eefdergi
Cilt 20 Sayı 1 Erzincan Üniversitesi Eğitim Fakültesi Dergisi
e-ISSN 2148-7510
http://eefdergi.erzincan.edu.tr | eefdergi@erzincan.edu.tr
Doi numarası: 10.17556/erziefd.297741

A Road Map for the Content Validity Used in Scale Development Studies*

Selami YEŞİLYURT**, Cüneyt ÇAPRAZ***

Received date: 13.03.2017 Accepted date: 26.01.2018

Abstract

The aim of the present study is to present a road map to the researchers in which the answer of the question of how the content validity studies can be carried out is given with introducing content validity ratio (CVR) and Content Validity Index (CVI). The road map is presented through a sample study within regard to the technique developed by Lawshe (1975) and the critical content ratio determined by Ayre and Scally (2014). 26 experts are joined to the sample study. As a data collection tool an 'expert form' in which 66 items are cross-examined, is used.

Keywords: Content validity ratio (CVR), Content Validity Index (CVI), critical content validity ratio, scale development, expert views.

* The aim of the present study is to present a road map to the researchers in which the answer of the question of how the content validity studies can be carried out is given with introducing content validity ratio and Content Validity Index.

** Atatürk University, Education Faculty, Department of Biology Education, Erzurum, Turkey; selamiy@atauni.edu.tr
*** Alaaddin Keykubat University Faculty of Education, Primary School Teaching, Alanya, Turkey; cuneyt.capraz@alanya.edu.tr

Ölçek Geliştirme Çalışmalarında Kullanılan Kapsam Geçerliği İçin Bir Yol Haritası *

Doi numarası: 10.17556/erziefd.297741

Selami YEŞİLYURT **, **Cüneyt ÇAPRAZ** ***

Geliş tarihi: 13.03.2017 **Kabul tarihi26.01.2018**

Öz

Bu çalışmanın amacı kapsam geçerlik oranları ve kapsam geçerlik indeksi tanıtılarak kapsam geçerlik çalışmalarının ne şekilde yapılabileceği ile ilgili çalışmacılara bir yol haritası sunmaktır. Yol haritası, örnek bir çalışma üzerinden, Lawshe (1975) tarafında geliştirilen teknik ile Ayre ve Scally (2014) tarafından belirtilen ölçütler dikkate alınarak verilmiştir. Örnek çalışmaya 26 uzman katılmıştır. Veri toplama aracı olarak 66 maddenin sorgulandığı "Uzman Formu" kullanılmıştır.

Anahtar kelimeler: Kapsam geçerlik oranı, kapsam geçerlik indeksi, kapsam geçerlik ölçütü, ölçek geliştirme, uzman görüşleri.

*Bu çalışmanın amacı kapsam geçerlik oranları ve kapsam geçerlik indeksi tanıtılarak kapsam geçerlik çalışmalarının ne şekilde yapılabileceği ile ilgili çalışmacılara bir yol haritası sunmaktır.

** Atatürk Üniversitesi, Kâzım Karabekir Eğitim Fakültesi, Biyoloji Eğitimi ABD, Erzurum, Türkiye; selamiy@atauni.edu.tr

*** Alaaddin Keykubat Üniversitesi Eğitim Fakültesi, Sınıf Eğitimi Anabilim Dalı, Alanya, Türkiye; cuneyt.capraz@alanya.edu.tr

