

Fen ve Teknoloji Dersi Öğretmen Adaylarının Bilimsel Süreç Becerilerinin Ölçülmesine İlişkin Bir Test Geliştirme Çalışması*

Fethiye KARSLI¹, Alipaşa AYAS²

¹ Yrd. Doç. Dr., Giresun Üniversitesi, Eğitim Fakültesi, Giresun-TÜRKİYE

² Prof. Dr., Bilkent Üniversitesi, Eğitim Fakültesi, Ankara-TÜRKİYE

Alındı: 02.02.2012

Düzeltildi: 20.02.2012

Kabul Edildi: 01.06.2013

Orijinal Yayın Dili Türkçedir (v.10, n.2, Haziran 2013, ss.66-84)

ÖZET

Bu çalışmada, fen ve teknoloji dersi öğretmen adaylarının bilimsel süreç becerilerini ölçmeye yönelik, geçerliği ve güvenilirliği sağlanmış çoklu formda bilimsel süreç becerileri testi (BİSBET) geliştirmek amaçlanmıştır. Bu amaçla ölçülecek davranışın niteliği dikkate alınarak, son zamanlarda program geliştirmeciler tarafından da öngörülen ölçme-değerlendirme tekniklerine uygun, 25'i çoktan seçmeli ve 11'i açık uçlu yapıda olmak üzere toplam 36 maddeden oluşan test geliştirilmiştir. Toplam 197 fen ve teknoloji dersi öğretmen adayına uygulanan testin geçerlik, güvenilirlik çalışmaları ve madde analizleri yapılmıştır. Bilimsel süreç becerilerini ölçmeye yönelik geliştirilen testin kapsam geçerliğine kanıt sağlamak için uzman görüşlerine, yapı geçerliğine kanıt sağlamak için ise hipotez testi yöntemine başvurulmuştur. Testin güvenilirliği; çoktan seçmeli test maddeleri için iç tutarlılık analizi yöntemi ile açık uçlu test maddeleri için ise iç tutarlılık ve gözlemciler arası tutarlılık yöntemleri ile sağlanmıştır. Testin geçerlik, güvenilirlik ve madde analizi sonuçlarına göre BİSBET'in fen ve teknoloji dersi öğretmen adaylarının BSB'lerinin ölçülmesi amacıyla kullanılabilir, geçerliği ve güvenilirliği sağlanmış bir test olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: Bilimsel Süreç Becerileri; Test Geliştirme; Geçerlik; Güvenirlik.

GİRİŞ

Fen eğitiminde öğrencilere bilimsel içeriğin kazandırılması mı yoksa Bilimsel Süreç Becerileri (BSB)'nin kazandırılması mı daha önemlidir? sorusu eğitimciler tarafından tartışılmaktadır. Oysaki bilimsel içerik bilimsel bilgiyi bilmeyi; BSB ise bilim yapmayı işe koşar. Aslında birinin öğrenilmesi, bir diğerinin öğrenilmesine yardımcı olur (Ayas, Çepni, Johnson & Turgut, 1997). Bu yüzden hem bilimsel içerik hem de BSB fen programlarında eşit öneme sahiptir. Amerikan Ulusal Fen Eğitimi Standartları'nda, öğrencilerin öğrenmelerindeki

* Bu çalışma KTÜ BAP 2009.116.002.1 Kodlu Proje kapsamında desteklenmiştir



öncelikli amacın, fen içeriğindeki önemli bilgilerle birlikte bilimsel araştırma yöntemlerinin öğrenilmesi, bilgiye ulaşma yollarının bilinmesi olduğu belirtilmiştir. Türkiye de yeniden yapılandırılan fen ve teknoloji, fizik, kimya ve biyoloji dersi öğretim programlarında da öğrencilerde fen içeriğindeki bilgilerin öğretilmesinin yanı sıra BSB'nin kazandırılmasının da önemi üzerine vurgu yapılmaktadır. BSB'nin üzerine vurgu yapılması akıllara BSB'nin eğitimde kullanılmasının ve geliştirilmesinin fen eğitiminde etkisinin nasıl olduğu sorusunu getirmektedir. Yapılan araştırmalarda da BSB ile öğrencilerin akademik başarıları arasında (Beaumont Walters & Soyibo, 2001; Kesamang & Taiwo, 2002), BSB ile öğrencilerin fene yönelik tutumları arasında (Downing & Filer, 1999), BSB ile bilimsel yaratıcılık arasında (Aktamış & Ergin, 2007; Roberts, 2003) ve BSB ile öğretim programının niteliği arasında (Bağcı Kılıç, Haymana & Bozyılmaz, 2008) pozitif ilişkiler olduğu ortaya konulmuştur. BSB'nin eğitimdeki bu pozitif etkilerinden hareketle birçok araştırmacı öğretmen adaylarının ve öğretmenlerin BSB hakkındaki farkındalıklarını belirlemeye yönelik çalışmalar yapmışlar (Farsakoğlu, Şahin, Karşlı, Akpınar & Ültay, 2008; İnan, 2010; Karşlı, Şahin ve Ayas, 2009) ve öğrencilere BSB'yi kazandırmaya yönelik etkinlikler geliştirmişlerdir (Karşlı, 2011; Duran & Özdemir, 2010; Karşlı & Şahin, 2009; Tifi, Natale & Lombardi, 2006). Farsakoğlu ve diğerlerinin (2008) 4. sınıf fen ve teknoloji dersi öğretmen adaylarının BSB farkındalıklarını belirlemek amacıyla yaptıkları araştırma sonucundan öğretmen adaylarının BSB'yi teorik olarak açıklayabilmelerine rağmen bu teorik bilgileri pratiğe dökemedikleri ortaya çıkmıştır. Bu bulgu Karşlı ve diğerlerinin (2010) son sınıf kimya öğretmen adayları ile yaptıkları çalışma sonuçlarıyla da paralellik göstermektedir. Karşlı, Şahin ve Ayas'ın (2009) fen bilgisi öğretmenleri ile bireysel mülakatlar yaparak onların BSB hakkındaki fikirlerini ortaya çıkarmak amacıyla yaptıkları araştırma sonucunda fen bilgisi öğretmenlerinin çoğunun BSB hakkında teorik bilgiye sahip olmadıkları ve bu kavramları Bloom Taksonomisi, Piaget'in Formal Operasyon Evresi ve problem çözme gibi kavramlarla açıklamaya çalıştıkları gözlenmiştir. Laçin Şimsek (2010)'te 4. sınıf öğretmen adaylarının deneylerle öğrencilere kazandırılması hedeflenen BSB'leri tespit edebilme yeterliliklerini belirlemek için çalışma yapmıştır. Araştırmada, öğretmen adaylarının deneylerdeki yapılan işlemlerde hangi BSB'lerin kazandırıldığını tespit etmede problem yaşadıkları ve BSB basamakları ile Bloom taksonomisi ve problem çözme yönteminin basamaklarını birbirine karıştırdıkları belirtilmiştir. Sinan ve Uşak (2011) araştırmalarında 27 biyoloji öğretmen adayının BSB'lerini değerlendirmek için biyokimya laboratuvar dersinde öğrencileri gruplar halinde deneylerini yaparken gözlemlemiş ve bu öğrencilerin 3 tanesi ile de yarı yapılandırılmış görüşmeler yapmışlardır. Araştırma sonucunda öğretmen adaylarının özellikle "işlemsel bilgi, deney yönergesini anlama, deney düzeneği kurma, hipotez oluşturma-sınama ve günlük yaşamla ve diğer alanlarla çapraz ilişkilerin kurulması" alanlarında zayıf olduğu tespit edilmiştir. Ayrıca bu araştırmada öğrencilerin yapılan deneylerle kavram ve günlük yaşam arasındaki ilişkiyi kurmada zayıf oldukları da gözlenmiştir. Fen ve teknoloji öğretmeni yetiştiren programların ders içerikleri incelendiğinde BSB'nin özel öğretim yöntemleri dersinin bir alt konusu olduğu ve fen öğretimi laboratuvar uygulamaları dersinde de bu konunun uygulamalarına yer verildiği görülmektedir. Fakat yapılan çalışmalarda öğretmen adaylarının mevcut sistemde bu dersleri almalarına rağmen öğretim programlarında özenle vurgulanan BSB'lerden çok haberdar olmadıkları görülmektedir. Mevcut sistemde halen yetişen öğretmen adaylarının ve öğretmenlerin öğretim programlarında özenle vurgulanan BSB'lerden haberdar olmadıkları ve ancak halihazırda yetişenlerin iyi yetişmesi halinde programın daha amaca yönelik olarak uygulanabileceği ve öğrencilerin BSB seviyelerinin ve başarılarının artacağı düşünüldüğünde öğretmen adaylarının BSB'lerinin gelişimine daha çok vurgu yapan çalışmaların yanı sıra onların bu becerilere ne derecede sahip olduğunun ya da bu öğrenme ortamlarında onların BSB gelişim seviyelerinin nasıl olduğunun belirlenmesi de o

derecede önem kazanmaktadır. Bunun için öğretim sürecinde öğretmen adaylarının BSB'sini ölçen geçerli, güvenilir ve uygun ölçme araçlarına duyulan ihtiyaç gün geçtikçe artmaktadır.

