

Seviye Belirleme Sınavı 7. Sınıf Matematik Alt Testinin Madde-Hedef Uyumu Açısından İncelenmesi

Hülya KELECİOĞLU*

Kübra ATALAY**

Neşe ÖZTÜRK***

Hacettepe Üniversitesi

Hacettepe Üniversitesi

Mehmet Akif Ersoy Üniversitesi

Özet

Bu çalışmada, Milli Eğitim Bakanlığı'nca (MEB) 2009 yılında uygulanan 7. Sınıf Seviye Belirleme Sınavı (SBS) Matematik alt testinin madde- hedef uyumu incelenmiştir. SBS'de yer alan maddelerin ilköğretim programında yer alan kazanımlara göre oluşturulduğu belirtilmektedir. Maddelerin kazanımlara uygun olarak hazırlanıp hazırlanmadığının belirlenmesi amacıyla Milli Eğitim Bakanlığı Eğitim Teknolojileri Genel Müdürlüğü'nden (EĞİTEK) her maddenin hangi öğrenme alanını ve kazanımı ölçmek amacıyla yazıldığına ilişkin bilgi alınmıştır. Araştırmacılar tarafından, her maddenin ölçmeyi amaçladığı geçerli kazanım ile maddenin öğrenme alanındaki kazanımlar listelenerek bir uzman formu hazırlanmıştır. Uzmanlardan, maddenin listedeki kazanımlardan her birini ölçme derecesini “ölçüyor”, “kısmen ölçüyor” ve “ölçmüyor” seçenekleri ile belirtmeleri istenmiştir. Her madde için uzmanların verdiği puanların ortalaması (μ_k) ve madde-hedef uyum katsayıları (I_{jk}) hesaplanmıştır. Değerlendirme, μ_k 'nin en az 0,75 olduğu maddelerin geçerli kazanımı ölçtüğü ve I_{jk} 'nin en az 0,70 olduğu maddelerin geçerli kazanımı ölçtüğü, diğer kazanımları ölçmediği kabul edilerek yapılmıştır. Maddelerin geçerli kazanımlara ilişkin μ_k değerlerinin -0,45 ile 1,00 arasında; I_{jk} değerlerinin de -0,26 ile 0,66 arasında değiştiği bulunmuştur. Uzmanlar, μ_k değerlerine göre 7 maddenin geçerli kazanımları ölçtüğü konusunda görüş birliğine varırken, I_{jk} değerlerine göre maddelerin geçerli kazanımları ölçtüğü; diğer kazanımları ölçmediği konusunda görüş birliğine varamamışlardır.

Anahtar sözcükler: Geçerlik, kapsam geçerliği, madde-hedef uyum katsayıları, Seviye Belirleme Sınavı

Abstract

The scope of this study was to evaluate item-objective congruence of second level elementary (7th grade) mathematic sub test of high school entrance exam which was administered in 2009. It is asserted that the items in SBS had been constructed in accordance with the objectives of elementary curriculum. In order to determine whether the questions were constructed in conformity with the objectives, the information about the questions and their respective objectives were obtained from Ministry of Education General Headship of Educational Technology (EĞİTEK). An Expert Form was prepared by researchers, listing the objectives aimed to be measured by each question and also the objectives considered to be respective in field of learning. The content experts evaluated each item by defining the item “clearly measuring”, “partially measuring” and “clearly not measuring”. The item-objective congruence indexes and average of expert scores were calculated for each item. The evaluation was carried out on the assumption that the items with a μ_k value of at least 0,75 measures the valid objectives, and that the items with a I_{jk} value of at least 0,70 measures the valid objectives but not the other objectives. For valid objectives the μ_k values vary from -0.45 to 1.00 and the I_{jk} values vary from -0.26 to 0.66. According to μ_k values; the content experts agreed that eight items had measured the defined objectives. With regard to I_{jk} values the content experts agreed that they hadn't measured the defined objectives but they could not agree on whether that they had measured other objectives or not.

