

## Ölçek Geliştirme Çalışmalarında Kapsam Geçerliği için Kapsam Geçerlik İndekslerinin Kullanılması

Dr. Halil Yurdugül  
Hacettepe Üniversitesi  
Eğitim Fakültesi  
yurdugul@hacettepe.edu.tr

### Özet:

Ölçek geliştirme çalışmalarında deneysel uygulamaların olanaklı olmadığı durumlarda kapsam geçerlik oranlarına başvurulur. Kapsam geçerlilik oranları, uzman görüşlerine dayalı nitel çalışmaları istatistiksel nicel çalışmalara dönüştürmek için kullanılan bir yöntemdir. Bu çalışmada kapsam geçerlik oranları ve kapsam geçerlik indekslerinin kullanımı tanıtılmış ve deneysel uygulama sonuçları ile kapsam geçerlik çalışmalarının tutarlığı incelenmiştir.

**Anahtar Kelimeler:** Ölçek geliştirme, uzman görüşleri, kapsam geçerlik oranları, kapsam geçerlik indeksleri

### Abstract:

When developing a scale in which an pilot study is not possible, the content validity ratios are used. The content validity ratio is a method to convert the expert's judgments based on qualitative analysis to statistical quantitative analysis. The purpose of this study is to explore the content validity ratio usage and the content validity index. In addition, in scale development process, the coherent between the results of experimental application and content validity application have been investigated.

**KeyWords:** Scale development, expert judgments, content validity ratio, content validity index

### Giriş

Eğitim ve psikolojide ölçek geliştirme çalışmalarında, geliştirilmek istenilen ölçek ile ölçülmek istenilen özellik arasındaki bağıntının tutarlı olması amaçlanır. Bu nedenle; ölçeği oluşturan maddelerin yüksek düzeyli psikometrik özellikler göstermesi gerekir. Ancak ölçek maddelerinin psikometrik özelliklerini belirlemek için deneme uygulamasının (pilot study) olanaklı olmadığı durumlarda ölçmeye konu olan özellik kapsamında uzman görüşlerine başvurulur. Uzman görüşleri üzerine yapılan çalışmalar özünde nitel bir çalışmadır. Kapsam geçerlik oranları ve kapsam geçerlik indeksleri, uzman görüşlerine dayalı nitel çalışmaları, istatistiksel nicel çalışmalara dönüştürmek amacıyla kullanılan bir yöntemdir. Bu çalışmada kapsam geçerlik oranlarının tutarlılığı ele alınmıştır.

### Ölçek Geliştirme Çalışmaları

Ölçek kavramı özünde ölçme sonuçlarının matematiksel niteliklerini göstermekle birlikte (Turgut ve Baykul, 1992) eğitim ve psikoloji gibi davranış bilimlerinin bir çok alanında, hedeflenen kişi(ler), sistem, konu ya da içerik açısından veri/bilgi toplamak amacıyla kullanılır. Torgerson (1958), ölçekleri uygulama biçimine göre iki grupta toplamaktadır; (a) yanıtlayıcı merkezli ve (b) gözlemleyici merkezli ölçekler. Çoktan seçmeli testler, tutum ölçekleri, anketler gibi ölçme araçları yanıtlayıcı merkezli ölçekler kapsamındadır. Diğer taraftan gözlemleyici merkezli ölçekler için gözlem formları, rubrikler vb. örnek olarak verilebilir.

Ölçek geliştirme çalışmaları, genellikle deneysel süreç ya da kuramsal süreçler ile gerçekleştirilir (EK-1). Deneysel süreçte literatür ya da uzman yaklaşımları sayesinde aday ölçek formu elde edilir ve hedef kitle ile benzer özellikler taşıyan bir örneklem grubuna deneme uygulaması yapılarak ölçek maddelerine ilişkin psikometrik özellikler belirlenerek ideal maddelerden nihai form elde edilir. Bu sürecin karakteristik özellikleri ise; nicel bir çalışma özelliği taşıması, genellikle faktör analizlerinin kullanılması ve büyük örneklem gerektirmesidir. Bu sürece ilişkin genel yaklaşım ve bilgiler Tezbaşaran (1989) ve Torgerson (1958) tarafından ayrıntılı bir şekilde açıklanmıştır.