Literatürde Geliştirilmiş ve Kullanılmış Bilimsel Süreç Beceri Testleri

İlgili literatür incelemesi yapıldığında, öğrencilerin BSB'sini ölçmek için geliştirilmiş pek çok testin olduğu görülmektedir. Bu tür testler başlangıçta 1960–1970 yılları arasında ABD'de geliştirilen Biological Sciences Curriculum Study (BSCS), Science Curriculum Improvement Study (SCIS), Intermediate Science Curriculum Study (ISCS) ve Science A Process Approach (SAPA) gibi programlarda kullanılmak üzere geliştirilmiştir. SAPA programı için, Walbesser (1965) K-3 seviyesindeki, Ludeman (1974) 6. sınıf seviyesindeki ve McLeod, Berkheimer, Fyffe ve Robison (1975) ise 6. ve 7. sınıf seviyesindeki öğrenciler için bilimsel süreç testleri geliştirmişlerdir. Daha sonraki yıllarda geliştirilen BSB testleri ise özel bir programa dayandırılarak değil, bir ihtiyaca yanıt olacak şekilde farklı sınıf seviyesindeki öğrencilere yönelik hazırlanmıştır. Tannenbaum (1968) 8. sınıf öğrencileri için, çoktan seçmeli formatta, sekiz farklı beceriyi; Fyffe (1971) ilköğretim seviyesindeki öğrenciler için, çoktan seçmeli formatta, iki farklı beceriyi; Molitor ve Kenneth (1976) 4- 6. sınıf seviyesindeki öğrenciler için, iki farklı beceriyi; Dillashaw ve Okey (1980) 7–12. sınıf seviyeleri için, çoktan seçmeli (multiple choice) formatta beş farklı beceriyi; Tobin ve Capie (1982) ortaokul ve lise seviyesindeki öğrenciler için, çoktan seçmeli formatta dört farklı beceriyi; Berger (1982), 12-14 yaş grubundaki öğrenciler için, bilgisayar destekli test formatında tek bir beceriyi; Burns, Okey ve Wise (1985) ortaokul ve lise seviyesindeki öğrenciler için, çoktan seçmeli formatta beş farklı beceriyi; Enger ve Yager (1998) ilköğretim seviyesindeki öğrenciler için; Solano-Flores (2000) 5. ve 6. sınıf seviyesindeki öğrenciler için, basit araç gereç aktiviteleri ile BSB'yi ölçme testleri geliştirmişlerdir. Beaumont Walters ve Soyibo (2001) 10. sınıf seviyesindeki öğrenciler için, çoklu formatta beş farklı beceriyi; Temiz, Taşar ve Tan (2006) 9. sınıf seviyesindeki öğrenciler için, çoklu formatta 12 farklı beceriyi; Temiz (2007) 9. sınıf seviyesindeki öğrenciler için, çoklu formatta altı farklı beceriyi; Monica (2005) 10–12. sınıf seviyelerindeki öğrenciler için beş farklı beceriyi; Aydoğdu ve Ergin (2009), Çalışkan ve Kaptan (2009) ve Hazır ve Türkmen (2008), ilköğretim seviyesindeki öğrenciler için, BSB'yi ölçme testleri geliştirmişlerdir. Bu çalışmalara ek olarak Feyzioğlu ve diğerleri (2012) 9. Sınıf öğrencilerinin BSB'lerini ölçmek için gözlem, sınıflama, ölçme, ilişki kurma, çıkarım yapma, tahminde bulunma, hipotez kurma, değişken belirleme ve kontrol etme, araştırma tasarlama, veri toplama ve kaydetme ve verilerin analizi alt faktörlerinden oluşan bir test geliştirmişlerdir. Geliştirilen bu testler incelendiğinde, testlerin genel olarak ilköğretim veya ortaöğretim seviyesindeki öğrencilerin BSB'sini ölçmek için hazırlandığı, daha büyük yaş gruplarındaki öğrenciler üzerinde kullanılabilecek özgün bir testin olmadığı dikkat çekmektedir. Özellikle fen eğitimi çalışmalarında, bilimsel süreç becerilerinin geliştirilmesiyle ilgili konularda çalışan araştırmacıların geliştirdikleri etkinlikleri uygulamadan önce ve sonra öğrencilerin BSB'lerini ölçmek için özellikle lisans düzeyindeki öğrenci seviyesine uygun ölçme araçlarına ulaşamamaktadır (Temiz, 2007; Karşlı, 2011).

Türkiye genelinde, BSB'yi ölçme amacıyla genel olarak yabancı literatürden alınan testlerin dil geçerliği yapıldıktan sonra kullanılması da dikkati çeken bir diğer noktadır (Ünal Coban, 2009; Tavukçu, 2008). Saraçoğlu, Büyük ve Tanık (2012) tarafından birleştirilmiş ve bağımsız sınıflarda öğrenim gören ilköğretim öğrencilerinin BSB düzeylerini belirlemek için yapılan çalışmada Burns, Okey ve Wise (1985) tarafından geliştirilip, Geban, Aşkar ve Özkan (1992) tarafından Türkçe'ye çevrilmiş olan BSB testi kullanılmıştır. Aynı test Ünal Coban (2009) tarafından modellemeye dayalı fen öğretiminin 7. Sınıf ilköğretim öğrencilerinin

BSB'lerine etkisini incelemek amacıyla da kullanılmıştır. Ayrıca, 8. Sınıf öğrencilerine uygun olarak geliştirilen ve 36 maddeden oluşan bu test Aktamış (2007) tarafından 7. sınıflara uyarlanarak 27 maddeye indirilerek de kullanılmıştır. Buna ek olarak aynı test "Fen Eğitiminde Bilgisayar destekli öğrenme ortamının öğrencilerin akademik başarı, BSB ve bilgisayar kullanmaya yönelik tutumuna etkisi adlı çalışma için de kullanılmıştır (Tavukçu, 2008). Benzer olarak dil geçerliği yapılarak kullanılan Smith (1995) tarafından hazırlanan "ilköğretim öğrencileri için Bilimsel Süreç Becerileri Değerlendirme" testi Şenyüz (2008) tarafından 2000 ve 2005 yılı ilköğretim fen programlarını bilimsel süreç becerileri açısından karşılaştırmak için yapılan çalışmada kullanılmıştır. Elli sorudan oluşan test ilköğretim öğrencilerinin 6, 7 ve 8. sınıfları için hazırlanmıştır. Testte ölçülmek istenen beceriler sırasıyla "gözlem, sınıflama, çıkarım yapma, tahmin, ölçme, iletişim, uzay-zaman ilişkisi, işlevsel tanımlama, hipotez oluşturma, deney yapma, değişkenleri belirleme, veri yorumlama ve model oluşturma" olarak toplam 13 beceriyi içermektedir.

İlgili literatürde örneklem grubu özelliklerinin, çalışma konularının ve çalışmaların farklı olmasına rağmen araştırmalarda aynı testlerin kullanıldığı görülmektedir (Ünal Coban, 2009; Şenyüz, 2008; Tavukçu, 2008; Aktamış, 2007). Bu kullanılan testler esasında BSB'leri ölçmelerine rağmen daha üst düzeyde BSB'lere sahip olması beklenen öğrencilere de uygulanması ve test içeriği oluşturulurken konu alan bilgisine dikkat edilmemesi geçerlik ve güvenilirlik konusunda tartışmalara neden olmaktadır (Feyzioğlu vd., 2012). Bu durumun oluşmasına gerek ulusal gerekse uluslararası literatürde özellikle üst düzeyde BSB'lere sahip öğrencilerin BSB seviyelerini ve onların öğretim öncesinden sonrasına BSB gelişim seviyelerini belirlemede kullanılabilir, çoklu formatta geçerli ve güvenilir bir ölçme aracının bulunmaması neden gösterilebilir.

Fen bilgisi dersinin doğal yapısı, çağın beklentileri ve yeni düzenlenen programların vizyonuyla ilişkilendirildiğinde, Fen ve Teknoloji dersi kapsamında kazandırılması gereken BSB'yi bütünsel olarak ele alan ve programdaki ünitelerle ilişkilendirilmiş bir testin bulunmaması konunun önemini ortaya koymaktadır. Aynı zamanda böyle bir testin, daha büyük yaş gruplarındaki öğrencilerin BSB'sini geliştirmek için tasarlanmış bir öğretim yönteminin etkili olup olmadığını tespit etmek ve süreç içinde öğrencilerin BSB seviyelerini izlemek için eğitimci ve eğitim araştırmacılarının ihtiyaçlarına cevap verebileceği düşünülmektedir. Bütün bu gerekçelerden hareketle fen ve teknoloji dersi öğretmen adaylarının BSB'lerini ölçmek için Fen ve Teknoloji Öğretim Programı'ndaki (FTÖP) ünitelerle ilişkilendirilmiş geçerli ve güvenilir bir testin geliştirilmesinin, literatürdeki bu eksikliği kapatacağına inanılmaktadır.

Bu çalışmanın amacı; FTÖP'te belirtilen BSB kazanımları ile paralel hazırlanmış, fen ve teknoloji dersi öğretmen adaylarının BSB'lerini ve BSB gelişim seviyelerini belirlemeye yönelik geçerliği ve güvenilirliği sağlanmış çoklu formda bir bilimsel süreç becerileri testi (BİSBET) geliştirmektir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. Araştırma kapsamında geliştirilen BİSBET fen ve teknoloji dersi öğretmen adaylarının BSB'lerinin ölçülmesinde ne derecede geçerli bir testtir?
2. Araştırma kapsamında geliştirilen BİSBET fen ve teknoloji dersi öğretmen adaylarının BSB'lerinin ölçülmesinde ne derecede güvenilir bir testtir?