Key words: validity, content validity, indexes of item-objective congruence, high school entrance exam

* Doç. Dr., Hacettepe Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Bölümü Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, hulyaebb@hacettepe.edu.tr

** Arş. Gör. Kübra ATALAY, Hacettepe Üniversitesi, katalay@hacettepe.edu.tr

*** Arş. Gör. Neşe Öztürk, Mehmet Akif Ersoy Üniversitesi, nozturk@mehmetakif.edu.tr

Geçerlik bir ölçme aracında aranan en önemli özelliklerden biridir. Geçerliğin klasik tanımı, bir testin ölçmek istediği değişkeni ölçmesi ve onu başka değişkenlerle karıştırmamasıdır (Thorndike ve Hagen, 1961, Turgut, 1987). Geçerliğe, testin kullanılış amacına uygun kanıtların ortaya konma süreci olarak da yaklaşılmaktadır. Bu açıdan bakıldığında testin kullanılış amacına uygun olarak hangi kanıtların toplanacağına önceden karar verilmelidir (Baykul, 2000).

Geçerlik 1930–1940 yılları arasında bir ölçüle test puanları arasındaki korelasyonun hesaplanmasına dayalı olarak tanımlanmıştır. 1950’lerin başlarında ise kapsam geçerliği tartışmaları başlamış ve Rulon tarafından geçerli bir testin, öğrencinin dersin amaçlarına göre aldığı eğitim sonunda, bu dersin amaçlarına uygun davranışlarını ölçmesi gerektiği belirtilmiştir. Böyle bir testin geçerlik düzeyini uzmanların belirleyebileceklerini ileri sürmüştür. Cureton 1951 yılında kapsam geçerliği düşüncesini ortaya atmış ve geçerliğin uzman kanısına dayalı olarak tanımlanabileceğini açıklamıştır (Aktaran; Lissitz ve Samuelsen, 2007).

Cronbach ve Meehl (1951) geçerliğin yordama, uyum, kapsam ve yapı geçerliği olmak üzere dört türü olduğunu belirtmişlerdir. Kapsam geçerliğini, test maddelerinin araştırmacıların ilgilendiği evrenin bir örnekleme olup olmadığının gösterilmesi olarak tanımlamışlardır. Kapsam geçerliğinin, testin yapı geçerliği için de önemli bir kanıt olduğu belirtilmektedir (Cronbach ve Meehl, 1951; Messick, 1995; Haynes, Richard ve Kubany, 1995; Mislevy, 2007). Özellikle eğitimde kullanılan testlerde yapı geçerliğine odaklanmak yerine testlerin alt birimlerinin irdelenmesi gerektiği vurgulanmaktadır. Geliştirilen testle ilgili bir teorinin olmadığı durumlarda ölçüt ve kapsam geçerliğinin öne çıkarılması gerektiği belirtilmektedir (Lissitz ve Samuelsen, 2007). Crocker (2003), eğer puanlar eğitimde bireylerle ilgili karar vermede kullanılacaksa geçerlikle ilgili tartışmaların kapsam geçerliği üzerine olması gerektiğine işaret etmektedir.

Kapsam geçerliği, bir testin testle ölçülmek istenen özellikleri kapsama derecesi (Baykul, 2000) olarak ele alındığında öncelikle testle ilgili ölçülecek özelliklerin ve testin kapsamının belirlenmesi gerekir. Daha sonra bunlara uygun maddeler yazılarak test oluşturulur. Kapsam geçerliği belirlemede yaygın olarak kullanılan yöntemler uzman görüşlerine dayalıdır (Thorndike ve Hagen, 1961; Crocker ve Algina, 1986; Aiken 1988; Murphy ve Davidshofer, 1991; Baykul, 2000). Uzmanlardan elde edilen görüşlerin nicel hale getirilmesine ilişkin pek çok yöntem bulunmaktadır (Crocker ve Algina, 1986).

Türkiye’de yükseköğretim kurumlarına ve birçok öğretim kurumuna öğrenci seçme ve yerleştirme işlemi merkezi olarak yapılan sınavlar sonucu olmaktadır. Bu amaçla yapılan sınavlardan biri de ortaöğretim kurumlarına geçiş sisteminde önemli yer tutan Seviye Belirleme Sınavı’dır (SBS).

