



**BİLİMSEL ÜRETKENLİK TESTİNİN
3, 4 ve 5. SINIF ÖĞRENCİLERİNE UYGUN
FORMUNUN GELİŞTİRİLMESİ VE ÖN PSİKOMETRİK
ÖZELLİKLERİNİN İNCELENMESİ**

Doktora Tezi

M. Bahadır AYAS

Eskişehir 2017

**BİLİMSEL ÜRETKENLİK TESTİNİN 3, 4 ve 5. SINIF ÖĞRENCİLERİNE
UYGUN FORMUNUN GELİŞTİRİLMESİ VE ÖN PSİKOMETRİK
ÖZELLİKLERİNİN İNCELENMESİ**

M. Bahadır AYAS

DOKTORA TEZİ

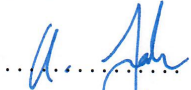
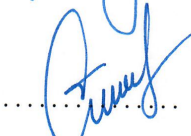

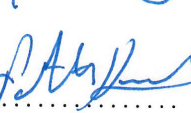

**Özel Eğitim Anabilim Dalı
Üstün Zekalıların Eğitimi Programı
Danışman: Prof. Dr. Uğur SAK**


**Eskişehir
Anadolu Üniversitesi
Eğitim Bilimleri Enstitüsü
Eylül 2017**

Bu tez çalışması BAP Komisyonunca kabul edilen 1508E578 no.lu proje kapsamında desteklenmiştir.

JÜRİ VE ENSTİTÜ ONAYI

M.Bahadır AYAS'ın "Bilimsel Üretkenlik Testinin 3, 4 ve 5. Sınıf Öğrencilerine Uygun Formun Geliştirilmesi ve Ön Psikometrik Özelliklerinin İncelenmesi" başlıklı tezi 29.09.2017 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği"nin ilgili maddeleri uyarınca Özel Eğitim Anabilim Dalı Üstün Zekalılar Öğretmenliği Programında, Doktora tezi olarak kabul edilmiştir.

	<u>Unvanı-Adı Soyadı</u>	<u>İmza</u>
Üye (Tez Danışmanı)	: Prof.Dr. Uğur SAK	
Üye	: Prof.Dr. Atilla CAVKAYTAR	
Üye	: Prof.Dr. Engin KARADAĞ	
Üye	: Yard.Doç.Dr. Fatih KARABACAK	
Üye	: Yard.Doç.Dr. Mehmet ERSOY	


Prof.Dr. Handan DEVECİ
Anadolu Üniversitesi
Eğitim Bilimleri Enstitüsü
Müdürü

ÖZET

BİLİMSEL ÜRETKENLİK TESTİNİN 3, 4 VE 5. SINIF ÖĞRENCİLERİNE UYGUN FORMUNUN GELİŞTİRİLMESİ VE ÖN PSİKOMETRİK ÖZELLİKLERİNİN İNCELENMESİ

M. Bahadır Ayas

Özel Eğitim Anabilim Dalı

Üstün Zekalıların Eğitimi Programı

Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Eylül 2017

Danışman: Prof. Dr. Uğur Sak

Yaratıcılık çalışmalarını olumsuz yönde etkileyen faktörlerin başında yaratıcılık ölçme araçlarının sınırlılıkları gelmektedir. Yaratıcılık ölçme araçlarındaki sınırlılıkların başında alana özgü modern yaratıcılık tanımları ile uyumlu ölçme araçlarının sayısının çok az olması ve bu ölçme araçlarının psikometrik özelliklerinin yetersiz olması gelmektedir. Bilimsel yaratıcılık bağlamında düşünüldüğünde ilkökul öğrencilerine yönelik çok az sayıda bilimsel yaratıcılık testi bulunmaktadır. Var olan çok az sayıdaki bu testlerinin ise psikometrik özelliklerinin yeterli düzeyde olmadığı görülmektedir. Bu bağlamda, mevcut çalışmanın amacı, bilimsel yaratıcılığı açıklamak için geliştirilmiş olan Bilimsel Yaratıcılığın İkili Arama Modeli (Scientific Discovery as Dual Search-*SDDS*) temelinde 3, 4 ve 5. sınıf öğrencilerine yönelik bir bilimsel yaratıcılık testi geliştirmektedir.

SDDS'ye göre geliştirilen Bilimsel Üretkenlik Testi 3, 4 ve 5. Sınıflara Uygun Formu (BÜT 3-5), hipotez geliştirme ve deney tasarlama bileşenlerinden oluşmaktadır. Testte fizik, kimya ve biyoloji alanlarında 4 açık uçlu madde yer almaktadır. BÜT 3-5'in geliştirilme sürecinde, 10 uzman öğretmenin katıldığı 10 oturumda, kuramsal alt yapıya uygun çok sayıda madde geliştirilmiştir. Madde havuzundan uzman görüşü, ön deneme uygulaması ve pilot uygulama ile seçilen maddelerden bir uygulama formu oluşturulmuştur. Uygulama formu 647 öğrenciye (166 beşinci sınıf; 251 üçüncü sınıf; 220 dördüncü sınıf) uygulanmış ve bulgular ışığında madde seçimi yapılarak testin son hali

elde edilmiştir. Araştırmanın bulguları ile BÜT 3-5'in iki bileşenli yapısı deneysel olarak kanıtlanmış, ayırt edicilik geçerliği ve uyum geçerliği değerleri yeterli düzeyde bulunmuştur. Testin iç tutarlık ve puanlayıcılar arası güvenilirliğinin ise yüksek olduğu görülmüştür.

Araştırmanın bulgularına göre BÜT 3-5 geçerliği ve güvenilirliği deneysel olarak kanıtlanmış bir bilimsel yaratıcılık testidir. BÜT 3-5 özel yeteneklilere yönelik programların etkilerinin değerlendirilmesinde, bu tür programlara öğrenci seçimlerinde ve bilimsel yaratıcılığın gelişimsel olarak incelenmesinde kullanılabilir.

Anahtar Sözcükler: Bilimsel yaratıcılık, Yaratıcılık, Bilimsel süreç becerileri, Hipotez geliştirme, Deney tasarlama.

ABSTRACT

DEVELOPING CREATIVE SCIENTIFIC ABILITY TEST FOR 3RD, 4TH, AND 5TH GRADERS AND ANALYSIS OF ITS PRE-PSYCHOMETRIC PROPERTIES

M. Bahadır Ayas

Department of Special Education

Gifted Education Programme

Anadolu University, Institute of Educational Sciences, September 2017

Advisor: Prof. Dr. Uğur Sak

One of the leading factors affecting creativity studies negatively is the limitations of creativity assessments. First of all, there are few instruments compatible with modern theories creativity. Secondly psychometric properties of these instruments are not sufficient. In the context of scientific creativity, there are fewer scientific creativity tests for primary school students. It is seen that the psychometric features of existing instruments are inadequate. Therefore, the aim of this study is to develop a scientific creativity test for 3rd, 4th, and 5th grade students on the basis of Scientific Discovery as Dual Search Model (SDDS).

Creative Scientific Ability Test for 3rd, 4th and 5th Grade (C-SAT 3-5), which was developed according to SDSS, consists of hypothesis development and experiment designing components. There are four open-ended items in science areas Physics, Chemistry, and Biology. In developmental process of C-SAT 3-5, many items that are suitable to theoretical background were produced in 10 sessions with 10 expert teachers. An implementation form was created by using these items. Items were chosen according to expert opinion, try-out and pilot studies. The implementation form was given to 647 students (166 5th grade students, 251 3rd grade students, 220 4th grade students) and according to the results the final version of the test was created. Two-component structure of C-SAT 3-5 was proved. It was also found that C-SAT 3-5 has adequate concurrent and discriminant validity. From the reliability analyses the internal consistency and the interrater reliability of the test was found to be high.

According to the results of the study C-SAT 3-5 is a valid and reliable scientific creativity test for primary school students. It can be used in evaluating the effects of educational interventions in science like gifted programs, identification and screening processes in gifted education, and the researches on scientific creativity, scientific reasoning, science process skills and scientific thinking.

Keywords: Scientific creativity, Creativity, Science process skills, Hypothesis formation, Experiment design.



TEŞEKKÜR

2007 yılından bu yana birlikte çalışmaktan onur duyduğum danışmanım, değerli Hocam Prof. Dr. Uğur Sak'a minnettarım. Tez izleme komitelerindeki katkılarından dolayı Prof. Dr. Atilla Cavkaytar ve Yrd. Doç. Dr. Fatih Karabacak'a teşekkür ederim. Savunma komiteme yoğunluklarına rağmen katılan ve önerileri ile bu çalışmaya katkı sunan kıymetli hocalarım Prof. Dr. Engin Karadağ ve Yrd. Mehmet Ersoy'a teşekkürlerimi sunarım. Çalışmaya sundukları destekten dolayı Anadolu Üniversitesine ve Anadolu Üniversitesi Bilimsel Araştırma Proje Birimine teşekkür ederim. Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü Müdürü Prof. Dr. Handan Deveci'ye ve lisansüstü eğitimimin tüm kritik aşamalarında yardımlarını esirgemeyen enstitü personeline yardımlarından ve hoş görülerinden dolayı teşekkür ederim.

Hazırlık aşamalarında beni yalnız bırakmayan öğretmen arkadaşlarım Asiye Polat, Tahir Polat, Hakan Arslan, Meltem Arslan, Yasin Aytav, Dilara Bozdağ, Funda Çıray Özkara, Özlem Atalay ve Sezgin Sarı'ya; veri toplama aşamasında bana yardım eden Araştırma Görevlileri Hatice Kübra Sözel, Ömer Faruk Tamul, Esranur Dülger, Saadet Bayar, Mevlüt Cırık, Ercan Öpengin, Şule Demirel, Ferhat Köprü ve Deniz Arslan'a; okullardaki uygulamalarda benden desteğini esirgemeyen idareciler Seyit Yıldırım, Enver Yılmaz, Numan Köseer, İbrahim Taş, Murat Karaca, İsmail Hakkı Ceran ve değerli büyüğüm Aydın Çöken'e; madde seçimi ve okuyucular arası güvenilirlik çalışmalarında yardımlarını esirgemeyen arkadaşlarım Araştırma Görevlisi N. Nazlı Özdemir ve Araştırma Görevlisi A. Arda Atakaya'ya; tezimin dil düzeltmesinde yardımcı olan Araştırma Görevlisi Musa Tömen ve Dr. Ferdi Bozkurt'a; uygulama kitapçıklarının çoğaltılmasında yardımcı olan Çakmak Ailesi'ne teşekkür ederim.

Doktora eğitimimin başında henüz konuşamayan oğlum Bahadır Ali 3. Sınıfa başladı ve kızım Muhterem Umay Neva mart ayında dünyaya gelecek. Bu süreçte bir baba olarak kendilerinden esirgediğim her an için özür diler ve beni sabırla bekledikleri için teşekkür ederim. Ailemden, çocuğumdan ve evimden çaldığım zamanın telafisi olan, bu günleri mümkün kılan eşim Nimet'e heyecanımı paylaştığı için minnettarım. Anam ve babama eğitim hayatım boyunca gösterdikleri özverilerinden ve hoşgörülerinden dolayı teşekkür ederim. İlkokuldan doktora savunmama kadar, tüm öğretmenlerime ve hocalarıma sonsuz teşekkür ederim.

M. Bahadır AYAS

Eskişehir 2017

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan “bilimsel intihal tespit programı”yla tarandığını ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.


M. Bahadır Ayas

İÇİNDEKİLER

Sayfa

BAŞLIK SAYFASI	i
JÜRİ VE ENSTİTÜ ONAYI.....	ii
ÖZET	iii
ABSTRACT.....	v
TEŞEKKÜR	vii
ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	viii
İÇİNDEKİLER	ix
TABLolar DİZİNİ	xi
ŞEKİLLER DİZİNİ	xii
KISALTMALAR DİZİNİ	xiii
1. GİRİŞ.....	1
1.1. Sorun	1
1.2. Araştırmanın Amacı	5
1.3. Önem	6
1.4. Varsayımlar	7
1.5. Sınırlılıklar.....	8
1.6. Tanımlar.....	8
2. ALANYAZIN	9
2.1. Yaratıcılık	9
2.1.1. Üç bileşenli bağlamsal yaratıcılık modeli	10
2.1.2. Yaratıcılığın yatırım modeli.....	12
2.1.3. Dinamik sitemler modeli	12
2.2. Bilimsel Yaratıcılık.....	14
2.2.1. Bilimsel yaratıcılığın ikili arama modeli.....	16
2.2.2. Bilimsel yaratıcılık yapı modeli	18
2.2.3. Bilimsel yaratıcılığın gelişimi	19
2.2.4. Bilimsel yaratıcılık, motivasyon ve bilime yönelik tutumlar	22
2.3. Yaratıcılığın Ölçülmesi	24
2.3.1. Bilimsel yaratıcılığın ölçülmesi	26
2.3.2. Bilimsel yaratıcılığın ölçülmesinde çoğul düşünme testlerinin kullanılması	27
2.3.3. Bilimsel yaratıcılığın ölçülmesinde kişilik envanterleri.....	35
2.3.4. Bilimsel yaratıcılığın ölçülmesinde uzman değerlendirmesi.....	35
3. YÖNTEM	40
3.1. Araştırma Modeli.....	40
3.2. Çalışma Grubu	40
3.3. Veri Toplama Araçları.....	41
3.3.1. Bilimsel üretkenlik testi 3-5 (BÜT 3-5)	41

3.3.2.	BÜT 3-5 uygulama formu	59
3.3.3.	BÜT 3-5'in puanlanması	59
3.3.4.	BÜT 3-5'in uygulanması	61
3.3.5.	Fen bilimleri dersi akademik benlik algısı ölçeği (FABAÖ)	61
3.3.6.	Bilime yönelik tutum ölçeği (BYTÖ).....	62
3.3.7.	Karne notları	67
3.4.	Verilerin Çözümlemesi ve Yorumlanması.....	68
3.5.	Veri toplama süreci	69
4.	BULGULAR	71
4.1.	BÜT 3-5'in Geçerliğine İlişkin Bulgular	71
4.1.1.	BÜT 3-5'in yapı geçerliğine ilişkin bulgular	74
4.1.2.	BÜT 3-5'in ayırt edicilik geçerliğine ilişkin bulguları.....	81
4.1.3.	BÜT 3-5'in uyum geçerliğine ilişkin bulgular	86
4.2.	BÜT 3-5'in Güvenirliği	88
4.2.1.	BÜT 3-5'in iç tutarlılığına ilişkin bulgular.....	88
4.2.2.	BÜT 3-5'in okuyucular arası güvenirliğine ilişkin bulgular.....	92
5.	SONUÇ, TARTIŞMA ve ÖNERİLER	93
5.1.	Tartışma ve Sonuç	93
5.1.1.	BÜT 3-5'in geçerliğine ilişkin tartışma ve sonuç	93
5.1.2.	BÜT 3-5'in güvenirliğine ilişkin tartışma ve sonuç	106
5.2.	Öneriler	110
5.2.1.	İleri araştırmalara yönelik öneriler	111
5.2.2.	Eğitim uygulamalarına yönelik öneriler	113
	KAYNAKÇA	115
	EKLER	
	ÖZGEÇMİŞ	

TABLolar DİZİNİ

Sayfa

Tablo 2.1. Üç bileşenli bağlamsal yaratıcılık modeli bileşenleri.....	11
Tablo 2.2. BÜT Teorik Yapısı.....	28
Tablo 3.1. Katılımcıların sınıf düzeylerine ve cinsiyete göre dağılımı.....	41
Tablo 3.2. Birinci aşamada seçilen maddelerin göreve ve disipline göre dağılımları....	52
Tablo 3.3. FABAÖ maddeleri.....	62
Tablo 3.4. Geliştirilen fen bilimlerine yönelik tutum maddeleri.....	65
Tablo 4.1. Uygulama formunun betimsel bulguları.....	72
Tablo 4.2. Uygulama formu maddeler arası korelasyon katsayıları	73
Tablo 4.3. Monte Carlo simülasyonu ile yapılan paralel analiz sonuçları	76
Tablo 4.4. BÜT 3-5 uygulama formu PCA için açıklanan toplam varyans	77
Tablo 4.5. BÜT 3-5 uygulama formu maddelerinin bileşenlerdeki faktör yükleri	77
Tablo 4.6. BÜT 3-5 PCA bulguları	78
Tablo 4.7. BÜT 3-5 için hesaplanan DFA Uyum Ölçütleri ve Kabul Kriterleri	81
Tablo 4.8. BÜT 3-5 maddelerinin betimsel istatistikleri	82
Tablo 4.9. Madde ayırt ediciliğine ilişkin bağımsız örneklem t-testi sonuçları	83
Tablo 4.10. BÜT 3-5 puanlarının sınıf düzeylerine göre betimsel istatistikleri	84
Tablo 4.11. BÜT 3-5 puanlarının sınıf düzeylerine göre ANOVA sonuçları	85
Tablo 4.12. BÜT 3-5'in karne notlarına göre uyum geçerliği bulguları	86
Tablo 4.13. BÜT 3-5'in BYTÖ ve FABAÖ puanlarına göre uyum geçerliği bulguları	88
Tablo 4.14. BÜT 3-5 maddeler arası ve madde-toplam korelasyon bulguları.....	89
Tablo 4.15. BÜT 3-5 alt testler arası korelasyon katsayıları.....	91
Tablo 4.16. BÜT 3-5'in okuyucular arası güvenilirliği.....	92

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1. Amabile'ın (1983) Üç bileşenli bağlamsal yaratıcılık modeli.....	11
Şekil 2.2. Yaratıcılığın dinamik sistemler modeli.....	13
Şekil 2.3. SDDS modeli ve temel bileşenleri	17
Şekil 2.4. Bilimsel yaratıcılık yapı modeli.....	18
Şekil 2.5. BÜT'ün teorik yapısı.....	28
Şekil 3.1. BÜT 3-5 teorik yapısı.....	45
Şekil 4.1. BÜT 3-5 DFA sonuçları.....	80



KISALTMALAR DİZİNİ

- BÜT:** Bilimsel Üretkenlik testi
- BÜT 3-5:** Bilimsel Üretkenlik Testi 3, 4 ve 5. Sınıf Öğrencilerine Uygun Formu
- SDDS:** Bilimsel Yaratıcılığın İkili Arama Modeli
- SSCM:** Bilimsel Yaratıcılık Yapı Modeli
- SCT:** Liseler için Bilimsel Yaratıcılık Testi
- FBYT:** Okul Öncesi Görsel Bilimsel Yaratıcılık Testi
- BYT-5:** 5. Sınıflar İçin Bilimsel Yaratıcılık Testi
- Y-BÇT:** Yaratıcı Bilimsel Çağrışımlar Testi
- FABAÖ:** Fen Bilimleri Akademik Benlik Algısı Ölçeği
- BYTÖ:** Bilime Yönelik Tutum Ölçeği

1. GİRİŞ

1.1. Sorun

Yaratıcılığın geliştirilmesindeki en önemli engelin yaratıcılık ölçümlerindeki eksiklik olduğu düşünülebilir (Plucker and Makel, 2010). Çünkü uygun ölçme yapılmadan yaratıcılık bağlamında bireysel farklılıklara karar verme, ölçmeden elde edilen bilgiler ışığında farklılaştırma yapma ve yapılan farklılaştırmaların etkililiklerinin ortaya koyulması mümkün değildir. Ayrıca yaratıcılık gibi kompleks bir kavramın anlaşılabilmesi için de uygun ölçme süreçlerinin devreye sokulması gerekmektedir.

Bazı kaynaklarda *deha* olarak da tanımlanan üstün yaratıcı performans sergileyen bireylerin medeniyetimiz üzerindeki etkisi tartışılmaz bir gerçektir. Bu bireylerin çocukluk yıllarına ilişkin yapılan çalışmalarda, çalışma alanlarına duydukları ilgi ve ilk yaratıcılık işaretlerini çocukluk yıllarında verdikleri bilinmektedir (Csikszentmihalyi, 1997). Çağdaş üstün zekâ ve yaratıcılık modellerinde üstün zekâ ve üstün yaratıcılığın gelişiminde çevre ve eğitim önemli etkenler olarak sunulmaktadır. Bu bağlamda üstün yaratıcı bireylerin olabildiğince erken keşfedilmesi ve gerekli eğitim olanaklarının sunulması bireysel ve toplumsal yarar bağlamında çok önemlidir. Üstün yetenekli veya üstün zekâlı bireylerin, eğitsel gereksinimlerini karşılayacak eğitim programlarına hassas ve objektif bir tanılama sonunda dahil edilmeleri gerekmektedir. Bu amaca ulaşılabilmesi için geçerliği, güvenilirliği ve psikometrik özellikleri yeterli düzeyde olan ölçme araçlarının kullanılması gerekmektedir.

Modern üstün zekâ kuramlarının bazılarında yaratıcılık zekânın bir alt bileşeni olarak değerlendirilirken bazı modellerde zekâ yaratıcılığın bir alt bileşeni olarak sunulmaktadır. Bazı modellerde ise yaratıcılık ve zekâ birbirlerinden ayrı kavramlar olarak ele alınmaktadır. Zekâ ve yaratıcılık arasındaki ilişkiyi araştıran çeşitli araştırmalarda üstün yaratıcı bireyler ile üstün zekâlı bireylerin belirli bir zekâ düzeyinden sonra iki farklı grup oluşturdukları rapor edilmiştir (Sak, 2014). Bu bağlamda yaratıcı bireylerin tanılanmasında IQ testlerinin yanında yaratıcılık testlerinin de tanılama süreçlerine dâhil edilmesi gerektiği düşünülebilir. Ancak geçerliği ve güvenilirliği yüksek, kuramsal olarak tutarlı yaratıcılık testlerinin eksikliğinden dolayı birçok programın tanılama sürecinde yaratıcılık testleri kullanılmamaktadır (Hennessey and Amabile, 1999, s. 352).

Yaratıcılık çalışmaları, yaratıcılığın alana özgü bir yapıya sahip olduğunu desteklemektedir. Baer (1994), yaptığı araştırmalar neticesinde aynı alanda olsa dahi,

herhangi bir test maddesindeki yaratıcı performansın başka bir test maddesindeki yaratıcı performansın habercisi olmayabileceğini belirtmiştir. Genel yaratıcılık testlerinin ve ölçümlerinin alana özgü yaratıcılık ölçümlerinde kullanılmasının uygun olmadığı bazı araştırmacı tarafından vurgulanmaktadır (Lubart and Guignard, 2004; Kaufman, Plucker and Baer, 2008). Alana özgü ölçme yöntemleri incelendiğinde performans ölçümleri, alana özgü yaratıcılık testleri ve çeşitli envanterlerin alana özgü yaratıcılık ölçümlerinde kullanıldığı görülmektedir. Ancak bu yöntemler de kendi içlerinde farklı sınırlılıklar barındırmaktadır. Örneğin Plucker (1998) ve Runco (1987), alana özgü yaratıcılık ölçümlerinde performans ölçümlerinin, genel yaratıcılık ölçümlerinde ise kontrol listeleri ve kişilik envanterleri gibi araçların kullanılmasının daha yararlı olacağını belirtmişlerdir. Benzer şekilde Kaufman ve Baer (2004) öz değerlendirme ve kişilik envanterlerinin genel yaratıcılıkla ilgili, performans değerlendirmesinin ise alana özgü yaratıcılığın ölçülmesiyle ilgili olduğunu belirtmiştir. Performans ölçümlerinin ise müzik ve sanat gibi alanlara daha uygun olacağı, fen bilimleri ve matematik gibi akademik alanlarda kullanımının sınırlı olacağı düşünülebilir. Alana özgü yaratıcılık testlerinin ise sayılarının sınırlı olduğu bilinmektedir (Frederiksen and Ward, 1978; Hu and Adey, 2002; Kaufman and Baer, 2002; Runco, 1987; Plucker, 1998; Baer, 1994; Ayas and Sak, 2014; Sak and Ayas, 2013).

Yaratıcılık ölçümlerinde çoğul düşünme testleri, kişilik envanterleri ve ürün değerlendirme yaklaşımları sıklıkla kullanılmaktadır. Çoğul düşünme testleri açık uçlu sorulardan oluşmaktadır ve katılımcılardan çok sayıda çözüm üretmeleri istenmektedir. Bu testler yaratıcılığın ölçülmesinde diğer yöntemlerle kıyaslandığında daha sık kullanılmaktadır. Kişilik envanterlerinin kullanımında, özdeğerlendirme, akran değerlendirmesi, öğretmen değerlendirmesi ve aile değerlendirmesi gibi öznel görüşler üzerinden bireylerin yaratıcılıkları değerlendirilmektedir. Ürün değerlendirme yönteminde ise ölçümün yapılacağı alanın uzmanları, bireylerin ortaya koydukları ürünleri yaratıcılıklarına göre değerlendirmektedir. Sıralanan bu yaratıcılık ölçme yöntemleri, uygun ölçme kriterlerinin eksikliği, öznel değerlendirme ve test maddelerindeki teorik tutarsızlık gibi çeşitli nedenlerden dolayı eleştirilmektedir (Baer, 1994; Frederiksen and Ward, 1978; Hocevar, 1979; Kaufman, Plucker and Baer, 2008).

Amabile'a (1982) göre yaratıcılık ölçümlerindeki temel problem, uygun kriterlerin ölçümlerde kullanılmamasıdır. Amabile (1982, 1983), yaratıcılık tanımlarında yer alan "yenilik, uygunluk ve yararlılık" gibi kriterlerin yaratıcılık testlerinden elde edilen

puanlar üzerinden değerlendirilemeyeceğini belirterek, yaratıcılık ölçümlerinde ürünlerin yaratıcılıklarının alan uzmanlarının öznel görüşleri ile değerlendirilmesini önermektedir. Birçok araştırmacı tarafından önerilen bu yöntem (Carson, Peterson and Higgins, 2005; Baer, 1994; Baer and McKool, 2009; Kaufman et.al., 2008) ekonomiklik gibi sınırlılıklardan dolayı eleştirilmektedir (Baer, 2016, s. 135). Belirli bir sayıda ve uygun uzmanların bulunması, ürünlerin değerlendirilmesinde yaşanan zaman kaybı ve uzmanların maddi olarak desteklenmesi ekonomik sorunların başında gelmektedir. Ayrıca kimlerin uygun olabileceğine dair de net bir fikir bulunmamaktadır. Bilimsel yaratıcılık bağlamında düşünüldüğünde uzman değerlendirme yaklaşımının çeşitli sınırlılıkları olduğu görülmektedir. Örneğin fen bilimlerinde fizik, kimya ve biyoloji gibi disiplinlerdeki bilgi ve yöntemler birbirlerinden farklılık gösterdiğinden, uzmanların özellikleri araştırma bulguları ciddi şekilde etkileyecektir. Disiplinler arasındaki bu fark ikinci olarak uzmanların puanlamaları arasındaki tutarlılığın da düşmesine neden olacaktır. Bu bakımdan uzman da olsa, öznel değerlendirmelerin güvenilirlik ve inandırıcılık sorunları yaratabileceği düşünülebilir (Frederiksen and Ward, 1978).

Frederiksen ve Ward (1978) ise yaratıcılık testlerini içerdikleri maddeler bakımından eleştirmektedirler. Yazarlara göre testlerde yer alan maddeler yaratıcılık gibi karmaşık bir yapıyı ölçmek için yetersizdir ve birçok alana özgü test maddesi gerekir yaşam problemlerini yansıtmadığı için gereksiz veya çok basittir. Alan yazında rastlanan çok az sayıdaki bilimsel yaratıcılık testi bu görüşü destekler niteliktedir. Frederiksen ve Ward (1978) bu gerekçelerle bilim insanlarının gerçek yaşamda bilimsel yaratma sürecinde başvurdukları becerilere yönelik maddelerden oluşan bir yaratıcılık testi geliştirmişlerdir. Testte yer alan maddeler içerik olarak bilim ve yaratıcılıkla ilgili olmakla birlikte sadece lisans öğrencileri için uygun bir testtir. Lise öğrencileri için Hu ve Adey (2002) tarafından geliştirilen bilimsel yaratıcılık testinde de aynı şekilde bilimle ilgili maddeler yer almaktadır. Ancak bu test maddeleri detaylı incelendiğinde maddelerin genel yaratıcılık ölçümünde kullanılan araçlara çok benzediği ve basit görevler içerdiği görülmektedir. Örneğin genel yaratıcılık maddesi “bir tuğla için farklı kullanım alanı yazınız” bu testte “cam için bilimsel kullanım alanı yazınız” şeklinde verilmektedir.

Yukarıda açıklanan her iki bilimsel yaratıcılık testinin de geçerlik ve güvenilirliklerinin pilot uygulamalarından elde edilen verilere göre yeterli seviyede olduğu söylenebilir. Ancak her iki testte de fizik, kimya ve biyoloji gibi fen bilimleri alanlarının teorik ayrımından söz etmek mümkün değildir. Farklı bir ifade ile bu testlerde fen

bilimleri alanlarına özgü test maddeleri yer almamaktadır. Ayrıca arařtırmalarda ve tanılama işlemlerinde kullanılabilmesi için her iki prototip testle ilgili daha fazla arařtırmanın yapılması gerekmektedir.

Son 50 yılda bilimsel yaratıcılığın ölçülmesine yönelik çok az sayıda olmakla birlikte çeşitli çalışmalara rastlanmaktadır. Örneğın Fredriksen ve Ward (1978) ve Hu ve Adey (2002) bilimsel yaratıcılığın ölçülmesi amacıyla prototip bilimsel yaratıcılık testleri geliřtirmişlerdir. Fredriksen ve Ward (1978, s. 2), bilim insanının önemli özelliklerinden olan hipotez geliřtirme, önerileri deęerlendirme, metodolojik problemleri çözüme ve ölçme görevlerinin yer aldığı lisans öğrencilerine yönelik bir test geliřtirmişlerdir. Hu ve Adey (2002) ise Bilimsel Yapı Yaratıcılık Modelini (*Scientific Structure Creativity Model-SSCM*) kullanarak ortaöğretim öğrencileri için bir bilimsel yaratıcılık testi geliřtirmişlerdir (SCT). Kullanılan modelde hem akıcılık ve esneklik gibi genel yaratıcılık bileşenleri hem de hayal etme ve düşünme gibi bilimsel yaratıcılıkla ilgili bileşenler yer almaktadır.

Yaratıcılık ölçümlerindeki bir dięer sorun ise yaratıcılık ölçümlerinde kullanılan testlerin teorilerle uyumluluğudur (Ayas and Sak, 2014). Genel yaratıcılığın ölçülmesinde teorilerle uyumlu farklı yöntemler ve ölçekler kullanılmaktadır. Alan yazına bakıldığında farklı eğitim kademeleri için çeşitli alana-özgü yaratıcılık testlerinin geliřtirildięi görülmektedir. Ancak bilimsel yaratıcılık üzerine çeşitli çalışmalara ve teorik modeller olmasına rağmen (Hu and Adey, 2002; Klahr, 2000; Kanlı, 2014) bu teorilerle uyumlu çok az bilimsel yaratıcılık testinin geliřtirilmiş olduęu söylenebilir. Geliřtirilmiş olan testler ile temel alınan kuramsal yapılar arasında uyum sorunları olduęu görülmektedir. Örneğın üç boyutlu SSCM modeli temel alınarak Hu ve Adey (2002) tarafından geliřtirilen testin geçerlik çalışmaları sınırlı bir örneklem üzerinden yapılmış ve tek faktörlü bir yapı elde edilmiştir. SSCM modeli kullanılarak Chin ve Siew (2015) tarafından geliřtirilen testte tek faktörlü bir yapı, Siew, Chong ve Chin (2014) tarafından geliřtirilen testte ise 2 faktörlü bir yapı elde edilmiştir. Benzer şekilde üç bileşenli SDDS temel alınarak geliřtirilen ortaokul öğrencilerine uygun Bilimsel Üretkenlik Testin'nin (BÜT) (Ayas ve Sak, 2008) yapı geçerlięi çalışmalarında tek faktörlü bir yapı elde edilmiştir (Sak and Ayas, 2013; Ayas and Sak, 2014). Kanlı (2014) tarafından geliřtirilen Yaratıcı Bilimsel Çağrışımlar Modeli (Y-BÇM), bu model temel alınarak geliřtirilen Yaratıcı Bilimsel çağrışımlar Testi (Y-BÇT) ile deneysel olarak desteklenmiştir. Ancak Y-

BÇM'nin çağrışımlar, analogik muhakeme ve analogik problem çözme gibi daha genel bilişsel süreçleri içerdiği de göz önünde bulundurulmalıdır.

Bilimsel yaratıcılığın ölçülmesinde bir diğer sınırlılık ise farklı yaş grupları ve eğitim kademelerine yönelik testlerin yetersizliğidir. Geliştirilmiş olan çok az sayıdaki bilimsel yaratıcılık testleri ortaokul, lise ve üniversite düzeyindeki öğrenci grubuna uygundur. Araştırmalar fen bilimleri alanında alan bilgisi ve tecrübesi sınırlı olan öğrencilerin bilimsel olmayan yanıtlar üretme eğiliminde olduğunu göstermektedir (Pekmez, Aktamış ve Can, 2009; Kuhn, 1989). İlkokul öğrencilerinin bilimsel süreç becerilerindeki başarılarının görevin türüne göre değiştiğini ortaya koyan çeşitli araştırma bulgularına da rastlanmaktadır (Bullock, 1991, s. 13; Sodian, Zaitchik and Carey, 1991, s. 763). Bu bağlamda ilkokul öğrencilerinin bilimsel yaratıcılıklarını değerlendirmek için öğrencilerin gelişimsel özelliklerini ve fen bilimlerindeki tecrübelerinin dikkate alınması gerekmektedir. Alan yazında gelişimin hızlı olduğu ilkokul düzeyinde herhangi bir bilimsel yaratıcılık testine rastlanmamıştır.

Sonuç olarak bilimsel yaratıcılığın ölçülmesine yönelik çalışmaların sınırlı olduğu düşünülebilir. Alan yazında var olan bilimsel yaratıcılık ölçümlerinin ise genelinin teorik alt yapısının olmadığı görülmektedir. Bilimsel yaratıcılık teorilerinin ise bu teoriler temel alınarak geliştirilmiş testlerle kanıtlanamadığı veya uyumsuz olduğu, bu testlerin psikometrik özelliklerinin ise yeterli düzeyde olmadığı görülmektedir. SDDS modelinin, bilimsel süreç becerileri kapsamı bakımından diğer bilimsel yaratıcılık modellerinden ayrıldığı görülmektedir. Ancak bu model temel alınarak geliştirilen test ise ortaokul öğrencilerine uygundur. Alan yazında ise ilkokul öğrencilerine uygun bir bilimsel yaratıcılık testinin olmadığı görülmektedir. Bu bağlamda bu çalışmada ilkokul öğrencilerine yönelik bir bilimsel yaratıcılık testi geliştirilmesi amaçlanmaktadır.

1.2. Araştırmanın Amacı

Bu araştırmanın amacı 3, 4 ve 5. sınıf öğrencilerine yönelik SDDS modeli temel alınarak Bilimsel Üretkenlik Testi 3-5'i (BÜT 3-5) geliştirmek ve bu testin psikometrik özelliklerini ortaya koymaktır.

Araştırmanın amacına bağlı olarak aşağıdaki sorulara yanıt aranmıştır:

1. BÜT 3-5'in geçerliği nasıldır?
 - 1.1. BÜT 3-5'in yapı geçerliği nasıldır?
 - 1.2. BÜT 3-5'in ayırt edicilik geçerliği nasıldır?

- 1.2.1. BÜT 3-5'in alt, orta ve üst bilimsel yaratıcı yetenek gruplarını ayırt etme düzeyi nasıldır?
- 1.2.2. BÜT 3-5'in farklı sınıf düzeylerindeki öğrencileri ayırt etme düzeyi nasıldır?
- 1.3. BÜT 3-5'in uyum geçerliği (*convergent validity*) nasıldır?
 - 1.3.1. BÜT 3-5'in ders notlarına bağlı uyum geçerliği nasıldır?
 - 1.3.2. BÜT 3-5'in fen bilimlerine yönelik tutumlara bağlı uyum geçerliği nasıldır?
 - 1.3.3. BÜT 3-5'in fende akademik benlik algısına bağlı uyum geçerliği nasıldır?
2. BÜT 3-5'in güvenilirliği nasıldır?
 - 2.1. BÜT 3-5'in iç tutarlığı nasıldır?
 - 2.2. BÜT 3-5'in puanlayıcılar arası güvenilirliği nasıldır?

1.3. Önem

Alanyazında genel yaratıcılıkla ilgili kuramsal modellere (örneğin Guilford, 1950) ve modellerle ilgili yaratıcılık testlerine rastlanmaktadır (örneğin Torrance Yaratıcı Düşünme Testi-TTCT). Alana özgü yaratıcılık ölçümlerinde ise genellikle durumsal yaratıcılık ölçümlerinin yapıldığı görülmektedir (örneğin Fredrikssen and Ward, 1978; Kaufman, Baer and Gentile, 2004). Bilimsel yaratıcılık bağlamında düşünüldüğünde ise var olan kuramsal çerçevelerle uyumlu çok az sayıda bilimsel yaratıcılık testi olduğu görülmektedir. Az olmakla birlikte bu eksikliği gidermeye yönelik çeşitli bilimsel yaratıcılık testleri geliştirildiği görülmektedir (Kanlı, 2014; Hu and Adey, 2002, Ayas ve Sak, 2008). Var olan bu testler lise ve ortaokul öğrencilerine uygun testlerdir. Bu tez araştırması kapsamında, SDDS modeli temel alınarak geliştirilen BÜT 3-5, ilkökul öğrencilerine yönelik ilk bilimsel yaratıcılık testi olması bakımından önemlidir.

Yaratıcılık (Csikszentmihalyi, 1997; Amabile, 1983) ve üstün zekâ (Sak, 2013) gibi kavramların kültürel yapı oldukları birçok araştırmacı tarafından dile getirilmektedir. Bu bağlamda yaratıcılık ve üstün zekâ ölçme süreçlerinde, Türkiye'de geliştirilmiş ölçme araçlarının kullanılması gerekmektedir. Ulusal düzeyde ilkökul öğrencileri için geliştirilmiş standart bir yaratıcılık testi bulunmamaktadır. Ancak Torrance Yaratıcı Düşünme Testi gibi bazı testlerin Türk kültürüne uyarlamaları birçok araştırmada ve program tanılamalarında kullanılmaktadır. Uluslararası düzeyde ise bu yaş grubuna uygun bir bilimsel yaratıcılık testi yoktur. Bu bakımdan geliştirilecek olan test ulusal ve

uluslararası düzeyde bir ilk olması bakımından çok önemlidir. Ayrıca geliştirilecek olan test arařtırmacılar için de bir veri toplama aracı olması bakımından alanda var olan bir boşluęu dolduracağı düşünölebilir.

Çaędař üstün zekâ ve yetenek kuramlarına göre üstün zekâ ve yaratıcılık çok boyutlu bir yapıdadır (Gardner, 1993; Sternberg and Lubart, 1991) ve öęrencilerin yetenek alanlarının geliştirilmesi için çok boyutlu bir tanılama önerilmektedir (Sak, 2013). Çünkü tanılamadan elde edilen bilgiler ışığında eęitsel farklılaştırma yapılması esastır. Ancak özel yeteneklilere yönelik eęitim programlarında, tanılama süreçlerinin IQ testleri veya başarı testleri kullanılarak yapıldığı bilinmektedir. Bu durumun nedeni psikometrik özellikleri yeterli düzeyde olan yaratıcılık testlerinin eksikliğidir (Chen et al., 2002, s. 172). BÜT 3-5'in ayırt edicilik geçerliği, yapı geçerliği ve güvenirlik bulgularından dolayı, alandaki bu eksilik için umut vaat eden bir test olduęu düşünölebilir.

Alan yazında var olan bilimsel yaratıcılık testlerinin geliştirilme süreçlerinde çeşitli sınırlılıklar olduęu görölmektedir. Bu testlerin geliştirilme süreçlerinde küçük madde havuzları ile az sayıdaki uzman veya tek seferlik uzman görüşünün alındığı bilinmektedir (örneğin Siew, Chong and Chin, 2014). BÜT 3-5, test geliştirme süreci bakımından dięer testlerden ve test geliştirme süreçlerinden ayrılmaktadır. BÜT 3-5, yaratıcılık, bilimsel yaratıcılık ve SDDS modeli eęitimi verilmiş bir uzman ekip tarafından geliştirilmiştir. Bu test geliştirme yönteminin dięer test geliştirme yöntemlerine örnek olacağı düşünölmektedir. Ayrıca bu test geliştirme yöntemi sadece yaratıcılık veya bilimsel yaratıcılık alanına özgü değildir. Dięer alanlarda veya farklı ölçek geliştirme süreçlerine örnek teşkil etmesi bakımından önem arz etmektedir.

1.4. Varsayımlar

1. Arařtırmanın katılımcıları test uygulamasında en iyi performanslarını ortaya koymuşlardır.

2. Test geliştirme ekibinde yer alan öęretmenler, maddelerin deęerlendirmesinde objektif davranmışlardır.

3. Arařtırmada görüşlerine başvuru alan uzmanlar içten ve objektif deęerlendirmeler yapmışlardır.

4. Uygulama yapılan okullarda sınıfların homojen olduęu kabul edilmiştir.

1.5. Sınırlılıklar

Bu araştırma;

- Eskişehir il merkezinde faaliyet gösteren bir ortaokul ve iki ilkokul
- Eskişehir il merkezinde görev yapan öğretmenler
- İlkokul 3, 4 ve ortaokul 5. sınıf öğrencileri
- BÜT 3-5 testinde ölçülen özellikler ile sınırlıdır.

1.6. Tanımlar

Yaratıcılık: Yeni, özgün, etkili ve yararlı ürünler veya fikirler ortaya koyma becerisi (Ayas and Sak, 2014, s. 195).

Bilimsel Yaratıcılık: Deney tasarlama ve yapma, verilerden teoriler çıkarma, teorileri modifiye etme ve icat yapma (Kulkarni and Simon, 1988).

2. ALANYAZIN

2.1. Yaratıcılık

Dâhilerde ve küçük çocuklarda gözlenebilen yaratıcılık (Kaufmann and Sternberg, 2010), psikoloji, eğitim, sosyoloji, felsefe, sanat, tarih ve diğer tüm disiplinlerde çalışılmıştır ve çalışılmaya da devam etmektedir. Yaratıcılık ilkçağlardan günümüze kültürel, sosyal ve teknolojik değişimlerin tamamında kendisini göstermiş bir kavramdır. İnsanoğlunun ağaç kovuklarında başlayıp uzayın sonsuz derinliklerinde devam eden medeniyet dediğimiz serüvenindeki en temel yakıtının insanoğlunun yaratma eylemi olduğu söylenebilir. Yaratma eyleminin, tüm toplumsal ve sosyal değişimlerdeki etkisi bu kavramın çok farklı disiplinlerde ele alınmasına ve birçok yaratıcılık tanımının ortaya atılmasına neden olmuştur.

Yaratıcılık eski bir kavram değildir ancak ilgi çekici bir bilimsel çalışma alanı olarak ele alınmasının, alanı en çok etkileyen bilim insanı olarak kabul edilen J.P. Guilford'ın Zihinsel Yapı Modeli (Structure of Intellect-SOI) ile başladığı düşünülebilir (VanTassel-Baska, 1998, s. 365). Guilford (1956, 1967) yaratıcılığı; akıcı, esnek, detaylı ve orijinal fikir üretme olarak tanımlamıştır. SOI modeline göre akıcılık çok sayıda fikir üretme olarak tanımlanırken, esneklik farklı alanlarda veya kategorilerde fikir üretme becerisi olarak tanımlanmaktadır. Modelde yer alan orijinallik, üretilen fikirlerin az rastlanırlığının bir ölçüsüdür ve yaratıcılığın en önemli göstergesidir.

Üç boyutlu bir model olan SOI modelinde ürün (üniteler, sınıflar, ilişkiler, sistemler, dönüşümler, uygulamalar), işlem (değerlendirme, tekil üretim, çoğul üretim, hafıza, biliş) ve içerik (figüral, sembolik, semantik, davranışsal) boyutlarının farklı kombinasyonlarından oluşan 120 beceri (4x5x6) yer almaktadır. Yaratıcılıkla ilgili olan akıcı, esnek, orijinal ve detaylı fikir üretme becerileri, *çoğul üretim* bilişsel fonksiyonu altında toplanmıştır. Guilford, yaratıcılığın çoğul üretime dayalı, açık uçlu problem durumlarına üretilen yanıtlar üzerinden ölçülebileceğini ileri sürmüş ve çoğul üretim modeline dayalı ilk yaratıcılık testlerini geliştirmiştir. Yaratıcılığın ölçülmesine yönelik uygulamalarından dolayı, SOI modeli, psikoloji ve eğitim alanında yaratıcılık çalışmalarının gelişimine bir ivme kazandırmıştır (Rhodes, 1961; Gaut, 2010; Micheal, 1999). Günümüzde yaygın bir şekilde kullanılan yaratıcılık testlerinin birçoğu, çoğul üretim testleridir. Çağdaş yaratıcılık tanımlarının ve teorilerinin birçoğunda orijinalliğe diğer farklı bileşenlerle birlikte yaratıcılığın en temel göstergesi olarak yer verilmektedir.

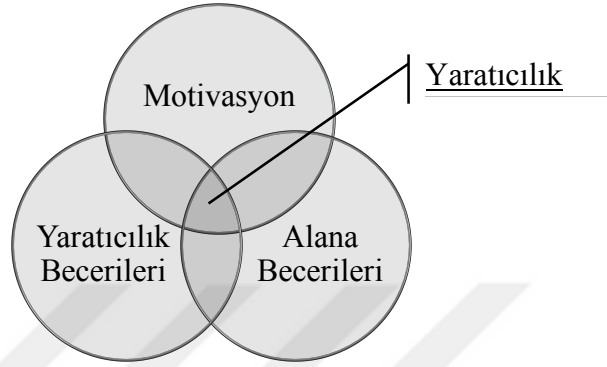
Guilford'dan (1950) günümüze birçok farklı yaratıcılık görüşü ortaya atılmıştır. Örneğin Repucci (1960) alan yazında 60 civarında yaratıcılık tanımı olduğunu belirtirken, Welsh (1981) yaratıcılık tanımlarının çok çeşitli olmasının, kavramın tanımlanmasını zorlaştırdığını vurgulamıştır (akt. Parkhaust, 1999). Ebert'e (1994) göre tek fikir birliği, yapılan tanımların yazarlara özgü olduğudur. Her ne kadar fikir birliğinden söz etmek mümkün olmasa da alan yazında rastlanan yaratıcılık tanımlarında bazı kavramlardan sıklıkla bahsedildiği görülmektedir. Birçok tanımda yenilik, problem durumuna uygunluk, etkililik, değer ve sürpriz yaratıcılık tanımlarının merkezinde yer almaktadır (Boden, 2004; Cropley, 1999; Sternberg and Lubart, 1992; Plucker, Beghetto and Dow, 2004; Runco ve and Jaeger, 2012; Kaufman and Sternberg, 2007). Bu özellikler ne tür fikirlerin yaratıcı değer taşıyacağına ilişkin ölçüt oluşturulması bakımından önemlidir.

Yaratıcılık tanımlarındaki çeşitlilik kavramın ne kadar kompleks olduğunun bir göstergesidir. Bu kompleks yapıyı açıklamaya çalışan bazı araştırmacılar yaratıcılıkla ilgili özelliklerin veya becerilerin setler hâlinde farklı bileşenler altında toplandığı *bileşensel yaratıcılık modellerini* ortaya atmışlardır. Bileşensel yaratıcılık modellerinin en büyük avantajları, yaratıcılığın doğasına ilişkin çok sayıda araştırma bulgusunu bütüncül şekilde kapsamaları ve yaratıcılığın alana özgü açıklanmasına olanak sağlamasıdır (Lubart, 1999). Bu modellerden “Üç Bileşenli Bağlamsal Yaratıcılık Modeli”, “Yaratıcılığın Yatırım Modeli” ve “Dinamik Sistemler Modeli” izleyen kısımlarda incelenmiştir.

2.1.1. Üç bileşenli bağlamsal yaratıcılık modeli

Bileşensel yaratıcılık modelleri arasında belki de en bilinen Amabile'in (1983) *üç bileşenli bağlamsal yaratıcılık* modelidir. Modele göre yaratıcılık üç bileşenin fonksiyonudur; alana özgü beceriler, yaratıcılığa özgü beceriler ve motivasyon. Motivasyon göreve karşı tutumları ve bireyin kendi motivasyonuna ilişkin algılarını içerir. İçsel ve dışsal olmak üzere ikiye ayrılan motivasyon, bireyi göreve bağlayan değişkenleri veya göreve karşı tutumlarını ifade etmektedir. İçsel motivasyon görevin doğasına ilişkin olup yaratıcılıkla ilgilidir. Dışsal motivasyon ise görevin dışında kalan ve görevle bireyin ilişkisini etkileyen durumlarla ilgilidir. Bu bakımdan sosyal ortam görev motivasyonunu etkiler. Alana özgü beceriler görevle ilgili olan bilgi, teknik beceriler ve özel yetenekleri kapsar. Müzik alanında nota bilgisi, herhangi bir müzik aletini kullanmaya ilişkin teknik bilgiler ve müzik kulağı gibi özel yetenekler alana özgü becerilere örnek olarak verilebilir. Yaratıcılığa özgü beceriler ise orijinal düşünme veya

akıcı düşünme gibi tüm alanlarda aynı olan becerilerdir. Yaratıcılığın sergilenebilmesi için bu bileşenlerin tamamının var olması gerekmektedir. Farklı bir ifade ile herhangi biri olmadığı zaman yaratıcılıktan da söz edilemez. Amabile'ın (1983, s. 68) bağlamsal yaratıcılık modeli ve modelin içeriği Şekil 2.1'de ve Tablo 2.1.'de verilmiştir.



Şekil 2.1. Amabile'ın (1983) Üç bileşenli bağlamsal yaratıcılık modeli

Tablo 2.1. Üç bileşenli bağlamsal yaratıcılık modeli bileşenleri

Yaratıcılığa Özgü beceriler	Motivasyon	Alana Özgü Beceriler
<p>İçeriği:</p> <ul style="list-style-type: none"> -Uygun bilişsel stil ve çalışma stili -Yeni fikir üretmeye yönelik içsel veya dışsal bilgi <p>Etkilendikleri:</p> <ul style="list-style-type: none"> -Alıştırma -Fikir üretme tecrübesi -Kişilik özellikleri 	<p>İçeriği:</p> <ul style="list-style-type: none"> -Göreve yönelik tutumlar -Bireyin kendi motivasyonuna yönelik algıları <p>Etkilendikleri:</p> <ul style="list-style-type: none"> -Göreve yönelik içsel motivasyon -Sosyal çevrede dışsal motivasyon kaynakları -Dışsal kısıtları bilişsel olarak minimize edecek yetenek 	<p>İçeriği:</p> <ul style="list-style-type: none"> -Alan bilgisi, alan hakkında bilgi -Gerekli teknik bilgiler -Alana özgü özel beceriler <p>Etkilendikleri:</p> <ul style="list-style-type: none"> -Doğuşla gelen bilişsel yetenekler -Doğumla gelen algısal ve motor beceriler -formel ve enformel eğitim

Daha sonraki çalışmalarında Amabile (2013, s. 135) Şekil 2.1'de verilen ve Tablo 2.1.'de verilen bileşenleri bireyle ilgili bileşenler olarak tanımlamış ve sosyal çevreyi dördüncü bir bileşen olarak modele eklemiştir. Modelde yer alan buluşsal bilgi ve alana özgü bilgi birbirinden farklılaşmaktadır. Buluşsal bilgi, bireyin problemi ele alma tarzıyla ilişkilidir ve dışsal kaynaklardan öğrenildiği gibi, içsel olarak da var olabilmektedir. Motivasyon bireyi göreve bağlamakla birlikte, dış faktörlerden etkilenmektedir. Bu bakımdan motivasyon aynı zamanda risk alam, belirsizliği tolere etme ve sebat/azim gibi bireysel özelliklerle de etkilidir. Bu özellikler dışsal kısıtların üstesinden gelinmesinde birincil öneme sahiptir.

