

SEVİYE BELİRLEME SINAVI (SBS) 2010 6.SINIF FEN VE TEKNOLOJİ ALT TESTİ MADDELERİNİN GEÇERLİ KAZANIMLARI ÖLÇME DERECELERİNİN BELİRLENMESİ*

Zekeriya NARTGÜN**
Melike ÖZER***

ÖZET

Bu çalışmanın genel amacı, Milli Eğitim Bakanlığı'nın 2010 yılında uyguladığı Seviye Belirleme Sınavı (SBS) 6. sınıf Fen ve Teknoloji alt testi maddelerinin geçerli kazanımları ölçme derecelerini alan uzmanlarının görüşlerine göre belirlemektir. Çalışmanın araştırma grubunu, Türkiye'nin çeşitli üniversitelerinde görev yapan 10 fen bilgisi eğitimi alan uzmanı oluşturmaktadır. Araştırmada kullanılacak veriler, araştırmacılar tarafından oluşturulan "Seviye Belirleme Sınavı 6. Sınıf Fen ve Teknoloji Alt Testi Uzman Görüşleri Formu" ile sağlanmıştır. Uzman görüşleri formu oluşturulmadan önce, Eğitim Teknolojileri Genel Müdürlüğü'nden (EĞİTEK), SBS 2010 Fen ve Teknoloji alt testindeki her bir maddenin ilköğretim 6. Sınıf Fen ve Teknoloji programında yer alan hangi öğrenme alanını ve hangi kazanımları ölçmeye yönelik yazıldığını gösteren kazanım tablosu alınmıştır. Uzman görüşleri formunda fen bilgisi eğitimi alan uzmanlarından, maddelerin listede yer alan kazanımların her birini ölçme derecesini "tamamen ölçüyor", "belirsiz" ve "kesinlikle ölçmüyor" seçenekleri ile belirtmeleri istenmiştir. Testte yer alan her bir madde için uzmanların verdiği puanların ortalaması ve madde-hedef uyum katsayıları hesaplanmıştır. Araştırma sonucunda elde edilen bulgulara göre uzmanlar, SBS Fen ve Teknoloji alt testinde yer alan 16 maddeden 10'unun geçerli kazanımları ölçtüğü konusunda hemfikir olup; diğer kazanımları ölçmediği konusunda görüş birliğine varamamışlardır.

Anahtar sözcükler: Geçerlik, Kapsam geçerliği, Madde-Hedef Uyumu, Seviye Belirleme Sınavı (SBS), Fen ve Teknoloji.

DETERMINING THE DEGREE OF MEASUREMENT OF VALID OUTCOMES IN THE 2010 6TH GRADE SCIENCE AND TECHNOLOGY SUBTEST ITEMS IN THE HIGH SCHOOL PLACEMENT EXAM (SBS)

ABSTRACT

The general aim of this study is to determine the degree of measurement of valid outcomes in the 6th grade science and technology subtest items in the high school placement exam administrated by the Ministry of National Education in 2010, based on the views of experts in the field. The research sample of the study consists of 10 experts in the field of science and technology education who work for various universities in Turkey. The data used in the study were obtained via the "Expert Opinions Form for the 6th Grade Science and Technology Subtest in the High School Placement Exam" developed by the researchers. Before the Expert Opinions Form was developed, an outcomes table was obtained from the General Directorate of Educational Technologies (EĞİTEK) that indicated which particular item in the 2010 6th grade science and technology subtest had been written to measure which particular learning domain and which

* Bu makale AİBÜ Eğitim Bilimleri Enstitüsünde tamamlanan tezden türetilmiştir.

** Doç. Dr. Abant İzzet Baysal Üniversitesi. Eğitim Fakültesi. Eğitim Bilimleri Bölümü.

nartgun@yahoo.com

*** Araş. Gör. Artvin Çoruh Üniversitesi. Eğitim Fakültesi. İlköğretim Bölümü.

melikeozer@hotmail.com

particular learning area in the 6th grade science and technology curriculum. In the Expert Opinions Form, the experts in the field of science and technology education were asked to evaluate the extent each subtest item measured the outcomes listed on the table, by choosing “clearly measuring”, “partially measuring,” and “clearly not measuring”. The average of expert scores and item-objective congruence indexes were calculated for each item. Based on the results, the experts expressed agreement that 10 out of the 16 items in the science and technology subtest measured the valid outcomes, yet could not reach agreement on whether or not they measured other outcomes.

Key words: Validity, Content validity, Item-Objective Congruence, High School Placement Exam (HSPE), Science and Technology.

1. GİRİŞ

1.1. Problem Durumu

Ölçme uygulamalarından elde edilen sonuçlar eğitim sistemi içinde çok farklı amaçlara hizmet etmektedir. Bu amaçlardan birkaçı aşağıda sıralanmıştır.

