değeri ve maddelerin doğrulanmış madde toplam korelasyon değerleri hesaplanmıştır. 11 madde için ölçeğin Cronbach Alfa değeri .913 olarak bulunmuştur. Bu değer ölçeğin iç tutarlılığının oldukça iyi olduğunu göstermektedir. Diğer taraftan herhangi bir maddenin ölçekten çıkarılması durumunda bu değer artmayacağı görülmüştür. Doğrulanmış madde toplam korelasyon değerlerinin ise .507 ile .723 arasında olduğu bulunmuştur. Pallant (2005) bu değerlerin .30 ve üzerinde olması durumunda madde ayırt ediciliğinin yüksek olacağını belirtmektedir. Bu bağlamda test maddelerinin ayırt ediciliğinin iyi olduğu ifade edilebilir.

Veri analizi sürecinde son olarak maddelerin testin teorik yapısını ne seviyede yansıttığını belirlemek amacıyla açımlyıcı faktör analizi yapılmıştır. Madde seçim sürecinde ve faktör sayısını ortaya koymada faktörleştirme tekniklerinden maksimum olasılık faktör analizinden yararlanılmıştır. Bu teknik, küçük örneklem gruplarında diğer tekniklere kıyasla daha güçlü sonuçlar ortaya koymakta (Tanaka, 1987) ve model uyum endekslerinin geniş bir yelpazede hesaplanmasına izin vermektedir (Fabrigar, vd., 1999) Açımlyıcı faktör analizini yapmadan önce kayıp değer, uç değerler, örneklem büyüklüğü, normallik, doğrusallık, çoklu doğrusal bağlantı ve tekliklik (Çokluk, Şekercioğlu ve Büyüköztürk, 2010) sayıltıları test edilmiştir.

Yapılan betimsel analizlerde veri setinde kayıp değer olmadığı bulunmuştur. Veri setindeki normallik incelemeleri için Mahalanobis uzaklığı, Cook uzaklık değeri, merkezi manivela gücü (centered leverage) değerleri ve maddelerin z puanları değerlendirilmiştir. 11 değişken için Mahalanobis kritik değerinin 31.264 ($p<.001$) olduğu (Tabachnick ve Fidell, 2007), Cook uzaklık değerlerinin 1'den yüksek olmaması, merkezi manivela gücü değerlerinin çoğunluğunun .02'nin altında olması ve hiçbir değer .05'in üstünde olmaması gerektiği (Akbulut, 2010, s. 69), z puanlarının ise -3.3 ile +3.3 aralığında olması gerektiği (Tabachnick ve Fidell, 2007) kabul edilmiştir. Bu değerleri karşılamayan 5 öğrenci tespit edilmiş ve veri setinden çıkarılmıştır. Bu durumda öğrenci sayısı 121'den 116'ya düşmüştür. Kline'a göre (1994) örneklem büyüklüğünün değişken sayısına oranı en az 10 olmalıdır. Pilot uygulamada test 11 maddeden oluşmaktadır. Bu durumda örneklem büyüklüğünün faktör analizi için yeterli olduğu ifade edilebilir ($116/11 > 10$). Ayrıca örneklem büyüklüğünün faktör analizine uygunluğunu test etmek için Kaiser-Meyer-Olkin (KMO) değeri incelenmiştir. Pilot verileri ile KMO değeri .921 olarak hesaplanmıştır. Bu değer örneklem büyüklüğünün faktör analizi yapmak için yeterli olduğunu

göstermektedir (Hutcheson ve Sofroniou, 1999). Veri setinin çok değişkenli normallik sayılıştısını test etmek için Bartlett Küresellik Testi yapılmış ve bu değer ($\chi^2_{(55)} = 645.334$; $p < .001$) anlamlı bulunmuştur. Bu bulgu çoklu normallik ölçütünün karşılandığını ortaya koymaktadır. Ancak bu değer tek başına yorumlanmaması, tek değişkenli normalliğin de test edilmesi gerektiği belirtilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Tek değişkenli normallik sayılıştısını test etmek amacıyla çarpıklık ve basıklık katsayıları incelenmiştir. Çarpıklık değerlerinin 0.315 ile 1.284 arasında, basıklık değerlerinin ise -0.103 ile 3.030 arasında değiştiği görülmüştür. Fabrigar ve arkadaşları (1999) maksimum olasılık faktör analizi tekniği kullanıldığında çarpıklık değerinin 2'den basıklık değerinin ise 7'den büyük olmaması gerektiği belirtmektedir. Bu durumda veri setinin tek değişkenli normallik sayılıştısını karşıladığı söylenebilir. Doğrusallık sayılıştısı için saçılma diyagramı (scatter plot) oluşturulmuş ve değişken çiftleri arasında doğrusal bir ilişki olduğu sonucuna ulaşılmıştır. Son olarak çoklu doğrusal bağlantı ve tekillik sayılıştısını incelemek için tolerans değerleri, varyans büyütme faktörü (VIF) değerleri ve maddeler arası korelasyon katsayıları hesaplanmıştır. Tolerans değerlerinin 0.337 ile 0.617 arasında, VIF değerlerinin ise 1.620 ile 2.972 arasında, madde korelasyon katsayılarının ise .201 ile .683 arasında olduğu bulunmuştur. Tolerans değerlerinin 0.10'dan büyük ve VIF değerlerinin 10'dan küçük olması (Akbulut, 2010), korelasyon katsayılarının .90'dan küçük olması (Çokluk, Şekercioğlu ve Büyüköztürk, 2010) veri setinde çoklu doğrusal bağlantı ve tekillik problemi olmadığını göstermektedir. Sonuç olarak pilot uygulama veri setinin faktör analizi için uygun olduğu belirlenmiştir.

Sayıltıların karşılandığı ortaya koyulduktan sonra testin faktör desenini ortaya koyma aşamasına geçilmiştir. Bu aşamada ilk olarak ölçeğin faktör sayısı belirlenmiştir. Faktör sayısının belirlenmesinde Kaiser ölçütü, Catell'in yamaç-birikinti grafiği (scree test) ve paralel analizden yararlanılmıştır. Kaiser ölçütüne göre özdeğeri (eigenvalue) 1 ve üzerinde olan faktörlerin korunması gerekmektedir (Akbulut, 2010). Pilot uygulama verileri ile yapılan açımlayıcı faktör analizi sonucunda, 11 madde için öz değeri 1'in üzerinde olan iki bileşen olduğu görülmüştür. Bu bileşenlerin toplam varyansa yaptıkları katkı %54.927'dir. Söz konusu bileşenlerin varyansa yaptıkları katkı sırasıyla %49.720 ve %5.277'dir. Catell'in yamaç-birikinti grafiğinde ise keskin düşüş birinci noktadan itibaren yaşanmaktadır. Bir başka ifadeyle grafiğe göre test tek faktörlü bir yapıya sahiptir. Son olarak paralel analiz yönteminden yararlanılmıştır. Çalışma kapsamında paralel analiz için Monte Carlo PCA paralel analiz programı kullanılmıştır (Watkins, 2000). Bu program ile 116 katılımcı ve 11 madde için rassal olarak 100 öz değer üretilmiştir. Maksimum olasılık yönteminden elde edilen başlangıç özdeğerleri ile paralel analizden

üretilmiş özdeğerleri karşılaştırılmıştır. Karşılaştırmaların sonucunda, maksimum olasılık yönteminden elde edilen bir özdeğerin paralel analizden üretilen özdeğerden büyük olduğu görülmüştür. Bu bulgu testin tek faktörlü bir yapıya sahip olduğunu göstermiştir. ÇABİYAT'ın teorik çerçevesine göre, deney tasarlama ve hipotez geliştirme olmak üzere iki faktörlü bir yapıya sahiptir. Bu durumda Kaiser ölçütü ve testin teorik çerçevesi dikkate alınarak faktör sayısı iki olarak belirlenmiştir. Faktör sayısı belirlendikten sonra verilerin daha kolay yorumlanabilmesi için eğik döndürme yöntemlerinden direct oblimin yöntemi uygulanmıştır. Hipotez geliştirme ve deney tasarlama becerilerinin ilişkili olduğu varsayılarak bu yöntem tercih edilmiştir. Açımlayıcı faktör analizi iki faktörlü bir yapı için direct oblimin yöntemi seçilerek yinelenmiştir. Analizin sonucunda faktörlerin varyansa yaptığı katkı birinci faktör için %37.467, ikinci faktör için ise %17.230 olarak bulunmuştur. Faktörlerin varyansa yaptıkları ortak katkı ise %54.697'dir.

Faktör desenini ortaya koymak amacıyla gerçekleştirilen faktör analizinde maddelerin ölçekte kalıp kalmamasına yönelik değerlendirmelerde madde toplam korelasyonları, faktör yük değeri, binişik olma durumu, iç tutarlılık katsayısındaki değişiklikler, açıklanan varyans, betimsel bulgular, madde analizleri ve teorik çerçeveye uygunluk açısından değerlendirilmiştir. Faktör yük değeri için kabul değeri .30 olarak belirlenmiştir (Şencan, 2005). Madde faktör yükleri arasındaki farkın .10'un altında olması durumunda maddenin binişik olacağı kabul edilmiştir (Akbulut, 2010).

Yapılan analiz sonucunda. *Boz Ayı ile Kutup Ayısı* maddesinin binişik bir madde olduğu görülmektedir çünkü madde her iki faktör için yakın bir yük değerine sahiptir ($\lambda_1 = .41, \lambda_2 = .48$). Benzer şekilde *Köy* maddesinin de binişik bir madde olduğu görülmektedir ($\lambda_1 = .4, \lambda_2 = .52$). Ayrıca bu iki maddenin çıkarılmasıyla 2 faktörün açıkladığı toplam varyansın %56.945'e yükseldiği görülmüştür. Pilot uygulama sürecinde araştırmacı bu iki maddenin öğrenciler tarafından daha zor anlaşıldığını ve daha fazla açıklama istediklerini gözlemlemiştir. Bu maddeler silinerek analiz yinelenmiştir. Analizin sonucunda ikinci faktör altında deney tasarlama maddesi olarak geliştirilen *Hamsterlar*, *Buzdan Küpler*, *Oyuncak Tünel* ve *Kum Havuzu* maddelerinin yer aldığı görülmüştür. Diğer taraftan birinci faktörün altında hipotez geliştirme maddesi olarak tasarlanan *Terazi*, *Oyuncak Gemi*, *Bataklıktaki Sinекler* ve *Suluklar* maddeleri yer almıştır. Bunlara ek olarak deney tasarlama maddesi olarak geliştirilen *Islak Tişört* maddesinin de bu faktör altında olduğu görülmüştür ($\lambda_1 = .64, \lambda_2 = .48$). Öğrencilerin yanıtları incelendiğinde, bu madde için genellikle benzer ve sınırlı sayıda cevap üretilebildiği fark edilmiştir. İstatistiksel bulgulara ve uygulamaya dönük gözlemlere dayanarak bu maddenin ölçekten çıkarılmasına karar verilmiştir.

İki faktör için tekrarlanan ve üç maddenin analiz dışı bırakılması sonucunda faktörlerin toplam varyansa yaptıkları katkının deney tasarlama faktörü için %28.195, hipotez geliştirme için ise %27.893 olduğu görülmüştür. Bu iki faktörün varyansa yaptıkları toplam katkı ise %56.88 olarak bulunmuştur. Deney tasarlama alt ölçeğinde yer alan dört maddenin faktör yükleri .594 ile .687 arasında, hipotez geliştirme alt ölçeğinde yer alan diğer dört maddenin faktör yükleri ise .610 ile .773 arasında değişmektedir. Deney tasarlama alt ölçeğinin Cronbach Alfa değeri .795, hipotez geliştirme alt ölçeğinin Cronbach Alfa değeri ise .845 olarak bulunmuştur. Toplam ölçek için Cronbach Alfa değeri ise .88 olarak hesaplanmıştır.

Pilot uygulamanın nicel bulguları ve uygulamada yapılan gözlemler doğrultusunda çeşitli değişiklikler yapılmıştır. Animasyonlarda, her maddenin ana fikrinden uzak olan görsel ve sözel kirlenici etkenler, senaryo metinlerinden ve görsellerden çıkarılmıştır. Her animasyonun sonunda sorulan soru anlatıcı tarafından “Sence...” ifadesiyle başlayarak sorulmuştur. Animasyonların her biri yeniden seslendirilmiş ve kurguları güncellenmiştir. Pilot uygulama sürecinde öğrencilerin ürettikleri doğru cevaplar, cevap havuzuna eklenmiştir. Yaratıcılık alanında çalışan bir uzmanla beraber her bir maddenin doğru cevap havuzu temalara ayrılmıştır. Bunlara ek olarak ÇABİYAT formu güncellenmiştir. Animasyonlarda yapılan revizyonlar ve testin uygulama yönergesi, üstün yeteneklilerin eğitimi ve yaratıcılık konusunda çalışmalar yürütmüş beş uzmanın görüşüne sunulmuştur. Uzmanların animasyonlara yönelik görüşleri doğrultusunda animasyonlarda ve ÇABİYAT formunda düzenlemeler yapılarak, test ana uygulama için hazır hale getirilmiştir. ÇABİYAT’ın geliştirilmesi sürecinde kuramsal altyapı, öğrencilerin yaşı ve testin uygulama süresi dikkate alınarak pilot çalışmadan sonra 6 maddenin test için ideal olacağı düşünülmüştür. Pilot uygulama sürecinde gerçekleştirilen aşamalarda 11 maddeden sadece 3’ü çıkarılabilmektedir. Bu nedenle ana uygulamada madde seçimine devam edilmesine karar verilmiş ve daha büyük bir katılımcı grubuna ulaşılması hedeflenmiştir.

Ana Uygulama

ÇABİYAT’ın geliştirildiği bu çalışmanın son aşaması ana uygulamadır. Ana uygulamada veri toplama sürecine başlamadan önce Eskişehir Millî Eğitim Müdürlüğü’nden gerekli izinler alınmıştır. Okulların belirlenmesinde ikili sisteme göre (sabahçı-öğlenci) eğitim verilmesine ve anasınıfı grubunun olmasına dikkat edilmiştir. Aynı gün içerisinde daha fazla sayıda öğrenciyi ulaşmak için ikili sisteme göre örgün eğitim veren okullar tercih edilmiştir.

Ana uygulamanın veri toplama süreci akademik takvime göre 2018-2019 güz döneminde 1 Ekim 2018-15 Ocak 2019 tarihleri arasında 3.5 aylık bir süreçte gerçekleştirilmiştir. Öğrencilerin yanıtlarının puanlanması, yanıt havuzunun

genişletilmesi, ÇABİYAT kayıt formunun yani el kitapçığının oluşturulması ve verilerin elektronik ortama aktarılması ise 2018-2019 bahar döneminde yaklaşık 5 aylık bir süreçte tamamlanmıştır.

Ölçeğin ana uygulama süreci, Eskişehir/Odunpazarı İlçe Millî Eğitim Müdürlüğü'ne bağlı A ve B okullarında araştırmacı tarafından gerçekleştirilmiştir. Araştırmacı veri toplama sürecine başlamadan önce A ve B okullarının idarecileri ile görüşmüş uygulama yapmak için uygun oda talebinde bulunmuştur. Okul idaresinin öğretmenleri bilgilendirmesi talep edilmiştir. Uygulamalar A okulunda kütüphanede, B okulunda ise bireysel çalışma odasında gerçekleştirilmiştir. A okulunda kütüphane güz döneminde öğrencilerin kullanımına kapalı olduğu için kullanılmıştır. Uygulama yapılan alanların ikisi de yeteri kadar aydınlık, havası temiz ve ses yalıtımı uygundur. Uygulama alanlarına bir masa ve iki sandalye yerleştirilmiştir. Araştırmacı uygulama yapılan odaları uygulama sürecinden önce uygulamaya hazır hale getirmiştir.

Araştırmacı, veri toplama sürecinde hafta içi her gün 7.00 ile 18.00 saatleri arasında uygulama okullarında bulunmuş, çalışma katılımcılarının her biri ile bireysel uygulama yapmıştır. Uygulama sürecinde öğrencilerin her biri ders sırasında sınıftan alınarak uygulama odalarına götürülmüştür.

Uygulamaların gerçekleştirilmesinde, ÇABİYAT uygulama yönteminde yer alan adımların hepsi sırasıyla izlenmiştir. Uygulamalarda, pilot uygulama sonrasında revize edilen 8 madde için animasyonlar tek oturumda öğrencilere izletilmiş ve yanıtları kaydedilmiştir. Uygulama sürecinde önce deney tasarlama alt ölçeğinde yer alan animasyonlar, sonra hipotez geliştirme alt ölçeğinde yer alan animasyonlar izletilmiştir. Gerçekleştirilen uygulamaların tamamında animasyonlar aynı sırada izletilmiştir. Bir öğrenci ile uygulama süresi yaklaşık 25 dakika sürmüştür. ÇABİYAT uygulaması tamamlandıktan sonra ortalama üç dakika süren fen bilimleri başarı testi uygulanmıştır. Teste ilişkin bilgilere bir sonraki başlıkta yer verilmiştir.

Ana uygulama kapsamında öncelikle A okulunda yer alan tüm anasınıfı, birinci ve ikinci sınıf şubelerinde uygulama yapılmıştır. A okulunda 5 tane ikinci sınıf, 5 tane birinci sınıf ve 4 tane anasınıfı şubesi bulunmaktadır. Öncelikle 145 ikinci sınıf öğrencisiyle, sonra da 147 birinci sınıf öğrencisi ile uygulamalar gerçekleştirilmiştir. En son olarak da 72 anasınıfı öğrencisi ile uygulama yapılmıştır. Anasınıfı öğrencilerinin araştırmacıya alışması amacıyla uygulamalar gerçekleştirilmeden önce araştırmacı bir ders saatinde sınıf öğretmeni ile birlikte derse katılmıştır. Uygulama sırası sınıf yoklama listesine göre yapılmıştır.

A okulundaki uygulamalar tamamlandıncaya B okulunda uygulamaya devam edilmiştir. B okulunda yer alan birinci ve ikinci sınıf şubeleri arasından beşer

şube rassal olarak belirlenmiştir. B okulunda yer alan 7 anasınıfı şubesinin tamamı ile uygulama yapılmıştır. A okulundaki uygulamalara benzer şekilde öncelikle 157 ikinci sınıf öğrencisiyle, sonra da 131 birinci sınıf öğrencisi ile uygulamalar gerçekleştirilmiştir. En son olarak da 149 anasınıfı öğrencisi ile uygulama yapılmıştır. Anasınıfı öğrencileri ile uygulama gerçekleştirilmeden önce araştırmacı yemek saatlerinde öğrencilerle vakit geçirmiştir. A ve B okullarında uygulama yapılan şubelere kayıtlı Türkçe bilmeyen yabancı uyruklu öğrenciler çalışma grubuna dâhil edilmemiştir. Sonuç olarak 801 öğrenci veri toplama süreci araştırmacı tarafından tamamlanmıştır.

Ana Uygulama Çalışma Grubu

Çalışma grubunun belirlenmesinde olasılıklı olmayan örnekleme yöntemlerinden amaçlı örnekleme ve elverişlilik örneklemesinden yararlanılmıştır. Amaçlı örneklemede araştırmacı belirli özelliklere sahip bireylere ulaşır (Johnson ve Christensen, 2014). Elverişlilik örneklemesinde ise araştırmacı daha kısa sürede, daha ekonomik yollarla ve daha kolay bir şekilde ulaşabileceği bireyleri araştırmaya dâhil eder (Böke, 2017). Bu çalışmada geliştirilen yaratıcılık testinin anasınıfı, birinci sınıf ve ikinci sınıf öğrencilerine yönelik olması nedeniyle bu sınıf düzeyindeki öğrencilere ulaşmada amaçlı örneklemeden, bu öğrencilerin öğrenim gördükleri okulların belirlenmesinde ve bu okullarda öğrencilerin seçiminde elverişlilik örneklemesinden yararlanılmıştır.

Araştırmanın ana uygulaması Eskişehir/Odunpazarı ilçesinde bulunan iki ilkokulda öğrenim gören 221 anasınıfı, 278 birinci sınıf ve 302 ikinci sınıf öğrencisi olmak üzere 801 öğrenciyle gerçekleştirilmiştir. Çalışma grubunun sınıf düzeyi, cinsiyet ve yaş değişkenlerine göre dağılımına Tablo 3.8’de yer verilmiştir.

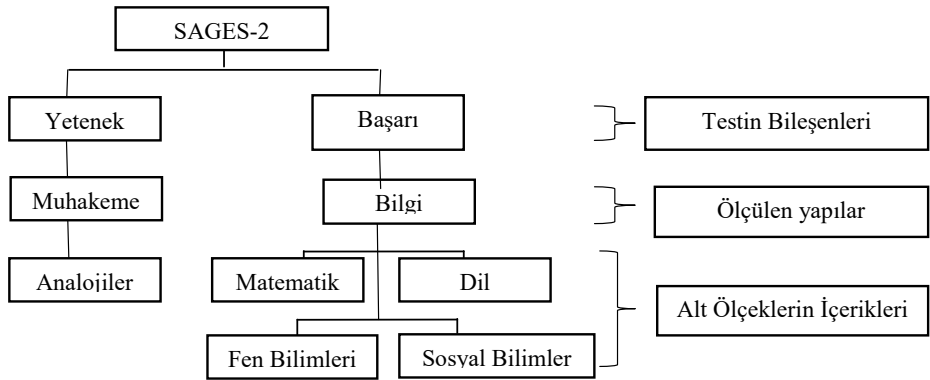
Tablo 3.8. Katılımcıların Sınıf Düzeyi, Cinsiyet ve Yaş Değişkenlerine Göre Dağılımları

Sınıf	Cinsiyet	Sıklık (f)	Yüzde (%)	YAŞ	Sıklık (f)	Yüzde (%)
Anasınıfı	Erkek	115	52	4 (54-59 Ay)	12	5.4
	Kız	106	48	5 (60-71 Ay)	207	93.7
	Toplam	221	100	6 (72-83 Ay)	2	0.9
1. Sınıf	Erkek	152	54.7	5 (60-69 Ay)	8	2.9
	Kız	126	45.3	6 (72-83 Ay)	265	95.3
	Toplam	278	100	7 (85-90 Ay)	5	1.8
2. Sınıf	Erkek	149	49.3	6 (76 Ay)	1	0.3
	Kız	153	50.7	7 (84-95 Ay)	291	96.4
	Toplam	302	100	8 (96-98 Ay)	10	3.3
Toplam	Erkek	416	51.9	4 (54-59 Ay)	12	1.5
	Kız	385	48.1	5 (60-71 Ay)	215	26.8
	Toplam	801	100	6 (72-83 Ay)	268	33.5
				7 (84-95 Ay)	296	37
			8 (96-98 Ay)	10	1.2	

Tablo 3.8 incelendiğinde anasınıfı öğrencilerinin 115'ü erkek, 106'sı kız, birinci sınıf öğrencilerinin 152'si erkek, 126'sı kız, ikinci sınıf öğrencilerinin 149'u erkek, 153'ü kız öğrenci olduğu görülmektedir. Çalışma grubunun %51.9'u (416) erkek, %48,1'i (385) ise kız öğrencidir. Anasınıfı öğrencilerinin yaşları 4 ile 6 arasında (54-74 ay) değişmektedir. Yıl olarak yaş ortalamaları 4.95 (ss=0.248), ay olarak yaş ortalamaları ise 62'dir (ss=2.864). Birinci sınıf öğrencilerinin yaşları 5 ile 7 arasında (60-90 ay) değişmektedir. Yıl olarak yaş ortalamaları 5.99 (ss=0.216), ay olarak yaş ortalamaları ise 76.62'dir (ss=3.848). İkinci sınıf öğrencilerinin yaşları 6 ile 8 arasında (76-98 ay) değişmektedir. Yıl olarak yaş ortalamaları 7.03 (ss=0.189), ay olarak yaş ortalamaları ise 88.75'tir (ss=3.516). Anasınıfı öğrencilerinin çoğunluğu 5 yaşında, birinci sınıf öğrencilerinin çoğunluğu 6 yaşında, ikinci sınıf öğrencilerinin çoğunluğu ise 7 yaşındadır.

Ana Uygulama-Fen Bilimleri Başarı Testinin Uyarlanması

Araştırma kapsamında geliştirilen ÇABİYAT'ın ölçüt geçerliğini test etmek amacıyla katılımcılara SAGES 2'nin (Screening Assessment for Gifted and Middle School Students-Second Edition) fen bilimleri alt testi uygulanmıştır. SAGES-2 5-14 yaş aralığındaki çocuklara bireysel veya grup olarak uygulanabilmektedir (Johnsen ve Corn, 2001). SAGES-2'nin bileşenleri, ölçülen beceriler ve alt ölçek içerikleri Şekil 3.5'te görselleştirilmiştir.