Tüm bileşenlerden etkilenen yaratma süreci birtakım alt-bileşenlerden oluşmaktadır: problemi analiz etme, problemle ilgili bilgi toplama ve gerekli becerileri geliştirme, çözümler ve fikirler üretme, çözümü değerlendirme, çözümü sosyal çevre ile paylaşma. Model, yaratma sürecini, bireye ve çevreye ait özellikler ve aynı zamanda birey ve çevre arasındaki etkileşimle kapsamlı şekilde açıklaması bakımından diğer yaratıcılık modellerinden ayrılmaktadır.

2.1.2. Yaratıcılığın yatırım modeli

Üç bileşenli bağlamsal yaratıcılık modelinden etkilenen fakat farklılıklar içeren yatırım modeline göre yaratıcılık birbirinden bağımsız ama etkileşim içinde olan 6 bileşenden oluşur. Sternberg ve Lubart'a (1992) göre zekâ, bilgi, motivasyon, bilişsel stil, kişilik ve çevre bireylerin yaratıcılığının kaynağıdır. Yaratıcılık için üç zihinsel yetenek gerekmektedir: a) problemleri tanımlamak ve yeni biçimlerde sunmak için gereken *sentetik* yetenek, b) hangi fikirlerin değerli olduğunu belirlemeye yarayan *analitik* yetenek ve c) çevredeki diğer bireyleri yeni fikirler konusunda ikna etmeye yarayan *pratik* yetenek. Farklı bir ifade ile birey kaynakları analiz ve sentez yapmak ve fikirlerini pazarlamak için kullanır. Yaratıcı kişi ise fikirleri ucuza alıp pahalıya satan bireylerdir.

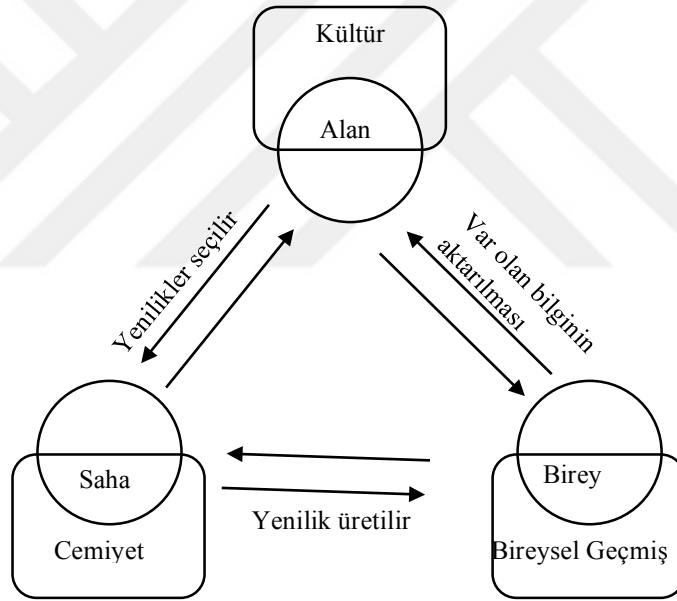
Modele göre yaratıcılık bileşenlerin basitçe toplamından daha fazlasıdır. Yaratıcılığın gerçekleşmesi için öncelikle kaynakların belirli eşik seviyelerini geçmesi gerekmektedir. Örneğin bilgi birikimi belirli seviyenin altında olduğunda yaratıcılık sergilenemez. İkinci olarak bileşenlerden birisindeki güçlü olması, zayıf olan bir bileşeni telafi edebilir (her iki bileşenin de belirli bir eşik değerinin üzerinde olmak şartı ile). Örneğin yüksek motivasyon, bilgi birikiminin az olmasını telafi edebilir. Üçüncü olarak yaratıcılık bileşenlerin basitçe toplamından öte, bileşenlerin etkileşimi ile de ilgilidir. Örneğin hem zekâ hem de motivasyonun yüksek olması, yaratıcılık üzerinde katlayıcı etkiye (çarpımsal etki) neden olabilir.

2.1.3. Dinamik sistemler modeli

Csikszentmihalyi'ye (1997) göre yaratıcılık, birey (person), çalışma alanı (field) ve disiplin (domain) olarak tanımlanan üç sistem arasındaki etkileşimin sonucunda ortaya çıkmaktadır. Birinci sistem olan *birey* çalışma sahasında var olan bilgiyi bilişsel süreçler, kişilik ve motivasyon gibi özellikleri ile dönüştürür veya geliştirir. İkinci sistem olarak değerlendirilen *çalışma alanı* ise o alanda çalışan ve alan üzerinde etkisi olan bireylerden oluşur ve yeni fikirlerin değerlendirilmesi ve seçilmesinde etkin rol oynar. Üçüncü sistem

olarak tanımlanan *çalışma disiplini*, kültürel olarak tanımlanmış, yaratıcı ürünleri de içeren ve kişiler arasında aktarılan bilgi birikiminden oluşur. Bu bilgi çalışılan disiplinindeki sembolik kurallar ve prosedürleri ifade eder.

Dinamik sistemler modeline göre bireyler çalışılan disiplin ve çalışma alanından etkilenirler, ancak aynı zamanda bu sistemlerde değişikliklere de neden olabilirler. Örneğin Van Gogh yaşamı boyunca sadece bir yağlı boya tablo satabilmiştir. Sistemler modeline göre resim alanında etkili olan sanatçılar, sanat eleştirmenleri, galeri sahipleri veya koleksiyonerler, Van Gogh'un eserlerini yaratıcı bulmamışlardır. Daha sonraki zamanlarda, alandaki değişimler sonucunda Van Gogh yaratıcı olarak kabul görmeye başlamıştır. Yaratıcı birey olan Van Gogh'un kabul görmesi ise alanda bir değişiklik yapmış ve kendinden sonra gelenleri etkilemiştir. Sistemler Modeli Şekil 2.2'de verilmiştir.



Şekil 2.2. Yaratıcılığın dinamik sistemler modeli

Yukarıda değinilen tanımlarda olduğu gibi günümüzde birçok yaratıcılık tanımında ilgi, motivasyon ve bilgi yaratıcılık için gerekli bileşenler arasında sıralanmaktadır. Bu bileşenlerin belirli bir alanla ilişkili olmasından dolayı yaratıcılığın alana özgü olduğu düşünülebilir. Alana özgü yaratıcılık fikri yeni değildir (Baer, 1998) ancak bilimsel kanıtlar görece yenidir. Alana özgü yaratıcılıkla ilgili ilk çalışmalar Einstein gibi yaratıcılığı herkes tarafından kabul görmüş bireylerin yaratma eylemleri ile ilgili retrospektif çalışmalardır. Bu çalışmalarda birey sadece kendi alanı içerisinde değerlendirilmektedir. Yeni çalışmalarda ise bireylerin birden fazla alandaki yaratıcı

performansları arasındaki ilişkiye bakılmaktadır. Yaş, alan/disiplin ve ölçme yöntemleri gibi değişkenler dikkate alınarak yapılan bu araştırmalarda, farklı alanlardaki yaratıcılık ölçümleri arasında düşük ilişki rapor edilmiştir (Runco, 1989; Sternberg and Lubart, 1992; Baer, 1991, 1996,1998; Kaufman and Baer, 2004). Hatta bazı araştırmalarda yaratıcılığın göreve özgü olduğu ve aynı alanda dahi farklı görevler arasında düşük ilişki bulmuşlardır (Almeida et. al., 2008; Ayas and Sak, 2014; Baer, 1994).

Alana özgü yaratıcılığa ilişkin deneysel kanıtlar, metodolojik bazı sınırlılıkları içermeleri bakımından eleştirilmektedir (Plucker and Zebelina, 2009). Bu eleştirilerin başında ölçme yöntemine ilişkin eleştiriler gelmektedir. Performans ölçmeleri ve uzman değerlendirmesine ilişkin eleştiriler alana özgü yaratıcılık kanıtlarına yöneltilen en kuvvetli eleştirilerdir. Genel yaratıcılığa özgü deneysel kanıtlarda azımsanmayacak kadar çoktur (Achter, Lubinski and Benbow, 1996; Lubinski, Benbow and Ryan, 1995; Torrance, 2004). Son çalışmalarda ise yaratıcılığın hem alana özgü hem de genel bileşenlerden oluştuğuna dair araştırmacıların görüşleri yaygınlaşmaktadır (Ayas ve Sak, 2014; Sak and Ayas, 2013; Kaufman and Baer, 2005; Plucker and Zabelina, 2009). Örneğin Alana özgü yaratıcılığın savunucularından Kaufman ve Baer (2005) ortaya attıkları Lunapark Teorisinde yaratıcılığın genel tematik ve alana özgü bileşenlerden oluştuğunu önerirken, genel yaratıcılık savunucularından Plucker, Beghetto ve Dow (2004) Gelişimsel Perspektif modellerinde yaratıcılığın genel yaratıcılıktan alana özgü yaratıcılığa doğru evrildiğini vurgulamışlardır. Bu bağlamda yaratıcılığın hem alana özgü bileşenlerden oluştuğunun hem de genel yaratıcılığa özgü bileşenlerinin olduğunun kabulü işlevsel tanımların yapılmasına olanak tanyacaktır.

2.2. Bilimsel Yaratıcılık

Yaratıcılık psikoloji alanında uzun yıllardan bu yana araştırılmasına rağmen, sanat ve edebiyat gibi diğer alanlarla karşılaştırıldığında, bilim insanlarının bilimsel yaratıcılıkları hakkında daha az sayıda çalışmanın yapıldığı söylenebilir. 1957 yılında Sputnik uzay aracının yörüngesine başarılı bir şekilde yerleştirilmesi ile bilim ve teknoloji alanında üstün yetenekli bireylere olan ilgi artmış ve bilimsel yaratıcılık çalışmaları da hızlanmıştır (Cropley and Cropley, 2005).

Bilimsel yaratıcılığın doğasına ilişkin bazı araştırmalarda bilimsel yaratıcılığın diğer yaratıcılık türleriyle aynı zihinsel süreçlerde kendisini gösterdiği fikri öne çıkmaktadır (Dumber, 1999). Ancak yaratıcı bilim insanlarının yaratma süreçleri incelendiğinde genel yaratıcılığa özgü bileşenlerle birlikte bilimsel yaratıcılığa özgü bazı

bileşenlerin de süreçte etkili olduğu görülür. Bu bakımdan bilimsel yaratıcılık, bilime özgü beceriler ile genel yaratıcılığa özgü becerilerin birleşimi olarak da görülebilir. Problem bulma ve çözme, hipotez oluşturma ve test etme, deney tasarlama, kanıt değerlendirme ve alan bilgisi bilimsel yaratıcılık sürecinde öne çıkan beceriler olarak sıralanabilir. Çoğul düşünme, çağrışımsal düşünme ve analogi kullanımı ise bilimsel yaratma sürecinde öne çıkan düşünme türleri olarak sıralanabilir.

Bilimsel yaratıcılık; zekâ, yaratıcılığa özgü beceriler, motivasyon, ilgi, konsantrasyon, bilgiyi arama ve zihinsel öğelerin olasılık permutasyonları gibi birtakım değişkenlerin bileşiminin bir sonucudur (Heller, 2007; Simonton, 1988). Hatta bazılarında göre fen bilimleri alanındaki üstün yetenek, sanatsal yetenekle bilimsel yeteneğin birleşimi olarak da görülmektedir (Innamorato, 1998). Yaratıcı potansiyelin alana özgü doğasından dolayı (Lubart and Guignard, 2004) bu görüşlerin bilimsel yaratıcılığın ölçülmesinde kullanılabilmesi için işlevsel bileşenler şeklinde tanımlanması gerekmektedir.

Yaratıcı bilimsel fikirlerin çok kapsamlı teorik, teknik ve deneysel bir süreçte oluşturulan bilgiye dayanması, bilimi diğer disiplinlerden ayırmaktadır (Dunbar, 1999). Daha farklı bir ifadeyle bilimsel yaratıcılık, deney tasarlama ve yapma, verilerden teoriler çıkarma, teorileri modifiye etme ve icat yapma (Kulkarni and Simon, 1988); hipotez üretme ve problem çözme (Klahr, 2000; Klahr and Dunbar, 1988; Newell and Simon, 1972); beklenmedik şeyler (Dunbar, 1993) ve zıt fikirler üzerinde çalışma (Rothenberg, 1971, 1996) gibi birçok farklı eylemlerin tümüyle ilgilidir.

Simon (1977), bilimsel yaratıcılığın farklı problem alanlarındaki problem çözme arayışı olarak düşünülebileceğini belirtmiştir. Bir problem alanı, bir problemin bütün hâllerini ve insanların bir hâlden bir diğer hâle geçmek için kullandıkları bütün eylemleri içerir. Bu bakımdan bilimsel yaratıcılığın, insanların bir durumdan başka bir duruma geçmek için kullandıkları süreç ve belirttikleri fikirler analiz edilerek anlaşılması mümkündür (Dunbar, 1999). Ancak sürecin tam anlamıyla anlaşılabilmesi için üretilen fikirler ve başvuru eylemlerin analizleri tek başına yeterli değildir. Bu eylemlerin ve fikirlerin nelerden beslendiğinin de ortaya koyulması gerekmektedir. Farklı bir ifade ile bilimsel yaratıcılığın nelerden oluştuğunun ortaya koyularak, bilimsel üretimlerin bu bileşenler bağlamında değerlendirilmesinin işlevsel olacağı düşünülebilir.

Bilimsel yaratıcılık çalışmaları, yaratıcı bilim insanlarının yaşamları (Rothenberg, 1995), laboratuvar gözlemleri (Dunbar, 1993) veya psikoloji laboratuvarlarında yapılan

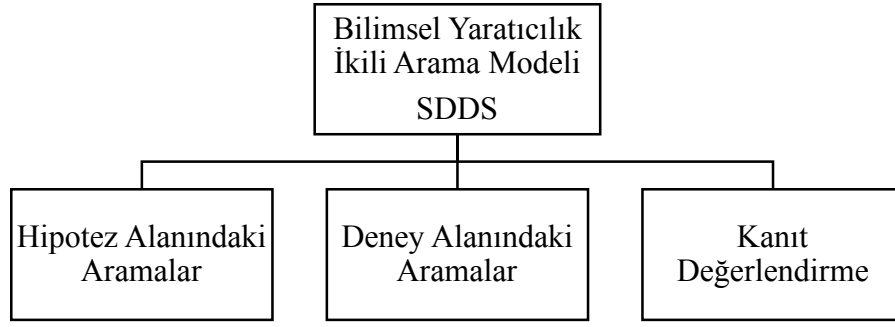
deneyle (Klahr, 2000) içermektedir. Bu çalıřmalardan bazıları bilimsel yaratıcılığın doğasına ilişkin kuramsal model önermeleri bakımından önemlidir. Bilimsel yaratıcılığın determinantlarının ortaya koyulması bakımından bu modellerin açıklanması faydalı olacaktır.

2.2.1. Bilimsel yaratıcılığın ikili arama modeli

Klahr (2000), bilimsel yaratıcılığın bileşenlerinin iki boyutlu bir sınıflandırmasını önermiştir. Birinci boyut alana özgü ve genele özgü bilgiyi kapsamaktadır. İkinci boyut ise buluş yapmada büyük öneme sahip hipotez oluşturma, deney tasarlama ve kanıt değerlendirme süreçlerini kapsamaktadır. Alana özgü bilgi edinimi çok önemlidir, çünkü bu sadece alandaki bilginin yapısını değil aynı zamanda o alandaki yeni hipotezlerin oluşturulması ve değerlendirilmesinde kullanılan süreçleri de kapsamaktadır. Bilimsel buluş esnasında alana özgü bilgi ve genele özgü deneyimler, deneylerin tasarlanmasında, yeni hipotezlerin seçilmesinde ya da oluşturulmasında ve kanıtların değerlendirilmesinde kullanılmaktadır.

Daha önceden değinilen Simon'un önerisini (1977) temel alan Klahr ve Dunbar (1988), hipotez alanı ve deney alanı olmak üzere iki özel alanı içeren bilimsel yaratıcılık görüşünü ortaya atmışlardır. Bu iki özel alandaki eylemlerde hipotez geliştirmek için başvurulan iki farklı strateji tanımlamışlardır: 1) olası hipotezler için hafızanın taranması (hipotez alanı), 2) verilerden yeni hipotezler oluşturuluncaya kadar deneyler yapılması (deney alanı). Klahr (2000), bu iki stratejiyi hipotez ve deney alanlarında aramalar olarak tanımlamış ve bilimsel buluş yapma sürecini açıklamak için SDDS'ye dâhil etmiştir.

Klahr (2000, s. 29), SDDS modelinde bilimsel buluşun iki farklı alandaki aramaların bir sonucu olduğunu iddia etmektedir. Bu alanlar hipotez alanı ve deney alanıdır. Bu alanlar ve bu alanlarda çözülmesi gereken problemler, farklı işlemler ve temsiller gerektirmeleri bakımından birbirlerinden farklılık gösterir. Hipotez alanındaki ürünler hipotezler olarak temsil edilirken, bu süreçte daha önceki bilgiler taranarak olası hipotezler üretilir. Deney alanındaki aramalar deneylerle temsil edilir ve deney verilerinden yeni hipotezler üretilene kadar deneyler yapılır. SDDS, hipotez ve deney alanlarındaki aramalara öncülük eden üç temel bileşenden oluşmaktadır. Bu bileşenler hipotez alanındaki aramalar, deney alanındaki aramalar ve bu iki alana aracılık eden kanıt değerlendirme bileşenleridir. SDDS modeli Şekil 2.3'te verilmiştir.



Şekil 2.3. SDDS modeli ve temel bileşenleri (Klahr, 2000, s. 31)

SDDS modeline göre bilimsel yaratma sürecinde öncelikle hipotez geliştirilir, daha sonra hipotezler deneylerle test edilir ve son olarak deney sonuçlarından elde edilen kanıtlar değerlendirilerek hipotezin kabul edilip edilmeyeceğine karar verilir.

Hipotez Alanı (Hipotez Oluşturma): Yeni hipotezlerin oluşturulması bir tür problem çözme sürecidir (Klahr, 2000). Bilimsel yaratma süreci bir problemle başlar. Bu ilk evrede problem hakkında birtakım bilgiler ve problemle ilişkili birtakım hipotezler vardır. Benzer şekilde hipotez alanının ilk aşamasında da birtakım bilgiler vardır ve bu ilk aşamadaki hedef bu bilgiyi açıklayacak bir hipotezdir. Hipotezler oluşturulduktan sonra ise uygunlukları kontrol edilir. Bu süreçte ön bilgiler ve deneysel veriler kullanılır. Hipotez alanındaki aramanın sonucunda iyi tanımlanmış bir hipotez geliştirilmiş olur. Bu yeni hipotez ise deney tasarımı için girdi olarak kullanılır.

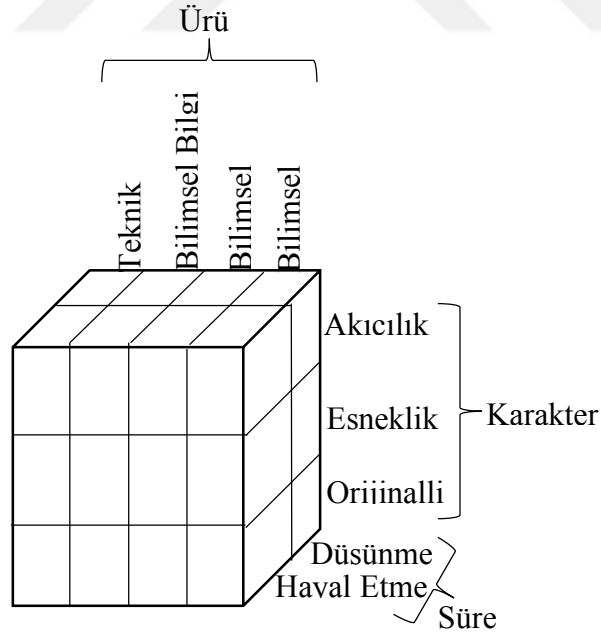
Deney Alanı (Hipotez Test Etme): Bilim insanları, ele aldıkları hipotezleri kanıtlayacak ya da çürütecek uygun deneyleri tasarlamak ve yapmak için deney alanında aramalar yaparlar. SDDS'ye göre hipotez test etme, hipotezlere uygun bir deneyin tasarlanması, tahminde bulunma, deneyi çalıştırma ve deneyin çıktıları ile tahminlerin birleştirilmesini içermektedir (Klahr, 2000). Bu süreçte, test edilen hipotezi destekleyecek kanıtlar üretilir. Bu kanıtlar, kanıt değerlendirme sürecinde girdi olarak kullanılır.

Kanıt Değerlendirme (Doğrulama): Kanıt sürecinde hipotezlerle açıklanan tahminler ile deney sonuçları karşılaştırılır (Klahr, 2000). Kanıt değerlendirme bileşeni, deneylerden elde edilen kanıtların hipotezin kabulü ya da retti için yeterli olup olmadığını belirler. Eğer kanıtlar yeterli değilse süreç hipotez alanından ya da hipotez test etme alanından baştan başlar. Bu bakımdan kanıt değerlendirme süreci hipotez ve deney alanlarındaki aramaların her ikisine de arabuluculuk eder ve teori ile kanıt arasındaki uyumun değerlendirilmesinde kullanılır.

SDDS modelinin iki önemli özelliği bu tez çalışması için önem arz etmektedir: bilime özgü bileşenlerden oluşması ve bu bileşenlerin psikolojik ölçümlere uygun şekilde tanımlanmış olması. Bu nedenle SDSS tez kapsamında geliştirilen bilimsel yaratıcılık testinin teorik yapısını da oluşturmuştur.

2.2.2. Bilimsel yaratıcılık yapı modeli

Bilimsel yaratıcılığa özgü bir diğer model ise Hu ve Adey (2002) tarafından önerilmiştir. Bilimsel yaratıcılığın ölçülmesinde ve geliştirilmesinde temel teşkil etmesi amacıyla geliştirilen Bilimsel Yaratıcılık Yapı Modeli (*Scientific Structure Creativity Model-SSCM*) üç boyuttan oluşmaktadır: ürün boyutu, karakter boyutu ve süreç boyutu. SSCM'ye göre bilimsel yaratıcılık sonucunda bireyler bilimsel problemlere yönelik çözümler, bilimsel olgulara yönelik açıklamalar ve bilimsel bilgi ve teknik *ürünler* ortaya koyarlar. Modelin *süreç* boyutu düşünme süreçleri ile ilgilidir ve yaratıcı düşünme ve hayal gücü bu boyutun alt bileşenleridir. Modelin üçüncü boyutu ise karakter özellikleri boyutudur ve akıcı, esnek ve orijinal düşünme yeteneği ile ilgilidir. SSCM Şekil 2.4'te verilmiştir.



Şekil 2.4. Bilimsel yaratıcılık yapı modeli

Hu ve Adey (2002, s. 392) bilimsel yaratıcılığı verilen bilgileri kullanarak belirli amaçlar doğrultusunda geliştirilmiş, bireysel ve toplumsal değeri olan, orijinal ürün

üretme potansiyeli olarak tanımlamışlardır. Bu tanımın detaylandırılması için yazarlar aşağıdaki aksiyomları önermektedirler:

- Bilimsel yaratıcılık deney yapma, bilimsel problemler bulma ve çözme gibi bilimsel eylemlerle ilgili olduğundan diğer yaratıcılık türlerinden farklılaşmaktadır.
- Bilimsel yaratıcılık bir yetenek türü olduğundan zihinsel olmayan faktörleri içermez, ancak zihinsel olmayan faktörler bilimsel yaratıcılığı etkileyebilir.
- Bilimsel yaratma eylemi bilimsel bilgi ve becerilerle ilgilidir.
- Bilimsel yaratıcılık statik ve gelişimsel yapıların bir kombinasyonudur. Çocuklar ve bilim insanları aynı temel bilimsel yaratıcılık yapılarına sahiptir, ancak bilim insanlarında bu yapı daha çok gelişmiştir.
- Yaratıcılık ve analitik zekâ, zihinsel yetenekten kaynaklanan tekil bir fonksiyonun iki ayrı faktörüdür.

SSCM üç boyutlu bir yapı sunması bakımından Guilfor'un SOI modeli ile benzerlikler taşımaktadır. Modelde bileşenlerin farklı kombinasyonları ile 24 farklı hücre elde edilmektedir.

2.2.3. Bilimsel yaratıcılığın gelişimi

Yaratıcılık, birtakım içsel ve dışsal değişkenlerin etkileşimi sonucunda sergilenmesinden dolayı nadir rastlanan bir özelliktir. Yaratıcılığın gelişimi, bilişsel gelişimle birlikte başlar ve yaşam boyu devam eder (Sak ve Maker, 2006, s. 279). Feldman'a (1999, s. 171) göre bu gelişim bilişsel ve sosyal/duygusal gelişimden, aile dinamiklerinden, genetik faktörlerden, eğitim ve tarihsel zorunluluklardan etkilenmektedir. Bu gelişimin seyri yaratıcılık alanlarına, bireyden bireye ve dışsal faktörlere göre değişmektedir. Bu bakımdan yaratıcılık gelişiminin doğrusal bir yol izlemediği düşünülebilir. Yapılan araştırmalarda çocuklarda (Runco, 1990) ve yetişkinlerde (Simonton, 1988) yaratıcılığın yaşa bağlı olarak eğrisel bir gelişim sergilediğini ortaya koymaktadır. Örneğin Simonton (1988) yetişkinlerde yaratıcılığın ters-J şeklinde geliştiğini, fen bilimlerinde 30-40 arası tavan yapan yaratıcılığın zamanla azaldığını ortaya koymuştur. Torrance (1968) çocuklarda yaratıcılığın sınıf düzeyine bağlı olarak geliştiğini, okulun ilk yıllarında hızlı bir gelişim aşamasından sonra 4. sınıfta yaratıcılıkta bir düşüş olduğunu ortaya koymuştur. Bulgular farklı alanlarda farklı

sonuçlar vermektedir. Bu çalışmanın amaçları doğrultusunda ilerleyen kısımda sadece bilimsel yaratıcılığın gelişimine odaklanılmıştır.

Bilimsel yaratıcılığın gelişimi üzerine yapılan çalışmalar ikiye ayrılabilir. Birinci tür araştırmalar ünlü bilim insanlarının yaratma eylemlerine anlamaya yönelik araştırmalardır. İkinci tür araştırmalarda ise çocuklarda bilimsel yaratıcılık ve bilimsel yaratıcılıkla ilgili bileşenlerin gelişimi üzerinde durulmaktadır. Ünlü bilim insanlarının bilimsel yaratıcılıkları üzerine yapılan araştırmalara göre (Rothenberg, 1996; Csikszentmihalyi, 1997; Simonton, 1984; Feist, 1997) bilimsel yaratıcılık zamana bağlı olarak doğrusal olmayan bir gelişim sergilemektedir. Örneğin Simonton'a göre (1984) ünlü bilim insanlarının yaratıcılıkları doğrusal bir rotada seyretmemektedir. Çalışmanın bulgularına göre ünlü bilim insanlarının yaratıcılıkları eğitim yaşamları boyunca ters-U şeklinde bir gelişim göstermektedir. Örgün eğitim ile hızlı bir yükselişe geçen yaratıcılık, eğitim hayatının ortalarında en üst seviyeye çıkmakta ve eğitimin geri kalanında azalmaktadır (Aktaran, Simonton, 1988, s. 252). Weisberg (1999, s. 226) bu durumun sebebinin eğitim olmadığını, yaratıcılık için alan bilgisinin gerektiğini ancak fazlasının üretimi sıradanlaştıracağını belirtmektedir.

Çocukların bilimsel gelişimi üzerine yapılan çalışmaların genellikle bilim insanlarının kullandığı hipotez geliştirme, kanıt değerlendirme veya deney tasarımı gibi bilimsel becerilerin gelişimine odaklandığı görülmektedir. Klahr ve Simon'a (1999, s. 525) göre çocuklar, bilim insanları gibi, etraflarındaki dünyayı anlamak için çeşitli hipotezler kurar, denemeler yapar, sonuçları yorumlar. Bu görüşü destekleyecek şekilde çeşitli araştırma bulguları da mevcuttur. Örneğin Samarapungavan (1992, s. 5-24) ilkökul öğrencilerinin benzer bilimsel fikirleri/teorileri seçme kriterlerinin bilim insanları ile benzerlik taşıyıp taşımadığını araştırmışlardır. Araştırmacılar bu çalışmada öğrencilerin bilim insanları ile benzer kriterleri kullanabildiklerini bulmuşlardır. Araştırmanın bulgularına göre ön bilgileri ile çelişmeyen durumlarda öğrenciler rakip teorilerin seçiminde deneysel ve mantıksal tutarlılık kriterlerine göre seçim yapabilmekte, sınıf düzeyi ile birlikte (1. sınıf, 3. sınıf ve 5. sınıf) bu kriterlerin kullanımı da artmaktadır.

Alan yazında çocuklarda bilimsel süreç becerilerinin gelişimine ilişkin çalışmaların daha çok tek tip beceriye odaklandığı görülmektedir. Örneğin Sodian ve diğerleri (1991) 1. ve 2. sınıflarla yaptıkları çalışmada öğrencilerin verilen hipotezleri ve deney düzeneklerini doğru şekilde eşleştirebildiklerini, verilen hipoteze uygun deneyler tasarlayabildikleri ve bu görevlerde 2. sınıfların 1'lerden daha başarılı olduklarını

bulmuşlardır. Ruffman ve diğerleri (1993) deney sonuçlarına ilişkin hipotez geliştirme becerisinin 6-7 yaşından itibaren gözlenebileceğini araştırma bulguları ile ortaya koymuşlardır. Tchirgi (1980, s. 9) ise öğrencilerin 2. sınıftan itibaren deney düzeneğindeki değişkenleri manipüle etmeye başladıklarını, değişken sayısı arttığında başarımın düştüğünü ortaya koymuşlardır. Bu araştırma bulguları Klahr ve Simon'un (1999) çocuk ve bilim insanı benzerliğini desteklemektedir. Ancak Kuhn (1989), bilimsel düşünmenin hipotez üretmek, deney tasarlamakla sınırlı olmadığını, teoriler ve kanıtlardan yeni kavramlar üretmekle ilgili olduğunu savunmaktadır. Kuhn yaptığı çalışmada, çocukların (3, 6 ve 9. sınıf) ve yetişkinlerin (Lisans öğrencileri, yüksek lisans mezunları ve lisans okumayan yetişkinler) bilim insanları gibi teori ve kanıtları eşleyemediklerini ve teorileri revize edemediklerini bulmuştur. Araştırmanın bulgularına göre kanıta dayalı yanıtlar, eğitim düzeyine bağlı olarak artmaktadır.

Alan yazında bilimsel yaratıcılığın gelişimi ile ilgili diğer çalışmalar ise bilimsel yaratıcılık testlerinden elde edilen verilere dayanmaktadır. Yang ve diğerleri (2016) yaptıkları çalışmada 3-6 sınıf öğrencilerinin bilimsel yaratıcılıkları ile hipotez geliştirme, deney yapma ve çıkarımda bulunma becerileri arasında sırasıyla .39, .49 ve .39 ($p<.001$) düzeylerinde anlamlı korelasyon katsayıları bulmuşlardır. Araştırmanın bulgularına göre öğrencilerin bilimsel yaratıcılık puanları sınıf düzeyine göre düzenli bir şekilde artmamaktadır. 3 ve 4. sınıflar arasında büyük farklılıklar bulunurken, 5. sınıfların 4. sınıflardan daha düşük puan aldıkları görülmüştür. Aynı çalışmada öğrencilerin bilimsel süreç becerilerin sınıf düzeyi ile arttığı, 3 ve 4. sınıflar arasında büyük farklılıklar olduğu, 4 ve 5. sınıflar arasında ise 5. sınıflar lehine küçük bir fark olduğu görülmüştür. Araştırmacılar 3 ve 4. sınıflar arasındaki büyük farkın 3. sınıfların sınırlı bilgi birikiminden kaynaklanabileceğini, 4. sınıfların 5. sınıflardan yüksek puan almalarını ise 4. sınıftaki öğrenci merkezli ve proje tabanlı eğitimden kaynaklanabileceğini belirtmişlerdir. Ayverdi ve diğerleri (2012, s. 655) ise 6 ve 8. sınıf öğrencileri ile yaptıkları çalışmada 8. sınıfların bilimsel yaratıcılık düzeylerinin en düşük, 7. sınıfların ise en yüksek puan ortalamalarına sahip olduklarını rapor edilmişlerdir. Araştırmacılar 8. sınıfların düşük bilimsel yaratıcılık puanlarını lise geçiş sınavları ile açıklamışlardır. Ancak araştırmaların tamamında bulunan sınıflar arası beklenmedik farklar tamamen örneklem hatalarından da kaynaklanabilir.

Bilimsel yaratıcılığın gelişimine ilişkin yapılan çalışmaların ağırlıklı olarak ünlü bilim insanlarının yaşamlarına odaklandığı görülmektedir. Çocuklarla yapılan çalışmalar ise

bilimsel süreç becerilerinin gelişimi ve bilimsel yaratıcılık testlerindeki performanslara dayanmaktadır. Araştırma bulgularına göre bilimsel yaratıcılığı besleyen bilimsel süreç becerileri 6-7 yaşlarından itibaren gözlenmektedir. Klahr, Fay ve Dunbar'a (1993) göre bilimsel yaratıcılığı besleyen diğer bir unsur olan alan bilgisi ve bilimsel süreç becerileri arasında ise çift taraflı bir ilişki vardır. Alan bilgisi bilimsel süreç becerilerinin kullanımını olumlu yönde etkilerken, çocuklar gelişimsel olarak bu becerileri kullanarak bilgi birikimlerini de arttırmaktadırlar. Bilimsel yaratıcılık test performanslarından elde edilen araştırma bulguları ise sabit değildir. Bu durumun birincil nedeni bilimsel yaratıcılığın ölçülmesinden kaynaklanabileceği gibi dışsal faktörlerin kontrol edilemediği araştırma desenlerinde de kaynaklanabilir.

2.2.4. Bilimsel yaratıcılık, motivasyon ve bilime yönelik tutumlar

Yukarıda yer verilen yaratıcılık tanımlarının ve açıklanan yaratıcılık kuramlarının birçoğunda motivasyon veya ilgi önemli bileşenler olarak karşımıza çıkmaktadır. Bu tanımlarda motivasyon veya ilgi bireyi yaratma eylemine veya yaratıcılık alanına bağlayan bir özellik olarak tanımlanmaktadır (Csikszentmihalyi, 1997; Amabile, 1983; Renzulli, 2002; Sternberg and Lubart, 1992). Diğer bileşenler ile kıyaslandığında motivasyon belki yaratıcılığın en önemli bileşeni olarak düşünülebilir (Conti and Amabile, 1999, s. 251; Sawyer, 2006). Örneğin Rothenberg'e (1987) göre Einstein'ı akranlarından ayıran en önemli özelliği çalışma alanına olan tutkusudur. Csikszentmihalyi (1997) ise yaratıcı bilim insanları ile yaptığı çalışmada bu bireylerin çalışma alanlarına karşı tutkularını ortaya koyarak bilime karşı ilgilerinin erken yıllarda başladığını belirtmektedir.

Amabile'a (1988, s. 133) bileşensel yaratıcılık modelinde motivasyon iki bileşenden oluşmaktadır: 1) göreve karşı temel seviyede tutumlar ve 2) bireyin girişimlerine yönelik öz algıları. Göreve karşı temel tutumlar bireyin göreve yatkınlığı ile ifade edilebilir. Çünkü tutumlar bireyin belirli bir görevi sevip sevmediği ile ilgidir. Bu bağlamda çocukların fen bilimleri gibi alanlarda tecrübeleri sınırlı olduğundan, fen bilimlerine yönelik tutumları (sevme, ilgilenme) motivasyonun göstergesi olarak kullanılabilir. Fen bilimlerinde yaratıcılık bağlamında düşünüldüğünde çocukların fen bilimlerine yönelik olumlu tutumları, bu alandaki eylemlerini arttıracak ve öğrenmelerine katkı sağlayacaktır. Bilimsel yaratıcılık ve alan bilgisi arasındaki ilişki düşünüldüğünde bilime yönelik tutumların yaratıcılıkla ilişkisi daha önemli hâle gelmektedir. Tutumlar bireylerin insanlara, nesnelere, konulara ve olaylara karşı sergiledikleri zihinsel

yatkınlıklar olarak tanımlanabilir (Tosun ve Genç, 2015, s.946). Bu bağlamda tutumların öğrenme üzerinde etkileri olduğu düşünülebilir. Örneğin fen bilimlerine yönelik olumlu tutuma sahip öğrencilerin fen konularını daha kolay ve kalıcı şekilde öğrenebilecekleri söylenebilir.

Fen bilimleri dersi bağlamında düşünüldüğünde tutumların öğrenme üzerindeki etkisini ortaya koyan çok sayıda araştırma bulgusuna rastlanmaktadır. Örneğin, Weinburg (1995, s. 396) 1970- 1991 yılları arasında fenne yönelik tutum ve akademik başarı arasındaki ilişkiyi araştırdığı meta analiz çalışmasında, öğrencilerin fen bilimlerine yönelik tutumları ile fen bilimleri dersindeki başarı arasında anlamlı ilişkiler rapor etmiştir. Germann (1988, s. 700) ise çeşitli başarı testleri ile fenne yönelik tutumlar arasında düşük korelasyon katsayıları bulmuş ancak, ödev ve sınıf çalışmasına dayalı başarı ile fen tutumları arasında orta ve yüksek düzeyde korelasyon katsayıları bulmuştur. Bu bağlamda fenne yönelik olumlu tutumların öğrencilerin öğrenmeleri, öğrenilenlerin kalıcılığı ve ders içi katılımı olumlu yönde etkilediği düşünülebilir. Farklı çalışmalarda ise fen bilimleri dersine yönelik tutumların gelişimsel olarak değişimi araştırılmıştır. Örneğin Smail (1993), fen bilimleri dersine yönelik tutumların öğrencilerin ilerleyen yıllarda aldıkları fen derslerini ve fen bilimleri alanında kariyer tercihlerini etkilediğini rapor etmiştir (aktaran Pell and Jarvis, 2001, s. 847). Bu bağlamda fen bilimleri dersine yönelik olumlu tutumların öğrencilerin gelişim çizgisi üzerinde bir etkisi olduğu düşünülebilir.

Çeşitli yaratıcılık tanımlarında ilgi veya motivasyon yaratıcılık için gerekli bir bileşen olarak değerlendirilmektedir (Amebile, 1983; Renzulli, 2002). Alana özgü yaratıcılık teorilerinde, alan bilgisi ve alana özgü beceriler birincil öneme sahiptir. Weisberg'e (1999) göre bireylerin yeni bir şeyler yaratabilmek için en azından o alanda neyin yeni olduğuna dair bilgi birikimi olması gerekmektedir. Dunbar (1999) ise fen bilimlerinde yaratıcılığın derin alan bilgisine bağlı olduğunu belirtmektedir. Alana özgü bu tür yaratıcılık tanımlarında motivasyon ve ilgi bireyi yaratımda bulunduğu alana bağlayan bir bileşen olarak değerlendirilmektedir. Örneğin Csikzentmihalyi (1997) fen alanında yaratıcılıkları ile bilinen 300'e yakın bilim insanı ile yaptığı çalışmada, bu bireyleri benzerlerinden ayıran en önemli özelliğin çalışma alanlarına olan tutkuları olduğunu belirtmiştir. Aynı araştırmanın bulgularına göre olağan üstü yaratıcı bilim insanları çocukluk yıllarında fen bilimleri alanında ilgilerinin yüksek olduğu rapor edilmiştir. Bu bağlamda fen bilimlerine yönelik olumlu tutumların bireylerin alandaki

çalışmalarını ve dolaylı olarak alan bilgilerini arttırdığı, alan bilgisinin ise bilimsel yaratıcılıklarını arttırdığı sonucuna ulaşılabilir.

Birçok araştırmacı ilgi ve yaratıcılık arasında güçlü ilişkiler önermektedirler. Duke'e (1972, s. 383) göre yaratıcılık ile zekâ, kişilik ve tecrübe arasında açık bir ilişki olmadığını ancak yaratıcılık ile ilgi ve özgüven arasındaki ilişkinin yaratma sürecinde çok önemli olduğunu vurgulamaktadır. Hatta yaratıcılığın olmamasını, ilgili yetersizliği olarak tanımlamaktadır. Farklı bir ifade ile ilgimizin olmadığı bir alanda yaratıcı olmamız söz konusu değildir. Çünkü ilgimizi çeken konularda daha fazla okur, daha fazla zaman ayırır ve daha çok çalışırız. Liang (2002) araştırmasında Duke'ü (1972) destekleyecek nitelikte bulgulara ulaşmıştır. Araştırmada bilimsel yaratıcılık ile yaratıcılık ölçümleri (orijinallik, detaylandırma ve açıklık), fen bilimlerine özgü ölçümler (hipotez test etme, bilimin doğası, fen akademik başarı) ve bilime yönelik tutumlar arasındaki ilişkiler incelenmiştir. Araştırmanın bulgularına göre en yüksek korelasyon değerleri bilimsel yaratıcılık ile bilime yönelik tutumlar arasında bulunmuş ve bilime yönelik tutumların tek başlarına bilimsel yaratıcılık puanlarının %35'ini anlamlı bir şekilde yordadığı görülmüştür. Bu bağlamda fen bilimlerine yönelik tutumlar bilimsel yaratıcılığın değerlendirilmesinde destekleyici veriler olarak kullanılabilir.

2.3. Yaratıcılığın Ölçülmesi

Yaratıcılık insanoğlunun en temel kaynağı olarak değerlendirilebilir (Toynbee, 1964, akt. Batey, 2012). Bu özelliğinden dolayı birçok farklı disiplinde ilgi çekici bir çalışma alanıdır. Ancak zekâ ve kişilik gibi benzer diğer kavramlarla karşılaştırıldığında görece daha az çalışılan bir konudur (Guilford, 1950; Sternberg and Lubart, 1992). Batey ve Furnham'a (2006) göre yaratıcılık çalışmalarının önündeki en büyük engel, açık ve net bir yaratıcılık tanımının olmayışıdır.

Zaman içerisinde yaratıcılık tanımları çeşitlenmiş ve alan bilgisi, motivasyon, çevre, kişilik özellikleri gibi çok farklı bileşenler yaratıcılık modellerine dâhil edilmiştir. Yaratıcılık tanımlarındaki bu değişimin yaratıcılığın ölçüme yöntemlerini de etkileyebileceği düşünülebilir. Ancak alan yazında rastlanan testlerin birçoğu genel yaratıcılığa özgü çoğul düşünme testleridir. Bu durumun başlıca nedeni, yaratıcılık tanımlarının işlevsel olmamasıdır (Runco, 2004; Parkhurst, 1999). İşlevsel tanımdan kasıt ise yaratıcılığın ölçülebilir, geliştirilebilir bileşenlerle ifade edilmesidir. Örneğin Kekule'nin yaratma eylemi gündüz düşü (day dreaming) ile açıklanması (Rothenberg,

1995) veya Leonardo da Vinci'nin yaratıcılığının mistik öğelerle açıklanması (Shlain, 1999, s. 508) işlevsellikten çok uzak girişimler olarak düşünülebilir.

Yaratıcılık ölçümlerine bakıldığında ölçme yöntemleri farklı sınıflar altında toplanabilir. Bu sınıflamalar ölçme araçlarının özelliklerine göre veya ölçülen özelliğe göre oluşturulmaktadır. Örneğin Weisberg (2006) yaratıcılık ölçümlerini kullanılan araçlara göre başarı ölçme araçları ve potansiyel ölçme araçları olmak üzere iki gruba ayırırken Davis (1997) ise çoğul düşünme testleri ve kişilik ölçme envanterleri olmak üzere iki gruba ayırmıştır. Hocevar ve Bachelor (1989) ise yaratıcılık ölçümlerine ilişkin bir taksonomi vermiştir. Bu taksonomiye göre yaratıcılık ölçümlerini sekiz kategoriye ayırmışlardır: çoğul düşünme testleri, ilgi ve tutum ölçekleri, kişilik envanterleri, biyografik envanterler, ürün değerlendirme, diğerleri tarafından değerlendirmeler, ünlülük (eminence) ve özdeğerlendirme anketleri. Yaratıcılık ölçümlerinde en yaygın yapılan sınıflama ise yaratıcı kişi, yaratıcı süreç, yaratıcı ürün ve yaratıcı çevrenin ölçülmesidir (Kaufman, Plucker and Baer, 2008; Houtz and Krug, 1995; Batey and Furnham, 2006; Starko, 2005).

Yaratıcı kişilik ölçümlerinde kontrol listeleri veya kişilik envanterleri gibi öznel veri toplama araçları ile bireylerin yaratıcılıkları değerlendirilmektedir. “herhangi bir alanda alınan ödül” veya “herhangi bir alana ilgi düzeyi” bu tür ölçümlerde kullanılan maddelere örnek olarak verilebilir. Yaratıcı süreç ölçümlerinde yaratıcılıkla ilgili olduğu kabul edilen akıcılık, esneklik, orijinallik veya detaylandırma gibi düşünme becerileri değerlendirilmektedir. Çoğul düşünme testleri süreç ölçümlerine örnek olarak verilebilir. Yaratıcılık için ürün ölçümünde bireylerin yaratıcılıkları ortaya koydukları ürünler değerlendirilerek yapılmaktadır. Sanat yarışmalarında yapılan değerlendirmeler bu tür ölçümlere örnek olarak verilebilir. Yaratıcı çevre ölçümünde ise bireylerin sosyal ortamları yaratıcılık bağlamında değerlendirilmektedir.

Yaratıcılık teorilerinin ölçme türlerini şekillendirdiği düşünüldüğünde Guilford'un (1950) çoğul düşünme testlerinin yaratıcılık ölçümlerindeki etkisi anlaşılabilir. Yaratıcılık çalışmalarında büyük eforların sadece çoğul düşünme testlerine sevk edilmesi gibi bir ironi vardır. Farklı bir ifade ile yaratıcılık çalışmalarında fazla bir çoğul ürüne rastlanmamaktadır (Kaufman, Plucker and Baer, 2008). Çoğul düşünme testlerinde katılımcıları açık uçlu sorulara veya görevlere birden fazla yanıt üretirler. Bu yanıtlardan bazıları diğer yanıtlarla karşılaştırıldıklarında az rastlanır olmaları bakımından orijinal olarak kabul edilirler ve yaratıcılığın göstergesi olarak kullanılırlar. Yaratıcılık

ölçümlerinde en sık kullanılan Torrance Yaratıcı Düşünme Testi bir çoğul üretim testidir. Hem alana özgü hem de genel yaratıcılığa özgü birçok yaratıcılık testinin olması, uygulama kolaylığı ve ulaşılabilir olmaları bakımından çoğul düşünme testleri alanda hâkim yaratıcılık testleridir.

Davis'in (1997) yaptığı sınıflamaya dönecek olursak ikinci tür yaratıcılık ölçme araçları kişilik envanterleridir. Birey tarafından doldurulabilen örnekleri olduğu gibi öğretmen, akran veya ebeveyn gibi bireyi yakından tanıyanlar tarafından da doldurulabilen çok çeşitli örnekleri vardır. Bu araçlarda yaratıcı kişilik özelliklerini tanımlayan maddelere bireylerin verdiği yanıtlar üzerinden değerlendirme yapılmaktadır. Bu testlerin kullanımındaki teorik duruş ise yaratıcılığın zekâ gibi bilişsel bir özellikten çok bir kişilik özelliği olarak değerlendirilmesidir. Yapılan bazı araştırmalarda kişilik ölçeklerinin yüksek ayırt edicilik ve tahmin geçerliği rapor edilmiştir (Feist, 1998; Carson, Peterson and Higgins, 2005;). Örneğin Gough (1979) tarafından geliştirilen Yaratıcı Kişilik Ölçeğinde (Creative Personality Scale) katılımcılar *akıllı, bireysellik, iç-görü, orijinallik, özgüven, sıradışılık* gibi özellikler ve bu özelliklerin kavramsal olarak tersi özellikler sıralanarak kendileri ile ilgili olanları işaretlemektedirler. Elde edilen veriler ile kişilik profilleri çıkartılmaktadır.

Diğer bir yaratıcılık ölçme yöntemi ise ürün değerlendirme yöntemidir. Günümüzde yapılan birçok yaratıcılık çalışmasında ölçümler katılımcıların ürünleri üzerinden yapılmaktadır. Bu yöntem gerçek performansın değerlendirilmesini de içerdiğinden birçok araştırmacı tarafından yeğlenmektedir. Alana özgü değerlendirmeye uygun olması yöntemin en önemli avantajlarından birisidir. Ancak değerlendirmeyi yapacak kişilerin özellikleri ve ekonomik kaygıları yöntemin sınırlılıkları arasında sayılabilir. Portfolyo değerlendirmesi ve proje yarışmaları bu kategori altında yer almaktadır. Yönteme ilişkin en önemli eleştiri ise sübjektif değerlendirmenin yapılmasıdır. Ancak Amabile (1982) geliştirdiği Konsensüs Ölçme Tekniği ile ürünlerin alan uzmanlarının sübjektif yaratıcılık görüşleri üzerinden değerlendirildiğinden ölçme geçerliğinin yüksek olduğunu vurgulamıştır. Bu tekniğe göre uzmanlar çok sayıdaki ürünü birbirleri ile kıyaslayarak yaratıcılıklarına göre gruplamaktadırlar. Bu gruplar daha sonra puana dönüştürülebilmektedir.