1. Eğitim ve öğretim programlarının onarılması bakımından ilgililere fırsatlar sunar (Ertürk, 1984),
2. Öğrencilerden beklenen bilgi, beceri ve tutumların gelişip gelişmediğinin belirlenmesini (Yetkin ve Daşcan, 2006), bir başka deyişle ulaşılan ya da ulaşılamayan hedef davranışların neler olduğunun tespitini sağlar (Yılmaz, 2010),
3. Eğitim öğretim süreci içindeki eksikliklerinin belirlenmesi ve giderilmesi (Murtha et al., 2006; Özçelik, 2010) ve öğrenciler hakkında geçti-kaldı, başarılı-başarısız vb. kararların verilmesi için gerekli verileri sağlar (Atılgan, 2009).

Ölçme uygulamalarının eğitim sistemi içindeki temel amaçlarından biri de seçme ve yerleştirme çalışmalarında karar vericilere veri sunmasıdır (“What are the diagnostic assessments,” 2009). Özellikle çok sayıda öğrencinin katıldığı Yüksek Öğretme Geçiş Sınavları, Lisans yerleştirme Sınavları, Orta Öğretime Geçiş Sınavları ve Kamu Personeli Seçme Sınavları seçme ve yerleştirme çalışmalarına örnek olarak sunulabilir.

Kısaca “Ölçme, olaylar ya da nesnelere hakkında değerlendirme yapmak ve elde edilen değerlendirme sonuçlarına dayanarak belli kararlar vermek amacıyla yapılır” (Tekin, 2004). Alınan kararların etkililiği, yerindeliği ve isabetliliği, bu kararların alınmasına dayanak oluşturan ölçme sonuçlarının güvenilirliğinin yanı sıra ölçme aracının geçerliği ile de yakından ilgilidir (Stiggins ve diğ., 1992: akt Nitko, 2004).

Testler özellikle çok sayıda bireyin katıldığı seçme ve yerleştirme amaçlı sınavlarda kullanılan ölçme araçlarından biridir. Geçerlik testlerin en önemli psikometrik özelliklerinden biridir. Geçerlik, test puanlarından yapılan özel yorumların uygunluğunu, anlamlılığını ve faydalılığını belirler (Tekindal, 1997). Geçerlik, kullanılan ölçüm aracının ölçülmek istenen özelliğe uygun olması, verilerin ölçülmek istenen özelliği tam olarak yansıtması ve aynı zamanda verilerin amaca yönelik olarak katkıda bulunmasıdır (Baykul, 2010; Şencan, 2005).

Amerikan Psikologlar Derneği tarafından yapılan sınıflamada geçerlik yordama geçerliği, yapı geçerliği ve kapsam geçerliği olmak üzere üç başlık altında toplanmaktadır (Cronbach, 1971). Bunlardan bu araştırmaya da konu olan kapsam geçerliği, bir bütün olarak testin ve testteki her bir maddenin amaca ne derece hizmet ettiği ile ilgilidir (Haynes, Richard ve Kubany, 1995; Tekin, 2004). Bir testteki toplam maddelerin ölçülecek davranışı ve konu içeriğini örnekleme derecesi, testteki her bir maddenin ölçmek istediği davranışı ne derece iyi ölçtüğüne bağlıdır. Herhangi bir dersteki başarıyı ölçmek için düzenlenen bir testin, o dersin hedeflerini ve içeriğini yeterince kapsayıp kapsamadığı önemli bir sorundur (Tekin, 2004). Kapsam geçerliği, tanımlanan alanla ilgili çıkarımlar yapan test içeriği örneklerinin incelenmesinde ve sınıflanmasında kullanılır (Cole ve Moss, 1993).

Bir dersin programı, birkaç üniteden ve her bir ünite de iki veya daha fazla konudan oluşur. Ayrıca bir dersin gerçekleştirmek istediği belli sayıda hedefleri, yani öğrencinin davranışında oluşturmak istediği değişiklikler ya da öğrenciye kazandırmak istediği yeni davranışlar mevcuttur. Bir dersteki başarı ölçülmek istendiğinde, o dersin konularının ve öğrenciye kazandırmak istediği davranışların tümü dikkate alınmalıdır (Tekin, 2004). Bu açıdan bakıldığında zaman, kapsam geçerliği, bir testin davranış evrenini temsil etme derecesi olarak tanımlanabilir (Demircioğlu, 2007). “Bir kapsam geçerliği çalışması yapılacaksa, öncelikle ölçme aracının kapsamı beklenen davranışların belirlenmiş olması gereklidir” (Kan, 2009). Başarı testlerinde bu davranışları belirlemek ve kapsam geçerliğini garantilemek için belirtke tablolarından yararlanılmaktadır. Belirtke tablosu, bir boyutunda, hazırlandığı ders ya da alanla ilgili konular ya da içerik, diğer boyutunda ise bu içerikle öğrencilere kazandırılması beklenen hedef ve davranışların yer aldığı iki boyutlu bir tablodur (Demircioğlu, 2007). Eğer bir başarı testi, belirtke tablosundaki davranışlara yönelik hazırlanmamış ve bu davranışların bazılarını kapsamıyorsa bu durum kapsam geçerliğini düşürür. Crocker ve Algina (2006) kapsam geçerliği belirlenirken şu adımların izlenmesi gerektiğini vurgulamışlardır:

1. İlgili konu alanı tanımlanmalıdır.
2. Ölçmeye konu olan ilgili alandaki nitelikli uzmanlar seçilerek bir panel oluşturulmalıdır.
3. Kapsama ilişkin belirlenen hedef davranışları ve ölçme aracının maddelerini eşleştirebilmek için bir çerçeve oluşturulmalıdır.
4. Bu eşleştirme işlemi ile ilgili bilgi toplanmalı ve özetlenmelidir.

“Bir testin kapsam geçerliğini sağlamak çok zor ve zahmetli bir süreçtir. Bununla birlikte kapsam geçerliği diğer geçerlilik türleri ile karşılaştırıldığında en önemli geçerlik türüdür” (Demircioğlu, 2007).

Seviye Belirleme Sınavı (SBS) liselere öğrenci seçmek ve yerleştirmek amacıyla öğretim programındaki kazanımları esas alarak Millî Eğitim Bakanlığı tarafından merkezi olarak yapılan bir sınavdır. Seviye Belirleme Sınavı; zorunlu temel dersler olan Türkçe, Matematik, Fen ve Teknoloji, Sosyal Bilgiler, Yabancı Dil, Din Kültürü ve Ahlak Bilgisi derslerinin içeriğini kapsamaktadır. Sınavda sorulan soru sayıları sınıf düzeyine ve derslere göre farklılık göstermektedir. Altıncı sınıflarda fen ve teknoloji alt testinde yer alan soru sayısı 16’dır (MEB, 2007). Günümüzde liselere geçiş için farklı

bir seçme ve yerleştirme süreci işletilmektedir. Böyle olmakla birlikte, bu çalışma, seviye belirleme ve benzeri amaçlarla kullanılacak testlerin kapsam geçerliğinin testi oluşturan maddelerin ilgili konu alanındaki kazanımlarla uyumlarının tespiti çerçevesinde analiz edilmesine örnek teşkil etmesi ve bu testleri geliştirenlere bilgi sunması bakımından önemli görülmektedir.

1.2. Araştırmanın Amacı

Bu araştırmanın amacı SBS (2010) 6.sınıf fen ve teknoloji alt testi maddelerinin uzman görüşlerine dayalı olarak madde-hedef uyum katsayılarının belirlenmesi ve testte yer alan maddelerin belirtilen öğrenme alanındaki geçerli kazanımları ölçüp ölçmediğinin tespitidir. Bu amacın gerçekleştirilebilmesi için aşağıdaki sorulara cevap aranmıştır.

1. Bir kazanımı ölçmek amacıyla hazırlanan maddeler, geçerli kazanımı ne derece ölçmektedir?
2. İki kazanımı ölçmek amacıyla hazırlanan maddeler, geçerli kazanımı ne derece ölçmektedir?
3. Üç kazanımı ölçmek amacıyla hazırlanan maddeler, geçerli kazanımı ne derece ölçmektedir?

2. YÖNTEM

Bu bölümde araştırma grubu, verilerin toplanması ve analizi ile ilgili açıklamalara yer verilmiştir.

2.1. Çalışma Grubu

Betimsel nitelikli bu araştırmanın çalışma grubu SBS fen ve teknoloji alt testi maddelerinin geçerli kazanımları ne derece ölçtüğü konusunda görüşleri alınan 10 alan uzmanı oluşturmaktadır. Uzmanların tamamı fen bilgisi lisans eğitimine sahiptirler. Çalışma grubundaki uzmanlardan üçünün fen bilgisi eğitimi alanında doktora vardır. Ayrıca fen bilgisi eğitiminde doktora yapan altı, ölçme değerlendirme alanında yüksek lisans yapmış bir uzman bulunmaktadır.

2.2. Veri Toplama Aracı

Araştırmada veri toplama aracı olarak “ SBS 6. Sınıf Fen ve Teknoloji Alt Testi Uzman Görüşleri Formu” kullanılmıştır. Belirtilen araç geliştirilmeden önce Eğitim Teknolojileri Genel Müdürlüğü’nden (EĞİTEK), SBS 2010 fen ve teknoloji alt testindeki her bir maddenin ilköğretim 6. Sınıf fen ve teknoloji programında yer alan hangi öğrenme alanını ve hangi kazanımları ölçmeye yönelik yazıldığını gösteren kazanım tablosu yazılı izinle sağlanmıştır. Buradan hareketle geliştirilen uzman görüşleri formunda her bir madde için, o maddenin ölçmeyi amaçladığı kazanım (geçerli kazanım) da dâhil olmak üzere, maddenin öğrenme alanı ile ilgili kazanımları listelenmiştir. Fakat bu maddelerin geçerli kazanımının hangisi olduğu belirtilmemiştir. Alan uzmanlarından, maddelerin listede yer alan ilgili kazanımların her birini ölçme derecesini “tamamen ölçüyor”, “belirsiz” ve “kesinlikle ölçmüyor” seçenekleri ile belirtmeleri istenmiştir. İlgili seçeneklerin puanlanması “tamamen ölçüyor” için 1, “belirsiz” için 0 ve “kesinlikle ölçmüyor” için -1 biçimindedir.