Şekil 3.5. SAGES-2'nin Teorik Alt Yapısı ve İçeriği (Johnsen ve Corn, 2001, s. 15)

Şekil 3.5'te de görüldüğü üzere SAGES-2 yetenek ve başarı olmak üzere iki alana odaklanmaktadır. Bu iki alanın ölçümü için 3 alt test geliştirilmiştir. Muhakeme becerisinin ölçülmesi için bir alt test oluşturulmuştur. Bu alt test analogi maddelerini kapsamaktadır. Bilgi iki alt test ile ölçülmektedir. Bir alt test fen bilimleri ve matematik başarı testlerinden, diğer alt test dil bilimleri ve sosyal bilimler başarı testlerinden oluşmaktadır. SAGES-2'de yer alan 3 alt testin iki formu bulunmaktadır. Bir formu anasınıfı-3.sınıf düzeyleri

arasındaki öğrencilere, diğer formu ise 4-8. sınıf düzeyleri arasındaki öğrencilere uygun olarak tasarlanmıştır. Bataryada yer alan tüm alt ölçekler çoktan seçmeli maddelerden oluşmaktadır. Bataryanın hem bireysel hem de grup şeklinde uygulamasında uygulayıcı maddeleri okumakta öğrencilerden ise doğru yanıtı seçmeleri beklenmektedir. Bireysel uygulamalarda uygulayıcı öğrencilerin yanıtlarını kaydetmekte, grup uygulamalarında ise öğrenci kendi yanıtını işaretlemektedir. Uygulamalardan sonra öğrencinin doğru yanıtlarının sayısı standart puanlara dönüştürülerek öğrenciye yönelik değerlendirmeler yapılmaktadır. Testin geliştirilmesi kapsamında 5313 öğrenci ile madde analizi ve güvenilirlik çalışmaları yapılmış, ölçeğin geçerli ve güvenilir olduğu sonucuna ulaşılmıştır (Johnsen ve Corn, 2001).

SAGES-2'nin anasınıfı-3. sınıf düzeyleri arasındaki öğrencilere uygun formunda yer alan fen bilimleri başarı testi 14 maddeden oluşmaktadır. Öğrenciden her bir madde için doğru yanıtı beş seçenek arasından bulması beklenmektedir. Seçenekler 14 maddenin onunda görsel formatta, dördünde ise metin şeklindedir. Testin başlangıcında öğrenciye örnek madde uygulanmaktadır.

SAGES-2 genel olarak değerlendirildiğinde, testin uygulamasının pratik olduğu görülmektedir. Tarama değerlendirme bataryasında yer alan başarı testleri öğrenci potansiyeline yönelik değerlendirmelere olanak sağlamaktadır. Testin maddeleri incelendiğinde, maddelerin evrensel bilgiye dayandığı dikkat çekmektedir. Bu bağlamda bataryada yer alan fen bilimleri başarı testinin uyarlanarak kullanılmasının araştırmaya katkı sağlayacağı düşünülmüştür. Bu doğrultuda SAGES-2'nin anasınıfı-3. sınıf düzeylerindeki öğrencilerine yönelik fen bilimleri başarı testi uyarlanmıştır.

SAGES 2'nin uyarlama süreci

Anadolu Üniversitesi Üstün Yeteneklilerin Eğitimi Uygulama ve Araştırma Merkezi (ÜYEP), SAGES-2 bataryasının kullanım haklarına sahiptir. Bu bataryada yer alan matematik-fen bilimleri başarı alt testinin kullanım alanı araştırma amaçlarına uygun bulunmuştur. Dolayısıyla alt ölçekte yer alan 14 maddeden oluşan fen bilimleri başarı alt testi uyarlanmıştır. Testin çevirisi, pilot çalışma ve geçerlik-güvenirlik çalışmaları olmak üzere üç aşamada testin uyarlanması tamamlanmıştır (Hambleton ve De Jong, 2005; Harkness, Villar ve Edwards, 2010).

Çeviri, ölçek uyarlama sürecindeki en önemli aşamadır. Bu adımda ölçek, kaynak dilden hedef dile çevrilmektedir. Ölçeğin hedef dile çevrilmesi sürecinde bir çeviri ekibinin oluşturulması önerilmektedir (Harkness, Villar ve Edwards, 2010). Fen bilimleri başarı testinin orijinal dili İngilizcedir. Testin uyarlanması sürecinde İngilizce ve Türkçeyi anadili gibi bilen üç uzman, bir İngiliz Dili Edebiyatı uzmanı ve araştırmacı olmak üzere beş kişilik bir çeviri ekibi oluşturulmuştur. Ekip üyeleri eş zamanlı olarak test maddelerini

Türkçeye çevirmiştir. Çeviriler bir tabloda araştırmacı tarafından birleştirilmiş ve karşılaştırılmıştır. Karşılaştırmada aşağıdaki sorular göz önünde bulundurulmuştur (Seçer, 2015, s. 69):

- Çeviri metinler arasında farklılıklar var mı?
- Çeviriler anlamsal ve kuramsal olarak uygun mu?
- Kültürel uyarlamalar yapılmış mı?

Karşılaştırmının sonucunda madde çevirilerinde bazı farklılıklar olduğu görülmüştür. Ekip üyelerinin de görüşleri doğrultusunda ölçek çevirisi ortaklaşa tamamlanmıştır. Sonrasında Türkçeye uyarlanmış ölçek anasınıfı ve ilkökul düzeyinde İngilizce eğitimi veren bir dil uzmanı tarafından İngilizceye çevirmiştir. İngilizce çevirinin orijinal ölçek ile tutarlı olduğu sonucuna ulaşılmıştır. Çeviri aşaması tamamlandıktan sonra, yaratıcılık alanında çalışan iki uzmandan, bir fen bilimleri öğretmeninden ve bir okul öncesi öğretmeninden ölçek maddelerinin anlaşılabilirliği, hedef yaşa uygunluğu, görsellerin uygunluğu ve kültürel uygunluk kapsamında değerlendirmeleri istenmiştir. Uzmanlardan gelen dönütler doğrultusunda seçeneklerde yer alan iki resmin Türk kültürüne uygun olmadığı gerekçesiyle değiştirilmiştir. Ayrıca iki maddenin soru kökü, yaş grubunun özelliklerine göre düzenlenmiştir. Uzmanlardan gelen düzenlemelerle beraber Fen Bilimleri Testi, pilot uygulamaya uygun hale getirilmiştir.

Pilot çalışma, ölçek uyarlamasının ikinci adımıdır (Hambleton ve De Jong, 2005). Fen bilimleri başarı testinin uyarlanması kapsamındaki pilot çalışma Eskişehir/Tepebaşı ilçesindeki bir anaokulunda öğrenim gören 15 öğrenci (5 yaş grubu) ile gerçekleştirmiştir. Test bir anasınıfı öğretmeni tarafından uygulanmıştır. Araştırmacı uygulamadan önce anasınıfı öğretmene testin uygulamasına yönelik bir eğitim vermiştir. Testin uygulama sürecinde, uygulayıcı bir soruyu okur ve öğrenciden 5 seçenek arasından doğru olanı seçmesini ister. Bu durum tüm maddeler için tekrar eder. Pilot uygulama sırasında uygulayıcıdan öğrencilerin verdikleri cevapları ve yaptıkları yorumları not alması istenmiştir. Uygulayıcı pilot çalışmayı yaklaşık bir saat içerisinde tamamlamıştır. Uygulayıcı pilot çalışma sonunda tüm maddelerin öğrenciler tarafından anlaşıldığını ve testin uygulamasının kolay olduğunu belirtmiştir. Uygulayıcı testin uygulama yönergesine yönelik çeşitli önerilerde bulunmuştur. Öğretmenin önerileri dikkate alınarak çeşitli düzenlemeler yapılmış ve ölçek hazır hale getirilmiştir.

Fen bilimleri başarı testinin uyarlanmasının son adımında geçerlik-güvenirlilik çalışmaları yapılmıştır. Bu aşamada ölçek araştırmacı tarafından 801 öğrenciye uygulanmıştır. Veri analizi kapsamında madde analizi ve iç tutarlılık analizi yapılmıştır. Yapılan analizler sonucunda madde güçlük değerlerinin .29 ile .83 arasında değiştiği, testin toplam güçlük endeksinin ise

.58 olduğu bulunmuştur. Maddelerin ayırt edicilik endekslerinin .30 ile .59 arasında değiştiği görülmüştür. Testin iç tutarlılığının test edilmesinde Kuder-Richardson-20 (KR-20) değerinden yararlanılmıştır. Çünkü maddelerin 0-1 (doğru-yanlış) şeklinde kodlandığı durumlarda KR-20 değerinin kullanılması önerilmektedir (Pallant, 2005). Huck (2012) maddelerin 0-1 (doğru-yanlış) şeklinde kodlandığı testlerde, KR-20 değerinin Cronbach Alfa değeri ile eşit çıkacağını ifade etmektedir. Bu nedenle testin Cronbach Alfa değeri hesaplanmış ve .664 değerine ulaşılmıştır. Özdamar (2004) Cronbach Alfa değerinin .60 ile .80 aralığında olması durumunda ölçeğin güvenilir olduğunu belirtmektedir. Bu bağlamda fen bilimleri başarı testinin güvenilir olduğunu ifade edebiliriz. Testin norm çalışması yapılmamıştır. Bu nedenle fen bilimleri başarı testinden elde edilen ham puanların ÇABİYAT puanları ile karşılaştırılmasına karar verilmiştir.

Testin Psikometrik Özelliklerinin Belirlenmesi

ÇABİYAT'ın psikometrik özelliklerinin belirlenmesi kapsamında testin yapı geçerliği ve ölçüt geçerliği test edilmiştir. Yapı geçerliği için öncelikle açımlayıcı faktör analizi yapılmış ve madde seçim süreci tamamlanmıştır. Sonrasında testin 6 maddelik son formu üzerinden geçerlik ve güvenilirlik analizleri yapılmıştır. Analizlerden elde edilen bulgulara Bölüm 4'te yer verilmiştir.

BÖLÜM 4: ÇABİYAT'IN PSİKOMETRİK ÖZELLİKLERİ

Bu bölümde ÇABİYAT'ın psikometrik özelliklerine yönelik bulgular sunulmuştur. Öncelikle betimsel bulgular ve yapılan açımlayıcı faktör analizi ile madde seçimine yer verilmiş, sonrasında ise testin altı maddelik son formu üzerinden testin psikometrik özellikleri incelenmiştir.

BETİMSEL analizler

Madde seçimi gerçekleştirilmeden önce maddelerin akıcılık puanına ilişkin betimsel bulgular incelenmiştir. 801 katılımcının yanıtları göz önünde bulundurularak ölçekte yer alan maddelere üretilen en düşük puan, en yüksek puan, madde ortalamaları ve standart sapmaları Tablo 4.1'de verilmiştir.

Tablo 4.1. ÇABİYAT'ın Betimsel Bulguları

Madde	En Düşük	En Düşük Sıklık	En Yüksek	En Yüksek Sıklık	\bar{X}	ss
DT1: Hamsterlar	1	12	13	1	6.08	2.10
DT2: Kum Havuzu	0	3	12	2	5.60	2.15
DT3: Oyuncak Tünel	0	1	13	1	5.54	2.00
DT4: Buz Küpleri	0	159	10	1	2.41	1.80
HG1: Bataklıktaki Sinekler	0	1	11	2	4.65	1.68
HG2: Suluklar	0	4	12	1	4.42	1.81
HG3: Oyuncak Gemi	0	5	14	1	4.35	1.85
HG4: Terazı	0	149	9	1	2.11	1.63

DT: Deney Tasarlama. HG: Hipotez Geliştirme.

Tablo 4.1'de yer alan değerler incelendiğinde maddeler için üretilen en yüksek doğru yanıt sayısının 9 (HG4) ile 14 (HG3) arasında değiştiği görülmektedir. En az yanıt üretilen maddenin HG4 (2.11), en çok yanıt üretilen maddenin ise DT1 (6.08) olduğu görülmüştür. Tüm öğrenciler DT1 maddesine en az 1 doğru yanıt üretmiştir. DT4 ve HG4 maddelerine hiç yanıt üretemeyen öğrenci sayısının birbirine yakın ve diğer maddelere yanıt üretemeyen öğrenci sayısından oldukça fazla olduğu görülmektedir. Bu nedenle iki maddenin (DT4 ve HG4) diğerlerine göre daha zor olabileceği düşünülebilir.

Geçerlik analizleri

ÇABİYAT'ın geçerliğini test etmek için ölçeğin yapı geçerliği, ölçüt geçerliği ve maddelerin yanlılığı incelenmiştir. Testten akıcılık, esneklik, orijinallik ve yaratıcılık bölümü olmak üzere dört puan türü elde edilmektedir. Maddelerin esneklik ve orijinallik puanları öğrencilerin doğru yanıt sayısına (akıcılık) bağlıdır. Yaratıcılık bölümü puanı ise akıcılık ve esneklik puanlarından elde edilmektedir. Bütün puan türleri aynı maddeler üzerinden elde edildiğinden ölçeğin yapı geçerliğine yönelik analizler akıcılık puanı üzerinden yapılmıştır. Diğer analizler ise tüm puanlar üzerinden gerçekleştirilmiştir.

Yapı Geçerliği

Yapı geçerliği kapsamında araştırmacının önerdiği yapıya yönelik varsayımlarında yanılma olasılığı bulunmaktadır. Bu durum göz önünde bulundurularak yapının irdelenmesi ve ölçek maddelerinin niteliğinin daha güçlü bir şekilde ortaya koyulması için açımlayıcı faktör analizinden sonra doğrulayıcı faktör analizinin yapılması önerilmektedir (Carpenter, 2018). Açımlayıcı faktör analizi ile elde edilen yapı, farklı bir veri seti üzerinden doğrulayıcı faktör analizi ile doğrulanmalıdır (Costello ve Osborne, 2005; Kline, 2013). Bu duruma yönelik çeşitli öneriler bulunmaktadır. Akbulut (2010) veri setini ikiye bölerek, bir yarısıyla açımlayıcı diğer yarısıyla da doğrulayıcı faktör analizlerinin yapılabileceğini belirtmektedir. Ölçek geliştirme sürecinde yeterli örneklem sayısına ulaşıldığında örneklem farklı gruplara ayrılarak bu gruplarda ölçekle ilgili farklı analizler de yapılabilmektedir (DeVellis, 2017). Whortington ve Whittaker (2006) madde seçim sürecine göre ölçek uzunluğunun belirlenmesini önermektedirler. Bu çalışmada testte yer alan 8 maddenin azaltılmasına karar verilmiştir. Bu nedenle çalışma kapsamında öncelikle açımlayıcı faktör analizi ile madde seçimi yapılmış sonrasında ise ölçeğin kuramsal yapısı doğrulayıcı faktör analizi ile incelenmiştir. Bu amaç doğrultusunda veri seti ikiye bölünmüştür. Her iki grubun da benzer özelliklere sahip olması için öncelikle veri setindeki katılımcılar sınıf düzeylerine göre üç gruba ayrılmıştır. Her bir gruptan ilk 10 öğrenci birinci veri setinde kalacak, sonrasındaki 10 öğrenci ise ikinci veri setinde olacak şekilde veriler ikiye ayrılmıştır. Böylece ilk veri setinde 401 öğrenci, ikinci veri setinde ise 400 öğrenci olmak üzere iki veri seti oluşturulmuştur. İlk veri seti ile açımlayıcı faktör analizi, ikinci veri seti ile de doğrulayıcı faktör analizi yapılmıştır. Yapı geçerliğinin incelenmesi kapsamında son olarak doğrulayıcı faktör analizi bulguları göz önünde bulundurularak ölçeğin yakınsak ve ıraksak geçerliği incelenmiştir.

Açımlayıcı faktör analizi

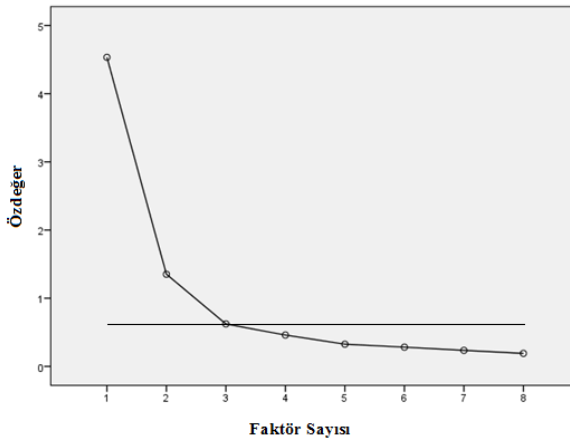
ÇABİYAT'ın madde seçim sürecinde ve faktör sayısını ortaya koymada faktörleştirme tekniklerinden maksimum olasılık faktör analizi yapılmıştır. Açımlayıcı faktör analizi yapmadan önce kayıp değer, uç değerler, örneklem büyüklüğü, normallik, doğrusallık, çoklu doğrusal bağlantı ve tekillik sayıtlarının (Çokluk, Şekercioğlu ve Büyüköztürk, 2010) karşılanıp karşılanmadığı kontrol edilmiştir.

İlk olarak betimsel analizler yapılmış ve sonucunda veri setinde kayıp değer olmadığı bulunmuştur. Veri setindeki çok yönlü uç değerleri tespit etmek için Mahalanobis uzaklığı, Cook uzaklık değeri, merkezi manivela gücü değerleri ve z puanları incelenmiştir. Tabachnick ve Fidell (2007) $p < .001$ güven aralığında 8 değişken için Mahalanobis kritik değer 26.125 olduğunu ve z puanlarının -3.3 ile +3.3 değerleri arasında olması gerektiğini belirtmektedir. Akbulut (2010) Cook uzaklık değerlerinin 1'den yüksek olmaması, merkezi manivela gücü değerlerinin çoğunluğunun .02'nin altında olması ve hiçbir değer .05'in üstünde olmaması gerektiğini ifade etmektedir. Bu değerleri karşılamayan 20 öğrencinin puanları analize dâhil edilmemiş örneklem büyüklüğü 381'e düşmüştür. Field (2009) faktör analizi için en az 300 katılımcının gerekli olduğunu ve bu büyüklüğe ulaşıldıktan sonra ideal değerlere ulaşılabileceğini belirtmektedir. Dolayısıyla örneklem büyüklüğünün faktör analizi için yeterli olduğu söylenebilir. Diğer taraftan örneklem büyüklüğünün faktörleştirmeye uygunluğunu test etmek için KMO değeri incelenmiştir. Çalışma kapsamında KMO değeri .857 olarak hesaplanmıştır. Bu bulgu doğrultusunda örneklem büyüklüğünün faktör analizi yapmak için "çok iyi derecede yeterli" olduğu sonucuna ulaşılmıştır (Hutcheson ve Sofroniou, 1999). Veri setinin çok değişkenli normallik sayıtlarını test etmek amacıyla Bartlett Küresellik Testi yapılmıştır. Analiz sonucunda bu değer $(\chi^2(28)=1856.794; p < .001)$ anlamlı bulunmuştur. Bu değer anlamlı olması hem veri setinin çoklu normallik ölçütünü karşıladığı hem de korelasyon matrisinin faktör çıkarmaya uygun olduğu anlamına gelmektedir. Bu değer örneklem büyüklüğünden etkilenmektedir. Bu nedenle tek değişkenli normalliğin de test edilmesi önerilmektedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Tek değişkenli normallik sayıtlarını test etmek amacıyla çarpıklık ve basıklık katsayıları incelenmiştir. Çarpıklık değerlerinin 0.231 ile 1.129 arasında; basıklık değerlerinin ise -0.810 ile 0.482 arasında değiştiği görülmüştür. Trochim ve Donnelly (2006) verilerin normal dağılım göstermesi için basıklık ve çarpıklık değerlerinin -2 ile +2 arasında olması durumunun kabul edilebilir olduğunu belirtmektedir. Diğer taraftan Fabrigar ve arkadaşları (1999) maksimum olasılık faktör analizi tekniği kullanıldığında çarpıklık değerinin 2'den basıklık değerinin ise 7'den büyük olmaması gerektiği belirtmektedir. Bu nedenle veri setinin tek değişkenli normallik sayıtlarını karşıladığı söylenebilir. Doğrusallık

sayılısı için saçılma diyagramı oluşturulmuş ve değişken çiftlerinin oluşturduğu matraste yer alan şekillerin elipse yakın olduğu görülmüştür. Bu nedenle doğrusallık sayılısının karşılandığı belirtilebilir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Çoklu doğrusal bağlantı ve teklik sayılısını test etmek için maddeler arası korelasyon katsayıları, tolerans değerleri ve VIF değerleri incelenmiştir. Tolerans değerlerinin .302 ile .564 arasında, VIF değerlerinin ise 1.774 ile 3.309 arasında, madde korelasyon katsayılarının ise .236 ile .794 arasında olduğu bulunmuştur. Tolerans değerlerinin .10'dan büyük ve VIF değerlerinin 10'dan küçük olması (Akbulut, 2010), korelasyon katsayılarının .90'dan küçük olması (Çokluk, Şekercioğlu ve Büyüköztürk, 2010) veri setinde çoklu doğrusal bağlantı ve teklik problemi olmadığını göstermektedir. Sonuç olarak veri setinin açılımlayıcı faktör analizi için uygun olduğu görülmüştür.

Faktör analizi kapsamında ilk olarak maddelerin ortak varyans değerleri (communalities) incelenmiştir. Maddelerin ortak varyans değerlerinin .50 ile .79 arasında değiştiği ve ortalamalarının ise .65 olduğu bulunmuştur. Maddelerin ortak varyans değerlerinin .30'dan büyük olması (Field, 2009), ölçek maddelerinin hedeflenen yapıyı ölçmeye hizmet ettiğini göstermektedir. Faktör sayısının belirlenmesinde Kaiser ölçütü, Catell'in yamaç-birikinti grafiği ve paralel analizden yararlanılmıştır. Kaiser ölçütüne göre özdeğeri 1 ve üzerinde olan faktörlerin sayısı temel alınmaktadır (Akbulut, 2010). Yapılan açılımlayıcı faktör analizi sonucunda, sekiz madde için öz değeri 1'in üzerinde olan iki bileşen olduğu görülmüştür. Bu bileşenlerin toplam varyansa yaptıkları katkı %65.387'dir. Söz konusu bileşenlerin varyansa yaptıkları katkı sırasıyla %51.959 ve %13.429'dur.

Faktör sayısını belirlemede kullanılan bir diğer yöntem Catell'in yamaç-birikinti grafiğidir. Analiz sonucunda elde edilen yamaç-birikinti grafiğine Şekil 4.1'de yer verilmiştir.



Şekil 4.1. Yamaç-Birikinti Grafiği

Şekil 4.1’de görüldüğü üzere grafikte üçüncü noktada öz değerlerin monoton dağılımı bozulmaktadır. Grafikte üçüncü noktaya kadar eğimde kırılmaların olduğu ancak bu noktadan sonra dağılımın kısmen düz bir şekilde olduğu görülmektedir. Bu durum ÇABİYAT’ın iki faktörlü bir yapıya sahip olduğunu ortaya koymaktadır.

Faktör sayısını belirlemede son olarak paralel analiz yönteminden yararlanılmıştır. Bileşen sayısının belirlenmesinde, bu yöntemin Kaiser ölçütü ve Catell’in yamaç-birikinti grafiğine göre daha güvenilir olduğu kabul edilmektedir (Zwick ve Velicer, 1986). Paralel analiz yönteminde çalışma kapsamında elde edilen öz değerler ile çalışmanın örneklem büyüklüğü ve değişken sayısı temel alınarak rassal olarak üretilen öz değerler karşılaştırılmaktadır (Pallant, 2005). Çalışma kapsamında paralel analiz için Monte Carlo PCA paralel analiz programı kullanılmıştır (Watkins, 2000). Bu program ile 381 katılımcı ve sekiz madde için rassal olarak 1000 öz değer üretilmiştir. Maksimum olasılık yönteminden elde edilen başlangıç özdeğerleri ile paralel analizden üretilmiş özdeğerlerinin karşılaştırılmasına Tablo 4.2’de yer verilmiştir. Karşılaştırmaların sonucunda, maksimum olasılık yönteminden elde edilen iki özdeğerin, paralel analizden üretilen özdeğerlerden büyük olduğu görülmüştür. Bu durum iki faktörlü bir yapıya işaret etmiştir.