2.3.1. Bilimsel yaratıcılığın ölçülmesi

Genel yaratıcılık ölçümlerinde olduğu gibi bilimsel yaratıcılığın ölçülmesinde de benzer yöntemler ve araçlar kullanılmaktadır. Bu yöntemlerin başında çoğul düşünme

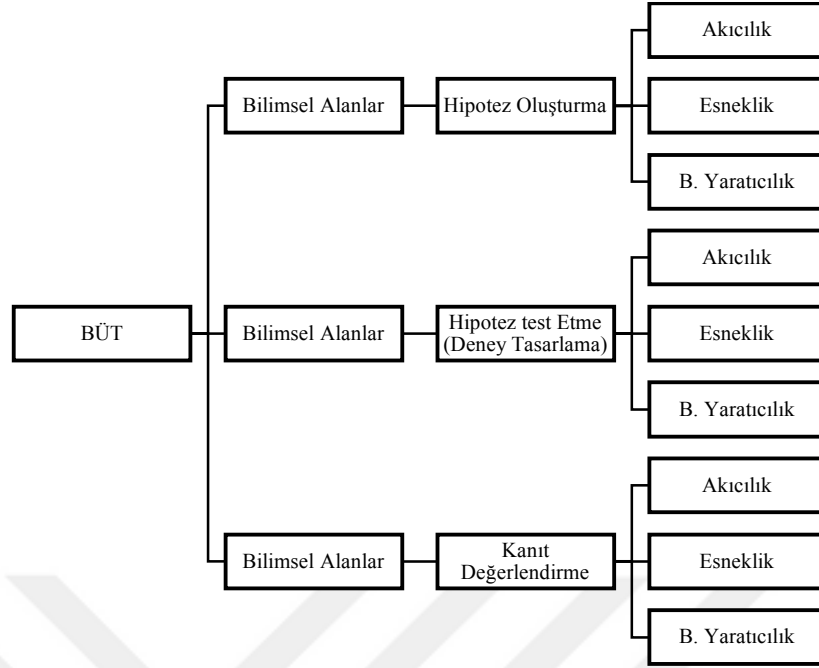
testleri, kişilik envanterleri ve ürün değerlendirme yaklaşımları gelmektedir. Müzik, resim, yazın gibi alana özgü diğer yaratıcılık türleri ve genel yaratıcılık düşünüldüğünde bilimsel yaratıcılığın ölçülmesine yönelik yöntemlerin sınırlılığı göze çarpmaktadır (Sak and Ayas, 2013; Hu and Adey, 2002, s. 392).

2.3.2. Bilimsel yaratıcılığın ölçülmesinde çoğul düşünme testlerinin kullanılması

Bilimsel yaratıcılığın ölçülmesinde çoğul düşünme testlerinin sayısı oldukça azdır. Alan yazında geçen testler teorik altyapı, ölçme modeli, geçerlik ve güvenirlik bakımından sınırlılıklara sahiptir. Hu ve Adey (2002) tarafından geliştirilen Liseler için Bilimsel Yaratıcılık Testi (Scientific Creativity Test for Secondary School Students-SCT), Ayas ve Sak (2014) tarafından geliştirilen Bilimsel Üretkenlik testi (BÜT), Kanlı (2014) tarafından geliştirilen Yaratıcı Bilimsel Çağrışımlar Testi (Y-BÇT), Chin ve Siew (2015) tarafından geliştirilen Okul Öncesi Öğrenciler İçin Görsel Bilimsel Yaratıcılık Testi (FBUYT) ve Siew, Chong ve Chin (2014) tarafından geliştirilen 5. Sınıflar için Bilimsel Yaratıcılık Testi (BYT-5) alan yazında var olan bilimsel yaratıcılık testleridir.

2.3.2.1. Bilimsel üretkenlik testi (BÜT)

BÜT, 6-8. sınıflar için geliştirilmiş bilimsel yaratıcılık testidir. Grup uygulama testi olan BÜT 40 dakikalık bir oturumda uygulanabilmektedir. Test Klahr'ın (2000) Bilimsel Yaratıcılığın İkili Arama Modeli (SDDS) temel alınarak geliştirilmiştir. İlk uygulaması 2008 yılında yapılmış ve ilerleyen yıllarda test üzerine yeni bilimsel araştırmalar yapılmış ve yayınlanmıştır. Test maddelerinin puanlanmasında akıcılık, esneklik ve bileşik yaratıcılık puanları hesaplanmaktadır. Akıcılık doğru yanıtların sayısını, esneklik ise yanıtların yer aldığı kavramsal kategorilerin sayısını ifade etmektedir. Bileşik yaratıcılık puanı ise akıcılık ve esneklik puanlarının Snyder ve arkadaşları (2004) tarafından önerilen bir bağıntı ile birlikte kullanılması elde edilmektedir.



Şekil 2.5. BÜT'ün teorik yapısı (Ayas and Sak, 2014, s. 198)

Ayas ve Sak (2014) tarafından SDDS modelini temel alarak geliştirilen BÜT'ün kuramsal yapısı Şekil 2.5'te verilmiştir. Kuramsal çerçevede hem fen bilimlerine özgü bileşenler hem de genel yaratıcılığa özgü bileşenler yer almaktadır. Fen bilimlerine özgü bileşenler bilimsel alanlar (fizik, kimya, biyoloji ve ekoloji) ve bilimsel becerilerdir (hipotez oluşturma, deney tasarlama, kanıt değerlendirme). Modelde yer alan akıcılık, esneklik ve bileşik yaratıcılık ise genel yaratıcılığa özgü bileşenlerdir. Bu kuramsal yapıya göre alt testlerin dağılımı Tablo 2.2'te verilmiştir.

Tablo 2.2. BÜT Teorik Yapısı (Sak and Ayas, 2013, s. 320)

Maddeler	Alanlar		
	Fen Alanları	Bilimsel Süreçler	Yaratıcılık Süreçleri
1. Sinek Deneyi	Biyoloji	Hipotez Oluşturma	Akıcılık, Esneklik, Yaratıcılık
2. Etkileşim grafiği	Disiplinler arası	Hipotez Oluşturma	Akıcılık, Esneklik, Yaratıcılık
3. Şeker Deneyi	Kimya	Hipotez Test Etme	Akıcılık, Esneklik, Yaratıcılık
4. Yay Deneyi	Fizik	Hipotez Test Etme	Akıcılık, Esneklik, Yaratıcılık
5. Besin zinciri	Ekoloji	Kanıt Değerlendirme	Akıcılık, Esneklik, Yaratıcılık

BÜT'te beş madde yer almaktadır. Bu maddeler fen bilimleri alanlarına göre fizik, kimya, biyoloji, ekoloji ve disiplinler arası alan olmak üzere sıralanmaktadır. Beş madde bilimsel süreçlere göre ise 2 deney tasarım (hipotez test etme), 2 hipotez geliştirme ve 1

kanıt değerlendirme görevleri olarak yer almaktadır. Sinek deneyinde öğrencilerden verilen bir deney düzeneğine uygun hipotezler üretmeleri istenirken, etkileşim grafiğinde değişime uygun etkilenen çiftleri üretmeleri istenmektedir. Şeker deneyinde öğrencilere verilen bir deney düzeneğini belirli bir hipotezi test edecek şekilde düzenlemeleri istenirken, yay deneyinde öğrencilerin verilen görevi yerine getirecek şekilde düzenekte değişiklik yapmaları istenmektedir. Besin zinciri maddesinde ise öğrencilere bir düzenek ve sonuç verilmekte, sonuç ve düzeneği karşılaştırarak kanıtları değerlendirmeleri istenmektedir. Testte öğrenci yanıtları yaratıcılık süreçlerinden akıcılık, esneklik ve bileşik yaratıcılık bağlamında değerlendirilmektedir. Örneğin “sinek deneyi” görevinde öğrencilerden *biyoloji* alanında *akıcı, esnek ve orijinal hipotezler geliştirmeleri* istenirken, “besin zinciri” görevinde öğrencilerden *ekoloji* alanında *akıcı, esnek ve orijinal* bir şekilde *kanıtları değerlendirmeleri* istenmektedir.

Testin geliştirilme sürecinde önce bir madde havuzu oluşturulmuş, çeşitli ön deneme ve pilot uygulamalarda maddeler denenmiş ve uzman görüşü alınarak testin son hali verilmiştir. Testin psikometrik özelliklerinin incelendiği araştırmalarda geçerlik ve güvenilirlik bulgularının yüksek olduğu rapor edilmiştir. Testin yapı geçerliği için temel bileşenler analizi (*Principal Component Analyses-PCA*) yapılmış, tek faktörlü bir yapı elde edilmiş ve toplam %34,45 varyans açıklanmıştır (Sak and Ayas, 2013). Yapı geçerliği için yapılan ikinci çalışmada tek faktörlü yapı doğrulayıcı faktör analizi (DFA) ile doğrulanmıştır (Ayas and Sak, 2014). Ayırt edicilik geçerliğine ilişkin çalışmalarda testin alt %27, orta %46 ve üst %27’lik grupları ayırt ettiği (Ayas and Sak, 2014); üstün yetenekli ve normal çocukları %87,5 oranında doğru sınıfladığı (Ayas, 2010); ve 7. sınıf üstün yetenekli öğrencilerin 6. sınıf üstün yetenekli öğrencilerden anlamlı şekilde daha yüksek puan aldıkları (Ayas ve Sak, 2008) rapor edilmiştir.

Testin güvenilirlik analizlerinde Cronbach Alpha iç tutarlık katsayısı hesaplanmış ve puanlayıcılar arası güvenilirlik analizlerinde Pearson korelasyon katsayısı hesaplanmıştır. Bu çalışmalarda testin iç tutarlık analizlerinde Cronbach Alfa değerleri .85-.87 arasında, puanlayıcılar arası güvenilirlik katsayıları ise .89-.98 arasında bulunmuştur (Ayas, 2010; Ayas and Sak, 2014). Farklı kültürlere adaptasyonu yapılan testin İspanyolca adaptasyonunda tek faktörlü yapı doğrulanmış, benzer yapı geçerliği bulguları rapor edilmiştir, farklı yetenek seviyelerinin (üst %27, %47 ve alt %27) BÜT puan ortalamaları arasında anlamlı bulunmuş, testin iç tutarlık katsayısının ve puanlayıcılar arası güvenilirlik katsayılarının yüksek olduğu rapor edilmiştir (Bermejo et. al., 2015).

BÜT kuramsal olarak SDDS modeline dayanması bakımından ümit vaat eden bir testtir. Ancak SDDS Modelinde yer alan 3 bileşen yapı geçerliği çalışmalarında elde edilememiştir. Testin iç tutarlık analizlerinde 5 maddenin akıcılık, esneklik ve bileşik yaratıcılık puanları analize birlikte dahil edilmiştir. Esneklik ve bileşik yaratıcılık puanlarının akıcılık puanlarından üretilmesi puanlar arasındaki korelasyonun yüksek olmasına sebep olmaktadır. İç tutarlık bulgularının yorumunda bu durum dikkate alınmalıdır. Ancak alternatif güvenilirlik bulguları (düzeltilmiş madde-toplam korelasyonları, maddeler arası ortalama korelasyon katsayıları) testin iç tutarlılığının tatmin edici seviyelerde olduğunu göstermektedir.

2.3.2.2. Liseler için bilimsel yaratıcılık testi (SCT)

Hu ve Adey (2002) tarafından geliştirilen SCT, SSCM modeli temel alınarak geliştirilmiş lise öğrencileri için bir kalem-kâğıt bilimsel yaratıcılık testidir. Test çoğul üretim modeline dayalı 60 dakikada uygulanabilen 7 maddeden oluşmaktadır. SSCM modeli üç boyutlu bir modeldir. Ürün boyu 4 bileşenden (bilimsel problem, bilimsel bilgi, bilimsel fenomen ve teknik ürün), süreç boyutu iki bileşenden (hayal gücü ve düşünme) ve kişilik özelliği boyutu üç bileşenden oluşmaktadır (akıcılık, esneklik ve orijinallik). Üç boyutlu bu yapı ile 24 farklı hücre kombinasyonu elde edilmektedir (3x4x2). Her bir hücre ise bilimsel yaratıcılığın ölçümüne yönelik farklı bir özelliği ifade etmektedir.

SCT'nin geliştirilme sürecinde SSCM'de yer alan 24 hücre için ikişer madde geliştirilmiş ve 48 maddelik bir soru havuzu oluşturulmuştur. Maddelerin lise öğrencilerinin bilimsel yaratıcılıklarını ölçmede kullanılıp kullanılmayacağına ilişkin Çinli 50 fen eğitimcisinin ve fen öğretmeninin görüşü alınmıştır. Eğitimcilerin en önemli bulduğu 9 madde seçilerek 60 kız öğrencinin katıldığı bir pilot uygulama yapılmıştır.

Testteki yedi maddeden altı tanesi, fen alanında olağan dışı kullanımlar, problem bulma, problem çözme, ürün geliştirme, hayal gücü ve deney tasarımı becerilerini çoğul düşünme modeline göre ölçerken son maddede fen alanında bir ürün tasarımı istenmektedir. İlk dört madde için akıcılık, esneklik ve orijinallik puanları hesaplanmaktadır. Akıcılık puanı doğru yanıt sayısı ile, esneklik puanı üretilen yanıtların kavramsal olarak kategoriler altında toplanması ile, orijinallik puanı ise frekansa dayalı olarak yanıtın rastlanma sıklığından (üst %5 için 2 puan, %5-%10 arası 1 puan, %10'dan daha fazla rastlanmış ise 0 puan) elde edilmektedir. 5. madde için sadece orijinallik puanı hesaplanırken (üst %5 için 3 puan, %5-%10 arası için 2 puan, diğerleri 1 puan) 6. maddenin puanı esneklik ve orijinallik puanlarının toplanması ile elde edilmektedir. 7.

maddenin puanlanmasında tasarlanan tek ürünün her bir özelliği için 3 puan verilerek akıcılık puanı elde edilmekte, orijinallik puanı için de 1-5 arasında uzman görüşüne göre değerlendirme yapılmaktadır. Teste verilen yanıtlardan öğrencilerin akıcılık, esneklik ve orijinallik puanları elde edilmektedir.

160 öğrenci ile yapılan uygulamada SCT'nin Cronbach Alpha iç tutarlılığı .893, puanlayıcılar arası güvenilirliği ise .793-.913 arasında bulunmuştur. En düşük puanlayıcılar arası güvenilirlik katsayısı ise ürün tasarım görevi olan son madde için hesaplanmıştır. Yapı geçerliği için temel bileşenler analizi (PCA) yapılmış ve 7 maddenin, tek bileşen altında toplanarak %63 varyans açıkladığı rapor edilmiştir. Maddelerin faktör yük değerlerinin ise yüksek olduğu görülmüştür. Testin ayırt edicilik çalışmalarında ise 7 ve 8. sınıf öğrencilerini birbirinden ayırt edemediği ancak 10 sınıfların her iki sınıf seviyesinden de anlamlı şekilde daha yüksek puanlar aldıkları belirtilmiştir. Testin yüz görünüm geçerliği için (*face validity*) İngiltere ve Çin'den 35 öğretmen ve fen eğitimcisinin görüşleri alınmış. Yüz görünüm geçerliği için eğitimcilere maddenin bilimsel yaratıcılığı ölçüp ölçmeyeceği sorulmuş ve evet-hayır yanıtlarının sıklığı/sayı üzerinden değerlendirme yapılmıştır. Ayrıca öğrencilere de testi ilginç bulup bulmadıkları sorularak öğrencilerin gözünden yüz görünüm geçerliği verisi toplanmıştır (Hu and Adey, 2002, s. 397-401).

Test geliştirme sürecinin kuramsal alt yapıda yer alan her bir bileşenin madde havuzunda düzgün bir şekilde temsil edilmesi önerilmektedir (DeVilles, 2003; Hinkin, 1995). SSCM'de yer alan 24 bileşenden her birinin sadece 2 madde ile temsil edilmesi madde havuzunun yeterince geniş olmadığı şekilden yorumlanabilir. Testin son halinde yer alan maddelerin çok genel maddeler olduğu düşünülebilir. Örneğin birinci maddede öğrencilerden *bir parça cam için cam için çok sayıda bilimsel kullanım şekli yazmaları* istenmektedir (örnek olarak *deney tüpü* verilmiştir). Ayrıca görüşüne başvuru uzmanların değerlendirmeleri nasıl yaptıklarına dair yeterli bilgi de verilmemektedir. Test geliştirme süreci ile ilgili diğer bir konu ise Çinli eğitimcilerin görüşlerinin alınmasına rağmen pilot uygulama ve asıl uygulamanın İngiltere'de yapılmış olmasıdır. Testin geçerlik ve güvenilirlik analizleri 160 öğrencinin verileri üzerinden yapılmış olması elde edilen bulguların güvenilirliğini düşürmektedir. Ayrıca modelde yer alan 3 boyut yerine PCA bulgularında 7 maddenin tek bileşen altında toplandığı rapor edilmiş ve elde edilen yapı ek kanıtlar ile (doğrulayıcı faktör analizi-DFA- gibi) desteklenmemiştir. SCT'nin geçerlik ve güvenilirliğine ilişkin ek araştırmalar yapmak yerine farklı kültürlerle

adaptasyonu yapılmış, fakt bu adaptasyonlarda da geçerlik ve güvenilirliğine ilişkin sınırlı bulgular rapor edilmiştir (örneğin Kadayıfçı, 2008). Testin puanlamasında bir standardın olmadığı görülmektedir. Bu tür sınırlılıklarından dolayı SCT prototip bir test olarak değerlendirilebilir. Ancak testin belirli bir teorik alt yapıya göre hazırlanmış olması güçlü yanı olarak değerlendirilmelidir.

2.3.2.3. Okulöncesi öğrenciler için görsel bilimsel yaratıcılık testi (FBYT)

SSCM temel alınarak okul öncesi öğrenciler için geliştirilen FBYT, figüral bir testtir. Testin geliştirilme sürecinde SSCM (Hu and Adey, 2002) revize edilerek özellik boyutundan esneklik bileşeni çıkartılarak *detaylandırma*, *başlıkların soyutluğu* ve *erken kapanmaya direnç* bileşenleri eklenmiştir. SSCM'nin ürün boyutundan ise *teknik ürün* bileşeni çıkartılmıştır (Chin and Siew, 2015). Okulöncesi müfredatında yer alan kazanımlar temel alınarak revize edilmiş SSCM modeline göre geliştirilmiştir. Testte 6 görsel maddeden oluşan bir kalem-kâğıt testidir. Test geliştirme sürecinde uzman görüşü alınarak test oluşturulmuştur. Test maddelerinden 2 tanesi bilimsel bilgi, 2 tanesi bilimsel problem ve 2 tanesi de bilimsel fenomen bileşenleri ile ilgilidir. Testin puanlanmasında 4 madde için akıcılık, esneklik, orijinallik puanları SCT'de olduğu gibi puanlanmaktadır. Modele sonradan eklenen detaylandırma 1-5 detay için 1 puan, 6-12 detay için 2 puan ve 13-19 detay için 3 puan şeklinde hesaplanmaktadır. Başlıkların soyutluğu ise betimleyici, hayali ve soyut olmak üzere sırasıyla 1, 2 ve 3 puan ile değerlendirilmektedir. Erken kapanmaya karşı direnç ise şeklin karmaşıklığına göre 0, 1 veya 2 puanla değerlendirilmektedir. 2 maddenin puanlamasında ise erken kapanmaya karşı direnç puanları hesaplanmamaktadır.

Geliştirilen test 30 okulöncesi öğrencisine uygulanmış ve geçerlik ve güvenilirlik analizleri yapılmıştır. Testin Cronbach Alpha iç tutarlık güvenilirliği .806 ve puanlayıcılar arası güvenilirliği .780-.933 arasında bulunmuştur. Testin yapı geçerliği çalışmasında 6 madde tek faktör altında toplanarak toplamda %50,85 varyans açıklanmıştır. Testin alt ve üst %27'lik grupları ayırt etme gücü ise yeterli seviye bulunmuştur.

FBYT okul öncesi öğrenciler için geliştirilmiş bir kalem-kâğıt testi olduğundan sınırlı okuma ve yazma becerisine sahip öğrencilere uygun görünmemektedir. Örneğin bu testin Türk kültüründe uygulanması mümkün görünmemektedir. SCT'de olduğu puanlama standartı oluşturulmamıştır. Geliştirilme sürecinde sadece 30 öğrencinin verilerinin kullanılmış olması bulguların güvenilirliğini düşürmektedir. SSCM modeline sonradan eklenen detaylandırma ve erken kapanmaya direnç bileşenlerinin nasıl

yorumlanacağı ise yazarlar tarafından açıklanmamıştır. Ayrıca test geliştirme sürecine ilişkin detaylı bilginin yer almadığı çalışmada madde geliştirme sürecinde sadece 6 madde üretilmesi eleştirilebilir. Sıralana eksikliklerinden dolayı FBYT prototip bir test olarak gelişiminin ilk aşamalarında olduğu düşünülebilir.

2.3.2.4. Beşinci sınıflar için bilimsel yaratıcılık testi (BYT-5)

SSCM modeli temel alınarak Siew, Chong ve Chin (2014) tarafından 5. sınıf öğrencileri için geliştirilen BYT-5'nin, 4'er maddelik iki paralel formu vardır. Test geliştirme sürecinde SSCM'de yer alan her hücre için bir madde geliştirilmiş (24 madde) ve fen bilgisi öğretmenlerinin görüşleri alınmıştır. Pilot uygulama sonrasında teknik ürün, bilgi, bilimsel fenomen ve bilimsel problem bileşenlerinin 5. sınıf öğrencileri için daha uygun olduğu sonucuna ulaşılmıştır. Testten SCT gibi akıcılık, esneklik ve orijinallik puanları elde edilmektedir. İlk üç madde için akıcılık, esneklik ve orijinallik puanları toplanmakta, dördüncü madde için esneklik ve orijinallik puanları toplanarak öğrencilerin toplam puanları hesaplanmaktadır.

BYT-5, 206 beşinci sınıf öğrencisine uygulanarak geçerlik ve güvenilirlik analizleri yapılmıştır. Geçerlik bulgularına göre maddelerden bir tanesinin alt ve üst %27'lik grupları ayırt etmediği rapor edilmiştir. Yapı geçerliği için yapılan PCA bulgularında A ve B formları için iki faktörlü yapılar bulunmuş ve sırasıyla %61 ve %65 varyans açıklayan yapılar elde edilmiştir. Cronbach Alpha iç tutarlık katsayıları A formu için .77 ve B formu için .68 olarak hesaplanmıştır. Testin okuyucular arası güvenilirlikleri için hesaplanan Pearson momentler çarpımı korelasyon katsayıları ise .652 ve .984 arasında değişmektedir.

BYT-5 aynı kuramsal alt yapıya sahip olmasına rağmen SCT ve FBYT'den farklı olarak PCA bulgularına göre iki faktörlü bir yapıya sahiptir. Test geliştirme çalışmalarında modelde yer alan 24 hücre için birer test maddesi hazırlanmış ve sadece öğretmenlerin görüşlerine göre 4 madde seçilmiştir. SCT, FBYT ve BYT-5'in geliştirilme sürecinde ön deneme veya pilot uygulamaya rastlanmazken sadece uzman görüşüne göre madde seçimi yapılmıştır. Ayrıca SSCM temel alınarak geliştirilen testlerin yapı geçerliği çalışmalarında farklı yapılar elde edilmiş ve bu yapılardan hiç birisinde 3 boyutlu model elde edilememiştir. Bu bakımdan SSCM modelinin ve bu model temel alınarak geliştirilen testlerin prototip testler olarak kabul edilmesi daha uygun olabilir.

2.3.2.5. *Yaratıcı bilimsel çağrışımlar testi (Y-BÇT)*

Kanlı (2014) tarafından Mednick'in (1962) Çağrışımsal Modeline göre Yaratıcı Bilimsel Çağrışımlar Modeli'ni (Y-BÇM) oluşturmuş ve bu modeli temel alarak Yaratıcı Bilimsel Çağrışımlar Testi'ni (Y-BÇT) geliştirmiştir. Y-BÇT ortaokul öğrencilerine uygun bir çoğul düşünme kalem-kâğıt bilimsel yaratıcılık testidir. Y-BÇT üç bileşenden oluşmaktadır: çağrışımlar, analogik muhakeme ve analogik problem çözme. Bu üç alandaki problemler aracılık ve benzerlikler yolu ile çözülmektedir. Testin çağrışımlar boyutunda 10 madde, analogik muhakeme boyutunda 10 madde ve analogik problem çözme boyutunda ise 5 madde olmak üzere Y-BÇT'de toplam 25 madde yer almaktadır. Testin puanlanmasında sadece akıcılık puanları hesaplanmaktadır.

Teste yer alan çağrışımlar boyutunda öğrencilere fen bilimleri alanlarından birtakım kavramlar verilmekte, daha sonra ise bu kavramlarla ilişkili yeni kavramlar bulmaları istenmektedir. Testin analogik muhakeme alt boyutunda öğrencilere birbirleri ile ilişkili iki kavram verilmekte ve üçüncü bir kavram ise tek başına verilmektedir. Öğrencilerden üçüncü kavram ile benzer ilişki kurabilecekleri dördüncü bir kavram bulmaları istenmektedir. Testin analogik problem çözme boyutunda ise öğrencilere verilen iki problem durumu arasında analogi kurarak problemi çözmeleri istenmektedir.

Test geliştirme sürecinde 50 maddelik bir havuz oluşturulmuştur. Havuzdaki maddeler iki aşamada alan uzmanlarının görüşleri alınarak revize edilmiştir. Uzman görüşü sonucunda 29 madde elde edilmiş ve küçük bir grupta ön deneme uygulaması yapılmıştır. Ön deneme uygulaması sonunda 27 madde ile 160 öğrencinin katılımıyla pilot uygulama yapılmıştır. Pilot uygulamada açımlayıcı faktör analizi, geçerlik ve güvenilirlik analizleri yapılarak testte çeşitli revizyonlar yapılmıştır. 385 normal ve 293 üstün zekâlı öğrenci ile asıl uygulama yapılmıştır.

Yapılan asıl uygulamada Y-BÇM'de öngörülen 3 boyut elde edilmiş (açıklanan varyans %47,35) ve DFA ile bu boyutlar doğrulanmıştır. Testin puanlayıcılar arası güvenilirlik katsayıları .973-.984 arasında değiştiği, Cronbach Alpha iç tutarlık katsayısının .890 olduğu ve test-tekrar test güvenilirlik katsayısının .935 olduğu rapor edilmiştir. Testin ayırt edicilik geçerlik çalışmalarında testin bazı boyutlarında 5. ve 6. sınıfların, bazı boyutlarında ise 6. ve 7. sınıfların puan ortalamalarının farklılaşmadığı görülmüştür.

Y-BÇT kuramsal alt yapıya göre geliştirilmiş ve bu kuramsal alt yapı faktör analizi bulguları ile desteklenmiştir. Ancak AFA ve DFA aynı veri seti üzerinden yapıldığından

ek kanıtlarla desteklenmesi gerektiği düşünülebilir. Testte yer alan çoğul üretime dayalı 25 madde için 50-75 dakikanın kısa olduğu düşünülebilir. Testin puanlanmasında sadece akıcılık puanlarına yer verilmesi Y-BÇT'nin bir sınırlılığı olarak düşünülebilir. Mednick'in (1962) Uzak Çağrışımlar Testinde çağrışımların uzaklığı ile yaratıcılık arasında pozitif bir ilişki olduğunu belirtmektedir. Bu bağlamda Y-BÇT'nin puanlanmasında yanıtların uzaklığının hesaba katılmasının gerekliliği düşünülebilir. Y-BÇM'de analogiler, çağrışımlar aracılıklar ve benzerlikler yolu ile gerçekleşmektedir. Ancak bu bileşenlere modelde yer verilmemiştir.

2.3.3. Bilimsel yaratıcılığın ölçülmesinde kişilik envanterleri

Kişilik envanteri türünde bilimsel yaratıcılığın ölçülmesine yönelik bir ölçme aracına alan yazında rastlanmazken genel yaratıcılığın ölçülmesinde kullanılan ölçme araçlarında alt test veya bilimsel yaratıcılıkla ilgili maddelere yer verilmektedir. Runco (1987) tarafından geliştirilen Yaratıcı Aktiviteler Kontrol Listesi, farklı alanlarda yaratıcılık ölçümlerinde kullanılan popüler bir araçtır (Liang, 2002). Farklı yaş grupları için uyarlanmış araçta edebiyat, fen, müzik, drama, sanat ve teknik alanlarda 50 maddeden yer almaktadır. Maddelerde herhangi bir yaratıcı aktiviteyi bireyin gösterme sıklığı sorulmaktadır. Her alt alan birer puan ve ölçek toplam puanı hesaplanmaktadır. Listede verilen durumlar asla, 1 ya da 2 defa, 3-5 arası, 5-10 arası, 10'dan fazla şekilde değerlendirilmektedir. Bilimsel yaratıcılıkla ilgili olarak hazırlanan alt testte yer alan konu başlıkları şunlardır: ödül kazanmış ürünler, yayınlanmış makaleler, patentler, burslar, olimpiyat kamplarına katılım, deney tasarımı, bilimsel cihaz yapımı, hayvan ve bitki incelemeleri, bilimsel yarışlara katılım, kabul görmüş projeler, yaz bilim kamplarına davet, bilimsel kulüplere üyelik, bilimsel kulüp başkanlığı ve bilimsel dergi kitap ve makale okuma şeklinde sıralanmaktadır.

2.3.4. Bilimsel yaratıcılığın ölçülmesinde uzman değerlendirmesi

Amabile'nin (1983) Bileşensel Yaratıcılık Modelinde ve Csikszentmihalyi'nin (1997) Sistemler Modelinde yaratıcı ürünlerin sosyal bağlamında değer kazandığı vurgulanmaktadır. Modellerde belirtilen sosyal çevreyi çalışma alanındaki uzmanlar oluşturmaktadır. Herhangi bir ürünün yaratıcı ürün olarak değer kazanması için alanı en iyi tanıyan bu uzmanların onayına gerek vardır (Csikszentmihalyi, 1997). Günümüzde Nobel Ödülleri gibi bilim dalındaki prestijli ödüller veya Oscar gibi sanat alanlarında verilen ödüller, uzman değerlendirmesine örnek olarak verilebilir.

Yaratıcılık ölçümünde kullanılan uzman değerlendirmesinin özel bir türü olan, Amabile (1982) tarafından geliştirilmiş *konsensüs ölçme tekniği* (Consensual Assessment Technique-CAT), alana özgü yaratıcılık ölçümlerinde yeni yeni kullanılmaya başlanılan bir yöntemdir. CAT temelde, ürün değerlendirme yaklaşımı olarak, yaratım alanındaki uzmanların ürünleri değerlendirmesine dayanmaktadır. Yaratıcılık ölçümlerinde “altın standart” olarak değerlendirilmesinden dolayı, yaratıcılık ölçümlerinden kabul gören bir yaklaşımdır (Kaufman, Baer, Cole ve Sexton, 2008).

Uzman değerlendirmesinde katılımcıların ürünleri uzmanlara dağıtmakta ve yaratıcılıklarına göre ürünlerin puanlanması istenmektedir. Ürünlerin değerlendirilmesi iki farklı şekilde yapılabilmektedir. Birinci yöntemde uzmanlar tüm kâğıtları belirlenen puanlama sistemine göre tüm ürünleri puanlamaları istenebilir. İkinci yöntemde ise uzmanlar ilk aşamada ürünleri “sıradan”, “yaratıcı” ve “çok yaratıcı” olarak gruplayıp ikinci aşamada her bir gurubu kendi içerisinde tekrardan gruplamaları istenebilir. Uzmanlar bir defa ürünleri puanladıktan her bir ürün için toplam yaratıcılık puanı veya ortalama yaratıcılık puanı hesaplanarak değerlendirme sonlandırılabilir. Bu yöntem ile herhangi bir ürün için birden fazla uzmanın görüşü değerlendirmeye katılmış olmaktadır. CAT uygulamalarında öznel bir değerlendirme yapılsa da yöntemin özünde öznel görüşlerin sayısal olarak ifade edilmesi yatmaktadır (Nusbaum ve Silvia, 2011, s. 38).

Yukarıda basitçe anlatılan CAT uygulamasında araştırmacıların dikkat etmeleri gereken bazı noktalar vardır. Öncelikle uzmanların öznel değerlendirme yapmaları çok önemlidir. Bunun için uzmanların birbirlerinden bağımsız şekilde puanlama yapması gerekmektedir. Değerlendirmenin öznel olabilmesi için uzmanlara herhangi bir eğitim verilmemeli, puanlama kriterleri konusunda sorgulanmamalı, aksine uzmanlar öznel değerlendirme yapmaları konusunda cesaretlendirilmelidir. CAT uygulamasında önemli olan diğer bir nokta ise değerlendirilmeye alınan ürünlerin birbirleri ile karşılaştırarak değerlendirmeleridir. Farklı bir ifade ile CAT’de bağıl bir değerlendirme yapılması gerekmektedir. Bundan dolayı CAT’den elde edilen puanlar sadece grup içindeki farklılığın bir göstergesi olup dış bir kriterden bağımsızdır. Ayrıca bu değerlendirme sadece ürünle ilgili olup, teorik olarak yaratıcılıkla ilgisi olduğu düşünülen bilişsel beceriler, kişilik özellikleri veya diğer bileşenlerden değerlendirme dışıdır. Bahsedilen bağıl değerlendirmenin sağlanabilmesi için ürünler puanlanırken, tüm skalayı kullanılmalıdır. Örneğin 5’li puanlama sistemi kullanılacaksa uzmanların 1, 2, 3, 4 ve 5 puanı en az birer defa kullanması gerekmektedir.

CAT uygulamalarında uzman gurubunun 10 ve daha fazla uzmandan oluşması önerilmektedir (Kaufman et. al., 2008, s. 172). Bu durum “kimler uzman olarak kabul edilebilir?” sorusunu akla getirmektedir. Hennesey ve Amabile’a (1999, s. 351) göre değerlendirme yapılacak alana yabancı olamayan alanda çalışan öğretmenler, üniversite çalışanları veya meslek erbapları olabilir uzman olarak sürece dahil edilebilir. Bazı araştırmacılar ise uzmanlar yerine akran değerlendirmesi kullanırken (Runco, 1989), bazıları uzman olmayan kişileri değerlendirme ekibine dahil etmektedir (Nui and Sternberg, 2001). Bu bağlamda CAT uygulamalarında uzmanın kimler olabileceği ve uzman seçiminde ne tür kriterler kullanılacağına ilişkin net bir bilgi bulunmamaktadır. Uzmanlarla ilgili ikinci bir konu ise uzmanlar arası tutarlıktır. Hennesey ve Amabile (1999) tutarlılığın sağlanması için uzmanların tecrübelerinin birbirlerine yakın olması gerektiğini vurgulamaktadır. Alan yazında CAT’de uzmanlar arası tutarlılığın araştırıldığı birçok çalışmada yüksek uzmanlar arası tutarlık katsayıları rapor edilmiştir (Amabile, 1982; Baer, Kaufman and Gentile, 2004; Hickey, 2001; Kaufman et. al., 2008). Ancak CAT uygulamasının geçerliği uzman ekibin özellikleri ile doğrudan ilişkilidir. Tutarlık ise güvenilirliğe ilişkin bir sorun olup sadece veriler üzerinden değerlendirilmektedir. Uzman olmayan bireyler arasında yüksek tutarlık oranlarının rapor edildiği çalışmalar olsa da (Ayas, 2013; Runco, 1994; Nui and Sternberg, 2001) bu araştırmaların geçerlikleri sorgulanmalıdır.

Resim, müzik, yazın ve genel yaratıcılık gibi birçok alanda CAT uygulamasına rastlanmaktadır. Bilimsel yaratıcılık ve çoğul düşünme testlerinde kullanımına ilişkin uygulamalar ise sınırlıdır. Ayas (2013) çoğul üretime dayalı bilimsel yaratıcılık testi olan Bilimsel Üretkenlik Testi (BÜT) kullanarak CAT uygulaması yapmıştır. Uzman ve uzman olmayanların yaratıcılık puanlamalarının karşılaştırıldığı çalışmanın uzman ekine fen bilimleri alanında çalışan 12 akademisyen dahil edilmiştir. Uzman olmayan ekip ise sosyal bilimler alanında çalışan 12 akademisyenden oluşmaktadır. Araştırmada 286 altıncı sınıf öğrencisine BÜT uygulanmış ve öğrencilerin ürettiği 234 farklı yanıt BÜT formuna kaydedilmiştir. Daha sonra puanlayıcı gruplara test maddesi ve maddeye üretilen yanıtlar birlikte verilmiştir. Puanlayıcılardan her bir yanıtın yaratıcılık düzeyini 1-10 arasında puanlanması istenmiştir. Araştırmanın sonucuna göre uzman olmayan değerlendiriciler arasındaki tutarlılık daha yüksek olmasına rağmen, uzman ekibin puan varyanslarının daha düşük olduğu görülmüştür. Ancak uzmanlardan 3 katılımcının analiz dışı bırakılması durumunda tutarlık oranlarının .57’den .86’ya çıktığı rapor edilmiştir.

Yapılan analizlerde uzman olmayanların orijinal olan yanıtlara daha yüksek puan verdikleri, uzmanların ise alana uygun olan yanıtlara daha yüksek puan verdikleri görülmüştür. Araştırmada uzmanlar arası tutarlılığın düşük çıkmasının başlıca nedeni ise tecrübe farklılığı ve 3 farklı alandan uzmanların aynı ekibe dahil edilmesine bağlanmıştır.

CAT gibi bilim ve proje yarışmaları uzmanlar tarafından ürünlerin değerlendirilmesine dayandığından bir tür ürün değerlendirme yaklaşımı olarak ele alınabilir. Intel Bilim Yetenek Yarışması bu gruba örnek olarak verilebilir. Yarışmaya lise öğrencileri fizik, kimya, biyoloji, matematik, psikiyatri, psikoloji ve sağlık alanlarında orijinal bir araştırma ve bu araştırmanın yazılmış makalesi ile katılmaktadır. Değerlendirmede çalışmalar “A: aşırı derecede orijinal; D: sıradan; E: zayıf deneme” şeklinde sıralanır. Bu sıralama nümerik olarak 14,00 (A) ile 1,00 (E) arası puanlarla değerlendirilir. Feist’in (2006) yaptığı çalışmaya göre finalistlerin %81’i doktora, %10’ u mastır ve %9’ u ise lisans eğitimlerini tamamlamışlardır. Ayrıca 1942-1999 yılları arasındaki finalistler 5 Nobel Ödülü, 2 Field Madalyası, 3 Ulusal Bilim Madalyası (*National Medal of Science*), 9 MacArthur Vakıf Madalyası, 56 Sloan Araştırma Bursu, 30 Ulusal Bilimler Akademisi (*National Academy of Sciences*) üyeliği ve 4 Ulusal Mühendisler Akademisi (*National Academy of Engineering*) üyeliğine sahip olmuşlardır.

Bilimsel yaratıcılığın ölçülmesi ile ilgili yapılan uygulamalara bakıldığında kişilik envanterleri veya kontrol listelerinin geçerlik ve güvenilirlik sorunlarının olduğu bilinmektedir. Ayrıca bu tür araçların genel yaratıcılık ölçümlerine daha uygun olduğu önerilmektedir (Kaufman and Baer, 2004; Runco, 1987; Plucker, 1998). Resim, müzik ve yazın gibi alanlarda daha etkili olan performans ölçümlerinin ise bilimsel yaratıcılık ölçümünde kullanımı sınırlıdır. Uzman değerlendirme yaklaşımı veya CAT uygulamaları geçerlik ve güvenilirlik açısından önerilen yöntemler olmakla birlikte çeşitli sınırlılıkları beraberinde getirmektedir. Örneğin CAT uygulamasında 10 veya daha fazla uzman gerektirmesi, bu uzmanların alan tecrübelerinin benzer olması, uzman belirleme kriterlerinin net olmaması, bu uzmanlarla yapılacak değerlendirmenin sınırlı sayıdaki veri üzerinden yapılabilir olması, uzmanlar arasındaki tutarlık gibi konular araştırmacıların baş etmeleri gereken sınırlayıcılarıdır.

Bilimsel yaratıcılığın ölçülmesinde yaratıcılık testlerinin (çoğul düşünme testleri) kullanımı, yöntemin kendisinden çok testlere ilişkin sınırlılıklara sahiptir. Örneğin yukarıda incelenen bilimsel yaratıcılık testleri belirli bir kuramsal model çerçevesinde geliştirilmiş testlerdir. Ancak bu kuramsal yapıların teste ilişkin araştırma bulguları ile

desteklenmediđi görlmektedir. İkinci olarak bu kuramsal modellerin çođu bilimsel yaratıcılıđı açıklamaktan uzak, test geliştirme sürecine rehberlik eden modellerdir. Örneđin Y-BÇM ve SSCM modelleri sadece geliştirilen testlerin geçerlik çalışmaları ile desteklenmiş modellerdir. SDDS modeli bu bakımdan diđer modellerden ayrılmakla birlikte, bu modeli temel alan BÜT çalışmalarında yapı desteklenememiştir. Farklı bir ifade bu testlerin teorilerle uyumlu olmadığı düşünlebilir. Bilimsel yaratıcılık testleri ile ilgili diđer sorunlar ise kapsam ve hedef kitleye ilişkin sınırlılıklarıdır. Farklı eğitim düzeylerine bakıldığında ilkokul öğrencilerine yönelik bir bilimsel yaratıcılık testinin olmadığı görlmektedir.



3. YÖNTEM

Bu bölümde çalışmanın araştırma modeli, araştırmanın çalışma grupları, veri toplama araçları, araştırmanın odağında yer alan Bilimsel Üretkenlik Testi 3, 4 ve 5. Sınıf Formunun geliştirilme sürecine, veri çözümlene ve istatistiksel analiz basamaklarına yer verilmiştir.

3.1. Araştırma Modeli

Bu çalışma bir yaratıcılık testi geliştirme çalışmasıdır. Çalışmanın amacı 3, 4 ve 5. sınıf öğrencilerinin fen bilimleri alanında yaratıcılık potansiyellerini değerlendirmek amacıyla Klahr ve Dunbar'ın (1988) Bilimsel Yaratıcılığın İkili Arama Modeline dayalı Bilimsel Üretkenlik Testi (BÜT 3-5) geliştirilerek psikometrik özelliklerini incelemektir. Bu bakımdan araştırmada tarama modeli kullanılmıştır. Tarama modelleri, var olan bir durumu herhangi bir müdahalede bulunmadan betimlemeyi amaçlayan (Karasar, 2009, s.77), bir konuya ilişkin katılımcıların ilgi, tutum, beceri ve yetenek gibi özelliklerinin betimlendiği (Fraenkel, Wallen and Hyun, 2012, s. 393), genellikle büyük örneklerde çalışan araştırma modelidir (Büyüköztürk ve diğerleri, 2010). Bu bağlamda çalışmanın modeli kesitsel tarama modelinde betimsel ve bağlantısal bir çalışma olarak tasarlanmıştır.

3.2. Çalışma Grubu

Araştırmanın çalışma grubunun belirlenmesinde seçkisiz olmayan örnekleme türlerinden uygun örnekleme yöntemi kullanılmıştır. Bu örnekleme yönteminde evren içerisinden uygun olan birimler araştırmaya dâhil edilir (Fraenkel, Wallen ve Hyun, 2012, 100). Uygun örnekleme yöntemi zaman, işgücü ve maliyet açısından var olan sınırlılıkların üstesinden gelebilmek amacıyla (Büyüköztürk ve diğerleri, 2010, s. 177) katılımcıların erişim kolaylığı ve uygunluğu nedeniyle başvurulan (Cresswell, 2013, 158) uygulamada ekonomiklik ve pratiklik sağlayan bir yöntemdir (Böke, 2010, s. 112).

Araştırma 2015-2017 yılları arasında yürütülmüştür. Araştırmanın veri toplama aşaması ise Eskişehir il merkezinde, 2016-2017 eğitim öğretim yılı bahar yarıyılında dönemin ortalarında (Mart-Nisan ayları) gerçekleştirilmiştir. Araştırmaya Eskişehir il merkezindeki iki farklı ilkokula ve bir ortaokula devam eden toplam 647 öğrenci dâhil edilmiştir. Araştırmaya katılan öğrencilerin sınıf düzeyi ve cinsiyete göre dağılımları Tablo 3.1'de verilmiştir.

Tablo 3.1. Katılımcıların sınıf düzeylerine ve cinsiyete göre dağılımı

Sınıf Düzeyi	Cinsiyet	F	%	% Yığılımı
3	Erkek	130	51,8	38,8
	Kız	121	48,2	
	Toplam	251	100	
4	Erkek	122	55,5	34,0
	Kız	98	44,5	
	Toplam	220	100	
5	Erkek	66	37,5	27,2
	Kız	110	62,5	
	Toplam	176	100	
Toplam	Erkek	318	49,1	100
	Kız	329	50,9	
	Toplam	647	100	

Tablo 3.1’de görüldüğü üzere araştırmaya katılan öğrencilerden 328’inin kız (%50,9) ve 318’inin (%49,1) erkek öğrencilerden oluştuğu görülmektedir. Katılımcıların 251’i (%38,8) 3. sınıf, 220’i (%34,0) 4. sınıf ve 176 tanesi (%28,2) 5. sınıf öğrencileridir.

3.3. Veri Toplama Araçları

Bu çalışma bir test geliştirme çalışması olduğundan birincil olarak kullanılan veri toplama aracı geliştirilen BÜT 3-5’tir. Geliştirilen bu testin psikometrik özelliklerini ortaya koymak için öğrencilerin notları, fen bilimlerine yönelik tutumları ve fen bilimlerine yönelik akademik benlik algıları ile BÜT 3-5’in ilişkisine bakılmıştır. Bu süreçte karne notları, fenne yönelik tutum ölçeği ve fende akademik benlik algısı ölçeği kullanılarak veriler toplanmıştır.

3.3.1. Bilimsel üretkenlik testi 3-5 (BÜT 3-5)

Bu çalışma ile geliştirilen BÜT, 3, 4 ve 5. sınıf seviyesindeki öğrencilere uygun bir bilimsel yaratıcılık testidir. BÜT 3-5, SDDS (Klahr, 2000) temel alınarak çok aşamalı bir yaklaşımla geliştirilmiştir. BÜT 3-5’in geliştirilme, geçerlik ve güvenilirlik çalışmaları tamamlandıktan sonra, son sürümü iki bileşenden oluşmuştur: deney tasarımı ve hipotez geliştirme. Her bileşende ikişer tane olmak üzere toplam 4 madde yer almaktadır. BÜT 3-5’in geliştirilme süreci izleyen kısımda verilmiştir.

3.3.1.1. BÜT 3-5 geliştirilme süreci

BÜT 3-5 geliştirilme çalışmalarında DeVilles (2003) tarafından önerilen ölçek geliştirme basamakları takip edilmiştir:

- BÜT 3-5'in kuramsal dayanaklarının belirlenmesi
- Alan uzmanlarından oluşan bir ekip ile çok sayıda test maddesinin geliştirilmesi
- Madde havuzundan uygun maddelerin seçilerek taslak formun oluşturulması
- Ön deneme uygulaması, revizyon yapılması, madde seçimi ve pilot formun oluşturulması
- Pilot uygulama ve revizyon yapılması, madde seçimi ve asıl formun oluşturulması
- Büyük örneklem ile asıl uygulamanın ve revizyon yapılması, madde seçimi ve son formun oluşturulması
- Son formun psikometrik özelliklerinin incelenmesi

Yukarıda sıralanan işlem basamaklarından ilk beşi test geliştirme basamakları olarak aşağıda detaylı şekilde açıklanmıştır. Son iki basamakta yer alan büyük örneklem ve psikometrik analizler ise bulgular bölümünde yer verilmiştir.

1.1.1.1.1. BÜT 3-5'in kuramsal yapısı

Ölçek geliştirme üzerine yapılmış birçok yayında sürecin ilk aşamasında kavramsal çerçevenin oluşturulması önerilmektedir (Hinkin, 1995, s. 969; Clark and Watson, 1995, s. 310; Domino and Domino, 2006, s. 159; Kaplan and Saccuzo, 2009, s. 159). BÜT 3-5 ilkokul ve ortaokul öğrencilerinin bilimsel yaratıcılıklarını değerlendirmek için geliştirilmiş bir test olduğundan kuramsal çerçevenin oluşturulmasında yaratıcılık, bilimsel yaratıcılık ve hedef gurubun özellikleri dikkate alınarak kapsamlı bir alan yazın çalışması yapılmıştır. Aşağıda alan yazın çalışmasından elde edilen bulgulardan hareketle geliştirilmiş kuramsal yapı detaylı şekilde verilmiştir.

Bilimsel yaratıcılık genel yaratıcılıkla ilgili beceriler, bilimsel beceriler ve kişilik özellikleri gibi birçok farklı bileşenden oluşmaktadır (Dunbar, 1999; Klahr, 2000; Puccio, 1991; Roe, 1952, 1961; Subotnik, 1993; Torrance, 1992). Çoğul düşünme üzerine yapılan araştırmalar (Guilford, 1950, 1956, 1967; Torrance, 1962, 1988) ve SDDS (Klahr, 2000) temel alınarak 6, 7 ve 8. sınıf öğrencilerinin bilimsel yaratıcılıklarını değerlendirmek amacıyla geliştirilen BÜT'ün (Ayas and Sak, 2014; Sak and Ayas, 2013) kuramsal altyapısı bu çalışma kapsamında geliştirilen BÜT 3-5'in kuramsal çerçevesini

oluşturmaktadır. Bu süreçte BÜT'ün kuramsal yapısı ile SDDS modeli ile tekrar karşılaştırılarak revize edilmiş ve BÜT 3-5'in kuramsal alt yapısı geliştirilmiştir.

Bu çalışma kapsamında SDDS modeli ve BÜT test modeli daha detaylı incelenmiş ve yeniden yorumlanmıştır. Bu süreçte *kanıt değerlendirme* bileşenine BÜT 3-5 geliştirme sürecinde yer verilmemiştir. Kanıt değerlendirme bileşeni aşağıda sıralanan üç nedenden dolayı kuramsal çerçeveden çıkartılmıştır.

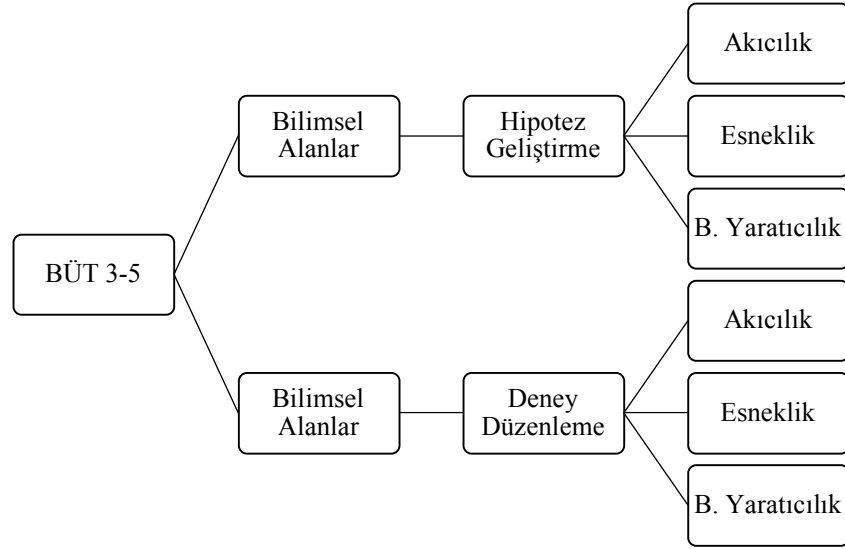
1. *Kuramsal* olarak ele alındığında kanıt değerlendirme bileşeninin SDDS modelinde diğer iki bileşen (hipotez oluşturma ve deney tasarlama) arasında arabulucu bir bileşen olduğu ve her iki alandaki aramalara öncülük ettiği görülmektedir Kanıt değerlendirme sonucunda buluş süreci hipotez alanından veya deney alanından tekrardan başlamaktadır (Klahr, 2000, s. 38). Farklı bir ifade ile kanıt değerlendirme süreci sonucunda ya yeni hipotezler oluşturulmakta ya da deney tasarımında değişiklikler yapılmaktadır. Bu bağlamda kanıt değerlendirme sürecinin temsilinin deney tasarımları veya hipotezler olarak gözlenebileceği düşünülebilir (Zimmerman, 2000, s. 104). Bullock'un (1991, s. 7) lamba tasarımı görevi kanıt değerlendirme sonucunda deney tasarımına örnek olarak verilebilir. Bu görevde öğrencilerden lambanın dış yüzeyi, içine yerleştirilecek mum ve askı parçasına ait ikişer alternatiften uygun kombinasyonları bularak rüzgârlı havada güvenli şekilde yanacak bir lamba oluşturmaları istenmektedir. Amsel ve Brock'un (1996, s. 533) bitki yetiştirme görevi ise hipoteze dönük kanıt değerlendirme görevine örnek olarak verilebilir. Bu görevde katılımcılardan farklı ortamlara bırakılan ve farklı şekillerde büyüme gösteren bitkilerdeki farklılığın nedenleri sorulmaktadır. Örneklerden de anlaşıldığı üzere geliştirilmesi planlanan testte yer alacak kanıt değerlendirme görevlerine üretilecek yanıtlar deney tasarımı veya hipotezler olacağından, kanıt değerlendirme bileşenine BÜT 3-5 teorik çerçevesinde ayrıca yer verilmemiştir.