YÖNTEM

a) Araştırmanın Yöntemi

Bu araştırma bir test (ölçek) geliştirme çalışmasıdır. Eğitim araştırmalarında yapılan test geliştirme çalışmaları incelediğinde test geliştirme süreci hazırlık, uygulama ve raporlaştırma aşamalarından oluşmakla birlikte (Çalışkan & Kaplan, 2009) bu aşamalarda çeşitli adımların

takip edildiği görülmektedir. Bu araştırmada Burns ve diğerleri (1985), Temiz ve diğerleri (2006) ve Feyzioğlu ve diğerleri (2012) tarafından da test geliştirme çalışmalarında kullanılan, test maddelerinin yazılması, pilot uygulamalar, geçerlilik, güvenilirlik ve madde analizi şeklinde adımlar takip edilmiş ve sırasıyla sunulmuştur. Bu araştırma klasik bir araştırmadan ziyade bir test geliştirme çalışması olduğu için takip edilen adımların eğitim araştırmaları için uygun olduğu söylenebilir.

b) Örneklem

Araştırmanın evreni fen bilgisi öğretmenliği bölümünde okuyan öğrencilerden oluşmaktadır. Araştırmanın örnekleme ise, 2010–2011 eğitim-öğretim yılında Karadeniz Teknik Üniversitesi ve Giresun Üniversitesi Eğitim Fakültelerinin İlköğretim Bölümü Fen Bilgisi Öğretmenliği programının üçüncü sınıflarında öğrenim gören toplam 197 öğretmen adayından oluşmaktadır. Testin birinci pilot uygulaması Karadeniz Teknik Üniversitesi'nde öğrenim gören öğrencilerle (N=87); ikinci pilot uygulaması Giresun Üniversitesi'nde öğrenim gören öğrencilerle (N=110) yapılmıştır. Araştırmalarda örneklem büyüklüğü arttıkça, gerçek puanlara daha fazla yaklaşılabileceği, daha doğru tahminler yapılabileceği bilinmektedir. Grup büyüklüğünün belirlenmesinde faktör analizi, madde analizi gibi işlemler dikkate alınarak madde sayısının en az iki kat, hatta tercihen 10 kat olması önerilir (Kline, 1994). Bu araştırma kapsamında ilk etapta 45 madde halinde hazırlanan test, yaklaşık olarak madde sayısının 4 katından daha fazla sayıda örnekleme uygulanabilmiştir. Test maddelerinin performansa yönelik soruları da içerdiği başka bir deyişle cevaplanmasının uzun süre aldığı da düşünüldüğünde örneklem sayısının yeterli olduğu söylenebilir. Araştırmaya katılan öğrencilerin fen bilgisi öğretmenliği üçüncü sınıf öğrencilerinden seçilmesinde bu öğrencilerin özel öğretim yöntemleri ve fen öğretimi laboratuvar uygulamaları dersleri kapsamında BSB ile ilgili teorik bilgileri almaları etkili olmuştur. Öğrencilerin teorik olarak bilmedikleri veya fikir sahibi olmadıkları bir konu hakkında test sorularını yanıtlamasındansa ilgili konuda teorik bilgiyi almış olan 3. sınıf öğrencilerin örneklem olarak seçilmesi yoluna gidilmiştir.

c Veri Toplama Aracı

Bu çalışmada BSB testi geliştirilirken ölçülecek davranışın niteliği dikkate alınarak, son zamanlarda program geliştirmeciler tarafından da öngörülen ölçme-değerlendirme tekniklerine uygun, çoklu formda bir test hazırlanmaya çalışılmıştır.

d) BİSBET'in Geliştirilme Sürecinde Takip Edilen Aşamalar:

1. Aşama: Bu bölümde araştırma kapsamında test geliştirme sürecinde testin kullanılış amacı belirlendikten sonra, ölçülmesi istenilen BSB'lerin neler olacağına karar verilmiştir. Bu çalışmaya temel olan BSB; gözlem yapma, ölçme, sınıflama, verileri kaydetme, verileri kullanma ve model oluşturma, önceden kestirme, değişkenleri belirleme, değiştirme ve kontrol etme, hipotez kurma, verileri yorumlama, sonuç çıkarma ve deney tasarlama-yapma becerilerinden oluşmaktadır. Söz konusu becerilere sahip olan bir öğrenciden beklenen davranışların neler olduğunu belirlemek için ise bu becerilere yönelik FTÖP'te belirtilen öğrenci kazanımlarından yararlanılarak bu kazanımlar öğrenci seviyelerine uygun olarak düzenlenmiştir (MEB, 2006). FTÖP'teki öğrenci kazanımlarından yararlanılmasında hizmet öncesi öğretim programlarında BSB kavramının öğretildiği Özel Öğretim Yöntemleri ve Fen Öğretimi Laboratuvar Uygulamaları I derslerinin ders içeriğinde, deney tasarlama ve geliştirme, BSB ve nasıl kazandırıldıkları ile ilgili uygulamaların yapılmasını gerektiren ifadelerin olması ve bu derse yönelik hazırlanan ders kitaplarının özellikle FTÖP'teki konu ve

BSB kazanımlarına odaklanılarak hazırlanması etkili olmuştur (Bahar vd., 2008; Bozkurt vd., 2008).

2. Aşama: Ölçülmesi hedeflenen davranışların neler olduğunun belirlenmesinin ardından bu becerilerin ölçülmesinde kullanılacak soru tiplerinin özelliklerine karar verme aşamasına geçilmiştir. Testte yer alacak soru tiplerine karar verilirken, soruların öğrenciden beklenen BSB kazanımlarına uygun formatta olmasına özen gösterilmiştir. Örneğin performansa dayalı becerileri ölçmek için açık uçlu soru maddelerinin kullanılması tercih edilmiştir. Bu düşünceden hareketle açık uçlu, çoktan seçmeli ve kağıt-kalem materyaline dayalı performansa yönelik soru formatlarından oluşan maddeler yazılmıştır. Her bir maddede yer alan örneklemelerin veya deney senaryolarının günlük hayatta karşılaşılan ve FTÖP'teki bazı ünitelerde öğrenilen konularla ilişkili temel kavramlardan oluşmasına dikkat edilmiştir. Çünkü öğrencilerin bilmedikleri veya fikir sahibi olmadıkları konu ile BSB'nin değerlendirilmesi doğru değildir (Bozkurt & Olgun, 2005). Testin maddeleri yazılırken literatürde BSB'leri ölçmek için geliştirilmiş test çalışmaları incelenmiş, bu araştırmalardan bazıları araştırmacılar için yol gösterici olmuştur (Örneğin; Burns vd., 1985; Temiz, 2007). Farklı soru formatlarında, ölçülmesi istenen 11 farklı beceriye yönelik 30'u çoktan seçmeli ve 15'i açık uçlu yapıda olmak üzere toplam 45 madde hazırlanmıştır. Çoktan seçmeli her bir madde beş seçenektan oluşmaktadır.

3. Aşama: Test maddelerinin yazılması işleminden sonra, maddelerin tekrar gözden geçirilmesi, test maddelerinin okunabilirliği ve testte yer alan şekil ve çizimlerin anlaşılabilirliği hakkında fikir sahibi olmak, anlaşılmasında güçlük çekilen terimleri testten çıkarmak ve testin cevaplandırılması için gereken süreyi tespit edebilmek için 45 maddelik test, Karadeniz Teknik Üniversitesi, Fatih Eğitim Fakültesi Fen Bilgisi Öğretmenliği bölümünün 3. sınıfında öğrenim gören toplam 87 öğrenciye uygulanmıştır. Yapılan birinci pilot uygulama esnasında öğrencilerin testte anlamakta güçlük çektikleri bazı ifadeler ve şekiller üzerinde çeşitli düzeltme ve düzenlemeler yapılmıştır. Örneğin değişkenleri belirleme ve hipotez kurma becerisini ölçmek için başlangıçta "Onur, annesini kayısı reçeli yaparken izler ve 2 litre kayısı-su karışımında 3 kg kadar şekerin çözünebildiğini fark eder. Onur şekerin çözünürlüğündeki artışın sebeplerini araştırmaya karar verir. Aşağıdaki soruları Onur'un yaptığı deneyi dikkate alarak cevaplandırınız" şeklinde ifade edilen soru kökünde öğrenciler yapılan deneyin tam olarak açık olmadığını söyleyerek soru kökü "Onur, annesini kayısı reçeli yaparken izler ve 2 litre kayısı-su karışımında 3 kg kadar şekerin çözünebildiğini fark eder. Onur şekerin çözünürlüğündeki artışın sebeplerini araştırmaya karar verir. Onur'un deney yapmasına yardım etmek için I, II, III, IV ve V alt sorularını çözünürlüğü etkileyen faktörleri dikkate alarak cevaplandırınız." şeklinde düzenlenmiştir (BİSBET'in 18. sorusu). Buna benzer olarak verileri yorumlama becerisini ölçmeye yönelik hazırlanan soruda verilen sütun grafiğinde bölmelerin ayrımı tam olarak yapılamamış, bunun neticesinde şekildeki bölmeler siyah ve gri renkleri kullanılarak daha belirgin bir halde sunumu için düzenleme yapılmıştır (BİSBET'in 22. sorusu). Örneğin BİSBET'in, aşağıdaki verilen bilgilerden hangisine gözlemler sonucunda ulaşamaz şeklindeki 1. sorusunun C seçeneği başlangıçta "Bir asit çözeltisi elektrik akımını iletir." şeklindeyken öğrencinin asit çözeltisine daldırdığı iletkenle bu ifade verilen durumu gözlemleyebileceğinden yola çıkılarak bu seçenek "Asit, baz ve tuz çözeltileri suda iyonlarına ayrılarak çözünürler." şeklinde değiştirilmiştir. Buna benzer olarak BİSBET'in sonuç çıkarma becerisiyle ilgili sorulan 30. sorusunda doğru seçenek diğer çeldiricilerden daha uzun olduğu için bu seçenek daha sadeleştirilerek diğer seçeneklerle eşdeğer uzunlukta düzenlenmiştir. Pilot uygulama esnasında, testin cevaplandırılması için, 75-90 dk. arasında bir sürenin yeterli olduğu gözlemlenmiştir. Bu şekilde öğrencilerin uygulama esnasında anlayamadıkları teknik terimler testten çıkarılarak testin bütün yönüyle anlaşılabilir olmasına çalışılmıştır. Pilot uygulama esnasında öğrencilerin

çoğunluğu tarafından açıklanması beklenen ve öğrencilerce anlaşılmadığı gözlenen üç maddenin de testten çıkarılmasına karar verilmiştir. Test maddelerinin çeldiricileri ve cevap anahtarı tekrar gözden geçirilmiştir.