Tablo 4.2. Paralel Analiz Sonuçları

Bileşen No	Maksimum Olasılık Yönteminden Elde Edilen Birincil Özdeğerler	Paralel Analizden Elde Edilen Özdeğerler	Karar
1	4.533	1.2211	Kabul
2	1.351	1.1383	Kabul
3	0.623	1.0761	Ret
4	0.460	1.0219	Ret
5	0.325	0.9697	Ret
6	0.282	0.9165	Ret
7	0.235	0.8623	Ret
8	0.191	0.7941	Ret

ÇABİYAT’ın kuramsal çerçevesine göre, deney tasarlama ve hipotez geliştirme olmak üzere iki faktörlü bir yapıdan oluşmaktadır. Maksimum olasılık yöntemi kapsamında faktör sayısını belirlemede kullanılan üç yöntem de iki faktörlü bir yapıyı desteklemektedir. Faktör sayısı belirlendikten sonra açıklayıcı faktör analizi bulgularını daha kolay yorumlayabilmek için döndürme işlemi yapılmıştır. Analizde eğik döndürme yöntemlerinden direct oblimin yöntemi seçilmiştir. Eğik döndürme yöntemleri, faktörlerin birbirleriyle ilişkili olduğu durumlarda kullanılmaktadır (Tabachnick ve Fidell, 2007). Kuramsal olarak hipotez geliştirme ve deney tasarlama becerilerinin ilişkili olduğunun kabul edilmesi nedeniyle eğik döndürme yöntemi tercih

edilmiştir. Direct oblimin yöntemi seçilerek iki faktörlü bir yapı için analiz tekrarlanmıştır. Analizin sonucunda faktörlerin varyansa yaptığı katkı birinci faktör için %51.959 ikinci faktör için ise %13.429 olduğu bulunmuştur. Faktörlerin varyansa yaptıkları ortak katkı %65.387'dir.

Faktör desenini ortaya koymak amacıyla gerçekleştirilen faktör analizinde döndürme sonucunda yapı ve model matrisleri elde edilmektedir. Model matrisi faktör ile değişkenler arasındaki ilişkilere göre, yapı matrisi ise maddeler arasındaki korelasyona göre oluşturulmaktadır. Genellikle yapı matrisinde yer alan faktör yük değerlerinin yorumlanması önerilmektedir (Tabachnick ve Fidell, 2007). Bu doğrultuda analizlerde yapı matrisinde yer alan faktör yük değerleri göz önünde bulundurulmuştur. Maddelerin ölçekte kalıp kalmamasına yönelik değerlendirmelerde faktör yük değeri için kabul değeri .30 olarak belirlenmiştir (Şencan, 2005). Faktör yükleri arasında .10'un altında fark olan maddeler binişik madde olarak kabul edilmiştir (Akbulut, 2010). İki faktör için yapılan analizde maddeler; binişiklik, faktör yük değerlerinin kabul düzeyini karşılayıp karşılamama, açıklanan varyans, betimsel bulgular ve kuramsal çerçeveye uygunluk açısından değerlendirilmiştir.

Analiz sonucunda teorik olarak tanımlanan HG4 maddesi dışındaki tüm maddelerin kendi faktörleri altında toplandığı görülmüştür. Hipotez geliştirme becerisi olarak tasarlanan bu maddenin deney tasarlama faktörü için yük değerinin ($\lambda_{DT}=.666$) hipotez geliştirme faktörü için yük değerinden ($\lambda_{HG}=.623$) daha büyük olduğu bulunmuştur. Ancak faktör yükleri arasındaki farkın .023 olması, maddenin binişik bir madde olduğunu göstermektedir. Bu maddenin ortak faktör varyansının en düşük değere sahip olması nedeniyle ölçüğe en az hizmet eden madde oluşu düşünülebilir. Deney tasarlama becerisi olarak tasarlanan DT4 maddesinin de iki farklı faktörde yakın yük değerleri ($\lambda_{DT}=.631$; $\lambda_{HG}=.583$) taşıması nedeniyle binişik olduğu görülmüştür. HG4 maddesi için üretilen maksimum doğru yanıt sayısının 9, madde ortalamasının ise 2.11 (ss= 1.63) olduğu bulunmuştur. DT4 maddesi için üretilen maksimum doğru yanıt sayısı 10, maddenin ortalaması ise 2.41'dir (ss=1.8). Bu iki madde dışındaki diğer maddeler için üretilen maksimum yanıt sayısı 11 ile 14 arasında, maddelerin ortalamalarının ise 4.35 ile 6.08 arasında değiştiği görülmektedir. Bu nedenle HG4 ve DT4 maddelerinin diğer maddelere göre daha zor olduğu ve açık uçluluk bağlamında daha sınırlı olduğu ifade edilebilir. Bu maddelerin ölçekten çıkarılmasıyla iki faktörün açıkladığı toplam varyansın %65.387'den %70.677'ye yükseldiği görülmüştür. Maddelerin analiz dışı bırakılması durumunda KMO ve Cronbach Alfa değerlerinde oldukça küçük de olsa bir düşüş olmuştur. Akbulut (2010) binişik maddelerin mutlaka analiz dışında tutulmaları gerektiğini belirtmektedir. İstatistiksel bulgular ve kuramsal çerçeve incelendiğinde HG4 ve DT4 maddelerinin zayıf olduğuna veya teorik

modele uymadığına karar verilmiş ve bu maddeler ölçekten çıkarılmıştır. Bu maddelerin analiz dışı bırakılması sonucunda ulaşılan faktör deseni ve bu desende yer alan maddelerin betimsel değerlerine Tablo 4.3'te yer verilmiştir.

Tablo 4.3. ÇABİYAT'ın Faktör ve Maddelere İlişkin Betimsel Değerleri

Faktörler ve maddeler	Açıklanan varyans (%)	\bar{X}	SS	Doğrulanmış madde toplam r	Faktör yükü
Deney Tasarlama ($\alpha = .85$)					
DT1: Hamsterlar	52.95	6.31	2.34	.72	.89
DT2: Kum Havuzu		5.77	2.36	.69	.90
DT3: Oyuncak Tünel		5.68	2.04	.72	.83
Hipotez Geliştirme ($\alpha = .84$)					
HG1: Bataklığıtaki Sinekler	17.73	4.57	1.65	.60	.78
HG2: Suluklar		3.90	1.68	.61	.92
HG3: Oyuncak Gemi		4.00	1.73	.48	.70
Toplam ($\alpha = .90$)	70.68				

DT: Deney Tasarlama. HG: Hipotez Geliştirme.

Tablo 4.3'te görüldüğü üzere alt ölçeklerler düzeyinde deney tasarlama alt ölçeği için faktör yük değerleri deney tasarlama alt ölçeği için .83 ile .90 arasında, hipotez geliştirme alt ölçeği için ise .70 ile .92 arasında değişmektedir. Faktör yük değerleri büyüklük açısından incelendiğinde maddelerin tamamının mükemmel seviyede olduğu görülmektedir (Comrey ve Lee, 1992).

İki faktör için yinelenen ve analiz dışı bırakılan maddelerin ardından faktörlerin toplam varyansa yaptıkları katkının deney tasarlama faktörü için %52.95, hipotez geliştirme için ise %17.73 olduğu görülmüştür. Her iki faktörün toplam varyansa katkılarının önem derecesinin birbirine yakın olduğu ifade edilebilir. Bu iki faktörün varyansa yaptıkları toplam katkı ise %70.68 olarak bulunmuştur. Sosyal bilimler alanında çok faktörlü desenlerde açıklanan varyansın %40 ile %60 arasında olması yeterli olarak kabul edilmektedir (Dunteman, 1989). Çalışma kapsamında iki faktörün toplam varyansa yaptığı katkının oldukça iyi seviyede olduğu görülmektedir.

ÇABİYAT'ın iç tutarlık düzeyi ise .90'dır. Cronbach Alfa güvenilirlik değeri deney tasarlama alt ölçeği için .85, hipotez geliştirme alt ölçeği için ise .84 olarak bulunmuştur. Özdamar (2004, s. 632-633) .80 ile 1 arasında Cronbach Alfa güvenilirlik değerine sahip ölçeklerin yüksek derece güvenilir olduğunu belirtmiştir. ÇABİYAT'ın ve alt ölçeklerinin yüksek derecede güvenilir olduğu görülmektedir. Bu değer madde sayısından oldukça etkilenmektedir.

Bu nedenle madde sayısı 10'dan küçük olan ölçekler için maddeler arası korelasyon değerlerinin incelenmesi önerilmektedir (Pallant, 2005). Clark ve Watson (1995) maddeler arası korelasyon katsayılarının ortalama değerlerinin .15 ile .50 arasında olması gerektiğini belirtmiştir. ÇABİYAT'ın maddeler arası korelasyon katsayılarının ortalaması ($r_{\min}=.24$, $r_{\max}=.79$, $r_{\text{ort}}=.49$) ölçeğin güvenilir bir ölçek olduğunu desteklemektedir. Ölçeğin tüm maddelerinin doğrulanmış madde toplam korelasyonlarının .30'un üzerinde olması ($r_{\min}=.48$, $r_{\max}=.72$) maddelerin benzer beceriyi ölçtüğü ve bireyleri iyi derecede ayırt ettiğini göstermektedir (Büyüköztürk, 2014).

Sonuç olarak, açımlayıcı faktör analizi bulgularının ÇABİYAT'ın iki faktörlü kuramsal yapısını desteklediği görülmektedir. Maddelerin deney tasarlama ve hipotez geliştirme faktörleri altında toplanması, ölçeğin yapısının deneysel bir kanıtı olarak düşünülebilir. Analiz sonucunda deney tasarlama alt ölçeğinde üç madde, hipotez geliştirme alt ölçeğinde üç madde olmak üzere toplam altı maddeden oluşmaktadır.

Doğrulayıcı faktör analizi

ÇABİYAT'ın 2 faktörlü ve 6 maddeli yapısını incelemek için doğrulayıcı faktör analizi gerçekleştirilmiştir. Doğrulayıcı faktör analizi 400 katılımcı ile AMOS 22 programı kullanılarak yapılmıştır. Doğrulayıcı faktör analizini gerçekleştirmeden önce veri setinin analiz için uygunluğu test edilmiştir. Bu amaçla ilk olarak kayıp değerler kontrol edilmiş ve kayıp değer olmadığı görülmüştür. İkinci olarak veri setindeki çok yönlü uç değerleri belirlemek için Mahalanobis uzaklığı, Cook uzaklık değeri, merkezi manivela gücü değerleri ve maddelerin "z" puanları incelenmiştir. Tabachnick ve Fidell (2007) $p<.001$ güven aralığında 6 değişken için Mahalanobis kritik değerinin 22.46 olduğunu ve "z" puanlarının -3.3 ile +3.3 değerleri arasında olması gerektiğini belirtmektedir. Cook uzaklık değerlerinin 1'den yüksek olmaması merkezi manivela gücü değerlerinin çoğunluğunun .02'nin altında olması ve hiçbir değer .05'in üstünde olmaması gerekmektedir (Akbulut, 2010). Çalışma kapsamında bu değerler göz önünde bulundurulduğunda 19 katılımcı uç değer olarak belirlenmiş ve veri setinden çıkarılmıştır. Uç değerler çıkarıldıktan sonra örneklem büyüklüğü 381'e düşmüştür. Faktör analizinin sayılıtlarından birisi olan normallığın incelenmesi için dağılım grafikleri incelenmiştir. Bu grafikler normal dağılıma yakın görüntüler vermiştir. Ayrıca her bir maddenin basıklık ve çarpıklık değerleri incelenmiştir. Bu değerlerin -1 ile +1 arasında olmasından dolayı da normallik sayılıtlarının karşılandığı sonucuna ulaşılmıştır (Huck, 2012). Doğrusallık sayılıtları için saçılma diyagramı incelenmiş ve değişken çiftlerinin oluşturduğu matriste yer alan şekillerin elipse yakın olduğu görülmüştür. Bu durum doğrusallık sayılıtlarının karşılandığına dair ipucu vermiştir. Son olarak çoklu doğrusal bağlantı ve tekillik sayılıtlarını test etmek için tolerans

değerleri, VIF değerleri ve maddeler arası korelasyon katsayıları incelenmiştir. Tolerans değerlerinin 0.425 ile 0.524 arasında, VIF değerlerinin ise 1.908 ile 2.351 arasında olduğu bulunmuştur. Tolerans değerlerinin .10'dan büyük ve VIF değerlerinin 10'dan küçük olması çoklu doğrusal bağıntı bağlamında sorun olmadığını göstermektedir (Akbulut, 2010). Maddeler arası korelasyon katsayılarının .359 ile .668 arasında değiştiği ve bu değerlerin .90'nın altında olmasından dolayı (Çokluk, Şekercioğlu ve Büyüköztürk, 2010) veri setinde çoklu doğrusal bağlantı ve teklik problemi olmadığı görülmüştür. Sonuç olarak veri setinin doğrulayıcı faktör analizi için uygun olduğu saptanmıştır.

ÇABİYAT'ın teorik altyapısına göre bilimsel yaratıcılık hipotez geliştirme ve deney tasarlama olmak üzere iki faktörden oluşmaktadır. Her bir faktörde üç madde yer almaktadır. Bu yapıyı doğrulamak için ikinci düzey faktör analizinin yapılması gerekmektedir. Kline (2011) ikinci düzey faktör analizinin yapılabilmesi için birinci düzeyde en az üç faktör olması gerektiğini belirtmektedir. Ancak alanyazında birinci düzeyde iki faktör olan ikinci düzey faktör analizinin yapıldığı çalışmalara rastlanmaktadır (ör. Milfont ve Duckitt, 2004; Sak, vd., 2019). ÇABİYAT'ın faktör yapısını incelemek için ikinci veri setinde yer alan 381 katılımcı ile AMOS 22 programı kullanılarak doğrulayıcı faktör analizi gerçekleştirilmiştir. Doğrulayıcı faktör analizi süreci ölçek modelinin sınanması ve değerlendirilmesi olmak üzere iki aşamada gerçekleştirilmiştir. Ölçek modelinin sınanması kapsamında birinci düzey, ölçek modelinin değerlendirilmesi kapsamında ikinci düzey faktör analizi yapılmıştır.

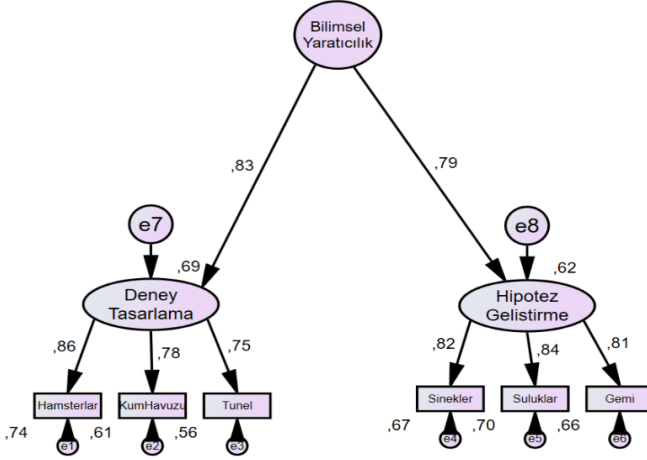
Ölçek modelinin sınanması kapsamında iki farklı model sınanmıştır. İlk ölçüm modeli altı maddenin tek bir yapıyı ölçtüğüne yönelik bir model, ikinci model ise altı maddenin iki farklı (deney tasarlama ve hipotez geliştirme) yapıyı ölçtüğüne yönelik bir modeldir. Bu ölçme modellerinin birinci düzey doğrulayıcı faktör analiz sonucunda elde edilen uyum iyiliği değerlerine Tablo 4.4'te yer verilmiştir.

Tablo 4.4. ÇABİYAT Verilerinin Modellerinin Uyum İyiliği Değerleri

	Uyum İyiliği Endeksleri (Kabul Edilebilir Uyum için Kesme Noktaları)					
	χ^2 /sd (≤ 3)	RMSEA (≤ 0.08)	RMR (≤ 0.05)	NFI ($0.90 \leq$)	CFI ($0.90 \leq$)	GFI ($0.90 \leq$)
Model I: Tek Faktörlü Model	24.463	0.248	0.273	0.807	0.813	0.799
Model II: İki Faktörlü Model	1.067	0.013	0.045	0.993	1	0.992

Tablo 4.4'te görüldüğü üzere öncelikle altı maddenin tek bir faktörü ölçtüğüne yönelik bir model sınanmış ve uyum iyiliği değerlerinin kabul edilebilir aralıkta olmadığı sonucuna ulaşılmıştır. Bu bulgu ölçeğin birden fazla alt boyutu olduğunun bir kanıtı olarak değerlendirilebilir. Daha sonra altı maddenin iki farklı yapıyı (deney tasarlama ve hipotez geliştirme) ölçtüğüne yönelik model test edilmiş ve elde edilen sonuçlara göre modelin uyum iyiliği değerlerinin oldukça yüksek olduğu sonucuna ulaşılmıştır. Bu bulgular deney tasarlama ve hipotez geliştirme alt ölçeğinin birlikte bilimsel yaratıcılığı ölçtüğüne yönelik kanıt oluşturmaktadır. Bu nedenle ÇABİYAT'ın iki faktörlü ve altı maddeli yapısını incelemek için ikinci düzey (second order) faktör analizi gerçekleştirilmiştir. Bu aşamada ölçek modelinin değerlendirilmesi yapılmıştır.

Ölçek modelini değerlendirme aşamasında standardize edilmiş katsayıları dikkate alınarak ÇABİYAT'ın teorik alt yapısındaki ilişkiler Şekil 4.2'de görselleştirilmiştir. Sonrasında beklenen kovaryans matrisi ile gözlenen kovaryans matrisleri arasındaki farkın (χ^2 değerinin) anlamlılığı hakkında bilgi veren p değeri incelenmiş ve anlamlı olmadığı bulunmuştur ($\chi^2_{(8)} = 8.536, p = .383$). Doğrulayıcı faktör analizinde bu değer anlamlı olmaması beklenmektedir (Tabachnick ve Fidell, 2007). Son olarak ölçek modelinin uyum iyiliği değerleri incelenmiştir. ÇABİYAT'ın kuramsal modelinin, uyum iyiliği değerleri göz önünde bulundurulduğunda mükemmel bir uyuma sahip olduğu kabul edilebilir ($\chi^2 /sd = 1.067$; RMSEA = .013; RMR = .045; SRMR = .016; NFI = .993; CFI = 1; GFI = .992; AGFI = .980). RMSEA değerinin mükemmel uyum gösterdiği durumlarda Yakın Uyum P değerinin (P Close Fit, PCLOSE) yorumlanması önerilmektedir. Bunun nedeni RMSEA değerinin örneklem büyüklüğünden etkilenmesidir. PCLOSE değeri ile RMSEA değerinin anlamlılığı test edilir (Browne ve Cudeck, 1993). İkinci düzey doğrulayıcı faktör analizinin sonunda PCLOSE değeri .861 olarak hesaplanmıştır. Hu ve Bentler (1999) bu değer .08'in üzerinde olması gerektiğini belirtmektedir. RMSEA değerinin mükemmel uyum göstermesinin anlamlı olduğu ifade edilebilir.



Şekil 4.2. ÇABİYAT'ın Teorik Altyapı Modeli

Sonuç olarak doğrulayıcı faktör analizinden elde edilen ölçek modeline ilişkin uyum endeksleri ve temel parametre tahminleri modelin verilerle uyum içerisinde olduğunu göstermektedir. Bu durum iki düzeyli ÇABİYAT'ın yapısının doğrulandığını ortaya koymaktadır.

Yakınsak ve iraksak geçerliği

Yakınsak geçerlik, yapı geçerliliğinin test edilmesinde kullanılan yaklaşımlardan biridir (Anastasi ve Urbina, 1997). Campbell ve Fiske (1959) yakınsak geçerliği, ölçek maddelerinin ölçülmek istenen yapıyı doğru bir şekilde ölçme seviyesi olarak tanımlamaktadırlar. Iraksak geçerliği ise ölçek modelinde yer alan bir faktörün ilişkisiz bir faktörden farklılaşma seviyesi olarak betimlemektedirler. Araştırmacılar bu geçerlik türlerinin test edilmesi için Çoklu Özellik Çoklu Yöntem (Multi trait-multi method) kullanımını önermişlerdir. Fornell ve Larcker (1981) yakınsak ve iraksak geçerliğin test edilmesinde bu yaklaşıma benzer fakat daha kolay uygulanan bir yöntem önermişlerdir. ÇABİYAT'ın yakınsak ve iraksak geçerliliğinin test edilmesinde Fornell ve Larcker'in önerdiği yöntem kullanılmıştır. Bu yöntemde göre her bir faktör yapısı için ortalama açıklanan varyans (Average Variance Extracted, AVE), bileşik güvenilirlik (Composite Reliability, CR) ve faktörler arası korelasyon katsayısı değerleri hesaplanmıştır. Adı geçen değerlerin hesaplanmasında doğrulayıcı faktör analizinin bulgularından yararlanılmıştır. AVE değeri, bir faktör altında yer alan maddelerin faktör yüklerinin karelerinin toplamının madde sayısına bölünmesi ile elde edilmiştir. Cronbach Alfa değerini kontrol etmek amacıyla kullanılan CR değeri ise bir faktör altında yer alan maddelerin faktör yüklerinin toplamının, bu değer ile hata varyanslarının toplamına bölümü ile hesaplanmaktadır. Faktörler arası korelasyon katsayısı ise Pearson momentler çarpımı korelasyon analizi ile elde edilmiştir.

ÇABİYAT'ın yakınsak geçerliği için AVE ve CR değerleri doğrulayıcı faktör analizinden elde edilen faktör yükleri göz önünde bulundurularak hesaplanmıştır. ÇABİYAT maddelerinin faktör yüklerine, faktörlerin AVE ve CR değerlerine Tablo 4.5'te yer verilmiştir.

Tablo 4.5. ÇABİYAT Faktör Yükleri ve Faktörlerin AVE ve CR Değerleri

Faktörler (Bileşenler)	Madde	λ	AVE	CR
Deney Tasarlama	DT1: Hamsterlar	0.858	0.635	0.839
	DT2: Kum Havuzu	0.780		
	DT3: Oyuncak Tünel	0.749		
Hipotez Geliştirme	HG1: Bataklıkta Sinekler	0.818	0.676	0.862
	HG2: Suluklar	0.837		
	HG3: Oyuncak Gemi	0.812		

Tablo 4.5 incelendiğinde ölçek maddelerinin faktör yüklerinin .749 ile .858 arasında değiştiği görülmektedir. Bu değerlerin .70'in üzerinde olması test maddelerinin mükemmel seviyede ölçğe hizmet ettiğini göstermektedir (Comrey ve Lee, 1992). Deney tasarlama ve hipotez geliştirme alt ölçekleri için AVE değerleri sırasıyla .635, .672, CR değerleri ise sırasıyla .839, .862 olarak hesaplanmıştır. Fornell ve Larcker (1981) Yakınsak geçerliğin sağlanabilmesi için AVE ve CR değerlerin sırasıyla .50'nin ve .70'in üzerinde olması gerektiğini belirtmektedirler. Bu bağlamda her bir alt ölçekte yer alan maddelerin ölçülmek istenen yapıyı temsil ettiğine yönelik kanıt oluşturduğu ifade edilebilir.

İraksak geçerliğin incelenmesinde her bir faktör için AVE değerinin karekökünün hesaplanması ve bu değerler arası korelasyon değerleri ile karşılaştırılması gerekmektedir. ÇABİYAT'ın iraksak geçerliğine yönelik yapılan analiz bulgularına Tablo 4.6'da yer verilmiştir.

Tablo 4.6. ÇABİYAT'ın İraksak Geçerliğine Yönelik Bulgular

N = 381	\bar{X}	ss	Deney Tasarlama	Hipotez Geliştirme	\sqrt{AVE}
Deney Tasarlama	16.136	4.442		.561**	0.797
Hipotez Geliştirme	13.843	4.359	.561**		0.822

**p<.01

Tablo 4.6 incelendiğinde deney tasarlama ve hipotez geliştirme alt ölçekleri arasındaki ilişki pozitif, anlamlı ve yüksek olarak bulunmuştur ($r = .561$, $p < .01$). Deney tasarlama ve hipotez geliştirme alt ölçekleri için hesaplanan AVE değerinin karekökünün faktörler arası korelasyon katsayısından büyük olması, deney tasarlama ve hipotez geliştirme yapılarının birbirleri ile ilişkili ancak birbirlerinden farklı yapılar olduğunu ortaya koymaktadır (Fornell ve Larcker, 1981). Ölçeğin yakınsak ve iraksak geçerliğine ilişkin yapılan analizlerin sonuçları, ÇABİYAT'ın yapı geçerliliğinin iyi düzeyde olduğunu göstermektedir.

Ölçüt Geçerliği

Ölçek geliştirme sürecinde, geçerliği incelemede yalnızca ölçeğin kendi içindeki geçerliği aramak yeterli değildir. Geçerliğin incelenmesinde yaş, yetenek, başarı, eğitim düzeyi gibi dış ölçütler de göz önünde bulundurulmalıdır (Anastasi ve Urbina, 1997; DeVellis, 2017). ÇABİYAT'ın ölçüt geçerliğinin incelenmesi kapsamında eğitim düzeyi ve başarı değişkenleri göz önünde bulundurulmuştur.