1. Giriş

Eğitim ve psikolojide tutum ve görüşlerin doğru bir şekilde tespiti büyük önem arz etmektedir (Özbek, Kahya ve Özgen, 2007; Yurdugül, 2005). Bu durum ölçek geliştirme çalışmalarını öne çıkarmaktadır. Ancak ölçek geliştirmede öncelikli dikkat edilmesi gereken husus ölçme aracının geçerliğidir (Brains, Willnat, Manheim ve Rich, 2011). Zira bir bilimsel çalışmanın inandırıcılığı ve genellenebilmesi için geçerlik önemli bir ölçüttür (Basham ve Sedlacek, 2009; Yıldırım ve Şimşek, 2013). Geçerlik çalışmalarında geliştirilmek istenen ölçek ile ölçülmek istenilen özellik arasındaki bağıntının tutarlı olmasına özen gösterilmelidir. Kapsam geçerliği olarak ifade edilen bu tip bir çalışmayla her bir maddenin amaca ne derece hizmet ettiği ortaya konulur. Bu sayede geliştirilen bir ölçekte, çalışılan konu harici ilgisiz ifadeler yerine konuyu temsil gücü yüksek ifadelere yer verilecektir (Ayre ve Scally, 2014; Basham ve Sedlacek, 2009; Brinkman, 2009; Frank-Stromberg ve Olsen, 2004; Wilson, Pan ve Schumsky, 2012). İçerik, mantıksal veya rasyonel geçerlilik olarak da adlandırılan kapsam geçerlik çalışmalarında (Shuttleworth, 2016), ölçek maddesinin ölçülmesi amaçlanan özelliği kapsama gücünü belirlemeye yönelik yeterli sayıda uzman görüşlerinin alındığı bir ön çalışma mutlaka yapılmalıdır (Allen ve Yen, 2002; Basham ve Sedlacek, 2009; Brinkman, 2009; Büyüköztürk, 2005; Ercan ve Kan, 2004; Lawshe, 1975; McMillan ve Schumacher, 2010, Otrar ve Argın, 2015; Rubio, ve diğ., 2003). Çünkü bu tür ön çalışmalarda uzman sayısı yeterli sayıda tutulduğu takdirde (5-40 arası) hazırlanacak ölçeğin geçerliği de yüksek olacaktır (Ayre ve Scally, 2014; Lawshe, 1975; Veneziano ve Hooper, 1997; Wilson, ve diğ., 2012).

Kapsam geçerliği ile ölçek maddelerinin araştırmacıların ilgilendiği çalışma evrenin bir örneklemi olup olmadığının tespitinden sonra kapsam çerçevesinde uygun maddeler yazılarak ölçek oluşturulabilir (Cronbach ve Meehl, 1951; Rubio, ve diğ., 2003). Ancak uzman görüşleri üzerine yapılan kapsam geçerlilik çalışmaları nitel çalışmalardır (Shuttleworth, 2016). Bundan dolayı elde edilen verilerin, kapsam geçerlik oranları (KGO) ve kapsam geçerlik indeksi (KGİ) hesaplanarak nicel verilere dönüştürülmesi gerekir (Ayre ve Scally, 2014; Büyüköztürk, 2005; Demiralp ve Kazu, 2012; McGartland, ve diğ., 2003). Kapsam çözümlemesi olarak ifade edilen bu işlem sonrasında, ölçeğin ölçmeye yöneldiği yapıya ilişkin yapı geçerliğini artırma, önsel madde analizleri ve madde faktör sayılarını kontrol edebilme gibi avantajlar sağlayan kapsama yeterliliğine sahip kullanışlı bir ölçek hazırlanabilir (Büyüköztürk, 2005; Demiralp ve Kazu, 2012). Kapsam geçerliliğinin tespitine yönelik farklı teknikler geliştirilmiştir (Cohen, 1960; Davis, 1992; Lawshe, 1975; Lindell ve Brandt, 1999; Lindell, Brandt ve Whitney, 1999; James, Demaree ve Wolf, 1993; Tinsley ve Weiss, 1975). En yaygın olarak kullanılanı Lawshe (1975) tarafından geliştirilen tekniktir (Ayre ve Scally 2014; Veneziano ve Hooper 1997; Wilson, ve diğ., 2012). Bu teknik alternatifleriyle karşılaştırıldığında basit ve kullanışlı olması ile öne çıkar (Wilson, ve diğ., 2012). Ancak, yöntem kısmında ayrıntılı olarak bahsedildiği üzere, Wilson ve arkadaşları (2012) ve Ayre ve Scally (2014) bu tekniğin kapsam geçerlik ölçütü ile ilgili eksikliklerinden bahsederek "Lawshe tekniğini" bu noktada yeniden revize etmişlerdir.

Sonuç olarak hazırlanacak ölçeklerde, ölçekte yer alan maddelerin, uzman görüşleri ile ihtiyaç duyulan olgusal ve/veya yargısal verileri kapsamada ve toplamada yeterli olup olmadıklarının belirlenmesi önem arz etmektedir.

Bu çalışmada amaç örnek bir çalışma üzerinden kapsam geçerlik oranları ve kapsam geçerlik indeksi tanıtılarak kapsam geçerlik çalışmalarının ne şekilde yapılabileceği ile ilgili bir yol haritası sunmaktır.

2. Yöntem

Kapsam çözümlemesi yapılan örnek çalışma kuramsal süreçler ile kuramsal formun elde edildiği nitel bir çalışmadır. Ancak elde edilen nitel veriler, uzman görüşleri arasındaki uyumluluğun test edilebilmesi ve çözümlemenin daha kolay yapılabilmesi için nicel verilere dönüştürülmüştür.