4. Aşama: Birinci pilot uygulamanın ardından testteki soru senaryolarında ve çeldiricilerde bilimsel hata olup olmadığını belirlemek, test maddelerinin kapsam ve görünüş geçerliğini test etmek için hazırlanan test, beşi kimya eğitiminde öğretim üyesi ve ikisi fen ve kimya eğitiminde araştırma görevlisi olmak üzere toplam 7 fen eğitimcisinin incelemesine sunulmuştur. Geçerlik için görüşlerine başvurulmuş uzmanlardan birisi birçok doktora tezi ve yüksek lisans tezleri yürütmüş olup yıllardır doktora öğrencilerine test geliştirme dersi vermektedir. Bununla birlikte alan uzmanlarından bir diğeri de lisansüstü eğitimde nicel veri analizi derslerini yürütmektedir. Uzman görüşlerinin alınmasında açık uçlu sorulardan oluşan bir uzman değerlendirme formundan yararlanılmıştır. Değerlendirme formunda testte yer alan maddelerin BSB'leri ölçmedeki uygunluğuna, maddelerde kullanılan ifadelerin doğruluğuna ve bilimselliğine, öğrenci seviyesine uygunluğuna, test maddelerinin okunabilirliğine ve testte yer alan şekil ve çizimlerin anlaşılabilirliğine yönelik sorular yer almaktadır. Test maddeleri ile ilgili bu forma verilen uzman görüşleri arasında uyuşma görülen noktalarda düzenlemeler yapılmıştır.

5. Aşama: Uzmanlardan alınan geri dönütler doğrultusunda uygun görülmeyen maddelerin testten çıkarılmasına karar verilmiştir. Bunun ardından teste son şeklini verebilmek, testin geçerlik, güvenilirlik ve madde analizlerini yapabilmek için ikinci pilot uygulama; Giresun Üniversitesi, Eğitim Fakültesi Fen Bilgisi Öğretmenliği programı 3. sınıfında öğrenim gören toplam 110 öğrenciye uygulanmıştır. Uygulama sonuçları çoktan seçmeli test maddeleri için doğru cevaplar 1, yanlış ve boş bırakılan cevaplar ise 0 olacak şekilde puanlanmıştır. Farklı beceri türlerine yönelik olarak testin açık uçlu maddelerinin puanlanmasında ise Temiz (2007) tarafından geliştirilen dereceli puanlama anahtarı ve kontrol listesi kullanılmıştır.

BİSBET'in geçerlik ve güvenilirliğine ilişkin işlemler çoktan seçmeli ve açık uçlu test maddeleri için ayrı ayrı yapılmıştır. İkinci pilot uygulamadan elde edilen verilere göre geçerlik, güvenilirlik ve madde analizi çalışmaları için sırasıyla aşağıdaki işlemler yapılmıştır.

BİSBET'in Geçerliğine İlişkin İşlemler: Hazırlanan testin görünüş ve kapsam geçerliğini sağlamak için test maddeleri uzmanların görüşlerine sunulmuştur.

Belirli bir konuyu ölçmek amacıyla hazırlanan ölçeğin yapı geçerliğine kanıt sağlamak için başvurulmuş yöntemler faktör analizi, hipotez testi ve iç tutarlılık analizi teknikleridir (Büyüköztürk, 2007: 168). Faktör analizi, tüm veri yapıları için uygun olmayabilir. Verilerin faktör analizine uygun olabilmesi için gözlem verilerinin eşit aralıklı veya oranlı ölçek verisi niteliğinde olması ve verilerin normal dağılım göstermesi gerekir. Buna ek olarak Kaiser-Meyer-Olkin (KMO) katsayısının 0,60'dan yüksek ve Barlett Küresellik Testi sonucunun anlamlı çıkması şartları sağlanmalıdır (Büyüköztürk, 2007: 126; Şencan, 2005: 375). Geliştirilen testin çoktan seçmeli test maddelerine ilişkin verilerin normal dağılım göstermemesi ve KMO katsayısının 0,60'dan küçük çıkmasından (0,47) dolayı faktör analizi yapılamamıştır. Bu nedenle çoktan seçmeli test maddelerinin yapı geçerliğine ilişkin kanıt sağlamak amacıyla hipotez testi (grup farklılıklarıyla yapı geçerliğinin analizi) ve iç tutarlılık analizi tekniklerinden faydalanılmıştır.

BİSBET'in Güvenirliğine İlişkin İşlemler: Test geliştirme sürecinin en önemli aşamalarından birisi de testin güvenilirliğidir. Güvenirlik deneklerin test maddelerine verdikleri cevaplar arasındaki tutarlılık olarak tanımlanmaktadır. Bir ölçme aracının güvenilirliği için aranılan iki temel ölçüt, "değişik zamanlarda elde edilen puanlar arasındaki tutarlılık" ve "aynı zamanda elde edilen cevaplar arasındaki tutarlılık" olarak açıklanabilir (Büyüköztürk, 2007:

170). Bir testin bu ölçütleri karşılama düzeyini incelemek için başvuru olan güvenilirlik türleri: 1) İç tutarlılık güvenirligi, 2) Test-tekrar test güvenirligi, 3) Paralel formlar güvenirligi ve 4) Gözlemciler arası güvenirlilik olarak sıralanabilir (Şencan, 2005). İlgili literatürde test-tekrar test ve paralel formlar yöntemlerinin kullanılması konusunda güçlüklerin (birden fazla uygulama, birden fazla forma gerektirme vb. sebeplerle) olabileceği üzerine vurgu yapılmıştır (Şencan, 2005; Çepni, 2007). BSB'leri ölçmek amacıyla öğrencilere uygulanan test maddelerinin sayısının fazla olması ve cevaplanmasının uzun sürmesi nedeniyle öğrencilerin ikinci kez testi aynı istekle cevaplamama ihtimalleri düşünülerek test-tekrar test ve paralel formlar yöntemleri kullanılmamıştır. Bu araştırmada BİSBET'in güvenirligini test etmek için iç tutarlılık analizi ve gözlemciler arası tutarlılık yöntemlerine başvurulmuştur.

Test puanları arasındaki iç tutarlılığı incelemek amacıyla, Kuder Richardson-20,21 (KR-20,21) ve Cronbach tarafından geliştirilmiş olan alfa katsayıları kullanılır (Büyüköztürk, 2007). KR-21 formülü çoktan seçmeli maddeler ve ölçekler için kullanılır (Şencan, 2005). Bu yüzden geliştirilen BİSBET'in çoktan seçmeli maddelerinin iç tutarlılık güvenirlilik analizi, Kuder Richardson-21 (KR-21) ve Cronbach alfa güvenirligi ile kestirilirken açık uçlu test maddeleri için ise Cronbach alfa güvenirlilik katsayısı ve her bir performansa dayalı açık uçlu test maddesi için gözlemciler arası tutarlılık katsayısı kullanılmıştır. Bunun için ikinci pilot uygulamaya katılan öğrenciler arasından tesadüfi olarak seçilen 15 öğrencinin cevap kâğıtları puanlanmıştır. Bu araştırmada testin açık uçlu maddeleri (7., 8., 9., 11., 12., 18., 19., 35. ve 36. maddeler) için dereceli puanlama anahtarı ve kontrol listeleri gibi ölçüm araçları kullanılmıştır. Açık uçlu soruların her biri kendi içinde farklı becerilere karşılık geldiği için değerlendirmede farklı ölçüm araçları kullanılmıştır (hangi maddelerin hangi puanlama aracı ile değerlendirildiği Tablo 5'te görülmektedir). Bu maddelere öğrencilerin verdiği cevaplar ölçüm araçları kullanılarak, birisi ilk yazar olan araştırmacı ve diğeri fen eğitiminde öğretim üyesi olan bağımsız iki araştırmacı tarafından puanlandırılmış ve gözlemcilerin arasındaki tutarlılığa bakılarak Kappa katsayısı hesaplanmıştır. Gözlemcilerin verdiği kategorik nitelikteki puanlar arasındaki tutarlılık için Kappa katsayısı hesaplanmıştır.

BİSBET'in Madde Analizine İlişkin İşlemler: Madde analizi yaparken ikinci pilot uygulamanın çoktan seçmeli test maddelerinden her bir öğrencinin aldığı puanlar hesaplanmış ve bu puanlar büyükten küçüğe doğru sıralanmıştır. Testteki maddeleri cevaplayan öğrencilerin toplam sayısının %27'si kadar alt ve üst gruplar oluşturulmuştur. Bunun ardından madde güçlüğü, $p=(Dü+Da)/2N$ formülünden, ayırt edicilik ise $d=(Dü-Da)/N$ ($Dü=$ maddeyi doğru cevaplayan üst grup öğrenci sayısı; $Da=$ maddeyi doğru cevaplayan alt grup öğrenci sayısı; $N=$ Tüm grubun % 27'sidir) formülünden yararlanılarak hesaplanmıştır.

BULGULAR

Bu bölümde araştırmanın amacına uygun olarak test geliştirme sürecinde yapılan geçerlik, güvenirlilik ve madde analizi çalışmalarından elde edilen bulgular sunulmaktadır.