Gelişimsel geçerlik

ÇABİYAT'ın sınıf düzeylerine göre ayırt ediciliğini sınamak için ana uygulamada toplanan tüm veriler üzerinden bağımsız gruplar için tek faktörlü ANOVA gerçekleştirilmiştir. Bağımsız gruplar için tek faktörlü ANOVA gerçekleştirilmeden önce testin normal dağılım ve varyansların eşleşliği sayıltıları incelenmiştir. Farklı sınıf düzeylerinde yer alan katılımcılardan elde edilen puanların çarpıklık değerlerinin -0.741 ile 1.541 aralığında, basıklık değerlerinin ise -0.389 ile 1.408 aralığında değiştiği görülmüştür. Bu değerlerin -2 ile +2 aralığında olması verilerin normal dağıldığı sonucuna ulaşılmıştır (Trochim ve Donnelly, 2006). Diğer taraftan varyansların eşleşliğini test etmek amacıyla Levene testi yapılmıştır. Levene testi sonuçlarına göre deney tasarlama alt ölçeğinden elde edilen akıcılık ve orijinallik puanlarında ve hipotez geliştirme alt ölçeğinden elde edilen akıcılık puanlarında grupların varyanslarının eşleştiği ($p > .05$), diğer puanlarda ise grupların varyanslarının eşleşmediği ($p < .05$) bulunmuştur. Varyansların eşleştiği durumuna göre de izleme testleri yapılmıştır (Huck, 2012). Varyansların eşleştiği durumlarda Scheffe izleme testi, olmadığı durumlarda ise Tamhane izleme testi yapılmıştır.

Gelişimsel geçerlik kapsamında katılımcıların yaşları ile ölçek puanları arasındaki ilişki incelenmektedir (Anastasi ve Urbina, 1997). Araştırma kapsamında katılımcıların yaşları ile sınıf düzeyleri arasında yüksek ve anlamlı bir ilişki bulunmuştur ($r = .968, p < .001$). Bu nedenle ölçeğin gelişimsel geçerliğini incelemek için farklı sınıf düzeyindeki öğrencilerin ÇABİYAT puanları karşılaştırılmıştır. ÇABİYAT'ın deney tasarlama alt ölçeği, hipotez geliştirme alt ölçeği ve toplamından akıcılık, esneklik, orijinallik, yaratıcılık bölümü puanları elde edilmektedir. Sınıf düzeylerine göre ÇABİYAT puanlarının betimsel istatistiklerine Tablo 4.7'de yer verilmiştir.

Tablo 4.7 incelendiğinde alt ölçeklerden ve toplam testten elde edilen tüm puan ortalamalarının sınıf düzeyi arttıkça arttığı dikkat çekmektedir. Söz konusu farklılığın anlamlılığını test etmek amacıyla 801 katılımcı ile bağımsız gruplar için tek faktörlü ANOVA yapılmıştır. Tablo 4.8'de analiz sonuçları yer almaktadır.

Tablo 4.7. ÇABİYAT Puanlarının Sınıf Düzeylerine Göre Betimsel İstatistikleri

Bileşen	Puan Türü	Sınıf	N	En Düşük	En Yüksek	\bar{X}	SS
Deney Tasarlama	Akıcılık	Anasınıfı	221	1	30	15.10	5.33
		1	278	6	31	17.04	5.21
		2	302	6	34	18.95	5.57
	Esneklik	Anasınıfı	221	1	13	7.30	1.86
		1	278	4	13	8.78	1.29
		2	302	3	13	10.26	2.33
	Orijinallik	Anasınıfı	221	0	2127	787.93	392.61
		1	278	135	2177	918.19	390.65
		2	302	204	2195	1078.93	425.93
	Yaratıcılık Bölümü	Anasınıfı	221	1.00	17.74	10.99	3.05
		1	278	5.58	19.34	12.86	2.68
		2	302	5.58	21.74	14.54	3.51
Hipotez Geliştirme	Akıcılık	Anasınıfı	221	2	22	10.82	4.32
		1	278	4	35	13.59	4.14
		2	302	6	30	15.17	4.49
	Esneklik	Anasınıfı	221	2	10	5.21	2.01
		1	278	3	10	7.26	1.40
		2	302	3	12	8.46	2.40
	Orijinallik	Anasınıfı	221	38	1709	618.76	318.94
		1	278	159	2502	790.30	324.88
		2	302	232	2058	908.10	343.19
	Yaratıcılık Bölümü	Anasınıfı	221	2.00	14.74	7.96	2.78
		1	278	3.58	16.28	10.33	2.17
		2	302	4.90	19.80	11.76	3.02
Toplam Test	Akıcılık	Anasınıfı	221	4	49	25.91	8.67
		1	278	11	60	30.63	7.32
		2	302	18	56	34.13	8.32
	Esneklik	Anasınıfı	221	3	22	12.51	3.21
		1	278	9	22	16.04	2.18
		2	302	6	25	18.72	3.80
	Orijinallik	Anasınıfı	221	48	3175	1406.69	614.04
		1	278	355	3983	1707.91	548.92
		2	302	668	3618	1987.03	625.99
	Yaratıcılık Bölümü	Anasınıfı	221	4	31.38	18.95	5.31
		1	278	10.16	35.34	23.18	3.99
		2	302	12.48	41.44	26.30	5.51

Tablo 4.8. ÇABİYAT Puanlarının Sınıf Düzeylerine Göre ANOVA Sonuçları

	Puan Türü	Varyansın Kaynağı	Kareler Toplamı	sd	Kareler Ortalaması	F	p	η^2
Deneysel Tasarlama	Akıcılık	G.arası	1914.63	2	957.32	33.04	.000	.076
		G.içi	23120.92	798	28.97			
		Toplam	25035.55	800				
	Esneklik	G.arası	1129.85	2	564.92	158.05	.000	.284
		G.içi	2852.24	798	3.57			
		Toplam	3982.08	800				
	Orijinallik	G.arası	11064864.06	2	5532432.03	33.75	.000	.078
		G.içi	130637702.64	797	163911.80			
		Toplam	141702566.70	799				
	Yaratıcılık Bölümü	G.arası	1616.94	2	808.47	83.36	.000	.173
		G.içi	7739.32	798	9.70			
		Toplam	9356.26	800				
Hipotez Geliştirme	Akıcılık	G.arası	2430.31	2	1215.16	65.05	.000	.145
		G.içi	14907.06	798	18.68			
		Toplam	17337.37	800				
	Esneklik	G.arası	1351.46	2	675.73	170.65	.000	.300
		G.içi	3159.86	798	3.96			
		Toplam	4511.32	800				
	Orijinallik	G.arası	10687145.52	2	5343572.76	48.98	.000	.109
		G.içi	87067081.71	798	109106.62			
		Toplam	97754227.23	800				
	Yaratıcılık Bölümü	G.arası	1843.28	2	921.64	127.83	.000	.243
		G.içi	5753.37	798	7.21			
		Toplam	7596.66	800				

Tablo 4.8. ÇABİYAT Puanlarının Sınıf Düzeylerine Göre ANOVA Sonuçları

	Puan Türü	Varyansın Kaynağı	Kareler Toplamı	sd	Kareler Ortalaması	F	p	η^2
Toplam Test	Akıcılık	G.arası	8605.59	2	4302.80	65.71	.000	.141
		G.içi	52253.42	798	65.48			
		Toplam	60859.01	800				
	Esneklik	G.arası	4922.71	2	2461.35	247.90	.000	.383
		G.içi	7923.34	798	9.93			
		Toplam	12846.04	800				
	Orijinallik	G.arası	43187793.19	2	21593896.59	60.59	.000	.132
		G.içi	284064549.68	797	356417.25			
		Toplam	327252342.87	799				
	Yaratıcılık Bölümü	G.arası	6889.51	2	3444.75	139.03	.000	.258
		G.içi	19772.38	798	24.78			
		Toplam	26661.89	800				

Tablo 4.8’de yer alan ANOVA testi sonuçları tüm test puanlarının sınıf düzeyine göre farklılaştığını göstermektedir ($p < .001$). Ayrıca tabloda yer

alan eta kare (η^2) etki büyüklüğü değerlerinin .076 ile .383 arasında değiştiği görülmektedir. Deney tasarlama alt ölçeği akıcılık ve orijinallik puan türlerinde ve toplam testten elde edilen orijinallik puan türünde eta kare etki büyüklüğü değerlerinin .07 ile .14 arasında olması gruplar arasındaki farkın orta düzeyde, diğer puan türlerinde ise eta kare etki büyüklüğü değerlerinin .14 ve üzerinde olması gruplar arasındaki farkın yüksek düzeyde bir etkiye sahip olduğunu göstermektedir (Cohen, 1988). Gruplar arasındaki farkın kaynağını bulma amacıyla izleme testleri yapılmıştır. İzleme testlerinin sonuçları Tablo 4.9’da yer almaktadır.

Tablo 4.9. Sınıf Düzeylerine Göre İzleme Testi Bulguları

Bileşen	Puan Türü	İzleme Testi	İ	j	xi-xj	Sh	p
Deney Tasarlama	Akıcılık	Scheffe	Anasınıfı	1	-1.95	.49	.000
				2	-3.86	.48	.000
			1	2	-1.91	.45	.000
	Esneklik	Tamhane	Anasınıfı	1	-1.48	.15	.000
				2	-2.96	.18	.000
			1	2	-1.48	.15	.000
	Orijinallik	Scheffe	Anasınıfı	1	-130.27	36.52	.002
				2	-291.00	35.84	.000
			1	2	-160.73	33.68	.000
	Yaratıcılık Bölümü	Tamhane	Anasınıfı	1	-1.87	.26	.000
				2	-3.55	.29	.000
			1	2	-1.69	.26	.000
Hipotez Geliştirme	Akıcılık	Scheffe	Anasınıfı	1	-2.77	.39	.000
				2	-4.35	.38	.000
			1	2	-1.58	.36	.000
	Esneklik	Tamhane	Anasınıfı	1	-2.05	.16	.000
				2	-3.25	.19	.000
			1	2	-1.20	.16	.000
	Orijinallik	Tamhane	Anasınıfı	1	-171.54	28.98	.000
				2	-289.34	29.16	.000
			1	2	-117.80	27.74	.000
	Yaratıcılık Bölümü	Tamhane	Anasınıfı	1	-2.36	.23	.000
				2	-3.79	.26	.000
			1	2	-1.43	.22	.000
Toplam Test	Akıcılık	Tamhane	Anasınıfı	1	-4.72	.73	.000
				2	-8.21	.76	.000
			1	2	-3.50	.65	.000
	Esneklik	Tamhane	Anasınıfı	1	-3.53	.25	.000
				2	-6.21	.31	.000
			1	2	-2.68	.25	.000
	Orijinallik	Tamhane	Anasınıfı	1	-301.22	52.86	.000
				2	-580.34	54.81	.000
			1	2	-279.12	48.84	.000
	Yaratıcılık Bölümü	Tamhane	Anasınıfı	1	-4.23	.43	.000
				2	-7.35	.48	.000
			1	2	-3.12	.40	.000

Tablo 4.9 incelendiğinde akıcılık, orijinallik ve yaratıcılık bölümü puanları bağlamında alt ölçeklerde ve toplam test puanlarında ikinci sınıf öğrencilerinin anasınıfları ve birinci sınıf öğrencilerinden anlamlı bir şekilde daha yüksek puan aldıkları görülmüştür ($p<.01$). Benzer şekilde birinci sınıf öğrencilerin puanlarının anasınıfları öğrencilerinin puanlarından anlamlı bir şekilde yüksek olduğu bulunmuştur ($p<.01$). Akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları bağlamında üst sınıfların lehine anlamlı bir farkın olduğu görülmüştür ($p<.001$).

ÇABİYAT'ın gelişimsel geçerliğini incelemek için sınıf düzeylerinin karşılaştırılması sonucunda, testten elde edilen puanların tamamında üst sınıfların lehine farklılıklar bulunmuştur. Alt testlerde ve toplam ölçekten elde edilen akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanlarında orta ve yüksek düzeyde bir etkiye sahip olması, testin ayırt ediciliğinin uygulamada ve kuramda orta ve yüksek düzeyde anlamlı olduğunu göstermektedir (Huck, 2012). Analiz bulguları ölçeğin ayırt edicilik geçerliğinin yeterli düzeyde olduğuna yönelik kanıt oluşturmaktadır.

Uyum geçerliği

ÇABİYAT'ın uyum geçerliğini incelemek için SAGES 2 ile olan ilişkisi incelenmiştir. ÇABİYAT'ın alt testlerinden ve toplamından elde edilen yaratıcılık puanları ile SAGES 2 puanları arasındaki korelasyonlar hesaplanmıştır. ÇABİYAT'ın uyum geçerliğinin incelenmesinde fen bilimleri başarısı göz önünde bulundurulmuştur. Bu doğrultuda 801 katılımcının ÇABİYAT puanlarıyla SAGES 2 puanı arasındaki ilişki Pearson momentler çarpımı korelasyon analizi ile incelenmiş ve elde edilen korelasyon katsayıları Tablo 4.10'da verilmiştir.

Tablo 4.10. ÇABİYAT Puanları ve SAGES 2 Puanı Arasındaki Bağlıntı Katsayıları

Bileşen	Puan Türü	SAGES 2
Deney Tasarlama	Akıcılık	.651**
	Esneklik	.593**
	Orijinallik	.641**
	Yaratıcılık Bölümü	.684**
Hipotez Geliştirme	Akıcılık	.540**
	Esneklik	.526**
	Orijinallik	.491**
	Yaratıcılık Bölümü	.594**
Toplam	Akıcılık	.706**
	Esneklik	.642**
	Orijinallik	.690**
	Yaratıcılık Bölümü	.722**

** $p<.001$

Tablo 4.10 incelendiğinde deney tasarlama bileşenine ait akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları ile SAGES 2 puanı arasındaki korelasyon katsayılarının .593 ile .684 arasında; hipotez geliştirme bileşenine ait akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları ile SAGES 2 puanı arasındaki korelasyon katsayılarının ise .491 ile .594 arasında değiştiği görülmektedir ($p < .001$). Toplam testten elde edilen puanlarla SAGES 2 puanı arasındaki ilişki katsayısının .642 ile .722 arasında değiştiği görülmektedir. Elde edilen korelasyon değerleri göz önünde bulundurulduğunda ÇABİYAT puanları ile SAGES 2 puanı arasında yüksek düzeyde ($r > .50$) bir ilişkinin olduğu belirtilebilir (Cohen, 1988). ÇABİYAT puanları ile SAGES 2 puanı arasında anlamlı ve yüksek düzeyde bir ilişki bulunmasının ölçeğin ölçüt geçerliğine ilişkin bir kanıt oluşturduğu söylenebilir.

Maddelerinin Yanlılık Analizi

Madde yanlılığı cinsiyet, sosyo-ekonomik düzey ve etnik köken gibi değişkenler bağlamında farklı gruplarda yer alan bireylerin doğru yanıt üretme olasılığının aynı olmamasıdır (Mellenbergh, 1989). Geçerlik çalışmaları kapsamında incelenen madde yanlılık analizlerinde farklı gruplarda yer alan bireylerin puan ortalamaları karşılaştırılmaktadır. Karşılaştırmalar sonucunda gruplar arasında anlamlı bir farkın bulunmaması beklenmektedir (Shepard, Camilli ve Williams, 1985). ÇABİYAT maddelerinin yanlılığı cinsiyet değişkeni bağlamında incelenmiştir. ÇABİYAT'ın puanlarındaki kız ve erkek öğrencilerin puanlarını karşılaştırmak için bağımsız örneklem t testi gerçekleştirilmiştir. Analiz gerçekleştirilmeden önce testin normal dağılım ve varyansların eşleşliği sayıltıları incelenmiştir. Kız ve erkek öğrencilerinin ölçek maddelerinden aldıkları puanların basıklık ve çarpıklık değerlerinin -2 ile +2 aralığında olması nedeniyle verilerin normal dağıldığı sonucuna ulaşılmıştır (Trochim ve Donnelly, 2006). Varyansların eşleşliğini test etmek amacı ile gerçekleştirilen Levene testi sonuçlarına göre Oyuncak Tünel maddesinin orijinallik puanında, Suluklar maddesinin akıcılık ve esneklik puanlarında, Oyuncak Gemi maddesinin akıcılık ve orijinallik puanlarında varyansların eşleşmediği ($p < .05$), diğer puanlarda ise grupların varyanslarının eşleştiği ($p > .05$) bulunmuştur.

ÇABİYAT maddelerinin yanlılık analizi kapsamında cinsiyet değişkeni göz önünde bulundurulmuştur. Bu doğrultuda kız ve erkek öğrencilerin ölçek maddelerinden aldıkları akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanlarını karşılaştırmak için bağımsız örneklem t testi kullanılmıştır. Karşılaştırmalar için 24 (6 madde x 4 farklı puan türü) analiz yapılmıştır. Aynı veri seti üzerinden birden çok analiz yapılması durumunda Tip I hataya düşüleceği ön görülmektedir. Bu nedenle Bonferroni uyarlamasının yapılması önerilmektedir (Huck, 2012). Analizde geleneksel anlamlılık düzeyi .05 yerine

.008 (.05/24) değeri kabul edilmiştir. Gerçekleştirilen bağımsız örneklem t testi sonuçlarına Tablo 4.11’de yer verilmiştir. Tablo 4.11 incelendiğinde ölçek maddelerinden elde edilen tüm puan türlerinde kız ve erkek öğrencilerin puan ortalamalarında anlamlı bir farklılık olmadığı görülmektedir ($p>.004$). Bu bulgu, ölçek maddelerinin cinsiyet bağlamında yanlı olmadığına ilişkin bir kanıt oluşturmaktadır.

Tablo 4.11. Cinsiyete Göre ÇABİYAT Madde Puanlarına Yönelik T Testi Sonuçları

Madde	Puan Türü	Cinsiyet	N	\bar{X}	SS	sd	t	p
DT1: Hamsterlar	Akıcılık	Erkek	416	6.14	2.11	0.80	799	.426
		Kız	385	6.02	2.10			
	Esneklik	Erkek	416	2.87	0.83	1.45	799	.147
		Kız	385	2.78	0.84			
	Orijinallik	Erkek	416	315.62	181.54	1.14	799	.253
		Kız	385	301.23	173.57			
	Yaratıcılık Bölümü	Erkek	416	4.40	1.22	0.52	799	.602
		Kız	385	4.36	1.29			
DT2: Kum Havuzu	Akıcılık	Erkek	416	5.72	2.15	1.57	799	.117
		Kız	385	5.48	2.14			
	Esneklik	Erkek	416	2.99	0.81	0.78	799	.434
		Kız	385	2.95	0.80			
	Orijinallik	Erkek	416	342.57	160.08	1.68	798	.094
		Kız	384	323.40	162.93			
	Yaratıcılık Bölümü	Erkek	416	4.29	1.24	1.09	799	.274
		Kız	385	4.19	1.27			
DT3: Oyuncak Tünel	Akıcılık	Erkek	416	5.65	2.03	1.63	799	.103
		Kız	385	5.42	1.96			
	Esneklik	Erkek	416	3.17	0.95	1.06	799	.288
		Kız	385	3.10	0.91			
	Orijinallik	Erkek	416	309.76	145.45	1.85	798.82	.065
		Kız	385	291.62	132.60			
	Yaratıcılık Bölümü	Erkek	416	4.40	1.31	1.20	799	.231
		Kız	385	4.29	1.31			
HG1: Bataklıktaki Sinekler	Akıcılık	Erkek	416	4.68	1.67	0.52	799	.601
		Kız	385	4.62	1.70			
	Esneklik	Erkek	416	2.28	0.84	0.81	799	.417
		Kız	385	2.23	0.82			
	Orijinallik	Erkek	416	238.99	135.66	1.08	799	.280
		Kız	385	228.80	131.06			
	Yaratıcılık Bölümü	Erkek	416	3.41	0.99	0.41	799	.685
		Kız	385	3.38	1.02			

Madde	Puan Türü	Cinsiyet	N	\bar{X}	SS	sd	t	p
HG2: Suluklar	Akıcılık	Erkek	416	4.45	1.88	0.57	798.97	.570
		Kız	385	4.38	1.73			
	Esneklik	Erkek	416	2.30	0.98	1.66	798.81	.098
		Kız	385	2.19	0.92			
	Orijinallik	Erkek	416	303.53	146.17	0.06	799	.950
		Kız	385	302.89	139.67			
	Yaratıcılık Bölümü	Erkek	416	3.35	1.25	1.00	799	.316
		Kız	385	3.27	1.18			
HG3: Oyuncak Gemi	Akıcılık	Erkek	416	4.41	1.90	0.86	798.45	.390
		Kız	385	4.30	1.80			
	Esneklik	Erkek	416	2.68	0.94	0.87	799	.384
		Kız	385	2.62	0.91			
	Orijinallik	Erkek	416	253.95	141.47	0.82	797.79	.411
		Kız	385	245.88	136.11			
	Yaratıcılık Bölümü	Erkek	416	3.53	1.25	0.63	799	.530
		Kız	385	3.48	1.21			

GÜVENİRLİK ANALİZLERİ

ÇABİYAT'ın güvenilirliğinin incelenmesine yönelik olarak iç tutarlık güvenilirliği ve okuyucular arası güvenilirlik analizleri yapılmıştır. İç tutarlık güvenilirliği için Cronbach Alfa katsayısı hesaplanmıştır. Madde analizleri için maddeler arası ve madde-toplam korelasyon katsayıları incelenmiştir.

ÇABİYAT'ın İç Tutarlılığı

ÇABİYAT deney tasarlama ve hipotez geliştirme olmak üzere iki alt ölçekten oluşmaktadır. Her bir alt ölçekte ise üç madde yer almaktadır. Her bir madde için akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Bu nedenle her bir alt ölçekten ve ölçeğin tamamından elde edilen puanların her biri için güvenilirlik katsayıları hesaplanmıştır. İç tutarlılık analizi kapsamında öncelikle Cronbach Alfa katsayısı hesaplanmıştır. Tablo 4.12'de elde edilen değerler yer almaktadır.

Tablo 4.12 incelendiğinde, ÇABİYAT'ın toplamından elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin .80'in üzerinde olduğu görülmüştür ($\alpha_{akıcılık} = .88$; $\alpha_{esneklik} = .85$; $\alpha_{orijinallik} = .81$; $\alpha_{y.bölümü} = .88$). Benzer şekilde deney tasarlama alt ölçeğinden elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin de .80'in üzerinde olduğu görülmüştür ($\alpha_{akıcılık} = .84$; $\alpha_{esneklik} = .83$; $\alpha_{orijinallik} = .85$; $\alpha_{y.bölümü} = .87$). Hipotez geliştirme alt ölçeğinden elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin .70'in üzerinde olduğu görülmüştür ($\alpha_{akıcılık} = .84$; $\alpha_{esneklik} = .85$; $\alpha_{orijinallik} = .79$; $\alpha_{y.bölümü} = .87$). Bulgular ölçeğin güvenilirliğinin yüksek olduğunu göstermektedir.

Tablo 4.12. ÇABİYAT Toplam Test ve Alt Testlerin Cronbach Alfa Katsayıları

Ölçek	Madde Sayısı	Puan Türü	Cronbach Alfa katsayısı
Deney Tasarlama	3	Akıcılık	.84
		Esneklik	.83
		Orijinallik	.85
		Yaratıcılık Bölümü	.87
Hipotez Geliştirme	3	Akıcılık	.84
		Esneklik	.85
		Orijinallik	.79
		Yaratıcılık Bölümü	.87
Toplam Test	6	Akıcılık	.88
		Esneklik	.85
		Orijinallik	.81
		Yaratıcılık Bölümü	.88

ÇABİYAT'ın iç tutarlık güvenilirliğinin incelenmesinde ikincil olarak madde analizi gerçekleştirilmiştir. Madde analizleri kapsamında ilk olarak maddeler arası ve madde-toplam korelasyon katsayıları incelenmiştir. Çok faktörlü ölçeklerde iç tutarlılığın incelenmesinde Cronbach Alfa katsayısının yorumlanmasının yanı sıra maddeler arası korelasyon katsayılarının da incelenmesi gereklidir (Cortina, 1993). Maddeler arası ve madde toplam katsayıları Pearson Çarpım Moment Korelasyon analizi ile elde edilmiştir. Tablo 4.13'te maddelerin kendi aralarında, alt testler ve toplam test puanlarıyla korelasyon katsayıları yer almaktadır.