Türkiye’de ortaöğretime geçiş 2007–2008 eğitim-öğretim yılından itibaren ilköğretim 6, 7 ve 8. sınıfların sonunda uygulanan SBS ve yılsonu başarı puanına dayalı olarak hesaplanan sınıf puanına göre yapılmaktadır. Sınıf puanının %70’ini SBS puanı, %30’unu öğrencilerin okuldaki yılsonu başarı puanı oluşturmaktadır. Ortaöğretime geçişte önemli bir ağırlığı olan SBS’nin soruları öğrenci kazanımlarına yöneliktir ve soru sayıları haftalık ders saatlerine göre belirlenmiştir. SBS ile ilköğretimin 6, 7 ve 8. sınıflarında öğrencinin derslerden o yılın öğretim programlarında belirtilen kazanımları elde etme durumlarının ölçüleceği belirtilmektedir (<http://oges.meb.gov.tr>).

Berk (1984), madde-hedef uyumunun incelenmesinin, kapsam geçerliği değerlendirme sürecinde önemli olduğuna dikkat çekmektedir. Rovinelli ve Hambelton (1976) seçme ve yerleştirme amacıyla kullanılan sınavlarda, maddelerin ölçülmek istenilen kazanımların geçerli bir temsili olmasının önemini vurgulanmıştır. Rovinelli ve Hambelton (1976), maddelerin hedeflenen kazanımları ölçtüğünün bir göstergesi olarak tek boyutlu madde hedef uyum katsayısını geliştirmiştir. Geliştirdikleri formül Crocker ve Algina (1986) tarafından sadeleştirilmiştir.

$$I_{ik} = \frac{N}{2N-2} (\mu_k - \mu) \quad (1)$$

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

Bu formül bir maddenin sadece bir kazanımı ölçtüğü varsayımına dayalıdır. Uzmanlardan madde kazanımla eşleşiyorsa (+1), belirsizlik varsa (0), kazanımla eşleşmiyorsa (-1) puan vermeleri 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 ölçtüğünü gösterir (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 ortalamaların uzmanlar arası görüş birliğini gösterebileceğini belirtmişlerdir.

Rovinelli ve Hambleton (1976) tarafından geliştirilen formül, maddenin sadece bir kazanımı ölçtüğü durumlarda kullanılmaktadır. Bu nedenle katsayının yüksek değer vermesi, uzmanların maddenin geçerli kazanımı ölçtüğü ve diğer kazanımları ölçmediği konusunda fikir birliğinde bulunmalarına bağlıdır. Eğer bir madde birden fazla hedefi ölçüyorsa bu katsayı ile 1'den küçük değerler elde edilecektir. Turner ve Carlson (2003), bir maddenin birden fazla hedefi ölçtüğü durumlar için düzeltilmiş madde-hedef uyum katsayısını geliştirmişlerdir. Formülün sadeleştirilmiş hali aşağıdadır.

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

I'_{ik} : Madde-hedef uyum katsayısı
 p: Maddenin ölçtüğü belirtilen geçerli hedef sayısı
 N: Toplam hedef sayısı
 μ_k : i maddesinin geçerli hedefleri ölçme derecesine ilişkin puanların ortalaması
 μ_i : i maddesinin geçerli olmayan hedefleri ölçme derecesine ilişkin puanların ortalaması

Uzmanların puanlaması ve formül sonucunda elde edilen değerlerin aralığı ve yorumlanması formül (1)'deki gibidir.

SBS'nin, ilköğretim programındaki kazanımların öğrenciler tarafından kazanılma durumlarını ölçmek amacıyla hazırlandığından, madde-hedef uyumunun değerlendirilmesi ön plana çıkmaktadır. Eğer madde-hedef uyumuna ilişkin bir kanıt bulunamamışsa, bu maddelerin karar verme amacıyla kullanılması uygun değildir. Bu nedenle, bu çalışmada SBS'nin madde-hedef uyumu ele alınmıştır.

Bu çalışmada, 2009 yılında uygulanan 7. sınıf SBS matematik alt testine ait 18 maddenin madde- hedef uyumu her madde için madde-hedef uyum katsayısı ve uzmanların her bir hedefe verdiği puanların ortalaması hesaplanarak incelenmiştir. Böylece matematik alt testinin madde –hedef uyumuna ve kapsam geçerliğine ait bir kanıt elde edilmeye çalışılmıştır.