2.3. Verilerin Analizi

2010 SBS 6. Sınıf fen ve teknoloji alt testinde toplam 16 madde bulunmaktadır. Testte bulunan maddelerin geçerli kazanımları ölçme düzeyini belirlemek amacıyla, fen bilgisi eğitimi alan uzmanlarının maddelerin öğrenme alanında her bir ilgili kazanım için verdikleri cevapların ortalaması ve öğrenme alanındaki her bir ilgili kazanım için madde-hedef uyum katsayısı hesaplanmıştır. Bir kazanımı ölçtüğü belirtilen maddeler için tek boyutlu madde-hedef uyum katsayısı, birden fazla kazanımı ölçtüğü belirtilen maddeler için ise düzeltilmiş madde-hedef uyum katsayısı kullanılmıştır. Bu hesaplamalar yapılırken Rovinelli ve Hambleton (1976) tarafından geliştirilip, Crocker ve Algina (2006) tarafından sadeleştirilen tek boyutlu madde-hedef uyum katsayısı formülü ve Turner ve Carlson (2003) tarafından geliştirilen düzeltilmiş madde-hedef uyum katsayısı formülü kullanılmıştır. İlgili formüller ve açıklamalar aşağıda görüldüğü gibidir.

Tek boyutlu madde-hedef uyum katsayısı formülü;

$$I_{ik} = \frac{N}{2N - 2} (\mu_k - \mu)$$

- I_{ik} : i maddesinin k hedefini ölçüp ölçmediğine ilişkin madde-hedef uyum katsayısı
 N : hedef sayısı ($i= 1, 2, \dots, N$)
 μ_k : i maddesinin k hedefini ölçme derecesine ilişkin uzmanların verdiği puanların ortalaması
 μ : i maddesinin tüm hedeflerine verilen puanların ortalaması

Bu formül bir maddenin sadece bir hedefi ölçtüğü durumlar için geçerlidir. Alan uzmanlarından maddenin hedefle eşleştiği durumlarda (+1), belirsizlik varsa (0), maddenin hedefle eşleşmediği durumlarda ise (-1) şeklinde puanlama yapmaları istenir. Bir madde için madde-hedef uyum katsayısı -1 ile +1 aralığında değer alır. Madde-hedef uyum katsayısının 0,70 ve daha büyük değerleri maddenin ilgili hedefi ölçtüğünün bir kanıtıdır (Rovinelli ve Hambleton, 1976).

Maddelerin kazanımlara uygun olarak hazırlanıp hazırlanmadığının belirlenmesinde I_{ik} 'nin yanı sıra maddelerin geçerli kazanımlarda aldığı ortalama puan (μ_k) da hesaplanabilir. Bir hedef için ortalama da I_{ik} gibi -1 ve +1 arasında değer alır. Turner ve Carlson (2003) değişen bir ölçüt olmakla beraber 0,75 ve daha büyük ortalamının uzmanlar arası görüş birliğini gösterebileceğini belirtmişlerdir.

Bir maddenin birden fazla hedefi ölçtüğü durumlardaki madde-hedef uyum katsayısının hesaplanmasında ise Turner ve Carlson (2003) tarafından geliştirilen düzeltilmiş madde-hedef uyum katsayısı formülü kullanılmıştır. İlgili formülün sadeleştirilmiş hali aşağıda görüldüğü gibidir.

Düzeltilmiş madde-hedef uyum katsayısı formülü;

$$I'_{ik} = \frac{(N)\mu_k - (N - p)\mu}{2N - p}$$

I'_{ik}	: düzeltilmiş madde-hedef uyum katsayısı
N	: hedef sayısı
p	: geçerli hedef sayısı
μ_k	: i maddesinin geçerli hedeflerine ilişkin uzmanların verdiği puanların ortalaması
μ_l	: i maddesinin geçerli olmayan hedeflerine ilişkin uzmanların verdiği puanların ortalaması

3. BULGULAR

3.1. Birinci araştırma sorusuna ilişkin bulgular

Bir kazanımı ölçmek amacıyla hazırlanan 8 maddenin geçerli kazanımlarındaki aldıkları ortalamalar (μ_k), madde-hedef uyum katsayıları (I_{ik}) ve beş madde için düzeltilmiş madde-hedef uyum katsayıları (I'_{ik}) Ek 1'de verilmiştir.