Kapsam geçerliliğinin tespiti konusunda yapılan çalışmalar göz önüne alındığında (Ayre ve Scally 2014; Cohen, 1960; Davis, 1992; James, Demaree, ve Wolf, 1993; Lawshe, 1975; Lindell ve Brandt, 1999; Lindell, Brandt, ve Whitney, 1999; Tinsley ve Weiss, 1975; Veneziano ve Hooper 1997; Wilson, ve diğ., 2012), Lawshe (1975) tarafından geliştirilen teknik hem basit hemde kullanışlı olmasıyla öne çıkmaktadır (Ayre ve Scally 2014; Veneziano ve Hooper 1997; Wilson, ve diğ., 2012). Çalışmada bu durum dikkate alınmıştır.

Veneziano ve Hooper (1997) kapsam geçerliliği ile ilgili yaptığı derlemede Lawshe (1975)'nin ortaya koyduğu tekniğin objektif ve kesin sonuçlar verdiğini nedenleri ile ortaya koymuştur. Sonrasında Wilson ve arkadaşları (2012) ve Ayre ve Scally (2014) yaptıkları çalışmalarla bu durumu teyit etmişlerdir. Ancak Wilson ve arkadaşları (2012) ve Ayre ve Scally (2014), Lawshe (1975)'nin yayınladığı kapsam geçerlilik ölçütü (critical CVR) değerleri ile ilgili eleştirilerde bulunmuşlar, sonrasında bu değerleri yeniden hesaplayarak revize etmişlerdir. Çalışmada bu eleştiriler dikkate alındığında "modifiye Lawshe tekniğini" olarak isimlendirebileceğimiz, düzenlemesi yeniden yapılan kapsam geçerlilik hesaplamalarına ait bir yol haritası, örnek çalışma üzerinden, aşağıda sunulmuştur.

Çalışmada, temelde aşağıdaki aşamalar takip edilmiştir;

1. Uzman grubunun oluşturulması.

2. Aday ölçek formunun hazırlanması ve uzman görüşlerinin eldesi.

3. Verilerin analizi.

- **a.** Kapsam geçerlik oranlarının (KGO=CVR=Content Validity Ratio) hesaplanması.
- **b.** Kapsam geçerlik indeksinin (KGİ=CVI= Content Validity Index) hesaplanması.

4. Kapsam geçerlik oranları ve kapsam geçerlik indeksi ölçütlerine göre her bir maddenin ölçekte yer alıp almamasına karar verilmesi (Bu kısım bulgular kısmında gösterilmiştir.).

2.1. Uzman Grubunun Oluşturulması

Kapsam geçerliliğinin tespiti için yapılacak hesaplamalarda objektif sonuçlar elde edilebilmesinde uzmanların niteliği ve sayısı (5-40 arası) büyük önem taşımaktadır (Ayre ve Scally, 2014; Lawshe, 1975; Veneziano ve Hooper, 1997; Wilson, ve diğ., 2012). Örnek çalışmada bu durum dikkate alınmıştır. Çalışmaya katılan uzmanlar uygun örnekleme yöntemiyle, erişim kolaylığı açısından, Atatürk Üniversitesi Kâzım Karabekir Eğitim Fakültesi'nde görev yapan öğretim elemanlarından (26 kişi) seçilmiştir. Çalışmaya katılan uzmanların cinsiyet ve unvanlarına göre dağılımı Tablo 1 de verilmiştir.

Tablo 1. Çalışmaya Katılan Uzmanların Cinsiyet ve Unvanlarına Göre Dağılımı

	Akademik Unvan					
Cinsiyet	**Prof. Dr.**	**Doç. Dr.**	**Dr. Öğretim Üyesi**	**Öğretim Görevlisi**	**Arş. Gör. (Dr.)**	**Toplam**
Bayan	0	1	3	0	1	5
Bay	5	8	5	2	1	21
Toplam	5	9	8	2	2	26

2.2. Aday Ölçek Formunun Hazırlanması ve Uzman Görüşlerinin Eldesi

Lawshe (1975) tekniğinde her bir madde için uzman görüşleri "madde hedeflenen yapıyı ölçüyor", "madde yapı ile ilişkili ancak gereksiz" ve "madde hedeflenen yapıyı ölçmüyor" şeklinde üçlü derecelendirilmektedir. Ancak Lawshe (1975) tekniğinde yer alan uzman görüşlerine ait derecelendirmeler "Uygun", "Uygun Ancak Düzeltilmeli" ve "Çıkartılmalı" şeklinde yeniden düzenlenmiştir. Benzer derecelendirmeler literatürde (Alisinanoğlu ve Şimşek, 2013; Ateş Çobanoğlu, 2013) mevcuttur.