1) BİSBET'in Geçerliğine İlişkin Bulgular

a) BİSBET'in Kapsam Geçerliğine İlişkin Bulgular

Testin kapsam geçerliğine ilişkin kanıt sağlamak için uzman görüşlerine başvurulmuştur. Birinci pilot uygulamadan önce test maddeleri, testi geliştiren araştırmacılar tarafından tekrar tekrar incelenmiş ve test maddelerinin BSB kazanımlarının tümünü kapsayacak şekilde olmasına özen gösterilmiştir. Birinci pilot uygulamadan sonra gerekli düzenlemelerin ardından 42 (28'i çoktan seçmeli ve 14'ü açık uçlu) maddeye düşürülen test

uzmanların incelemesine sunulmuştur. Uzmanlar test maddelerinin, öğrencilerin düzeyine uygunluğu, öğretim programlarında belirtilen BSB kazanımlarını ölçüp ölçmediği, testteki soru senaryolarında ve çeldiricilerde bilimsel hata olup olmadığı boyutlarında testi incelemişlerdir. İncelemeler sonucunda 42 maddelik testte ifadelerin anlaşılabilirliği açısından ve özel alan bilgisi gerektirmesinden dolayı uygun görülmeyen biri çoktan seçmeli ve üçü açık uçlu yapıda olmak üzere dört maddenin testten çıkarılmasına karar verilmiştir. Çıkarılan sorulardan 2'sinde "basınca dayanıklı bir kabın içindeki gaz üzerine yapılan etki ile basınç, hacim ve sıcaklık değerlerini gösteren bir grafikten yola çıkılarak öğrencilerden verileri yorumlamaları istenmektedir. Bu sorularda uzmanlardan birinin grafikteki basınç, hacim ve sıcaklıkla ilgili bütün verilerin karşılığında bilimsel değerlerin olması gerektiğine yönelik dönütünden ve bu soru tipine benzer soruların BİSBET'te fazla sayıda bulunmasından (Tablo 16, s. 94) yola çıkılarak bu sorular testten çıkarılmıştır. Buna benzer olarak testten çıkarılan diğer 2 açık uçlu soru ise deney tasarlama becerisi ile ilgilidir. Bu sorulardan birisi "Bir telin kalınlığı arttıkça direnci azalır" şeklinde ifade edilen hipotezi diğeri "Bir gazın üzerine ne kadar basınç uygulanırsa, o derecede sıvıda fazla çözünür" şeklinde ifade edilen hipotezi test etmek isteseydiniz nasıl bir deney tasarladınız?" şeklindedir. Testte deney tasarlama becerisi ile ilgili 2 tane daha açık uçlu soru bulunmaktadır. Fakat uzmanların uzun cevaplı olan bu beceriye yönelik soru sayısının fazla olmasının öğrencilerin soruları cevaplama isteğini düşürür şeklindeki önerisinden yola çıkılarak bu sorular da testten çıkarılmıştır. Uzmanlar, bunların dışında kalan 38 maddenin BSB kazanımlarını kapsadığı yönünde görüş belirtmişlerdir.

b) BİSBET'in Yapı Geçerliğine İlişkin Bulgular

Geliştirilen ölçme aracının yapı geçerliğine kanıt sağlamak için hipotez belirlenmiş ve bu hipotez test edilmiştir. "BSB'ye dayalı laboratuvar dersi almayan öğrencilerin (G1), geliştirilen BİSBET ortalama puanı, BSB'ye dayalı laboratuvar dersi alan öğrencilerin (G2) ortalama puanından daha düşüktür" şeklindeki hipotezi test etmek için, yedi hafta boyunca BSB'ye vurgu yapılarak laboratuvar dersi işlenen toplam 26 öğrencinin test ortalama puanı ile bu şekilde ders işlenmeyen toplam 25 öğrencinin test ortalama puanları arasındaki fark, ilişkisiz ölçümler için Mann Whitney U-testi ile analiz edilerek karşılaştırılmıştır. Bu analizde kullanılan veriler, testin ikinci pilot uygulamasına katılan öğrencilerin test sonuçlarından elde edilmiştir.

Tablo 1. BİSBET Puanlarının Grup Farklılıklarına Göre Mann Whitney U-testi Sonuçları

Test türü	Grup	N	Sıra Ort.	Sıra Top.	U	p
BİSBET'in çoktan seçmeli kısmı	G1	25	20.66	516	191	.029
	G2	26	29.52	708		
BİSBET'in açık uçlu kısmı	G1	25	13.76	344	19	.000
	G2	26	36.71	881		

Tablo 1'de BSB'ye yönelik eğitim alan öğrencilerin BİSBET'in hem çoktan seçmeli, hem de açık uçlu maddelerinden aldıkları toplam BSB puanlarının, eğitim almayan öğrencilere göre anlamlı derecede daha yüksek olduğu görülmektedir ($p < .05$). Sıra ortalamaları dikkate alındığında da, BSB'ye vurgu yapılarak laboratuvar dersi işlenen grubun (G2) testin çoktan seçmeli ve açık uçlu kısımlarından aldıkları puanlarının G1 grubundaki öğrencilerin puanlarından daha yüksek olduğu anlaşılmaktadır.

2) BİSBET'in Güvenirliğine İlişkin Bulgular

Bu çalışmada BİSBET'in güvenirlğini test etmek için iç tutarlılık analizi ve gözlemciler arası tutarlılık yöntemlerine başvurulmuştur.

b) BİSBET'in İç Tutarlılık Güvenirliğine İlişkin Bulgular

Testin çoktan seçmeli kısmının iç tutarlılık analizi için öncelikle öğrenci cevapları doğru seçenek işaretlenmiş ise 1 ile yanlış seçenek işaretlenmiş ise 0 olarak puanlanmıştır. Bu şekilde elde edilen test verileri için KR-21 ve Cronbach alfa katsayıları hesaplanmıştır. Testin açık uçlu kısmının iç tutarlılık analizinde ise öğrenci cevapları Temiz (2007) tarafından geliştirilen dereceli puanlama anahtarları ve kontrol listesi aracılığı ile puanlanmıştır. Her bir madde keline özgü bir beceriyi ölçtüğü ve bu farklı beceri türlerini ölçen maddeler farklı birer değerlendirme aracı ile puanlandığından, her beceriye yönelik maddeler üzerinde ayrı ayrı iç tutarlılık, Cronbach alfa güvenirlği katsayıları hesaplanmıştır. Geçerlik çalışması sonucunda toplam 27 çoktan seçmeli ve 11 açık uçlu madde olmak üzere 38 maddeye indirgenen ölçek için belirlenen iç tutarlılık Cronbach alfa katsayıları Tablo 2'de sunulmuştur.

Tablo 2. BİSBET'in Çoktan Seçmeli ve Açık Uçlu Maddeleri için Hesaplanan İç Tutarlılık Katsayıları

Test türü	Ölçülmek İstenen Beceri Türleri	N	Madde sayısı	Cronbach alfa	KR-21
BİSBET'in çoktan seçmeli kısmı	Gözlem, Ölçme, Sınıflama, Önceden kestirme, Değişkenleri belirleme, değiştirme ve kontrol etme, Hipotez kurma, Verileri yorumlama, Sonuç çıkarma, Deney yapma-tasarlama	110	25	0,78	0,75
BİSBET'in açık uçlu kısmı	Sınıflama	110	2	0,71	-
	Verileri kaydetme	110	3	0,83	-
	Verileri kullanma ve model oluşturma (Grafik çizme)	110	2	0,90	-
	Değişkenleri belirleme ve hipotez kurma	110	2	0,71	-
	Deney tasarlama	110	2	0,90	-

BİSBET'in 27 maddeden oluşan çoktan seçmeli kısmı farklı beceri türlerini ölçmeye yönelik tek formattan (soru tipinden) oluşmaktadır. Testin çoktan seçmeli maddelerine ilişkin güvenirlk kestiriminde iki madde Cronbach alfa katsayısını düşürmesi nedeniyle testten çıkarılmıştır. Tablo 2'de görüldüğü gibi BİSBET'in 25 maddeye indirgenen çoktan seçmeli kısmı için Cronbach alfa katsayısı 0,78 ve KR-21 değeri 0,75 olarak hesaplanmıştır. BİSBET'in 11 maddeden oluşan açık uçlu kısmında sınıflama gerektiren maddeler arası güvenirlk Cronbach alfa katsayısı 0,71; verileri kaydetme için 0,83; verileri kullanma ve model oluşturma (grafik çizme) için 0,90; değişkenleri belirleme ve hipotez kurma için 0,71; deney tasarlama için 0,90 olarak kestirilmiştir.

b. BİSBET'in Gözlemciler Arası Tutarlılık Analizine İlişkin Bulgular

Gözlemciler arası tutarlılık, testin açık uçlu maddeleri (7., 8., 9., 11., 12., 18., 19., 35. ve 36. maddeler) için dereceli puanlama anahtarları ve kontrol listesi kullanılarak yapılan puanlamalarda gözlemciler arası tutarlılık hesaplanarak incelenmiştir. Açık uçlu maddelerin her biri kendi içinde farklı becerilere karşılık geldiği için değerlendirmede farklı ölçme araçları kullanılmıştır. Tablo 5'te hangi maddelerin hangi puanlama aracı ile değerlendirildiğine ilişkin bilgi ayrıntılı olarak sunulmaktadır. Bunun için ikinci pilot uygulamaya katılan öğrenciler arasından tesadüfi olarak seçilen 15 öğrencinin cevap kağıtları

birbirinden bağımsız iki araştırmacı tarafından, dereceli puanlama anahtarı ve kontrol listeleri aracılığı ile puanlanmıştır. Tablo 3'te araştırmacıların, 15 öğrencinin cevaplarının her birine verdikleri puanlar arasındaki tutarlılık oranı, Kappa katsayıları, hesaplanarak sunulmuştur.

Tablo 3. BİSBET'in Açık Uçlu Maddeleri İçin Gözlemciler Arası Tutarlılık Oranı (Kappa Katsayıları)

Madde numarası	G1-G2 arasındaki Kappa katsayısı	Madde numarası	G1-G2 arasındaki Kappa katsayısı	Madde numarası	G1-G2 arasındaki Kappa katsayısı
7. Madde	0,857	11. Madde	0,826	19. Madde	0,831
8. Madde	0,789	12. Madde	0,762	35. Madde	0,757
9. Madde	0,732	18. Madde	0,729	36. Madde	0,754

G1: 1. Gözlemci; G2: İkinci Gözlemci

Tablo 3'te performans dayalı BSB'leri ölçen dokuz madde için, ölçüm değerlerinin güvenilirliğine ilişkin iki farklı gözlemcinin puanları arasındaki Kappa katsayıları sonuçları görülmektedir. Ölçüm değerlerinin güvenilirliğine ilişkin Kappa istatistiği sonuçları incelendiğinde en düşük katsayının 0.72 ve en büyüğünün ise 0.85 olduğu ve Kappa katsayıları ortalamasının yaklaşık 0,78 olduğu görülmektedir.