Ek 1 incelendiğinde, maddelerin geçerli kazanımlara ilişkin μ_k değerlerinin 0,87 ile 1,00 arasında; I_{ik} değerlerinin de 0,28 ile 0,61 arasında değiştiği gözlemlenmektedir. 8 maddenin tümünde geçerli kazanıma ilişkin μ_k değerlerinin 0,75'in üzerine çıktığı görülmektedir. I_{ik} değerlerine göre ise hiçbir maddenin 0,70 ölçütünün üzerine çıkamadığı görülmektedir. μ_k değerlerine göre, uzmanlar 8 maddenin tümünün de geçerli kazanımı ölçtüğü konusunda hemfikirdirler. 8 maddenin madde-hedef uyum katsayısı değerlerine baktığımız zaman uzmanlara göre maddeler geçerli kazanımı ölçmektedir. Fakat diğer kazanımları ölçmediğine dair görüş birliği sağlanamamıştır. Buna göre, uzmanların 8 maddenin geçerli kazanımları ölçtüğü bununla birlikte öğrenme alanı ile ilgili kazanımlardan bazılarını da ölçtüğü görüşünde olduğu sonucuna varılabilir. Ancak madde-hedef uyum katsayıları, uzmanların hangi maddelerin, öğrenme alanı ile ilgili verilen kazanımlardan hangilerini ölçtüğü konusunda görüş birliğine varamadıklarını göstermektedir.

Uzmanlardan elde edilen ortalamalar incelendiğinde 10. madde hariç, tüm maddelerin geçerli kazanımla birlikte en az iki kazanımda pozitif ortalama aldığı görülmektedir. 10. maddeye baktığımız zaman, bu maddenin geçerli kazanımda pozitif yüksek ortalama ve madde-hedef uyum katsayısı değerine sahip olduğu görülmektedir. Bu durum, 10. maddenin geçerli kazanımı ölçtüğü sonucunu göstermektedir.

Ek 1'de 1., 6. ve 9. maddenin geçerli kazanımda ve başka bir kazanımda μ_k değerinin aynı olması, bu maddelerin geçerli kazanımla birlikte başka bir kazanımı ölçtüğünü göstermektedir. Ayrıca 7. ve 11. maddelerin geçerli kazanımlarındaki μ_k değerlerinin, başka bir kazanımdaki μ_k değerinden düşük olduğu gözlenmiştir. Bu durum, uzman görüşlerinden elde edilen ortalamalara göre 7. ve 11. maddelerin geçerli kazanımdan başka bir kazanımı ölçtüğü sonucunu ortaya koymaktadır. 7. ve 11. maddelerin, iki kazanımı birden ölçüp ölçmediğini anlamak için (I'_{ik}) değerleri hesaplanmıştır. Beş madde için hesaplanan çok boyutlu madde-hedef uyum katsayıları, bu maddelerin tek boyutlu madde-hedef uyum katsayılarından yüksek olmasına rağmen yine de 0,70 ölçütüne ulaşamamıştır.

3.2. İkinci araştırma sorusuna ilişkin bulgular

İki kazanımı ölçmek amacıyla hazırlanan 5 maddenin geçerli kazanımlarındaki aldıkları ortalamalar (μ_k), madde-hedef uyum katsayıları (I_{ik}) ve beş madde için düzeltilmiş madde-hedef uyum katsayıları (I'_{ik}) Ek 2'de verilmiştir.

Ek 2 incelendiğinde, maddelerin geçerli kazanımlarına ilişkin μ_k değerlerinin 0,12 ile 1,00 arasında; I_{ik} değerlerinin de -0,13 ile 0,51 arasında değiştiği görülmektedir. 2 maddenin (3. ve 4. maddeler) geçerli kazanımlarına ait μ_k değerleri 0,75'in üzerindedir. I'_{ik} değerlerine baktığımız zaman, hiçbir madde 0,70 ölçütünü karşılamamaktadır. μ_k değerlerine göre, uzmanlar 5 maddenin 2'sinin geçerli kazanımı ölçtüğünü belirtmişlerdir. Bununla birlikte uzmanlar belirtilen iki maddenin geçerli kazanımları ölçtüğünü belirtmekle beraber; diğer kazanımları ölçmediği hususunda görüş birliğine varamamışlardır. Buna göre, uzmanların 2 maddenin geçerli kazanımları ölçtüğü bununla birlikte öğrenme alanı ile ilgili kazanımlardan bazılarını da ölçtüğü görüşünde olduğu sonucuna varılabilir. Ancak madde-hedef uyum katsayıları, uzmanların hangi maddelerin, öğrenme alanı ile ilgili verilen kazanımlardan hangilerini ölçtüğü konusunda görüş birliğine varamadıklarını göstermektedir.