Tablo 4.13. ÇABİYAT Maddeler Arası ve Madde-Alt Test Korelasyon Katsayıları

		DT2	DT3	HG1	HG2	HG3	DT	HG	Toplam
Akıcılık	DT1	.727**	.707**	.389**	.351**	.315**	.908**	.402**	.797**
	DT2		.666**	.365**	.335**	.281**	.896**	.374**	.774**
	DT3			.377**	.368**	.343**	.879**	.416**	.786**
	HG1				.655**	.604**	.421**	.856**	.727**
	HG2					.655**	.393**	.886**	.725**
	HG3						.349**	.871**	.688**
Esneklik	DT1	.654**	.634**	.477**	.463**	.503**	.876**	.549**	.813**
	DT2		.593**	.349**	.367**	.395**	.855**	.424**	.727**
	DT3			.297**	.311**	.360**	.869**	.370**	.703**
	HG1				.735**	.602**	.429**	.880**	.761**
	HG2					.614**	.436**	.899**	.775**
	HG3						.482**	.848**	.771**
Orijinallik	DT1	.686**	.670**	.303**	.312**	.237**	.908**	.337**	.782**
	DT2		.606**	.290**	.308**	.214**	.874**	.322**	.751**
	DT3			.267**	.306**	.243**	.847**	.324**	.734**
	HG1				.589**	.517**	.328**	.828**	.669**
	HG2					.578**	.351**	.864**	.703**
	HG3						.263**	.831**	.627**

	DT1	.718**	.709**	.348**	.381**	.386**	.903**	.562**	.835**
	DT2		.671**	.378**	.435**	.420**	.889**	.495**	.791**
Yaratıcılık Bölümü	DT3			.338**	.411**	.441**	.891**	.489**	.789**
	HG1				.625**	.586**	.545**	.885**	.795**
	HG2					.681**	.502**	.902**	.779**
	HG3						.502**	.891**	.773**

* $p < .001$ DT1: Hamsterlar. DT2: Kum Havuzu. DT3: Oyuncak Tünel. HG1: Bataklıktaki Sinekler. HG2: Suluklar. HG3:Oyuncak Gemi. DT: Deney Tasarlama. HG: Hipotez Geliştirme.

Tablo 4.13 incelendiğinde tüm korelasyon katsayılarının anlamlı olduğu bulunmuştur ($p < .001$). Akıcılık puan türüne göre maddeler arasındaki korelasyon katsayılarının .281 ile .727 arasında, esneklik puan türüne göre maddeler arasındaki korelasyon katsayılarının .349 ile .735 arasında, orijinallik puan türüne göre maddeler arasındaki korelasyon katsayılarının .214 ile .686 arasında, yaratıcılık bölümü puan türüne göre maddeler arasındaki korelasyon katsayılarının .332 ile .709 arasında değiştiği görülmüştür. Tüm puan türlerine göre deney tasarlama maddelerinin, deney tasarlama alt ölçeği ile arasındaki korelasyon katsayılarının .847 ile .908 arasında, hipotez geliştirme alt ölçeği ile arasındaki korelasyon katsayılarının .322 ile .562 arasında, toplam test ile arasındaki korelasyon katsayılarının .669 ile .835 arasında değiştiği bulunmuştur. Tüm puan türlerine göre hipotez geliştirme maddelerinin, deney tasarlama alt ölçeği ile arasındaki korelasyon katsayılarının .263 ile .545 arasında, hipotez geliştirme alt ölçeği ile arasındaki korelasyon katsayılarının .831 ile .902 arasında, toplam test ile arasındaki korelasyon katsayılarının .627 ile .795 arasında değiştiği görülmüştür. Sonuç olarak test maddelerinin, ait oldukları alt testle ve toplam testle aralarında yüksek, diğer alt testle ise aralarında düşük ya da orta düzey bir ilişki olduğu görülmüştür. Bir ölçeğin iç tutarlılığının ortaya koyulmasında maddeler arası korelasyon katsayılarının ortalamasının .15 ile .50 arasında olması beklenmektedir (Clark ve Watson, 1995). Maddeler arasındaki korelasyon katsayılarının ortalaması akıcılık için .48, esneklik için .49, orijinallik için .41, yaratıcılık bölümü için .49 olarak hesaplanmıştır. Dolayısıyla maddeler arası korelasyon katsayılarına yönelik bulgular ölçeğin iç tutarlılığını destekler niteliktedir.

ÇABİYAT'ın iç tutarlık güvenilirliğinin test edilmesi amacıyla gerçekleştirilen madde analizi kapsamında ikincil olarak madde kalan analizi incelenmiştir. Madde kalan analizi ile puan türlerine göre maddelerin ortalamaları, standart sapmaları, madde silindiğinde ölçek ortalaması, madde silindiğinde ölçeğin varyansı, doğrulanmış madde toplam korelasyonu ve madde silindiğinde Cronbach alfa değeri hesaplanmıştır. Elde edilen değerlere Tablo 4.14'te yer verilmiştir.

Tablo 4.14. ÇABİYAT Madde Kalan Analizi

Puan Türü	Madde	\bar{X}	ss	Madde Silindiğinde Ölçek Ortalaması	Madde Silindiğinde Ölçek Varyansı	Doğrulanmış Madde Toplam Korelasyonu	Madde Silindiğinde Cronbach Alfa
Akıcılık	DT1	6.08	2.10	24.56	51.23	0.68	.81
	DT2	5.60	2.15	25.04	51.68	0.64	.82
	DT3	5.54	2.00	25.11	52.66	0.67	.81
	HG1	4.65	1.68	26.00	57.60	0.61	.82
	HG2	4.42	1.81	26.23	56.45	0.60	.82
	HG3	4.35	1.85	26.29	57.26	0.55	.83
Esneklik	DT1	2.83	0.84	13.25	11.30	0.72	.81
	DT2	2.97	0.81	13.11	12.00	0.61	.83
	DT3	3.13	0.93	12.95	11.69	0.55	.84
	HG1	2.25	0.83	13.83	11.67	0.65	.82
	HG2	2.24	0.95	13.83	11.05	0.65	.82
	HG3	2.65	0.93	13.43	11.19	0.65	.82
Orijinallik	DT1	308.59	177.87	1421.47	263256.19	0.63	.76
	DT2	333.37	161.63	1396.70	280299.82	0.60	.77
	DT3	300.92	139.68	1429.14	297835.12	0.60	.77
	HG1	234.16	133.55	1495.90	313070.39	0.53	.78
	HG2	303.09	143.03	1426.97	301342.92	0.56	.78
	HG3	249.92	138.91	1480.14	317371.00	0.47	.79
Yaratıcılık Bölümü	DT1	4.38	1.26	18.81	22.80	0.75	.85
	DT2	4.24	1.26	18.95	23.42	0.68	.86
	DT3	4.35	1.31	18.84	23.11	0.67	.86
	HG1	3.39	1.00	19.80	25.11	0.72	.86
	HG2	3.31	1.21	19.88	23.88	0.67	.86
	HG3	3.51	1.23	19.68	23.86	0.66	.86

DT1: Hamsterlar. DT2: Kum Havuzu. DT3: Oyuncak Tünel. HG1: Bataklıktaki Sinekler HG2: Suluklar. HG3:Oyuncak Gemi. DT: Deney Tasarlama. HG: Hipotez Geliştirme.

Tablo 4.14 incelendiğinde her bir puan türüne göre maddelerden birinin çıkarılması durumunda ölçeğin toplam puan ortalamasına ve toplam varyansa etkisinin benzer olduğu ifade edilebilir. Her bir alt ölçek için maddelerin birbirleri ile uyumlu olduğu düşünülebilir. Doğrulanmış madde toplam korelasyonu değerleri .47 ile .75 arasında değişmektedir. Bu değerlerinin .30 üzerinde olması ölçekte yer alan maddelerin ölçeğe hizmet ettiğini göstermektedir (Field, 2009). Tüm puan türlerine göre ölçekte yer alan herhangi bir maddenin çıkarılmasında Cronbach Alfa değerinde bir artış olmadığı görülmektedir ($\alpha_{akıcılık} = .88$; $\alpha_{esneklik} = .85$; $\alpha_{orijinallik} = .81$; $\alpha_{y.bölümü} = .88$). Bu bulgu ölçek maddelerini amacına hizmet ettiğine yönelik bir kanıt oluşturmaktadır. Madde kalan analizleri ölçeğin iç tutarlılığının yeterli düzeyde olduğunu göstermektedir.

ÇABİYAT'ın Puanlayıcılar Arası Güvenirliği

Açık uçlu sorulardan oluşan yaratıcılık testlerinin puanlamasında öznel faktörlerin etkisi söz konusudur. Bu nedenle yaratıcılık testleri gibi ölçüm araçlarında puanlayıcıdan kaynaklı hatalar olabilir. Puanlayıcıdan kaynaklı ölçüm hata oranının düşük olduğunu ortaya koymak için farklı puanlayıcıların verdiği puanların birbiriyle karşılaştırılması gereklidir. Puanlayıcılar arası güvenirlilik analizlerinde iki ya da daha fazla puanlayıcının yapmış oldukları puanlamalar arasındaki tutarlılık düzeyi belirlenmektedir (Anastasi ve Urbina, 1997; Güler ve Taşdelen Teker, 2015). ÇABİYAT'ın bir yaratıcılık testi olması nedeniyle puanlayıcılar arasındaki tutarlılık incelenmiştir. Bu amaç doğrultusunda ana uygulamada araştırmacı 103 öğrencinin tüm yanıtlarını öğrencilerin ifade ettiği şekilde not almıştır. Öğrenci yanıtları araştırmacı ve başka bir puanlayıcı tarafından birbirinden bağımsız olarak puanlanmıştır. Diğer puanlayıcı, madde havuzunun geliştirilmesinde yer alan bilimsel yaratıcılık ve BÜT puanlama yöntemi eğitimi almış bir fen bilimleri öğretmenidir. Fen bilimleri öğretmeni ve araştırmacının puanları arasındaki tutarlılığı incelemek için sınıf içi korelasyon analizinden yararlanılmıştır. Analiz bulgularına Tablo 4.15'te yer verilmiştir.

Tablo 4.15. ÇABİYAT'ın puanlayıcılar arası güvenirlilik sonuçları

Madde	Puan Türü	r _{icc}	F	sd	p
DT1: Hamsterlar	Akıcılık	0.997	1111.297	102	.000
	Esneklik	0.987	228.647	102	.000
	Orijinallik	0.997	1070.895	102	.000
	Yaratıcılık Bölümü	0.994	490.049	102	.000
DT2: Kum Havuzu	Akıcılık	0.991	310.241	102	.000
	Esneklik	0.911	21.508	102	.000
	Orijinallik	0.993	392.746	102	.000
	Yaratıcılık Bölümü	0.991	746.983	102	.000
DT3: Oyuncak Tünel	Akıcılık	0.993	1532.843	102	.000
	Esneklik	0.991	340.020	102	.000
	Orijinallik	0.994	508.572	102	.000
	Yaratıcılık Bölümü	0.996	762.238	102	.000
HG1: Bataklıkta Sinekler	Akıcılık	0.976	120.006	102	.000
	Esneklik	0.914	11.656	102	.000
	Orijinallik	0.977	129.783	102	.000
	Yaratıcılık Bölümü	0.963	26.848	102	.000
HG2: Suluklar	Akıcılık	0.998	1599.588	102	.000
	Esneklik	0.992	182.529	102	.000
	Orijinallik	0.997	1084.619	102	.000
	Yaratıcılık Bölümü	0.994	478.966	102	.000
HG3: Oyuncak Gemi	Akıcılık	0.992	380.573	102	.000
	Esneklik	0.975	115.455	102	.000
	Orijinallik	0.995	563.294	102	.000
	Yaratıcılık Bölümü	0.983	174.796	102	.000
Toplam Test	Akıcılık	0.991	2298.962	102	.000
	Esneklik	0.988	688.681	102	.000
	Orijinallik	0.981	2372.518	102	.000
	Yaratıcılık Bölümü	0.990	96.319	102	.000

BÖLÜM 5: ÇABİYAT'IN PSİKOMETRİK ÖZELLİKLERİNE İLİŞKİN TARTIŞMA VE SONUÇ

Bu çalışma kapsamında anasınıflı, birinci ve ikinci sınıf öğrencilerinin bilimsel yaratıcılık becerisini ölçmek amacıyla Bilimsel Keşfin İkili Arama Modeli (Klahr ve Dunbar, 1988) temel alınarak, çocuklar için animasyonlu bilimsel yaratıcılık testi geliştirilmiş ve testin psikometrik özellikleri incelenmiştir. Testin madde havuzu, alanında uzman bir grup ile oluşturulmuştur. Madde havuzundan 12 madde uzman görüşü alınarak seçilmiştir. 12 madde ön deneme uygulaması kapsamında 20 öğrenciye uygulanmıştır. Ön deneme uygulaması sonunda bir maddenin testten çıkarılmasına diğer maddelerde de çeşitli revizyonlar yapılmasına karar verilmiştir. Sonrasında 11 maddelik test 121 öğrenciye uygulanarak pilot çalışma gerçekleştirilmiştir. Pilot çalışma bulgularına göre üç madde daha testten çıkarılmıştır. Sekiz maddelik test 801 öğrenciye uygulanmış ve ölçeğin psikometrik özellikleri incelenmiştir. Bu kapsamda ilk olarak açılımlayıcı faktör analizi ile madde seçimi yapılmış ve testin altı maddelik son formu oluşturulmuştur. Altı maddelik ölçeğin geçerlik analizleri için yapı geçerliği, ayırt edicilik geçerliği ve ölçüt geçerliği; güvenirlik analizleri için iç tutarlılığa yönelik analizler ve puanlayıcılar arası güvenirlik analizleri yapılmıştır.

ÇABİYAT'ın Geçerliğine Yönelik Tartışma ve Sonuç

ÇABİYAT, fen bilimleri alanındaki yaratıcı düşünme becerilerini ölçmek amacıyla geliştirilmiş bir testtir. Yaratıcılık testlerinin büyük bir çoğunluğunda sınırsız doğru cevap bulunmaktadır. Önemli bir sorun olarak kabul edilen bu durum yaratıcılık testlerin puanlanmasını zorlaştırmaktadır (Kaufman, Plucker ve Baer, 2008). Ön deneme, pilot ve ana uygulamada ulaşılan 942 öğrencinin yanıtlarına göre oluşturulan ÇABİYAT'ın doğru cevap havuzunda 175 doğru yanıt yer almaktadır. Bu açıdan ÇABİYAT'ın puanlamasının diğer yaratıcılık testlerine göre daha pratik ve kolay olabileceği ifade edilebilir. Diğer taraftan Plucker ve Zabelina (2009) yaratıcılık testlerinin yeteri kadar zor sorulardan oluşmamasından dolayı yaratıcılık düzeyleri yüksek öğrencilerin gerçek düzeylerinin belirlenemediğini iddia etmektedirler. Bu durum tavan etkisi olarak adlandırılmaktadır. ÇABİYAT'ın son formunda yer alan maddelerin

ortalamaları 4.35 (HG3: Oyuncak Gemi) ile 6.08 (DT1: Hamsterlar) arasında değişmektedir. Maddelere ait doğru yanıt havuzunda ortalama 30 yanıt olduğu düşünüldüğünde maddelerin çok kolay olmadığı söylenebilir. Diğer taraftan madde ortalamalarının düşük olması taban etkisine neden olabilir. Yaratıcılık testlerinin çok zor sorulardan oluşması durumunda bireylerin çoğunluğu düşük puan alır ve düşük puan potansiyele sahip bireylerin ayırt edilmesini güçleştirir (Perry ve Karpova, 2017). ÇABİYAT maddelerine verilen en düşük doğru cevap sayısı 0 ve 1'in frekansı çok düşüktür (Bkz. Tablo 4.1). Bu durum maddelerin çok zor olmadığını göstermektedir. Bu bağlamda ÇABİYAT'ın taban ve tavan sorununun olmadığı söylenebilir.

ÇABİYAT'ın Yapı Geçerliğine Yönelik Tartışma ve Sonuç

ÇABİYAT'ın teorik çerçevesine göre bilimsel yaratıcılık deney tasarlama ve hipotez geliştirme olmak üzere iki bileşenden oluşmaktadır. Bu teorik çerçeveyi test etmek amacıyla açımlayıcı faktör analizi ve doğrulayıcı faktör analizi yapılmıştır. Analizleri gerçekleştirmek için farklı iki veri seti kullanılmıştır. Yapılan analiz sonucunda teorik olarak tanımlanan altı maddenin kendi faktörleri altında toplandığı ve iki faktörlü yapıyı desteklediği görülmüştür. Deney tasarlama ve hipotez geliştirme faktörlerinin varyansa yaptıkları toplam katkı %70.68 olarak bulunmuştur. Dunteman (1989) çok faktörlü desenlerde açıklanan varyansın %40 ile %60 arasında olmasının yeterli olduğunu belirtmektedir. Bu bağlamda iki faktörün toplam varyansa yaptığı katkının oldukça iyi seviyede olduğu görülmektedir. Diğer taraftan maddelerin faktör yüklerinin .68 ile .90 arasında değerler alması, maddelerin buldukları faktörleri iyi ve mükemmel seviyede temsil ettiğini göstermiştir (Comrey ve Lee, 1992).

ÇABİYAT'ın yapı geçerliğini test etmek amacıyla gerçekleştirilen doğrulayıcı faktör analizi kapsamında tek faktörlü ve iki faktörlü model sınanmıştır. Elde edilen bulgulara göre ölçeğin iki faktörlü modelinin uyum iyiliği değerlerinin mükemmel düzeyde olduğu sonucuna ulaşılmıştır. Bu nedenle ÇABİYAT'ın iki faktörlü ve altı maddeli yapısını incelemek için ikinci düzey faktör analizi yapılmış ve teorik alt yapının mükemmel bir uyuma sahip olduğu görülmüştür.

ÇABİYAT'ın yapı geçerliğini test etmek amacıyla son olarak ölçeğin yakınsak ve ıraksak geçerliği incelenmiştir. Bu kapsamda doğrulayıcı faktör analizinde elde edilen faktör yük değerleri kullanılarak ($\lambda_{\text{en düşük}} = .749$; $\lambda_{\text{en yüksek}} = .858$) AVE, CR ve karekök AVE değerleri hesaplanmıştır. Deney tasarlama alt ölçeği için AVE değeri .635, CR değerinin .839; hipotez geliştirme alt ölçeği için AVE değerinin .676, CR değerinin ise .862 olduğu bulunmuştur. AVE değerlerinin .50'nin, CR değerlerinin ise .70'in üzerinde olması ölçeğin yakınsak geçerliğinin iyi bir düzeyde olduğunu göstermiştir (Fornell ve

Larcker, 1981). Diğer taraftan deney tasarlama ve hipotez geliştirme alt ölçekleri arasındaki korelasyon katsayısının AVE değerlerinin karekökünden küçük bulunması, deney tasarlama ve hipotez geliştirme yapılarının birbirleri ile ilişkili ancak birbirlerinden farklı yapılar olduğunu ortaya koymuştur. Sonuç olarak ÇABİYAT'ın teorik alt yapısı doğrulanmıştır.

ÇABİYAT'ın teorik alt yapısının doğrulanması ve açıklayıcı ve doğrulayıcı faktör analizinde yüksek faktör yük değerlerinin bulunmasının bir nedeni ölçek maddelerinin teoriye uygun olarak geliştirilmiş olması, diğer bir ifadeyle içerik geçerliğinin sağlanmış olmasıdır. ÇABİYAT'ın içerik geçerliğini sağlamak için maddeler uzman bir ekiple testin teorik yapısını uygun olacak şekilde geliştirilmiştir. Ayrıca içerik geçerliğini oluşturmak için maddeler uzman görüşüne sunulmuştur. Uzman dönütleri incelendiğinde de maddelerin ölçtüğü beceri konusunda %90'nın üzerinde uyum yakalanmıştır.

ÇABİYAT'ın teorik alt yapısının doğrulanması ve açıklayıcı ve doğrulayıcı faktör analizinde yüksek faktör yük değerlerinin bulunmasının bir nedeni Bilimsel Keşfin İkili Arama Modeli'nin sadeleştirilmesidir. Bilimsel Keşfin İkili Arama Modeli'nin üç bileşenli yapısını temel alan, 6-8. sınıf öğrencilerinin bilimsel yaratıcılığını ölçme amacıyla geliştirilen BÜT'ün yapı geçerliğinin incelemesi kapsamında Sak ve Ayas (2013), Temel Bileşenler Analizi yapmış ve tek faktörün ölçeğin %34.45'ini açıkladığını bulmuşlardır. Ayas ve Sak (2014) bir başka çalışmada tek faktörlü yapıyı doğrulamışlardır. Bermejo, vd. (2016) ölçeğin İspanyol kültürüne uyarlamışlar ve çalışmalarında tek faktörlü yapının doğrulandığı sonucuna ulaşmışlardır. Ayas (2017) BÜT'ün 3-5. sınıf öğrencilerine yönelik formunu geliştirirken Bilimsel Keşfin İkili Arama Modeli'ni sadeleştirmiştir. Ayas, bu çalışmada benzer şekilde, kanıt değerlendirme becerisini teorik yapıdan çıkarmıştır. Araştırmacı geliştirdiği ölçeğin yapı geçerliğini incelemek için Temel Bileşenler Analizi ve doğrulayıcı faktör analizi yapmıştır. Temel bileşenler analizi sonucunda iki faktörlü yapının %78.99'unu açıkladığı bulunmuş ve doğrulayıcı faktör analizi ile de iki faktörlü yapı doğrulanmıştır. Ancak madde seçimlerinde temel bileşenler analizinin yerine açıklayıcı faktör analizinin yapılması önerilmektedir (Akbulut, 2010). Çünkü bu analiz ile veri setinden elde edilebilecek azami varyansa ulaşmak hedeflenmektedir. Bu nedenle faktörlerin doğasına yönelik fikrin olmadığı durumlarda kullanılması önerilmektedir (Tabachnick ve Fidell, 2007). Bu çerçevede açıklanan varyansın daha düşük olacağı düşünülmektedir. Bu değer düşmesine rağmen sosyal bilimler alanı için yeterli düzeyde olacağı söylenebilir (Dunteman, 1989). Bu bağlamda iki faktörlü yapının doğrulandığı belirtilebilir. BÜT'ün iki faktörlü teorik yapısının anasınıfı-beşinci sınıf öğrencileri için uygun olduğu ifade edilebilir.

Alanyazında var olan bilimsel yaratıcılık testleri incelendiğinde yapı geçerliğinin incelenmediği ölçek geliştirme çalışmalarına (Sharma ve Mahrshi, 2017; Yang, vd., 2016, 2019) rastlanmaktadır. Sınırlı sayıda da olsa geliştirilen ölçeklerde yapı geçerliğinin incelendiği görülmektedir. Bu testlerin çoğunluğunun kâğıt-kalem testi olduğu ve çoğul düşünme becerilerini ölçmeye odaklanıldığı görülmektedir. Bilimsel Yaratıcılığın Yapı Modeli'ni temel alarak geliştirilen testlerde üç bileşenli bir yapı önerilmesine rağmen tek faktörlü yapı elde edildiği görülmektedir (Chin ve Siew, 2005; Çeliker ve Balım, 2012; Hu ve Adey, 2002; Samsudin vd., 2018). Diğer taraftan ortaokul öğrencilerinin (5-8 sınıf) bilimsel yaratıcılıklarını ölçmek için Yaratıcı Bilimsel Çağrışımlar Testi'ni geliştiren Kanlı (2014a), 19 maddenin üç faktör altında toplandığını ve üç faktörün varyansa yaptığı katkının %47.4 olduğunu bulmuştur. Ayrıca Kanlı önerdiği modelin yapısını doğrulayıcı faktör analizi ile doğrulamıştır.

Sonuç olarak alanyazında var olan bilimsel yaratıcılık testleri genel olarak değerlendirildiğinde genellikle tek faktörlü yapıya ulaşıldığı görülmektedir. ÇABİYAT'ın ise iki faktörlü yapısı açımlayıcı faktör analizi bulguları ile desteklemiştir. Ayrıca ikinci düzey doğrulayıcı faktör analizi bulguları ile de ölçeğin iki faktörlü iki düzeyli yapısı doğrulanmıştır. Ölçeğin iraksak ve yakınsak geçerliğinin de yeterli düzeyde olması, ÇABİYAT'ın kuramsal yapısını desteklemektedir.

ÇABİYAT'ın Ölçüt Geçerliğine Yönelik Tartışma ve Sonuç

ÇABİYAT'ın ölçüt geçerliğini değerlendirmek için ölçeğin gelişimsel geçerliği ve uyum geçerliğine yönelik analiz bulguları incelenmiştir.