Kuramsal süreçte ise; büyük örneklem gruplarına ulaşamama durumlarında aday ölçek formundaki maddelere ilişkin uzman görüşleri alınarak nitel çalışma yapılmaktadır. Nitel çalışmada uzman görüşleri arasındaki uyumluluklar test edilebilmektedir. Ancak kapsam geçerlik oranları yardımı ile bu nitel süreç istatistiksel nicel bir sürece dönüştürebilmektedir (McGartland vd., 2003).

### Kapsam Geçerlik Oranları ve İndeksleri

Bilindiği gibi ölçülmek istenilen özellik ile ölçek maddeleri arasındaki bağıntı, ölçme aracının geçerliğine ilişkindir. Ölçek maddesinin ölçülmesi amaçlanan özelliği kapsama (kapsam geçerliği) ya da maddenin ilgili yapıyı yorma (yapı geçerliği) gücünü belirlemek amacıyla önsel çalışmalara ihtiyaç vardır (McGartland vd., 2003). Yine ölçme aracının geçerliğini etkileyen diğer faktörler de ölçek geçerliği için göz önüne alınması gereken noktalardır; ölçek

maddesinin anlaşılabilir olması, hedef-kitleye uygunluğu vb. Önsel çalışmalarda elde edilen uzman görüşleri arasındaki uyum/uyumsuzluk aynı zamanda kapsam ya da yapı geçerliği için birer kestirim niteliğinde kullanılmaktadır.

Kapsam geçerlik oranları, Lawshe (1975) tarafından geliştirilmiştir. Bu nedenle Lawshe tekniği olarak bilinen bu yaklaşım 6 aşamadan oluşmaktadır.

- Alan uzmanları grubunun oluşturulması
- Aday ölçek formlarının hazırlanması
- Uzman görüşlerinin elde edilmesi
- Maddelere ilişkin kapsam geçerlik oranlarının elde edilmesi
- Ölçeğe ilişkin kapsam geçerlik indekslerinin elde edilmesi
- Kapsam geçerlik oranları/indeksi ölçütlerine göre nihai formun oluşturulması

Lawshe tekniğinde, en az 5 en fazla ise 40 uzman görüşüne ihtiyaç vardır. Her bir madde uzman görüşleri “madde hedeflenen yapıyor ölçüyor”, “madde yapı ile ilişkili ancak gereksiz” ya da “madde hedeflenen yapıyı ölçmez” şeklinde derecelendirilmektedir. Kapsam geçerliğinin yanı sıra benzer şekilde maddenin anlaşılabilirliği, hedef-kitleye uygunluğu vb. amacıyla da uzman görüşleri derecelendirilebilir.

Buna göre, uzmanların herhangi bir maddeye ilişkin görüşleri toplanarak kapsam geçerlik oranları elde edilir (Tablo 2). Kapsam geçerlik oranları (KGO), herhangi bir maddeye ilişkin “Gerekli” görüşünü belirten uzman sayılarının, maddeye ilişkin görüş belirten toplam uzman sayısına oranının 1 eksiği ile elde edilir.

$$KGO = \frac{N_G}{N/2} - 1 \quad (1)$$

Burada;  $N_G$ , maddeye “Gerekli” diyen uzmanları sayısını ve  $N$  ise maddeye ilişkin görüş belirten toplam uzman sayısını göstermektedir.

Eşitlik 1’e göre; uzmanların yarısı maddeye ilişkin “Gerekli” şeklinde görüş bildirdiklerinde  $KGO=0$ , yarısından fazlası “Gerekli” şeklinde görüş bildirmiş ise  $KGO>0$  ve uzmanların yarısından fazlası “Gerekli” şeklinde görüş bildirmemiş ise  $KGO<0$  olacaktır.