Yöntem

Bu araştırma, 2009 yılında ilköğretimin 7. sınıflarında uygulanan SBS matematik alt testinin kapsam geçerliğini incelemek amacıyla yapılmış betimsel bir çalışmadır.

Çalışma grubu

Çalışma grubu SBS matematik alt testinin kapsam geçerliği konusunda görüşleri alınan 11 uzmandan oluşmaktadır. Uzmanlarının tümü matematik eğitimi lisansına sahiptir. Uzmanlardan üçünün matematik eğitimi ve ikisinin ölçme ve değerlendirme alanında doktora vardır. Ayrıca çalışma grubunda, matematik eğitiminde doktora yapan dört, ölçme ve değerlendirme alanında yüksek lisans yapan iki uzman bulunmaktadır.

Veri toplama aracı

Araştırmada veri toplama aracı olarak araştırmacılar tarafından geliştirilen Uzman Formu kullanılmıştır. Uzman Formu geliştirilmeden önce, Eğitim Teknolojileri Genel Müdürlüğü'nden (EĞİTEK), SBS 2009 matematik alt testindeki her bir maddenin ilköğretim 7. Sınıf Matematik Programında yer alan hangi öğrenme alanını ve kazanımları ölçmek için geliştirildiğini gösteren madde formları alınmıştır. Bir madde için, o maddenin ölçmeyi amaçladığı kazanım (geçerli kazanım) da dâhil olmak üzere, maddenin öğrenme alanı ile ilgili kazanımları listelenmiştir. Uzman formunda, maddelerin hangi kazanımı ölçmek amacıyla yazıldığı belirtilmemiştir. Uzmanlardan, maddenin listede yer alan kazanımlarına göre her birini ölçme derecesini “tamamen ölçüyor”, “kısmen ölçüyor” ve “ölçmüyor” seçenekleri ile belirtmeleri istenmiştir. Bu şekilde 18 madde için ayrı ayrı hazırlanan formlar, görüşleri alınmak üzere uzmanlara verilmiştir. Uzmanlardan alınan cevaplara göre “tamamen ölçüyor” için 1, “kısmen ölçüyor” için 0 ve “ölçmüyor” için -1 verilerek puanlanmıştır.

Verilerin analizi

Seviye Belirleme Sınavı 7. sınıf matematik alt testinde toplam 18 madde bulunmaktadır. Testin kapsam geçerliğini değerlendirmek amacıyla, uzmanların maddelere her bir kazanım için verdikleri puanların ortalaması ve maddelerin her kazanımda aldıkları madde-hedef uyum katsayısı incelenmiştir. 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 hesaplanmıştır.

Turner ve Carslon (2003) tarafından 0,75 ve daha büyük ortalamının uzmanlar arası görüş birliğini gösterebileceğini belirtilmiştir. Bu nedenle bir geçerli kazanımı olan maddelerden μ_k değeri geçerli kazanımın dışındaki bir kazanım için 0,75 ve üzeri olanlar için ayrıca düzeltilmiş madde-hedef uyum katsayıları (I'_{jk}) da hesaplanmıştır. Bu hesaplama, bu maddelerin birden fazla kazanımı ölçüp ölçmediğini belirlemek amacıyla yapılmıştır.

Bulgular

Her maddenin öğrenme alanı ile ilgili kazanımları için uzmanlardan aldığı ortalama ve madde-hedef uyum katsayısı hesaplanmıştır. Maddeler, ortalama ve madde-hedef uyum katsayısı ölçütlerine göre değerlendirilmiştir.

Bir kazanımı ölçmek amacıyla hazırlanan 15 maddenin geçerli kazanımlarında aldıkları ortalamalar (μ_k), madde-hedef uyum katsayıları (I_{jk}) ve üç madde için düzeltilmiş madde-hedef uyum katsayıları (I'_{jk}) Tablo 1'de verilmiştir.