ÇABİYAT'ın gelişimsel geçerliğine yönelik tartışma ve sonuç

Gelişimsel geçerliğe yönelik analizlerde ölçek puanları ile katılımcıların yaşları arasındaki ilişki incelenmektedir. ÇABİYAT'ın katılımcılarının yaşları 4 ile 8 arasında değişmektedir. Araştırma katılımcılarının yaşları ile sınıf düzeyleri arasında yüksek ve anlamlı bir ilişki bulunmuştur ($r = .968, p < .001$). Bu nedenle ölçeğin gelişimsel geçerliğini incelemek için anasınıfı, birinci ve ikinci sınıf düzeyindeki öğrencilerin ÇABİYAT puanları karşılaştırılmıştır. ÇABİYAT'ın alt ölçeklerinden ve toplamından akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları elde edilmektedir. Testten elde edilen tüm puan ortalamalarının sınıf düzeyi arttıkça arttığı bulunmuştur ($p < .01$). Söz konusu farklılığın anlamlılığını test etmek amacıyla bağımsız gruplar için tek faktörlü ANOVA yapılmıştır. Analiz sonucunda alt testlerde ve toplam testte akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları bağlamında üst sınıfların lehine anlamlı bir farkın olduğu görülmüştür ($p < .01$). Diğer bir deyişle sınıf düzeyi ile puan türleri arasında doğrusal bir ilişki bulunmuştur. Bu bulgu alanyazındaki diğer bilimsel yaratıcılık testlerinin sınıf düzeyine göre ayırt

edicilik analizleri çalışmalarından ayrılmaktadır. Diğer çalışmalarda belirli yaş aralıklarında yaratıcılık puanlarında bir düşüş olduğu tespit edilmiştir. Ayas (2017) dördüncü sınıflarla üçüncü sınıflar arasında, Hu ve Adey (2002) 13 ile 15 yaşındaki öğrenciler arasında, Kanlı (2014a) beşinci ve altıncı sınıflar arasında anlamlı farklılık olmadığını ortaya koymuşlardır. Farklı disiplinlere özgü yaratıcılık testlerinde sınıf düzeyi ile yaratıcılık puan türleri arasında doğrusal bir ilişki bulunduğu görülmektedir. Örneğin Sak ve Maker (2005), 1.-5. sınıflarda 841 öğrencinin matematiksel yaratıcılıklarını karşılaştırmışlar ve üst sınıfların lehine bir farklılık olduğunu bulmuşlardır. ÇABİYAT'ın sınıf düzeyi ile alt ölçeklerinden ve toplamından elde edilen akıcılık, orijinallik ve yaratıcılık bölümü arasında doğrusal bir ilişki bulunmasının nedeni yaratıcılığın alan bilgisi ve deneyimle olan ilişkisi olabilir. Sınıf düzeyi arttıkça öğrencilerin bilgi düzeyleri ve deneyimleri artmaktadır. Alanyazında yaratıcılığın alan bilgisi ve deneyimle ilişkili olduğu iddia edilmektedir. (Amabile, 1987; Feldhusen, 1995; Lubart, 1994; Sternberg ve Lubart, 1991).

ÇABİYAT'ın akıcılık, orijinallik ve yaratıcılık bölümü puanları ile yaş arasında doğrusal bir ilişki söz konusudur. Alanyazında yaş ile genel yaratıcılık arasındaki ilişkiye yönelik bir fikir birliği söz konusu değildir (Starko, 2017). Warren, vd. (2018) 4-5 yaş grubundaki çocuklarla, 6-7 yaş grubundaki çocukların çoğul düşünme becerilerini karşılaştırmışlar ve üst yaş grubunun lehine anlamlı fark bulmuşlardır. Ayrıca cinsiyet, zekâ ve yaş değişkenleri arasından yaş değişkeninin çoğul düşünme becerisini en iyi yordayan değişken olduğu sonucuna ulaşmışlardır. Urban (1991) 4 ile 8 yaş aralığındaki çocukların yaratıcılıklarını karşılaştırmıştır. Çalışmasının sonunda 4-5 ve 6-8 yaş aralığındaki çocuklarda yaş düzeyi arttıkça yaratıcılığın da arttığını ancak 5 yaşındaki çocukların 6 yaşındaki çocuklara göre yaratıcılık testinden anlamlı bir şekilde yüksek puan aldığını bulmuştur. Araştırmacı 6 yaşındaki çocukların okula başlaması nedeniyle yaratıcılıklarının düştüğünü iddia etmektedir. Smith ve Carlsson (1983) 7 ile 11 yaş arasındaki öğrencilerin yaratıcılık düzeylerini karşılaştırmışlar ve yaş düzeyi arttıkça yaratıcılığın da arttığını bulmuşlardır. Bu çalışmanın devamı niteliğindeki bir başka araştırmada Smith, Carlsson ve Danielsson, (1985) 14 ile 16 yaş aralığındaki öğrencilerin yaratıcılıklarını karşılaştırmışlar ve yaratıcılık ile yaş arasında pozitif ve doğrusal bir ilişki olduğu sonucuna ulaşmışlardır. Findlay ve Lumsden (1988) ise yedi yaşından erken ergenlik dönemine kadar olan dönemde akran baskısı ve çevresel faktörlerden dolayı bireylerin yaratıcılık düzeylerinin doğrusal bir şekilde düştüğünü iddia etmektedir.

Torrance (1967), okulöncesi dönemden erken yetişkinliğe kadar olan süreçte yaratıcılık ile yaş arasında doğrusal bir ilişki olmadığını belirli yaş dönemlerinde düşüş olduğunu boylamsal ve kültürlerarası araştırmalarla ortaya

koymuştur. 5, 9 ve 12 yaşlarında yaratıcılığın düştüğünü ileri sürmüştür. Runco (1989) 4.-6. sınıflarda öğrenim gören 104 öğrenci ile yaptığı çalışmada yaş ile yaratıcılık, orijinallik, teknik beceriler gibi alanlar arasında bir ilişki olmadığını ortaya koymuştur. Bu yaş aralığında olan öğrencilerle başka bir araştırma yapan Runco (1991) yaratıcılıkla yaş arasında doğrusal bir ilişki olmadığını, belirli yaş dönemlerinde düşüş olduğu sonucuna ulaşmıştır. Bu iddiayı destekleyen diğer çalışmalarda çocukluk ile yetişkinlik arası dönemde yaş ile yaratıcılık arasındaki ilişkinin ters U şeklinde parabolik bir yapıda olduğu iddia edilmektedir (Davis, 1993; Pariser ve Van Den Berg, 1997). Smolucha ve Smolucha (1985) ise yaş ile yaratıcılık arasındaki ilişkinin ters J şeklinde olduğunu ileri sürmektedirler. Araştırmacılara göre yaratıcılık düzeyinde 6 yaşında küçük keskin bir artış, 20'li yaşlarda ise daha büyük keskin bir artış görülmektedir.

Sonuç olarak ÇABİYAT'ın geneli dikkate alındığında alt testlerde ve toplam ölçekten elde edilen puan türlerinde sınıflar arasında anlamlı farklılıklar bulunmuştur. Ayrıca alt testlerde ve toplam ölçekten elde edilen akıcılık, esneklik orijinallik ve yaratıcılık bölümü puanlarında orta ve yüksek düzeyde bir etkiye sahip olmasından (Huck, 2012) dolayı testin ayırt ediciliğinin uygulamada ve kuramda orta ve yüksek düzeyde etki yarattığı şeklinde yorumlanmıştır. Bu çerçevede bilimsel yaratıcılık düzeyi kapsamında ÇABİYAT'ın farklı sınıf düzeyindeki öğrencileri ayırt edebildiği ifade edilebilir.

ÇABİYAT'ın uyum geçerliğine yönelik tartışma ve sonuç

Bilgi birikiminin yaratıcılığı pozitif yönlü olarak etkilediği iddiası pek çok araştırmacı tarafından desteklenmektedir. Bu araştırmacılar yaratıcılık sürecinde bilgi ediniminin önemine vurgu yapmışlar ve bireylerin bilgilerinin yetersiz olduğu alanlarda yaratıcılık gösteremeyeceklerini belirtmişlerdir (örn. Baer, 2012; Boden, 2001; Craft, 2005; Csikszentmihalyi, 1996; Feldhusen, 1995; Gero ve Maher, 2013; Glaser, 1984; Hayes, 1989; Ivcevic, 2007; Kulkarni ve Simon, 1988). Bu bağlamda, alana özgü bilgi birikimi yaratıcı çalışmaların ön koşullarından biri olarak kabul edilmektedir (Weisberg, 2009). Alana özgü bilgi birikimi bir alandaki başarı düzeyi ile değerlendirilebilir. Bu nedenle ÇABİYAT'ın uyum geçerliğini değerlendirmek için ölçek puanlarıyla SAGES 2 puanı arasındaki ilişki incelenmiştir. Analiz sonucunda ölçeğin alt ölçeklerinden ve toplam testten elde edilen akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanları ile SAGES 2 puanı arasında pozitif, anlamlı ve yüksek düzeyde bir ilişki bulunmuştur ($r_{\min}=.49$, $r_{\max}=.72$, $p<.001$).

Alanyazında yaratıcılık ve bilgi arasındaki ilişkiyi ortaya koymak amacıyla pek çok araştırma gerçekleştirilmiştir. Gajda, Karwowski ve Beghetto (2016), 1960 yılından 2016 yılına kadar yaratıcılık ve başarı arasındaki ilişkiyi inceleyen 120 çalışmanın bulguları ile bir meta-analiz çalışma gerçekleştirmişlerdir.

Yaratıcılığın akıcılık, esneklik, orijinallik ve değerlendirme bileşenlerini kapsadığı kabul edilmiştir. Yaratıcılık ve başarı arasındaki korelasyon katsayılarının ortalamasının .22 olduğu bulunmuştur. Araştırmacılar bu değer düşük olduğunu ancak yaratıcılık ve akademik başarı arasında güçlü bir ilişki olduğunu vurgulamışlardır. Bu değer düşük olmasında zaman, eğitim düzeyi ve ölçüm araçları gibi değişkenlerin etkisinin olabileceğini ileri sürmüşlerdir. Araştırmacılar farklı zaman dilimlerinde gerçekleştirilen araştırma bulgularını karşılaştırmış ve analizleri sonucunda yaratıcılık ve başarı arasındaki ilişkinin zamana göre değişmediğini ortaya koymuşlardır. Yaratıcılık ve akademik başarı arasındaki ilişki katsayısının tüm eğitim kademelerinde benzer olduğu sonucuna ulaşılmıştır. Araştırmacılar kullanılan ölçüm araçlarına göre yaratıcılık ve başarı arasındaki ilişkiyi karşılaştırmışlardır. Yaratıcılık ölçümlerinde öz değerlendirme yönteminin kullanıldığı araştırmalarda yaratıcılık ve başarı arasındaki ilişki katsayısının .12, yaratıcılık ölçümlerinde testlerin kullanıldığı araştırmalarda yaratıcılık ve başarı arasındaki ilişki katsayısının .23, yaratıcılık ölçümlerinde sözel olarak uygulanan testlerin kullanıldığı araştırmalarda yaratıcılık ve başarı arasındaki ilişki katsayısının .30 olduğu rapor edilmiştir. Ayrıca başarı ölçümlerinde okul puanları yerine standart bir test kullanıldığında daha güçlü ilişkilerin bulunduğu sonucuna ulaşılmıştır. Araştırmanın sonunda yaratıcılık ve akademik başarı arasında doğrusal ve güçlü bir ilişki olduğunu ortaya koymuşlardır. Feldhusen (2002) de yaratıcılık ve başarı arasında doğrusal bir ilişki olduğu görüşünü desteklemektedir. Feldhusen erken çocukluk döneminde olan ve hızlandırma eğitimine katılan iki çocuğun yaratıcılık süreçlerini gözlemlemiştir. Araştırmacı çocuklarla yaptığı görüşmelerde çocukların akıcılık, esneklik ve değerlendirme becerilerinin bilgi düzeyine göre şekillendiği sonucuna ulaşmıştır.

ÇABİYAT'ın puanları ile fen bilimleri başarısı arasındaki uyumun alanyazında bu iki değişken arasında anlamlı ilişki bulan çalışma bulguları ile paralellik göstermektedir. Ayas ve Sak (2014) altıncı sınıf öğrencilerinin BÜT puanları ile matematik notları arasındaki korelasyon katsayısını .31, BÜT puanları ile fen bilimleri notları arasındaki korelasyon katsayısını .35, BÜT ile MYT (Matematiksel Yetenek Testi) puanları arasındaki korelasyon katsayısını .52 olarak rapor etmişlerdir. Testin 3-5. sınıflara öğrencilere yönelik formunda da benzer sonuçlara ulaşılmıştır (Ayas, 2017). Ayas dördüncü sınıf öğrencilerin bilimsel yaratıcılık düzeyleri ile fen bilimleri ders notları arasındaki korelasyon katsayısının .41, bilimsel yaratıcılık düzeyleri ile matematik ders notları arasındaki korelasyon katsayısının .40, beşinci sınıf öğrencilerin bilimsel yaratıcılık düzeyleri ile fen bilimleri ders notları arasındaki korelasyon katsayısının .32, bilimsel yaratıcılık düzeyleri ile matematik ders notları arasındaki korelasyon katsayısının .35 olduğunu rapor etmiştir. Ayrıca sınıf düzeyine göre bilimsel yaratıcılık ile ders notları arasındaki

ilişkinin farklılaşmadığı sonucuna ulaşılmıştır ($p>.05$). Kanlı (2014a) ortaokul öğrencilerinin fen bilimleri notlarının arttığına bilimsel yaratıcılık puanlarının da arttığını ortaya koymuştur ($r = .390$, $p<.001$). Öğrencilerin çağrışım kurma puanları ile fen bilimleri notları arasındaki korelasyon katsayısının .34, öğrencilerin analogik muhakeme puanları ile fen bilimleri notları arasındaki korelasyon katsayısının .32 olduğu, öğrencilerin analogik problem çözme puanları ile fen bilimleri notları arasındaki korelasyon katsayısının .35 olduğu rapor edilmiştir. Sharma (2015) da lise öğrencilerinin öğrencinin bilimsel yaratıcılık düzeyleri ile fen bilimleri başarıları arasındaki ilişkiyi araştırmış fen bilimleri başarıları yüksek öğrencilerin bilimsel yaratıcılık puanlarının diğer öğrencilerin puanlarından anlamlı bir şekilde yüksek olduğu sonucuna ulaşmıştır.

ÇABİYAT'ın puanları ile fen bilimleri başarıları arasındaki uyumun alanyazında bu iki değişken arasında anlamlı ilişki bulan çalışma bulguları ile paralellik gösterdiği ifade edilebilir. ÇABİYAT'ın puanları ile fen bilimleri başarıları arasındaki korelasyon katsayısı diğer çalışma bulgularından daha yüksek olduğu görülmüştür. Bu durumun bir nedeni kullanılan ölçüm araçları olabilir. Diğer çalışmalarda fen bilimleri başarıları için genellikle okul puanları kullanılmaktadır. Ancak bu çalışma kapsamında standart bir ölçüm aracı kullanılmıştır.

ÇABİYAT'ın Madde Yanlılığına Yönelik Tartışma ve Sonuç

Ölçme değerlendirme uzmanları ölçek maddelerinin yanlı olmaması gerektiği konusunda hemfikirdirler (Shepard, Camilli ve Williams, 1985). ÇABİYAT'ın madde yanlılığını incelemek için kız ve erkek öğrencilerin puanları karşılaştırmıştır. Analiz sonucunda ölçek maddelerinden elde edilen akıcılık, esneklik, orijinallik ve yaratıcılık bölümü puanlarında kız ve erkek öğrencilerin puan ortalamalarında anlamlı bir farklılık olmadığı bulunmuştur ($p>.004$). Bu bulgu ölçek maddelerinin cinsiyet bağlamında yanlı olmadığını göstermektedir.

Alanyazında yaratıcılık alanındaki cinsiyet farklılıkları oldukça dikkat çeken ve araştırılan bir konudur. Ancak araştırma bulguları konusunda bir tutarsızlık söz konusudur. Baer ve Kaufman (2008) yaratıcılık alanındaki cinsiyet farklılıklarını inceleyen araştırmalara yönelik bir meta analiz çalışma gerçekleştirmişlerdir. Çalışmaları kapsamında yaratıcılığın ölçümünde çoğul düşünme testlerin kullanıldığı 78 araştırmaya ulaşılmıştır. Bu araştırmaların çoğunluğunda akıcılık, esneklik, değerlendirme ve orijinallik bileşenlerinde cinsiyet farklılıklarının olmadığı rapor edilmiştir. Bu bağlamda çoğul düşünme testlerinde cinsiyet farklılıklarının olmaması beklenebilir.

Bilimsel yaratıcılık alanındaki cinsiyet farklılıklarını inceleyen araştırmalar yaratıcılık alanına göre daha sınırlıdır. Ancak sınırlı araştırma bulguları bu

çalışmanın bulgusu ile paralellik göstermektedir. Araştırmalarda erkeklerin akıcılık, esneklik, orijinallik puan ortalamalarının kız öğrencilerin puan ortalamalarından daha yüksek olduğu ancak aradaki farkın anlamlı olmadığı ortaya konulmuştur (Ayverdi, vd., 2012; Baysal, Kara ve Üçüncü, 2013; Hu, vd., 2010; Kanlı, 2017; Mohamed, 2006; Sansanwal ve Sharma, 1993; Shukla ve Sharma, 1986). Özdemir ve Sak (2013) altıncı sınıfta öğrenim gören 704 öğrencinin BÜT puanlarındaki cinsiyet farklılıklarını araştırmışlardır. Çalışmalarının sonucunda hipotez geliştirme, akıcılık ve toplam yaratıcılık becerilerinde erkeklerin lehine bir farklılık bulmuşlar ancak farkın küçük bir etkiye sahip olduğunu rapor etmişlerdir. Ayrıca deney tasarlama, kanıt değerlendirme ve esneklik becerilerinde cinsiyet bağlamında bir farklılık olmadığı sonucuna ulaşmışlardır. Alanyazında bilimsel yaratıcılık alanındaki cinsiyet farklılıklarına yönelik araştırma bulguları incelendiğinde bilimsel yaratıcılık alanında cinsiyet bağlamında bir farklılık olmadığı konusunda bir fikir birliği olduğu belirtilebilir. Sonuç olarak ÇABİYAT maddelerinde cinsiyet farklılıklarının olmadığı bulunması cinsiyet bağlamında maddelerin yanlı olarak hazırlanmadığını göstermektedir.

ÇABİYAT'ın Güvenirliğine Yönelik Tartışma ve Sonuç

ÇABİYAT'ın güvenilirliğini test etmek için iç tutarlılığa ve puanlayıcılar arası güvenilirliğe ilişkin analizler yapılmıştır.

İç Tutarlılığa Yönelik Tartışma ve Sonuç

ÇABİYAT'ın iç tutarlılığını incelemek için Cronbach Alfa katsayısı incelenmiş ve madde analizi yapılmıştır. Madde analizi kapsamında maddeler arası ve madde-toplam korelasyon katsayıları ve madde kalan analizleri incelenmiştir.

ÇABİYAT'ın toplamından elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin .80'in üzerinde olduğu görülmüştür ($\alpha_{\text{akıcılık}} = .88$; $\alpha_{\text{esneklik}} = .85$, $\alpha_{\text{orijinallik}} = .81$, $\alpha_{\text{y.bölümü}} = .88$). Benzer şekilde deney tasarlama alt ölçeğinden elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin de .80'in üzerinde olduğu görülmüştür ($\alpha_{\text{akıcılık}} = .84$, $\alpha_{\text{esneklik}} = .83$, $\alpha_{\text{orijinallik}} = .85$, $\alpha_{\text{y.bölümü}} = .87$). Hipotez geliştirme alt ölçeğinden elde edilen tüm puan türlerinin Cronbach Alfa değerlerinin .79'un üzerinde olduğu görülmüştür ($\alpha_{\text{akıcılık}} = .84$, $\alpha_{\text{esneklik}} = .85$, $\alpha_{\text{orijinallik}} = .79$, $\alpha_{\text{y.bölümü}} = .87$). Bulgular ölçeğin güvenilirliğinin yüksek olduğunu göstermektedir.

Alanyazında BÜT ile yapılan çalışmalar incelendiğinde toplam test için Cronbach alfa değerinin .76 ile .91 arasında değiştiği görülmektedir (Ayas, 2010; Ayas ve Sak, 2008, 2014; Sak, 2010; Sak ve Ayas, 2013). Bu çalışmalarda Cronbach alfa değerinin hesaplanmasında tüm alt testlerden elde

edilen akıcılık, esneklik ve yaratıcılık bölümü puanları dâhil edilmiştir. Ayas (2017) bu durumun Cronbach alfa değerlerini yükselteceğini ileri sürmektedir. Bu nedenle 3-5. sınıflara yönelik formunu geliştirdiği çalışmasında, toplam test ve alt testlerin iç tutarlılık değerlerini ayrı ayrı hesaplamıştır. Ayas, BÜT 3-5 için Cronbach alfa değerlerinin toplam testten elde edilen puan türlerinde $\alpha_{\text{akıcılık}} = .743$, $\alpha_{\text{esneklik}} = .741$, $\alpha_{\text{y.bölümü}} = .736$ olduğu, deney tasarlama alt ölçeğinden elde edilen puan türlerinde $\alpha_{\text{akıcılık}} = .828$, $\alpha_{\text{esneklik}} = .813$, $\alpha_{\text{y.bölümü}} = .825$ olduğu, hipotez geliştirme alt ölçeğinden elde edilen puan türlerinde ise $\alpha_{\text{akıcılık}} = .594$, $\alpha_{\text{esneklik}} = .433$, $\alpha_{\text{y.bölümü}} = .583$ olduğunu bulmuştur. Özellikle hipotez alt testinin düşük iç tutarlılığa sahip olduğunu dikkat çekmektedir. Alanyazında diğer bilimsel yaratıcılık testleri (Chin ve Siew, 2005; Çeliker ve Balım, 2012; Hu ve Adey, 2002; Siew, Chong ve Chin, 2014; Samsudin vd., 2017; Usta ve Akkanat, 2015) incelendiğinde Cronbach alfa değerinin .68 ile .89 arasında değiştiği görülmektedir. Alanyazında var olan ölçeklerin iç tutarlılık değerleri ve istatistiksel olarak kabul edilen değerler göz önünde bulundurulduğunda ÇABİYAT'ın iç tutarlılığının iyi düzeyde olduğu ifade edilebilir.

ÇABİYAT'ın iç tutarlılık güvenirliliğinin bir diğer göstergesi olarak madde analizi yapılmıştır. Madde analizi kapsamında ilk olarak maddeler arasındaki korelasyon katsayıları incelenmiştir. Bu değerler .214 ile .735 arasında değiştiği bulunmuştur. Maddeler arasındaki korelasyon katsayılarının ortalamasının akıcılık puanına göre .48, esneklik puanına göre .49, orijinallik puanına göre .41, yaratıcılık bölümü puanına göre .49 olduğu sonucuna ulaşılmıştır. Ayrıca test maddelerinin, ait oldukları alt testle ve toplam testle aralarında yüksek, diğer alt testle ise aralarında düşük ya da orta düzey bir ilişki olduğu görülmüştür. İkinci olarak madde kalan analizi yapılmıştır. Madde toplam korelasyon değerlerinin ise .47 ile .75 arasında değiştiği bulunmuştur. Field (2009) bir ölçekte yer alan maddelerin doğrulanmış madde toplam korelasyon değerlerinin .30 ve üzerinde olması gerektiğini belirtmektedir. Bu bağlamda ölçekte yer alan maddelerin ölçeğe hizmet ettiği ifade edilebilir. Ayrıca tüm puan türlerine göre ölçekte yer alan herhangi bir maddenin çıkarılmasında Cronbach Alfa değerinde bir artış olmadığı görülmektedir. Alanyazında bilimsel yaratıcılığın ölçülmesine yönelik geliştirilen testlerin (Ayas ve Sak, 2014; Sak, 2010; Sak ve Ayas, 2013; Chin ve Siew, 2005; Çeliker ve Balım, 2012; Hu ve Adey, 2002; Siew, Chong ve Chin, 2014; Samsudin vd., 2017; Usta ve Akkanat, 2015) madde analizleri incelendiğinde elde edilen değerlerin ideal olarak kabul edilen değerler arasında olduğu görülmüştür. Sonuç olarak ÇABİYAT'ın güvenirliliği iç tutarlılığını incelemek için gerçekleştirilen analiz bulguları ölçeğin iç tutarlılığının yeterli düzeyde olduğunu ortaya koymuştur.