2. *Test formatı* açısından değerlendirildiğinde kanıt değerlendirme görevlerinin yaratıcılık kuramları ile örtüşmediği düşünülebilir. Her ne kadar gerçek hayatta farklılıklar olsa da araştırmalarda kullanılan kanıt değerlendirme görevleri, hipotezlerin kabul edildiği veya reddedildiği binari yanıtlara (ikili yanıtlar: kapalı uçlu evet/hayır türü yanıt) dayanmaktadır. Çünkü kanıtları değerlendirme süreci nedensel ilişkiler üzerinden yapılmaktadır (Zimmerman, 2007, s. 180) ve nedensel ilişkilerde bir neden yalnızca bir sonuçla ilişkilidir. Bu durum ise tekil üretim olarak değerlendirilebilir (Guilford, 1956).

Bu durumun yaratıcılık değerlendirmelerinin merkezinde yer alan çoğul üretime dayalı (açık uçlu maddeler: çoklu yanıt) kanıt değerlendirme görevlerinin geliştirilmesini engelleyeceği düşünülebilir. Shaklee ve Paszek'in (1985, s. 1233) sayı matrisleri ile geliştirdikleri görevler bu duruma örnek olarak verilebilir. Bu görevde hipotezler ve kanıtlar birlikte verilip katılımcılardan kanıtların hangi hipotezi desteklediğini bulmaları istenmektedir.

3. *Uygulama grubunun özellikleri göz önünde bulundurulduğunda çocuklarda deneye dayalı kanıt değerlendirme becerisine ilişkin araştırma bulgularının kararlı olmadığı görülmektedir.* Bulgular araştırmada kullanılan görevlere, katılımcıların yaş gurubuna ve bilimsel olmayan inanışlarına göre değişmektedir. Çeşitli araştırmalarda kanıt değerlendirme görevlerinin çocuklar için uygun olmadığı ortaya koyulurken, çocukların sadece tutarlı kanıtlar verildiğinde kanıt değerlendirme görevlerinde başarılı oldukları rapor edilmiştir (Amsel and Brock, 1996, s. 525). Çocukların başarı olduğu çalışmalarda ise yanıtlara üretilen gerekçelerin tutarsız olduğu ve inanışlardan etkilendiği rapor edilmiştir (Bullock, 1991, s. 13; Kuhn et. al., 1988; Ruffman et. al., 1993, s. 1617). Araştırma bulgularına kanıt değerlendirme becerisinin 11-12 yaşından itibaren tutarlı şekilde ölçülebileceğini ortaya koymaktadır (Sodian, Zaitchik and Carey, 1991, s. 764; Bullock, 1991, s. 16; Kuhn et. al., 1988). Ruffman ve diğerleri (1993, s. 1621 ve 1627) ise yaptıkları araştırmada basitleştirilmiş ve hipotezlerle mükemmel uyumlu deney çıktıkları verildiği kanıt değerlendirme görevlerinin 6-7 yaşındaki çocuklar için uygun olabileceğini rapor etmişlerdir.

SDDS modelinde ve BÜT teorik çerçevesinde yer alan kanıt değerlendirme bileşeni hipotez oluşturma veya deney tasarımı şeklinde olmasından, ilkökul öğrencilerine uygun bir görev türü olmamasından ve yaratıcılığın çoğul üretim görevlerine uygun olmamasından dolayı BÜT'ün teorik çerçevesi revize edilerek Şekil 2.3'te verilen BÜT 3-5'in teorik çerçevesi oluşturulmuştur.



Şekil 3.1. BÜT 3-5 teorik yapısı

BÜT 3-5 geliştirilme sürecinde Şekil 3.1’de verilen teorik çerçeve kapsamında açık uçlu hipotez geliştirme ve deney düzenleme görevleri üretilmiştir. Madde yazma sürecine geçmeden önce araştırmacı tarafından Şekil 3.1’de yer alan kuramsal yapının bileşenlerine ilişkin örnekler araştırılmıştır. Bu süreçte yapılan alan yazın taramasında farklı deney tasarımı ve hipotez oluşturma görevlerine rastlanmıştır. Alan yazında 2 farklı hipotez oluşturma ve 2 farklı deney tasarımı görevine ulaşılmıştır:

- 1. Tip Hipotez Oluşturma Görevi: Bu tür hipotez oluşturma maddelerinde öğrencilere bir deney düzeneği sunularak deneyin yapılma nedenleri sorularak, deneyle uyumlu hipotezler üretmeleri istenmektedir. (Ayas ve Sak, 2014, s. 198)
- 2. Tip Hipotez Oluşturma Görevi: Bu tür hipotez oluşturma maddelerinde öğrencilere deneyin sonucu verilmekte, sonucun nedenlerine ilişkin, sonuçla ve deneyle uyumlu hipotezler üretmeleri istenmektedir (Liang, 2002, s. 162; Ayas and Sak, 2014, s. 198; Amsel ve Brock, 1996, s. 535).
- 1. Tip Deney Tasarımı Görevi: Bu tür deney tasarım görevinde öğrencilere bir hipotez ve bu hipotezi test etmek için uygun olmayan bir deney düzeneği sunularak deney düzeneğinde hipotezi düzgün bir şekilde ispatlayabilecekleri düzenlemeler yapmaları istenmektedir. Verilen hipoteze uygun deney düzenekleri istenmektedir (Ayas and Sak, 2014, s. 198; Mohamed, 2006, s. 264; Bullock, 1991, s. 5)

- 2. Tip Deney Tasarımı Görevi: Bu tür deney tasarım görevinde öğrencilere bir deney düzeneği sunulur ve bir sonuç verilir ve öğrencilerden sonuca ulaşacak şekilde deney düzeneğinde değişiklikler yapmaları istenir. Deney düzeneğini sonuca göre düzenlemeleri istenmektedir (Ayas and Sak, 2014, s. 198; Bullock, 1991, s. 5; Kuhn and Angelev, 1976, s. 699)

Yukarıda açıklanan deney tasarımı ve hipotez oluşturma görevlerine ilişkin birer tane prototip test maddesi geliştirilmiştir. Geliştirilen prototip maddeler bilimsel yaratıcılık ve yaratıcılık ölçümünde tecrübeli iki alan uzmanı tarafından değerlendirilmiş ve gerekli düzenlemeler yapılarak test görselleri ve yönergeler eklenmiştir.

Murphy ve Davidshofer (2004) test geliştirme aşamalarını testin oluşturulması, testin standardizasyonunun yapılması ve zaman içerisinde revizyonların yapılması olarak sıralamış ve testin oluşturulması aşamasında iki farklı yöntem önermiştir. Bunlardan birincisi teori temelli madde oluşturmayı öngören *rasyonel yöntemdir*. İkinci yöntem ise farklı ve çok sayıda maddenin yer aldığı bir havuz oluşturup bu maddelerin belirli bir özelliğe göre iyi tanımlanmış grupları birbirlerinden ayırıp ayırmadıklarına bakılarak maddelerin belirlendiği *deneysel yöntemdir*. BÜT 3-5 geliştirilme sürecinde rasyonel yöntem benimsenmiştir. Bunun için testin son halinde hedeflenen sayıdan (4 madde) daha fazla sayıda madde oluşturularak ön deneme ve pilot uygulama aşamalarında elde edilen bulgular ışığında madde seçimi yapılarak teste son hali verilmiştir.

3.3.1.1.1. BÜT 3-5 madde geliştirme süreci

Ölçek geliştirme sürecinin belki de en önemli aşaması maddelerin yazılmasıdır. Çünkü madde yazım aşamasını önemli kılan kapsam geçerliği ile ilgili olmasıdır (Hinkin, 1995, s. 969). Kapsam geçerliği, ölçülmek istenen yapıyı temsil eden maddelerin olup olmadığı ile ilgilidir ve bu maddelerin uygun şekilde geliştirilip geliştirilmediğinin bir göstergesidir (Kaplan and Saccuzo, 2009, s. 136). Bu bağlamda testin kapsam geçerliğinin sağlanması için BÜT 3-5 kuramsal yapısında yer alan her bir bileşen ile ilgili yeterli sayıda maddenin elde edilmesi için geniş bir madde havuzunun oluşturulmasına karar verilmiştir. Madde havuzunda kaç madde olması gerektiğine ilişkin net bir sayı olmamakla beraber, testin son halinde yer alması planlanan madde sayısının 3-4 katı kadar madde geliştirilmesi önerilmektedir (DeVilles, 2003, s. 66; Kaplan and Saccuzo, 2009, s. 158). BÜT 3-5 bir ders saatinde uygulanacak (yaklaşık 40 dakika) bir yaratıcılık testi olarak tasarlandığından testin son halinde 2 deney ve 2 hipotez geliştirme maddesi

yer alması planlanmıştır. Bu bakımdan 12-16 arası maddeden oluşan bir havuz hedeflenmiştir. Hedeflenen sayıdaki maddelerin ise daha geniş bir havuzdan seçilmesine karar verilmiştir.

Madde yazma sürecinde iki farklı yöntem izlenebilir: tümevarım ve tümünden gelim (Kline, 2005, s. 29; Clark, 1995, s. 311). Tüm dengelim yönteminde maddeler tanımlanmış bir kuramsal yapı çerçevesinde yapıyı temsil edecek şekilde yazılırken, tümevarım yönteminde net bir kavramsal çerçeve olmadığı durumlarda konu hakkında bilgi sahibi olanların görüşleri sorularak maddeler yazılır. BÜT 3-5 madde havuzunun oluşturulmasında tüm dengelim yöntemi benimsenmiştir. Kapsam geçerliğinin sağlanması amacıyla madde yazım sürecinde uzmanlardan oluşan bir ekip ile çalışılmıştır (Hinkin 1995, s. 969).

Madde yazma sürecinde Eskişehir il merkezindeki ortaokullarda ve ilkokullarda görev yapan sınıf öğretmeni ve fen bilgisi öğretmenleri ile birlikte çalışılmıştır. Öğretmenler en az 10 yıl öğretmenlik tecrübesine sahip öğretmenler arasından kolay ulaşılabilir amaçlı örnekleme yöntemi (Yıldırım ve Şimşek, 2008) ile seçilmiştir. Bu aşamada öğretmenlere araştırmanın amacı, sürecin nasıl işleyeceği, öğretmenlerden beklentiler ve süreçte öğretmenlerin kazanımları açıklanmış, gönüllülük esasına dayalı 10 fen bilgisi öğretmeni ve 2 sınıf öğretmeninden oluşan madde geliştirme ekibi oluşturulmuştur. Fen bilgisi öğretmenlerinden bir tanesi süreçte beklentileri karşılayamayacağı gerekçesi ile bir tanesi de iş yoğunluğundaki artış nedeni ile ikinci haftadan itibaren çalışmadan ayrılmışlardır. Araştırmanın test maddesi geliştirme gurubu Eskişehir il Milli Eğitim Müdürlüğü'ne bağlı 2 farklı ilkokul ve 8 farklı ortaokulda görev yapan, 6'sı kadın ve 4'ü erkek ve hizmet süreleri 6-17 yıl ($\bar{X}=12,3$ yıl) arasında değişen 10 öğretmeninden oluşmaktadır. Hizmet yılı 10 yılın altında olan bir öğretmen ise yüksek lisans derecesine sahip olduğu için madde geliştirme ekibine dâhil edilmiştir.

Madde yazma süreci haftada bir olmak üzere 10 oturumda gerçekleştirilmiştir. Her bir oturum standart bir formatta araştırmacının moderatorlüğünde 3-5 saatlik oturumlarda yürütülmüştür. Oturumlar Anadolu Üniversitesi Eğitim Fakültesinde Mart, Nisan ve Mayıs aylarında (2015-2016 eğitim öğretim yılı), hafta sonları akıllı tahta olan sınıflarda yürütülmüştür. Her oturum için öğretmenlere MEB ek-ders yönetmeliğinde belirtilen rakamdan az olmamak üzere 4 saat üzerinden ücret ödenmiş, sorularını taşımaları ve kayıt etmeleri için birer hafıza kartı verilmiş ve süreç sonunda teşekkür belgesi ve katılım belgesi takdim edilmiştir.

Sürecin birinci haftası öğretmen eğitimine ayrılmıştır. Bu eğitimde öğretmenlere öncelikle araştırmanın amacı, sürecin nasıl işleyeceği, öğretmenlerden beklentiler ve süreçte öğretmenlerin kazanımları tekrardan hatırlatılmıştır. Toplantılara katılma ve devam konuları açıklanmış ve öğretmenlerin görüşleri alınarak tüm ekip elemanları için uygun olan bir gün ve saat belirlenmiştir. Katılımcılara sadece katıldıkları ve test maddesi geliştirdikleri toplantılar için ödeme yapılacağı, 7 kişinin altında katılımcı olması hâlinde toplantının ertelenerek telafi toplantısı yapılacağı belirtilmiş, öğretmenlerin süreçte yaptıkları çalışmalarını kaydetmeleri için hafıza kartları verilmiştir. Son olarak katılımcılara araştırmacı tarafından hazırlanan BÜT 3-5 madde hazırlama eğitimi verilmiştir. Eğitimde aşağıdaki konu başlıklarına yer verilmiştir.

- Yaratıcılık ve yaratıcılığın ölçülmesi
- Bilimsel yaratıcılık ve bilimsel yaratıcılığın ölçülmesi
- SDDS modeli
 - Hipotez oluşturma
 - Deney tasarımı
 - Kanıt değerlendirme (bu bileşen açıklanmadan verilmiştir)
- BÜT teorik yapısı
- BÜT 3-5 teorik yapısı
 - Hipotez görevi türleri
 - Deney görevi türleri
- BÜT puanlama
 - Akıcılık, esneklik, orijinallik, bileşik yaratıcılık
- Prototip test maddelerinin analizi
 - Kötü test maddelerinin prototip maddeler ile karşılaştırılması
 - Madde değerlendirme kriterleri: Prototip ve kötü test maddelerinin değerlendirilmesi
- Örnek uygulama
 - Örnek uygulamanın değerlendirilmesi

Yaklaşık 2 saat süren eğitimin ardından kalan 9 haftada toplantıların nasıl yürütüleceği hakkında bilgi sahibi olmaları için örnek bir uygulama yapıdır. Örnek uygulamada öğretmenlere ikişer tane madde geliştirme formu (EK-3a ve EK-3b) verilmiş

ve bir hipotez ve bir deney tasarım maddesi geliřtirmeleri için yarım saat zaman verilmiřtir. Yarım saat sonunda öğretmenlerden madde geliřtirme formu toplanarak dijital ortama aktarılmıř (fotoğrafları çekilmiř) ve her katılımcının hazırladıđı madde sunulmuřtur. Sunum yapan katılımcının geliřtirdiđi maddeler diđer katılımcılar tarafından ařađıda verilen madde deđerlendirme kriterlerine gre deđerlendirilmiř ve dzeltmeler nerilmiřtir.

- Bilgi Dzeyi: Konu hakkında sınırlı birikimi olan đrencilerde bireysel tecrbelerinden hareketle yanıt retebilir mi? Belirli bir disiplindeki bilgiye dayanmalı ancak yanıt retebilmek için konunun tamamen bilinmesine gerek duyulmamalı. İerik olarak 3, 4 ve 5. sınıf fen bilimleri mfredatı ile uyumlu olmalı.
- Aık Ululuk: Maddeye 10-15 arası yanıt retebiliyor mu? Bu yanıtlar farklı kavramsal kategoriler altında toplanıyor mu? Maddeler ođul retim yaklařımına uygun olmalı.
- Greve Uygunluk: Hipotez geliřtirme grevlerinde yanıtlar hipotezler olarak retebiliyor mu? Deney tasarım grevlerine retilen yanıtlar rnekler ile uyumlu mu? Hipotez grevlerinde hipotez, deney grevlerinde deneysel dzenlemeler retilmeli.
- Teknik Uygunluk: Madde alan bilgisi ile tutarlı mı? Maddenin retildiđi disiplin ile tutarlı olmalı.
- Grsel Uygunluk: Madde grsel ieriyor mu, grselleřtirilebilir mi? Madde đrencilerin ilgisini eker mi? Her madde đrencilerin ilgisini ekecek řekilde grsel iermeli ve mmkn ise gnlk yařam problemleri ile iliřkili olmalı.
- Madde Formunun Uygunluđu: Madde geliřtirme formu uygun řekilde doldurulmuř mu? Maddeler madde geliřtirme formuna kaydedilmeli ve formdaki bilgiler eksiksiz doldurulmalı ve dijital ortama aktarılmalı.

Eđitim sonunda đretmenlere eđitim materyalleri ve madde geliřtirme formları dađıtılmıřtır. Geliřtirilecek maddelerin alan bilgisi gerektirmesinden dolayı 3, 4 ve 5. sınıf MEB Fen bilimleri dersi mfredatı dađıtılmıř ve geliřtirilecek maddelerin ieriklerinin mfredat ile uyumlu olması gerektiđi tekrardan hatırlatılmıřtır. retilen maddelerin disiplinlere gre (fizik, kimya, biyoloji) dengeli řekilde dađılması amacıyla đretmenlerden gelen neri zerine fen bilgisi đretmenleri kendi aralarında 2 fizik, 3

kimya ve 3 biyoloji olmak üzere disiplinlere ayrılmışlar ve her hafta aynı disiplinden madde hazırlamayı önermişlerdir. Bu öneri kabul edilmiştir. Sınıf öğretmenlerine ise istedikleri alandan madde geliştirebilecekleri söylenmiştir. Son olarak öğretmenlerden örnek uygulamada geliştirdikleri maddeleri verilen dönütlere göre düzenlemeleri ve takip eden toplantı için bir hipotez ve bir deney tasarımı sorusu hazırlamaları istenmiştir.

İkinci toplantıya gelen 10 öğretmen toplanda 20 yeni madde ve 20 düzenlenmiş maddeyi sunmuşlardır. Maddelerin değerlendirilme dönüt süreci çok zaman almasından dolayı öğretmenler geriye kalan haftalarda birer madde getirmeyi, 3, 4, 5 ve 6. haftalar için deney tasarımı maddesi; 7, 8, 9 ve 10. haftalarda ise hipotez geliştirme maddeleri hazırlamayı önermişlerdir. Öğretmenlerden gelen bu öneri de kabul edilerek devam eden oturumlarda öncelikle düzeltme verilen maddeler, daha sonra ise yeni maddeler sunulmuştur.

Oturumlarda Nisan ve Mayıs aylarında okullardaki programların yoğunlaşmasından dolayı devam problemleri olmakla birlikte en az 7 öğretmenin katıldığı 10 oturum tamamlanmıştır. Süreçte 118 madde geliştirilmiştir. Bu maddelerden bazıları birbirleri ile benzerliklerinden dolayı, bazıları orijinal olmadıkları için, bazıları ise madde geliştirme kriterleri dikkate alınarak havuzdan çıkartılmış ve 85 madde elde edilmiştir. Bu maddelerden 23 tanesi fizik, 28 tanesi kimya ve 33 tanesi biyoloji alanından maddelerdir.

3.3.1.1.2. Havuzdan madde seçimi ve taslak formun oluşturulması

Madde havuzu oluşturulduktan sonra iki aşamalı bir süreçle madde havuzundan 12 madde elde edilmiştir. Birinci aşamada araştırmacı ve ikinci bir uzman tarafından havuzdan 24 madde seçilmiştir. Bu süreçte, kuramsal alt yapıda yer alan ölçüm hücrelerine (bakınız Tablo 2.5) ikişer madde gelecek şekilde, yukarıda açıklanan madde değerlendirme kriterlerine göre seçim yapılmıştır.

Madde havuzundaki maddeler öncelikle Tablo 2.5'te verilen hücrelere göre sınıflanmıştır. Sınıflama işleminde maddeler öncelikle deney ve hipotez olarak ayrılmış, daha sonra deney ve hipotez soruları kendi içlerinde fen alanlarına göre sınıflanmış (fizik, kimya, biyoloji), son olarak bu maddeler kendi içlerinde görev türlerine (1. Tip/2. Tip) göre sınıflandırılmıştır. Bu şekilde uzmanlar arasında tutarlılığın artırılarak seçim işleminde pratiklik sağlanmıştır. Örneğin havuzdaki fizikte birinci tip hipotez oluşturma görevi 4 tanedir ve uzmanlar birbirlerinden bağımsız şekilde bunlardan 2 tanesini

seçmişlerdir. Uzmanlar daha sonra bir araya gelerek seçimlerini karşılaştırmışlar, ortak olan maddeler doğrudan seçilmiş, diğerleri için uzlaşmaya varılmıştır. Seçilen maddelerin kuramsal alt yapıdaki ölçüm hücrelerine göre dağılımı Tablo 3.2’de verilmiştir.

Tablo 3.2. Birinci aşamada seçilen maddelerin göreve ve disipline göre dağılımları

Görev	Görev Türü	Fen Bilimleri Alanları			Toplam	Toplam
		Fizik	Kimya	Biyoloji		
Hipotez Oluşturma	1. Tip Hipotez	2	2	2	6	12
	2. Tip Hipotez	2	2	2	6	
Deney Tasarımı	1. Tip Deney	2	2	2	6	12
	2. Tip Deney	2	2	2	6	
Toplam		8	8	8	24	

Madde seçiminin ikinci aşamasında birinci aşamada seçilen 24 madde araştırmacı ve farklı bir uzman tarafından birinci aşamadaki yöntem ile değerlendirilmiş, her bir hücreden 1 madde olmak üzere 12 madde elde edilmiştir. Bu maddeler görev türüne göre 6 hipotez geliştirme (3 tane 1. Tip + 3 tane 2. Tip), 6 deney tasarımı (3 tane 1. Tip + 3 tane 2. Tip) maddesi olacak şekilde dağılırken fen alanlarına göre 4 fizik, 4 kimya ve 4 biyoloji şeklinde dağılmaktadır.

Madde seçim sürecinde elde edilen 12 maddenin görselleri bir grafik tasarımcısı tarafından çizilmiş ve yönergeler ve açıklamalar bir dil uzmanı tarafından kontrol edilerek düzenlenmiştir. Düzenlenen maddeler ile uzman görüş formu oluşturulmuş ve 11 alan uzmanına uzman değerlendirmesi için gönderilmiştir. Uzmanların 6’sı fen eğitimi ve yaratıcılık alanında çalışan akademisyen, 4’ü madde geliştirme sürecine görev alan fen bilgisi öğretmeni ve 1’i fen eğitimi alanında doktora derecesine sahip fen bilgisi öğretmenidir. Akademisyenlerden 5’i fen eğitimi ve biri matematikte yaratıcılık ve yetenek alanında çalışan uzmanlar olup bunların 3’ü bilimsel yaratıcılık ve fen yeteneği üzerine çalışmaktadırlar. Öğretmenler ise madde geliştirme ekibindeki motivasyonu yüksek ve devamsızlığı en az olan öğretmenlerdir. Ancak bu süreçte değerlendirme formunda öğretmenlerin yönerge, görsel ve görev türünün ne olduğuna dair görüşleri alınmıştır

Madde geliştirme ekibindeki öğretmenlerin kendi maddelerine karşı öznel yaklaşacağı düşünülebilir. Ancak uzmanlar, öncelikle maddeleri yönergelerin ve görsellerin uygunluğu bakımından değerlendirmiş ve dönütler vermişlerdir. İkinci olarak

içerik geçerliğinin sağlanması amacıyla maddelerin kuramsal yapıya uygunlukları değerlendirilmiştir. Bunun için her maddenin hangi fen alanında olduğunu (fizik, kimya biyoloji) ve hangi beceriyi (hipotez geliştirme, deney tasarımı) ölçtüğünü bulmaları istenmiştir. Uzmanların görsel ve yönergeler hakkındaki dönütleri incelenerek yönergelerde küçük değişiklikler yapılmış, bir maddenin görseli yeniden düzenlenmiştir. Uzmanların %80'i tarafından (9 uzman) aynı görev veya disiplin için sınıflandırılan maddeler olduğu gibi kabul edilmiş, daha düşük uyumlar için madde araştırmacı tarafından yeniden kontrol edilmiş ve düzeltme yapılmıştır (Hinkin, 1995, s. 970). Uzmanların maddeleri hipotez ve deney görevleri olarak sınıflamaları yüksek düzeyde uyumlu bulunurken, kimya maddesi olarak seçilen maddelerden birinin 3 uzman tarafından fizik maddesi olarak kodlandığı görülmüştür. Madde incelendiğinde hem fizik hem de kimya alan bilgisi gerektiren ısı-sıcaklık konusunda bir madde olduğu görülmüş ve bu madde kimya maddesi olarak kabul edilmiştir.

Uzman değerlendirmesi sonucunda 12 madde ile her birinde 6 madde yer alan (3 deney ve 3 hipotez) iki *taslak form* oluşturulmuştur. Birinci taslak forma birinci tip hipotez ve birinci tip deney maddeleri, ikinci taslak forma ise ikinci tip hipotez ve ikinci tip deney maddeleri koyulmuştur.

3.3.1.1.3. Ön deneme uygulaması ve pilot formun oluşturulması

Geliştirilen taslak formda yer alan maddelerin, görsellerin ve yönergelerin anlaşılıp anlaşılmadığını, maddelerin öğrencilerin ilgisini çekip çekmediğini ve her bir madde için gereken yaklaşık zamanın belirlenmesi amacıyla 4. sınıfa devam eden 18 öğrenci ile ön deneme uygulaması yapılmıştır. 4. sınıf öğrencilerinin 3 ve 5. sınıfların ortasında yer almasından dolayı bu sınıf düzeyi ile ön deneme uygulamasının yapılmasına karar verilmiştir.

Yapılan uygulamada 10 öğrenciye birinci taslak form, 8 öğrenciye ise ikinci taslak form dağıtılmış ve iki uygulayıcının katılımı ile uygulama yaklaşık bir ders saatinde gerçekleştirilmiştir. Öğrencilerden yönergeleri dikkatlice okumaları, görselleri incelemeleri ve anlamadıkları yerleri sormaları istenmiştir. Uygulama esnasında öğrenciler anlamadıkları yerleri sormuşlar, uygulayıcılar ise öğrencilerin anlamadıkları noktaları ve uygulamaya ilişkin gözlem notları almışlardır.

Uygulama notları ve öğrenci kağıtları incelendiğinde ön deneme uygulamasına ilişkin bulgular test formatı, zaman ve içerik olmak üzere üç başlık altında toplanmıştır. Bulgular temalara göre aşağıdaki gibi sıralanabilir:

- Öğrencilerin *test formatına* uyum sağlayamadıkları, maddelere kısa ve az sayıda yanıt ürettikleri gözlenmiştir. Uygulayıcıların “Bu soruların birden çok yanıtı var, hatta bir sürü yanıtı var” şeklindeki açıklamalarının ardından üretimin arttığı görülmüştür.
- Testte ayrılan *zaman* ile ilgili olarak iki farklı nokta ön plana çıkmaktadır: ilk olarak öğrencilerin maddelere ortalama 5-7 dakika zaman ayırdıkları, ikinci olarak da testte maddelere eşit zaman ayırmadıkları gözlenmiştir. Öğrencilerin ilgilerini çeken maddelere diğerlerine göre daha fazla zaman ayırma eğiliminde oldukları, dolayısı ile fazla yanıt ürettikleri görülmüştür.
- *Testin içeriğine* ilişkin uygulama notları ve öğrencilerin kâğıtları incelendiğinde bazı yönergelerin ve görsellerin anlaşılmadığı görülmüştür. En fazla sorunun birinci taslak formdaki iki birinci tip hipotez (deney düzeneğine ilişkin hipotez üretme görevi) maddelerine geldiği görülmüştür. Bu maddelerden birisi fizik alanında elektrik devreleri ile ilgili olup, diğeri kimya alanında ısı-sıcaklık konusu ile ilgilidir. Fizik maddesinde 6 farklı elektrik devresi ile ilgi hipotez oluşturulması istenirken, kimya maddesinde 6 farklı erime düzeneğine ilişkin hipotez oluşturulması istenmektedir. Öğrenci kâğıtları incelendiğinde bu iki maddeye hiçbir öğrencinin doğru yanıt üretemediği görülmüştür. İkinci olarak bazı öğrencilerin deney tasarımı görevini hipotez görevine genelleyerek, hipotez görevi için deney düzenlemeye ilişkin yanıtlar ürettikleri, bazılarının ise tam tersini yaptıkları görülmüştür.

Ön deneme uygulamasından elde edilen bilgiler ile testin uygulanması ve testin içeriğinde bazı değişiklikler yapılmıştır. Öğrencilerden gelen dönütlere göre bir maddenin görselinde küçük değişiklikler yapılmış ve bazı maddelerin yönergeleri düzenlenmiştir. Testin içeriğine ve uygulamaya yönelik aşağıdaki kararlar alınmıştır.

- Öğrenciler tarafından anlaşılmayan iki adet birinci tip hipotez sorusu testten çıkartılmıştır.
- Öğrencilerin çoğul üretim ve tekil üretim arasındaki farkı anlayabilmeleri ve çoğul üretime alışmaları için uygulama öncesi basit bir çoğul üretim görevinin çözülmesine karar verilmiştir.
- Her bir madde için belirli süre ayrılmasının uygun olacağına karar verilmiştir. Bu zamanın her soruya eşit kullanılması için maddelerin teker teker ve sırasıyla

çözülmesine ve süre sonunda diğer soruya geçilmesine karar verilmiştir. Test bittikten sonra isteyen öğrencilere ek 10 dakika verilmesine karar verilmiştir.

Ön deneme uygulamasından elde edilen bulgular doğrultusunda maddelerde yapılan değişiklikler sonucunda elde edilen 10 madde ile (4 hipotez ve 6 deney) iki *pilot form* geliştirilmiştir. Her bir pilot formda 2 hipotez ve 3 deney görevi yer almaktadır.

3.3.1.1.4. Pilot uygulama

Daha önceki aşama sonucunda elde edilen 2 pilot form kolay ulaşılabilir örnekleme yöntemi ile belirlenen bir ilkokul ve bir ortaokula devam eden 3 ve 5. sınıf 121 öğrenciye (57 üçüncü sınıf, 64 beşinci sınıf) uygulanmıştır. Pilot uygulamanın yapıldığı ilkokulda 4. sınıf öğrencilerine 23 Nisan hazırlıklarından dolayı ulaşılamamıştır. Pilot formlar birer gün ara ile birer ders saatinde uygulanmıştır. Öğrencilerin uygulamada ürettikleri doğru yanıtların sayısı (akıcılık) kullanılarak analizler yapılmıştır. Uygulama esnasında öğrencilerin sordukları sorular not alınmış ve nicel bulgularla birlikte değerlendirilmiştir. Öğrencilerin kağıtlarından elde edilen yanıtlarla testin doğru yanıt anahtarı güncellenmiştir. Pilot uygulama verileri ile betimsel istatistikler, korelasyon katsayıları, maddelerin sınıf düzeyleri arasında ayırt edicilik düzeyleri, iç tutarlık istatistikleri hesaplanmış ve açımlayıcı faktör (AFA) yapılmıştır. Betimsel istatistiklerden sınıfların her madde için ortalama, standart sapma değerleri hesaplanmıştır. Madde ayırt edicilik analizleri için sınıf ortalamaları arasındaki farkın anlamlı olup olmadığına bakılmıştır. Maddelerin ölçek toplamı ile korelasyon katsayıları hesaplanmış ve Cronbach Alpha iç tutarlık analizleri yapılmıştır. Son olarak maddelerin ölçeğin kuramsal alt yapısını yansıtmayı yansıtmadığını belirlemek amacıyla küçük örneklemler için önerilen (Tanaka, 1987, s.138) maksimum olasılık (maximum likelihood) (ML) faktörleştirme yöntemi ve direct oblimin döndürme işlemi ile AFA yapılmıştır. Faktör belirlemede ML metodunun kullanılmasının bir diğer nedeni ise analiz sonucunda elde edilen faktör çözümüne ilişkin istatistik değerlendirmelerine olanak sağlaması (χ^2 uyum iyiliği testi gibi) (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s. 199) ve faktör sayısını daha sıkı ve uygun yöntemlerle belirlemesinden dolayı tercih edilmiştir (Kroonenberg and Lewis, 1982'den aktaran Akbulut, 2010, s. 103).

Veri setinin normallik sınaması için dağılım grafiklerine ve basıklık-çarpıklık istatistiklerine bakılmıştır. Dağılım grafiklerinin normal dağılıma yakın olduğu, basıklık ve çarpıklık değerlerinin tüm maddeler için +2 ve -2 arasında olmasından dolayı normal

dağılımdan aşırı sapma olmadığı sonucuna ulaşılmıştır (Trochim and Donnelly, 2006, s. 48). Kline (1994, s. 74) faktör analizi için gerekli örneklem büyüklüğünün madde sayısının 10 katı kadar olmasını gerektiğini önermektedir. Faktör analizine dâhil edilen 10 madde başına 12 katılımcı düştüğünden dolayı örneklem büyüklüğünün yeterli olduğu kabul edilebilir. Ayrıca veri setinin faktörleştirmeye uygunluğunu test etmek için KMO örneklem uygunluğu istatistiği ve Bartlett küresellik testleri kontrol edilmiştir. KMO test istatistiği .820 olarak hesaplanmış ve Bartlett testi anlamlı bulunmuştur [$\chi^2(45) = 473,61$; $p < .001$]. KMO değerinin .80 üzerinde olmasından dolayı örneklem büyüklüğünün faktör analizi için yeterli olduğu sonucuna ulaşılmıştır. Bartlett testinin anlamlı çıkması veri setinin çok değişkenli normal dağılım varsayımını karşıladığı anlamına gelmektedir (Pallant, 2011, s. 183). Faktör yük değeri için kabul edilebilir düzey .32 (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s.194), binişik madde yükler arası fark .10 olarak belirlenmiştir (Akbulut, 2010, 93). AFA’da faktör belirleme yöntemi olarak Kaiser Kriteri kullanılmış, ayrıca yamaç birikinti grafiğine bakılmıştır. Ancak Henson ve Robert (2006, s. 399; Akbulut, 2010, s. 88) bu iki yöntemin yeterince güvenilir olmadığını belirterek ek verilerin de kullanılmasını önermektedir. Bu bakımdan paralel analizi yapılarak elde edilen öz değerler (eigenvalues) AFA’dan elde edilen öz değerler ile karşılaştırılmıştır.

Yapılan açımlayıcı faktör analizi sonucunda analize temel olarak alınan 10 madde için öz değerleri 1’in üzerinde olan iki faktör olduğu görülmüştür. Yamaç birikinti grafiğinde ise üçüncü noktadan sonra eğimin plato yaptığı görülmüştür. Son olarak paralel analiz için Monte Carlo PCA simülasyon programı (Watkins, 2004) kullanılarak 1000 adet rastgele öz değer üretilmiştir. AFA sonucunda çıkan öz değerlerden sadece iki tanesinin bu öz değerlerden büyük olduğu görülmüştür. Bu bakımdan 10 maddenin iki faktör altında toplandığı ve sonuçların kuramsal alt yapı ile de uyumlu olduğu sonucuna ulaşılmıştır. Elde edilen iki faktörlü yapının uyum iyiliği testi anlamlı çıkmadığından [$\chi^2(26) = 34,453$; $p = .124$] iki faktörlü model için beklenen ve gözlenen kovaryans matrislerinin farklı olmadığı şeklinde yorumlanmıştır (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s. 363). Bu faktörlerin toplam varyansa yaptıkları katkı %44,60’tır. İki faktörün açıkladığı toplam varyans sosyal bilimlerde çok faktörlü desenler için kabul edilebilir değer olan %40’ın üzerindedir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s. 239). 1. Tip deney sorusu olan 1, 2 ve 3. maddeler birinci faktörün altında toplanmış ve faktör yük değerleri sırasıyla .744, .974 ve .900 olarak bulunmuştur. 2. Tip deney soruları

ise ikinci faktör altında toplanmış ve faktör yük değerlerinin .486 ve 723 arasında değiştiği görülmüştür. Dört adet hipotez maddesinin ise faktör yük değerlerinin .371 ve .566 arasında değiştiği görülmüştür.

Analize dahil edilen 10 madde için Cronbach Alpha iç tutarlık katsayısı .832 olarak hesaplanmıştır. Grup ayırt ediciliği için 3 ve 5. sınıf öğrencilerinin ortalama puanları madde bazında bağımsız örneklem t-testi ile karşılaştırılmış, 10 karşılaştırma yapıldığından analizlerde Bonferoni düzeltmesi yapılmış ve anlamlılık değeri 10'a bölünerek .005 olarak belirlenmiştir (Akbulut, 2010, s. 125).

Faktör analizi bulgularında 10. maddenin ortak varyansının .130 olarak hesaplanması bu maddenin ölçeğin tamamına hizmet etmediği şeklinde yorumlanabilir (Field, 2009, s. 639). Hair ve diğerleri (2010, s. 115) 120 ve üzeri örneklem büyüklüklerinde en düşük faktör yük değerinin .50 olması gerektiğini belirtmektedir. 10. maddenin faktör yük değeri (.377) bu kritik değer altında olduğu görülmüştür. Çünkü bu madde varyansın yaklaşık olarak %14'ünü açıkladığından zayıf madde olarak değerlendirilebilir (Tabachnick and Fidell, 2012, s. 654). Maksimum olasılık analizinde maddelerin faktörler ile korelasyonu yapı matrisine (structure matrix) bakılarak (Tabachnick and Fidell, 2012, s. 614) incelenmiş ve 10. maddenin her iki faktör ile korelasyon değerlerinin ($r_1=.211$, $r_2=.360$) düşük olduğu görülmüştür (Pallant, 2011, s. 134; Huck, 2009, s. 55; Büyük Öztürk, 2011, s. 32). Betimsel bulgulara göre hipotez sorusu olan 10. madde için ortalama 2,17 (SS=1,05) ve en fazla 4 doğru yanıt üretilmiştir. 3 ve 5. sınıfların ortalamaları sırası ile 2,02 (SS=1,06) ve 2,31 (SS=1,02) olarak bulunmuş, maksimum puan ise her iki sınıf için de 4,00 olarak hesaplanmıştır. Yapılan bağımsız örneklem t-testi sonuçlarında 10. madde için 3 ve 5. sınıf öğrencilerinin ortalama puanları arasında anlamlı bir farklılık olmadığı görülmüştür ($t_{119}=-1,56$; $p=.112$). Bu bakımdan 10. maddenin bu iki gurubu ayırt etmediği sonucuna ulaşılmıştır. Güvenirlik analizlerinde 10. madde için hesaplanan madde-toplam korelasyon değeri .299 olup bu değer .30'un altında olduğundan maddenin ölçeğin tamamına hizmet etmediği düşünülebilir (Pallant, 2011, s. 100; Akbulut, 2010, s. 81). Daha kesin deliller için maddenin diğer maddeler ile korelasyon katsayıları hesaplanmış ($r_{min}=.112$; $p>.05$; $r_{max}=.273$; $p<.01$), değerlerin düşük olduğu (Pallant, 2011, s. 134; Huck, 2009, s. 55; Büyük Öztürk, 2011, s. 32) ve 10. maddenin silinmesi durumunda Cronbach Alpha katsayısının .834'e yükseleceği görülmüştür. Her ne kadar kolay bir madde gibi gözükse de betimsel bulgular, güvenirlik katsayıları ve AFA bulgularına göre çoğul üretime uygun

olmadığı anlaşılan 10. maddenin, testin güvenilirliğini düşürdüğü ve faktör yapısına uymadığı düşünülerek testten çıkartılmasına karar verilmiştir.

10. madde çıkartıldıktan sonra AFA tekrar edilmiş ve 9 maddenin (6 deney ve 3 hipotez) iki faktör altında toplanarak %48,07'lik varyans açıkladığı görülmüştür. İlk bulgularla benzer şekilde 3 adet birinci tip deney maddesi birinci faktörü oluştururken, 3 adet ikinci tip deney maddesi ve 3 adet hipotez maddesi ikinci faktör altında toplanmıştır. Maddelerin ortak varyans değerleri .173 ve .872 arasında, faktör yük değerlerinin ise .369 ve .985 arasında değiştiği görülmüştür.

Kuramsal alt yapı merkeze alınarak AFA sonuçları, güvenilirlik bulguları ve betimsel istatistikler birlikte değerlendirilerek 9 madde hakkında çeşitli kararlar alınmıştır. Alınan her kararın ardından analizler yenilenmiştir. Bu kararlar ve gerekçeleri şu şekildedir:

- Kuramsal yapıya uymayan ikinci tip deney sorusu olan 6. madde ortak varyansının .208 olması ve bu değer .30'un altında olmasından dolayı ölçeğin tamamına hizmet etmediği düşünülebilir (Field, 2009, s. 639). İkinci faktördeki faktör yük değeri ise .447 olarak bulunmuştur. Örneklem sayısının 121 kişi olmasından dolayı .50'in altında yük değerlerine sahip maddelerin pratikte anlamlı olmadığı düşünülebilir (Hair et. al., 2010, s. 115). Çünkü bu madde varyansın yaklaşık olarak %23'ünü açıkladığından zayıf madde olarak değerlendirilebilir (Tabachnick and Fidell, 2012, s. 654). Güvenirlik analizleri 6. madde için madde-toplam korelasyon değeri .399 olup bu değer .40'un altında olduğundan maddenin ölçeğin tamamına hizmet etmediği düşünülebilir (Gliem and Gliem, 2003, s.86; Hair et. al., 2010, s. 123). Daha kesin deliller için maddenin diğer maddeler ile korelasyon katsayıları hesaplanmış ($r_{min}=.139$; $p>.05$; $r_{max}=.321$; $p<.01$), değerlerin genel olarak düşük olduğu (Pallant, 2011, s. 134; Huck, 2009, s. 55; Büyük Öztürk, 2011, s. 32) ve 6. maddenin silinmesi durumunda Cronbach Alpha katsayısının değişmediği görülmüştür. Yapılan bağımsız örneklem t-testi sonuçlarında 6. madde için 3 ve 5. sınıf öğrencilerinin ortalama puanları arasında anlamlı bir farklılık olduğu görülmüştür ($t_{119}=-3,142$; $p<.05$; $\eta^2=.07$). Bulunan etki büyüklüğünün küçük olduğu görülmüştür (Cohen, 1988, s. 80) Eta-kare etki büyüklüğü bir korelasyon katsayısı olduğundan 6. maddede puanlarının gözlenen varyansının %7'sinin sınıf düzeyine bağlı olması gruplar arasındaki farkın kuramsal olarak ve pratikte her zaman anlamlı

olmayacağı şeklinde yorumlanabilir (Akbulut, 2010, s.114). Bu gerekçelerle 6. maddenin testten çıkartılmasına karar verilmiştir.

- Betimsel bulgulara bakıldığında maddeler için hesaplanan ortalamaların 0,98 (SS=1,35) ve 3,06 (1,65) arasında değiştiği görülmektedir. Ortalaması en düşük olan 7. Madde için maksimum puanın 5,00 olduğu görülmüştür. Bu madde için 3 ve 5. sınıfların ortalamaları ayrı ayrı hesaplanmış ve sırası ile 0,702 (SS=1,03) ve 1,234 (SS=1,54) olarak hesaplanmış, maksimum puan ise 3. sınıflar için 4,00, 5. sınıflar için 5,00 olarak bulunmuştur. Sınıf ortalamaları arasındaki farkın ise istatistiksel olarak anlamlı olmadığı görülmüştür ($t_{111} = -2,25$; $p > .005$). Bu maddenin 1. Tip hipotez görevi olduğu ve ön deneme uygulamasında diğer 1. Tip hipotez maddelerinin öğrenciler tarafından anlaşılmadığı için testten çıkartılması betimsel bulgularla örtüşmektedir. Buradan hareketle 7. maddenin her iki sınıf seviyesi için de anlaşılması zor bir madde olduğu sonucuna varılmış ve testten çıkartılmıştır.

Pilot uygulama formlarında yer alan 10 madde (6 deney ve 4 hipotez) ile hesaplanan betimsel istatistikler, güvenilirlik analizleri, ayırt edicilik sonuçları ve faktör analizi bulgularından hareketle deney görevi olan 6. madde ve hipotez görevi olan 7. ve 10. maddeler testten çıkartılmıştır. Geride kalan 7 madde ile yapılan AFA bulgularına göre 2 faktör altında toplanan maddelerin toplan açıkladığı varyansın %54,22 olduğu ve bu değer sosyal bilimler için yeterli olduğu kabul edilmiştir (Akbulut, 2010, s. 103). Kalan 7 madde için Cronbach Alpha iç tutarlık katsayısı ise .826 olarak hesaplanmıştır. Bu değer elde edilen 7 madde ile ölçeğin güvenilirliğinin yüksek olduğu anlamına gelmektedir (Pallant, 2011, s. 299).

Pilot uygulamadan elde edilen 7 maddenin 5'i deney sorusu, 2'si ise hipotez sorusudur. Deney görevlerinin 3 tanesi 1. Tip deney sorusu, 2 tanesi ise 2. Tip deney sorusudur. Deney ve hipotez sorularının dengeli bir şekilde dağıtılması amacıyla 3. maddenin basitleştirilmesine ve hipotez sorusu şeklinde yönergenin yeniden düzenlenmesine karar verilmiştir. 1. Tip deney tasarımı görevi olan bu maddeden hipotez çıkartılmış ve bunun yerine deneyin sonucu verilerek bu durumun nedenlerinin sorulduğu 2. Tip hipotez görevine dönüştürülmüştür. Maddeler düzenlendikten sonra alan uzmanlarına tekrar gönderilmiş ve dönütlere göre yapılan son düzenlenmeler ile testin uygulama formuna eklenmişlerdir. Son olarak betimsel bulgulardan ($\bar{X}=1.12$; $SS= 1.64$)

hareketle 1. madde için öğrenci kağıtları incelenmiş ve madde görselinin yeniden düzenlenmesine karar verilmiştir.

Uygulama esnasında alınan notlara göre uygulama öncesi yapılan çoğul üretim alıştırmalarının öğrencilerin görevleri anlamalarına yardımcı olduğu görülmüştür. Ancak pilot uygulamadaki öğrenci kağıtları incelendiğinde öğrencilerin bir kısmının hipotez görevlerinden sonra gelen deney görevlerine de hipotez ürettikleri, bir kısmının ise deney görevlerinden sonra gelen hipotez görevlerine de deney düzenleme yanıtları ürettikleri görülmüştür. Öğrencilerin farklı görevlerden etkilenmemeleri için hipotez ve deney görevleri ayrı kitapçıklarda toplanmıştır.

BÜT 3-5 geliştirme sürecine kuramsal yapının oluşturulması ile başlanmış, ikinci aşamada ise alan uzmanlarından oluşan bir ekip ile kuramsal yapıya uygun çok sayıda madde geliştirilmiş ve bir madde havuzu oluşturulmuştur. Madde havuzundan seçilen 12 madde ile bir taslak form oluşturulmuş, taslak form ön deneme uygulamasından sonra revize edilerek 10 maddelik pilot form elde edilmiştir. Pilot form, pilot uygulamadan sonra revize edilerek 7 maddelik uygulama formu geliştirilmiştir.

3.3.2. BÜT 3-5 uygulama formu

Uygulama formundaki maddeler iki kitapçık şeklinde tasarlanmıştır. Birinci kitapçıkta 3 hipotez görevi yer alırken ikinci kitapçıkta 4 deney görevi yer almıştır. *Hipotez* görevlerinde bir deney düzeneği ve bu düzeneğin çalışmasıyla elde edilen sonuçlar verilmektedir. Öğrencilerden verilen deney düzeneğinin elemanlarını ve sonuçlarını kullanarak neden-sonuç ilişkilerinin yer aldığı hipotezler üretmeleri istenmektedir. Hipotez alanında fizik, kimya ve biyoloji alanlarından birer görev yer almaktadır. *Deney alanında* bir fizik, bir biyoloji ve ikisi kimya alanında olmak üzere iki farklı deney düzenleme görevi yer almaktadır. Birinci tip deney görevlerinde öğrencilere bir hipotez ve bu hipotezi test etmek için kurulmuş fakat hipotezi test etmek için uygun olmayan bir deney düzeneği verilmektedir. Öğrencilerden hipotezin test edilebileceği deneyi düzenlemeleri istenmektedir. İkinci tip deney görevlerinde bir deney düzeneği ve düzeneğin çalıştırılması ile elde edilen sonuçlar verilmektedir. Öğrencilerden verilen farklı bir sonucu elde edecek şekilde deneyi yeniden düzenlemeleri istenmektedir.

3.3.3. BÜT 3-5'in puanlanması

BÜT 3-5'in puanlanması ile yaygın olarak kullanılan yaratıcılık indeksleri ve yeni bir bileşik yaratıcılık indeksi elde edilmektedir. Her madde için öğrenci yanıtlarından

akıcılık, esneklik ve bileşik yaratıcılık puanları hesaplanmaktadır. *Akıcılık* puanı bir problem için üretilen doğru çözümlerin sayısıdır. *Esneklik* puanı ise çözümlerin üretildiği kavramsal kategorilerin sayısını ifade etmektedir. Yaratıcılık ölçümlerinde yaygın olarak kullanılan orijinallik puanı yerine BÜT 3-5'te *bileşik yaratıcılık (CQ)* puanları hesaplanmaktadır.