KGO değerleri negatif ya da 0 değer içeriyorsa böyle maddeler ilk etapta elenen maddelerdir. KGO değerleri pozitif olan maddeler için istatistiksel ölçütler ile anlamlılıkları test edilirler. Elde edilen KGO’ların istatistiksel olarak anlamlılığını test etmek için kapsam geçerlik ölçütleri için ilgili literatürde önceleri birikimli normal dağılımdan yararlanılırken, hesaplama kolaylığı açısından  $\alpha=0,05$  anlamlılık düzeyinde KGO’ların minimum değerleri (kapsam geçerlik ölçütleri) Veneziano ve Hooper (1997) tarafından tabloya dönüştürülmüştür. Buna göre, uzman sayısına ilişkin minimum değerler aynı zamanda maddenin istatistiksel anlamlılığını vermektedir.

**Tablo 1:**  $\alpha=0,05$  anlamlılık düzeyinde KGO’ları için minimum değerler.

Uzman Sayısı	Minimum Değer	Uzman Sayısı	Minimum Değer
5	0.99	13	0.54
6	0.99	14	0.51
7	0.99	15	0.49
8	0.78	20	0.42
9	0.75	25	0.37
10	0.62	30	0.33
11	0.59	35	0.31
12	0.56	40+	0.29

Tablo 2’de ise kapsam geçerlik oranlarının elde edilmesine ilişkin örnek uygulama verilmiştir. Toplam 12 uzmanın maddelere ilişkin belirtmiş oldukları görüşler üzerinden, Eşitlik 1 ile verilen ifade yardımı ile kapsam geçerlik oranları elde edilmiştir. Daha sonra, bu oranların istatistiksel olarak anlamlılığı, Tablo 1’de yer alan değerlerle (0,59) karşılaştırılarak belirlenmiştir. Testin tümüne ilişkin kapsam geçerlik indeksi de yine Tablo 1’deki değerler ile karşılaştırılmıştır.

Kapsam geçerlik indeksi (KGİ),  $\alpha=0,05$  düzeyinde anlamlı olan ve nihai forma alınacak maddelerin toplam KGO ortalamaları üzerinden elde edilir. Eğer ölçülmek istenilen özellik birden fazla boyutta toplanmış ise her bir boyut için KGİ elde edilmelidir. Tablo 2’deki örnekte tek bir boyut söz konusu olduğu varsayımı ile  $KGİ>0,67$  olduğundan ölçek istatistiksel olarak anlamlı olduğu söylenebilir. KGİ değerleri alt boyutlar için geçerli olup, her bir alt boyut için, alt boyutta yer alan maddeler dikkate alınarak elde edilmektedir.

**Tablo 2:** Uzman görüşlerinin toplanması ve kapsam geçerlik oranlarının elde edilmesi

	Gerekli	Yaratılı/Yetersiz	Gereksiz	Kapsam Geçerlik Oranları
Madde 1	12	0	0	1,00 <sup>+</sup>
Madde 2	10	1	1	0,67 <sup>+</sup>
Madde 3	8	1	3	0,33 <sup>-</sup>
:	:	:	:	:
Madde k	5	3	4	-0,167
<b>Uzman Sayısı</b>	12			
<b>Kapsam Geçerlik Ölçütü</b>	0,59			
<b>Kapsam Geçerlik İndeksi</b>	0,67*			

Kapsam geçerliği konusunda kullanılan en yaygın teknik Lawshe tekniği olmasına karşın başka tekniklerde geliştirilmiştir. Bunlardan Davis (1992) tekniği uzman görüşlerini (a) “Uygun”, (b) “Madde hafifçe gözden geçirilmeli”, (c) “Madde ciddi olarak gözden geçirilmeli” ve (d) “Madde uygun değil” şeklinde dörtlü derecelendirmektedir. Bu teknikte (a) ve (b) seçeneğini işaretleyen uzmanların sayısı toplam uzman sayısına bölünerek maddeye ilişkin “kapsam geçerlik indeksi” elde edilmektedir ve bu değer istatistiksel bir ölçütü karşılaştırmak yerine 0,80 değeri ölçüt olarak kabul edilmektedir.

## Yöntem

Bu çalışmada askeri kurumlarda çalışan personelin verimliliklerine ilişkin değerlendirme ölçeğinin geliştirilmesi ele alınmıştır. İlgili literatüre ve konu alanı uzmanlarına dayanılarak 5 alt boyutta toplam 90 adet aday ölçek maddesi üretilmiştir.