Tablo 1 incelendiğinde, maddelerin geçerli kazanımlara ilişkin μ_k değerlerinin -0,45 ile 1,00 arasında; I_{jk} değerlerinin de -0,26 ile 0,66 arasında değiştiği gözlenmektedir. 7 maddenin (2. , 3. , 5. , 6. , 8. , 12. Ve 18. maddeler) geçerli kazanımına ilişkin μ_k değeri 0,75'in üzerindedir. I_{jk} değerlerine göre ise hiçbir maddenin 0,70 ölçütünün üzerine çıkamadığı görülmektedir. μ_k değerlerine göre, uzmanlar 15 maddenin 7'sinin geçerli kazanımı ölçtüğü konusunda görüş birliğindedirler. 15 maddenin madde-hedef uyum katsayısı değerlerine göre, uzmanlar arasında maddelerin geçerli kazanımları ölçtüğü; diğer kazanımları ölçmediği konusunda görüş birliği oluşmamıştır. Bu durumdan, uzmanların 7

maddenin geçerli kazanımları ölçtüđü; ancak bu kazanımla birlikte öğrenme alanı ile ilgili kazanımlardan bazılarını da ölçtüđü görüşünde olduđu sonucu çıkartılabilir. Ancak, madde-hedef uyum katsayıları, uzmanların maddelerin öğrenme alanı ile ilgili verilen kazanımlardan hangilerini ölçtüđü konusunda da görüş birliđi içinde olmadıklarını göstermektedir.

Tablo 1. Bir kazanımı ölçtüđü belirtilen soruların μ_k ve I_{ik} deđerleri

Madde no	Kazanımlar												I'_{ik}
	1		2		3		4		5		6		
	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	
1	0,64	0,59	0,00*	0,21	-0,55	-0,12	-0,91	-0,34	-0,73	-0,23	-0,55	-0,12	
2	-0,18	-0,32	0,09	-0,15	1,00*	0,42	0,45	0,08	0,27	-0,03			
3	0,00	-0,02	0,91*	0,66	-0,82	-0,64							
4	0,64*	0,18	0,45	0,07	0,09	-0,16	0,82	0,30	-0,27	-0,39			0,42
5	0,09	-0,02	0,82*	0,43	0,36	0,15	-0,18	-0,19	-0,45	-0,36			
6	1,00*	0,39	0,18	-0,10	0,27	-0,05	1,00	0,39	0,09	-0,15	-0,45	-0,48	0,59
8	0,82*	0,64	0,09	0,15	-0,73	-0,39	-0,73	-0,39					
9	-0,45*	-0,26	0,64	0,47	-0,27	-0,14	-0,18	-0,08					
10	0,36*	0,11	0,55	0,25	-0,27	-0,36							
11	-0,73	-0,48	0,73*	0,61	-0,27	-0,14							
12	0,82*	0,27	0,73	0,21	-0,09	-0,33	0,18	-0,15					
13	0,73*	0,43	0,55	0,30	-0,82	-0,73							
14	0,91	0,27	0,73*	0,14	0,00	-0,41							0,62
16	-0,18*	-0,18	0,18	0,03	0,18	0,05							
18	0,09	-0,27	1,00*	0,41	0,27	-0,14							

*Geçerli kazanım

Uzmanlardan elde edilen ortalamalar incelediđinde, 3. ve 11. Maddeler hariç, tüm maddelerin geçerli kazanımla birlikte en az iki kazanımda pozitif ortalama aldıđı gözlenmektedir. Belirtilen bu iki madde de analizler sonucu en yüksek I_{ik} (0,66 ve 0,61) deđerine sahip olan maddelerdir. Belirtilen iki maddenin geçerli kazanımda pozitif yüksek ortalama ve madde-hedef uyum katsayısı deđerine sahip olması, bu maddelerin geçerli kazanımı ölçtüđü sonucunu göstermektedir.