Örnek çalışmada, veri toplama aracı olarak kullanılacak uzman formu toplam 69 maddeden oluşmaktadır. Bunlardan 66 madde kapsam geçerliği tespitine yöneliktir. Geri kalan 3 madde ise uzmanlara ait demografik özellikleri vermektedir. Uzmanlardan formda yer alan 66 maddenin her biri için yukarıdaki derecelendirmelerden birisini işaretlemeleri istenmiştir. Ölçeğin kapsam geçerlilik oranlarının hesaplanabilmesinde "Uygun" 3, "Uygun Ancak Düzeltilmeli" 2 ve "Çıkartılmalı" 1 olacak şekilde puanlanmıştır. Ayrıca araştırmacılar tarafından Lawshe (1975) tekniğindeki derecelendirmeye ilaveten uzmanlardan düzeltilmeli seçeneğini işaretlemiş ise "Cevabınız düzeltilmeli ise ne şekilde olması gerektiği ile ilgili öneriniz nedir?"; çıkartılmalı seçeneğini işaretlemiş ise "Cevabınız çıkartılmalı ise neden?" şeklinde her bir madde için görüşlerini yazmaları istenmiştir.

2.3. Verilerin Analizi

Çalışmada, ölçekte yer alacak maddelerin kapsam geçerliğinin tespiti için, uzman görüşleri doğrultusunda elde edilen nitel veriler KGO ve KGİ hesaplanarak nicel verilere dönüştürülmüştür. Bu dönüştürme işleminde önce KGO daha sonra KGİ hesaplanmıştır. Hesaplamalar Microsoft Excel 2013 programıyla yapılmıştır (Şekil 1).

2.3.a. Kapsam Geçerlik Oranlarının (KGO) Hesaplanması

KGO, maddelerin ölçekte olması ya da olmamasına ilişkin kapsam geçerliğine dayalı bir madde istatistiği olup aşağıdaki formüle göre hesaplanır (Lawshe, 1975);

$$KGO = \frac{Nu - N/2}{N/2} \quad \text{veya} \quad KGO = \frac{Nu}{N/2} - 1 \qquad \text{(Eşitlik.1)}$$

Burada; Nu, maddeye "Uygun" diyen uzman sayısını ve N ise maddeye ilişkin görüş belirten toplam uzman sayısını göstermektedir.

KGO, -1 (mutlak red) ile +1 (mutlak kabul) arasında bir değere sahiptir. Katılımcıların tamamı ölçekteki herhangi bir maddeyi "Uygun" olarak derecelendirirse o maddenin KGO değeri 1 olur. Ancak hata payı veya şans eseri olabilme göz önüne alınarak bu değer 0,99 kabul edilmiştir. Bu

durumda katılımcıların yarısından fazlası ölçekteki herhangi bir maddeyi "Uygun" olarak derecelendirirse KGO 0 ile 0,99 arası bir değer alacaktır. KGO eşitliğine göre; uzmanların yarısı ölçekteki maddeye ilişkin ''Uygun" şeklinde görüş bildirdiklerinde KGO=0, yarısından fazlası "Uygun" şeklinde görüş bildirmiş ise KGO>0 ve uzmanların yarısından azı "Uygun" şeklinde görüş bildirmiş ise KGO<0 olacaktır. Eğer KGO oranı 0 (sıfır) veya negatif (sıfırdan küçük) değer alıyorsa bu şekilde bir değere sahip maddenin kapsam geçerliliği yoktur. Dolayısıyla ölçekteki bu maddeler doğrudan elenir (Ayre ve Scally 2014; Lawshe, 1975; Wilson, ve diğ., 2012).