Uzmanlardan elde edilen ortalamalar incelendiğinde 13. ve 15. maddeler hariç, tüm maddelerin geçerli kazanımla birlikte en az iki kazanımda pozitif ortalama aldığı görülmektedir. Aynı zamanda belirtilen bu iki madde, en yüksek madde-hedef uyum katsayısı değerine sahip olan maddelerdir. Bu durum, 13. ve 15. maddenin geçerli kazanımları ölçtüğü sonucunu göstermektedir.

Ek 2'de 3. maddenin geçerli kazanımlarda ve başka bir kazanımda μ_k değerinin 1,00 olması, bu maddenin geçerli kazanımlarla birlikte başka bir kazanımı ölçtüğünü göstermektedir. Ayrıca 2. maddenin geçerli kazanımlardan birisi olan altıncı kazanımın μ_k değerinin, başka bir kazanımdaki μ_k değerinden düşük olduğu gözlenmiştir. Bu durum, uzman görüşlerinden elde edilen ortalamalara göre 2. maddenin geçerli kazanımlardan başka bir kazanımı ölçtüğü sonucunu ortaya koymaktadır.

3.3. Üçüncü araştırma sorusuna ilişkin bulgular

Üç kazanımı ölçmek amacıyla hazırlanan 3 maddenin geçerli kazanımlarındaki aldıkları ortalamalar (μ_k), madde-hedef uyum katsayıları (I_{ik}) ve üç madde için düzeltilmiş madde-hedef uyum katsayıları (I'_{ik}) Ek 3'de verilmiştir.

Ek 3 incelendiğinde, maddelerin geçerli kazanımlarına ilişkin μ_k değerlerinin -0,75 ile 0,87 arasında; I_{ik} değerlerinin de -0,61 ile 0,68 arasında değiştiği görülmektedir. 8., 14. ve 16. maddelerin geçerli kazanımlarındaki μ_k değerlerinden bazıları 0,75 ölçütünü karşılamaktadır. Bunlar 8.maddenin altıncı ve sekizinci, 14. maddenin dördüncü ve 16. maddenin yirminci kazanımlarıdır. I'_{ik} değerlerine baktığımızda, hiçbir maddenin 0,70 ölçütünü karşılamadığı görülmektedir. μ_k değerlerine göre, uzmanlar 3 maddenin 3'ünün de geçerli kazanımı tam olarak ölçmediğini belirtmişlerdir. Bununla birlikte, bu 3 maddenin madde-hedef uyum katsayıları incelendiğinde, uzmanların bu maddelerin geçerli kazanımları ölçtüğünü belirtmekle beraber; diğer kazanımları ölçmediği hususunda görüş birliği oluşturmadıkları görülmektedir. Buna göre, uzmanların 3 maddenin geçerli kazanımları ölçtüğü bununla birlikte öğrenme alanı ile ilgili kazanımlardan bazılarını da ölçtüğü görüşünde olduğu sonucuna varılabilir. Ancak madde-hedef uyum katsayıları, uzmanların hangi maddelerin, öğrenme alanı ile ilgili

verilen kazanımlardan hangilerini ölçtüğü konusunda görüş birliğine varamadıklarını göstermektedir.

Ek 3'de 8. maddenin geçerli kazanımlarının dışındaki bir kazanımın μ_k değerinin 1,00 olması, bu maddenin geçerli kazanımları değil başka bir kazanımı ölçtüğünü göstermektedir. Ayrıca 14. maddenin geçerli kazanımlarından olan ikinci ve üçüncü kazanımlarıyla, 16. maddenin geçerli kazanımlarından olan on birinci ve on üçüncü kazanımın μ_k değerinin, başka bir kazanımdaki μ_k değerinden düşük olduğu gözlenmiştir. Bu durum, uzman görüşlerinden elde edilen ortalamalara göre 14. ve 16. maddenin geçerli kazanımlardan başka bir kazanımı ölçtüğü sonucunu ortaya koymaktadır.

4. SONUÇ ve ÖNERİLER

4.1. Sonuç

Bu çalışmada, 2010'da ilköğretimin 6. sınıflarına uygulanan SBS'deki fen ve teknoloji alt **testi sorularının kapsam geçerliği uzman** görüşlerine dayanarak incelenmiş ve değerlendirilmiştir. Ortalaması en az 0,75 olan maddelerin belirtilen kazanımı ölçtüğü, madde-hedef uyum katsayısı en az 0,70 olan maddelerin de belirtilen kazanımı ölçtüğü, diğer kazanımları ölçmediği kabul edilmiştir (Rovinelli ve Hambleton, 1977; Turner ve Carlson, 2003). Araştırmada elde edilen sonuçlar aşağıdaki gibidir:

1. SBS fen ve teknoloji alt testinde yer alan 16 maddenin 8'i bir kazanımı ölçmek amacıyla hazırlanmıştır. Bu 8 maddeden 8'inde de ortalamalar 0,75 ölçütünü karşılamaktadır. Ancak 8 maddenin hiçbiri 0,70 ve üstü madde-hedef uyum katsayısı ölçütünü karşılayamamıştır.
2. SBS fen ve teknoloji alt testinde yer alan 16 maddenin 5'i iki kazanımı ölçmek amacıyla hazırlanmıştır. Bu 5 maddeden 2'sinde ortalamalar 0,75 ölçütünü karşılamaktadır. Ancak 5 maddenin hiçbiri 0,70 ve üstü düzeltilmiş madde-hedef uyum katsayısı ölçütünü karşılayamamıştır.
3. SBS fen ve teknoloji alt testinde yer alan 16 maddenin 3'ü üç kazanımı ölçmek amacıyla hazırlanmıştır. Bu 3 maddeden 3'ünde de ortalamalar 0,75 ölçütünü karşılayamamaktadır. Bununla birlikte 3 maddenin hiçbiri 0,70 ve üstü düzeltilmiş madde-hedef uyum katsayısı ölçütünü karşılayamamıştır.

Araştırmada ulaşılan bu sonuçlara göre, 2010'da ilköğretimin 6. sınıflarına uygulanan SBS'deki fen ve teknoloji alt testinde yer alan soruların geçerli kazanımları ölçtüğü, geçerli kazanımlarla birlikte öğrenme alanında listelenen diğer kazanımları ölçmediği konusunda uzmanlar görüş birliği sağlayamamışlardır. Bu durum, soruların geçerli kazanımlar dışında diğer kazanımları da ölçtüğü sonucunu verebilir. Bununla birlikte bazı maddelerde, bazı kazanımların geçerli kazanımla aynı ya da daha yüksek bir ortalamaya sahip olduğu görülmüştür. Bu bize SBS fen ve teknoloji alt testinde yer alan soruların birden fazla kazanımı ölçtüğünü göstermektedir. Birinci, ikinci ve üçüncü alt problem kapsamında ulaşılan sonuçlar, Kelecioğlu, Atalay ve Öztürk (2010)'ün çalışmaları ile benzerlik göstermektedir.

4.2. Öneriler

1. Kapsam geçerliğinin incelenmesi, sınav uygulanmadan önce yapılması gereken bir işlemdir. Bu sebeple EĞİTEK, her yıl soruların yazılması tamamlandıktan sonra alan uzmanlarının ve öğretmenlerin görüşlerinden faydalanarak madde-hedef uyumu çalışmasını gerçekleştirmelidir.
2. Bu çalışma SBS (2010) 6. sınıf fen ve teknoloji alt testinin kapsam geçerliğini incelemek amacıyla yapılmıştır. Diğer yıllarda yapılan SBS'nin diğer alt testlerine yönelik çalışmalar yapılabilir.

KAYNAKÇA

- Atılgan, H. (Ed.). (2009). *Eğitimde Ölçme ve Değerlendirme* (ss. 349-395). Ankara: Anı Yayıncılık.
- Baykul Y. (2010). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması (2.Baskı)*. Ankara: Pegem Akademi.
- Cole, N. S. ve Moss, P. A. (1993). Bias in test use. Robert L. Linn (eds.), *Educational Measurement* içinde (ss. 201-219). Phoenix: Oryx Press.
- Crocker, L ve Algina, J. (2006). *Introduction to classical and modern Test Theory*. Mason: Thomson Wadsworth.
- Cronbach, L.J. (1971). Test Validation. R.L.Thorndike (Ed). *Educational Measurement* içinde. Washington: American Council on Education.
- Demircioğlu, G. (2007). Geçerlik ve güvenilirlik. Emin Karip (Ed.), *Ölçme ve değerlendirme* içinde (ss.51-79). Ankara: Pegem A Yayıncılık.
- Ertürk, S. (1984). *Eğitimde program geliştirme*. Ankara: Yelkentepe Yayınları.
- Haynes, S. N., Richard, D. C. S., ve Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238-247.
- Kelecioğlu, H., Atalay, K. ve Öztürk, N. (2010). Seviye belirleme sınavı 7.sınıf matematik alt testinin madde-hedef uyumu açısından incelenmesi. *Eğitimde Ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 37-43.
- MEB. (2007). 64 soruda oges, ve seviye belirleme sınavı örnek sorular 02.05.2013 tarihinde <http://www.meb.gov.tr/duyurular/duyurular2007/64sorudaoges/oges64soru.pdf> sitesinden indirildi.
- Murtha, L., Stec, E., & Wilt, M. (2006). Using assessment as a tool to improve learning. An IFLA Workshop. *IFLA Journal*, 32(4), 294-3009.
- Nitko, A. J. (2004). *Educational assessment of students*. New Jersey, NJ: Pearson Merrill Prentice Hall.
- Özçelik D. A. (2010). *Ölçme ve değerlendirme (3.Baskı)*. Ankara: Pegem Akademi
- Rovinelli, R. J., & Hambleton, R. K. (1976). On the use of content specialists in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik*. Ankara: Seçkin Yayıncılık.
- Tekin H. (2004). *Eğitimde ölçme ve değerlendirme (17.Baskı)*. Ankara: Yargı Yayınevi