Tablo 1'de ayrıca 4. maddede geçerli kazanımının μ_k deđerinin 0,64; başka bir kazanımdaki μ_k deđerinin 0,82 ve 14. maddede geçerli kazanımda μ_k 'nın 0,73 ve başka bir kazanımda 0,91 olduđu gözlenmiştir. Bu bulgular uzman görüşlerinden elde edilen ortalamalara göre, 4. ve 14. maddelerin geçerli kazanımdan başka bir kazanımı ölçtüđüğünü ortaya koymaktadır. Ayrıca 6. maddenin geçerli kazanımda ve başka bir kazanımda da μ_k 'nın 1,00 deđerini alması, bu maddenin de geçerli kazanımla birlikte başka bir kazanımı da ölçtüđüğünü göstermektedir. 4., 6. ve 14. maddelerin iki kazanımı birden ölçüp ölçmediđini anlamak için I'_{ik} deđerleri hesaplanmış ve bu deđerlerin sırasıyla 0,43, 0,59 ve 0,62 olduđu görülmüştür. Üç madde için hesaplanan çok boyutlu madde-hedef uyum katsayıları, bu soruların tek boyutlu madde-hedef uyum katsayılarından yüksek olmasına rağmen yine de 0,70 ölçütünü karşılamamıştır.

İki kazanımı ölçmek amacıyla hazırlanan üç maddenin ortalamaları, madde-hedef uyum katsayıları ve düzeltilmiş madde-hedef uyum katsayıları Tablo 2'de verilmiştir.

Tablo 2. İki kazanımı ölçtüğü belirtilen soruların μ_k , I_{ik} ve I'_{ik} değerleri

Madde no	Kazanımlar												I'_{ik}
	1		2		3		4		5		6		
	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	μ_k	I_{ik}	
7	0,27	0,04	0,55*	0,20	0,18	-0,02	-0,09	-0,18	0,00	-0,13	0,36*	0,09	0,24
15	0,64*	0,16	0,64*	0,16	0,00	-0,32							0,48
17	0,27*	0,11	-0,09*	-0,14	0,09	-0,02	0,18	0,05					0,16

*Geçerli kazanımlar

Tablo 2 incelendiğinde, üç maddenin geçerli kazanımlardaki μ_k değerlerinin 0,75'in ve I'_{ik} değerlerinin 0,70'in altında kaldığı görülmektedir. Bu bulgular, uzmanların Tablo 2'deki maddelerin ölçtüğü kazanımlar konusunda görüş birliğinde olmadığını göstermektedir.

Tartışma

Bu araştırma kapsamında ilköğretim 7. sınıf SBS matematik alt testinin madde-hedef uyumu uzman görüşü alınarak incelenmiştir. Maddeler, geçerli kazanımları için uzmanlardan aldıkları ortalamalara ve madde-hedef uyum katsayılarına göre değerlendirilmiştir. Ortalaması en az 0,75 olan maddelerin geçerli kazanımı ölçtüğü; madde-hedef uyum katsayısı en az 0,70 olan soruların da geçerli kazanımı ölçtüğü, diğer kazanımları ölçmediği kabul edilmiştir (Turner ve Carlson, 2003; Hambleton, 1976). SBS matematik alt testinde yer alan 18 maddenin 15'i bir kazanımı ölçmek amacıyla hazırlanmıştır. 15 maddenin 8'inde ortalamalar 0,75 ölçütünü karşılamaktadır. Ancak 15 maddenin hiçbirinde 0,70 ve üzerinde madde-hedef uyum katsayısına ulaşamamıştır.

İncelenen maddelerin geçerli kazanımları ölçtüğü; geçerli kazanım dışındaki kazanımları ölçmediği konusunda uzmanların görüş birliğinde olmamaları bu maddelerin geçerli kazanımlar dışında başka kazanımları da ölçtüğünün bir göstergesi olabilir. Bazı maddelerde bazı kazanımların geçerli kazanımla aynı ya da daha yüksek ortalama aldığı görülmüştür. Bu durum, bu maddelerin birden fazla kazanımı ölçtüğüne işaret etmektedir. Birden fazla kazanımda pozitif ortalama alan 4, 6 ve 14. maddelerde I'_{ik} değerlerinin, I_{ik} değerlerinden yüksek olduğu gözlenmiştir; ancak I'_{ik} değerleri de 0,70 ölçütünü karşılamamıştır. Bu bulgular da soruların hangi kazanımları ölçtüğüne ilişkin uzmanlar arasında görüş birliği oluşmadığını göstermektedir.