Orijinallik puanları yerine CQ puanlarının kullanılmasının çeşitli nedenleri vardır. Çözümlerin yenilik düzeyini belirlemek amacıyla kullanılan orijinallik, çözümlerin örnekleme rastlanma sıklığının örneklem büyüklüğüne bölünmesi ile elde edilmektedir. Bu bakımdan orijinallik puanının örneklem büyüklüğünden etkilendiği düşünülebilir. Örneğin küçük örneklemelerde tüm yanıtların orijinallik puanının yüksek olacağı, büyük örneklemelerde ise düşük olacağı düşünülebilir. Ayrıca birçok araştırmada orijinallik puanları ve akıcılık arasında yüksek korelasyon değeri olduğu rapor edilmektedir (Mouchiroud and Lubart, 2001). Bu durum ise akıcılık puanlarının orijinallik puanlarını şişirmesine neden olmaktadır. Farklı bir ifade ile akıcılık puanları ile orijinallik puanları aynı şeyi ifade etmektedir. Diğer bir sorun ise yaratıcılık ölçümlerinde elde edilen akıcılık ve esneklik puanlarından esneklik puanlarının orijinallik puanlarının hesaplarında kullanılmamasıdır. Bu tür problemlerle baş etmek amacıyla Snyder, Mitchell, Bossomaier ve Pallier (2004) tarafından geliştirilen CQ formülü kullanarak kavramsal esnekliğin daha çok öne çıkarıldığı CQ puanlarının kullanılmasına karar verilmiştir. CQ puanları aşağıdaki formül ile hesaplanmaktadır:

$$CQ = \log_2(1+U_1) + \log_2(1+U_2) + \dots + \log_2(1+U_n)$$

Formülde verilen 1, 2, ..., n kategorileri, U_1, U_2, \dots, U_n ise herhangi bir kategoride üretilen doğru yanıtların sayısını ifade etmektedir. BÜT 3-5'te hesaplanan puanlar aşağıdaki örnek üzerinden daha kolay anlaşılabilir.

Örnek Görev: Yediğimiz bitkilere yazabileceğiniz kadar çok sayıda örnek yazınız.

Örnek Yanıt: Ayva, elma, armut, kiraz, domates, patlıcan, maydanoz, tere, roka

Örnek Puanlama:

Akıcılık: 9 (doğru yanıt sayısı)

Esneklik: 3 (3 kavramsal kategori: 4 meyve, 2 sebze, 3 ot)

$$CQ = \log_2(1+4) + \log_2(1+2) + \log_2(1+3) = 2,32 + 1,58 + 2$$

$$CQ = 5,90$$

Örnekteki öğrencinin akıcılık puanı 9, esneklik puanı 3 ve bileşik yaratıcılık puanı ise 5,90 olarak bulunmuştur. Görüldüğü üzere akıcılık ve CQ arasında doğrusal bir ilişki yoktur. Çünkü CQ formülü logaritmik bir fonksiyondan elde edilmektedir. Ayrıca “ $\log_2(1+U_i)$ ” bağıntısının her kategori için ayrı ayrı hesaplanması gerektiğinden CQ puanlarının kategorinin de bir fonksiyonu olduğu görülmektedir. Yani yaratıcılık puanının çok olması için akıcılık gibi esneklik puanlarının da artması gerekmektedir. Yukarıdaki örnekte meyve kategorisi için alınan puan 2,32 iken, yanıtların hepsi farklı kategorilerde üretilmiş olsaydı 4 puan elde edilecekti. Bu bakımdan CQ’nun akıcılığın şişirme etkisini ve örneklemeden kaynaklanan problemleri baskıladığı düşünülebilir.

BÜT 3-5’in geliştirilme sürecinde, ön deneme ve pilot uygulamalarda elde edilen yanıtlar ile bir cevap anahtarı oluşturulmuştur. Cevap anahtarında üretilebilecek yanıtlar kavramsal kategorilere ayrılmıştır. Cevap anahtarı bir araştırmacı ve uzman tarafından gözden geçirilerek düzenlenmiştir. Asıl uygulamada bu cevap anahtarı kullanılarak öğrenci kağıtları araştırmacı tarafından puanlanmıştır.

3.3.4. BÜT 3-5’in uygulanması

BÜT 3-5 bir ders saatinde bir uygulayıcı tarafından, öğrencilere sınıflarında grup uygulaması şeklinde uygulanmaktadır. Testteki maddelere geçmeden önce öğrencilere çoğul üretim ile ilgili 1-2 dakikalık bir alıştırmaya yapılmaktadır. Uygulama esnasında her maddeye en az 8 dakika olmak üzere eşit süre verilmektedir. Testteki maddeler sırasıyla (1, 2, 3 ve 4), verilen süre içerisinde (her madde için en az 8 dakika) yanıtlanmaktadır. Örneğin birinci madde için 8 dakika verilmekte, süre dolduğunda ise tüm sınıfın ikinci maddeye geçmesi istenmektedir. Diğer maddeler için de aynı uygulama yapılmaktadır. 4 maddenin de bu şekilde uygulanmasından sonra, öğrencilere düzeltme ve eklemeler için ek 5 dakika zaman verilmektedir.

3.3.5. Fen bilimleri dersi akademik benlik algısı ölçeği (FABAÖ)

Araştırmada geliştirilen BÜT 3-5’in uyum geçerliğini test etmek için Akademik Benlik Algısı Ölçeğinin (ABAÖ) (Tokmak, 2016) fen bilimleri alt ölçeği (FABAÖ) katılımcı öğrencilere uygulanmış ve ölçekten elde edilen toplam puanlar analize dâhil edilmiştir.

FABAÖ 6 maddeden oluşan Likert tipi bir ölçektir. FABAÖ’nin orijinal formunda kullanılan 6’lı değerlendirme sistemi (1=Kesinlikle Yanlış, 2=Çoğunlukla Yanlış, 3=Yanlışın Yakını, 4=Doğruya Yakın, 5=Çoğunlukla Doğru, 6=Kesinlikle Doğru)

katılımcıların yaşına uygun olarak 5'li değerlendirme sistemine (1=Kesinlikle Katılmıyorum, 2=Katılmıyorum, 3=Kararsızım, 4=Katılıyorum, 5=Kesinlikle Katılıyorum) dönüştürülmüş (EK-4). Ölçekten alınabilecek en düşük puan 6, en yüksek puan ise 30'dur. Puanların 5'e yakın olması o alandaki düşük akademik benlik algısını, 30'a yaklaşması ise yüksek akademik benlik algısını ifade etmektedir. FABAÖ'nin geliştirilme sürecinde kapsam geçerliği için uzman görüşü alınmış, güvenilirlik analizi için Cronbach Alpha katsayısı hesaplanmış ve .89 olarak bulunmuştur (Sak, Tokmak ve Akbulut, 2016 incelemede). FABAÖ maddeleri Tablo 3.3'te verilmiştir.

Tablo 3.3. FABAÖ maddeleri

-
1. Fen bilimleri dersinden iyi notlar alırım.
 2. Fen bilimleri dersinin konularını çabuk öğrenirim.
 3. Fen bilimleri dersinde yetersizim.
 4. Yaşıtlarıma göre fen bilimleri dersinde iyiyimdir.
 5. Bugüne kadar fen bilimleri dersinde her zaman başarılı olmuşumdur.
 6. Fen bilimleri dersi konuları benim için kolaydır.
-

3.3.6. Bilime yönelik tutum ölçeği (BYTÖ)

Araştırmada geliştirilen BÜT 3-5'in uyum geçerliğini test etmek amacıyla kullanılan bir diğer ölçek de BYTÖ'dür. BYTÖ çalışma kapsamında araştırmacı tarafından geliştirilmiş, 5'li değerlendirme sistemi kullanılan (1=Kesinlikle Katılmıyorum, 2=Katılmıyorum, 3=Kararsızım, 4=Katılıyorum, 5=Kesinlikle Katılıyorum) ve 6 maddeden oluşan likert tipi bir tutum ölçeğidir (EK-4). Ölçeğin geçerlik çalışmalarında içerik ve kapsam geçerliği için uzman görüşü alınmıştır. Yapı geçerliği için faktör analizi yapılmış tek faktör altında toplanan 6 maddenin toplam varyansın %42.75'ini açıkladığı bulunmuştur. Bu değer kabul edilebilir değer olan %35'in üzerindedir (Akbulut, 2010 s. 93). Güvenirlik analizlerinde Cronbach Alpha iç tutarlık katsayısı hesaplanmış ve .816 olarak bulunmuştur. Bu değer ölçeğin iç tutarlığının iyi olduğunu göstermektedir (Pallant, 20011, s. 100). BYTÖ'nün geliştirilme süreci aşağıda detaylı bir şekilde verilmiştir.

3.3.6.1. BYTÖ'nün kuramsal altyapısı

Alan yazın incelendiğinde çok sayıda fen tutum ölçeğine rastlanmasına rağmen küçük yaş gurubuna uygun ölçeklerin sınırlılığı göze çarpmaktadır (Pell ve Jarvis, 2010,

s. 849). İlköğretim öğrencilerine uygun olan sınırlı sayıdaki fen tutum ölçeğinin ise Türkçe'ye uyarlanmadığı görülmektedir. Bu bağlamda araştırmada kullanılmak üzere bir fen tutum ölçeği geliştirilmiştir.

Tutumlar farklı şekillerde ölçülebilmektedir. Osborn, Simon ve Collins (2003, s. 1054) tutum ölçümlerini sıralama ölçekleri, davranış ölçekleri, ilgi ölçekleri, katılım belirleme ölçekleri ve nitel yöntemler olmak üzere beş grupta toplamaktadır. Nitel yöntemler küçük gruplarla yürütülen çalışmalara daha uygun görülmektedir. Katılım belirleme ve davranış ölçekleri sosyal-kültürel değişkenlerden etkilendiği düşünülebilir. “Boş zamanlarımda bilim müzelerine giderim” şeklinde bir ölçek maddesi bu duruma örnek olarak verilebilir. Sıralama ölçeklerinde ise bağıl bir değerlendirme söz konusu olduğundan ilişkisel araştırmalara uygun değildir. Bu bakımdan ilişkisel tarama modelinde kullanılarak geliştirilen BÜT 3-5'in uyum geçerliğini test etmek için geliştirilen BYTÖ'nin öğrencilerin fen bilimlerine yönelik ilgilerinin ölçülmesine karar verilmiştir.

Tutum ölçümü kadar tutumun kendisi de alan yazında sıcaklığı koruyan bir tartışma alanıdır. Hala kabul gören bazı görüşler davranışların tutumları etkilediğini savunurken (Allport, 1935, s. 820), bazıları ise tutumların davranışları etkilediğini savunmaktadır (Ajzen ve Fishbien, 1980, s. 138). Günümüzde ise genel olarak tutumların üç boyutunun olduğu birçok araştırmacı tarafından kabul edilmektedir;

- Bilişsel bileşen: nesnelere, inanışlar veya fikirler hakkında bilgi
- Duyuşsal bileşen: nesneye yönelik sevmek ya da sevmemek şeklindeki duygular
- Davranış bileşeni: nesneye yönelik eylemler (Reid, 2006, s. 4)

BYTÖ geliştirme sürecinde tutum boyutuna karar verme aşamasında katılımcı öğrencilerin özellikleri dikkate alınmıştır. Araştırmanın katılımcıları fen bilimleri alanında sınırlı tecrübeleri olan ilkököl ve ortaoköl öğrencileridir. Bu durum göz önünde bulundurularak BYTÖ'de fen dersi ve bilimle ilgili duyuşsal tutum maddelerinin geliştirilmesine karar verilmiştir. Tutumun önemli bir diğere özelliğı ise “tutumun bir hedefi olması” gerekliliğidir. Reid (2006, s. 7) fen bilimlerine yönelik tutumların hedeflerine göre dört gruba ayırmıştır;

- Bir disiplin olarak fen bilimleri (fen dersinin/bilimin kendisi)
- Fen bilimleri konularının öğrenilmesi (dersin zor/kolay olması)

- Herhangi bir konu başlığı (küresel ısınma)
- Fen bilimlerinde kullanılan yöntemler (örneğin deney yapmak)

Alan yazında karşılaşılan ölçeklerin bazıları sadece bir hedefe odaklanırken (örneğin özel olarak laboratuvarın çalışmalarına odaklanan Shah, 2004, s. 85) çoğu ölçeğin birden çok hedefi olduğu görülmektedir (örneğin, Pell and Jarvis, 2001, s. 856, fen öğrenimine, bilim hakkındaki görüşlere ve fen deneylerine odaklanmıştır.). Gardner (1995, s. 285) fen de tutum ölçümlerinde sonuçların yorumlanabilmesi açısından iç tutarlığı yüksek tek boyutlu ölçümleri önermektedir. Aynı şekilde Gardner'a (1975, s. 12) göre ayrı klasörler altında toplanmış çok sayıda madden bir ölçek oluşmaz. Çünkü bu ölçekten anlamlı puanlar elde edilemez. Bu gerekçelerden hareketle BYTÖ'de sadece bir hedefe yönelik maddelerin olmasına ve bu hedefin de "bir disiplin olarak fen bilimleri" olmasına karar verilmiştir. Öğrencilerin fen bilimleri alanındaki sınırlı bilgi birikimleri ve tecrübeleri, öğretmen farklılıkları ve okulların fiziksel imkanları (laboratuvar, bilim müzesi/fuarı) göz önünde bulundurularak diğer hedeflere yönelik tutum ölçümlerinin uygun olmayacağı düşünülmüştür. Ayrıca BYTÖ, FABAÖ ile birlikte kullanılacağından dolayı da fen öğrenme ile ilgili maddelere yer verilmemiştir.

Sonuç olarak, yukarıda sıralanan gerekçelerden dolayı BYTÖ'nin, hedefinde fen bilimleri olan (disiplin olarak) ve fen bilimlerine yönelik tutumların duyuşsal boyutuyla ifade edildiği bir ölçek olmasına karar verilmiştir. BYTÖ'nün geliştirme çalışmaları aşağıda detaylı şekilde verilmiştir.

3.3.6.2. BYTÖ geliştirilme sürecinde yürütülen çalışmalar

BYTÖ, dört aşamada geliştirilmiştir: a) tutum maddelerinin oluşturulması, b) uzman görüşü, c) pilot uygulama ve d) asıl uygulama. Bu aşamalarda yürütülen faaliyetler aşağıda sırası ile açıklanmıştır.

a. Tutum Maddelerinin oluşturulması: Tutum maddelerinin oluşturulmasına alan yazın çalışması ile başlanmıştır. Alan yazında ilkökul öğrencilerine uygun fen bilimleri tutum ölçeklerinde "disiplin olarak fen bilimleri" hedefine uygun ölçek maddeleri incelenmiştir. Daha sonra hedef kitlenin özelliklerine ve ölçek formatına uygun Tablo 3.4'te verilen 12 tutum maddesi geliştirilmiştir.

Tablo 3.4. *Geliştirilen fen bilimlerine yönelik tutum maddeleri*

1. Gelecekte bilim insanı olmak isterim.
2. Fen bilgisi dersinde başarılıyım.
3. Fen alanında yetenekli olduğumu düşünüyorum.
4. Fen bilgisi dersini severim.
5. Fen bilgisi dersini diğer derslerden daha çok severim.
6. Fen bilgisi dersinin daha çok yapılmasını isterim.
7. Fen bilgisi derslerinde sıkılırım.
8. Evde sıklıkla fen deneyleri yaparım.
9. Bilimle ilgili belgeseller izlemekten hoşlanırım.
10. Bilimsel gelişmelerle ilgili haberleri okumaktan ve izlemekten hoşlanırım.
11. Bilim ile ilgili kitap ve dergileri okumaktan hoşlanırım.
12. Bilim müzelerini gezmekten hoşlanırım.

b. Uzman Görüşü: Geliştirilen tutum maddeleri öncelikle bir dil uzmanına gönderilerek dil düzeltmeleri yapılmıştır. Dil düzeltmesi yapılan maddelerden uzman değerlendirme formu oluşturulmuş ve 4 fen bilimleri eğitim uzmanı ve 2 sınıf öğretmenine gönderilmiştir. Uzmanlar ve öğretmenler maddeleri anlaşılabilirlik, fen bilimlerine yönelik tutumları ölçüp ölçmediği noktasında değerlendirmişlerdir. Uzman görüşü neticesinde sosyal değişkenlerden etkileneceği gerekçesi ile 8 ve 12. madde, ölçekten çıkartılmıştır. Uzman görüşünden sonra araştırmada kullanılması planlanan ABAÖ ölçeğinde yer alan maddeler ile benzerliklerinden dolayı 2 ve 3. maddenin çıkartılmasına karar verilmiştir. Maddeler Pell ve Jarvis'in (2001, s. 855) önerdiği şekilde gülücükler (smile) eklenerek basitleştirilmiş ve pilot uygulama formu oluşturulmuştur.

c. Pilot Uygulama: Pilot uygulama formu Eskişehir merkezdeki bir ilkokul ve orta okula devam eden 120 öğrenciye (3. sınıf 27, 4. sınıf 32, 5. sınıf 61 öğrenci) uygulanarak elde edilen veriler üzerinden güvenilirlik ve geçerlik analizleri yapılmıştır. Güvenirlik analizleri için Cronbach Alpha iç tutarlık katsayısı hesaplanırken geçerlik çalışmalarında yapı geçerliği için açımlayıcı faktör analizi yapılmıştır.

Geliştirilen ölçeğin faktör yapısını keşfetmek amacıyla AFA yapılmıştır. Analize geçmeden önce veri seti faktör analizinin varsayımları bakımından test edilmiştir. Normallik sınaması için basıklık ve çarpıklık değerleri kontrol edilmiş, değişkenlere ilişkin dağılım grafiklerine bakılmıştır. Çok değişkenli normallik sınaması için KMO (Kaiser-Mayer-Olkin) ve Barlett test istatistikleri kontrol edilmiştir. Kline (1994, s. 74)

faktör analizi madde başına 10 katılımcının yeterli olacağını önermektedir. Faktör analizi çalışması 8 madde için 120 katılımcının verisi üzerinden yapıldığından örneklemin faktör analizi için yeterli olduğu düşünülebilir (madde başına 15 katılımcı).

Normallik sınaması için kontrol edilen dağılım grafiklerinin normal dağılıma yakın olduğu ve basıklık ve çarpıklık değerlerinin tüm değişkenler için +2 ve -2 arasında olmasından dolayı normal dağılım şartının sağlandığı sonucuna ulaşılmıştır (Trochim and Donnelly, 2006, s. 48). Veri setinin normal dağılıma yakın olmasından dolayı faktörleştirme metodu olarak, örneklem yerine evren için üretilmiş korelasyon matrisleri üzerinden çalışarak değişkenler arasında daha doğru bir kestirim sunan ve uyum indeksleri hesaplanabilen ve küçük örneklerde diğer yöntemlere göre daha iyi sonuç veren maksimum olasılık (maximum likelihood) yöntemi kullanılmıştır (Thompson, 2004, s. 38; Tanaka, 1987, s. 136). Veri setinin faktörleştirmeye uygunluğunu test etmek için KMO örneklem uygunluğu istatistiği ve Bartlett küresellik testleri kontrol edilmiştir. KMO test istatistiği .81 olarak hesaplanmış ve Bartlett testi anlamlı bulunmuştur [$\chi^2(28)=268,48$; $p<.001$]. KMO değerinin .80 üzerinde olmasında ve Bartlett testinin anlamlı çıkmasından dolayı veri setinin faktörleştirmeye uygun olduğu sonucuna ulaşılmıştır (Pallant, 2011, s. 183).

Direct oblimin döndürme işlemi ve maksimum benzerlikler metodu ile gerçekleştirilen faktör analizinde madde ayıklanmasında öz değerleri 1'in üzerinde olan maddeler ve faktör yük değerleri .45 ve daha yüksek olan maddeler önemli maddeler olarak kabul edilmiş ve binişiklik için .15 faktör yükü kriter olarak belirlenmiştir. Ayrıca ortak varyans yük değerleri düşük maddelerin iç tutarlılığı düşüreceği gerekçesi ortak varyans (communalities) için minimum yük değeri .30 olarak belirlenmiştir (Akbulut, 2010, s. 95). Bunun dışında madde ayıklamada Cronbach Alpha iç tutarlık katsayılarına da dikkat edilerek Gardner'ın (1975, s. 12; 1995, s. 285) tavsiye ettiği şekilde iç tutarlığın yüksek olacağı şekilde tek faktörlü bir yapının elde edilmesine özen gösterilmiştir.

Faktör analizi bulgularına göre 8 maddenin 2 faktör altında toplam %46,03 varyans açıklayacak şekilde toplandığı görülmüştür. Ancak ortak varyans tablosunda 1 ve 7 numaralı maddelerin ortak varyans yük değerlerinin sırasıyla .275 ve .255 olduğu görülmüştür. Ortak varyans yük değerlerin .30'un altında olduğundan bu iki maddenin ölçeğin tamamına hizmet etmediği düşünülebilir (Field, 2009, s. 638). Bu durumun ölçeğin iç tutarlılığını olumsuz yönde etkileyip etkilemediğini belirlemek için 8 madde için Cronbach Alpha iç tutarlık katsayısı hesaplanmış ve .789 olarak bulunmuştur. Bu

değer ölçek geliştirme çalışmaları için kabul edilebilir değer olan .80'in altındadır (Pallant, 2011, s. 100). Bu duruma neden olan maddelerin belirlenmesi için düzeltilmiş madde-toplam korelasyonlarına bakılmış ve 1. ve 7. maddeler için düşük değerler saptanmıştır ($r_1 = .263$ ve $r_7 = .343$). Son olarak madde-toplam bulgularında Cronbach Alpha değerleri kontrol edilmiş ve 1. ve 7. maddelerin ölçekten çıkartılması halinde Cronbach Alpha iç tutarlık katsayısının .80'in üzerine çıkacağı görülmüştür. Bu bulgulara dayanarak 1 ve 7. maddelerin ölçekten çıkartılmasına karar verilmiştir.

1 ve 7. maddeler ölçekten çıkartılarak geriye kalan 6 madde ile analizler tekrarlanmıştır. 6 madde için analiz sonuçlarına göre tüm maddeler tek bir faktörde toplanmıştır. Elde edilen bu yapı toplam varyansın %42,75' ini açıkladığı görülmüştür. Maddelerin faktör yük değerleri ise .56 ile .74 arasında değişmektedir. Stevens (2002) örneklem sayısına bağlı olarak minimum faktör yük değerini 100 kişilik bir örneklem için .512 olarak hesaplamıştır. Bu bakımdan ölçeği oluşturan maddelerin faktör yüklerinin kabul edilebilir seviyede olduğu söylenebilir. Ölçeğin son halinde 6 madde için bulunan Cronbach Alpha iç tutarlık katsayısı ise .82 olarak hesaplanmış, düzeltilmiş madde-toplam korelasyonlarının ise .52-.640 arasında değerler aldığı görülmüştür. Bu bulgular Gardner (1995, s.284) tarafından önerilen tek boyutlu yapının kanıtları olarak değerlendirilebilir. İç tutarlık bulgularına göre ölçeğin iç tutarlık güvenilirliğinin yüksek olduğu söylenebilir (Briggs and Cheek, 1986, s. 115).

3.3.7. Karne notları

Araştırma kapsamında geliştirilen BÜT 3-5'in uyum geçerliğini test etmek için kullanılan verilerden biri öğrenci karne notlarıdır. Öğrenci karne notları uygulama yapılan sınıflar için okul yönetiminden izin kapsamında temin edilmiştir. Araştırmanın veri toplama aşaması 2016-2017 eğitim öğretim yılı, bahar yarıyılında gerçekleştirildiğinden öğrencilerin birinci yarıyıl karne notları kullanılmıştır.

MEB tarafından 1, 2 ve 3. sınıf öğrencilerinin ders notları üçlük not sistemi ile değerlendirilerek karneye “geliştirilmeli”, “iyi” ve “çok iyi” şeklinde işlenmektedir. Bu bakımdan 3. sınıf katılımcıların fen bilimleri ve matematik ders notları analiz dışı bırakılmıştır. İlkokul 4. sınıftan itibaren öğrencilerin ders notları 100 puan üzerinden değerlendirilmektedir. Öğrenci karne notları dönem içinde yapılan sınav puanları ve ders içi etkinliklere katılım puanlarının aritmetik ortalaması alınarak hesaplanmakta ve öğrenci not sistemine işlenmektedir.

3. sınıflar öğrencilerinin karne notları analiz dışı bırakılmıştır. Bu durumun birinci nedeni not sistemindeki farklılıktır. 4 ve 5. sınıfların karne notları 100'lük sisteme göre değerlendirilirken, 3. sınıfların karne notları 3'lük sisteme göre değerlendirilmektedir. İkinci neden ise 3. sınıfların ders notlarında değişkenlik olmamasıdır ($\bar{X}_{Fen} = 2,96$; $SS = 0,19$; $\bar{X}_{Matematik} = 2,91$; $SS = 0,90$). 251 öğrencinin karne notları incelendiğinde fen karne notlarının 241 tanesinin "3" olduğu, matematik karne notlarının ise 232 tanesinin "3" olduğu görülmüştür. Bu dağılımın korelasyon analizi için uygun olmadığı ve bulguların yanlış yorumlanacağı (Goodwin and Leech, 2006, s. 252; Wilcox, 2016, s. 215) gerekçesi ile 3. sınıf öğrencilerinin karne notları uyum geçerliği analizlerinde kullanılmamıştır. Bu bakımdan sadece 4. ve 5. sınıf öğrencilerinin fen bilgisi ve matematik ders notları 100'lük not sistemi üzerinden analizlere dâhil edilmiştir.

3.4. Verilerin Çözümlemesi ve Yorumlanması

Testin puanlanması araştırmacı tarafından yapılmıştır. Kağıtlar okunurken madde geliştirme, ön deneme ve pilot uygulama aşamalarında elde edilen cevap anahtarı kullanılmıştır. Cevap anahtarı bir uzman tarafından kontrol edilmiş ve düzeltmeler yapılmıştır. Yeni cevap gelmesi durumunda ise aynı uzmandan görüş alınarak cevap anahtarına eklenmiştir. Çözümlenen veriler analizinde IBM SPSS ve LISREL istatistik programları kullanılmıştır.

Birinci araştırma sorusu olan BÜT 3-5 Kuramsal Yapısının geçerliği temel bileşen analizi (PCA) ve doğrulayıcı faktör analizleri (DFA) ile incelenmiştir. Testin son halinde 4 madde (2 deney ve 2 hipotez) olmasına karar verildiğinden PCA bulguları aynı zamanda madde seçiminde de kullanılmıştır. Elde edilen 4 madde ile BÜT 3-5'in kuramsal yapısı LISREL istatistik programı kullanılarak DFA ile test edilmiştir. Testin sınıf düzeylerine ve yetenek seviyelerine göre ayırt edicilik geçerliğini test etmek için ANOVA istatistikleri kullanılmıştır. Uyum geçerliği analizleri ise test puanları ile ders notları, BYTÖ ve FABAÖ puanları arasında hesaplanan Pearson momentler çarpımı korelasyon katsayıları üzerinden yapılmıştır.

BÜT 3-5'in güvenilirlik analizleri için öncelikle Cronbach Alpha iç tutarlık katsayısı hesaplanmış, madde-toplam ve maddeler arası Pearson momentler çarpımı korelasyon katsayıları değerlendirilmiştir. Puanlayıcılar arası güvenilirlik analizleri için rastgele seçilen 100 kâğıt okunduktan sonra ikinci bir okuyucuya cevap anahtarı ile birlikte verilmiş ve bağımsız olarak ikinci okuyucu tarafından tekrar puanlanmıştır. İkinci

okuyucu yaratıcılık ve bilimsel yaratıcılık konusunda uzman, BÜT puanlama eğitimi olan tecrübeli bir alan uzmanıdır. Puanlayıcılar arası güvenilirliğin belirlenmesi için iki okuyucunun puanları arasında sınıf içi korelasyon korelasyon katsayıları hesaplanmıştır.

3.5. Veri toplama süreci

Veri toplama sürecine başlanmadan önce T.C. Eskişehir İl Millî Eğitim Müdürlüğü'nden gerekli izinler alınarak uygulama yapılacak okullar belirlenmiştir. Okul idarecileri ile görüşülerek araştırmanın amacı ve uygulama hakkında bilgi verilmiştir. Uygun olan iki ilkokul ve bir ortaokulda uygulamanın yapılmasına karar verilmiştir. Okullardaki uygulamalarda önce BÜT 3-5 uygulanmış, daha sonra öğrencilerin karne notları okul idaresinden temin edilmiştir.

BÜT 3-5'in asıl uygulaması 2016-2017 eğitim-öğretim yılı bahar yarı yılında ders saatleri içerisinde öğrencilerin kendi sınıflarında uygulanmıştır. Test iki kitapçık hâlinde iki oturumda verilmiştir. İlk oturumda hipotez kitapçığı uygulanmış, bir ya da iki gün sonra ikinci oturumda deney kitapçığı uygulanmıştır. Uygulama esnasında her madde öğrencilere okunmuş ve görseller açıklanmıştır. Her soruya öğrencilerin yanıtlaması için 8 dakika zaman verilmiş, tüm öğrenciler aynı maddeyi birlikte yanıtlamış, 8 dakika dolduğunda ise sıradaki maddeye geçilmiştir. Öğrencilerden de yönergeleri dikkatlice okumaları istenmiştir. Testten önce öğrencilerle çoğul üretim alıştırmaları yapılmıştır. Bu alıştırmada aşağıda verilen görev üzerinden öğrencilerin tekil ve çoğul üretim arasındaki farkı anlamaları ve çoğul üretime yönlendirilmeleri hedeflenmiştir:

Uygulayıcı: *“Karşılaştığımız bazı soruların ve problemlerin bir tek yanıtı vardır. Örneğin Türkiye'nin başkenti neresidir?”*

Öğrenciler: *Ankara*

Uygulayıcı: *Peki bu sorunun başka yanıtı var mı? Bu soru tek yanıtı olan sorulardan. Peki bildiğimiz yuvarlak şeyler nelerdir?*

Öğrenciler: *top, portakal, ampul.....*

Uygulayıcı: *peki bu sorunun kaç yanıtı?*

Öğrenciler: *bir sürü*

Uygulayıcı: *Evet haklısınız. Bazı soruların bir sürü yanıtı olabilir. İşte bugün burada yapacağımız çalışmadaki sorularında bir sürü doğru yanıt. O bakımdan soruları dikkatlice okuduktan sonra üretebileceğiniz kadar çok ve farklı yanıtlar üretmenizi istiyorum.*

Ürettiğiniz her yanıt için bir puan, çok farklı yanıtlar için de ek puan alacaksınız.

Veri toplama aşamasında ekonomik olması açısından 8 uygulayıcı eğitilmiştir. Uygulama yapılacak okula gidildiğinde uygulayıcıların hepsi aynı anda birer sınıfta uygulama yapmışlardır. Bu sayede öğrencilerin birbirlerinden etkilenmelerinin önüne geçilmeye çalışılmıştır. Diğer veri toplama aracı olan BYTÖ ilk oturumda 3 madde olduğu için hipotez kitapçığı ile birlikte öğrencilere dağıtılmış ve testten önce uygulanmıştır. Uygulama iki farklı günde yapıldığından iki oturumdan herhangi birine katılmayan öğrenciler araştırmaya dâhil edilmemiştir.



4. BULGULAR

Bulgular kısmında öncelikle madde analizleri ve arkasından PCA analizi sonucunda yapılan madde seçimi sunulmuştur. Yapılan madde seçimi sonrasında elde edilen 4 madde üzerinden geliştirilen testin son formunun psikometrik özellikleri incelenmiştir. Madde seçimi ve psikometrik özelliklerin farklı örneklemelerden gelen veriler üzerinden yapılması amacıyla veri seti rastsal olarak iki guruba ayrılmıştır. DFA, 200 katılımcıya duyarlı bir analiz yöntemidir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s. 268). Katılımcı sayısının 200'ün üzerine olması test edilen modelin uyumunu güçleştirmekte ve sonuçların yorumlanmasını zorlaştırmaktadır. Bu sınırlılığı gidermek için Floyd ve Widaman (1995, s. 290) örneklemin bölünerek analizlerin yapılmasını önermektedirler. Ayrıca DeVilles, (2003, s. 99) ölçek geliştirme sürecinde yeterli büyüklükteki örneklemelerin bölünerek birinci örneklemin geliştirme aşamasının bir parçası olarak, ikinci örneklemin ise geliştirme sürecindeki bulguların test edilmesinde kullanılmasını önermektedir. Bu bağlamda SPSS programı kullanılarak katılımcılar rastsal 211 ve 436 (yaklaşık %32) olmak üzere iki guruba ayrılmıştır. Birinci gurubun verileri ile PCA yapılmış ve uygulama formunda 4 madde seçilmiştir. İkinci grupta ise ölçeğin son hali olan BÜT 3-5'in geçerlik ve güvenirlik analizleri yapılmıştır.

4.1. BÜT 3-5'in Geçerliğine İlişkin Bulgular

BÜT 3-5'in geçerliğine ilişkin bulgular yapı geçerliği, ayırt edicilik geçerliği ve uyum geçerliğine ilişkin bulguları kapsamaktadır. Yapı geçerliği için PCA ve DFA bulgularından, ayırt edicilik geçerliği için t-testi ve ANOVA fark testi bulgularından, uyum geçerliği içinde korelasyon katsayılarından yararlanılmıştır.

Geçerlik çalışmaları için öncelikle madde analizleri yapılmıştır. Madde analizleri için uygulama formundaki maddelerin betimsel istatistikleri çıkartılmış ve maddeler arası korelasyon katsayıları hesaplanmıştır. Uygulama formuna ilişkin betimsel bulgular Tablo 4.1'de verilmiştir.

Tablo 4.1. *Uygulama formunun betimsel bulguları*

Madde	Puan	\bar{X}	SS	Minimum	Maksimum
1. Hipotez	Akıcılık	1,742	1,549	,00	7,00
	Esneklik	1,280	0,941	,00	4,00
	B. Yaratıcılık	1,519	1,206	,00	5,94
2. Hipotez	Akıcılık	2,139	1,593	,00	8,00
	Esneklik	1,474	0,9775	,00	4,00
	B. Yaratıcılık	1,824	1,239	,00	5,58
3. Hipotez	Akıcılık	3,424	1,890	,00	10,00
	Esneklik	2,167	0,9621	,00	5,00
	B. Yaratıcılık	2,838	1,388	,00	7,32
4. Deney	Akıcılık	2,019	2,635	,00	12,00
	Esneklik	1,366	1,548	,00	5,00
	B. Yaratıcılık	1,711	2,082	,00	8,16
5. Deney	Akıcılık	1,892	2,433	,00	13,00
	Esneklik	1,196	1,350	,00	4,00
	B. Yaratıcılık	1,559	1,867	,00	8,22
6. Deney	Akıcılık	2,629	1,812	,00	9,00
	Esneklik	1,769	0,9849	,00	4,00
	B. Yaratıcılık	2,217	1,343	,00	6,00
7. Deney	Akıcılık	3,434	1,914	,00	11,00
	Esneklik	2,241	1,101	,00	5,00
	B. Yaratıcılık	2,890	1,477	,00	7,58

Tablo 4.1’de akıcılık puanlarına baktığımızda maddeler için üretilen maksimum yanıtların 7 (1. Madde) ile 13 (5. Madde) arasında değiştiği görülmektedir. Bu bulguya göre 5. madde için en az 13 farklı yanıt, 1. madde içinde en az 7 yanıt üretilmiştir. Tablodan maksimum esneklik puanlarının 5 olduğu görülmektedir. Buna göre maddelere üretilebilecek yanıtlar 5 farklı kategorinin altında toplanmaktadır. Bileşik yaratıcılık puanları ise akıcılık puanlarından düşük ama esneklik puanlarından yüksektir. Bu bağlamda betimsel bulgulardan hareketle maddelerin çoğul üretime uygun olduğu düşünülebilir.

Maddeler arası korelasyon katsayıları hesaplanmış ve Tablo 4.2’de verilmiştir.

Tablo 4.2. Uygulama formu maddeler arası korelasyon katsayıları

Madde	Akıcılık							Esneklik							Bileşik Yaratıcılık						
	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7	
Akıcılık	1	,45**	,43**	,26**	,30**	,42**	,36**	,86**	,36**	,37**	,28**	,32**	,39**	,37**	,97**	,43**	,43**	,27**	,32**	,43**	,38**
	2	1	,43**	,29**	,26**	,33**	,34**	,37**	,83**	,38**	,29**	,28**	,30**	,35**	,43**	,97**	,44**	,30**	,27**	,34**	,36**
	3		1	,41**	,39**	,42**	,44**	,37**	,38**	,83**	,38**	,37**	,37**	,45**	,42**	,43**	,98**	,41**	,39**	,42**	,46**
	4			1	,71**	,29**	,34**	,25**	,27**	,31**	,94**	,66**	,29**	,31**	,26**	,30**	,38**	,99**	,70**	,30**	,34**
	5				1	,31**	,28**	,28**	,23**	,32**	,68**	,92**	,31**	,29**	,30**	,27**	,37**	,71**	,99**	,33**	,30**
	6					1	,51**	,42**	,31**	,37**	,28**	,31**	,82**	,48**	,44**	,34**	,42**	,29**	,31**	,97**	,51**
	7						1	,33**	,33**	,40**	,33**	,28**	,45**	,85**	,36**	,35**	,44**	,34**	,28**	,51**	,98**
Esneklik	1						1	,31**	,33**	,28**	,31**	,39**	,35**	,95**	,36**	,37**	,27**	,29**	,41**	,35**	
	2							1	,35**	,28**	,26**	,27**	,32**	,35**	,93**	,39**	,28**	,24**	,31**	,34**	
	3								1	,31**	,29**	,34**	,41**	,36**	,39**	,92**	,32**	,31**	,38**	,41**	
	4									1	,67**	,29**	,33**	,29**	,30**	,36**	,98**	,69**	,30**	,34**	
	5										1	,32**	,31**	,33**	,29**	,35**	,67**	,97**	,327**	,30**	
	6											1	,45**	,40**	,30**	,37**	,29**	,32**	,93**	,46**	
	7												1	,37**	,35**	,45**	,32**	,30**	,49**	,93**	
Bileşik Yaratıcılık	1													1	,41**	,42**	,28**	,32**	,44**	,38**	
	2														1	,44**	,31**	,28**	,31**	,36**	
	3															1	,38**	,37**	,42**	,46**	
	4																1	,71**	,30**	,35**	
	5																	1	,33**	,31**	
	6																		1	,52**	
	7																			1	

**p<.001

Tablo 4.2'ye göre maddeler arası akıcılık için .258-.708 arasında, esneklik için .255-.671 arasında ve bileşik yaratıcılık için .276-.706 arasında anlamlı korelasyon katsayıları bulunmuştur. Buna göre uygulama formundaki maddelerin birbirleri ile ilişkili oldukları ve aynı yapıya hizmet ettikleri düşünülebilir. Aynı maddelerin akıcılık ve esneklik puanları arasında .825-.939, akıcılık ve yaratıcılık arasında .971-.990 ve bileşik yaratıcılık ve esneklik arasında .919-.977 düzeyinde anlamlı korelasyonlar bulunmuştur. Bu bulguya göre öğrencilerin doğru yanıtları arttıkça yanıtların üretildiği kategoriler ve bileşik yaratıcılık puanları da artmaktadır. Bu bağlamda geçerlik ve güvenirlik analizlerinde akıcılık, esneklik ve bileşik yaratıcılık puanlarının benzer sonuçlar vereceği düşünülebilir.

4.1.1. BÜT 3-5'in yapı geçerliğine ilişkin bulgular

Yapı geçerliği analizlerinde öncelikle *Uygulama Formunda* yer alan 7 madde için PCA bulgularından yararlanılarak madde seçimi yapılmış ve BÜT 3-5'in en uygun 4 maddesi belirlenmiştir. BÜT 3-5'in son formu oluşturulmuştur. Daha sonra BÜT 3-5'in yapı geçerliğini ortaya koymak amacıyla PCA bulguları incelenmiş, arkasından elde edilen yapının DFA ile doğrulanıp doğrulanmadığı kontrol edilmiştir. PCA ve DFA analizlerinde yalnızca akıcılık puanları kullanılmıştır çünkü akıcılık puanları ile diğer puan türleri arasında yüksek korelasyon bulunmakta, varyans daha yüksek ve akıcılığın yorumlanması daha kolaydır.

4.1.1.1. BÜT 3-5'in yapı geçerliğine ilişkin PCA bulguları

Ölçek geliştirme sürecinde testin uzunluğunun madde seçimi sürecinde belirlenmesi önerilmektedir (DeVilles, 2003; Whortington and Whittaker, 2006, s. 809; Clark and Watson, 1995, s. 314). Kuramsal alt yapı, öğrencilerin yaşı ve testin uzunluk süresi dikkate alınarak testin son halinde 2 deney sorusu ve 2 hipotez sorusu olmasına karar verilmiştir. Bu bakımdan uygulama formunda yer alan 3 hipotez sorusundan 2'sinin ve 4 deney sorusundan 2'sinin seçilmesi için 436 katılımcının verisi ile PCA yapılmıştır.

PCA fazla sayıdaki değişkeni, küçük sayıdaki değişken altında azaltarak toplamak için kullanılan (Çokluk, Şekercioğlu, Büyüköztürk, s. 198) ve kuramsal yapının parçalarını ortaya çıkartan bir yöntemdir (Şencan, 2005). Her ne kadar alan yazında AFA yöntemleri önerilse de PCA ve AFA sonuçlarının farklılaşmadığını ortaya koyan, özellikle bileşen sayıları eşit olan analizlerde yüksek benzerlikler rapor eden birçok

araştırmaya rastlanmaktadır (Velecier and Jackson, 1990, s. 5; Arrindell and Van der Ende, 1985, s. 174; Pallant, 2011, s.182). Field (2009, s. 636) analizin amacına uygun yöntemin belirlenmesini önermektedir. Bu çalışma bir ölçek geliştirme çalışmasıdır ve burada kullanılan analizin amacı 7 madde içinden kuramsal yapı ile uyumlu 4 maddenin belirlenmesidir. Analiz sonuçları öncelikle madde seçiminde kullanılacağı için PCA yöntemi ilk tercih olarak değerlendirilebilir (Widaman, 1993, s. 308; Floyd and Widaman, 1995, s. 287; Matsunaga, 2010, s. 99). İkinci olarak faktör belirsizliklerinden arınmış, basit ve yorumlanması kolay çözümler sunması bakımından PCA (Tabachnick and Fidell, 2012, s. 613) yöntemi kullanılmaya karar verilmiştir. Çünkü bu yöntemde değişkenler arasında doğrusal ilişkiler olduğu varsayılır ve bundan dolayı gözlenen varyanslar maksimize edilmiş olur. Varyansın çoğunu açıklayan çözümler üretmesinden dolayı PCA ile elde edilen bileşenler, takip eden analizlerde tahmin veya kriter değişken olarak kullanılabilir. Elde edilen yapılar ile öğrencilerin bilimsel yaratıcılık düzeyleri belirlenmesi hedeflendiğinden PCA yapılmasına karar verilmiştir.

Analize başlamadan önce veri seti PCA varsayımları bakımından test edilmiştir. Öncelikle veri seti normal dağılım bağlamında değerlendirilmiştir. PP ve QQ grafikleri ve histogram grafiklerine bakılmış ve değişkenlerin normal dağılımdan aşırı sapma göstermedikleri görülmüştür. Basıklık (kurtosis) ve çarpıklık (skewness) değerlerinin tüm değişkenler için +2 ve -2 arasında olmasından dolayı (Trochim and Donnelly, 2006, s. 48) veri setinin tüm değişkenler bakımından normal dağılıma yakın bir dağılım gösterdiği sonucuna varılmıştır. Veri setinin faktörleştirmeye uygunluğunu test etmek için KMO örneklem uygunluğu istatistiği ve Bartlett küresellik testleri kontrol edilmiştir. KMO test istatistiği .81 olduğu Bartlett testinin anlamlı çıktığı görülmüştür [$\chi^2(21)=528,63$; $p<.001$]. KMO değerinin .80 üzerinde olmasında ve Bartlett testinin anlamlı çıkmasından dolayı veri setinin faktörleştirmeye uygun olduğu sonucuna varılmıştır (Pallant, 2011, s. 183). Karşıt görüntüyle örneklem yeterliğinin ölçülmesi (*Anti-image Measures of Sampling Adequacy, MSA*) için her madde için MSA değerleri hesaplanmış ve tamamının .60'ın üzerinde olmasından dolayı maddelerin tamamının PCA için uygun olduğu söylenebilir (Tabachnick and Fidell, 2012, s. 661). Bu bağlamda verilerin çok değişkenli normal dağılımdan geldiği kabul edilmiş ve PCA işlemlerine geçilmiştir.

Yapılan PCA'de öncelikle ortak varyans (communalities) değerleri kontrol edilmiştir. Maddelerin ortak varyans değerlerinin .474 ve .851 arasında değiştiği

görülmüştür. Ortak varyans değerlerinin tamamının .30'un üzerinde olması analize dahil edilen maddelerin tamamının ölçeğe hizmet ettiği şeklinde yorumlanmıştır (Field, 2009, s. 639). Bileşen sayısını belirlemek için sıklıkla kullanılan Kaiser kriterine göre öz değeri (eigenvalue) 1'in üzerinde olan bileşenlere bakılmıştır. PCA çözümünde öz değeri 1'in üzerinde iki bileşen tespit edilmiş ve bu iki bileşenin %61,94 varyans açıkladığı görülmüştür. Bileşenlerin varyansa yaptıkları katkıların sırasıyla %44,94 ve %17 olduğu saptanmıştır. Bileşenlerin önemini belirleyerek bileşen sayısına karar vermek amacıyla bileşenlerin açıkladığı varyans değerleri yamaç birikinti grafiği (scree plot) ile birlikte değerlendirilmiştir. Yamaç birikinti grafiğinin 3. noktadan itibaren düzleştiği görülmüştür. Bileşen sayısını netleştirmek için Kaiser kriterinden daha iyi sonuçlar verdiği düşünülen (Thompson, 2004, s. 22) paralel analiz yapılmıştır. Monte Carlo simülasyon programı (Watkins, 2004) ile 436 katılımcı ve 7 madde için rast gele 1000 öz değer üretilmiştir. Daha sonra bu öz değerler Tablo 3.3'de verilen PCA'den elde edilen öz değerlerle karşılaştırılmıştır. PCA'den elde edilen 2 öz değer simülasyon ile elde edilen öz değerlerden daha yüksek olduğu ve 2 bileşenli bir yapıyı işaret ettiği görülmüştür (Pallant, 2011, s. 193).

Tablo 4.3. Monte Carlo simülasyonu ile yapılan paralel analiz sonuçları

Bileşen No	PCA Öz değerler	Monte Carlo Öz değerler	Karar
1	3,1467	1,183	Kabul
2	1,190	1,106	Kabul
3	,791	1,047	Ret
4	,594	,996	Ret
5	,534	,945	Ret
6	,472	,839	Ret
7	,284	,827	Ret

BÜT 3-5 kuramsal yapısı hipotez oluşturma ve deney tasarlama olmak üzere iki bileşenden oluşmaktadır. PCA'da belirlenen bileşen sayısı kuramsal yapı ile uyumludur. Bu bakımdan PCA bulgularının daha kolay yorumlanabilmesi için maksimum değişkenlik (varimax) dik döndürme yöntemi döndürme işlemleri için seçilmiş ve bileşen sayısı 2 olarak belirlenip analiz tekrarlanmıştır. Tekrarlanan PCA bulgularına iki bileşene ait açıklanan toplam varyans değerleri Tablo 3.4'te verilmiştir.

Tablo 4.4. BÜT 3-5 uygulama formu PCA için açıklanan toplam varyans

Bileşen	Başlangıç Öz Değerleri			Yüklerin Kareleri Toplamında Çıkarılan			Yüklerin Kareler Toplamının Döndürülmesi		
	Toplam	Varyans %	Kümülatif	Toplam	Varyans %	Kümülatif	Toplam	Varyans %	Kümülatif
			Varyans %			Varyans %			
1	3,146	44,942	44,942	3,146	44,942	44,942	2,522	36,026	36,026
2	1,190	16,994	61,936	1,190	16,994	61,936	1,814	25,910	61,936
3	0,791	11,295	73,231						
4	0,594	8,491	81,722						
5	0,524	7,482	89,204						
6	0,472	6,739	95,943						
7	0,284	4,057	100,000						

Tablo 4.4'te görüldüğü üzere iki bileşen toplam varyansın %61,94'ünü açıklamaktadır. İlk bileşenin varyansa katkısı %44,94 ve ikinci bileşenin katkısı %17 olarak bulunmuştur. Her ne kadar birinci bileşenden ikinci bileşene doğru açıklanan varyansta sert bir düşüş gözlemlense de döndürme işleminden sonra açıklanan varyans oranlarının sırasıyla %36,03 ve %25,91 olduğu görülmektedir. Döndürme işleminde sonra maddelerin bileşenlere göre dağılımı ve bileşen yük değerleri Tablo 4.5'te verilmiştir.

Tablo 4.5. BÜT 3-5 uygulama formu maddelerinin bileşenlerdeki faktör yükleri

Madde	Alan	Görev	Bileşenler	
			Hipotez Oluşturma	Deney Düzenleme
1	Fizik	Hipotez Oluşturma	,725	
2	Fizik/Kimya	Hipotez Oluşturma	,703	
3	Biyoloji	Hipotez Oluşturma	,669	
6	Fizik	Deney Düzenleme ^a	,750	
7	Kimya	Deney Düzenleme ^a	,654	
4	Biyoloji	Deney Düzenleme ^b		,905
5	Kimya	Deney Düzenleme ^b		,889

^a 1. Tip deney düzenleme görevi; ^b 2. Tip deney düzenleme görevi

Tablo 4.5'te görüldüğü üzere deney ve hipotez bileşenleri altında toplanan maddelerin bileşen yük değerleri .654 ve .905 arasında değişmektedir. 1, 2, 4, 5 ve 6. maddelerin yük değerlerinin “mükemmel” 3 ve 7. maddelerin yük değerleri ise “çok iyi” olduğu görülmektedir (Comrey and Lee, 1992, s. 243).

Tablo 4.5'te 1. Tip deney görevlerinin hipotez bileşeni altında toplandığı ve iyi faktör yüklerine sahip olduğu görülmektedir. Bu bulgu pilot uygulama bulgularıyla da uyumludur. 2. Tip deney görevleri ise ayrı bir bileşen olarak deney bileşeninin altında mükemmel bileşen yüklerine sahiptir. Yüksek bileşen yükleri maddelerin deney bileşeninin saf birer ölçümü olduğu anlamına gelmektedir (Tabachnick and Fidell, 2012, s. 654). Bu bağlamda 4 ve 5. maddelerin BÜT 3-5'e deney düzenleme görevleri olarak seçilmişlerdir.

6 ve 7. maddeler testten çıkartıldıktan sonra PCA tekrar edilmiş ve geriye kalan maddelerin hipotez (1, 2 ve 3. maddeler) ve deney (4 ve 5. maddeler) bileşenlerinin altında toplanarak %72,68 varyans açıkladığı görülmüştür. Ortak varyan değerlerinin ise .602 ve .845 arasında değiştiği, faktör yük değerlerinin ise .710 ve .900 arasında değiştiği görülmüştür. 2 hipotez maddesinin belirlenmesi için faktör yük değerlerinin ve ortak varyansların birbirlerine yakın olduğu görülmüştür. Madde seçimi için beş madde için Cronbach Alpha iç tutarlık katsayısı (.754) hesaplanmıştır (Akbulut, 2010, s. 90; Hinkin 1995, s. 977; DeVilles, 2003, s. 94). 1, 2 ve 3. maddelerin testten çıkartılması hâlinde Cronbach Alpha değeri sırasıyla .726, .735 ve .706 olarak hesaplanmıştır. Farklı bir ifade ile 2. maddenin çıkartılması halinde testin iç tutarlığının daha yükse olacağı söylenebilir. Bu bağlamda 1 ve 3. maddeler BÜT 3-5'in hipotez maddeleri olarak seçilmiştir.