Lawshe (1975)'ye göre pozitif bir değere sahip her bir madde için α=0,05 anlamlılık düzeyinde kapsam geçerlilik ölçütüne (KGÖ=CVRcritical=critical CVR) bakılması gerekir. KGÖ=CVRcritical ölçekteki her bir maddeye uygun denilme oranının şans eseri olma durumunun ortadan kalkması ve bir maddenin gerçekten uygun olup olmadığına karar verilebilmesi için ihtiyaç duyulan KGO değeridir. KGÖ değerleri bir maddenin gerekli veya gereksiz olduğuna karar verilebilmesi için en az kaç uzmanın onay vermesi gerektiğini belirlemede kullanılmaktadır. Ancak KGÖ değeri uzman sayısına göre farklılık arz etmektedir. Lowell Schipper bu kritik değerleri belli ölçülere göre hesaplamış, Lawshe (1975) de Schipper'ın bu hesaplamalarını kullanarak KGO'ların Minimum/Kritik Değerlerine ait bir tablo hazırlamış, ancak yaptığı çalışmada bu tablodaki değerlerin nasıl hesaplandığı ile ilgili bilgi vermemiştir (Ayre ve Scally, 2014; Lawshe, 1975; Veneziano ve Hooper, 1997; Wilson, ve diğ., 2012). Daha sonra Wilson ve arkadaşları (2012) Lawshe (1975)'nin yayınladığı bu tabloda hesaplama hatası tespit ederek farklı yöntemlerle bu kritik KGO değerlerini (KGÖ) yeniden hesaplamışlardır. Ayre ve Scally (2014) ise gerek Wilson ve arkadaşları (2012)'nın gerekse Lawshe (1975)'nin kritik KGO değerleri üzerine yapılan çalışmalarını değerlendirmişlerdir. Bu değerlendirmede Wilson ve arkadaşlarının (2012) yapmış olduğu kritik değer hesaplamalarından yola çıkılarak maddeler üzerine gerekli veya gereksiz şeklinde karar verilemeyeceğini, bu kritik değerlerin güvenilir olmadığını ifade etmişlerdir. Buna karşılık Lawshe (1975)'nin makalesinde gösterilen kritik değerlerin ise maddeler ile ilgili karar verme konusunda güvenilir olduğunu ancak yeterli olmadığını belirtmişlerdir. Ayre ve Scally (2014), Lawshe'den farklı olarak çalışmaya katılacak uzman sayısının 1 kişi dahi artması veya azalması durumunda KGO kritik değerlerinin değişeceğine dikkat çekmişler ve uzman sayısına göre α=0,05 anlamlılık düzeyinde KGO'ların minimum/kritik değerleri (KGÖ) için, farklı istatistiksel analizler yaparak, yeni bir tablo hazırlamışlardır (Tablo 2).

Sonuç olarak Lawshe (1975), Wilson ve arkadaşları (2012) ve Ayre ve Scally (2014)'nın KGÖ'ye yönelik çalışmalarında ortaya koyduğu savlar beraberce değerlendirildiğinde örnek çalışmada Ayre ve Scally (2014)'nin ortaya koyduğu KGÖ değerlerinin esas alınmasına karar verilmiştir. Örnek çalışmaya bu esas dikkate alınarak bakıldığında, α=0,05 anlamlılık düzeyinde 26 uzman için KGÖ=CVRcritical değerinin 0.385 olduğu Tablo 2'de görülmektedir.