Tablo 3: Aday ölçek maddelerinin alt boyutlara göre dağılımı

Alt Boyut	Madde Sayısı
Takım Ruhu	18
Sorumluluk	16
Planlama	22
Mesleki Deneyim	16
Problem Çözme ve Yaratıcılık	18

Aday ölçek maddeleri öncelikle kapsam geçerliği açısından uzman görüşlerine başvurularak (uzman sayısı 14) her bir madde için kapsam geçerlik oranları elde edilerek maddelerin kalitesi belirlenmiştir. Ancak aday ölçek için, 90 maddeye ilişkin KGO'lara dayalı madde analizi yapılmadan (olumsuz maddeler elimine edilmeyerek) tüm maddeleri içeren ölçek taslağına yönelik deneme uygulamasına gidilmiştir. Elde edilen sonuçlar Baykal (1994) tarafından eğitim alanına uygulanan “yapısal geçerlik katsayısı” yardımı ile tutarlılıkları incelenmiştir.

$$H(X) = \log K - \frac{1}{K} \sum_{i=1}^B N(X) \log N(X)$$

$$H(Y) = \log K - \frac{1}{K} \sum_{i=1}^B N(Y) \log N(Y)$$

Burada; K, toplam madde sayısını, B, testte tanımlanan boyut sayısını, H(X), beklenen belirsizlik değerini ve H(Y) ise gözlenen belirsizlik değerini göstermektedir. N(X) ve N(Y) ise sırasıyla; tanımlanan boyutlardaki beklenen ve gözlenen boyut sayılarıdır. Aynı zamanda çapraz belirsizlik ise;

$$H(X, Y) = \log K - \frac{1}{K} \sum_{X=1}^B \sum_{Y=1}^B N(X, Y) \log N(X, Y)$$

şeklinde ifade edilmektedir. Bu belirsizlik değerleri uyarında yapısal geçerlik katsayısı ise;

$$g = \frac{H(X) + H(Y) - H(X, Y)}{H(X, Y)}$$

şeklinde tanımlanmaktadır. Yapısal geçerlik katsayısının istatistiksel anlamlılığı ki-kare (chi-square) test istatistiği ile test edilmektedir (Baykal, 1994).

### Bulgular

Kapsam geçerlik oranlarına dayalı kapsam geçerlik çözümlemesi ve deneme uygulaması sonucu faktör çözümlemesi ile elde edilen maddelerin ölçeğin alt boyutlarına göre dağılımları Tablo 4'te verilmiştir.

Tablo 4: Kapsam geçerlik çözümlemesi ve faktör çözümlemesine göre maddeleri dağılımı

		Beklenen (Kapsam Geçerlik Çözümlemesi)						
Gözlenen (Faktör Çözümlemesi)		Takım Ruhu	Sorumluluk	Planlama	Mesleki Deneyim	Yaratıcılık	Elenen Maddeler	Toplam
	Takım Ruhu	12						12
	Sorumluluk		12				2	14
	Planlama			15				15
	Mesleki Deneyim				10			10
	Yaratıcılık					12	2	14
	Elenen Maddeler	1	2	1		2	19	25
	<b>Toplam</b>	<b>13</b>	<b>14</b>	<b>16</b>	<b>10</b>	<b>14</b>	<b>23</b>	<b>90</b>

Tablo 4'e göre; kapsam geçerlik oranları kullanılarak belirlenen maddeler ile faktör analizi yöntemi ile özellikleri saptanan maddelerin dağılımı ve her iki yaklaşıma ilişkin boyutlarda toplanan maddelerin marjinal dağılımları görülmektedir. Bu verilere göre elde edilen belirsizlik katsayıları ve beklenen ile gözlenen madde/boyut tutarlıklarına ilişkin yapısal geçerlik katsayıları elde edilmiştir.

Kapsam Geçerlik Belirsizliği:  $H(X)=0,78$   
 Faktör Çözümleme Belirsizliği:  $H(Y)= 0,78$   
 Çapraz Belirsizlik:  $I(X, Y)=0,93$   
 Yapısal Geçerlik Katsayısı:  $g=0,69$

Bu bulgulara göre her iki yaklaşımın birbirleri ile tutarlı olduğu gözlenmiştir.