Yapılan PCA ile uygulama formunda yer alan 7 madde içerisinde 2 deney ve 2 hipotez görevi belirlenmiş ve BÜT 3-5 için seçilmiştir. Son 4 madde ile tekrarlanan PCA bulguları Tablo 4.6'da verilmiştir.

Tablo 4.6. BÜT 3-5 PCA bulguları

Madde	Hipotez Bileşeni	Deney Bileşeni	Ortak Faktör Varyansı (h^2)
1	.874	.103	.774
2	.789	.269	.695
3	.194	.901	.850
4	.185	.898	.840
Varyans Oranları	%36,45	%42,55	
Açıklanan Toplam Varyans	%78,99		

Tablo 4.6'da görüldüğü üzere hipotez ve deney bileşenleri toplam varyansın %78,99'unu açıklamaktadır. Açıklanan varyansın sosyal bilimler için yeterli düzeyde

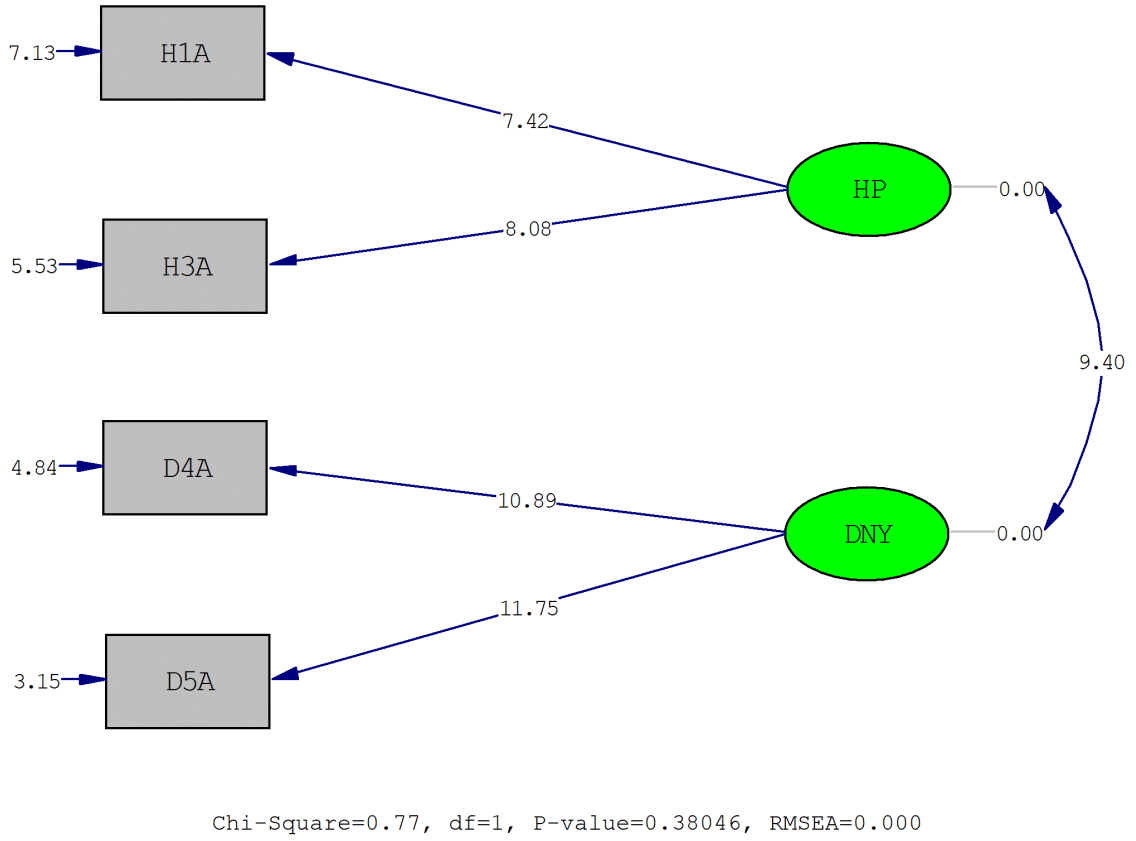
olduğu kabul edilebilir (Çokluk, Şekercioğlu ve Büyüköztürk, 2010, s. 197). Açıklanan toplam varyansa hipotez bileşeninin katkısı %36,45 ve deney bileşeninin katkısı ise %42,55'tir. Bu bakımdan varyansa katkılarının önem derecesinin birbirine yakın olduğu düşünülebilir. Maddelerin ortak faktör varyanslarının ise tamamının .30'un üzerinde olması maddelerin tamamının ölçeğin amacına hizmet ettiği şeklinde yorumlanabilir (Field, 2009, s. 639). Faktör yük değerlerinin ise tüm maddeler için mükemmel seviyede olduğu görülmektedir (Comrey and Lee, 1992, s. 243). Bu bağlamda PCA bulguları, BÜT 3-5'in kuramsal yapısına deneysel kanıtları olarak değerlendirilebilir.

4.1.1.2. BÜT 3-5'in yapı geçerliğine ilişkin DFA bulguları

BÜT 3-5'in PCA bulguları ile desteklenen kuramsal yapısı DFA analizi ile incelenmiştir. DFA işlemlerinde BÜT 3-5'te yer alan 2 deney ve 2 hipotez olmak üzere 4 madde üzerinden 211 katılımcının verisi üzerinden yürütülmüştür. Analizler LISREL programı kullanılarak öğrencilerin akıcılık puanları üzerinden yapılmıştır.

Veri seti öncelikle DFA varsayımları için kontrol edildiğinde kayıp değerler olmadığı görülmüştür. Uç değerlerin kontrol edilmesi için puanlar z-puanlarına dönüştürülmüş ve +3 ve -3 aralığının dışında kalan 6 katılımcı veri setinden çıkartılmıştır (Tabachnick and Fidell, 2012, s. 73). AFA 200 katılımcıya duyarlı olduğu, daha büyük örneklerde uyum endekslerinin bozulacağı düşünüldüğünde uç değerlerden arındırılmış 206 katılımcının AFA için yeterli olduğu düşünülebilir (Thompson, 2004, s. 129; MacCallum, Browne and Sugawara, 1996, s. 139). Dağılım grafiklerinin normal dağılıma yakın görüntüler vermesinden ve basıklık ve çarpıklık değerlerinin +2 ve -2 arasında olmasından dolayı normal dağılım şartının sağlandığı sonucuna ulaşılmıştır (Trochim and Donnelly, 2006, s. 48). Çoklu doğrusallık (multicollinearity) için 4 madde arasındaki korelasyon katsayıları hesaplanmıştır. Maddeler arasındaki korelasyon katsayılarının .258 ve .708 arasında değiştiği görülmüştür. Maddeler arası korelasyon katsayılarının .90'ın altında olmasından dolayı veri setinin çoklu doğrusallık varsayımını karşıladığı söylenebilir (Tabachnick and Fidell, 2012, s. 88).

DFA'da test edilen BÜT 3-5'in kuramsal modeli ve DFA modelin anlamlılığına ilişkin ki-kare, p değeri, yaklaşık hataların karekökü (*Root Mean Square Error of Approximation*-RMSEA) ve gözlenen değişkenlerin t -değerleri Şekil 4.1'de verilmiştir.



Şekil 4.1. BÜT 3-5 DFA sonuçları

Şekil 4.1’de görüldüğü üzere analize dâhil edilen dört madde için üretilen t-değerlerinin tamamı 2.56’nın üzerinde olup .01 düzeyinde manidardır. DFA sonuçlarına göre beklenen ve gözlenen kovaryans matrisleri arasında bir farklılık olmadığı görülmüştür [$\chi^2(1)= 0.77$; $p=0.380$]. Gözlenen değişkenlerin hata varyansları kontrol edildiğinde hipotez maddeleri için hata varyanslarının .63 ve .53 olduğu deney maddeleri için hata varyanslarının ise .27 ve .38 olduğu görülmüştür. Bu değerlere göre maddelerin hata varyanslarının düşük olduğu düşünülebilir (Çokluk, Şekercioğlu ve Büyüköztürk, s. 305). DFA sonucunda hesaplanan uyum ölçütleri ve bu ölçütler için kabul kriterleri Tablo 4.7’de verilmiştir.

Tablo 4.7. BÜT 3-5 için hesaplanan DFA Uyum Ölçütleri ve Kabul Kriterleri

İndeksi	Değer	Kesme Noktası/ Kriter	Kaynak
χ^2	0.77; p=.380	p>0.05	-
χ^2/sd	0.77	≤ 2 ; mükemmel uyum	Tabahchnick ve Fidell (2012, s. 720)
RMSEA	0.000	≤ 0.05 ; mükemmel uyum	Kline (2010, s. 206)
NFI	0.997	≥ 0.95 ; mükemmel uyum	Thompson (2004, s. 130)
RMR	0.037	≤ 0.05 ; mükemmel uyum	Tabahchnick ve Fidell (2012, s. 725)
SRMR	0.009	≤ 0.05 ; mükemmel uyum	Tabahchnick ve Fidell (2012, s. 725)
GFI	0.998	≥ 0.95 ; mükemmel uyum	Sümer (2000)
AGFI	0.981	≥ 0.95 ; mükemmel uyum	Sümer (2000)
CFI	1.000	≥ 0.95 ; mükemmel uyum	Thompson (2004, s. 130)
PGFI	0.099	≥ 0.95 ; mükemmel uyum	Tabahchnick ve Fidell (2012, s. 724)

Tablo 3.7’de görüldüğü gibi DFA sonucunda elde edilen uyum indeksleri, mükemmel uyum kriterleri karşılamaktadır. DFA bulgularına göre hipotez maddelerinin standardize edilmiş yük değerleri sırası ile .61 ve .68, deney maddelerinin standardize edilmiş yük değerleri ise .79 ve .86 olarak bulunmuştur. DFA çıktılarında ise herhangi bir modifikasyon önerine yer verilmemiştir. Sonuç olarak BÜT 3-5’in kuramsal yapısı DFA bulgularında doğrulanmıştır. Bu bağlamda BÜT 3-5’in yapı geçerliğinin iyi düzeyde olduğu söylenebilir.

4.1.2. BÜT 3-5’in ayırt edicilik geçerliğine ilişkin bulguları

PCA bulguları ile elde edilen BÜT 3-5’in DFA ile yapı geçerliği doğrulandıktan sonra testin madde bazında ayırt edicilik analizleri yapılmıştır. Bunun için öncelikle katılımcılar toplam bileşik yaratıcılık puanlarına göre alt ve üst %27’lik gruplara ayrılmış, daha sonra bu grupların ortalama puanları karşılaştırılmıştır. Ayırt edicilik geçerliğinin ikinci aşamasında ise farklı sınıf seviyesindeki öğrencilerin ortalama puanları karşılaştırılmıştır. Ayırt edicilik analizlerinin tamamı öğrencilerin akıcılık, esneklik ve bileşik yaratıcılık puanları üzerinden yapılmıştır. Alt ve üst grupların ayırt edicilik analizlerinden önce maddelerin betimsel istatistikleri çıkartılmıştır. Maddelerin betimsel istatistikleri Tablo 4.8’de verilmiştir.

Tablo 4.8. *BÜT 3-5 maddelerinin betimsel istatistikleri*

Madde	Puan	\bar{X}	SS	Minimum	Maksimum
1. Hipotez	Akıcılık	1,742	1,549	,00	7,00
	Esneklik	1,280	,941	,00	4,00
	B. Yaratıcılık	1,519	1,206	,00	5,94
2. Hipotez	Akıcılık	3,424	1,890	,00	10,00
	Esneklik	2,167	,9621	,00	5,00
	B. Yaratıcılık	2,838	1,388	,00	7,32
3. Deney	Akıcılık	2,019	2,635	,00	12,00
	Esneklik	1,366	1,548	,00	5,00
	B. Yaratıcılık	1,711	2,082	,00	8,16
4. Deney	Akıcılık	1,892	2,433	,00	13,00
	Esneklik	1,196	1,350	,00	4,00
	B. Yaratıcılık	1,559	1,867	,00	8,22
Test Toplam	Akıcılık	9,076	6,522	,00	33,00
	Esneklik	6,009	3,565	,00	14,00
	B. Yaratıcılık	7,627	5,001	,00	22,28

Veri seti toplam akıcılık puanına göre alt ve üst %27'lik gruplara bölüldüğünde her grupta 175 katılımcının yer aldığı görülmüştür. Daha sonra grupların her madde için hesaplanan akıcılık, esneklik ve bileşik yaratıcılık puanlarının ortalamaları bağımsız örneklem t-testi ile karşılaştırılmıştır. 12 karşılaştırma yapıldığından (4 madde X 3 farklı puan türü) Bonferoni düzeltmesi yapılarak anlamlılık değeri .004 olarak belirlenmiştir. Alt ve üst grupların madde puanlarına ilişkin bağımsız örneklem t-testi sonuçları Tablo 4.9'da verilmiştir.

Tablo 4.9. Madde ayırt ediciliğine ilişkin bağımsız örneklem t-testi sonuçları

Madde- Puan	Grup	\bar{X}	SS	T	Sd	p	η^2	
Madde 1 (Hipotez)	Akıcılık	Alt %27	0,491	0,633	-17,918	348	.001	0,480
		Üst %27	2,811	1,591				
	Esneklik	Alt %27	0,463	0,574	-18,732	348	.001	0,502
		Üst %27	1,869	0,809				
	B. Yaratıcılık	Alt %27	0,479	0,603	-19,347	348	.001	0,518
		Üst %27	2,350	1,127				
Madde 2 (Hipotez)	Akıcılık	Alt %27	1,806	1,153	-20,571	348	.001	0,549
		Üst %27	4,983	1,686				
	Esneklik	Alt %27	1,377	0,841	-17,539	348	.001	0,469
		Üst %27	2,817	0,687				
	B. Yaratıcılık	Alt %27	1,621	0,991	-20,700	348	.001	0,552
		Üst %27	3,936	1,098				
Madde 3 (Deney)	Akıcılık	Alt %27	0,074	0,322	-26,375	348	.001	0,667
		Üst %27	5,200	2,551				
	Esneklik	Alt %27	0,069	0,275	-36,334	348	.001	0,791
		Üst %27	3,223	1,119				
	B. Yaratıcılık	Alt %27	0,072	0,301	-30,879	348	.001	0,733
		Üst %27	4,258	1,768				
Madde 4 (Deney)	Akıcılık	Alt %27	0,051	0,222	-29,054	348	.001	0,708
		Üst %27	4,880	2,187				
	Esneklik	Alt %27	0,514	0,222	-39,193	348	.001	0,815
		Üst %27	2,794	0,899				
	B. Yaratıcılık	Alt %27	0,514	0,222	-35,114	348	.001	0,780
		Üst %27	3,868	1,421				
Test Toplam	Toplam Akıcılık	Alt %27	2,423	1,181	-43,297	348	.001	0,843
		Üst %27	17,874	4,571				
	Toplam Esneklik	Alt %27	1,960	0,899	-63,88	348	.001	0,921
		Üst %27	10,703	1,588				
	Toplam Yaratıcılık	Alt %27	2,223	1,031	-53,349	348	.001	0,891
		Üst %27	14,413	2,841				

Tablo 4.9’da alt ve üst yetenek grupları için tüm maddelerin ve bu maddeler için hesaplanan akıcılık, esnek ve bileşik yaratıcılık puanlarının ortalamaları arasında anlamlı farklılıklar verilmiştir. Toplam puanlar üzerinden yapılan bağımsız örneklem t-testi bulgularına göre de üst gurubun toplam puan ortalamaları, alt gurubun toplam puan ortalamalarından istatistiksel olarak anlamlı bir şekilde daha yüksektir. Hesaplanan etakare etki büyüklükleri ise .469 ile .921 arasında değişmektedir. Bu değerlerin tamamının büyük olmasından dolayı, BÜT 3-5 puanlarının varyansının büyük bir kısmını grup değişkeninden kaynaklandığı düşünülebilir (Büyüköztürk, 2011, s.44).

Sınıf düzeyleri arasındaki ayırt ediciliğini incelemek için tek yönlü varyans analizi (Oneway ANOVA) yapılmıştır. Analiz bileşen puanları ve testin toplam puanları

üzerinden yapılmıştır. Grupların analize dâhil edilen puanlarının ortalama, standart sapma, minimum ve maksimum değerlerine ilişkin betimsel bulgular Tablo 4.10'da verilmiştir.

Tablo 4.10. *BÜT 3-5 puanlarının sınıf düzeylerine göre betimsel istatistikleri*

		Sınıf	\bar{X}	SS	Minimum	Maksimum
Test Toplam	Akıcılık	3	5,91	4,51	0	20,00
		4	8,09	5,14	0	26,00
		5	14,81	6,80	0	33,00
	Esneklik	3	4,37	3,06	0	13,00
		4	5,56	3,01	0	14,00
		5	8,88	3,13	0	14,00
	B. Yaratıcılık	3	5,22	3,81	0	16,48
		4	6,92	4,05	0	19,58
		5	11,94	4,83	0	22,28
Hipotez Toplam	Akıcılık	3	3,54	2,13	0	10,00
		4	5,01	2,57	0	14,00
		5	7,66	2,54	0	15,00
	Esneklik	3	2,67	1,46	0	7,00
		4	3,46	1,43	0	7,00
		5	4,52	1,10	0	7,00
	B. Yaratıcılık	3	3,15	1,79	0	8,74
		4	4,30	1,96	0	10,84
		5	6,14	1,69	0	9,90
Deney Toplam	Akıcılık	3	2,37	3,42	0	13,00
		4	3,08	3,82	0	18,00
		5	7,14	5,59	0	22,00
	Esneklik	3	1,70	2,26	0	9,00
		4	2,10	2,257	0	9,00
		5	4,36	2,764	0	9,00
	B. Yaratıcılık	3	2,06	2,871	0	11,16
		4	2,62	3,041	0	12,80
		5	5,79	4,095	0	14,38

Tablo 4.10'da sınıfların deney ve hipotez alt testleri için ve testin tamamı için hesaplanan akıcılık, esneklik ve bileşik yaratıcılık ortalama puanları verilmiştir. Tabloda görüldüğü gibi alt testlerde, test toplamında tüm puan türleri için 5. sınıf öğrencilerinin ortalama puanları 4. sınıfların ortalama puanlarından, 4. sınıf öğrencilerinin ortalama puanları da 3. sınıf öğrencilerinin ortalama puanlarından daha yüksektir. Sınıfların akıcılık, esneklik ve bileşik yaratıcılık puanları ortalamaları arasındaki farkın istatistiksel olarak anlamlı olup olmadığını belirlemek için tek yönlü ANOVA analizleri yapılmıştır.

Analize geçmeden önce Levene varyans eşleşliği testi yapılmış ve toplam esneklik puanları dışındaki diğer tüm puanlar için Levene test istatistiği anlamlı çıktığından varyansların homojen olmadığı sonucuna varılmıştır. Bu bakımdan grup karşılaştırmalarında farkın kaynağını belirlemek için toplam esneklik puanları için

Scheffe, diğer puanlar için de Games-Howell izleme testi sonuçları incelenmiştir (Field, 2009, s. 374, s. 53; Akbulut, 2010, s. 123). Grupların test puanlarına ilişkin ANOVA bulguları Tablo 4.11’de verilmiştir.

Tablo 4.11. BÜT 3-5 puanlarının sınıf düzeylerine göre ANOVA sonuçları

Görev	Puan	Sınıf	Varyans Kaynağı	KT	sd	KO	F	p	Fark	η^2
Test Toplam	Akıcılık	3	G. Arası	8497,61	2	4248,80	144,11	.001	3-4,	0,309
		4	G. İçi	18987,67	644	29,48			3-5,	
		5	Toplam	27485,28	646				4-5	
	Esneklik	3	G. Arası	2167,19	2	1083,59	115,41	.001	3-4,	0,264
		4	G. İçi	6046,74	644	9,38			3-5,	
		5	Toplam	8213,94	646				4-5	
	B. Yaratıcılık	3	G. Arası	4834,42	2	2417,21	137,43	.001	3-4,	0,299
		4	G. İçi	11326,83	644	17,58			3-5,	
		5	Toplam	16161,25	646				4-5	
Hipotez Toplam	Akıcılık	3	G. Arası	1762,90	2	881,45	152,33	.001	3-4,	0,321
		4	G. İçi	3726,40	644	5,78			3-5,	
		5	Toplam	5489,30	646				4-5	
	Esneklik	3	G. Arası	354,01	2	177,01	94,68	.001	3-4,	0,227
		4	G. İçi	1203,89	644	1,86			3-5,	
		5	Toplam	1557,91	646				4-5	
	B. Yaratıcılık	3	G. Arası	924,63	2	462,31	138,02	.001	3-4,	0,300
		4	G. İçi	2157,09	644	3,35			3-5,	
		5	Toplam	3081,72	646				4-5	
Deney Toplam	Akıcılık	3	G. Arası	2.584,28	2	1292,14	71,74	.001	3-5,	0,182
		4	G. İçi	11598,51	644	18,01			4-5	
		5	Toplam	14182,80	646					
	Esneklik	3	G. Arası	802,50	2	401,25	69,15	.001	3-5,	0,177
		4	G. İçi	3736,711	644	5,80			4-5	
		5	Toplam	4539,215	646					
	B. Yaratıcılık	3	G. Arası	1579,153	2	789,57	72,43	.001	3-5,	0,184
		4	G. İçi	7020,885	644	10,90			4-5	
		5	Toplam	8600,039	646					

Tablo 4.11’de görüldüğü gibi test toplam puanları ve alt test puanları bağlamında sınıfların akıcılık, esneklik ve bileşik yaratıcılık puan ortalamaları arasında anlamlı farklılıklar bulunmuştur. Farkın kaynağını bulmak için izleme testlerine bakılmış ve iki karşılaştırmalarda elde edilen anlamlı farklılıklar Tablo 4.11’de *fark* sütununa işlenmiştir. Buna göre testin toplam puanları ve hipotez alt test puanları için 5. sınıfların ortalamalarının 3 ve 4. sınıfların ortalamalarından, 4. sınıfların ortalamalarının ise 3. sınıfların ortalamalarından anlamlı şekilde daha yüksek olduğu görülmüştür. Deney alt testi için 5. sınıfların ortalamalarının 3 ve 4. sınıfların ortalamalarından anlamlı şekilde daha yüksek olduğu, ancak 4. sınıfların ortalamalarının 3. sınıfların ortalamalarından yüksek olmasına rağmen aradaki farkın istatistiksel olarak anlamlı olmadığı görülmüştür.

Yapılan ayırt edicilik çalışmalarına göre BÜT 3-5'in alt testleri ve toplam test puanları için alt ve üst puan grupları arasında anlamlı farklılıklar çıktığı görülmüştür. Sınıf düzeyine göre yapılan ayırt edicilik analizleri ise deney maddelerinin 3 ve 4. sınıf öğrencilerini ayırt etmediğini göstermiştir. Ancak hipotez alt testi ve testin toplam puanı üzerinden yapılan analizlerde 5. sınıfların 4. sınıflardan, 4. sınıfların da 3. sınıflardan istatistiksel olarak anlamlı şekilde daha yüksek puan aldıkları görülmüştür. Ayırt edicilik analizlerinde elde edilen büyük etki büyüklükleri sonuçların kuramsal olarak hem de pratikte anlamlı olduğu şeklinde yorumlanabilir (Akbulut, 2010, s. 114). Bu bağlamda BÜT 3-5'in ayırt edicilik geçerliğinin yeterli düzeyde olduğu düşünülebilir.

4.1.3. BÜT 3-5'in uyum geçerliğine ilişkin bulgular

Psikolojik testlerde testin kendi içerisindeki geçerliğinin yanında dış değişkenler ile de ilişkisinin değerlendirilmesi gerekmektedir. Uyum geçerliği olarak tanımlanan bu işlemde ölçmenin merkezinde yer alan özellikle ilişkili olduğu varsayılan değişkenler kullanılır. BÜT 3-5'in dış geçerliğini belirlemek amacıyla fen bilimlerine karşı tutum, fen bilgisinde akademik benlik algısı, fen bilgisi ve matematik karne notları dış ölçütler olarak belirlenmiş ve test puanlarının bu ölçütlerle ilişkisi Pearson momentler çarpımı korelasyon katsayılarına bakılarak değerlendirilmiştir.

4.1.3.1. BÜT 3-5'in ders notlarına göre uyum geçerliği

Ders notları ile uyum geçerliği analizleri 4. ve 5. sınıf öğrencilerinin karne notları ve BÜT 3-5'teki performansları arasındaki ilişki incelenerek araştırılmıştır. Öğrencilerin karne notları aynı kalırken, BÜT 3-5 puanları 5. sınıflar lehine daha yüksek olduğundan analizler 4 ve 5. sınıflar için ayrı ayrı yapılmıştır. BÜT 3-5'in karne notları ile uyum geçerliği bulguları Tablo 4.12'de verilmiştir.

Tablo 4.12. BÜT 3-5'in karne notlarına göre uyum geçerliği bulguları

Sınıf	Puan	Fen Notu	Matematik Notu	Fisher r-z İstatistikleri
4	Akıcılık	.404**	.409**	Z=.06; p=.476
	Esneklik	.393**	.397**	Z=.05; p=.480
	B. Yaratıcılık	.405**	.404**	Z=.01; p=.496
	Matematik Notu	.808**		
5	Akıcılık	.318**	.332**	Z=.14; p=.444
	Esneklik	.310**	.333**	Z=.23; p=.409
	B. Yaratıcılık	.317**	.345**	Z=.28; p=.389
	Matematik Notu	.582**		

** p<.001

Tablo 4.12’de görüldüğü gibi 4. ve 5. sınıf öğrencilerinin fen ve matematik karne notları ile BÜT 3-5 toplam akıcılık, esneklik ve yaratıcılık puanları arasında .001 seviyesinde anlamlı pozitif korelasyon katsayıları bulunmuştur. BÜT 3-5 ve ders notları için elde edilen korelasyon katsayılarının tamamı orta düzeydedir (Cohen, 1977, s. 82). Tablodan görüldüğü üzere her iki sınıf düzeyi için de BÜT 3-5 puanları ve matematik ders notları için hesaplanan korelasyon katsayılarının fen ders notları için üretilen katsayılardan daha büyük olduğu görülmektedir. Korelasyon katsayıları arasındaki farkın anlamlı olup olmadığını belirlemek için Fisher r-z (*Fisher r to z transformation*) dönüşümü yapılmıştır (Meng, Rosenthal and Rubin, 1992, s. 172). Fisher r-z istatistiklerine göre matematik ve fen notlarının BÜT 3-5 ile ilişkileri arasındaki fark istatistiksel olarak anlamlı değildir.

Tablo 4.12’de verilmiş olan korelasyon katsayıları 4. sınıflar için daha yüksektir. Aradaki en yüksek fark ise bileşik yaratıcılık puanı ve fen bilgisi ders notlarında görülmektedir ($r_{4 \text{ fen-By}}=.405$ ve $r_{5 \text{ fen-By}}=.317$). Bu iki korelasyon katsayısı için yapılan Fisher r-z dönüşümüne göre bu iki değer arasındaki fark istatistiksel olarak anlamlı değildir ($z_{\text{yaratıcılık-fen}}= 0,98$; $p=.163$). Sınıfların fen ve matematik notları arasındaki korelasyon katsayılarına bakıldığında 4. sınıf öğrencileri lehine anlamlı bir farklılık görülmüştür ($z=4.39$; $p<.001$).

Anlamlı korelasyon değerleri alana özgü bir test olan BÜT 3-5’in uyum geçerliği kanıtları olarak değerlendirilebilir.

4.1.3.2. BÜT 3-5’in fen tutumuna ve fen bilgisi akademik benlik algısı ile olan ilişkisi

BÜT 3-5’in uyum geçerliği için test puanları ile BYTÖ ve FABAÖ puanları arasındaki ilişkiye bakılmıştır. Öğrencilerin BYTÖ ve FABAÖ puanları betimsel istatistikler bağlamında incelendiğinde 3. ve 4. sınıf öğrencilerinin her iki ölçek puanları için de değişkenliklerinin düşük olduğu görülmüştür. Her iki ölçek için de 3 ve 4. sınıf öğrenci puanlarının %85’ten fazlasının 20 ve üzerinde, yarıya yakını ise 25 puanının üzerindedir. Grupların betimsel istatistiklerine bakıldığında hem 3 hem de 4. sınıfların BYTÖ için modu 26, medyanı 26’dır. Ortalamalar ise sırasıyla 24,57 (SS=3,94) ve 23,93 (SS=4,04) ve varyansları ise 1,53 ve 16,36’tir. Ayrıca bu iki sınıf için hem BYTÖ hem de FABAÖ ölçümlerinin normal dağılımdan uzak olduğu gözlenmiştir. Normal dağılım

göstermeyen ve uç değerleri olan veri setlerinde ve değişkenliğin az olduğu ölçümlerde ilişki katsayılarının yanlış hesaplanacağı ve dolayısı ile sonuçların yanlış yorumlanacağı bilinmektedir (Goodwin and Leech, 2006, s. 252; Wilcox, 2016, s. 215). Bu bakımdan BÜT 3-5'in BYTÖ ve FABAÖ ile uyum geçerliği analizleri sadece 5. sınıf öğrencilerinden elde edilen verilerle yapılmıştır. Analize ilişkin bulgular Tablo 4.13'te verilmiştir.

Tablo 4.13. *BÜT 3-5'in BYTÖ ve FABAÖ puanlarına göre uyum geçerliği bulguları*

Puan	BYTÖ	FABAÖ
Akıcılık	.238**	.295**
Esneklik	.293**	.286**
B. Yaratıcılık	.259**	.297**
FABAÖ	.586**	

**p<.001

Tablo 4.13'e göre öğrencilerin fen bilimlerine yönelik tutumlar ve fen bilimleri dersi akademik benlik algısı ile BÜT 3-5 toplam akıcılık, esneklik ve yaratıcılık puanları arasında .001 seviyesinde anlamlı pozitif korelasyon katsayıları bulunmuştur. Hesaplanan korelasyon katsayıları düşük olmakla beraber, anlamlı korelasyon değerleri alana özgü bir test olan BÜT 3-5'in uyum geçerliği kanıtları olarak değerlendirilebilir.

4.2. BÜT 3-5'in Güvenirliği

BÜT 3-5'in güvenirliliğinin incelenmesi için iç tutarlık güvenirliliği ve okuyucular arası güvenirlilik analizleri yapılmıştır.

4.2.1. BÜT 3-5'in iç tutarlılığına ilişkin bulgular

BÜT 3-5 iki alt testten ve her alt testte 2 madde olmak üzere toplam dört maddeden oluşmaktadır. Testten toplam akıcılık, toplam esneklik ve toplam bileşik yaratıcılık puanları elde edilmektedir. Aynı puanlar hipotez (hipotez akıcılık, hipotez esneklik, hipotez bileşik yaratıcılık) ve deney alt testi (deney akıcılık, deney esneklik, deney bileşik yaratıcılık) içinde elde edilmektedir. Bu bakımdan testin güvenirliliği için hesaplanan Cronbach Alpha iç tutarlılık katsayısı akıcılık, esneklik ve bileşik yaratıcılık puanları için ayrı ayrı hesaplanmıştır. Ek olarak hipotez ve deney bileşenleri için akıcılık ve bileşik yaratıcılık puanları kullanılarak Cronbach Alpha iç tutarlık katsayıları hesaplanmış ve Tablo 4.14'te verilmiştir.

Tablo 4.14. BÜT 3-5 maddeler arası ve madde-toplam korelasyon bulguları

Ölçek	Puan Türü	Madde	Düzeltilmiş Madde Toplam Korelasyonu	Madde Silindiğinde Cronbach Alpha	Ölçek Cronbach Alpha
BÜT 3-5	Akıcılık	1	.387	.758	.743
		2	.505	.703	
		3	.645	.620	
		4	.660	.606	
	Esneklik	1	.378	.699	.701
		2	.386	.694	
		3	.611	.556	
		4	.629	.536	
	B. Yaratıcılık	1	.397	.744	.736
		2	.472	.709	
		3	.641	.611	
		4	.662	.589	
Hipotez	Akıcılık	1	.431	.594	
		2	.431		
	Esneklik	1	.277	.433	
		2	.277		
	B. Yaratıcılık	1	.415	.583	
		2	.415		
Deney	Akıcılık	1	.708	.828	
		2	.708		
	Esneklik	1	.686	.813	
		2	.686		
	B. Yaratıcılık	1	.706	.825	
		2	.706		

Tablo 4.14’te görüldüğü gibi BÜT 3-5 akıcılık, esneklik ve bileşik yaratıcılık puanları için hesaplanan Cronbach Alpha iç tutarlık katsayılarının tamamı kabul edilebilir değer olan .70’in üzerindedir (Nunnally, 1978’den aktaran Hinkin, 1995, s. 979; Cortina, 1993, s. 103, Field, 2009, s. 675). Alt ölçekler için üretilen Cronbach Alpha katsayıları hipotez alt ölçeği için .433 ve .594, deney alt ölçeği için de .813 ve .825 olarak bulunmuştur. Hipotez alt testi esneklik puanları dışında düzeltilmiş madde-toplam korelasyonlarının tamamının .30’un üzerinde olması maddelerin aynı şeyi ölçtüğü şeklinde yorumlanabilir (Field, 2009, s. 678). Madde silindiğinde elde edilecek Cronbach Alpha katsayılarına bakıldığında 1. madde silindiğinde akıcılık puanları için Cronbach Alpha katsayısının .743’ten .758’e, yaratıcılık puanları için .736’dan .744’e çıktığı

görülmektedir. Cronbach Alpha değerindeki bu değişimin anlamlı olup olmadığını test etmek için akıcılık puanları için (fark daha yüksek olduğu için yaratıcılık için gerek duyulmamıştır) hesaplanan katsayılar üzerinden Feldt Testi (Feldt, Woodroof and Salih, 1987) yapılmıştır. Feldt test sonucuna göre 1. maddenin testten çıkartılması durumunda Cronbach Alpha katsayısında anlamlı bir değişim olmayacağı görülmüştür ($W=1,062$; $p=.778$). Hipotez alt testinin iç tutarlık düzeyi düşük görünmekle birlikte, hipotez maddeleri için hesaplanan düzeltilmiş madde-toplam korelasyon değerleri yeterli düzeydedir. Bu bağlamda BÜT 3-5'in 4 madde için iç tutarlığının yeterli düzeyde olduğu düşünülebilir.

Cronbach Alpha katsayılarının yorumlanmasında Tablo 4. 14'de verilen değerlerin 4 madde ve alt ölçeklerde 2 madde için hesaplanan değerler olduğu göz önünde bulundurulmalıdır. Çok bileşenli ölçeklerde güvenilirlik incelemesinde Cronbach Alpha katsayısının yerine korelasyon bulgularının verilmesinin daha uygun olacağı yönünde görüşler vardır (örneğin Cortina, 1993, s. 102). Çünkü Cronbach Alpha katsayısının ölçekteki madde sayısı ve maddeler arası korelasyon ile doğru orantılı, bileşen sayısı ile ters orantılı olarak değiştiği bilinmektedir (Kline, 2005, s. 175; Urbina, 2004, s. 131). Her ne kadar BÜT 3-5 için üretilen Cronbach Alpha değerleri kabul edilebilir seviyelerde olsa da 2 faktör ve 4 maddeden oluşan BÜT 3-5'in iç tutarlığının incelenmesinde korelasyon analizi bulgularına da yer vermek ek kanıt sunması bakımından yararlı olacaktır. Bu bakımdan BÜT 3-5 için korelasyon katsayıları hesaplanmış ve Tablo 4.15'te verilmiştir.

Tablo 4.15. BÜT 3-5 alt testler arası korelasyon katsayıları

Puan Türü	Maddeler					Alt Ölçekler	
	H2	D3	D4	Düzeltilmiş Madde-Toplam Korelasyonu	Maddeler Arası Ortalama Korelasyon	Hipotez	Deney
Akıcılık	H1	,431*	,258*	,302*	,387*	0,417*	,302*
	H2		,414*	,390*	,505*		,435*
	D3			,708*	,645*		,405*
	D4				,660*		,413*
Esneklik	H1	,331*	,278*	,306*	,374*	0,365*	,318*
	H2		,311*	,293*	,386*		,331*
	D3			,671*	,611*		,361*
	D4				,629*		,367*
B. Yaratıcılık	H1	,415*	,278*	,317*	,397*	0,411*	,321*
	H2		,382*	,367*	,472*		,406*
	D3			,706*	,641*		,396*
	D4				,662*		,408*

*p<.001

Tablo 4.15'te görüldüğü gibi akıcılık puanları için maddeler arası korelasyon katsayıları .258 ve .708 arasında değişmektedir. Yaratıcılık puanları için maddeler arası korelasyon katsayıları ise .278 ve .706 arasında değişmektedir. Akıcılık puanlarının madde çıkartıldıktan sonra test toplamı ile korelasyon katsayıları .387 ve .660 arasında değişirken bu değerler yaratıcılık puanları için .397 ve .662 arasında değişmektedir. Akıcılık, esneklik ve bileşen yaratıcılık puanları için maddeler arası korelasyon katsayıları ortalaması sırası ile .417, .365 ve .411 olarak bulunmuştur. Güvenirlik incelemesinde maddeler arası korelasyon katsayılarının tamamının .15- .50 arasında olmasını ve ortalamalarının .40 civarında olması gerektiğini belirtmektedirler (Clark and Watson, 1995, s. 316; Briggs and Check, 1986, s. 115). Korelasyon katsayıları bağlamında BÜT 3-5'in iç tutarlık güvenilirliğinin yeterli seviyede olduğu düşünülebilir. Ayrıca iç tutarlık analizleri için hesaplanan korelasyon katsayıları, tablo 4.14'te verilen Cronbach Alpha iç tutarlık bulguları ile örtüşmektedir. Tablo 4.14 ve tablo 4.15'te sunulan bulgulardan hareketle BÜT 3-5'in ve alt testlerinin iç tutarlık güvenilirliğinin yeterli düzeyde olduğu söylenebilir.

4.2.2. BÜT 3-5'in okuyucular arası güvenirlğine ilişkin bulgular

BÜT 3-5, açık uçlu sorulardan oluşan çoğul üretime dayalı bir testtir. Öğrencilerin yanıtları okuyucu tarafından değerlendirilerek akıcılık, esneklik ve bileşik yaratıcılık puanları elde edilmektedir. Bu süreçte okuyucu öncelikle yanıtın doğru olup olmadığına karar vermekte, ikinci olarak da yanıtları kavramsal kategorilere göre sınıflandırmakta ve son olarak bir bağıntı kullanarak yanıtlardan bileşik yaratıcılık puanı elde etmektedir. Bu bakımdan BÜT 3-5'in puanlanmasında hatanın olabileceği düşünülerek puanlayıcılar arası güvenirlilik analizi yapılması uygun bulunmuştur.

BÜT 3-5'in okuyucular arası güvenirlilik analizleri için rastgele 100 katılımcının (27 üçüncü sınıf, 39 dördüncü sınıf, 34 beşinci sınıf) kâğıdı seçilmiş ve ikinci bir okuyucuya verilmiştir. Her iki okuyucu da madde geliştirme sürecinde ve pilot uygulamalarda geliştirilen doğru yanıt havuzunu kullanarak kâğıtları okumuştur. Okuyucular arası güvenirlilik analizleri için iki okuyucunun sınıf içi korelasyon (intraclass correlation) analizi yapılmıştır. Sınıf içi korelasyon analiz bulguları Tablo 4.16'da verilmiştir.

Tablo 4.16. *BÜT 3-5'in okuyucular arası güvenirlliği*

Madde	Akıcılık			Esneklik			B. Yaratıcılık		
	r_{ICC}	$F_{(99,99)}$	P	r_{ICC}	$F_{(99,99)}$	P	r_{ICC}	$F_{(99,99)}$	p
1	.975	79,96	.001	.917	22.97	.001	.965	56.05	.001
2	.965	56.32	.001	.942	33.51	.001	.966	57.16	.001
3	.994	314.90	.001	.989	179.73	.001	.994	329.88	.001
4	.998	1317.26	.001	.992	259.86	.001	.995	648.08	.001
BÜT 3-5 Toplam	.994	360.78	.001	.989	176.18	.001	.994	338.65	.001

Tablo 4.16'da görüldüğü gibi tüm maddeler ve puan türleri için yüksek derecede okuyucular arası güvenirlilik katsayıları bulunmuştur. Ayrıca deney maddeleri için hesaplanan okuyucular arası güvenirlilik katsayıları, hipotez maddeleri için hesaplanan okuyucular arası güvenirlilik katsayısından biraz yüksektir. Tabloda verilen bütün katsayıların ortalaması ise .974'tür. Bu bağlamda BÜT 3-5'in okuyucular arası güvenirliliğinin yüksek olduğu söylenebilir.

5. SONUÇ, TARTIŞMA ve ÖNERİLER

Bu araştırmanın amacı SDDS temel olarak 3, 4 ve 5. sınıf öğrencilerine yönelik bilimsel yaratıcılık testi geliştirmek ve bu testin geçerlik ve güvenilirlik incelemesini yapmaktır. Bunun için öğretmenlerden oluşan bir uzman ekip ile bir madde havuzu oluşturulmuştur. Uzman görüşünden yararlanılarak bu havuzdan seçilen maddeler ile bir taslak form oluşturulmuş, ön deneme uygulaması yapılmıştır. Ön deneme uygulamasının ardından çeşitli revizyonlarla 10 maddelik bir pilot form oluşturularak 120 öğrenci ile yapılan uygulamanın ardından çeşitli revizyonların ardından 7 maddelik uygulama formu oluşturulmuştur. Uygulama formu 647 öğrenciye uygulanmış ve temel bileşenler analizi ile kuramsal yapı ve yorumlanabilirlik merkeze alınarak 4 madde seçilerek Bilimsel Üretkenlik Testi Form 3-5 oluşturulmuştur. Elde edilen BÜT 3-5'in geçerlik analizlerinde doğrulayıcı faktör analizi, ayırt edicilik ve uyum geçerliği bulguları incelenmiştir. BÜT 3-5'in güvenilirliğinin belirlenmesi için tutarlık analizleri ve okuyucular arası güvenilirlik analizleri yapılmıştır. Araştırmanın sonuçları ve bulguların tartışılmasının ardından öneriler verilmiştir.

5.1. Tartışma ve Sonuç

5.1.1. BÜT 3-5'in geçerliğine ilişkin tartışma ve sonuç

Bu bölümde BÜT 3-5'in geçerliğine ilişkin bulgular, yapı geçerliği, ayırt edicilik geçerliği ve uyum geçerliği alt başlıkları altında tartışılmıştır.

5.1.1.1. *BÜT 3-5'in yapı geçerliğine ilişkin tartışma ve sonuç*

BÜT 3-5 geliştirme sürecine kuramsal alt yapının oluşturulması ile başlanmıştır. Klahr'ın (2000) geliştirmiş olduğu SDDS modeli temel alınarak BÜT 3-5 kuramsal yapısı geliştirilmiştir. SDDS orijinalinde 3 bileşenden oluşmaktadır: hipotez oluşturma, deney tasarlama ve kanıt değerlendirme. Bu araştırma için modelde yer alan kanıt değerlendirme bileşeni bir aracı bileşen olarak yorumlanmış ve bu nedenle modele ayrı bir bileşen olarak eklenmemiştir. Geliştirilen BÜT 3-5 kuramsal olarak iki temel bileşenden oluşmuştur.

Kuramsal yapının geçerliğini test etmek amacıyla PCA ve DFA yapılmıştır. PCA analizlerinde bir deneyin sonucuna ilişkin hipotezlerin üretildiği 2. Tip hipotez ve verilen hipoteze göre uygun olmayan deneylerin düzenlenmesini içeren 1. Tip deney görevlerinin kuramsal yapı ile daha uyumlu olduğu görülmüştür. Yapılan PCA bulgularına göre bir deneyin verilen sonuca göre yeniden düzenlenmesini içeren 2. Tip deney görevlerinin

hipotez oluřturma grevleri ile aynı bileřen altında toplandıđı grlmř ve bu maddelerin ıkartılmasına karar verilmiřtir.

2. tip deney grevlerinin hipotez grevleri ile aynı bileřen altında toplanmasının bařlıca nedeni grevlerin birbirine yakın olması olabilir. Hem 1. Tip deney grevinde hem de 2. Tip hipotez grevinde deney ve sonucu verilmektedir. Hipotez oluřturma maddelerinde sonucun nedenleri sorulurken, deney grevlerinde dzenekte sonuca uygun deđiřiklikler yapılması istenmektedir. Bu durumun ikinci nedeni ise madde glklerinin birbirlerine yakın olması olabilir (Solano-Flores, 1993, s. 29). Madde ortalamalarına bakıldıđında 1. Tip deney grevlerinin ortalamalarının 2. Tip deney ve 2. Tip hipotez grevlerinden nispeten daha dřk olduđu grlmřtir (bakınız. Tablo 3.1). Sonular SDDS modelini temel alarak ortaokul đrencileri iin benzer bir test geliřtiren Ayas ve Sak'ın (2014, s. 201) bulguları ile benzerlik tařımaktadır. Bu arařtırmanın bulgularına gre 2. Tip deney grevlerinin ortalaması ($\bar{X}=2,01$; $SS= 2,08$), 1. Tip deney ($\bar{X}=2,08$; $SS=1,80$) ve 2. Tip hipotez sorularının ortalamalarından ($\bar{X}=2,63$; $SS=2,57$) daha yksektir. Benzer řekilde Ayas (2010, s. 61) aynı testi kullandıđı arařtırmasında 2. Tip deney grevlerinin ortalaması ($\bar{X}=1,51$; $SS= 1,89$), 1. Tip deney ($\bar{X}=1,81$; $SS=1,71$) ve 2. Tip hipotez sorularının ortalamalarından ($\bar{X}=2,23$; $SS=2,65$) daha dřk olduđunu rapor etmiřtir. Maddelerin ortalamaları ve standart sapmalarındaki bu benzerlik, maddeler arası korelasyon katsayılarını etkilediđinden 1. Tip deney ve 2. Tip hipotez maddeleri aynı bileřen altında toplanmıř olabilir. Tablo 3.2'de 1. Tip deney maddelerinin hipotez maddeleri ile ortalama korelasyon katsayısı .37, 2. Tip deney maddeleri ile ortalama korelasyon katsayıları ise .29'dur. Benzer řekilde Ayas ve Sak (2014, s. 201) bu iki grev arasında daha yksek korelasyon deđerleri bulmuřlardır.

Madde seiminden sonra yapılan PCA bulguları BT 3-5'in iki bileřenli bir yapıda olduđunu desteklemektedir. Hipotez maddeleri hipotez bileřeninin altında, deney maddeleri ise deney bileřeninin altında toplanarak toplam %78,9 varyans aıklamıřtır. Aıklanan varyansın yksek olmasının nedeni pilot uygulamada ve madde seim ařamasında iki ayrı aımlayıcı faktr analizi yapılması ve her ařamada faktr yk deđerlerinin madde seiminde kriter olarak kullanılması olabilir (okluk, řekerciođlu ve Bykztrk, 2010, s. 197). Bileřenlerin faktr yk deđerleri ise .789 ve .901 arasında deđerismektedir. Maddelerin ortak faktr varyansları ise .695 ve .850 arasında deđerismektedir. Bileřenler iin hesaplanan varyans oranları ise hipotez bileřenini iin

%36,35, deney bileşeni içinse %42,55'tir. Bileşenlerin varyans oranları arasında büyük farklılıkların olmaması da iki faktörlü yapıya işaret etmektedir (Çokluk, Şekercioğlu, Büyüköztürk, 2010, s. 221). PCA bulgularından hareketle BÜT 3-5'in kuramsal yapısının deneysel olarak desteklendiği söylenebilir.

BÜT 3-5'in yapı geçerliği için ikinci aşamada PCA ile elde edilen yapı DFA ile farklı bir örneklem üzerinden test edilmiştir. Modelin uyumu için kullanılan indekslerin tamamının mükemmel uyuma işaret ettiği görülmüştür (Tablo 3.7). BÜT 3-5'in yapı geçerliğine ilişkin bulgular SDDS modeli temel alınarak geliştirilen BÜT 6-8'in (Sak and Ayas, 2013, s. 324; Ayas and Sak, 2014, s. 202) yapı geçerliği bulgularından ve Bilimsel Yaratıcılık Yapı Modeli (SSCM) (Hu and Adey, 2002) temel alınarak geliştirilmiş testlerin yapı geçerliği bulgularından farklılaşmaktadır. Bu durumun nedeni örneklem, analizde kullanılan puanlar ve testin içeriği olabilir. Sak ve Ayas (2013, s. 324) yaptıkları çalışmada PCA sonucunda tek bir bileşenli yapı bulunmuştur. %34,45 varyans açıklayan bu yapıda maddelerin faktör yük değerleri .47 ve .63 arasında değişmektedir. Bu bileşen yapısı daha sonraki çalışmalarında (Ayas and Sak, 2014) DFA ile de doğrulanmıştır. BÜT 6-8'in İspanyolca versiyonunda da benzer bulgular rapor edilmiştir (Bermejo et. al., 2016, 659). Bu durumun başlıca nedeni kuramsal yapı olabilir. BÜT 6-8'in kuramsal yapısında kanıt değerlendirme bileşeni tek bir maddeyle temsil edilmektedir. İkinci neden ise BÜT 6-8'in ortaokul öğrencilerine yönelik bir test olması olabilir. Örneğin Ayas ve Sak (2014) ve Sak ve Ayas (201) verilerini bir özel yetenek programına başvuru yapan öğrencilerden toplamışlardır. Bu durum veri setindeki çeşitliliği düşürmüş olabilir. Diğer bir neden ise analize dahil edilen puanlar olabilir. BÜT 6-8'in yapı geçerliği çalışmalarında 5 maddenin akıcılık, esneklik ve bileşik yaratıcılık puanları analize birlikte dahil edilmiştir. Ancak esneklik ve bileşik yaratıcılık puanları akıcılık puanlarından üretilmekte ve bu indeksler arasında yüksek korelasyon katsayıları çeşitli araştırmalarda rapor edilmiştir. BÜT 3-5'in kuramsal yapısında kanıt değerlendirme bileşenine yer verilmemiştir. İkinci olarak veri seti rastsal olarak seçilen 9-12 yaş arası öğrencilerden toplanmıştır. Son olarak BÜT 3-5'in yapı geçerliği analizleri akıcılık puanları üzerinden yapılmıştır. Çünkü akıcılık puanları için maddelerin ranjı 7-13 arasında değişirken, esneklik puanlarının ranjı 4-5 arasında ve bileşik yaratıcılık puanlarının ranjı 5,58-7,58 arasında değişmektedir. Kullanılan indekslerin faktör çıkarma işlemlerini olumlu yönde etkilediği düşünülebilir. SSCM temelinde geliştirilen testlerde test geliştirme sürecine temel oluşturan kuramsal

yapı elde dilememiş ve yine tek faktörlü bir yapı elde dilmiştir (Hu and Adey, 2002, s. 398; Chin and Siew, 2015, s. 1399; Siew, Chong and Chin, 2014, 119). Hu ve Adey (2002) 160 öğrenci ile yaptıkları çalışmada katılımcıların yetenek seviyelerine göre sınıflandırıldığı (%38 üst, %42 orta ve %20 alt yetenek seviyesi) bir okuldan 12-15 yaş aralığındaki öğrencilerden veri toplamış ve analizlerde akıcılık, esneklik ve orijinallik puanlarını topladıkları bir indeks puanı üzerinden yapmışlardır. BÜT 3-5 gibi Kanlı (2014, s. 104) da Yaratıcı Bilimsel Çağrışımlar Testinde testin kuramsal yapısıyla uyumlu çok faktörlü (3 faktörlü) bir yapı rapor etmiştir. Kanlı (2014) çalışmasında 12-15 yaş öğrencilerinden veri toplamış ve yapı geçerliği analizlerinde sadece akıcılık puanlarını kullanmıştır. Bu bağlamda kullanılan indekslerin ve katılımcı grubun özelliklerinin yapı geçerliğine ilişkin sonuçları etkileyebileceği düşünülebilir.