3) BİSBET Madde Analizine Yönelik Bulgular

BİSBET'in çoktan seçmeli 25 maddelik kısmı için alt ve üst grupta yer alan öğrencilerin seçeneklere göre dağılımları, güçlük ve ayırt edicilik indisleri hesaplanmıştır. İkinci pilot uygulama verileri kullanılarak yapılan madde analizi sonuçları Tablo 4'te sunulmuştur.

Tablo 4. Üst ve Alt gruptaki Öğrencilerin Doğru Cevap Sayısına Göre Madde Analizi Sonuçları

M	Grup	A	B	C	D	E	P	d	M	Grup	A	B	C	D	E	p	d
1	üst	0	0	23	0	0	0,78	0,43	14	üst	3	0	20	0	0	0,48	0,78
	alt	7	0	13	2	1				alt	16	0	2	1	4		
2	üst	0	0	2	21	0	0,69	0,45	15	üst	0	20	3	0	0	0,61	0,52
	alt	1	1	8	10	2				alt	0	8	13	1	1		
3	üst	1	4	18	0	0	0,61	0,34	16	üst	0	0	23	0	0	0,80	0,40
	alt	1	12	10	0	0				alt	2	6	14	1	0		
4	üst	0	0	3	20	0	0,64	0,46	17	üst	0	0	0	3	20	0,67	0,40
	alt	0	2	10	9	1				alt	1	8	2	1	11		
5	üst	1	1	21	0	0	0,63	0,56	18	üst	20	0	2	0	1	0,67	0,40
	alt	0	15	8	0	0				alt	11	2	5	1	4		
6	üst	6	2	2	0	13	0,39	0,34	19	üst	0	0	21	2	0	0,73	0,36
	alt	10	0	7	1	5				alt	3	1	12	3	3		
7	üst	0	22	0	0	1	0,76	0,39	20	üst	0	1	22	0	0	0,74	0,43
	alt	1	13	0	0	9				alt	0	1	12	1	9		
8	üst	1	21	1	0	0	0,74	0,34	21	üst	0	20	1	0	2	0,67	0,40
	alt	4	13	2	4	0				alt	2	11	0	2	8		
9	üst	0	1	22	0	0	0,70	0,52	22	üst	19	0	1	3	0	0,67	0,30
	alt	0	11	10	1	1				alt	12	1	4	5	1		
10	üst	0	0	2	1	20	0,59	0,56	23	üst	2	18	1	2	0	0,59	0,40
	alt	1	0	12	3	7				alt	0	9	9	3	2		
11	üst	1	2	1	16	3	0,53	0,33	24	üst	0	22	0	1	0	0,65	0,61
	alt	5	5	3	8	1				alt	1	8	12	2	0		
12	üst	0	5	1	16	1	0,48	0,43	25	üst	5	0	18	0	0	0,63	0,30
	alt	0	14	2	6	1				alt	10	0	11	2	0		
13	üst	1	4	1	2	15	0,37	0,56									
	alt	1	6	1	13	2											

M: Madde no; p: Madde güçlüğü; d: Ayırt edicilik

Tablo 4 incelendiğinde BİSBET'in geçerlik ve güvenilirlik analizleri yapıldıktan sonra kalan 25 çoktan seçmeli maddenin ortalama gücünün yaklaşık 0,60; ortalama ayırt ediciliğinin ise 0,42 civarında olduğu görülmektedir.

BİSBET'in pilot uygulamaları, geçerlik, güvenilirlik ve madde analizi sonucunda elde edilen son formunda hangi maddelerle, hangi becerilerin ölçüleceği ayrıntılı olarak Tablo 5'te sunulmuştur.

Tablo 5. BİSBET'te Ölçülecek BSB Türüne Göre Madde Yapısı, Sayısı, Numaraları ve Bu Maddelerin Puanlanmasında Kullanılan Anahtarların Türü

Ölçülecek BSB türü	Madde yapısı türü	N	Madde numaraları	Puanlama Aracı
Gözlem yapma	Çoktan seçmeli	1	1.	Cevap anahtarı
Ölçme	Çoktan seçmeli	2	2., 3.	Cevap anahtarı
Sınıflama	Çoktan seçmeli	1	4.	Cevap anahtarı
	Sınıflama gerektiren testler	2	5., 6.	Cevap anahtarı
Verileri kaydetme	Performansa dayalı	3	7., 8., 11.	Dereceli puanlama anahtarı
Önceden kestirme	Çoktan seçmeli	2	10., 13.	Cevap anahtarı
Değişkenleri belirleme ve hipotez kurma	Çoktan seçmeli	6	14., 15., 16., 17., 20.	Cevap anahtarı
	Açık uçlu	2	18., 19.	Dereceli puanlama anahtarı
Verileri yorumlama	Çoktan seçmeli	7	21., 22., 23., 24., 25., 26., 27.	Cevap anahtarı
Sonuç çıkarma	Çoktan seçmeli	3	28., 29., 30.	Cevap anahtarı
Verileri kullanma ve model oluşturma (Grafik çizme)	Performansa dayalı	2	9., 12.	Kontrol listesi
Değişkenleri değiştirme ve kontrol etme	Çoktan seçmeli	2	31., 32.	Cevap anahtarı
Deney tasarlama ve yapma	Çoktan seçmeli	2	33., 34.	Cevap anahtarı
	Performansa dayalı	1	35., 36.	Dereceli puanlama anahtarı

N: Madde sayısı

TARTIŞMA VE SONUÇ

Bu araştırmada üniversite öğrencilerinin BSB'lerini belirlemeye yönelik olarak geliştirilen çoklu formatta test geliştirme çalışmasına yer verilmiştir. Bu amaçla, literatür taraması sonucunda 45 maddelik bir test hazırlanmıştır. Ölçekten elde edilen sonuçların kapsam geçerliğinin sağlanmasında, uzman kanılarına başvuru yapılan çalışmalara literatürde de sıklıkla rastlanılmaktadır (Turan & Demirel, 2009; Feyzioğlu vd., 2012). Bu araştırmada da birinci pilot uygulamadan sonra gerekli düzenlemelerin ardından 42 (28'i çoktan seçmeli ve 14'ü açık uçlu) maddeye düşürülen testin, kapsam geçerliği çalışması için de, uzman görüşlerine başvurulmuş ve bu görüşler sonucunda testin kapsam geçerliği koşulunu sağladığı sonucuna ulaşılmıştır. Araştırma kapsamında geliştirilen BİSBET'in, yapı geçerliğini sınamak için yapılan hipotez testi sonucuna göre, BSB'ye dayalı laboratuvar dersi alan öğrencilerin testin hem çoktan seçmeli hem de açık uçlu kısmından aldıkları toplam BSB puanlarının, buna yönelik ders almayan öğrencilerin puanından daha yüksek olduğu saptanmıştır. Başka bir deyişle BSB'ye yönelik eğitim alan öğrenciler testte daha başarılı olmuşlardır. Bu bulgu kurulan hipotezin desteklendiğini ve test maddelerinin yapıyı ölçmeye hizmet ettiğini göstermektedir. Bu anlamda da geliştirilen test yapı geçerliğine bir kanıt oluşturmaktadır. Literatürde de araştırmacıların geliştirdikleri testin yapı geçerliğine kanıt oluşturmada hipotez testi tekniğine sıklıkla başvurduğu sıklıkla rapor edilmektedir (Şahin & Çepni, 2011; Temiz, 2007).

Güvenirlik katsayıları test ve ölçeğin niteliğine göre değişmekle birlikte, Alfa güvenilirlik katsayısının 0.70 ve yukarı olması test maddelerinin güvenilirliği için genel olarak yeterli

görülmektedir (Büyüköztürk, 2007). Şencan (2005) da bilimsel araştırmalarda güvenilirlik katsayısının büyüklüğü için alt sınırı 0.70 olarak belirtmiştir. Özdamar (2004), uygulanan bir ölçeğin alfa güvenilirlik katsayısının değerlendirilmesinde uyulan değerlendirme ölçütlerini belirtirken, $0.60 \leq \alpha < 0.80$ ranjını “oldukça güvenilir” olarak nitelendirmiştir. Bu araştırma kapsamında geliştirilen BİSBET’in hesaplanan Cronbach alfa güvenilirlik katsayılarının tümü 0.70 ve üzerinde değerler almaktadır. Ayrıca BİSBET’in çoktan seçmeli kısmının güvenilirliğine, güvenilirlik hesaplamalarının alternatif formu olan KR-21 değeri de hesaplanmış ve 0.75 olarak bulunmuştur. Alfa katsayısı iç tutarlılık güvenilirliğinin üst sınırını oluşturur. Alfa değeri eğer yüksek çıkmışsa alternatif güvenilirlik formlarından elde edilen korelasyon katsayıları da yüksek çıkar (Şencan, 2005: 163). Büyüköztürk (2007), Şencan (2005) ve Özdamar’ın (2004) güvenilirlik hakkındaki açıklamaları ve KR-21 değerinin 0.70’den büyük olması gerektiği koşulları dikkate alındığında, son hali 25 çoktan seçmeli ve 11 açık uçlu maddeden oluşan BİSBET’in güvenilir bir test olduğu sonucuna varılabilir. Benzer şekilde literatürde BSB ölçmek amacıyla geliştirilen testlerinde Cronbach alfa güvenilirlik katsayısının 0.70 ve üzeri olduğu ve güvenilir kabul edildiğini görülmektedir (Temiz, 2007; Monica, 2005; Aydoğdu & Ergin, 2009; Feyzioğlu vd., 2012).