Puanlayıcılar Arası Güvenirliğe Yönelik Tartışma ve Sonuç

Çoğul düşünme testlerinde karşılaşılan önemli sorunlardan biri puanlamada öznel faktörlerin etkisinin olmasıdır (Kaufman, Plucker ve Baer, 2008). Açık uçlu maddelerin puanlanması puanlayıcının görüş ve inanışlarına göre değişebilir. Bu nedenle puanlayıcıdan kaynaklı hata oranını düşürmek için önlemler alınmalıdır. ÇABİYAT'ta puanlayıcıdan kaynaklı hata oranını önlemek için ölçeğin cevap havuzu oluşturulmuş ve puanlamalar cevap havuzuna göre yapılmıştır. Ön deneme, pilot ve ana uygulamada ulaşılan 942 öğrencinin yanıtlarına göre oluşturulan ÇABİYAT'ın doğru cevap havuzunda 175 doğru yanıt yer almaktadır. Yanıtların temalaştırılması yaratıcılık alanında çalışan bir uzman ve araştırmacı tarafından yapılmıştır. Bir madde için 5, diğer maddelerin her biri için ise 4 tema oluşturulmuştur. ÇABİYAT'ın puanlayıcıdan kaynaklı hata oranının düşük olduğunu ortaya koymak için farklı puanlayıcıların verdiği puanlar karşılaştırılmıştır. ÇABİYAT'ın puanlayıcılar arası güvenirliliğini test etmek için rassal olarak belirlenen farklı sınıf düzeylerindeki 103 öğrencinin yanıtlarını iki farklı puanlayıcı değerlendirmiştir. Puanlayıcılar öğrenci yanıtlarını birbirlerinden bağımsız olarak doğru yanıt havuzuna göre değerlendirmiştir. Puanlamalar arasındaki tutarlılığı incelemek için sınıf içi korelasyon analizinden yararlanılmıştır. Analiz sonucunda sınıf içi korelasyon değerlerinin .911 ile .988 arasında değiştiği, ortalamalarının ise .983 olduğu bulunmuştur. Cicchetti ve Sparrow (1990) puanlayıcılar arası güvenirlilik katsayısının .90 ve üzerinde olması durumunda ölçeğin güvenirliliğinin yüksek olduğunu belirtmektedir. Bu bağlamda ÇABİYAT'ın puanlayıcılar arası güvenirliliğinin yüksek düzeyde olduğu ifade edilebilir.

Alanyazında bilimsel yaratıcılık testlerinin puanlayıcılar arası güvenirliliği incelemek için gerçekleştirilen analizler incelendiğinde bu değerlerin .793 ile .998 arasında değiştiği görülmektedir (Aktamış ve Ergin, 2008; Ayas, 2017; Ayas ve Sak, 2014; Sak, 2010; Sak ve Ayas, 2013; Hu ve Adey, 2002). Bu durum ÇABİYAT'ın puanlayıcılar arası güvenirlilik değerleri ile paralellik göstermektedir. Puanlayıcılar arası güvenirlilik değerlerinin yüksek olmasının nedeni, testlerin alan bilgisine dayalı olması olabilir. ÇABİYAT'ın puanlayıcılar arası güvenirliliğinin yüksek olmasının bir diğer nedeni ise yanıt havuzundaki yanıtların çeşitliliğinin sınırlı olması olabilir. Doğru yanıt havuzunda 175 yanıtın olması puanlamayı kolaylaştırabilir ve puanlayıcıdan kaynaklı ölçüm hata oranını düşürebilir. Puanlayıcılar arası güvenirlilik analizi sonuçları ölçek maddelerinin puanlanmasında öznelliğin en aza indirildiğine dair kanıt oluşturmaktadır. Sonuç olarak iç tutarlılığa ve puanlayıcılar arası güvenirlilik analizleri ölçeğin güvenirliliğinin yeterli düzeyde olduğunu ortaya koymuştur.

kaynakça

- Adamson, L. B., Foster, M. A., Roark, M.L., and Reed, D.B. (1998). Doing a science project: Gender differences during childhood. *Journal of Research in Science Teaching*, 35, 845-857.
- Akbulut, Y. (2010). *Sosyal bilimlerde SPSS uygulamaları*. İstanbul: Pasifik Ofset.
- Aktamış, H., and Ergin, Ö. (2008). The effect of scientific process skills education on students' scientific creativity, science attitudes and academic achievements. *Asia-Pacific Forum on Science Learning and Teaching*, 9(1), 1-21.
- Alencar, E. M. L. S., Fleith, D. S., and Bruno-Faria, M. F. (2014). The measurement of creativity: possibilities and challenges. In E. M. L. S. Alencar, M. F. Bruno-Faria and D.S. Fleith (Eds.), *Theory and practice of creativity measurement* (pp. 1-20). Texas: Prufrock Press Inc.
- Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychology*, 27, 33-51.
- Alisinanoğlu, F., Özbey, S. ve Kahveci, G. (2011). *Okul öncesinde fen eğitimi*. İstanbul: Maya Akademi.
- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5), 997-1013.
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, 45(2), 357-376.
- Amabile, T. M. (1987). The motivation to be creative. In S. G. Isaken (Ed.), *Frontiers of creativity research: Beyond the basics* (pp. 223-254). Buffalo, NY: Bearly Limited.
- Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in Organizational Behavior*, 10(1), 123-167.
- Amabile, T. M. (1996). *Creativity in context*. New York: Springer-Verlag.
- Amabile, T. M. (2012). Componential theory of creativity. *Harvard Business School*, 12(96), 1-10.
- Amabile, T. M., Conti, R., Coon, H., Lazenby, J., and Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal*, 39(5), 1154-1184.
- Amsel, E., and Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11, 523-550.
- Anastasi, A., and Urbina, S. (1997). *Psychological testing*. Upper Saddle River, NJ: Prentice-Hall.
- Astin, H. ve Bayer, A. E. (1979). Pervasive sex differences in the academic reward system: Scholarship, marriage and what else?" D.R. Lewis and W.E. Becker (Ed.), *Academic rewards in higher education* içinde (s.211-229). Cambridge, Mass.: Ballinger Publishing Company.
- Atakaya, M. A. (2018). *Yaraticılığı puanlama endekslerinin yordama gücünün karşılaştırılması üzerine bir araştırma: Bilimsel yaratıcılık örneği* (Yayınlanmamış yüksek lisans tezi). Anadolu Üniversitesi, Eskişehir.
- Ayas M. B., and Sak, U. (2008). Test of scientific creativity: Its development and psychometric properties. *Paper presented at the 4th International Conference on Intelligence and Creativity*, Münster, Germany.
- Ayas, M. B. (2010). *Bilimsel üretkenlik testinin 6. sınıf düzeyinde psikometrik özelliklerinin incelenmesi* (Yayınlanmamış yüksek lisans tezi). Anadolu Üniversitesi, Eskişehir.

- Ayas, M. B. (2017). *Bilimsel üretkenlik testinin 3, 4 ve 5. Sınıf öğrencilerine uygun formunun geliştirilmesi ve ön psikometrik özelliklerinin incelenmesi* (Yayınlanmamış doktora tezi). Anadolu Üniversitesi, Eskişehir.
- Ayas, M. B., and Sak, U. (2014). Objective measure of scientific creativity: Psychometric validity of the Creative Scientific Ability Test. *Thinking Skills and Creativity, 13*, 195-205.
- Ayverdi, L., Asker, E., Özaydın, S. ve Sarıtaş, T. (2012). İlköğretim öğrencilerinin bilimsel yaratıcılıkları ile fen ve teknoloji dersi akademik başarıları arasındaki ilişkinin belirlenmesi. *İlköğretim Online, 11*(3), 646-659.
- Ayyıldız Potur, A., and Barkul, Ö. (2009). Gender and creative thinking in education: A theoretical and experimental overview. *ITU A | Z, 6*(2), 44-57.
- Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4*(1), 23-39.
- Baer, J. (1994a). Divergent thinking is not a general trait: A multidomain training experiment. *Creativity Research Journal, 7*(1), 35-46.
- Baer, J. (1994b). Performance assessments of creativity: Do they have long-term stability? *Roeper Review, 17*(1), 7-11.
- Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal, 11*(2), 173-177.
- Baer, J. (1999). Domains of creativity. In M. Runco and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 591-596). San Diego, CA: Academic Press.
- Baer, J. (2012). Domain specificity and the limits of creativity theory. *The Journal of Creative Behavior, 46*(1), 16-29.
- Baer, J. (2016). *Domain specificity of creativity*. San Diego, CA: Academic Press.
- Baer, J., and Kaufman, J. C. (2008). Gender differences in creativity. *Journal of Creative Behavior, 42*(2), 75-105.
- Baer, J., Kaufman, J. C., and Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity Research Journal, 16*(1), 113-117.
- Baer, J., and McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In C. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65-77). Hershey, PA: IGI Global.
- Baysal, Z. N., Kaya, N. B., ve Üçüncü, G. (2013). İlkokul dördüncü sınıf öğrencilerinde bilimsel yaratıcılık düzeyinin çeşitli değişkenler açısından incelenmesi. *Eğitim Bilimleri Dergisi, 38*, 55-64.
- Bermejo, M. R., Ruiz-Melero, M. J., Esparza, J., Ferrando, M., and Pons, R. (2016). A new measurement of scientific creativity: The study of its psychometric properties. *Anales de Psicología, 32*(3), 652-661.
- Björk-Willén, P., and Aronsson, K. (2014). Preschoolers' "animation" of computer games. *Mind, Culture, and Activity, 21*(4), 318-336.
- Blok, H., Oostdam, R., Otter, M. E., and Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72*(1), 101-130.
- Boden, M. A. (2001). Creativity and knowledge. In A. Craft, B. Jeffrey, and M. Liebling (Eds.), *Creativity in education* (pp. 95-103). London: Continuum.
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. New York: Routledge.

- Böke, K. (2017). Örneklemeye. K. Böke (Ed.), *Sosyal bilimlerde araştırma yöntemleri içinde* (5. baskı, s. 105-147). İstanbul: Alfa Basım Yayım Dağıtım Ltd. Şti.
- Browne, M. W., and Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Thousand Oaks, CA: Sage Publications.
- Bullock, M. (1991). Scientific reasoning in elementary school: Developmental and individual differences. *Paper presented at the Symposium on Scientific Thinking, SRC, Seattle, WA, United States of America.*
- Bülbül, A. H. (2019). *Duygusal tasarımın çoklu ortam uygulamalarında kullanımının bilişsel yüke, motivasyona, konu ilgisine ve öğrenmeye etkisi* (Yayınlanmamış doktora tezi). Anadolu Üniversitesi, Eskişehir.
- Büyüköztürk, Ş. (2014). *Sosyal bilimler için veri analizi el kitabı* (19. baskı). Ankara: Pegem Akademi.
- Campbell, D. T., and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105.
- Carey, S. (2000). Science education as conceptual change. *Journal of Applied Developmental Psychology, 21*(1), 13-19.
- Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25-44.
- Carson, S. H., Peterson, J. B., and Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal, 17*(1), 37-50.
- Chambers, J. A. (1964). Relating personality and biographical factors to scientific creativity. *Psychological Monographs: General and Applied, 58*(4), 1-20.
- Chan, D. W. (2005). Self-perceived creativity, family hardiness, and emotional intelligence of Chinese gifted students in Hong Kong. *Journal of Secondary Gifted Education, 16*, 47-56.
- Chan, D., and Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*(1), 143-159.
- Charyton, C., and Snelbecker, G. E. (2007) General, artistic and scientific creativity attributes of engineering and music students. *Creativity Research Journal, 19*(3), 213-225.
- Chin, M. K., and Siew, N. M. (2015). The development and validation of a figural scientific creativity test for preschool pupils. *Creative Education, 6*(12), 1391-1402.
- Cicchetti, D. V., and Sparrow, S. S. (1990). Assessment of adaptive behavior in young children. In J. J. Johnson, and J. Goldman (Eds.), *Developmental assessment in clinical child psychology: A handbook* (pp. 173-196). New York: Pergamon Press.
- Clark, L. A., and Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.
- Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, L. M., and Ambrose, D. (1999). Adaptation and creativity. In M. Runco and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 9-22). San Diego, CA: Academic Press.
- Cohen, R. J., and Swerdlik, M. (2002). *Psychological testing and assessment: An introduction to test and measurement*. Boston: McGraw-Hill.

- Cole, S. (1979). Age and scientific performance. *American Journal of Sociology*, 84, 958-977.
- Cole, J. R., and Cole, S. (1973). *Social stratification in science*. Chicago: The University of Chicago Press.
- Cole, J. R., and Zuckerman, H. (1987). Marriage, motherhood and research performance in science. *Scientific American*, 25, 119-125.
- Comrey, A. L., and Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.
- Costello, A. B., and Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7), 1-9.
- Craft, A. (2005). *Creativity in schools: Tensions and dilemmas*. New York: Routledge.
- Creswell, J. W. (2014). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Cropley, A. J. (1999). Definitions of creativity. In M. Runco, and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 511-524). San Diego, CA: Academic Press.
- Csikszentmihalyi, M. (1996). *Creativity: Flow and the Psychology of Discovery and Invention*. New York: Harper Collins.
- Csikszentmihalyi, M. (1997). *Flow and the psychology of discovery and invention*. New York: Harper Collins.
- Csikszentmihalyi, M. (1999). Implications of a system perspective for the study of creativity. In R. Sternberg (Ed.), *Handbook of creativity* (pp. 313-328). New York: Cambridge University Press.
- Csikszentmihalyi, M. (2014). Society, culture, and person: A systems view of creativity. In M. Csikszentmihalyi (Ed.), *The systems model of creativity: The collected works of Mihaly Csikszentmihalyi* (pp. 47-61). Dordrecht: Springer.
- Csikszentmihalyi, M., and Sawyer, K. (2014). The social dimension of a solitary moment. In M. Csikszentmihalyi (Ed.), *The systems model of creativity: The collected works of Mihaly Csikszentmihalyi* (pp. 73-98). Dordrecht: Springer.
- Csikszentmihalyi, M., and Wolfe, R. (2014). New conceptions and research approaches to creativity: Implications of a systems perspective for creativity in education. In M. Csikszentmihalyi (Ed.), *The systems model of creativity: The collected works of Mihaly Csikszentmihalyi* (pp. 161-182). Dordrecht: Springer.
- Çeliker, H. D. ve Balım, A. (2012). Bilimsel yaratıcılık ölçeğinin Türkçeye uyarlama süreci ve değerlendirme ölçütleri. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 5(2), 1-21.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Ankara: Pegem Akademi.
- Dacey, J. (1999). Concepts of creativity: A history. In M. Runco and S. Pritzker (Eds.), *Encyclopedia of creativity* (Vol: 1-1; pp. 309-322). San Diego, CA: Academic Press.
- Dancy, M. H., and Beichner, R. (2006). Impact of animation on assessment of conceptual understanding in physics. *Physical Review Special Topics-Physics Education Research*, 2(1), 0101041-0101047.

- Davey, T., Pitoniak, M. J., and Slater, S. C. (2015). Designing computerized adaptive tests. In S. Lane, M. R. Raymond, and T. M. Haladyna (Eds.), *Handbook of test development* (pp. 19-34). New York: Routledge.
- Davis, G. A., and Belcher, T. L. (1971). How shall creativity be measured? Torrance tests, RAT, alpha biographical, and IQ. *Third Quarter*, 5(3), 153-161.
- Davis, J. (1993). *Drawing's demise: U-shaped development in graphic symbolization*. Unpublished manuscript, Harvard Project Zero, Harvard University, Cambridge, MA.
- De Vries, H. B., and Lubart, T. I. (2019). Scientific creativity: Divergent and convergent thinking and the impact of culture. *The Journal of Creative Behavior*, 53(2), 145-155.
- Dellas, M., and Gaier, E. L. (1970). Identification of creativity: The individual. *Psychological Bulletin*, 73(1), 55-73.
- Demir, S., and Şahin, F. (2014). Assessment of open-ended questions directed to prospective science teachers in terms of scientific creativity. *Procedia-Social and Behavioral Sciences*, 152, 692-697.
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Downing, S. M. (2006). Twelve steps for effective test development. In S. M. Downing and T.M. Haladyna (Eds.), *Handbook of test development* (pp. 3-25). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Dunbar, K. (1997). How scientists think: Online creativity and conceptual change in science. In T. Ward, S. Smith, and S. Vaid (Eds.), *Conceptual structures and processes: Emergence, discovery, and change* (pp. 461-492). Washington: American Psychological Association Press.
- Dunteman, G. H. (1989). *Principal component analysis: Quantitative applications in the social sciences series* (Vol. 69). Thousand Oaks, CA: Sage Publications.
- Eichenberger, R. J. (1978). Creativity measurement through use of judgment criteria in physics. *Educational and Psychological Measurement*, 38(2), 421-427.
- Erickson, G. A., and Erickson, L.J. (1984) Achievement: Evidence, explanations, and implications. *Science Education*, 68, 63-89.
- Erickson, G. A., and Farkas, S. (1991). Prior experience and gender differences in science achievement. *The Alberta Journal of Educational Research*, 37, 225-239.
- Erkuş, A. (2016). *Psikolojide ölçme ve ölçek geliştirme-I* (3.baskı). Ankara: Pegem Akademi.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-279.
- Feldhusen, J. F. (2002). Giftedness and creativity. In M. Runco and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 773-777). San Diego, CA: Academic Press.
- Feldhusen, J. F. (1995) Creativity: A knowledge base, metacognitive skill, and personality factor. *Journal of Creative Behavior*, 29, 255-268.
- Feldhusen, J. F. (2002). Creativity: the knowledge base and children. *High Ability Studies*, 13(2), 179-183.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Findlay, C. S., and Lumsden, C. J. (1988). The creative mind: Toward an evolutionary theory of discovery and innovation. *Journal of Social and Biological Structures*, 11, 3-55.

- Fishkin, A. S., and Johnson, A. S. (1998). Who is creative? Identifying children's creative abilities. *Roepfer Review*, 21(1), 40-46.
- Ford, L., and Dahinten, V. S. (2005). Use of intelligence tests in the assessment of preschoolers. In D. P. Flanagan and P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 487-503). New York: The Guilford Press.
- Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18 (1), 39-50.
- Fox, M. F. (2005). Gender, family characteristics, and publication productivity among scientists. *Social Studies of Science*, 35, 131-150.
- Frederiksen, N., and Ward, W. C. (1978). Measures for the study of creativity in scientific problem-solving. *Applied Psychological Measurement*, 2(1), 1-24.
- Frensch, P. A., and Sternberg, R. J. (1989). Expertise and intelligent thinking: When is it worse to know better. *Advances in the Psychology of Human Intelligence*, 5, 157-188.
- Friedlander, M. (1983). A natural science creativity test as a prediction of creative thinking in science [Abstract]. *Creative Child & Adult Quarterly*, 8(4), 211-215.
- Gajda, A., Karwowski, M., and Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*, 109(2), 269-299.
- Gardner, H. (2011). *Creating minds: An anatomy of creativity seen through the lives of Freud, Einstein, Picasso, Stravinsky, Eliot, Graham, and Ghandi*. New York: Basic Books.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170.
- Gero, J. S., and Maher, M. L. (2013). Introduction. In J. S. Gero, and M. L. Maher (Eds.). *Modeling creativity and knowledge-based creative design* (pp. 1-6). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., Publisher.
- Gick, M. L., and Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15(1), 1-38.
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39(2), 93-104.
- Gruber, H. E., and Davis, S. N. (1988). Inching our way up Mount Olympus: The evolving-systems approach to creative thinking. In R. Sternberg (Eds.), *The nature of creativity: Contemporary psychological perspectives* (pp. 243-270). New York: Cambridge University Press.
- Guilford, J. P. (1950). Creativity. *American Psychologist*, 5, 444-454.
- Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, 53(4), 267-293.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guilford, J. P. (1968). *Intelligence, creativity, and their educational implications*. San Diego: Robert R. Knapp.
- Gullo, D. F. (2005). *Understanding assessment and evaluation in early childhood education* (2nd ed.). New York: Teachers College Press.
- Güler, N. ve Taşdelen Teker G. T. (2015). Açık uçlu maddelerde farklı yaklaşımlarla elde edilen puanlayıcılar arası güvenilirliğin değerlendirilmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 12-24.

- Hadzigeorgiou, Y., Fokialis, P., and Kabouropoulou, M. (2012). Thinking about creativity in science education. *Creative Education*, 3(5), 603-611.
- Hambleton, R. K., and De Jong, J. H. A. L. (2003). Advances in translating and adapting educational and psychological tests. *Language Testing*, 20, 127-134.
- Han, K. S. (2000). *Varieties of creativity: Investigating the domain-specificity of creativity in young children* (Doctoral dissertation). Lincoln: The University of Nebraska. Retrieved from ProQuest (PQ number: 9973594).
- Harkness, J. A., Villar, A., and Edwards, B. (2010). Translation, adaptation, and design. In Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L. E., Mohler, P. P., Pennell, B. E., and Smith, T. W. (Eds.), *Survey methods in multinational, multiregional, and multicultural contexts* (pp. 115-140). Hoboken, NJ: John Wiley & Sons.
- Hayes, J. R. (1989). Cognitive processes in creativity. In J. A. Glover, R. R. Ronning, and C.R. Reynolds (Eds.), *Handbook of creativity* (pp. 135-145). Boston, MA: Springer.
- Heller, K. A. (2007). Scientific ability and creativity. *High Ability Studies*, 18(2), 209-234.
- Hennessey, B. A., and Amabile, T. M. (1999). Consensual assessment. In M. A. Runco, and S. R. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 347-359). San Diego, CA: Academic Press.
- Hitchcock, C. H., and Noonan, M. J. (2000). Computer-assisted instruction of early academic skills. *Topics in Early Childhood Special Education*, 20(3), 145-158.
- Hocevar, D. (1979). A comparison of statistical infrequency and subjective judgment as criteria in the measurement of originality. *Journal of Personality Assessment*, 43(3), 297-299.
- Hocevar, D. (1980). Intelligence, divergent thinking, and creativity. *Intelligence*, 4(1), 25-40.
- Hocevar, D., and Bachelor, P. (1989). A taxonomy and critique of measurements used in the study of creativity. In J. A. Glover, R. R. Ronning, and C. R. Reynolds (Eds.), *Handbook of creativity* (pp. 53-75). Boston, MA: Springer.
- Hong, E., and Milgram, R. M. (2010). Creative thinking ability: Domain generality and specificity. *Creativity Research Journal*, 22(3), 272-287.
- Houtz, J. C., and Krug, D. (1995). Assessment of creativity: Resolving a mid-life crisis. *Educational Psychology Review*, 7(3), 269-300.
- Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Hu, W., and Adey, P. A. (2002). Scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389-403.
- Hu, W., Shi, Q. Z., Han, Q., Wang, X., and Adey, P. (2010). Creative scientific problem finding and its developmental trend. *Creativity Research Journal*, 22, 46-52.
- Huck, S. W. (2012). *Reading statistics and research* (6th ed.). Boston: Pearson.
- Hutcheson, G. D., and Sofroniou, N. (1999). *The multivariate social scientist: Introductory statistics using generalized linear models*. Thousand Oaks, CA: Sage Publications.
- Ivcevic, Z. (2007). Artistic and everyday creativity: An act-frequency approach. *The Journal of Creative Behavior*, 41(4), 271-290.
- Jang, Y., and Ko, Y. (2017). Sources of scientific creativity: participant observation of a public research institute in Korea. *Journal of Open Innovation: Technology, Market, and Complexity*, 3(1), 1-16.

- Johnson, S. (1987). Gender differences in science: Parallels in interest, experience, and performance. *International Journal of Science Education*, 9, 467-481.
- Johnsen, S. K., and Corn, A. L. (2001). *Screening Assessment for Gifted Elementary and Middle School Students: Examiner's Manual*. Texas: PRO-ED, Inc.
- Johnson, R. B., and Christensen, L. (2014). *Educational research: Quantitative, qualitative, and mixed approaches* (5th ed.). Thousand Oaks, CA: Sage Publications.
- Kang, D., Park, J., and Hong, H. (2015). Changes in the number of ideas depending on time when conducting Scientific Creativity activities. *Journal of Baltic Science Education*, 14(4), 448-459.
- Kanlı, E. (2014a). *Yaratıcı bilimsel çağrışımlar testinin geliştirilmesi ve psikometrik özelliklerinin incelenmesi* (Yayınlanmamış doktora tezi). İstanbul Üniversitesi, İstanbul.
- Kanlı, E. (2014b). The associative basis of scientific creativity: A model proposal. *Turkish Journal of Giftedness & Education*, 4(1), 37-50.
- Kanlı, E. (2017). Üstün yetenekli öğrencilerin bilimsel yaratıcılık düzeyleri, cinsiyet ve bilimsel tutumları arasındaki ilişkilerin incelenmesi. *İlköğretim Online*, 16(4), 1792-1802.
- Kaplan, C. A., and Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, 22(3), 374-419.
- Kaplan, R., and Saccuzzo, D. (2012). *Psychological testing: Principles, applications, and issues*. Belmont, CA: Cengage Learning.
- Karademir, E. (2016). Investigation the scientific creativity of gifted students through project-based activities. *International Journal of Research in Education and Science*, 2(2), 416-427.
- Kaufman, J. C., and Baer, J. (2005). *Creativity across domains: Faces of the muse*. Mahwah, NJ: Lawrence Erlbaum.
- Kaufman, J. C., and Baer, J. (2009). Is one dimension enough? A response to Simonton's varieties of (scientific) creativity (2009). *Perspectives on Psychological Science*, 4(5), 453-454.
- Kaufman, J. C., Baer, J., Cole, J. C., and Sexton, J. D. (2008). A comparison of expert and nonexpert raters using the consensual assessment technique. *Creativity Research Journal*, 20(2), 171-178.
- Kaufman, J. C., Baer, J., and Gentile, C. A., (2004). Differences in gender and ethnicity as measured by ratings of three writing tasks. *Journal of Creative Behavior*, 39, 56-69.
- Kaufman, J. C., Plucker, J. A., and Baer, J. (2008). *Essentials of creativity assessment*. Hoboken, NJ: John Wiley & Sons.
- Ke, F., Lin Kun Shan, H., Ching, Y. H., and Dwyer, F. (2006). Effects of animation on multi-level learning outcomes for learners with different characteristics: A meta-analytic assessment and interpretation. *Journal of Visual Literacy*, 26(1), 15-40.
- Kelly, A. (1988). Sex stereotypes and school science. A three-year follow-up. *Educational Studies*, 14, 151-163.
- Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes*. Cambridge: The MIT Press.
- Klahr, D., and Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1-48.
- Klahr, D., Fay, A. L., and Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology*, 25(1), 111-146.