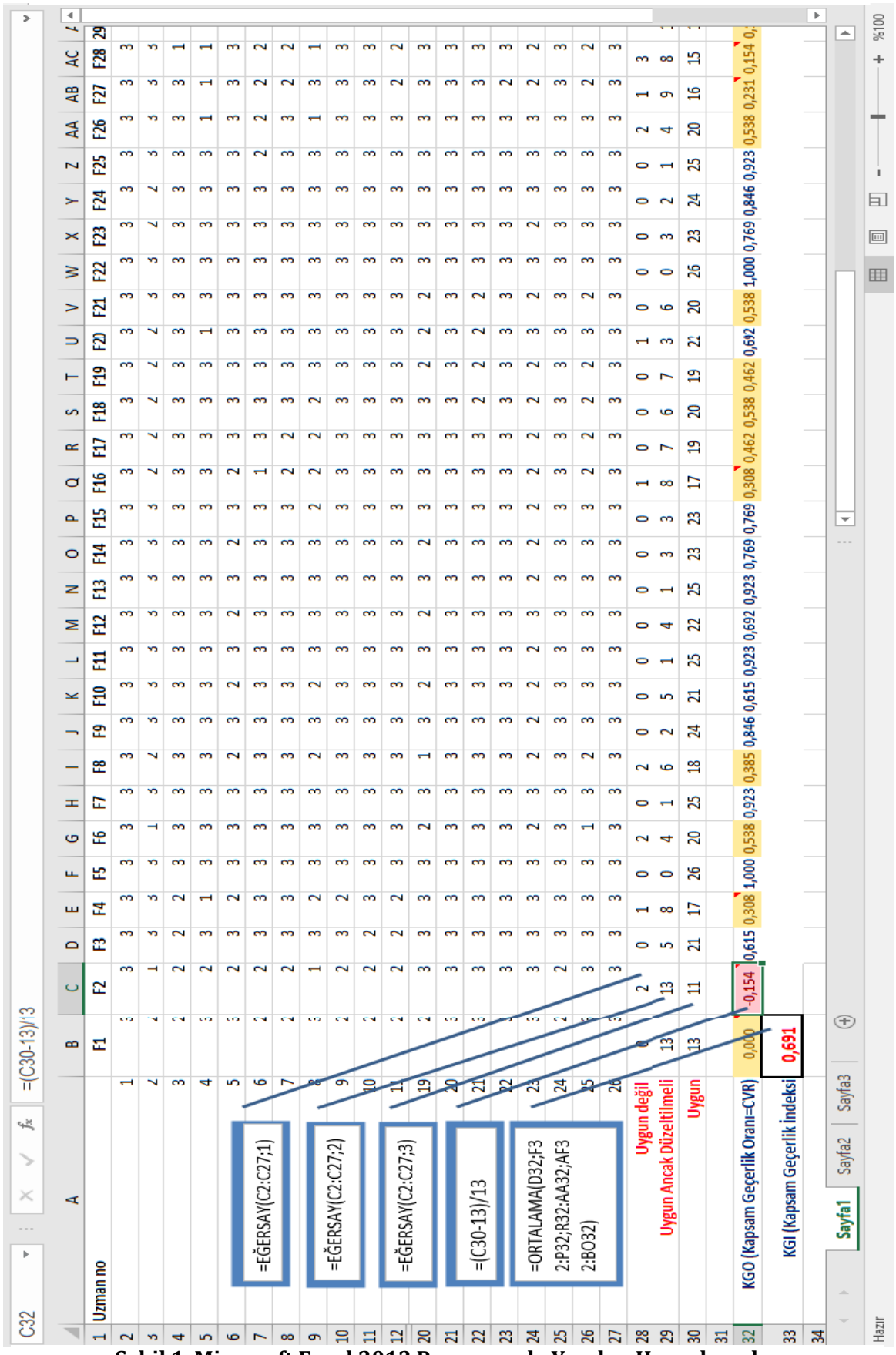
2.3.b. Kapsam Geçerlik İndeksinin (KGİ=CVI) Hesaplanması

KGO belirli maddelerin kabulünde veya reddinde kullanılan istatiksel bir araçtır. KGO'nun tespitiyle maddeler ölçeğe dâhil olmak üzere tanımlandıktan sonra, KGİ testin tamamı için hesaplanır. Bu durumda ölçekte yer almasına karar verilen maddelerin KGO değerlerinin ortalaması hesaplanarak KGİ değeri elde edilir. Kapsam geçerlilik indeksinin korelasyon katsayısı ile karıştırılmaması dikkat edilmesi gereken bir husustur. KGİ tanımlanmış bir iş performans alanının fonksiyon kapasitesi ile üzerinde inceleme yapılan testte gözlemlenen performans arasında fark edilen örtüşmenin boyutlarını temsil eder. İşlevsel olarak KGİ test maddeleri ile iş performans alanı arasındaki örtüşmenin ortalama yüzdesidir (Lawshe, 1975).

Yurdugül (2005) ölçülmek istenilen özelliğin kaç boyutta toplandığının önemli olduğunu belirterek, özellik birden fazla boyutta toplanmış ise her bir boyut için KGİ elde edilmesi gerektiğini vurgulamıştır. Bu vurgu dikkate alındığında, çalışmada kullanılan örnek ölçeğe ait tek bir boyut olduğu varsayımı ile KGİ değeri tek boyut için 0,691 olarak hesaplanmıştır (Tablo 3).

Tablo 2. α=0,05 Anlamlılık Düzeyinde KGO'ların Minimum/Kritik Değerleri (KGÖ= $CVR_{critical}$) (Ayre ve Scally, 2014).

Uzman Sayısı	Minimum Değer	Uzman Sayısı	Minimum Değer
5	1.000	23	0.391
6	1.000	24	0.417
7	1.000	25	0.440
8	0.750	**26**	**0.385**
9	0.778	27	0.407
10	0.800	28	0.357
11	0.636	29	0.379
12	0.667	30	0.333
13	0.538	31	0.355
14	0.571	32	0.375
15	0.600	33	0.333
16	0.500	34	0.353
17	0.529	35	0.314
18	0.444	36	0.333
19	0.474	37	0.297
20	0.500	38	0.316
21	0.429	39	0.333
22	0.455	40	0.300



Şekil 1. Microsoft Excel 2013 Programıyla Yapılan Hesaplamalar

3. Bulgular

Örnek çalışmaya ait uzmanlardan elde edilen görüşler doğrultusunda her bir madde için hesaplanan KGO değerleri ve örnek ölçeğin tamamına yönelik KGİ değeri Tablo 3'de verilmiştir.