İki veya daha fazla gözlemcinin yaptığı değerlendirmeler arasındaki tutarlılığı belirlemek için Kappa istatistiği kullanılır. Kappa katsayısı -1 ile +1 arasında değer alır. Sıfır değeri tesadüfi uyuşmayı, negatif değerler tesadüfi olmaktan daha kötü bir uyuşmayı, +1 ise mükemmel uyuşmayı temsil eder. Kappa katsayısı 0.40 ile 0.75 arasında ise makul bir uyuşma, 0.75 ve daha büyük ise mükemmel bir uyuşma olduğu anlamına gelmektedir (Şencan, 2005: 267). Bu çalışmada hesaplanan Kappa istatistiği sonuçları 0.72 ile 0.85 arasında değişmektedir. Kappa istatistik sonuçlarından gözlemciler arasında “makul” ve “mükemmel” arasında değişen uyuşmaların olduğu, başka bir deyişle gözlemci puanları arasında anlamlı bir fark olmadığı sonucuna ulaşılmaktadır.

Madde analizi sonucunda ayırt edicilik değerlendirirken aşağıdaki ölçütlere dikkat edilmiştir: Ayırt ediciliği sıfır veya negatif olan maddeler teste dâhil edilmez. Ayırt edicilik indisi 0.40 veya daha yüksek olan bir madde “çok iyidir ve düzeltilmesi gerekmez”; 0.30-0.40 arasında ise “iyidir ve düzeltilmesi gerekmez”. 0.20-0.30 arasında bir madde ise “zorunlu hallerde aynen kullanılabilir veya değiştirilebilir”; 0.20’den daha küçük bir değerde ise “madde kullanılmamalıdır veya yeniden düzenlenmelidir” (Turgut, 1992). BİSBET’in madde analiz sonuçları, testin ayırt edicilik gücünün oldukça iyi olduğu ve maddelerin kullanılabilirliğini göstermektedir.

Sonuç olarak geliştirilen BİSBET’in fen ve teknoloji dersi öğretmen adaylarının BSB’lerinin ölçülmesi amacıyla kullanılabilir, geçerliği ve güvenilirliği sağlanmış bir ölçek olması özellikleri ile araştırmacılar tarafından çalışmalarında rahatlıkla kullanılabilir.

ÖNERİLER

Bu çalışma fen bilgisi öğretmenliği öğrencilerinden oluşan bir örneklem üzerinde yürütülmüştür. Bundan sonraki çalışmalarda ölçeğin orta öğretim öğrencileri ve farklı bölümlerde öğrenim gören lisans öğrencileri üzerinde de geçerlik güvenilirlik çalışmaları yapılabilir. Ayrıca ölçeğin öğrenci kavram başarısını ölçen geçerli ve güvenilir ölçeklerle ilişkisine bakılabilir. Araştırmada geliştirilen BİSBET’te hem çoktan seçmeli hem de açık uçlu kısımların geçerlik ve güvenilirlik kanıtları ayrı ayrı yapıldığı için test mevcut haliyle kullanılabilirliği gibi çoktan seçmeli ya da açık uçlu kısımları ayrı ayrı da kullanılabilir. BİSBET’te fizik, kimya ve biyoloji konu başlıklarından seçilen deney senaryolarına ve kavramlarına ek olarak bu konu başlıklarının dışında konular da ele alınarak farklı sorular geliştirilebilir ve bu sorulardan BSB ölçme soru havuzu oluşturulabilir.



A Study on the Development of a Test about Measurement of Science and Technology Teachers' Science Process Skills

Fethiye KARSLI¹, Alipaşa AYAS²

¹ Assist.Prof.Dr., Giresun University, Educational Faculty, Giresun-TURKEY

² Prof.Dr., Bilkent University, Educational Faculty, Ankara-TURKEY

Received: 02.02.2012

Revised: 20.04.2013

Accepted: 01.06.2013

The original language of article is Turkish (v.10, n.2, June 2013, pp.66-84)

Key Words: Scientific Process Skills; Test Development; Validity; Reliability.

SYNOPSIS

INTRODUCTION

Which one is more important: Giving the scientific content to the students or Scientific Process Skills (SPS) in science education? Academically the scientific content requires knowing scientific knowledge and SPS requires doing science. Basically learning one of them helps to learn the other one (Ayas et.al., 1997). For this reason, both scientific content and SPS have the same importance in Science Curricula. The main goals for students' learning are teaching, learning scientific research methods and gaining SPS are highlighted in the curricula of science and technology, physics, chemistry and biology that are rearranged both in Turkey and American National Science Education Standards. There has been found a positive relationship among academically success of the students (Beaumont-Walters & Soyibo, 2001; Kesamang & Taiwo, 2002) and attitudes toward science and technology (Downing & Filer, 1999) and scientific creativeness with SPS (Aktamış & Ergin, 2007; Roberts, 2003) in the literature. Many researcher have developed activities encouraging the students to gain SPS (Karlı, 2011; Duran & Özdemir, 2010; Karlı & Şahin, 2009; Tifi, Natale & Lombardi, 2006). Education is known as an important mean to help students develop competencies stated in the curriculum. Equally important is to identify the degree to which students attain the competencies or SPS during the instruction. Therefore, there is a need for valid, reliable and suitable measurement instruments to measure students' SPS in the learning process.

When the literature is examined, it is observed that tests to measure students' SPS have been developed in 1960s. However, if these tests are examined then it is observed that these tests are generally developed to measure students' SPS in elementary or secondary schools, and there is no tests to be developed and used for the students in higher education. Another



point is that these tests are usually in multiple choice format. Also, the tests used in Turkey are transferred from foreign countries after translation and validation (Saraçoğlu et.al., 2012; Geban, 1990). In both national and international literature, there is no reliable and validated measurement instrument to test students' SPS at undergraduate level before and after instruction. For these reasons, to develop a valid and reliable instrument will fill the aforementioned gap in the literature in terms of measuring the SPS levels of teacher candidates.

PURPOSE OF THE STUDY

The purpose of this study is to develop and validate a Test of Scientific Process Skills in Multiple Format (TSPSMF) to be useful to measure developmental level of science teacher candidates' SPS in Science and Technology.

METHODOLOGY

The sample of this research are composed of 197 science teacher candidates trained in Elementary Science Education Department of two universities namely Karadeniz Technical (first pilot study; N=87) and Giresun Universities (second pilot study; N=110) in the fall term of 2010-2011.

FINDINGS

In this study, during developmental process of SPS tests a multiple form is used because it is suggested by program developers in recent years as a means of fair and adequate assessment. Therefore, to develop this test, firstly what the SPS are have been determined from the related literature. The test consists of SPS as observing, measuring, classifying, data recording, using data to develop models, prediction, defining, changing and controlling variables, making a hypothesis, interpreting the data, drawing conclusion, designing and carrying experiment. After determining the attributes to be measured, the next step is that to decide what type of question is necessary to measure these skills. Then if the skills related with performance essay type questions are prepared and if other types of skills are to be measured, then other types of questions such as multiple choice are prepared. As a result a 45 item test was developed. At first step, the test is applied to the 87 students to have some sense on the understandability of figures and drawings in the tests and readability of test items. Also, the total time required for answering the questions is determined. After the first pilot, 3 items that are not understood by students removed from the test. To know about content validity of the test, experts' views were taken. The experts stated that 38 items except 1 multiple choice and 3 essay type questions are suitable for the students' level, and includes the SPS attributes, there was no error noticed in the scenario of test items and scientific error in the options.

To test the construct validity of TSPSMF, hypotheses test techniques was used. The results showed the hypotheses can be acceptable and test items can measure the structure correctly as intended. To test the reliability of this test, internal consistency analysis and consistency among observers were done. Internal consistency of the multiple choice items was done with internal consistency reliability analysis, Cronbach Alpha and Kuder Richardson-21 (KR-21) reliability, and Cronbach Alpha coefficient was used for essay type and Kappa coefficient for essay type test items based on the performance. Consistency among categorical points of the observers was calculated with Kappa coefficient. Cronbach Alpha coefficient for multiple choice test items was 0.78, essay type test items required classification was 0.71, test items for saving the data was 0.83, test items for drawing graphs 0.90, test items

for determining the variables and making hypotheses 0.71 and test items for doing experiments was 0.90. Moreover, for reliability of multiple choice part of this test, alternative form KR-21 was calculated and it was 0.75. These values prove all forms of the TSPSMF are reliable. From the Kappa statistic between two observers independent from each other for essay type questions, the least Kappa coefficient was 0.72 and the highest was 0.85. This case shows that there is a logical and excellent agreement between the observers. In addition to validity and reliability of this test, from the results of item analysis of multiple choice part of this test, distinctiveness power of the test is quietly good and the items can be used.

DISCUSSION and CONCLUSION

In conclusion, TSPSMF developed in this study can be used with the aim of measurement of SPS of science teacher candidates and it is a contribution to the literature as it was found reliable and valid.