Her ne kadar alan yazında bileşen başına en az 3 madde önerilse de (Akbulut, 2010, s. 103) Hinkin (1995, s. 972) yaptığı araştırmada ulaştığı ölçeklerin %10'dan fazlasında 2 maddeden oluşan ölçeklerin olduğunu belirtmiştir. Raubenheimer (2004, s. 60) ise tek bileşenli ölçeklerde en az dört madde olması gerektiğini ve çok bileşenli ölçekler faktörlerin belirgin şekilde oluşması ve maddelerin faktörler altında anlamlı bir şekilde toplanması durumunda 2 maddelik bileşenlerin kullanılabileceğini belirtmiştir. Faktör analizi bulgularında BÜT 3-5'in bileşen yapısı için yapılan yamaç birikinti testi, paralel analiz ve Kaiser kriteri birbirlerini desteklemekte ve iki bileşenli bir yapıya işaret etmektedir. Ayrıca bileşenlerin varyans oranları arasında da büyük farklılıklar yoktur. Açıklanan toplam varyans oranının yüksek olması bu yapının güçlü bir yapı olduğunun göstergesi olarak değerlendirilebilir. Maddelerin bileşenlerdeki faktör yüklerine bakıldığında herhangi bir binişik maddeye rastlanmamıştır. Maddelerin bileşenler altındaki faktör yük değerleri mükemmel düzeydedir (Comrey and Lee, 1992, s. 243). Bu bağlamda BÜT 3-5'in bileşenleri birer alt ölçek olarak değerlendirilebilir.

BÜT 3-5'in geliştirilme sürecinde SDDS modeline göre geliştirilen BÜT'ün (Sak and Ayas, 2013; Ayas and Sak, 2014) kuramsal alt yapısı kullanılmıştır. Ancak BÜT'ün kuramsal alt yapısı revize edilerek basitleştirilmiştir. Benzer test geliştirme çalışmaları Chin ve Siew (2015) ve Siew, Chong ve Chin (2014) tarafından yapılmıştır. Araştırmacılar Hu ve Adey (2002) tarafından geliştirilen SSCM'yi temel alarak ilkökul öğrencileri için bir test geliştirmiştir. Araştırmacılar SSCM'ni revize ederek bileşenleri daha basit hale getirmişlerdir. Testin geçerlik ve güvenilirlik analizlerinde ise Hu ve

Adey'in (2002) lise öğrencileri ile elde ettikleri sonuçlara yakın sonuçlar elde etmişlerdir. Bu araştırma BÜT 3-5'in geliştirilme süreci ile benzerlik taşısa da temelde yapı geçerliğine ilişkin bulgular bağlamında farklılaşmaktadır. Yukarıda belirtildiği gibi Chin ve Siew (2015) orijinal çalışmadaki yapıyı tekrardan bulmuşlardır. Siew, Chong ve Siew (2014, s. 119) ise 5. sınıf öğrencileri için geliştirdikleri testte yine tek faktörlü bir yapı bulmuşlardır. Bu iki çalışmada da tek faktörlü bir yapı elde edilmiştir. SSCM açılımlayıcı faktör analizi çalışmalarında ortaya koyulmamış ve ayrıca elde edilen tek faktörlü yapı DFA ile test edilmemiştir. BÜT 3-5'in yapı geçerliği çalışmalarında kuramsal yapı PCA ile ortaya koyulmuş olup aynı zamanda DFA ile de doğrulanmıştır. Bu durumun sebebi küçük yaşlarda görevlerin (hipotez, deney) disiplinlere (fizik, kimya, biyoloji) göre daha baskın gelmesi olabilir. Alan yazında yapılan çalışmalarda yaratıcılık puanlarının yaratıcılığın bilişsel süreçlerinden ziyade görevlerin altında toplandığına dair çeşitli araştırma bulgularına rastlanmaktadır (Almeida et. al., 2008, s. 56; Kim, 2006, s. 7). Ancak alan yazında rastlanan diğer yaratıcılık testlerinde disipline özgü maddeler geliştirmek yerine süreç becerilerine yönelik maddeler geliştirilmiştir. Buna rağmen yapı geçerliği çalışmalarında kuramsal alt yapıyı ortaya koyulamamıştır. Bu bakımdan BÜT 3-5'in yapı geçerliğinin alan yazındaki diğer testlerle karşılaştırıldığında daha güçlü kanıtlarla desteklendiği düşünülebilir.

Alan yazındaki diğer testlerle karşılaştırıldığında kuramsal yapının PCA'de ortaya çıkmasının ve DFA ile desteklenmesinin temel nedeninin test geliştirme sürecindeki farklılıklar olduğu düşünülebilir. Ölçek geliştirme uzmanları (DeVilles, 2003; Hogan, 2002; Hinkin, 1995), test geliştirme sürecinde alan uzmanları tarafından kuramsal alt yapıya uygun çok sayıda maddenin geliştirilmesini önermektedirler. BÜT 3-5'in madde havuzu kuramsal alt yapıya ilişkin eğitim alan 10 alan uzmanından oluşturulan bir ekip tarafından geliştirilmiştir. Kuramsal alt yapının PCA ve DFA ile desteklenmesinin başlıca nedeninin test geliştirme süreci olduğu düşünülebilir. Bu süreç aynı zamanda testin kapsam geçerliğinin de sağlanması bakımından önemlidir.

Sonuç olarak SDDS modeli temel alınarak geliştirilen BÜT Form3-5'in kuramsal yapısı PCA bulguları ile deneysel olarak desteklenmiş ve DFA bulguları ile doğrulanmıştır. Elde edilen bileşenlerdeki maddelerin yüksek faktör yükleri maddelerin bileşenlerinin saf birer göstergesi olduğu düşünülebilir. Açıklanan yüksek varyans BÜT 3-5'in kuramsal yapısının delili olarak değerlendirilmelidir. Bileşenlerin toplam varyansa

katkıları birbirlerine yakın olması iki bileşenli yapıya ek kanıt olarak değerlendirilebilir. Maddeler için hesaplanan ortak faktör varyanslarının yüksek olması ise ölçeğin homojenliğinin göstergesidir. DFA bulgularına göre yüksek uyum indeksleri elde edilmiştir. Bu bağlamda PCA ve DFA bulgularından hareketle BÜT 3-5'in yapı geçerliğinin yeterli düzeyde olduğu söylenebilir.

5.1.1.2. BÜT 3-5'in ayırt edicilik geçerliğine ilişkin tartışma ve sonuç

BÜT 3-5'in ayırt edicilik geçerliği iki şekilde incelenmiştir. Öncelikle katılımcılar her madde için alt %27 ve üst %27'lik gruplara ayrılmış ve grupların ortalamalarının farklılaşıp farklılaşmadığına bağımsız örneklem t-testi ile bakılmıştır. İkinci olarak 3, 4 ve 5. sınıf öğrencilerinin madde, alt ölçek ve test toplam puan ortalamalarının farklılaşıp farklılaşmadığına tek yönlü ANOVO ile bakılmıştır.

Testte yer alan 4 madde akıcılık, esneklik ve bileşik yaratıcılık puanları bağlamında üst gurubun puan ortalamaları alt gruptan anlamlı bir şekilde daha yüksektir. Bu bulgu alan yazındaki diğer testlerin ayırt edicilik çalışmaları ile uyumludur (Ayas and Sak, 2013; Sak and Ayas, 2010; Kanlı ,2014; Hu and Adey, 2002; Chin and Siew, 2015; Siew, Chong and Chin, 2014). Sınıf düzeylerine göre ayırt edicilik analizlerinde test toplam puanına göre ve hipotez alt ölçeğine göre 5. sınıflar 4. sınıflardan, 4. sınıflar ise 3. sınıflardan daha yüksek puanlar almışlardır (Tablo 3.11). Deney alt ölçeğinde 5. sınıf öğrencileri 4. sınıf öğrencilerinden, 4. sınıf öğrencileri ise 3. sınıf öğrencilerinden yüksek puanlar almışlardır. Ancak 4 ve 3. sınıfların puan ortalamaları arasındaki fark deney maddeleri için anlamlı değildir.

Ayırt edicilik bulguları çelişkili olmakla birlikte alan yazında var olan diğer bilimsel yaratıcılık testlerinin yaşa veya sınıf düzeyine göre ayırt edicilik geçerlik çalışmaları ile benzerlik taşımaktadır. Örneğin Kanlı (2014, s.115) Y-BÇT'nin alt ve üst gruplar için ayırt edicilik düzeyinin iyi olmasına rağmen testin 5, 6, 7 ve 8. sınıf öğrenciler için ayırt edicilik çalışmasında 5 ve 6. sınıf öğrencilerinin ortalama puanları arasında anlamlı fark olmadığını belirtmiştir. Hu ve Adey (2002) ise geliştirdikleri testin yaşa göre ayırt edicilik çalışmalarını test toplam puanını kullanarak 12, 13 ve 15 yaş çocukları ile yapmışlar ve 13 ve 15 yaş arasında ayırt ediciliğin olmadığını rapor etmişlerdir. Elde edilen bu sonuçları her iki araştırmacı da yaratıcılığın gelişiminin lineer olmadığı şeklinde yorumlamışlardır.

Yukarıda ayırt edicilik bulgularından bahsedilen testlerin alan bilgisine dayanması tartışmanın yönünü değiştirebilir. Normal şartlar altında alan bilgisi sınıf düzeyi arttıkça artması beklenebilir. Ancak Simonton (1983, s. 149) eğitim ve yaratıcılık arasında doğrusal olmaktan çok u-şeklinde bir ilişkiden bahsetmektedir. Çocukluk yılları için düşünüldüğünde çocukluktan ergenliğe geçiş dönemlerinde grup kabulü ve sosyal beceriler çocukların yaratıcılıklarında düşüşe neden olabilmektedir (Findley and Lumsden, 1988, s. 44).

Son 50 yılda farklı kültürlerde ve farklı alanlarda yapılan birçok araştırmada rastlanan “4. Sınıf Düşüşü-Fourth Grade Slump” olarak bilinen fenomen bu durumun çıktılarında birisi olarak değerlendirilebilir. Bu fenomene göre akademik başarı ve yaratıcılık gibi özelliklerde 4. sınıfta bir düşüş yaşanmaktadır (Barbot, Lubart and Besançon, 2016, s. 36). Örneğin Torrance (1968, s. 197) 5 yıllık uzamsal çalışmasında yaratıcılıkta 4. Sınıfların %59’unda düşüş olduğuna ve sadece %11’inde gelişime rastlamıştır. Vernon (1948) yaptığı çalışmada ise çocukların hipotez geliştirme becerilerinin gelişerek 11 yaşında zirveye ulaştığını (Sor ve Tahmin Testi-Ask-and-Guess) belirtmiştir (aktaran Hu ve diğerleri, 2010, s. 48). Bilimsel yaratıcılık bağlamında düşünüldüğünde gelişimin doğrusal olmadığını ortaya koyan çeşitli çalışmalara rastlanmaktadır (Hu et. al., 2004, s. 718; Hu and Han, 2006; Hu et. al., 2010, s. 49). Örneğin Hu ve diğerleri (2004, s. 51) 3-11. sınıf öğrencilerinin bilimsel problem bulma becerilerini inceledikleri çalışmada bilimsel yaş veya sınıf düzeyi ile doğrusal bir ilişki olmadığı, 4. sınıfta bir düşüşün yaşandığını rapor etmişlerdir. Benzer çalışmalar diğer alanlarda da görülmektedir. Örneğin Sak ve Maker (2006, s. 287) matematikte yaratıcılıkta 3 ve 4. sınıf öğrencilerinde bir düşme veya durgunluk olabileceğini rapor etmişlerdir.

Çeşitli araştırmacılar ise 4. sınıf düşüşünü müfredat ile ilişkilendirmektedir (Best, Floyd and McNamara, 2004, s. 4, Chall and Jacobs, 2003, s. 13; Sanacore and Palumbo, 2009, s. 68). Ülkemizde bu konu üzerine yapılmış herhangi bir araştırma bulgusuna rastlanmamış olsa da 3. sınıfta öğretim programında 3 akademik ders yer almaktadır. 4. sınıfta birlikte bu derslerin sayısı sert bir şekilde artmaktadır. Ders sayısı ile birlikte öğrenme süreçleri de değişmektedir. Fen bilgisi, sosyal bilgiler ve yabancı dil gibi derslere alan öğretmenleri girmektedir. Ayrıca derslerde anlatı şeklindeki metinler, bilgilendirici metinlere dönüşmektedir. Uluslararası alan yazında sıralanan nedenlerin

ülkemiz içinde geçerli olduğunu düşünebiliriz. Dolayısı ile BÜT 3-5'in ayırt edicilik geçerliğinin 4. sınıf düşüşünden etkilendiğini –tek neden olmamakla birlikte- söylemek mümkündür.

Bu araştırmada ayırt edicilik çalışmalarında 3 ve 4. sınıf öğrencilerinin puanlarının farklılaşmamasının nedenlerinden biri de katılımcıların özellikleri olabilir. Araştırmada veriler iki ilkokul ve bir ortaokula devam eden öğrencilerden toplanmıştır. Bu ilkokullardan birisinde özel yetenekli öğrenciler için kaynak oda uygulamasında Anadolu Üniversitesi ile ortak bir proje kapsamında farklılaştırılmış fen bilimleri, matematik ve Türkçe dersleri verilmektedir. Öğrenciler bu derslerin yaklaşık %30'unu kaynak oda da almaktadırlar. 3. sınıflardan 18 öğrencinin 4. sınıflardan ise 10 öğrencinin bu programa devam ettiği belirlenmiştir. 3. sınıfa devam eden özel yetenekli öğrenci sayısının 4. sınıfa devam eden özel yetenekli öğrenci sayısından fazla olması bu durumun nedeni olabilir. Her ne kadar bulgular kısmında rapor edilmemiş olsa da yapılan ek analizlerde kaynak oda uygulamasının yapıldığı okulda 3 ve 4. sınıfların deney puan ortalamaları arasında anlamlı farklar bulunmamıştır. Ancak uygulama yapılan diğer ilkokulda 4. sınıf öğrencilerinin deney puan ortalamalarının 3. sınıfların deney puan ortalamalarından anlamlı şekilde daha yüksek olduğu bulunmuştur ($t_{313}=3,63$; $p<.005$; $\eta^2=.04$).

Sonuç olarak BÜT 3-5'in ayırt edicilik geçerliği çalışmalarında alt ve üst grupları ayırt ettiği sonucuna ulaşılmıştır. Test toplamı için akıcılık, esneklik ve bileşik yaratıcılık puanlarının 3, 4 ve 5. sınıf öğrencileri için farklılaştığı görülmüştür. Alt ölçeklerden deney maddelerinin ise 3 ve 4. sınıfların ortalamaları bağlamında farklılaşmadığı görülmüştür. Bu durumun nedeni ve birçok kültürden bağımsız olarak ortaya çıkan 4. sınıf düşüşü veya örneklemden kaynaklanmış olabilir. Örneklemden kaynaklanan sorunlar ise bu araştırmanın sınırlılıklarından birisi olarak değerlendirilebilir.

5.1.1.3. BÜT 3-5'in uyum geçerliğine ilişkin tartışma ve sonuç

BÜT 3-5'in uyum geçerliğini değerlendirmek için bilimsel yaratıcılıkla teorik olarak ilgili olduğu düşünülen öğrencilerin karne notları, akademik benlik algıları ve fen bilimlerine yönelik tutumları ve test puanları arasındaki korelasyon katsayıları incelenmiştir.

5.1.1.3.1. BÜT 3-5'in ders notları ile uyum geçerliğine ilişkin tartışma ve sonuç

Araştırmada öğrencilerin matematik ve fen bilgisi karne notları ve BÜT 3-5 puanları arasında orta düzeyde ve anlamlı korelasyon katsayıları bulunmuştur (bakınız Tablo 3.9). Sonuçların derslere, puan türlerine ve sınıf düzeylerine göre birbirlerine yakın çıkmalarından dolayı BÜT 3-5'in akademik başarı ile anlamlı ilişkisi olduğu söylenebilir. Bu sonuçlara göre öğrencilerin fen ve matematik ders notları arttıkça BÜT 3-5 puanları da artmaktadır. Ders notları ve BÜT 3-5'in ortak varyansları ise %9,6 ve %16 arasında değişmektedir.

Alan yazında genel yaratıcılık ve akademik başarı arasında ilişki Getzels ve Jackson'ın (1962) araştırmalarına kadar uzanmaktadır. BÜT 3-5'in ders notları ile olan ilişki, genel yaratıcılık testleri ve akademik başarı arasındaki ilişki ile benzerlikler taşımaktadır. Getzels ve Jackson (1962) çalışmanın sonuçlarını 120 IQ puanına kadar zekâ akademik başarıyı etkilemekle birlikte bu eşğin üzerinde zekânın devreden çıkarak yaratıcılığın akademik başarı üzerinde bir etkisinin olacağını belirtmişlerdir (aktaran Ai, 1999, s. 330). Geçen 50 yıl içerisinde birçok araştırmada yaratıcılık ve akademik başarı arasında anlamlı ilişkiler rapor edilmiştir. Örneğin Marjoribanks (1976) ise araştırmasında okul notları yerine standart bir başarı testi kullanmış ve farklı sonuçlar rapor etmiştir. Araştırmanın bulgularına göre matematik başarı ile TTCT arasında .36 ve fen (fizik ve biyoloji) ve TTCT arasında .37 ve .45 korelasyon katsayıları bulmuştur ($p < .001$). Araştırmanın ilginç olan diğer bir sonucu ise 120 IQ puanına kadar (eşik kuramı) yaratıcılık ve akademik başarı arasında bir ilişki varken, bu eşğin üzerinde ilişki olmadığı bulunmuştur. Ai (1999, s. 333) ortaokul öğrencilerinin TTCT puanları ve fen bilimleri ders notları arasında yaklaşık .33 ve matematik ders notu arasında ise .31 düzeyinde anlamlı korelasyon katsayıları bulmuş, yaratıcılık-akademik başarı ilişkisinin kızlarda ve erkeklerde farklılaştığını rapor etmiştir. Ai (1999) bulguları okullarda akademik başarıyı arttırmak için yürütülen faaliyetlerin bir kısmının yaratıcılığa katkısı olmadığı şeklinde yorumlamıştır. Gajda (2016, s. 256) akademik başarı ve yaratıcılık arasında zayıf fakat anlamlı ilişkiler bulmuş ve bu ilişkinin ortaokulda daha yüksek olduğunu belirtmekte ve ergenlik ile ilişkilendirmektedir. Gajda, Karwowski ve Beghetto (2017, s. 285-289) yaptıkları meta analiz çalışmasında ise akademik başarı ve yaratıcılık arasındaki ortalama korelasyon katsayısını .22 ($p < .05$) olarak hesaplamışlar, bu değerın zamandan ve kültürden bağımsız olduğunu ancak

öğretim kademesinden etkilendiğini (liseler için $r=.33$ ve ilköğretim için $r=.15$) rapor etmişlerdir. Araştırmanın bulgularına göre öz değerlendirme yerine yaratıcılık testleri, karne notları veya öğretmen görüşleri yerine standart başarı testleri, figüral testler yerine sözel yaratıcılık testleri kullanıldığında korelasyon güçlenmektedir. Genel yaratıcılık ve akademik başarı ilişkisinin rapor edildiği araştırma bulgularından hareketle, bu iki değişken arasındaki ilişkinin kullanılan ölçümlere, katılımcıların özellikleri ve disipline göre değiştiği şeklinde bir yoruma ulaşılabılır. Zamandan ve kültürden bağımsız şekilde bu iki değişken arasında düşük, nadiren orta düzeyde korelasyon katsayıları rapor edilmiştir. Daha önceden yapılmış çalışma bulgularından hareketle BÜT 3-5'in karne notları ile ilişki katsayılarının büyük ölçüde örtüştüğü ve dolayısıyla BÜT 3-5'in karne notları ile uyum geçerliğine ek kanıtlar olarak değerlendirilebilir.

Alan yazında bilimsel yaratıcılık ve akademik başarı arasındaki ilişkinin araştırıldığı çalışmalarda da benzer sonuçlar rapor edilmiştir. Örneğin Kanlı (2014, s. 121) ortaokul öğrencilerinin Y-BÇT puanları ile fen ders notları arasında $.390$ ($p<.001$), matematik ders notları ile $.487$ ($p<.001$) korelasyon katsayıları bulmuştur. Kanlı YBÇT puanlarına göre öğrencilerin farklı gruplara ayrıldığında fen ve YBÇT arasındaki ilişkinin de arttığını rapor etmiştir. Ayrıca okullardaki değerlendirme ölçütlerinden ve fen ve matematik dersine merkezi sınav sisteminde verilen önemden kaynaklanabileceğini belirtmiştir. SDDS modeli temelinde geliştirilen BÜT ve ders notları arasındaki ilişkinin incelendiği iki çalışmanın da benzer sonuçlar rapor ettiği görülmektedir. Ayas (2010, s.92) BÜT puanları ile fen ders notları arasında $.45$, matematik notları arasında $.47$ ($p<.001$) korelasyon katsayıları bulmuştur. Ayas ve Sak (2014) ise matematikle $.31$, fen bilimleri ile $.36$ ve Matematiksel Yetenek Testi ile $.52$ ($p<.001$) korelasyon katsayıları bulmuşlardır.

BÜT 3-5'in uyum geçerliği çalışmasında fen ve matematik ders notları ile orta düzeyde bir ilişki bulunmuştur. Alan yazında akademik başarı ve yaratıcılık üzerine yapılan araştırmalarda düşük ve orta düzeyde ilişkiler rapor edilmiştir. Bilimsel yaratıcılık ve ders notları arasındaki ilişkinin araştırıldığı çalışmalarda ise orta düzeyde korelasyon katsayıları rapor edilmiştir. Bu bağlamda alan yazındaki bulgular ile bu çalışmanın bulgularının büyük benzerlikler taşıdığı düşünülebilir. Sonuç olarak BÜT 3-5'in ders notları bağlamında uyum geçerliğinin beklendiği şekilde olduğunu söyleyebiliriz.

5.1.1.3.2. BÜT 3-5'in fen bilimlerine yönelik tutum ve fen akademik benlik algısına göre uyum geçerliğine ilişkin tartışma ve sonuç

BÜT 3-5'in uyum geçerliği çalışması için BÜT 3-5 ve fen akademik benlik algısı (FABAÖ) ve fen bilimlerine karşı tutum (BYTÖ) ikinci boyutunu fen bilgisi akademik benlik algısı arasındaki korelasyon katsayılarına bakılmıştır. Yapılan incelemede BYTÖ ve FABAÖ puanlarının 3 ve 4. Sınıf öğrencileri için normal dağılmadığı ve varyansının çok düşük olduğu görülmüştür. Bu durumun başlıca nedeni her iki ölçekte yer alan maddelerden kaynaklanabilir. Örneğin BYTÖ'de fen dersini sevme, belgesel izleme gibi maddeler yer almaktadır. Bu maddelerin tamamı öğrencilerin beğendiği şeyler olduğundan üst puanlarda bir yığılma görülmüş olabilir. FABAÖ'de ise “fen bilgisinde başarılıyım” gibi maddeler yer almaktadır.

Türkiye'de ortaokul yerleştirmelerinde karne notları informal bir şekilde etkin olarak kullanılmaktadır. Örneğin öğrenciler ortaokullarda karne notlarına göre derece sınıflarına yerleştirilebilmektedirler. Bu durumun öğretmenlerin öğrencilere verdikleri notların dağılımını öğrenci lehine arttırdığı düşünülebilir. Dolayısı ile öğrenciler kendilerini karne notları ile değerlendirdiklerinde olduğundan daha yetenekli algılamış olabilirler ve FABAÖ'ne yüksek puanlar vermiş olabilirler. Örneğin 3. sınıf öğrencilerinin ders notları “geliştirilmeli”, “iyi” ve “çok iyi” şeklinde değerlendirilmektedir. Araştırmaya dahil edilen 251 üçüncü sınıf öğrencisinden sadece 2 tanesinin matematik notu “geliştirilmeli”, 17 tanesinin notu ise “iyi” olarak karneye işlenmiştir. Fen bilgisi notlarına bakıldığında ise “geliştirilmeli” karne notuna hiç rastlanmamış ve 10 tane öğrenci “iyi” notu ile değerlendirilmiştir. 4. sınıflarda ise öğrencilerin %91,3'ünün fen değerlendirmesi karnelere “5” olarak işlenirken, matematikte bu oran 82,7'dir. Bu durumun bir sonucu olarak öğrenciler kendilerini gereğinden daha yetenekli algılamış olabilirler. Diğer bir sorun ise ölçeklerin 5-li Likert tipinde olmasıdır. Değişkenliğin artması için birçok araştırmacı 7 veya 9'lu Likert skalasını önerse de öğrencilerin yaşlarının küçük olmasından dolayı 5'li Likert tercih edilmiştir.

Yapılan Pearson momentler çarpımı korelasyon analizi sonucunda FABAÖ ve BÜT 3-5 puanları arasında anlamlı fakat düşük korelasyon katsayıları bulunmuştur (.286-.297 arasında; $p < .001$). Bulgular BÜT 3-5'in akademik benlik algısı ile uyum geçerliğinin kanıtları olarak değerlendirilmelidir. Ancak elde edilen korelasyon katsayılarının düşük olduğu gözden kaçırılmamalıdır. Alan yazında benzer bir çalışmada

Kanlı (2014, s. 122) ortaokul öğrencileri için çok benzer sonuçlar rapor etmiştir. Çalışmada Y-BÇT ile öğrencilerin yetenek algıları arasından .273 ve .240 ($p<.001$) korelasyon katsayıları bulunmuştur. Yakın bir disiplin olan matematik alanında Güçyeter (2013, s. 112) matematikte akademik benlik algısı ile yetenek seviyesi arasında anlamlı ve orta düzeyde ilişki bulmuştur.

Araştırmada BÜT 3-5 ile FABAÖ arasında anlamlı ilişkiler bulunmuş olsa da korelasyon katsayılarının zayıf ilişkiye işaret ettiği görülmektedir. Alan yazında yaratıcılık ve akademik benlik algısı arasında düşük ilişki rapor edilse de araştırmalar detaylı incelendiğinde çelişkili sonuçlar rapor edildiği görülmektedir. Çelişki tahmin geçerliği ve uyum geçerliği arasındaki farklılıktan kaynaklanıyor olabilir. Yaratıcı bireylerin benlik algılarının normal akranlarından daha yüksek olduğuna dair geçmişten günümüze istikrarlı araştırma bulgularına rastlanmaktadır (Getzels and Jackson, 1962; Fleith, Renzulli and Westberg, 2002; Bounelli, Makri and Mylonas, 2009; Al-Quasy and Turki, 2011). Örneğin Fleith, Renzulli ve Westberg (2002, s. 380) yaratıcılık programının öğrencilerin yaratıcılıkları üzerinde anlamlı etkileri olmasına rağmen, benlik algıları üzerinde olumlu bir etkisi olmadığını ancak olumsuz bir etkiye de rastlamadıklarını rapor etmişlerdir. Saurenman ve Micheal (1980, s. 84) üstün zekâlı tanısı almış 96 ortaokul öğrencisi ile yaptıkları çalışmada gurubu akademik başarılarına göre düşük başarılı ve yüksek başarılı olmak üzere ikiye ayırmışlardır. Yüksek başarılı gurubun yaratıcılık ve akademik benlik algılarının düşük başarılı gruptan anlamlı şekilde yüksek olduğunu ancak akademik benlik algısı ve yaratıcılık arasında ilişki olmadığını rapor etmişlerdir. Spaulding (1965, s. 39) ise 4-6. sınıf öğrencileri için uygulanan özel yetenek sınıfındaki öğrencilerin akademik benlik algılarının ve yaratıcılıklarının normal akranlarından yüksek olmasına rağmen bu iki kavram arasında bir ilişki bulamamışlardır. Schempp, Cheffer ve Zaichkowsky (1983) ise normal gelişim gösteren 3-5. sınıf öğrencileri ile yaptıkları çalışmada demokratik öğretmen davranışlarının öğrencilerin yaratıcılık ve akademik benlik algıları üzerinde olumlu etkileri olmasına rağmen bu iki özellik arasında $-.056$ ($p<.05$) korelasyon katsayısı bulmuşlardır. Al-Qaisy ve Turki (2011, s. 98) yaptıkları çalışmada yaratıcılık düzeylerine göre ayırdıkları öğrencilerden orta gruptakilerin akademik benlik algısının üst ve alt gruptan yüksek olduğunu, en düşük akademik benlik algısının ise yaratıcı grupta olduğunu belirtmiştir. Araştırmacılar ek olarak akademik benlik algısı düşük olan öğrenci sayısının beklenenden yüksek olduğunu

belirtmişler, akademik benlik algısı ile yaratıcılık arasında -0.065 ve -0.055 arasında korelasyon katsayıları bulmuş olmalıdır. Araştırmada düşük akademik benlik algısına sahip öğrencilerin sayılarının yüksek olanlardan daha fazla olduğunu bulmuşlardır. Bilimsel yaratıcılık alanında da benzer sonuçlara rastlanmıştır. Kanlı (2014, s. 123) düşük ancak anlamlı korelasyon katsayıları rapor etmiş ve kullanılan bilimsel yaratıcılık testinden yüksek puan alan öğrencilerin büyük çoğunluğunun akademik benlik algısı puanının yüksek olduğunu belirtmiştir. Verilen araştırma bulgularından hareketle yaratıcı bireylerin benlik algıları yüksek olabilir ancak aralarında doğrusal bir ilişkiden söz etmek için istikrarlı araştırma bulgularının olmadığı söylenebilir.

BÜT 3-5'in fen bilimlerine yönelik tutumlar ile uyum geçerliği için kullanılan BYTÖ puanları için $.238$ ve $.293$ arasında ve $.001$ 'da anlamlı korelasyon katsayıları bulunmuştur. Bulgular alana özgü yaratıcılık kuramları ile örtüşmektedir (Csikszentmihalyi, 1997; Amabile, 1983; Renzulli, 2002). BÜT 3-5 alana özgü bir test olmasından dolayı test puanları, alan bilgisi ve geçmiş tecrübelerden etkilenmektedir. Herhangi bir alandaki bilgi birikiminin o alandaki ilgi ile yakından ilişkili olduğu bilinmektedir. Örneğin Gardner (1975, s.21) fen bilimlerine yönelik ilgi üzerine yaptığı alan yazın taramasında akademik başarı ve ilgi arasında $.25$ -. 35 arasında korelasyon katsayılarına rastlandığını, ortaokuldan sonra bu ilişkinin $.40$ -. 60 seviyelerine çıktığını belirtmektedir. Bu bağlamda bilime yönelik tutumların ve akademik başarı arasındaki ilişkinin tecrübe ve alan bilgisi ile birlikte arttığı düşünülebilir. Bu tez çalışmasının katılımcılarının fen bilimleri alanındaki sınırlı tecrübe ve bilgi birikimlerinden dolayı BÜT 3-5 ve BYTÖ ve BÜT 3-5 ve fen notları arasında düşük ve orta düzeyde korelasyon katsayıları bulunduğu düşünülebilir.

Elde edilen korelasyon katsayılarının daha yüksek olması beklense de bulgular alan yazında rastlanan benzer çalışmaların sonuçları ile benzerlik taşımaktadır. Örneğin Kanlı (2014, s. 123) geliştirdiği bilimsel yaratıcılık testinin ortaokul öğrencilerinin fen bilimlerine olan ilgi ile $.142$ -. 287 ($p < .001$) arasında korelasyona sahip olduğunu bulmuştur. Ortaokul öğrencileri ile çalışan Güçyeter (2013, s. 144) matematik yetenek testi ve matematiği sevme arasında $.083$ ($p < .05$) ve $.267$ ($p < .001$) düzeyinde korelasyon katsayısı bulmuştur. Kılıç (2011, s. 86) ise 8. sınıf öğrencilerinin bilimsel yaratıcılık düzeyleri ile bilime yönelik tutumları arasında bir ilişki olmadığını belirtmiştir.

Yapılan arařtırmalarda fen biliminin doęasına ynelik algıların bilimsel yaratıcılık zerinde bir etkisinin olduęu belirtilmektedir (Liang, 2002). Bilimin doęasına ynelik algıların aynı zamanda fen bilimlerine ynelik ilgiyi de etkiledięi bilinmektedir. Fen biliminin doęasında ise gzlem yapma, hipotezler retme, deneyler yapma ve ıkarımda bulunma gibi eylemler yatmaktadır. Bu bakımdan fen bilimlerine ynelik ilgi ve tutumun yaratıcılıkla ilgisi dięer alanlara gre daha belirgindir (Usta and Akkanat, 2015, s. 1409). Birok ęrenci bu tr eylemleri tecrbe edemedikleri iin fen bilimlerine karřı tutumları, bilimin doęasına ynelik algıları ve dolayısı ile bilimsel yaratıcılıkları bu durumdan olumsuz etkilenmiř olabilir (Yager and McCormak, 1989, s. 49). nk lkemizdeki sınav sisteminden dolayı ęretmenler ve veliler bu tr beceriler yerine ulusal akademik sınavlarda yksek bařarıya odaklanmaktadır. Okulların biroęunda laboratuvar bulunmadıęı veya ęretmenlerin var olan laboratuvarları kullanmadıkları bilinmektedir. Bu baęlamda BYT ve BT 3-5 arasındaki iliřkinin okul sisteminden olumsuz etkilendięi dřnlebilir.

Bu arařtırma bir test geliřtirme alıřması olduęundan geerlik analizleri iin sonuların yorumlanmasında kolaylık saęlaması ve geici olarak ntr bir yaklařım sunması bakımından korelasyon katsayıları hesaplanmıřtır (DeVilles, 2003, s. 81). Korelasyon katsayıları uyum geerlięi iin faydalı olsa da tahmin geerlięi iin uygun olmayabilir. Bu baęlamda BT 3-5 ile FABA ve BYT arasındaki dřk korelasyon deęerleri uyum geerlięinin de dřk olduęu řeklinde deęerlendirilmemelidir. Daha nceden de belirtildięi gibi katılımcıların FABA ve BYT'ne yksek puan vermiřlerdir. Bu durum ise eřitlilięin ve varyansın dřk olmasına neden olmaktadır. Dřk eřitlilik ise korelasyon katsayılarının dřk ıkmasının nedeni olabilir. BT 3-5'in FABA ve FT ile dřk korelasyon katsayıları bulunmuřtur. Bulgular alan yazında var olan benzer alıřmalarla rtřmektedir. Bu baęlamda .001 seviyesinde anlamlı korelasyon deęerleri BT 3-5'in uyum geerlięinin kanıtı olarak deęerlendirilebilir.

5.1.2. BT 3-5'in gvenirlięine iliřkin tartıřma ve sonu

BT 3-5'in gvenirlięine iliřkin bulgular i tutarlık gvenirlięi ve okuyucular arası gvenirlik bulguları zerinden tartıřılarak sonuca ulařılmıřtır.

5.1.2.1. *BÜT 3-5'in iç tutarlık güvenilirliğine ilişkin tartışma ve sonuç*

BÜT 3-5'in iç tutarlık güvenilirliği için Cronbach Alpha katsayısı değerlendirilmiştir. Ek kanıt sağlaması bakımından madde-toplam ve maddeler arası korelasyon katsayıları da incelemeye dâhil edilmiştir.

BÜT 3-5 toplam akıcılık, esneklik ve bileşik yaratıcılık için hesaplanan Cronbach Alpha sırasıyla .743, .701 ve .736 olarak bulunmuştur. Hipotez alt testi için bu iç tutarlık katsayıları sırası ile .594, .433 ve .583; deney alt testi içinse sırasıyla .828, .813 ve .825 olarak hesaplanmıştır. BÜT 3-5 toplam puanları üzerinden yapılan analiz bulgularına göre BÜT 3-5'in güvenilirlik bulgularının eterli düzeyde olduğu düşünülebilir. Hipotez ve deney alt testlerindeki iç tutarlık katsayılarının ise ayrıca tartışılması gerekmektedir.

1. Madde silindiğinde Cronbach Alpha katsayısında bir yükselme olmaktadır (.758). Bu durumun başlıca nedeni 1. Maddenin ortalamasının diğer maddelerden düşük olmasıdır. Ayrıca madde için üretilen maksimum akıcılık puanının 7,00 olması çoğul üretime uygunluğunun düşük olduğu ve testteki diğer maddelerle uyumunun az olduğu şeklinde yorumlanabilir. Ancak bu maddenin silinmesi durumunda Cronbach Alpha katsayısındaki artışın istatistiksel olarak anlamlı olmayacağı görülmüştür.

Bilindiği gibi Cronbach Alpha katsayısı madde sayısından etkilenmektedir. BÜT 3-5'te 4 madde olmasından dolayı bu değerlerin tatmin edici olduğu düşünülebilir. Ayrıca BÜT 3-5 iki bileşenli bir testtir. Elde edilen değerler Cronbach Alpha katsayısının çok bileşenli ölçekler için uygun olmadığı yönündeki görüşler göz önünde bulundurularak yorumlanmalıdır (Kline, 2005, s. 175; Urbina, 2004, s. 131). Deney alt ölçeği için hesaplanan Cronbach Alpha iç tutarlık katsayısı iyi düzeyde (akıcılık için .828, yaratıcılık için .825) olarak hesaplanmıştır. Hipotez alt ölçeği için hesaplanan Cronbach Alpha katsayısının ise düşük olduğu görülmüştür (akıcılık için .594 ve yaratıcılık için .583). Alt ölçeklerin her ikisinde de 2'şer madde olmasına rağmen hipotez maddeleri için hesaplanan iç tutarlık katsayısı düşük olmasının nedeni madde varyansları olabilir. Hipotez maddeleri daha kolay maddeler olduğundan madde varyansları da düşük çıkmıştır. Değişkenliğin daha az olduğu hipotez alt testinde maddeler arası korelasyonun düşük çıktığı söylenebilir. Deney maddeleri daha zor olduğu için bu alt ölçek için toplam varyans daha yüksek çıkmıştır. Alt ölçeklerin ortalama, standart sapma ve alt ölçek maddeler arası korelasyon katsayıları bu görüşü desteklemektedir. Hipotez alt testi

akıcılık ortalaması 3,54 (SS=2,13) ve deney alt testi akıcılık ortalaması 2,37 (SS=3,42) olarak hesaplanmıştır. Hipotez alt testi maddelerinin akıcılık puanları arasındaki korelasyon katsayısı .431, deney maddelerinin akıcılık puanları arasındaki korelasyon katsayısı .708 olarak hesaplanmıştır. Alt testlerin toplam esneklik ve bileşik yaratıcılık puanları için de benzer sonuçlar bulunmuştur.

İç tutarlığın diğer bir göstergesi olan düzeltilmiş madde toplam korelasyonlarının tamamının .30'un üzerinde olduğu görülmüştür (Field, 2009, s. 675). Düzeltilmiş madde toplam korelasyonları en düşük madde ise daha önce tartışılan 1. Madde olduğu ve en yüksek olan maddelerin ise deney maddeleri olduğu görülmüştür. Bu bulgu da deney alt ölçeği için hesaplanan Cronbach Alpha değerinin iyi olmasının ve hipotez maddelerinin Cronbach Alpha katsayısının düşük olmasının nedeni olabilir. Daha önceden tartışıldığı gibi deney maddeleri için akıcılık ranjları 12 ve 13 iken hipotez maddeleri için bu değerler 7 ve 8'dir. Bu durum hipotez maddelerinin varyansının düşük olmasına ve Cronbach Alpha değerinin düşmesine neden olmaktadır.

İç tutarlık güvenilirliği kapsamında madde-toplam ve maddeler arası korelasyon katsayıları incelenmiştir. Maddeler arası korelasyon katsayılarının .258 ve .708 arasında değiştiği bulunmuştur. Her puan türü için hesaplanan ortalama maddeler arası korelasyon değerleri ise akıcılık için .411, esneklik için .365 ve bileşik yaratıcılık için .417 olarak bulunmuştur. Maddeler arası korelasyon katsayılarının .15- .50 arasında olması ve ortalama madde-madde korelasyon değerinin .40 civarında olmasından dolayı BÜT 3-5'in iç tutarlık güvenilirliğinin yeterli düzeyde olduğu söylenebilir (Briggs and Check, 1986, s. 115).

SDDS modeli temel alınarak geliştirilen BÜT 6-8'in 6. sınıflar için Cronbach Alpha iç tutarlık katsayısı Sak ve Ayas (2013, s.324) tarafından .848 olarak hesaplanmıştır. Bu çalışmada düzeltilmiş madde toplam korelasyonlarının .47 ve .631 arasında değiştiği rapor edilmiştir. BÜT üzerine yapılan bir diğer çalışmada Ayas ve Sak (2014, s. 201) Cronbach Alpha iç tutarlık katsayılarının .80 ile .90 arasında olduğunu belirtmişlerdir. Aynı çalışmada düzeltilmiş madde toplam korelasyonlarının .47 ile .70 arasında, maddeler arası korelasyon katsayılarının ise .17 ile .33 arasında değiştiğini rapor etmişlerdir. Ancak hesaplanan Cronbach Alpha değerlerinin akıcılık, esneklik ve yaratıcılık puanlarının analize aynı anda dahil edildiğinden 5 madde ve 3 puan türünde toplam 15 puanla hesaplandığından yüksek çıktığı düşünülebilir.

Hu ve Adey (2002, s. 397) geliřtirdikleri bilimsel yaratıcılık testinde yer alan 7 madde için Cronbach Alpha tutarlık katsayısının .893, düzenlenmiş madde toplam korelasyonlarının .576 ve .766 arasında olduğunu ve maddeler arası korelasyon katsayılarının ise .421-.729 arasında olduğunu rapor etmişlerdir. Aynı araç için Türk kültürüne adaptasyon çalışmalarında Çeliker ve Balım (2012, s. 16) madde toplam korelasyonlarını ise .369 ve .739 (düzeltilmemiş) olarak bulmuştur. Bu testlerin tek faktörlü olmaları faktör iç tutarlık katsayılarının BÜT 3-5 göre yüksek olmasının nedeni olabilir. Kanlı (2014, s.107) ise geliřtirdiđi yaratıcılık testi için iç tutarlık katsayısının .89, maddeler arası korelasyon katsayılarını ise .103 ve .441 arasında deđiřtiđini bulmuştur. Alt testler için .83 ve .73 Cronbach Alpha katsayıları rapor etmiştir. Adı geçen çalışmaların tamamı ortaokul öğrencileri ile yürütölmüş, madde sayıları 7-25 arasında deđişen testlerdir. 5. sınıf öğrencileri ile çalışan Chin ve Siew (2014, s. 116) SSCM'ne (Hu and Adey, 2002) göre 4 maddelik 2 paralel form hazırlamışlardır. Bu testin Cronbach Alpha katsayılarının paralel formlar için .68 ve .77 olduğunu, maddeler arası korelasyon katsayılarının .15 ve .33 arasında deđiřtiđini rapor etmişlerdir. Aynı çalışmada düzeltilmiş madde-toplam korelasyonları ise .21 ve .49 arasında deđiřtiđi görölmektedir. Bu çalışma 5. sınıf öğrencileri ile çalışılması ve testteki madde sayısı bakımından BÜT 3-5 ile karşılaştırılabilir. Benzerliklerine rağmen güvenilirlik deđerlerinin BÜT 3-5 için daha tatmin edici olduđu söylenebilir. SSCM temel alınarak geliřtirilen ikinci bir çalışmada Siew, Chong ve Chin (2015, s. 1397) 6 maddelik bir test geliřtirmiş ve 30 okulöncesi öğrencisi ile yapılan çalışmada Cronbach Alpha katsayısını .806, maddeler arası korelasyon katsayıları .050 ve .613 arasında, madde toplam korelasyonlarını ise .541-.842 arasında bulmuşlardır.

BÜT 3-5 için yapılan güvenilirlik incelemesinde Cronbach Alpha ve korelasyon katsayılarına göre BÜT 3-5'in güvenilirlik düzeyinin tatmin edici düzeyde olduđu söylenebilir. Az sayıdaki madde için önerilen alternatif iç tutarlık deđerlendirme metotları kullanıldığında BÜT 3-5'in iç tutarlık düzeyinin yeterli olduđu sonucuna ulařılabilir.

5.1.2.2. BÜT 3-5'in okuyucular arası güvenilirliđine ilişkin tartışma ve sonuç

BÜT 3-5 güvenilirlik analizleri için ikinci olarak okuyucular arası güvenilirlik analizlerine başvurulmuştur. Bunun için rastgele seçilen 100 öğrencinin bađımsız olarak iki puanlayıcı tarafından puanlanmıştır. Daha sonra sınıf içi korelasyon katsayıları hesaplanmıştır. Sınıf içi korelasyon katsayılarının .917 ve .998 arasında deđiřtiđi

görülmüştür. Bu değerler alan yazında bilimsel yaratıcılık için rapor edilen okuyucular arası güvenilirlik katsayılarından daha yüksektir (Hu and Adey, 2002; Ayas and Sak, 2014; Sak and Ayas, 2013; Kanlı, 2013).

BÜT 3-5 açık uçlu bir kalem-kâğıt testidir. Puanlama sürecinde okuyucu öncelikle yanıtın doğru olup olmadığına karar vermekte, ikinci olarak doğru olan yanıtları kategorilerin altına yerleştirerek bileşik yaratıcılık puanı hesaplamaktadır. Bu bağlamda BÜT 3-5'in puanlanmasının niteliksel olduğu, içerik analizi yapıldığı düşünülebilir. Okuyucular arası yüksek güvenilirlik katsayılarının nedeni doğru yanıt formunun BÜT 3-5 ile paralel olarak aynı zamanda geliştirilmesi olabilir. Çünkü test geliştirme sürecinde maddeyi hazırlayan uzman ekip aynı zamanda geliştirilen maddelerin yanıtlarını da değerlendirmişlerdir. Daha sonra yapılan ön deneme ve pilot uygulamalarda bu havuz revize edilmiştir. Asıl uygulama da dâhil olmak üzere havuza yeni yanıtlar eklenmiştir. Bu bağlamda BÜT 3-5'in cevap anahtarı geliştirilirken, her aşamada doğru yanıt havuzunun uzman görüşlerine de başvurulmuş genişletilmiş olması, okuyucular arasındaki uyumun yüksek çıkmasının nedeni olabilir. Uyum katsayılarının yüksek çıkmasının bir diğer nedeni ise BÜT 3-5'in alana özgü bir test olması olabilir. Üretilen yanıtlar alan bilgisi ile tutarlı olması gerekmektedir.

Sonuç olarak BÜT 3-5'in maddeleri ve puan türleri için bağımsız okuyucular arasındaki uyum katsayıları yüksek çıkmıştır. Bu değerler aynı zamanda testle birlikte geliştirilen doğru yanıt formunun geçerlik bulguları olarak değerlendirilebilir. Geliştirilen doğru yanıt formu kullanılarak farklı okuyucular testi güvenilir bir şekilde puanlayabilirler.

5.2. Öneriler

Araştırmanın bulguları, tartışma ve sonuçlarından hareketle ileride gerçekleştirilecek araştırmalara ve eğitim uygulamalarına yönelik öneriler geliştirilmiştir. Önerilere geçmeden önce araştırmanın sınırlılıklarının ortaya koyulması önerilerin değerlendirilmesinde yararlı olacaktır. Araştırma Eskişehir ilinde iki ilkokul ve bir ortaokulda yürütülmüştür. Bu durumun katılımcı çeşitliliğini sınırladığı düşünülebilir. Her sınıf seviyesinden 166-234 arasında katılımcıya ulaşılmıştır. Bu bakımdan daha büyük örneklerle bu sınırlılığın giderilebileceği düşünülebilir.

5.2.1. İleri arařtırmalara yönelik öneriler

Arařtırmanın bulguları, tartıřma ve ulařılan sonuçlar dikkate alındığında BÜT 3-5 ve alana özgü yaratıcılık testi geliřtirme çalıřmalarına yönelik birtakım öneriler geliřtirilmiř ve ařaęıda sıralanmıřtır.

- 1) BÜT 3-5'in geliřtirilme süreci kuramsal alt yapının oluřturulması ve bu kuramsal yapıya uygun test maddelerinin geliřtirilmesi ile bařlanmıřtır. Madde havuzu 10 alan uzmanı tarafından geliřtirilmiřtir. Dięer bilimsel yaratıcılık testlerinden farklı yapılan bu arařtırmada kuramsal alt yapı deneysel kanıtlarla desteklenmiřtir. Bu bakımdan rasyonel yöntemin benimsendięi test geliřtirme süreçlerinde, alan uzmanlarına hedeflenen testin kuramsal alt yapısına iliřkin eğitim verilmesi ve daha sonra bu uzmanların kuramsal alt yapıya uygun maddeler geliřtirmeleri önerilebilir.
- 2) BÜT 3-5'in kuramsal alt yapısı kullanılarak okulöncesi, ilkokul, ortaokul ve lise öęrencilerine yönelik bilimsel yaratıcılık testi geliřtirilebilir.
- 3) BÜT 3-5'in alana özgü yapısı kuramsal olarak kullanılıp matematik gibi akademik alanlarda ve resim gibi sanat alanlarından yaratıcılık testleri geliřtirilebilir. Bunun için her alana özgü alt disiplinler (fen için fizik, kimya ve biyoloji gibi), ve süreç becerileri (fen bilimleri için hipotez ve deney gibi) belirlenip yeni kuramsal alt yapılar geliřtirilebilir. Örneęin matematik alt disiplinleri fen alt disiplinlerinin yerine, matematik süreç becerileri de bilimsel süreç becerileri yerine koyularak SDDS modeli yeniden yorumlanabilir ve bu yeni modele göre matematik yaratıcılık testi geliřtirilebilir.
- 4) BÜT 3-5'in test geliřtirme modeli süreç olarak (kuramsal yapı, uzmanlara kuramsal yapı eğitimi, uzmanların yapıya uygun maddeleri geliřtirmesi, ön deneme, pilot ve asıl uygulamalar) farklı alanlarda, farklı ölçek geliřtirme süreçlerinde kullanılabilir.
- 5) Madde havuzu geliřtirme sürecinde öęretmenlerin sürece iliřkin dönütleri informal şekilde alınıp, süreçte deęiřiklikler yapılmıřtır. Bu deęerlendirmeler kayıt altına alınmadığından burada rapor edilmemiřtir. Bu bağlamda sürecin öęretmenlerin gözünden deęerlendirilmesi, öęretmenlere katkıları açısında da niteliksel yöntemlerle deęerlendirilmesi sürecin tekrarlanabilir olmasına katkı sunabilir.