Tablo 3. Uzman Görüşleri Doğrultusunda Örnek Ölçeğe Ait Kapsam Geçerlik Oranları ve İndeksi

Madde	Uygun	Düzeltilmeli	Çıkartılmalı	KGO	Madde	Uygun	Düzeltilmeli	Çıkartılmalı	KGO
1	13	13	0	0.000*	**34**	22	3	1	0.692
2	11	13	2	-0.154*	**35**	23	3	0	0.769
3	21	5	0	0.615	**36**	20	5	1	0.538
4	17	8	1	0.308*	**37**	21	5	0	0.615
5	26	0	0	1.000	**38**	19	7	0	0.462
6	20	4	2	0.538	**39**	22	3	1	0.692
7	25	1	0	0.923	**40**	22	3	1	0.692
8	18	6	2	0.385	**41**	22	3	1	0.692
9	24	2	0	0.846	**42**	21	4	1	0.615
10	21	5	0	0.615	**43**	24	2	0	0.846
11	25	1	0	0.923	**44**	24	2	0	0.846
12	22	4	0	0.693	**45**	20	6	0	0.538
13	25	1	0	0.923	**46**	19	7	0	0.462
14	23	3	0	0.769	**47**	19	5	2	0.462
15	23	3	0	0.769	**48**	21	3	2	0.615
16	17	8	1	0.308*	**49**	23	3	0	0.769
17	19	7	0	0.462	**50**	22	4	0	0.692
18	20	6	0	0.538	**51**	23	3	0	0.769
19	19	7	0	0.462	**52**	22	3	1	0.692
20	22	3	1	0.692	**53**	21	4	1	0.615
21	20	6	0	0.538	**54**	21	3	2	0.615
22	26	0	0	1.000	**55**	21	4	1	0.615
23	23	3	0	0.769	**56**	19	5	2	0.462
24	24	2	0	0.846	**57**	25	0	1	0.923
25	25	1	0	0.923	**58**	21	3	2	0.615
26	20	4	2	0.538	**59**	23	2	1	0.769
27	16	9	1	0.231*	**60**	21	3	2	0.615
28	15	8	3	0.154*	**61**	23	3	0	0.769
29	15	10	1	0.154*	**62**	23	2	1	0.769
30	15	8	3	0.154*	**63**	23	3	0	0.769
31	23	3	0	0.769	**64**	21	4	1	0.615
32	23	3	0	0.769	**65**	22	3	1	0.692
33	23	3	0	0.769	**66**	22	2	2	0.692
				Toplam Uzman Sayısı			**26**		
				Kapsam Geçerlik Ölçütü (KGÖ)			**0.385**		
				Kapsam Geçerlik İndeksi (KGİ)			**0.691**		

*KGÖ=CVRcritical değeri (0.385)'nin altındaki maddeler

Toplamda 26 uzmanın maddelere ilişkin belirtmiş oldukları görüşler üzerinden "Eşitlik 1"'de verilen ifade yardımıyla KGO değerleri elde edilmiş, KGO oranı 0 (sıfır) veya negatif (sıfırdan küçük) değere sahip Madde 1 ve Madde 2 ölçekten doğrudan çıkarılmıştır (Tablo 3). Daha sonra, sıfırdan büyük değere sahip maddelerin KGO değerlerinin istatistiksel olarak anlamlılığına, dolayısıyla ölçekte kalıp kalmayacağına Tablo 2'de yer alan KGÖ değerlerine bakılarak karar verilmiştir. Tablo 2'ye göre çalışmaya katılan 26 uzman için KGÖ değeri 0.385'dir. Buna göre Tablo

3 incelendiğinde ölçekte yer alan, iki madde çıkarıldıktan sonra (1 ve 2. Madde) kalan 64 maddeden 6'sının (4, 16, 27, 28, 29 ve 30. Maddeler) KGO değerinin KGÖ değerinden (=0.385) küçük olduğu görülmüştür. Dolayısıyla bu kritik değerin altında bir değere sahip bu 6 madde de geliştirilmekte olan örnek ölçekten çıkarılmıştır. Bu durumda örnek ölçekte 58 madde kalmıştır.