- Tekindal, S. (1997). *Ölçme ve değerlendirmeye ilişkin tutum aracı geliştirme*. Amasya: Cem Ofset.
- Turner, R. C. ve Carslon, L. (2003). Indexes of item-objective congruence for multidimensional items. *International journal of testing*. 3(2), 163-171
- Yetkin D. ve Daşcan Ö. (2006). *Son değişikliklerle ilköğretim programı (1-5 Sınıflar)*. Ankara: Anı Yayınları.
- Yılmaz, H. (2010). *Eğitimde ölçme ve Değerlendirme (8.Baskı)*. Konya: Çizgi Kitapevi.
- What are the diagnostic assessments. (2009). *Diagnostic test support*. Ohio. Retrieved FEBRUARY 12, 2009 from <http://www.diagnostictestsupport.org/about/whatare.asp>.

EK 1. Bir kazanımı ölçtüğü belirtilen maddelerin μ_k , I_{ik} ve I'_{ik} değerleri																									
Kazanımlar																									
Madde No	1		2		3		4		5		6		7		8		9		10		11		12		
	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	I'_{ik}
1	0,75	0,13	1,00	0,28	0,62	0,06	0,62	0,06	1,00*	0,28	-0,87	-0,83													0,46
5	0,00	0,06	0,50	0,33	0,75	0,46	1,00*	0,59	0,12	0,12	-0,12	-0,00	-0,87	-0,41	-0,37	-0,14	-1,00	-0,49	-0,87	-0,41	0,00	0,06	-0,50	-0,21	
6	-1,00	-0,68	-0,87	-0,60	0,50	0,22	1,00	0,52	1,00*	0,52															0,65
7	0,87	0,30	-0,75	-0,62	0,87*	0,30	-0,50	-0,48	0,87	0,30	1,00	0,38	0,50	0,09	-0,12	-0,26									0,40
9	0,87*	0,38	0,37	0,08	0,87	0,38	0,75	0,31	-0,50	-0,44	-1,00	-0,74													0,66
10	-1,00	-0,8	0,87*	0,37	0,00	-0,18	-0,75	-0,64	-0,50	-0,49															
11	1,00	0,58	-0,62	-0,44	-0,50	-0,36	-0,37	-0,28	0,87*	0,49															0,58
12	1,00*	0,61	-0,50	-0,33	-0,50	-0,33	0,75	0,45	-0,62	-0,41															

*Geçerli kazanımlar

EK 2. İki kazanımı ölçtüğü belirtilen maddelerin μ_k , I_{ik} ve I'_{ik} değerleri																										
Madde No	Kazanımlar																									
	1		2		3		4		5		6		7		8		9		10		11		12		I'_{ik}	
	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}		
2	0,37	0,02	0,25	-0,05	-0,50	-0,48	0,87	0,30	0,75	0,23	0,12*	-0,13	0,00	-0,19	0,87*	0,30										
3	1,00*	0,41	1,00*	0,41	1,00	0,41	-0,87	-0,68	0,00	-0,18	-0,50	-0,47	0,50	0,12												0,58
4	0,50	0,03	1,00*	0,31	0,50	0,03	0,87	0,24	0,87*	0,24	-0,50	-0,54	0,75	0,17	-0,37	-0,47										0,44
13	0,37*	0,51	-1,00	-0,34	0,25*	0,44	-0,87	-0,26	-1,00	-0,34																0,41
15	-0,62	0,02	0,25*	0,5	-1,00	-0,19	-0,87	-0,11	-0,75	-0,05	-1,00	-0,19	-1,00	-0,19	-1,00	-0,19	-0,25	0,22	0,25*	0,5	-0,87	-0,11	-1,00	-0,19		0,45

*Geçerli kazanımlar

Madde No	Ek 3. Üç kazanımı ölçtüğü belirtilen maddelerin μ_k , I_{ik} ve I'_{ik} değerleri																																				
	Kazanımlar																																				
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20																	
	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	I'_{ik}																
8	0,62	0,01	-0,62	-0,70	0,87	0,15	0,62	0,01	1,00	0,22	0,87*	0,15	0,62*	0,01	0,87*	0,15					0,35																
14	0,50	0,23	0,12*	-0,03	-0,75*	-0,61	0,75*	0,39													0,01																
16	-1,00	-0,24	-1,00	-0,24	-1,00	-0,24	-1,00	-0,24	-1,00	-0,24	-1,00	-0,24	-0,62	-0,04	-0,75	-0,11	-0,50*	0,02	-0,12	0,22	0,50*	0,55	0,62	0,61	0,62	0,61	-1,00	-0,24	-1,00	-0,24	-0,87	-0,17	-0,37	0,09	0,75*	0,68	0,40

*Geçerli kazanımlar