- 6) BÜT 3-5'in yapı geçerliği bulgularına göre hipotez bileşeni ve deney bileşeni belirgin bir şekilde elde edilmiştir. Ancak her bileşende 2'şer madde yer almaktadır. 3, 4 ve 5. sınıf öğrencileri için 4 madde uygun olmakla birlikte her bileşene en az 3 madde yer alacak şekilde daha fazla sayıda maddenin yer aldığı bir test geliştirilebilir. Elde edilecek 6 maddelik form ise 2 oturumda uygulanabilir. Bu durumun geçerlik ve güvenirlik düzeyini de olumlu şekilde attıracağı düşünülebilir.
- 7) Deney ve Hipotez bileşenlerinin belirgin şekilde ortaya çıkması ve bileşenlerin yüksek iç tutarlık değerlerinden dolayı, her bir bileşen ayrı ayrı uygulanarak deney tasarlama ve hipotez geliştirme üzerine yürütülecek çalışmalarda kullanılabilir.
- 8) Bu çalışma sadece iki ilkokul ve bir ortaokulda yürütülmüştür. Katılımcı sayılarına göre 5. sınıf öğrencilerin sayıları 4 ve 3. sınıflardan belirgin bir şekilde daha azdır. Bu bakımdan BÜT 3-5'in psikometrik özellikleri daha büyük katılımcı gruplarında, incelenebilir. Bu sayede çeşitlilik artmış olacağından geçerlik ve güvenirlik verileri daha sağlıklı yorumlanabilir.
- 9) Bu çalışmada uyum geçerliği için akademik başarı öğrencilerin karne notları üzerinde yapılmıştır. İleri araştırmalarda standart bir başarı testi kullanılarak akademik başarı ve BÜT 3-5 arasındaki ilişkiye bakılabilir.
- 10) BÜT 3-5 bir kalem-kâğıt testidir. BÜT 3-5 puanlarının sözel ve motor becerilerden etkilendiği düşünülebilir. Bu bakımdan ileri araştırmalarda bu durumun test edilmesi amacıyla Türkçe ders notları, okuma-yazma becerisi veya Türkçe akademik başarı puanlarının kontrol değişkeni olarak kullanıldığı çalışmalar yapılabilir. Bu sayede testin geçerlik ve güvenirlik değerleri saflaştırılmış olabilir.
- 11) Ayırt edicilik çalışmalarında deney maddeleri için 3 ve 4. sınıf öğrencilerinin puan ortalamalarının anlamlı şekilde farklılaşmadığı görülmüştür. Bu bakımdan etkililik çalışmalarında veya tanılama çalışmalarında değerlendirmelerin toplam puan üzerinden yapılması önerilebilir.
- 12) Dünyada farklı kültürlerde ve derslerde gözlemlenen "Yaratıcılıkta 4. Sınıf Düşüşü" üzerine ülkemizde herhangi bir çalışma yapılamamış olduğu görülmektedir. Bu bağlamda ileri araştırmalarda BÜT 3-5 uzamsal çalışmalarda kullanılarak bu fenomen bilimsel yaratıcılık alanında da incelenebilir.

- 13) BÜT 3-5'in uyum geçerliği için fen bilimlerine ilişkin tutum ve fen bilimlerinde akademik benlik algısı puanları kullanılmıştır. İleri araştırmalarda alan yazında var olan diğer bilimsel yaratıcılık testleri veya bilimsel yaratıcılık değerlendirme ölçekleri ile BÜT 3-5'in ilişkisine bakılabilir.
- 14) BÜT 3-5'in yapı geçerliği çalışmaları akıcılık puanları kullanılarak yapılmıştır. BÜT 3-5 kullanılarak yapılacak çalışmalarda yaratıcılık puanları kullanılarak yapı geçerliği çalışmaları tekrarlanabilir.
- 15) Bu çalışmada uyum geçerliği hariç diğer geçerlik ve güvenilirlik analizleri tüm grup üzerinden, bazen de sınıf düzeylerine göre yapılmıştır. İleri araştırmalarda bilimsel yaratıcılık seviyelerine göre ayrılmış aynı sınıf düzeyindeki öğrenci grupları için geçerlik ve güvenilirlik çalışmaları tekrarlanabilir. Örneğin 5. sınıf öğrencilerinden oluşan büyük bir örnekleme yapı geçerliği üst orta ve alt yetenek grupları için tekrarlanabilir veya ayırt edicilik geçerliği çalışmaları büyük bir 3. sınıf örnekleme üst, orta ve alt guruba ayrılmış öğrenciler üzerinden yapılabilir. Bu çalışmada örneklem sınırlılığından dolayı katılımcıların tamamı kullanılmıştır.
- 16) BÜT 3-5'in puanlanmasında sadece akıcılık, esneklik ve bileşik yaratıcılık puanları kullanılmıştır. İleri araştırmalarda daha kapsamlı bir örnekleme öğrenci yanıtlarının orijinalliği de hesaplanarak analizler bu puan üzerinden tekrarlanabilir. Böylece geleneksel yaklaşımda sıklıkla kullanılan orijinallik puanının bileşik yaratıcılık puanı ile karşılaştırılması yapılmış olur.
- 17) BÜT 3-5'in uyum geçerliği için uygulamadaki sınırlılıklardan dolayı 6 maddelik BYTÖ ve 6 maddelik FABAÖ araçları kullanılmıştır. Öğrencilerin fen bilimlerine tutumları için öğretmen, akran veya ebeveyn görüşleri alınarak bu çalışma tekrarlanabilir.
- 18) Bu çalışmada öğrencilerin BÜT 3-5'e ilişkin görüşleri sadece ön deneme uygulamasında değerlendirilmiştir. İleri araştırmalarda BÜT 3-5 kullanıldığında öğrencilerin teste ilişkin görüşleri alınarak görünüm geçerliğine (face validity) ilişkin kanıt sağlanabilir. Ayrıca görünüm geçerliği için test geliştirme ekibinden olmayan diğer öğretmenlerin de görüşleri alınabilir.

5.2.2. Eğitim uygulamalarına yönelik öneriler

BÜT 3-5'in eğitim uygulamalarında kullanımına yönelik öneriler aşağıda sıralanmıştır.

- 1) BÜT 3-5 özel yeteneklilere yönelik tanılama süreçlerinde diğer testlerle birlikte (zekâ ve genel yaratıcılık testleri, başarı testleri) ek veri sağlaması bakımından kullanılabilir.
- 2) BÜT 3-5'in uygulaması toplamda 40-45 dakikalık bir ders saatinde tamamlanmaktadır. Ancak öğrenci kağıtlarının puanlanması zaman almaktadır. Büyük gruplara tarama amaçlı uygulamalarda deney ve hipotez maddelerinden birer madde alınarak iki maddeden bir tarama formu oluşturulabilir.
- 3) BÜT 3-5 test geliştirme modeli fen bilimlerinde farklılaştırılmış öğretim etkinliklerinin geliştirilmesinde kullanılabilir.
- 4) BÜT 3-5 kuramsal alt yapısı temel alınarak fen bilgisi öğretmenlerine bilimsel yaratıcılık, yaratıcılık üzerine çeşitli hizmet içi eğitim verilebilir.
- 5) BÜT 3-5'in kuramsal yapısında temel alınan hipotez üretme, deney tasarlama gibi beceriler öğretmenler tarafından ölçme değerlendirme süreçlerinde kullanılabilir.
- 6) Fen bilimleri alanında yapılan farklılaşmaların veya eğitsel etkinliklerin belirlenmesinde öntest-sontest olarak kullanılabilir.

KAYNAKÇA

- Achter, J. A., Lubinski, D. and Benbow, C. P. (1996). Multipotentiality among the intellectually gifted: It was never there and already it's vanishing. *Journal of Counseling Psychology*, 43(1), 65.
- Ai, X. (1999). Creativity and academic achievement: An investigation of gender differences. *Creativity Research Journal*, 12(4), 329-337.
- Ajzen, I. and Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Akbulut, Y. (2010). *Sosyal bilimlerde SPSS uygulamaları: Sık kullanılan istatistiksel analizler ve açıklamalı SPSS çözümleri*. İstanbul: İdeal Kültür Yayıncılık.
- Al-Qaisy, L. M., and Turki, J. (2011). Adolescent' s creativity, self-concept and Achievement Motivation. *British Journal of Arts and Social Sciences*, 2(2), 88-101.
- Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.). *Handbook of social psychology* (pp. 798-844). New York: Harper & Row. Andersen.
- Almeida, L. S., Prieto, L. P., Ferrando, M., Oliveira, E., Ferrándiz, C. (2008). Torrance Test of Creative Thinking: The question of its construct validity. *Thinking Skills and Creativity*, 3(1), 53-58.
- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43, 997-1013.
- Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, 45(2), 357-377.
- Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in Organizational Behavior*, 10, 123–167.
- Amabile, T. M. (2013). Componential theory of creativity. In E. H. Kessler (Ed.), *Encyclopedia of Management Theory* (pp. 134-139). Thousand Oaks, CA: Sage Publications, 2013.
- Amsel, E. and Brock, S. (1996). The development of evidence evaluation skills. *Cognitive Development*, 11, 523–550.

- Arrindell, W. A. and Van der Ende, J. (1985). An Empirical-Test of the Utility of the Observationsto-Variables Ratio in Factor and Components-Analysis. *Applied Psychological Measurement*, 9(2), 165-178.
- Ayas M. B. and Sak, U. (2008). Test of scientific creativity: It's development and psychometric properties. *4th International Conference on Intelligence and Creativity*'de sunulan bildiri. Münster, Germany.
- Ayas, B. (2010). *Bilimsel Üretkenlik Testi'nin 6. sınıflar düzeyinde psikometrik özelliklerinin incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi. Eskişehir: Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Ayas, M. B. (2013). Scoring creativity: A comparison of different methods. *13th International ECHA Conference*'ta sunulan bildiri, Münster.
- 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.
- Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal*, 4(1), 23-39.
- Baer, J. (1994). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal*, 7(1), 35-46.
- Baer, J. (1994). Why you shouldn't trust creativity tests. *Educational Leadership*, 51(4), 80-83.
- Baer, J. (1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior*, 30(3), 183-187.
- Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal*, 11(2), 173-177.
- Baer, J. (2016). *Domain specificity of creativity*. San Diego, CA: Academic Press.

- Baer, J. and McKool, S. (2009). Assessing creativity using the consensual assessment technique. In G. Christopher, S. Schreiner (Eds.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65-77). New York: Information Science Publishing.
- 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.
- Barbot, B., Lubart, T. I. and Besançon, M. (2016). Peaks, slumps, and bumps: Individual differences in the development of creativity in children and adolescents. *New directions for child and adolescent development*, 2016(151), 33-45.
- Batey, M. (2012). The measurement of creativity: From definitional consensus to the introduction of a new heuristic framework. *Creativity Research Journal*, 24(1), 55-65.
- Batey, M. and Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social, and General Psychology Monographs*, 132(4), 355-429.
- Bermejo, M. R., Ruiz-Melero, M. J., Esparza, J., Ferrando, M., Pons, R. (2016). Una Nueva Medida de la Creatividad Científica: El estudio de sus propiedades psicométricas. *Anales de Psicología*, 32(3), 652-661.
- Best, R., Floyd, R. G. and McNamara, D. S. (2004). Understanding the fourth-grade slump: Comprehension difficulties as a function of reader aptitudes and text genre. *85th Annual Meeting of the American Educational Research Association*'da sunulan bildiri. San Diego, CA. 09/08/2017 tarihinde <ftp://129.219.222.66/Publish/pdf/Bestetal.pdf> adresinden ulaşıldı.
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms*. London: Routledge
- Bournelli, P., Makri, A. and Mylonas, K. (2009). Motor creativity and self-concept. *Creativity Research Journal*, 21(1), 104-110.
- Böke, K. (2014). *Sosyal bilimlerde araştırma yöntemleri* (4. Baskı). İstanbul: Alfa Yayıncılık.

- Briggs, S. R. and Cheek, J.M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106–48.
- Bullock, M. (1991). Scientific reasoning in elementary school: Developmental and individual differences. *Biennial Meeting of the Society for Research in Child Development*'da sunulan bildiri. Seattle.
- Büyüköztürk, Ş. (2017). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., Demirel, F. (2010). *Bilimsel araştırma yöntemleri*. Ankara: Pegem.
- Carson, S., Peterson, J. B. and Higgins, D. M. (2005). Reliability, validity, and factor structure of the Creative Achievement Questionnaire. *Creativity Research Journal*, 17, 37–50.
- Chall, J. S. and Jacobs, V. A. (2003). The classic study on poor children's fourth-grade slump. *American educator*, 27(1), 14-15.
- Chen, C., Kasof, J., Himsel, A. J., Greenberger, E., Dong, Q. and Xue, G. (2002). Creativity in drawings of geometric shapes: A cross-cultural examination with the consensual assessment technique. *Journal of Cross-Cultural Psychology*, 33(2), 171-187.
- 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.
- Clark, L. A. and Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309.
- Clark, L. A. and Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological assessment*, 7(3), 309- 3019.
- Cohen, J. (1977). *Statistical power for the behavioral sciences* (revised edition). Orlando, FL: Academic Press.
- Comrey, A. L. and Lee, H. B. (1992). *A first course in factor analysis* (2nd. ed.). Hillsdale, NJ: Erlbaum.

- Conti, R. and Amabile, T. (1999). Motivation/Drive. In M. A. Runco and S. R. Pritzker (Eds), *Encyclopedia of creativity* (pp. 251-259). San Diego, California: Academic Press.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*, 78(1), 98-104.
- Cresswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed approaches*. Thousand Oaks, CA: Sage.
- Cropley, A. (1999). Definitions of creativity. In S. R. Pritzker and M. A. Runco (Eds.), *Encyclopedia of creativity* (Vol. 2) (pp. 511–524). San Diego, CA: Academic Press.
- Cropley, D. and Cropley, A. (2005). Engineering creativity. In J. C. Kaufmann and J. Baer (Eds), *Creativity across domains: Faces of the muse* (pp. 169-185). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Csikszentmihalyi, M. (1997). *Flow and the psychology of discovery and invention*. New York: Harper Perennial.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Pegem Akademi.
- Davis, G. A. (1997). Identifying creative students and measuring creativity. In N. Colangelo and G. A. Davis (eds.), *Handbook of gifted education* (pp.245-266). Needham Heights, MA: A Viacom Company.
- Deniş, H. and Balım, A. G. (2012). Bilimsel yaratıcılık ölçeğinin Türkçeye uyarlama süreci ve değerlendirme ölçütleri. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 5(2), 1-21.
- Deville, R. F. (2003). *Scale development: Theory and applications*. London: Sage Publication.
- Domino, G. and Domino, M. L. (2006). *Psychological testing: An introduction*. Cambridge University Press.
- Duke, F. R. (1972). Creativity in science. *Journal of Chemistry Education*, 49(6), 382-384.

- Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17(3), 397-434.
- Dunbar, K. (1999). Science. In M. A. Runco and S. R. Pritzker (Eds), *Encyclopedia of creativity* (pp. 525-531). San Diego, CA: Academic Press.
- Ebert, E. S. (1994). The cognitive spiral: Creative thinking and cognitive processing. *The Journal of Creative Behavior*, 28(4), 275-290.
- Feist, G. J. (1997). Quantity, quality, and depth of research as influences on scientific eminence: Is quantity most important? *Creativity Research Journal*, 10(4), 325-335.
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4), 290-309.
- Feldman, D.H. (1999). The development of creativity. In R.J. Sternberg (Ed.), *Handbook of creativity* (pp. 169-186). New York: Cambridge University Press.
- Feldt, L. S., Woodruff, D. J. and Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103.
- Field, A. (2009). *Discovering statistics using SPSS*. London: SAGE.
- 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(1), 3-55.
- Fleith, D. D. S., Renzulli, J. S. and Westberg, K. L. (2002). Effects of a creativity training program on divergent thinking abilities and self-concept in monolingual and bilingual classrooms. *Creativity research journal*, 14(3-4), 373-386.
- Floyd, F. J. and Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, 7(3), 286-299.
- Fraenkel, J. R. and Wallen, N. E. (2009). *How to design and evaluate research in education* (7th ed.). McGraw, Boston: Hill Higher Education.
- Fraenkel, J. R., Wallen, N. E. and Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). New York: McGraw-Hill.

- Frederiksen, N. and Ward, W. C. (1978). Measures for the study of creativity in scientific problem solving. *Applied Psychological Measurement*, 2(1), 1-24.
- Gajda, A., Karwowski, M. and Beghetto, R. A. (2017). Creativity and academic achievement: A meta-analysis. *Journal of Educational Psychology*, 109(2), 269-299.
- 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. (1993). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gardner, P. L. (1975). Attitudes to science: A review. *Studies in Science Education*, 4, 1-41.
- Gardner, P. L. (1995). Measuring attitudes to science. *Research in Science Education*, 25, 283-289.
- Gaut, B. (2010). The philosophy of creativity. *Philosophy Compass*, 5(12), 1034-1046.
- Germann, P. J. (1988) Development of the attitude toward science in school assessment and its use to investigate the relationship between science achievement and attitude toward science in school. *Journal of Research in Science Teaching*, 25(8), 680-703.
- Gliem, J. A. and Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. *Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education*'da sunulan bildiri. Columbus, OH: The Ohio State University, pp. 81-88.
- Goodwin, L. D. and Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *The Journal of Experimental Education*, 74(3), 249-266.
- Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of personality and social psychology*, 37(8), 1398.
- Guilford, J. P. (1950) Creativity. *American Psychologist*, 5, 444-454.
- Guilford, J. P. (1956) The structure of intellect. *Psychological Bulletin*, 53, 267-293.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

- Güçyeter, Ş. (2013). *Matematiksel yeteneği tanılama modeli*. Yayınlanmamış Doktora Tezi. İstanbul: İstanbul Üniversitesi, Eğitim Bilimler Enstitüsü.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. (2010). *Multivariate data analysis* (7th Ed.). Harlow: Pearson Education Limited.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice hall.
- Heller, K. A. (2007). Scientific ability and creativity. *High Ability Studies*, 18(2), 209-234.
- Henson, R. K. and Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416.
- Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, 49(3), 234-244.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967-988.
- 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. 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. Reynolds (Eds.), *Handbook of creativity* (pp. 53-75). New York: Springer.
- Hogan, T. P. (2002). *Psychological testing: A practical introduction*. New Jersey: Wiley Global Education.
- Houtz, J. C. and Krug, D. (1995). Assessment of creativity: Resolving a mid-life crisis. *Educational Psychology Review*, 7(3), 269-300.
- Hu, W. and Adey, P. (2002). A test of scientific creativity for secondary school students. *International Journal of Science Education*, 24(4), 389-403.

- Hu, W. P. and Han, Q. (2006). A developmental study of creative scientific problem finding in elementary students. *Psychological Science*, 29, 944–946.
- Hu, W., Adey, P., Shen, J., Lin, C. (2004). The comparisons of the development of creativity between English and Chinese adolescents. *Acta Psychologica Sinica*, 36(6), 718–731.
- Hu, W., Shi, Q. Z., Han, Q., Wang, X., Adey, P. (2010). Creative scientific problem finding and its developmental trend. *Creativity Research Journal*, 22(1), 46-52.
- Huck, S. W. (2015). *Statistical misconceptions*. New York: Routledge.
- Innamorato, G. (1998). Creativity in the development of scientific giftedness: Educational implications. *Roeper Review*, 21(1), 54-60.
- Kadayıfçı, H. (2008) *Yaratıcı düşünmeye dayalı öğretim modelinin öğrencilerin maddelerin ayrılması ile ilgili kavramları anlamalarına ve bilimsel yaratıcılıklarına etkisi*. Yayınlanmamış Doktora Tezi. Ankara: Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Kanlı, E. (2014). *Yaratıcı bilimsel çağrışımlar testinin geliştirilmesi ve psikometrik özelliklerinin incelenmesi*. Yayınlanmamış Doktora Tezi. İstanbul: İstanbul Üniversitesi, Eğitim Bilimler Enstitüsü
- Kaplan, R. and Saccuzzo, D. (2012). *Psychological testing: Principles, applications, and issues*. Belmont CA: Cengage Learning.
- Kaufman, J. C. and Baer, J. (2004). Sure, I'm creative but not in mathematics! Self-reported creativity in diverse domains. *Empirical Studies of the Arts*, 22(2), 143-155.
- Kaufman, J. C. and Baer, J. (2005). The amusement park theory of creativity. In J.C. Kaufman and J. Baer (Eds.), *Creativity across domains: Faces of the muse* (pp. 321-328). New Jersey: Lawrence Earlbaum
- Kaufman, J. C. and Sternberg, R. J. (2007). Resource review: Creativity. *Change*, 39(4), 55-58.
- Kaufman, J. C., Baer, J. and Gentile, C. A. (2004). Differences in gender and ethnicity as measured by ratings of three writing tasks. *The Journal of Creative Behavior*, 38(1), 56-69.

- Kaufman, J. C., Baer, J., Cole, J. C., 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., Plucker, J. A. and Baer, J. (2008). *Essentials of creativity assessments*. Hoboken, New Jersey: John Wiley & Sons.
- Kaufman, J.C. and Sternberg, R.J. (2010). *The Cambridge handbook of creativity*. New York: Cambridge University Press.
- Kılıç, B. (2011). *İlköğretim sekizinci sınıf öğrencilerinin bilimsel yaratıcılık ve bilimsel tutum düzeylerinin belirlenmesi*. Yayınlanmamış Yüksek Lisans Tezi. Eskişehir: Osmangazi Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal*, 18(1), 3-14.
- 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. and Simon, H. A. (1999). Studies of scientific discovery: Complementary approaches and convergent findings. *Psychological Bulletin*, 125(5), 524.
- Klahr, D., Fay, A. L. and Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive psychology*, 25(1), 111-146.
- Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.
- Kline, R. B. (2010). *Principles and practice of structural equation modeling*. New York: Guilford publications.
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage.
- Kuhn, D. (1989). Children and adults as cognitive scientists. *Psychological Review*, 96(4), 674-689.
- Kuhn, D. (1989). Children and adults as cognitive scientists. *Psychological Review*, 96(4), 674-689.

- Kuhn, D. and Angelev, J. (1976). An experimental study of the development of formal operational thought. *Child Development*, 47(3), 697-706.
- Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., Yotive, W. (1988). *The development of scientific thinking skills*. San Diego, CA: Academic Press.
- Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., 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.
- Liang, J.C. (2002). *Exploring scientific creativity of eleventh grade students in Taiwan*. Unpublished Doctoral Dissertation, Austin, TX: University of Texas.
- Lubart, T. and Guignard, J. H. (2004). The generality-specificity of creativity: A multivariate approach. In R. J. Sternberg, E. Grigorenko and J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 43–56). Washington: A.P.A.
- Lubart, T. I. (1999). Componential models. *Encyclopedia of creativity* (pp. 295-300). New York: Academic Press.
- Lubinski, D., Benbow, C. P. and Ryan, J. (1995). Stability of vocational interests among the intellectually gifted from adolescence to adulthood: A 15-year longitudinal study. *Journal of applied Psychology*, 80(1), 196.
- MacCallum, R. C., Browne, M. W. and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, 1(2), 130-149.
- Marjoribanks, K. (1976) Academic achievement, intelligence, and creativity: A regression surface analysis. *Multivariate Behavioural Research*, 11(1), 105-118
- Marjoribanks, K. (1976). Academic achievement, intelligence, and creativity: A regression surface analysis. *Multivariate Behavioral Research*, 11(1), 105-118.
- Matsunaga, M. (2010). How to factor-analyse your data right: Do's, dont's, and how-to's. *International Journal of Psychological Research*, 3(1), 97–110.

- Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, 69(3), 220-232.
- Meng, X. L., Rosenthal, R. and Rubin, D. B. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111(1), 172.
- Micheal, W. (1999). Guilford's view. In M.A. Runco and S. Pritzer (Eds.), *Encyclopedia of creativity* (pp. 785-797). New York: Academic Press.
- Mohamed, A. H. (2006). *Investigating the scientific creativity of fifth-grade students*. Unpublished Doctoral Dissertation, Arizona: University of Arizona.
- Mouchiroud, C. and Lubart, T. (2001). Children's original thinking: An empirical examination of alternative measures derived from divergent thinking tasks. *The Journal of Genetic Psychology*, 162(4), 882-401.
- Murphy, K. R. and Davidshofer, C. O. (1988). *Psychological testing: Principles, and applications*. New Jersey: Pearson Englewood Cliffs.
- Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. New Jersey: Prentice-Hall.
- Niu, W. and Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *International Journal of Psychology*, 36, 225–241.
- Nusbaum, E. C. and Silvia, P. J. (2011). Are intelligence and creativity really so different? Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, 39(1), 36-45.
- Osborne, J., Simon, S. and Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9) 1049–1079.
- Pallant, J. (2011). *Survival manual: A step by step guide to data analysis using SPSS*. Crows Nest: Allen&Unwin.
- Parkhurst, H. B. (1999). Confusion, lack of consensus, and the definition of creativity as a construct. *The Journal of Creative Behavior*, 33(1), 1-21.
- Pekmez, E. S., Aktamis, H. and Taskin, B. C. (2009). Exploring scientific creativity of 7 th grade students. *Journal of Qafqaz University*, 26, 204-214.

- Pell, T. and Jarvis, T. (2001). Developing attitude to science scales for use with children of ages from 5 to 11. *International Journal of Science Education*, 23(8) 847-862.
- 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. and Makel, M. C. (2010). Assessment of creativity. In J.C. Kaufman and R.J. Sternberg (Eds.), *The Cambridge handbook of creativity*, (pp. 48-73). New York: Cambridge University Press.
- Plucker, J. and Zabelina, D. (2009). Creativity and interdisciplinary: One creativity or many creativities? Humanities. *ZDM: The International Journal on Mathematics Education*, 41(1-2), 5-11.
- Plucker, J., Beghetto, R. A. and Dow, G. (2004). Why isn't creativity more important to educational psychologists? Potential, pitfalls, and future directions in creativity research. *Educational Psychologist*, 39(2), 83-96.
- Puccio, G. J. (1991). William Duff's eighteenth-century examination of original genius and its relationship to contemporary creativity research. *Journal of Creative Behaviour*, 25(1), 1-10.
- Raubenheimer, J. (2004). An item selection procedure to maximize scale reliability and validity. *SA Journal of Industrial Psychology*, 30(4), 59-64.
- Reid, N. (2006). Thoughts on attitude measurement. *Research in Science & Technological Education*, 24(1), 3-27.
- Renzulli, J. S. (2002). Emerging conceptions of giftedness: Building a bridge to the new century. *Exceptionality*, 10(2), 67-75.
- Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan*, 47(2), 305-310.
- Roe, A. (1952). A psychologist examines 64 eminent scientists. *Scientific American*, 185(5), 21-25.
- Roe, A. (1961). The psychology of a scientist. *Science*, 134(3477), 56-459.
- Rothenberg, A. (1971). The process of janusian thinking in creativity. *Archives of General Psychiatry*, 24(3), 195-205.

- Rothenberg, A. (1987). Einstein, Bohr, and creative thinking in science. *History of science*, 25(2), 147-166.
- Rothenberg, A. (1995). Creative cognitive processes in Kekule' s discovery of the structure of the benzene molecule. *The American Journal of psychology*, 108(3), 419-438.
- Rothenberg, A. (1996). The janusian process in scientific creativity. *Creativity Research Journal*, 9(2-3), 207-209.
- Ruffman, T., Perner, J., Olson, D. R., Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis- evidence relation. *Child Development*, 64(6), 1617-1636.
- Ruffman, T., Perner, J., Olson, D. R., Doherty, M. (1993). Reflecting on scientific thinking: Children's understanding of the hypothesis- evidence relation. *Child Development*, 64(6), 1617-1636.
- 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*, 19(3), 177-190.
- Runco, M. A. (1990). The divergent thinking of young children: implications of the research. *Gifted Child Today Magazine*, 13(4), 37-39.
- Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, 55, 657–687.
- Runco, M. A. and Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92-96.
- Runco, M. A., McCarthy, K. A. and Svenson, E. (1994). Judgments of the creativity of artwork from students and professional artists. *The Journal of Psychology*, 128(1), 23-31.
- Sak, U. (2013). *Üstün zekâlılar özellikleri tanılanmaları eğitimleri*. Ankara: Vize
- Sak, U. (2014). *Yaratıcılık: Gelişimi ve geliştirilmesi*. Ankara: Vize.
- 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), 316–329.

- Sak, U. and Maker, C. J. (2006). Developmental variation in children's creative mathematical thinking as a function of schooling, age, and knowledge. *Creativity Research Journal*, 18(3), 279-291.
- Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45(1), 1-32.
- Sanacore, J. and Palumbo, A. (2008). Understanding the fourth-grade slump: Our point of view. *The Educational Forum*, 73, 67-74.
- Saurenman, D. A. and Michael, W. B. (1980). Differential placement of high-achieving and low-achieving gifted pupils in grades four, five, and six on measures of field dependence-field independence, creativity, and self-concept. *Gifted Child Quarterly*, 24(2), 81-86.
- Sawyer, R. K. (2006). *Explaining creativity; The science of human innovation*. New York: Oxford University Press.
- Schempp, P. G., Cheffers, J. T. and Zaichkowsky, L. D. (1983). Influence of decision-making on attitudes, creativity, motor skills and self-concept in elementary children. *Research Quarterly for Exercise and Sport*, 54(2), 183-189.
- Shah, I. (2004) *Making university laboratory work in chemistry more effective*. Unpublished PhD thesis, University of Glasgow.
- Shaklee, H. and Paszek, D. (1985). Covariation judgment: Systematic rule use in middle childhood. *Child Development*, 56(5), 1229-1240.
- Shlain, L. (1999). Leonardo da Vinci. In M. A. Runco and S. R. Pritzker (Eds)., *Encyclopedia of creativity* (pp. 501-509). San Diego, California: Academic Press.
- 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.
- Simon, H. A., (1977). *Models of discovery: And other topics in the methods of science*. Dordrecht: D. Reidel Pub. Co.
- Simonton, D. K. (1983). Formal education, eminence and dogmatism: The curvilinear relationship. *The Journal of Creative Behavior*, 17(3), 149-162.

- Simonton, D. K. (1984). Creative productivity and age: A mathematical model based on a two-step cognitive process. *Developmental Review, 4*, 77-111.
- Simonton, D. K. (1988). *Scientific genius: A psychology of science*. Cambridge: Cambridge University Press.
- Snyder, A., Mitchell, J., Bossomaier, T., 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*, 753-766.
- Solano-Flores, G. (1993). Item structural properties as predictors of item difficulty and item association. *Educational and Psychological Measurement, 53*(1), 19-31.
- Spaulding, R. L. (1965). Achievement, Creativity, and Self-Concept Correlates of Teacher-Pupil Transactions in Elementary School Classrooms. State College, PA: Ed & Psych Associates.
- Starko, A. J. (2005). *Creativity in the classroom: Schools of curious delight*. Mahwah, NJ: Lawrence Erlbaum.
- Sternberg, R. J. and Lubart, T. I. (1992). Buy low and sell high: An investment approach to creativity. *Current Directions in Psychological Science, 1*-5.
- Sternberg, R. J. and Lubart, T. I. (1991). An investment theory of creativity and its development. *Human development, 34*(1), 1-31.
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Hillsdale, NS: Erlbaum.
- Subotnik, R. F. (1993). Adult manifestations of adolescent talent in science. *Roeper Review, 15*(3), 164-169.
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk psikoloji yazıları, 3*(6), 49-74.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tabachnick, B. G. and Fidell, L. S. (2012). *Using multivariate statistics*. NJ: Pearson.

- 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.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- Tokmak, F. (2016). *Özel eğitim gruplamasının üstün yetenekli öğrencilerin akademik benlik algıları üzerindeki etkisi*. Yayınlanmamış Yüksek Lisans tezi. Eskişehir: Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Torrance, E. P. (1962). Guiding creative talent. Englewood Cliffs, NJ: Prentice-Hall.
- Torrance, E. P. (1968). A longitudinal examination of the fourth grade slump in creativity. *Gifted Child Quarterly*, 12(4), 195-199.
- Torrance, E. P. (1988). The nature of creativity as manifest in its testing. In R. J. Sternberg (Ed.), *The nature of creativity* (pp. 43-75). New York: Cambridge Univ. Press.
- Torrance, E. P. (1992). A national climate for creativity and invention. *Gifted Child Today*, 15(1), 10-14.
- Torrance, E. P. (2004). Predicting the creativity of elementary school children (1958–80) and the teacher who made a difference. In D. Treffinger (Ed.) *Creativity and giftedness* (pp. 35-49). California: Corwin Press
- Tosun, C. and Genç, M., (2015). Adaptation of science attitude scale developed for primary school students to Turkish: Validity and reliability studies. *Elementary Education Online*, 14(3), 946-960
- Trochim, W. M. and Donnelly, J. P. (2006). *The research methods knowledge base* (3rd ed.). Cincinnati, OH: Atomic Dog.
- Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development*, 51(1), 1-10.
- Urbina, S. (2004). *Essentials of psychological testing*. Hakoben, 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.
- VanTassel-Baska, J. (1998). Creativity. In VanTassel-Baska (Ed.), *Excellence in educating gifted and talented learners*. Denver: Love Publishing Company

- Velicer, W. F. and Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25(1), 1-28.
- Watkins, M.W. (2004). *Monte Carlo PCA for parallel analysis* [computer software].
- Weinburgh, M. (1995) Gender differences in student attitudes toward science: A metanalysis of the literature from 1970 to 1991. *Journal of Research in Science teaching*, 32(4), 387-398.
- Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (ed.), *Handbook of creativity* (pp.226-250). New York, NY: Cambridge University Press.
- Weisberg, R. W. (1999). Creativity and knowledge: A challenge to theories. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp.226-250). New York, NY: Cambridge University Press.
- Weisberg, R. W. (2006). *Creativity: Understanding innovation in problem solving, science, invention and the arts*. New Jersey: John Wiley & Sons.
- Widaman, K. F. (1993). Common factor-analysis versus principal component analysis: Differential bias in representing model parameters. *Multivariate Behavioral Research*, 28(3), 263-311.
- Wilcox, R. R. (2016). Comparing dependent robust correlations. *British Journal of Mathematical and Statistical Psychology*, 69(3), 215-224.
- 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.
- Yager, R. E. and McCormack, A. J. (1989). Assessing teaching/learning successes in multiple domains of science and science education. *Science Education*, 73(1), 45-58.
- Yang, K., Lin, S., Hong, Z., Lin, H. (2016). Exploring the assessment of and relationship between elementary students' scientific creativity and science inquiry. *Creativity Research Journal*, 28(1), 16-23.

Yıldırım, A. ve Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.

Zimmerman, C. (2000). The development of scientific thinking skills. *Developmental Review*, 20(2), 172-223.

Zimmerman, C. (2000). The development of scientific thinking skills. *Developmental Review*, 20(2), 172-223.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27(1), 99-149.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27(1), 99-149.

EKLER

EK-1. Kurul Kararı

Kayıt Tarihi: 23.12.2016

Protokol No: 137866



ANADOLU ÜNİVERSİTESİ ETİK KURULU KARARI

ÇALIŞMANIN TÜRÜ:	BAP Projesi-Doktora Tez Çalışması
KONU:	Eğitim Bilimleri
BAŞLIK:	Bilimsel Üretkenlik Testinin 3, 4 ve 5. Sınıf Öğrencilerine Uygun Formunun Geliştirilmesi ve Türkiye Normunun Çıkarılması
PROJE/TEZ YÜRÜTÜCÜSÜ:	Prof. Dr. Uğur SAK
TEZ YAZARI:	M. Bahadır AYAS
ALT KOMİSYON GÖRÜŞÜ:	-
KARAR:	Olumlu

ETİK KURUL ÜYELERİ

Prof. Dr. Aydın AYBAR
Rektör Yardımcısı / Etik Kurul Başkanı

Prof. Dr. Hayrettin TÜRK
Fen Bil. (Fen Fak.)

Prof. Dr. Yusuf ÖZTÜRK
Sağlık Bil. (Ecz. Fak.)

Prof. Dr. Esra CEYHAN
Eğitim Bil. (Eğitim Bil. Ens.)

Prof. Dr. Bülent GÜNŞOY
Sos. Bil. (İkt. Fak.)

Prof. Dr. Münevver ÇAKI
Güz. San. (Güz. San. Fak.)

İMZA/ TARİH

23.01.2017

(Handwritten signatures of Prof. Dr. Aydın Aybar, Prof. Dr. Hayrettin Türk, Prof. Dr. Yusuf Öztürk, Prof. Dr. Esra Ceyhan, Prof. Dr. Bülent Günsoy, and Prof. Dr. Münevver Çaki)

(Handwritten signature of Prof. Dr. Münevver Çaki)

(Handwritten signature of Prof. Dr. Bülent Günsoy)

(Handwritten signature of Prof. Dr. Münevver Çaki)

EK-2. İL Milli Eğitim Müdürlüğü Uygulama İzni

T.C
ESKİŞEHİR VALİLİĞİ
İl Milli Eğitim Müdürlüğü

ARAŞTIRMA DEĞERLENDİRME FORMU

ARAŞTIRMA SAHİBİNİN	
Adı Soyadı	M. Bahadır AYAS
Kurumu/Üniversitesi	Anadolu Üniversitesi
Araştırma Yapılacak Eğitim Kurumu ve Kademesi	İl Merkezindeki İlkokul ve Ortaokullar
Araştırmanın Konusu	Bilimsel Üretkenlik Testinin 3., 4. ve 5. Sınıf Öğrencilerine Uygun Formunun Geliştirilmesi ve Türkiye Normunun Çıkarılması
Üniversite / Kurum Onayı	Var
Araştırma/Proje/Ödev/ Tez Önerisi	Var
Veri Toplama Araçları	Bilimsel Üretkenlik Testi
Görüş İstenecek Birimler	-
KOMİSYON GÖRÜŞÜ	
Milli Eğitim Bakanlığı Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü'nün 2012/13 sayılı genelgesi gereğince 2016-2017 öğretim yılında uygulanmasında sakınca yoktur.	
Komisyon Kararı	KABUL (Oybirliği ile)
Muhalef Üyenin Adı ve Soyadı	Gerekçesi :

KOMİSYON

27/02/2017
Komisyon Başkanı
Barış HANCI
Milli Eğitim Müdür Yardımcısı

Üye
Ömer GARAN
Öğretmen

Üye
Dr. Seda AKKAYA
Baş Öğretmen

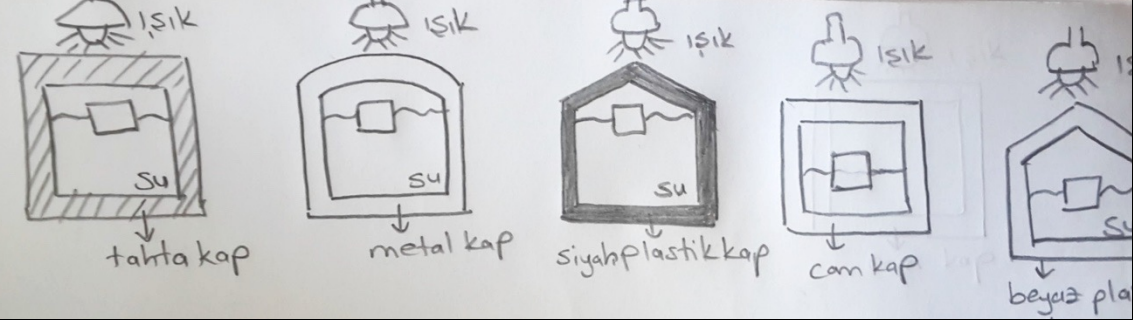
Üye
E. Senay KUTLU
Öğretmen

EK-3a. BÜT Form 3-5 Madde Geliştirme Formu

Madde (görsel ekleyiniz)*	
Yönerge (soru/görev metnini yazınız)	
Yanıtlar (Üretilbilecek başlıca yanıtlar) <ul style="list-style-type: none">••••••	
Hazırlayan:	Seviye (sınıf seviyesi 1,2,3. sınıf) :
Alan (fizik, kimya, biyoloji vb.) :	Görev (hipotez geliştirme, deney tasarlama) :
Bilgi (bu soruya cevap üretilbilmesi için gerekli bilgi/kavramları yazınız) :	

EK-3b. Örnek Madde

Madde (görsel ekleyiniz)*



Yönerge (soru/görev metnini yazınız)

Bir araştırmacı farklı malzemelerden yapılmış kapları kullanarak yukarıdaki deney düzeneklerini hazırlıyor. Kapların içerisine su ve özdeş buzlar koyarak, özdeş lambalar altında beklemeye bırakıyor. Araştırmacının bu düzeneği hazırlamaktaki amaçları neler olabilir?

Yanıtlar (Üretilebilecek başlıca yanıtlar)

- Su miktarlarının buzların erime süresi üzerine etkisini belirlemek
- Kullanılan kapların yapıldığı malzemenin buzların erime süresi üzerine etkisini belirlemek
- Kapların yüzey alanlarının buzların erime süresi üzerine etkisini belirlemek
- Kapların renginin buzların erime süresi üzerine etkisini belirlemek
- Kapların şeklinin buzların erime süresi üzerine etkisini belirlemek
-

Hazırlayan: B...a A...z İ...i

Seviye (sınıf seviyesi 1,2,3. sınıf) : 3

Alan (fizik, kimya, biyoloji vb.) : Kimya

Görev (hipotez geliştirme, deney tasarlama) :





























































1. Tip Hipotez geliştirme

Bilgi (bu soruya cevap üretilmesi için gerekli bilgi/kavramları yazınız) :

Madde, Isı Alışverişi

EK-4. Bilime Yönelik Tutum Ölçeği (BYTÖ) ve Fen Dersi Akademik Benlik Algısı Ölçeği (FBAÖ).

Fen bilgisi dersi hakkındaki düşüncelerinizi merak ediyoruz. Her ifadeyi dikkatlice okuyunuz ve size en uygun gelen seçeneği işaretleyiniz. Katılımınız için teşekkürler.

1. Fen bilgisi dersini severim					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
2. Fen bilgisi dersini diğer derslerden daha çok severim					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
3. Fen bilgisi dersinin daha çok yapılmasını isterim					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
4. Bilimle ilgili belgeseller izlemekten hoşlanırım					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
5. Bilimsel gelişmelerle ilgili haberleri okumaktan ve izlemekten hoşlanırım					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
6. Bilimle ilgili kitap ve dergileri okumaktan hoşlanırım					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
7. Fen bilimleri dersinde iyi notlar alırım					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
8. Fen bilimleri dersinin konularını çabuk öğrenirim					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
9. Fen bilimleri dersinde yetersizim					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
10. Yaşıtlarıma göre fen bilimleri dersinde iyiyimdir					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
11. Bugüne kadar fen bilimleri dersinde her zaman başarılı olmuşumdur					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
12. Fen bilimleri dersindeki konular benim için kolaydır					
Kesinlikle Katılmıyorum		Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum

ÖZGEÇMİŞ

Adı-Soyadı : M. Bahadır AYAS
Yabancı Dil : İngilizce
Doğum Yeri ve Yılı : Kadirli / 06.02.1978
E-Posta : mbayas@anadolu.edu.tr

Eğitim ve Mesleki Gelişim:

- 2010, Anadolu Üniversitesi, Özel Eğitim Bölümü, Üstün Zekalıların Eğitimi Ana Bilim Dalı Yüksek Lisans Programı
- 2001, Atatürk Üniversitesi, Fizik Öğretmenliği Lisans Programı
- 2002-2011, Öğretmen, Millî Eğitim Bakanlığı
- 2011-Devam ediyor, Araştırma Görevlisi, Anadolu Üniversitesi Eğitim Fakültesi

Makaleler:

- Sak, U., Ayas, M. B., Sezerel, B. B., Öpengin, E., Özdemir, N. N., Gürbüz, S. D. (2015). Gifted and Talented Education in Turkey: Critics and Prospects/Türkiye'de Üstün Yeteneklilerin Eğitiminin Elestirel Bir Değerlendirmesi. *Türk Üstün Zekâ ve Eğitim Dergisi*, 5(2), 110-132.
- 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.
- 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), 315-328.

Uluslararası Bilimsel Toplantılarda Sunulan Bildiriler:

- Ayas, M. B. (2017). Anadolu Sak zeka ölçeği: Geliştirilme süreci ve psikometrik özellikleri. 1st International Education research and teacher Education Congress, sözlü bildiri, 14-16 Eylül, Uşak, Turkey.
- Ayas, M. B. (2017). Perfectionism levels of gifted and talented students. 15th International Conference on Excellence in Education (ICIE), sözlü bildiri, 3-5 July, Lisboa, Portugal.
- Yılmaz, G. and Ayas, M. B. (2017). Examination of threshold theory with a new measure of intelligence. 15th International Conference on Excellence in Education (ICIE), sözlü bildiri, 3-5 July, Lisboa, Portugal.
- Ayas, M. B. (2016). Perceptions of the teachers of the gifted students on creativity. 14th Asia Pacific Conference on Giftedness, sözlü bildiri, 14-16 July, Macau, China.
- Ayas, M. B. (2016). Threshold theory in the area of science: creative potential of school children. 13th International Conference on Excellence in Education (ICIE), sözlü bildiri, 18-21 May, Rijeka, Croatia.

- Ayas, M. B. (2015). First European talent point in Turkey: EPTS. In proceedings of First International Conference for Teachers: Alternative Methods of Teaching, tam metin bildiri, 26-11.2015, Ljubljana, Slovenia (pp. 8-14). ISBN: 978-961-93-89-2-2.
- Sak, U. and Ayas, M. B. (2015). Research findings about creative scientific ability test., sözlü bildiri, 12th International Conference on Excellence in Education (ICIE), sözlü bildiri, 1-4 July, Krakow, Poland.
- Ayas, M. B., Opengin, E. and Tokmak, F. (2015). Blind variation view of creativity Scientific creativity of 6th graders. 21st World Council for Gifted and Talented Children, poster bildiri, 10-14 August, Odense, Denmark.
- Öpengin, E., Ayas, M. B. and Tokmak, F. (2015). Evaluation of EPTS identification model according to some socio demographic variables. 21st World Council for Gifted and Talented Children, sözlü bildiri, 10-14 August, Odense, Denmark.
- Tokmak, F., Opengin, E. and Ayas, M. B. (2015). Characteristics of Leaders of top 100 Universities in the World. 21st World Council for Gifted and Talented Children, sözlü bildiri, 10-14 August, Odense, Denmark.
- Ayas, M. B. (2014). Psychometric validity of the creative scientific ability test. 14th International ECHA Conference, sözlü bildiri, 17-20 September, Ljubljana, Slovenia.
- Ayas, M. B. (2014). Early identification and its effects in gifted children. International Conference on Early Childhood Intervention (ICECI2014), sözlü bildiri, 3-6 April, Antalya, Turkey.
- Ayas, M. B. (2014). Creativity as a subject area in Turkey: An analysis of dissertations, 11th International Conference on Excellence in Education (ICIE), sözlü bildiri, 7-10 July, Paris, France.
- Ayas, M. B. (2014). External validity of ASIS: Correlational studies with other measures of intelligence. 15th International Conference on Excellence in Education (ICIE), sözlü bildiri, 3-5 July, Lisboa, Portugal.
- Ayas, M. B. (2013). Scoring creativity: A comparison of different methods. 13th International ECHA Conference, sözlü bildiri, 12-15 September, Münster, Germany.
- Ayas, M.B. and Özdemir, N. N. (2013). ÜYEP fen bilimleri dersinin değerlendirilmesi. 3rd International Conference on Talent Development and Excellence, sözlü bildiri, 25-28 Eylül 2013, Antalya, Türkiye.
- Ayas, M. B. (2012). Assessment of scientific creativity. The International Conference on Talent Development and Excellence, sözlü bildiri, 19-11 April, Beijing, China.
- Özdemir N. N. and Ayas M. B., (2012). Gender differences in Creative Scientific Ability. 13th International ECHA Conference on Giftedness, sözlü bildiri, 12-15 September, Münster, Germany.
- Ayas, M. B. (2011). The investigation of the psychometric properties of creative scientific ability test at 6th graders. International Conference on Excellence in Education 2011 (ICIE2011), sözlü bildiri, 6-9 July, Istanbul, Turkey.

Ayas, M. B., Sak, U. and Demirel, S. (2008). Test of scientific creativity. 4th International Conference on Intelligence and Creativity (ICIC), sözlü bildiri, 9-11 october, Münster, Germany.

Ulusal Bilimsel Toplantılarda Sunulan Bildiriler:

Ayas, M. B. (2014). Türkiye’de üstün zekalıların eğitimi: yapılanlar, yapılmayanlar, yapılması gerekenler. 4. Ulusal Üstün Zekalı ve Yeteneklilerin Eğitimi Kongresi, sözlü bildiri, 22-25 Eylül, İstanbul, Türkiye

Ayas, M. B. (2013). ÜYEP fen bilimleri müfredatının değerlendirilmesi. 3. Türkiye Üstün yetenekliler Kongresi, sözlü bildiri, 14-16 Kasım, Ankara,

Ayas, M. B. (2012). Sınıfı içi performansın yaratıcı test performansı ile karşılaştırılması. 22. Ulusal özel Eğitim Kongresi, sözlü bildiri, 11-12 Ekim, Trabzon, Türkiye.

Ayas, M. B. and Sak, U. (2009). Bilimsel Üretkenlik Testi. Teorik alt yapısı, geliştirilme süreci ve psikometrik özellikleri. Üstün Yetenekli Çocuklar 2. Ulusal Kongresi, sözlü bildiri, 25-27 Mart, Eskişehir, Türkiye.

Kitap ve Kitap Bölümleri:

Sak, U., Bal Sezerel, B., Ayas, B., Tokmak, F., Özdemir, N. N., Demirel Gürbüz, Ş., Öpengin, E. (2016). Anadolu Sak Zeka Ölçeği (ASİS) uygulayıcı kitabı. Eskişehir: Anadolu Üniversitesi ÜYEP Merkezi.

Sak, U., Ayas, B., Bal-Sezerel, B., Öpengin, E., Özdemir, N. N., Demirel-Gürbüz, Ş. (2015). A critical assessment of the education for gifted and talented students in Turkey. In D. Dai & C. Kuo (Eds.), *Gifted education in Asia* (pp. 167–190). Charlotte, NC: Information Age Publishing.

Ayas, M. B. and Kirişçi, N. (2017). Özel yeteneklilerin özellikleri ve tanınması. İçinde M. A. Melekoğlu ve U. Sak (Eds.), *Öğrenme güçlüğü ve özel yetenek* (pp. 153-171). Ankara: Pegem.

Demirel, Ş. and Ayas, M. B. (2013). Üstün zekalılar ve eğitimleri. İçinde S. Vuran (Eds.), *Özel eğitim* (pp.363-397). Ankara: Maya.