İki kazanımı ölçmek amacıyla hazırlanan maddelerde I'_{ik} değerlerinin, ilgili kazanımlardaki I_{ik} değerlerinden yüksek olduğu, fakat 0,70'in üzerinde olmadığı gözlenmiştir.

Çalışma kapsamında incelenen 7. Sınıf 2009 SBS matematik alt testi maddelerinin ilgili olduğu bazı kazanımlarda “çözer ve kurar”, “açıklar ve gösterir”, “tahmin eder ve oluşturur” gibi birden fazla ifadenin bulunması, uzmanların soruların hangi kazanımı ölçtüğü konusunda görüş birliğine varamamasının bir nedeni olarak gösterilebilir. “Dairesel silindirin temel elemanlarını belirler, inşa eder ve açılımını çizer.” kazanımında olduğu gibi öğretim programındaki bazı kazanımların çoktan seçmeli sorularla ölçülebilir nitelikte olmaması SBS'nin madde- hedef uyumunu düşüren bir diğer etken olarak düşünülebilir.

Bu nedenle öncelikle öğretim programındaki kazanımların gözden geçirilerek kazanımların ölçmeye daha elverişli olacak şekilde yeniden düzenlenmesi uygun olacaktır.

Bu çalışmada, SBS'de yer alan maddelerin belirtilen öğrenme alanını ölçtüğü kabul edilmiştir. Bundan sonra yapılacak çalışmalarda, madde-kazanım uyumu ile birlikte maddelerin öğrenme alanı ile ilgili olup olmadığı da değerlendirilmelidir.

Kapsam geçerliğinin incelenmesi, sınav uygulanmadan önce yapılması gereken bir çalışmadır. Bu nedenle EĞİTEK tarafından her yıl soruların yazılması tamamlandıktan sonra madde-hedef uyum çalışması yapılmamıştır.

Kaynaklar

- 2009 Ortaöğretim kurumlarına geçiř sistemi tercih ve yerleřtirme e-kılavuzu. (7 Ağustos 2009) Ankara: Milli Eğitim Bakanlığı. İnternette 5 Ocak 2010’da elde edilmiřtir. <http://oges.meb.gov.tr>
- Aiken, L. R. (1988). *Psychological testing and assessment*. 6. baskı. USA: Allyn and Bacon, Inc.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme*. Ankara: ÖSYM Yayınları.
- Berk, R. A. (1984). *Conducting the item analysis*. In R. A. Berk (Ed.), *A guide to criterion referenced test construction* (pp. 97–143). Baltimore: The Johns Hopkins University Press.
- Crocker ve Algina (1986). *Introduction to classical and modern Test Theory*, Rinehart and inston, New York: Holt.
- Cronbach, J. L. ve Meehl, P. L. (1951). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- 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.
- Lissitz, R. W. ve Samuelsen, K. (2007). A suggested change in terminology and emphasis regarding validity and education. *Educational Researcher*, 36, 8, 437-448.
- Messick, S. (1995). Validity of psychological assessment. *American Psychologist*, 50, 9, 741-749.
- Mislevy, R. J. (2007). Validity by design. *Educational Researcher*, 36, 8, 463-469.
- Murphy, K. R. ve Davidshofer, C. O. (1991). *Psychological Testing: Principles and Applications*. New Jersey: Prentice Hall., 2. Baskı.
- Thorndike, R. L. ve Hagen, E. P. (1961). *Measurement and evaluation in psychology and education*. 2. baskı. USA: Wiley International Edition,
- Turgut, M. F. (1992). *Eğitimde ölçme ve deęerlendirme Metotları*. 9. Baskı. Ankara: Saydam Matbaacılık.
- Turner, R. C. , Mulvenon, S. W. , Thomas, S. P. ve Balkin, R. S. (2002). Computing indices of item congruence for test development validity assessments. *Proceedings of the SAS Users’ Group International Conference, USA*, 27, 255.1-255.5.
- Turner, R. C. ve Carlson, L. (2003). Indexes of item-objective congruence for multidimensional items. *International journal of testing*. 3(2), 163-171