- Kline, P. (1999). *The handbook of psychological testing* (2nd ed.). London: Routledge.
- Kline, R. B. (1994). *An easy guide to factor analysis*. New York: Routledge.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford publications.
- Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. Schatschneider, and D. L. Compton (Eds.), *Applied quantitative analysis education and the social sciences* (pp. 171–207). New York: Routledge.
- Kuhn, D., Amsel, E., O’Loughlin, M., Schauble, L., Leadbeater, B., and Yotive, W. (1988). *The development of scientific thinking skills*. San Diego, CA: Academic Press.
- Kulkarni, D., and Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science*, 12(2), 139-175.
- Küçüküran, G. (2003). Okulöncesi fen öğretiminde bir teknik: Anoloji. *Milli Eğitim Dergisi*, 157, 16-21.
- Laius, A., and Rannikmäe, M. (2011). Impact on student change in scientific creativity and socio-scientific reasoning skills from teacher collaboration and gains from professional in-service. *Journal of Baltic Science Education*, 10(2), 127-137.
- Lane, S., Raymond, M. R., Haladyna, T. M., and Downing, S. M. (2015). Test development process. In S. Lane, M. R. Raymond, & T. M. Haladyna (Eds), *Handbook of test development* (pp. 19-34). New York: Routledge.
- Lee, K. H. (2002). Creative thinking in real world situations in relation to gender and education of late adolescents. *Korean Journal of Thinking and Problem Solving*, 12, 59-70.
- Liang, J. C. (2002). *Exploring scientific creativity of eleventh-grade students in Taiwan* (Doctoral dissertation). Austin, TX: The University of Texas. Retrieved from ProQuest (PQ number: 3099482).
- Lin, C., Hu, W., Adey, P., and Shen, J. (2003). The influence of CASE on scientific creativity. *Research in Science Education*, 33(2), 143-162.
- Long, J., and Fox, E. (1995). Scientific careers: Universalism and particularism. *Annual Review of Sociology*, 21, 45-71.
- Lubart, T. I. (1994). Creativity. In R. J. Sternberg (Ed.), *Thinking and problem solving* (pp. 289-332). London: Academic Press.
- Lubart, T. I. (1999). Componential models. In M. Runco, and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 295-300). San Diego, CA: Academic Press.
- Lubart, T., Zenasni, F., and Barbot, B. (2013). Creative potential and its measurement. *International Journal of Talent Development and Creativity*, 1(2), 41-51.
- Maccoby, E. E. (1966). Sex differences in intellectual functioning. In E. E. Maccoby (Ed), *The development of sex differences* (pp.25-55). California: Stanford University Press.
- Majumdar, S. K. (1975). A systems approach to identification and nurture of scientific creativity [Abstract]. *Journal of Indian Education*, 1(2), 17-23.
- Maker, C. J. (1993). Creativity, intelligence, and problem solving: A definition and design for cross-cultural research and measurement related to giftedness. *Gifted Education International*, 9(2), 68-77.
- Matud, M. P., Rodrigues, C., and Grande, J. (2007). Gender differences in creative thinking. *Personality and Individual Differences* 43, 1137–1147.

- Mattern, N., and Schau, C. (2002). Gender differences in science attitude-achievement relationships over time among white middle- school students. *Journal of Research in Science Teaching*, 39, 324- 340.
- Mellenbergh, G. J. (1989). Item bias and item response theory. *International journal of educational research*, 13(2), 127-143.
- Michael, W. B. (1999). Guilford's view. In M. Runco and S. Pritzker (Eds.), *Encyclopedia of creativity* (pp. 785-797). San Diego, CA: Academic Press.
- Milfont, T. L., and Duckitt, J. (2004). The structure of environmental attitudes: A first-and second-order confirmatory factor analysis. *Journal of Environmental Psychology*, 24(3), 289-303.
- Milgram, R. M., and Milgram, N. A. (1976). Creative thinking and creative performance in Israeli students. *Journal of Educational Psychology*, 68(3), 255-259.
- Mohamed, A. H. (2006). *Investigating the scientific creativity of fifth-grade students* (Doctoral dissertation). Tucson, AZ: University of Arizona. Retrieved from ProQuest (PQ number: 3207635).
- Moran, J. D., Milgram, R. M., Sawyers, J. K., and Fu, V. R. (1983). Original thinking in preschool children. *Child Development*, 22(4), 921-926.
- Moravcsik, M. J. (1981). Creativity in science education. *Science Education*, 65(2), 221-227.
- Morris, T. R., Cho, C., Dilda, V., Shine, J. M., Naismith, S. L., Lewis, S. J., and Moore, S. T. (2013). Clinical assessment of freezing of gait in Parkinson's disease from computer-generated animation. *Gait & Posture*, 38(2), 326-329.
- Mukhopadhyay, R. (2013). Measurement of creativity in physics-a brief review on related tools. *Journal of Humanities and Social Science*, 6(5), 45-50.
- Nitko, A. J. (2004). *Educational assessment of students* (4th ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Nowell, A., and Hedges, L. V. (1998). Trends in gender differences in academic achievement from 1960 to 1994: An analysis of differences in mean, variance, and extreme scores. *Sex Roles*, 39, 21-43.
- Öner, N. (1997). *Türkiye'de kullanılan psikolojik testler: Bir başvuru kaynağı* (3. baskı). İstanbul: Boğaziçi Üniversitesi Yayınları.
- Özdamar, K. (2004). *Paket programlarla istatistiksel veri analizi*. Eskişehir: Kaan Kitabevi.
- Özdemir, N. N. (2013). *ÜYEP'e başvuran öğrencilerin bilimsel yaratıcılık bileşenlerindeki cinsiyet farklılıklarının incelenmesi* (Yayınlanmamış yüksek lisans tezi). Anadolu Üniversitesi, Eskişehir.
- Özdemir, N. N., and Sak, U. (2013). A componential analysis of gender differences in scientific creativity. *Turkish Journal of Giftedness & Education*, 3(2), 53-65.
- Özdemir, G., and Dikici, A. (2017). Relationships between scientific process skills and scientific creativity: Mediating role of nature of science knowledge. *Journal of Education in Science, Environment and Health*, 3(1), 52-68.
- Özdemir, Ş. A. ve Macaroğlu, E. (2000). İlköğretim matematik ve fen bilgisi öğretmen adaylarının bilimsel okur-yazarlık seviyelerinin tespiti, IX. *Ulusal Eğitim Bilimleri Kongresi. Erzurum: Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi*, 558-564.
- Pallant, J. (2005). *SPSS survival manual* (12th ed.). Australia: Allen & Unwin.

- Pariser, D., and Van Den Berg, A. (1997). The mind of the beholder: Some provisional doubts about the U-curved aesthetic development thesis. *Studies in Art Education*, 38(3), 158-178.
- Parshall, C. G., Spray, J. A., Kalohn, J. C., and Davey, T. (2002). *Practical considerations in computer-based testing*. New York: Springer-Verlag.
- Perry, A., and Karpova, E. (2017). Efficacy of teaching creative thinking skills: A comparison of multiple creativity assessments. *Thinking Skills and Creativity*, 24, 118-126.
- Pestonjee, D. M. (1997). *Third handbook of psychological and social instruments* (Vol. 1). New Delhi: Concept Publishing Company.
- Pfeiffer, S. I., Petscher, Y., and Jarosewich, T. (2007). The gifted rating scales-preschool/ kindergarten form: An analysis of the standardization sample based on age, gender, and race. *Roeper Review*, 29(3), 206-211.
- Piirto, J. M. (1999). A survey of psychological studies in creativity. In A. S. Fishkin, B. Cramond, and P. Olszewski-Kubilius (Eds.), *Investigating creativity in youth: Research and methods* (pp. 27-48). Cresskill, NJ: Hampton Press.
- Plucker, J. A. (1998). Beware of simple conclusions: The case for content generality of creativity. *Creativity Research Journal*, 11(2), 179-182.
- Plucker, J. A., Beghetto, R. A., and Dow, G. (2004). Why isn't creativity more important to educational psychologist? Potential, pitfalls, and future directions in creativity research. *Educational Psychologists*, 39(2), 83-96.
- Plucker, J. A., Callahan, C. M., and Tomchin, E. M. (1996). Wherefore art thou, multiple intelligences? Alternative assessments for identifying talent in ethnically diverse and low income students. *Gifted Child Quarterly*, 40(2), 81-91.
- Plucker, J. A., and Runco, M. A. (1998). The death of creativity measurement has been greatly exaggerated: Current issues, recent advances, and future directions in creativity assessment. *Roeper Review*, 21(1), 36-39.
- Plucker, J., and Zabelina, D. (2009). Creativity and interdisciplinarity: One creativity or many creativities? *ZDM Mathematics Education*, 41(1-2), 5-11.
- Popp, E. C., Tuzinski, K., and Fetzer, M. (2016). Actor or avatar? Considerations in selecting appropriate formats for assessment content. In F. Drasgow (Ed.), *Technology and testing: Improving educational and psychological measurement* (pp. 79-109). New York: Taylor & Francis.
- Puolakanaho, A., Poikkeus, A. M., Ahonen, T., Tolvanen, A., and Lyytinen, H. (2003). Assessment of three-and-a-half-year-old children's emerging phonological awareness in a computer animation context. *Journal of Learning Disabilities*, 36(5), 416-423.
- Rachelson, S. (1977). A question of balance: A wholistic view of scientific inquiry. *Science Education*, 61(1), 109-117.
- Renzulli, J. S., and Hartman, R. K. (1971). Scale for rating behavioral characteristics of superior students. *Exceptional children*, 38(3), 243-248.
- Renzulli, J. S., Smith, L. H., White, A. J., Callahan, C. M., and Hartman, R. K. (1976). *Scales for rating the behavioral characteristics of superior students*. Mansfield Center, CT: Creative Learning Press.
- Reskin, B. F. (1978). Scientific productivity, sex, and location in the institution of science. *American Journal of Sociology*, 83, 1235-1243.
- Rhodes, M. A. (1961). An analysis of creativity. *The Phi Delta Kappan*, 42(7), 305-310.

- Rocke, A. J. (2010). *Image and reality: Kekulé, Kopp, and the scientific imagination*. Chicago: University of Chicago Press.
- Rosen, Y. (2009). The effects of an animation-based on-line learning environment on transfer of knowledge and on motivation for science and technology learning. *Journal of Educational Computing Research*, 40(4), 451-467.
- Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16(3), 371-416.
- Rothenberg, A. (1995). Creative cognitive processes in Kekulé's discovery of the structure of the benzene molecule. *The American Journal of Psychology*, 108(3), 419-438.
- Ruf, D. L. (2005). *Losing our minds: Gifted children left behind*. Scottsdale, AZ: Great Potential Press.
- Runco, M. A. (1987). The generality of creative performance in gifted and nongifted children. *Gifted Child Quarterly*, 31(3), 121-125.
- Runco, M. A. (1989). The creativity of children's art. *Child Study Journal*, 19, 177-189.
- Runco, M. A. (1991). The evaluative, valuative, and divergent thinking of children. *Journal of Creative Behavior*, 25, 311-319.
- Runco, M. A. (1992). Children's divergent thinking and creative ideation. *Developmental Review*, 12(3), 233-264.
- Runco, M. A. (1999). Divergent thinking. In M. A. Runco, and S. Pritzker (Eds.), *Encyclopedia of creativity* (Vol. I, pp. 577-582). San Diego: Academic Press.
- Runco, M. A., and Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96.
- Runco, M. A., and Mraz, W. (1992). Scoring divergent thinking tests using total ideational output and a creativity index. *Educational and Psychological measurement*, 52(1), 213-221.
- Sak, U. (2010). Objective measure of creativity in identifying gifted students. *Paper presented at the 12th International ECHA Conference, Paris, France*.
- Sak, U. (2014). *Yaratıcılık gelişimi ve geliştirilmesi*. Ankara: Vize Yayıncılık.
- Sak, U. ve Ayas, M. B. (2009). Bilimsel üretkenlik testi: Teorik alt yapısı, geliştirilme süreci ve psikometrik özellikleri. *Türkiye Üstün Yetenekli Çocuklar II. Ulusal Kongresi'nde sunulan bildiri*, Eskişehir, Türkiye.
- Sak, U., and Ayas, M. B. (2013). Creative Scientific Ability Test (C-SAT): A new measure of scientific creativity. *Psychological Test and Assessment Modeling*, 55(3), 316.
- Sak, U., Bal-Sezerel, B., Dülger, E., Sözel, K., and Ayas, M. B. (2019). Validity of the Anadolu-Sak Intelligence Scale in the identification of gifted students. *Psychological Test and Assessment Modeling*, 61(3), 263-283.
- Sak, U., and Maker, C. J. (2005). Divergence and convergence of mental forces of children in open and closed mathematical problems. *International Education Journal*, 6(2), 252-260.
- Sax, L. J., Hagedorn, L. S., Arredondo, M., and Dicrisi, F. A. (2002). Faculty research productivity: Exploring the role of gender and family-related factors. *Research in Higher Education*, 43, 423-445.
- Samsudin, A., Setyadin, A. H., Suhendi, E., Chandra, D. T., and Siahaan, P. (2018). Seventh grade students' scientific creativity test: A preliminary-study on earth science context. In Ade Gafar Abdullah, Asep Bayu Dani Nandiyanto, and Isma Widiaty (Eds.), *The*

- 2nd Annual Applied Science and Engineering Conference (Vol. 288, pp. 1-5). Bandung, Indonesia: Universitas Pendidikan Indonesia. doi: 10.1088/1757-899X/288/1/012012.
- Sansanwal, D. N., and Sharma, D. (1993). Scientific creativity as a function of intelligence, self-confidence, sex and standard. *Indian Journal of Psychometry and Education*, 24(1), 37-44.
- Sawyer, R. K. (2012). *Explaining creativity: The science of human innovation* (2nd ed.). New York: Oxford University Press.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32(1), 102-119.
- Schunn, C. D., and Dunbar, K. (1996). Priming, analogy, and awareness in complex reasoning. *Memory & Cognition*, 24(3), 271-284.
- Seçer, İ. (2015). *Psikolojik test geliştirme ve uyarlama süreci: SPSS ve Lisrel Uygulamaları*. Ankara: Anı Yayıncılık.
- Shafiq, M. N. (2013). Gender gaps in mathematics, science and reading achievements in Muslim countries: a quantile regression approach. *Education Economics*, 21(4), 343-359.
- Sharma, A., and Mahrshi, M. V. (2017). Construction and standardization of scientific creativity test. *International Journal of Engineering Technology, Management and Applied Sciences*, 5(6), 380-387.
- Sharma, N. (2015). Scientific creativity in relation to cognitive style and achievement in Science of secondary school students. *Educational Quest*, 6(1), 25-29.
- Shepard, L. A., Camilli, G., & Williams, D. M. (1985). Validity of approximation techniques for detecting item bias. *Journal of Educational Measurement*, 22(2), 77-105.
- Shrager, J. (1987). Theory change via view application in instructionless learning. *Machine Learning*, 2(3), 247-276.
- Shukla, J. P., and Sharma, V. P. (1986). Sex differences in scientific creativity [Abstract]. *Indian Psychological Review*, 30(3), 32-35.
- Siew, N. M., Chong, C. L., and Chin, K. O. (2014). Developing a scientific creativity test for fifth graders. *Problems of Education in the 21st Century*, 62, 109-123.
- Silvia, P. J., Wigert, B., Reiter-Palmon, R., and Kaufman, J. C. (2012). Assessing creativity with self-report scales: A review and empirical evaluation. *Psychology of Aesthetics, Creativity, and the Arts*, 6(1), 19-34.
- Sinha, A. K., and Singh, C. (1987). Measurement of scientific creativity [Abstract]. *Indian Journal of Psychometry & Education*, 18(1), 1-13.
- Simon, H. A. (1977). *Models of discovery: And other topics in the methods of science*. Dordrecht: Reidel Publishing Company.
- Simon, H. A. (1989). The scientist as problem solver. In D. Klahr, and K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 375-398). Hillsdale, NJ: Erlbaum.
- Simonton, D. K. (1988). *Scientific genius: A psychology of science*. New York: Cambridge University Press.
- Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. New York: Cambridge University Press.

- Simonton, D.K. (2009). Varieties of (scientific) creativity: A hierarchical model of domain-specific disposition, development, and achievement. *Perspectives on Psychological Science*, 4(5), 441-452.
- Smith, G. J. W., and Carlsson, I. (1983). Creativity in early and middle school years. *International Journal of Behavioral Development*, 6, 167-195.
- Smith, G. J. W., Carlsson, I., and Danielsson, A. (1985). Identification with another person: Manipulated by means of subliminal stimulation. *Scandinavian Journal of Psychology*, 26(1), 74-87.
- Smolucha, L.W., and Smolucha, F. C. (1985). A fifth Piagetian stage: The collaboration between analogical and logical thinking in artistic creativity. *Visual Arts Research*, 11, 90-99.
- Snyder, A., Mitchell, J., Bossomaier, T., and Pallier, G. (2004). The creativity quotient: An objective scoring of ideational fluency. *Creativity Research Journal*, 16(4), 415-420.
- Sodian, B., Zaitchik, D., and Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, 62(4), 753-766.
- Spearman, C. (1930). *Creative mind*. London: Nisbet & Co., Ltd.
- Starko, A. J. (2017). *Creativity in the classroom: Schools of curious delight*. New York: Routledge.
- Starkweather, E. K. (1964). Problems in the measurement of creativity in preschool children. *Journal of Educational Measurement*, 1(2), 109-113.
- Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. New York: Cambridge University Press.
- Sternberg, R. J. (2010). Teaching for creativity. In R. A. Beghetto, and J. C. Kaufman (Eds.), *Nurturing creativity in the classroom* (pp. 394-414). New York: Cambridge University Press.
- Sternberg, R. J., and Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development*, 34, 1-31.
- Sternberg, R. J., and Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.
- Sternberg, R. J., and Lubart, T. I. (1996). Investing in creativity. *American Psychologist*, 51(7), 677-688.
- Stocking, M. L. (1994). *Three practical issues for modern adaptive testing item pools*. Princeton, NJ: Educational Testing Service.
- Stoltzfus, G., Nibbelink, B. L., Vredenburg, D., and Hyrum, E. (2011). Gender, gender role, and creativity. *Social Behavior and Personality: An International Journal*, 39(3), 425-432.
- Stumpf, H. (1995). Scientific creativity: A short overview. *Educational Psychology Review*, 7(3), 225-241.
- Swanston, M., Abraham, C., Macrae, W. A., Walker, A., Rushmer, R., Elder, L., and Methven, H. (1993). Pain assessment with interactive computer animation. *Pain*, 53(3), 347-351.
- Şahin Zeteroğlu, E. (2010). Eğitimde yaratıcılığı ölçme. E Çelebi Öncü (Ed.), *Erken çocukluk döneminde yaratıcılık ve geliştirilmesi* içinde (s. 100-133). Ankara: Pegem Akademi.

- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tabachnick, B. G., and Fidell, L. S. (2007). *Using multivariate statistics* (3rd ed.). Boston, MA: Allyn & Bacon. Inc.
- Tanaka, J. S. (1987). "How big is big enough?" Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 58(1), 134-146.
- Taylor, C. W., and Ellison, R. L. (1967). Biographical predictors of scientific performance. *Science*, 155(3766), 1075-1080.
- Tegano, D. W., Moran, J. D., and Godwin, L. J. (1986). Cross-validation of two creativity tests designed for preschool children. *Early Childhood Research Quarterly*, 1(4), 387-396.
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Herron, J., and Lindamood, P. (2010). Computer-assisted instruction to prevent early reading difficulties in students at risk for dyslexia: Outcomes from two instructional approaches. *Annals of dyslexia*, 60(1), 40-56.
- Torrance, E. P. (1967). The Minnesota studies of creative behavior: National and international extensions. *The Journal of Creative Behavior*, 1(2), 137-154.
- Torrance, E. P. (1972). Predictive validity of the Torrance tests of creative thinking. *The Journal of Creative Behavior*, 6(4), 236-262.
- Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. Sternberg (Eds.), *The nature of creativity: Contemporary psychological perspectives* (pp. 43-75). New York: Cambridge University Press.
- Torrance, E. P. (1998). *Torrance tests of creative thinking: Norms-technical manual figural (streamlined) forms A & B*. Illinois: Scholastic Testing Service, Inc.
- Travers, J. C. (2010). *Emergent literacy skills of young children with autism: A comparison of teacher-led and computer-assisted instruction* (Unpublished doctoral dissertation). University of Nevada, Las Vegas.
- Treffinger, D. J. (1986). Research on creativity. *Gifted Child Quarterly*, 30(1), 15-19.
- Trochim, W. M., and Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati, OH: Atomic Dog.
- Urban, K. K. (1991). On the development of creativity in children. *Creativity Research Journal*, 4(2), 177-191.
- Urbina, S. (2014). *Essentials of psychological testing*. Hoboken, NJ: John Wiley & Sons.
- Usta, E., and Akkanat, Ç. (2015). Investigating scientific creativity level of seventh grade students. *Procedia-Social and Behavioral Sciences*, 191, 1408-1415.
- Warren, F., Mason-Apps, E., Hoskins, S., Azmi, Z., and Boyce, J. (2018). The role of implicit theories, age, and gender in the creative performance of children and adults. *Thinking Skills and Creativity*, 28, 98-109.
- Watkins, M. W. (2000). Monte Carlo PCA for parallel analysis, State College. PA: Ed & Psych Associates [Computer software]. Unpublished Instrument. Retrieved from <http://www.softpedia.com/get/Tools/Home-Education/Monte-Carlo-PCA-for-Parallel-Analysis.shtml>
- Weisberg, R. W. (2009). Creativity and knowledge: A challenge to theories. In R. Sternberg (Ed.), *Handbook of creativity* (pp. 226-250). New York: Cambridge University Press.

- Woodman, R. W., and Schoenfeldt, L. F. (1990). An interactionist model of creative behavior. *The Journal of Creative Behavior*, 24(4), 279-290.
- Wortham, S. C. (2008). *Assessment in early childhood education* (5th ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Wortham, S. C. (2010). *Early childhood curriculum: Developmental bases for learning and teaching* (5th ed.). Boston: Pearson.
- Wortham, S. C., and Hardin, B. J. (2016). *Assessment in early childhood education* (5th ed.). Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Worthington, R. L., and Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838.
- Yang, K. K., Hong, Z. R., Lee, L., and Lin, H. S. (2019). Exploring the significant predictors of convergent and divergent scientific creativities. *Thinking Skills and Creativity*, 31, 252-261.
- Yang, K. K., Lin, S. F., Hong, Z. R., and Lin, H. S. (2016). Exploring the assessment of and relationship between elementary students' scientific creativity and science inquiry. *Creativity Research Journal*, 28(1), 16-23.
- Zachopoulou, E., Makri, A., and Pollatou, E. (2009). Evaluation of children's creativity: Psychometric properties of Torrance's 'Thinking Creatively in Action and Movement' test. *Early Child Development and Care*, 179(3), 317-328.
- Zhu, W., Shang, S., Jiang, W., Pei, M., and Su, Y. (2019). Convergent thinking moderates the relationship between divergent thinking and scientific creativity. *Creativity Research Journal*, 0(0), 1-9.
- Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review*, 20(1), 99-149.
- Zwick, W. R., and Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432-442.