KAYNAKLAR/REFERENCES

- Aktamış, H. & Ergin, Ö. (2007). Investigating the Relationship Between Science Process Skills and Scientific Creativity. *Hacettepe University Journal of Education*, 33, 11-23.
- Aktamış, H., 2007. *Fen Eğitiminde Bilimsel Süreç Becerilerinin Bilimsel Yaratıcılığa Etkisi: İlköğretim 7. Sınıf Fizik Ünitesi Örneği*, Doktora Tezi, Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.
- Ayas, A., Çepni, S., Johnson, D. & Turgut, M.F. (1997). *Kimya öğretimi, Kimya Öğretimi, Öğretmen Eğitimi Dizisi*, YÖK/Dünya Bankası Milli Eğitimi Geliştirme Projesi Yayınları, Bilkent, Ankara.
- Aydoğdu, B. & Ergin, Ö. (2009). The Development of Science Process Skills Scale Toward "Electricity in Our Life" Units of Science and Tehnology Course. *e-Journal of New World Sciences Academy, Education Sciences*, 4(2), 296-316.
- Bağcı Kılıç, G., Haymana, F. & Bozyılmaz, B. (2008). İlköğretim Fen ve Teknoloji Dersi Öğretim Programı'nın Bilim Okuryazarlığı ve Bilimsel Süreç Becerileri Açısından Analizi. *Eğitim ve Bilim*, 33, 52-63.
- Beaumont-Walters, Y. & Soyibo, K. (2001). An Analysis of High School Students' Performance on Five Integrated Science Process Skills. *Journal of Research in Science & Technological Education*, 19, 2.
- Berger, C. F. (1982). Attainment of Skill in Using Science Processes: Instrumentation, Methodology and Analysis. *Journal of Research in Science Teaching*. 19(3), 249-260.
- Bozkurt, O. & Olgun, Ö. S. (2005). *Fen ve Teknoloji Eğitiminde Bilimsel Süreç Becerileri*. M, Aydoğdu ve T. Kesercioğlu (Eds), *İlköğretimde Fen ve Teknoloji Öğretimi*, Ankara: Anı Yayıncılık.
- Burns, J. C., Okey, J. R. & Wise, K. C. (1985). Development of an Integrated Process Skill Test: TIPS II. *Journal of Research in Science Teaching*, 22(2), 169-177.
- Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı* (7. Baskı). Ankara: PegemA Yayıncılık.
- Çalışkan, İ. Ö. & Kaptan, F. (2009). Constructing Science Process Skills Test. *Çağdaş Eğitim Dergisi*, 34, 369, 27-34.
- Çepni, S. (2007). *Araştırma ve proje çalışmalarına giriş* (3. Baskı). Trabzon: Celepler Matbaacılık.
- Dillashaw, F. G. & Okey, J. R. (1980). Test of the Integrated Science Process Skills for Secondary Students, *Science Education*. 64, 601-608.
- Downing, J. E. & Filer, J. D. (1999). Science Process Skills and Attitudes of Preservice Elementary Teachers. *Journals of Elementary Science Education*. 11, 2, 57-64.
- Duran, M. & Özdemir, O. (2010). The Effects of Scientific Process Skills-Based Science Teaching on Students' Attitudes Towards Science. *US-China Education Review*, 7(3), 17- 28.
- Enger, S. K. & Yager, R. E. (Eds.). (1998). *Iowa Assessment Handbook*. Iowa City: Science Education Center, The University of Iowa.
- Farsakoğlu, Ö. F., Şahin, Ç., Karşlı, F., Akpınar, M. & Ultay, N. (2008). A Study on Awareness Levels on Prospective Science Teachers on Science Process Skills in Science Education. *World Applied Sciences Journal*, 4(2), 174-182.
- Feyzioğlu, B., Demirdağ, B., Akyıldız, M. ve Altun, E. (2012). Ortaöğretim Öğrencilerine Yönelik Bilimsel Süreç Becerileri Testi Geliştirilmesi: Geçerlik ve Güvenirlik Çalışması, *Kuram ve Uygulamada Eğitim Bilimleri*, 12(13), 1887-1906.
- Fyffe, D. W. (1971). "The Development of test items for the integrated science processes: formulating hypothesis and defining operationally." Unpublished doctoral dissertation, Michigan State University, East Lansing, Michigan.

- Geban, Ö. (1990). "Effects of two different instructional treatments on the students' chemistry achievement, science process skills and attitudes towards chemistry at the high school level." Unpublished doctoral dissertation, Middle East Technical University, Ankara.
- Geban, Ö., Aşkar, P. & Özkan, İ. (1992). Effects of computer simulation and problem solving approaches on high school. *Journal of Educational Research*, 86 (1), 5-10.
- Hazır, A. & Türkmen, L. (2008). İlköğretim 5. Sınıf Öğrencilerinin Bilimsel Süreç Beceri Düzeyleri. *S.Ü. Ahmet Keleşoğlu Eğitim Fakültesi Dergisi*, 26, 81-96.
- İnan, H. Z. (2010). Examining Pre-School Education Teacher Candidates' Content Knowledge and Pedagogical Content Knowledge. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 10(4), 2309-2323.
- Karşlı, F. & Şahin, Ç. (2009). Developing Worksheet Based On Science Process Skills: Factors Affecting Solubility, *Asia-Pacific Forum on Science Learning and Teaching*, 10,1, 15.
- Karşlı, F. (2011). *The effect of enriched laboratory guide materials on improving science process skills and conceptual change of prospective science teachers*, Unpublished doctoral dissertation, Karadeniz Technical University; Trabzon.
- Karşlı, F., Şahin, Ç. & Ayas, A. (2009). Determining Science Teachers' Ideas About The Science Process Skills: A Case Study. *Procedia Social and Behavioral Sciences*, 1, 890-895.
- Kesamang, M. E. E. & Taiwo, A.A. (2002). The Correlates of the Socio-Cultural Background of Botswana Junior Secondary School Students with Their Attitudes Towards and Achievements in Science. *International Journal of Science Education*, 24, 919-940.
- Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge.
- Ludeman, R. R. (1975). *Development of the science processes test (TSPT)*. Unpublished doctoral dissertation, Michigan State University, MI.
- McLeod, R. J., Berkheimer, G. D., Fyffe, D. W. & Robison, R. W. (1975). The Development of Criterion-Validated Test Items for Four Integrated Science Processes. *Journal of Research in Science Teaching*, 12, 3, 415-421.
- MEB (2006). *Fen ve teknoloji dersi programı*, ilköğretim 6, 7, ve 8. sınıf. Ankara.
- Molitor, L. L. & Kenneth, G. D. (1976). Development of a Test of Science Process Skills. *Journal of Research in Science Teaching*, 13, 5, 405-412.
- Monica, K. M. M. (2005). *Development and Validation of a test Integrated science Process Skills for the further Education and Training Learners*. Master Thesis, University of Pretoria, South Africa.
- Özdamar, K. (2004). *Paket programlar ile istatistiksel veri analizi 1*. (5. Baskı). Eskişehir: Kaan Kitabevi.
- Roberts, L. (2003). Creativity, *Tech Directions*, 63, 3.
- Saraçoğlu, S., Büyük, U. ve Tanık, N. (2012). Birleştirilmiş ve Bağımsız Sınıflarda Öğrenim Gören İlköğretim Öğrencilerinin Bilimsel Süreç Beceri Düzeyleri, *Türk Fen Eğitimi Dergisi*, 9(1), 83-100.
- Smith, K. (1995). *Science Process Assessments for Elementary and Middle School Students*. <http://www.scienceprocesstests.com/> 23 Şubat 2012.
- Solano-Flores, G. (2000). Teaching and Assessing Science Process Skills in Physics: The Bubbles Task. *Science Activities*, 37, 1, 31-37.
- Şahin, Ç. & Çepni, S. (2011). "Yüzme- Batma, Kaldırma Kuvveti ve Basıncı" Kavramları ile İlgili İki Aşamalı Kavramsal Yapılardaki Farklılaşmayı Belirleme Testi Geliştirilmesi. *Journal of Turkish Science Education (TUSED)*, 8, 1, 79-110.
- Şencan, H. (2005). *Sosyal ve Davranışsal Ölçümlerde Güvenirlilik ve Geçerlik*. Ankara: Seçkin Yayıncılık.

- Şenyüz, G. (2008). *2000 Yılı Fen Bilgisi ve 2005 Yılı Fen ve Teknoloji Dersi Öğretim Programlarında Yer Alan Bilimsel Süreç Becerileri Kazanımlarının Tespiti ve Karşılaştırılması*, Yüksek Lisans Tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Tannenbaum, R. S. (1968). *The development of the test of science processes*. Unpublished doctoral dissertation, Columbia University, New York: Teacher's College.
- Tavukçu, F. (2008). *Fen Eğitiminde Bilgisayar Destekli Öğrenme Ortamının Öğrencilerin Akademik Başarı, Bilimsel Süreç Becerileri ve Bilgisayar Kullanmaya Yönelik Tutuma Etkisi*, Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi, Sosyal Bilimler Enstitüsü, Zonguldak.
- Temiz, B. K. (2007). *Fizik öğretiminde öğrencilerin bilimsel süreç becerilerinin ölçülmesi*. Yayımlanmamış doktora tezi, Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Temiz, B. K., Taşar, M. F. & Tan, M. (2006). Development and Validation of a Multiple Format Test of Science Process Skills. *International Education Journal*, 7, 7, 1007-1027.
- Tifi, A., Natale, N. & Lombardi, A. (2006) Scientists at Play: Teaching Science Process Skills. *Science in School*, 1, 37-40.
- Tobin, K. G. & Capie, W. (1982). Development and Validation of a Group Test of Integrated Science Processes. *Journal of Research in Science Teaching*, 19, 2, 133-141.
- Turan, S. & Demirel, Ö. (2009). Probleme Dayalı Öğrenmeye İlişkin Tutum Ölçeği Geçerlik ve Güvenirlik Çalışması. *Eğitim ve Bilim*, 34 (152), 15-29.
- Turgut, M. F (1992). *Eğitimde ölçme ve değerlendirme* (9. Baskı). Ankara: Saydam Matbaacılık.
- Ünal Çoban, G. (2009). *Modellemeye Dayalı Fen Öğretiminin Öğrencilerin Kavramsal Anlama Düzeylerine, Bilimsel Süreç Becerilerine, Bilimsel Bilgi ve Varlık Anlayışlarına Etkisi: 7. Sınıf Işık Ünitesi Örneği*, Doktora Tezi, Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.
- Walbesser, H. H. (1965). *An evaluation model and its application*. Washington, DC: The American Association for the Advancement of Science.