### Sonuç ve Öneriler

Eğitim ve psikoloji alanındaki ölçek geliştirme çalışmalarında; deneysel uygulamaların olanaklı olmadığı durumlarda, uzman görüşlerine dayalı nitel çalışmalar ile ölçek ve ölçekte yer alan maddelerin kalitesi kestirilmektedir. Kapsam geçerlik çözümlemesi, uzman görüşlerine dayalı nitel çalışmaları istatistiksel nicel çalışmalara dönüştüren bir süreç olarak ele alınabilir.

Bu çalışmada deneysel uygulama yöntemiyle elde edilen maddelerin psikometrik özellikleri ile kapsam geçerlik çözümlemesinden elde edilen özelliklerin tutarlı olduğu bulgusuna ulaşılmıştır.

Son zamanlarda eğitim alanında süreç ya da ürün değerlendirmeye yönelik olarak gözlemleyici merkezli ölçme araçlarına ilişkin uygulamaların arttığı görülmektedir. Gözlemleyici merkezli yaklaşımlarda ölç(ül)meye konu olan özelliklerin daha özel alanlarda yer almasından ve bu alanda geniş örneklemelere ulaşmanın zorlukları göz önüne alındığında kapsam geçerlik çözümlemelerinin uygulanabilirliği görülmüştür. Örneğin proje tabanlı öğrenme ürünlerinde proje üye sayılarının çok sınırlı olması, bu konudaki ölçek geliştirme çalışmalarında deneme uygulamalarını olanaksız hale getirmektedir. Bu gibi durumlarda kapsam geçerlik çözümlemelerinin kullanımı deneme uygulaması ile tutarlı sonuçlar verebileceği görülmektedir. Kapsam çözümlemesinin tutarlı/yansız sonuçlar verebilmesinin bir ölçütü olarak uzmanların niteliği de önem taşımaktadır.

Kuramsal sürece, bir diğer ifade ile uzman görüşlerine dayalı çalışmalarda kapsam geçerlik çözümlemesi, geçerliği yüksek ölçekler geliştirmede, uzman görüşlerini nicel bulgulara dönüştürmede ve işlem-zaman kolaylığı açısından yararlıdır.

Kapsam geçerlik çalışmaları, deneysel süreçlerde de uygulanabilirliği gözlenmektedir. Bu durumda kapsam çözümlemesi; ölçeğin ölçmeye yöneldiği yapıya ilişkin yapı geçerliğini artırma, önsel madde analizleri ve madde-faktör sayılarını kontrol edebilme gibi avantajlar sağlamaktadır.

### **Kaynakça**

- Baykal, A. (1994). “*Davranış ölçümünde yapısal geçerlik göstergesi.*” Türk Psikoloji Dergisi, 33, 45-50.
- Davis, L.L.(1992). “*Instrument review: Getting the most from a panel of experts*”. Applied Nursing Research, 5, 194-197.
- Lawshe, C. H. (1975). “*A quantitative approach to content validity.*” Personnel Psychology, 28, 563–575.
- McGartland, R. D., Berg-Weger, M., Tebb, S., Lee, E. S., ve Rauch, S. (2003). “*Objectifying content validity: Conducting a content validity study in social work research*”. Social Work Research, 27(2), 94 - 104.
- Tezbaşaran, A. (1997). **Likert tipi ölçek geliştirme kılavuzu**. Ankara: Türk Psikologlar Derneği Yayınları.
- Togerson, W.S. (1958). **Theory and Methods of Scaling**, New York: Wiley.
- Turgut, M. F. ve Baykul, Y. (1992). **Ölçekleme Teknikleri**, ÖSYM Yayınları, Ankara.
- Veneziano L. ve Hooper J. (1997). “*A method for quantifying content validity of health-related questionnaires*”. American Journal of Health Behavior, 21(1):67-70.

# Ölçek Geliştirme Süreci

## Ölçek Oluşturma Süreci

## Psikometrik Özelliklerin Belirlenme süreci