Örnek çalışmaya ait KGİ değeri, ölçekten 8 madde çıkarıldıktan sonra hesaplanmış ve 0.691 olarak tespit edilmiştir. Elde edilen KGİ değerinin KGÖ değerinden büyük olması (KGİ>KGÖ) ölçekte kalan maddelerinin kapsam geçerliğinin istatistiksel olarak anlamlı olduğunu gösterir (Ateş Çobanoğlu, 2013; Batdı, 2013; Lawshe 1975; Öngöz, 2011). Diğer bir deyişle KGİ ve KGÖ değerleri karşılaştırıldığında, KGİ değeri KGÖ değerinden küçük ise ölçekte kalan maddelerin kapsam geçerliğine sahip olmadığı ifade edilir (Lawshe, 1975). Buna göre örnek çalışmada elde edilen değerlerden KGİ (0,691)> KGÖ (0,385) olduğundan hazırlanacak ölçekte kalan maddelere (58 madde) ait kapsam geçerliliği istatistiksel olarak anlamlı düzeydedir. Örnek çalışmanın kapsam geçerlik analizleri sonrasında kalan 58 maddenin en az bir uzman tarafından düzeltilmeli görüşü belirtilenleri tekrar gözden geçirilerek uzman görüşleri doğrultusunda düzeltilmiş ve nihayetinde kapsam geçerliliğine sahip bir ölçek elde edilmiştir.

4. Tartışma ve Sonuç

Bu örnek çalışmayla, kapsam geçerlik oranları ve kapsam geçerlik indeksi tanıtılmış, ölçek geliştirme çalışmalarında kapsam çözümlemesinin basit ve kullanışlı bir şekilde yapılabilmesi için bir yol haritası sunulmuştur.

Ölçek geliştirme çalışmalarında ölçme aracının güvenirliği ölçme aracının kararlılığı ile ilgilidir. Bu durumda yapılacak işlemlerde ölçüm aracıyla neyi ölçmek istenildiği, maddelerin amaç doğrultusunda doğru ölçme yapıp yapmadığı dikkate alan geçerlik çalışmaları öne çıkmaktadır. Ölçme araçlarından elde edilecek ölçümler hangi amaçla kullanılacaksa ölçme geçerliği de o amaca bağlı olarak değişecektir (Ercan ve Kan, 2004; Kalaycı, 2016; Leech, Barrett ve Morgan, 2005).

Ölçeğin ve ölçekteki her bir maddenin amaca ne derece hizmet ettiği ortaya koyan kapsam geçerlik çalışmaları, uzman görüşlerine dayalı nitel çalışmaları istatistiksel nicel çalışmalara dönüştüren bir süreçtir (Ercan ve Kan, 2004; Shuttleworth, 2016; Yurdugül, 2005). Ölçek geliştirme çalışmalarında kapsam çözümlemesi olarak ifade edilen bu süreçle ölçek ve ölçekte yer alan maddelerin kalitesi artırılmakta, işlem-zaman kolaylığı sağlanmaktadır (Büyüköztürk, 2005; Demiralp ve Kazu, 2012; Yurdagül, 2005). Uzmanların niteliği ise sonuçların tutarlı ve yansız olabilmesinde oldukça önemlidir (Shuttleworth, 2016; Yurdagül, 2005).

Sonuç olarak bu çalışmada belirtilen yol haritası izlenerek, herhangi bir ölçeğin ölçmeye yöneldiği yapıya ilişkin yapı geçerliğini artırma, önsel madde analizleri ve madde faktör sayılarını kontrol edebilme gibi avantajlar sağlayan, kapsama yeterliliğine sahip kullanışlı bir ölçek hazırlanabilir.

Kaynaklar

Alisinanoğlu, F., & Şimşek, Ö. (2013). Okul öncesi dönemdeki çocukların yazmaya hazırlık becerilerini değerlendirme kontrol listesinin geçerlik ve güvenirlik çalışması. *Kastamonu Eğitim Dergisi, 21(3), 1163-1176.*

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory* (2nd ed.). Prospect Heights, IL: Waveland Press. s.310.

Ateş Çobanoğlu, A. (2013). Eğitsel websitelerini değerlendirmeye yönelik bir ölçek önerisi. *Eğitim Teknolojileri Araştırmaları Dergisi, 4* (1), 17.09.2016 tarihinde https://s3.amazonaws.com/academia.edu.documents/31488732/etad-

2013.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1521720476&Signature=VKi6GcOkO8a7rw0Z%2BiSn8pKNda8%3D&response-content-disposition=attachment%3B%20filename%3DAtes_A._2013_._Egitsel_websitelerini_deg.pdf adresinden alınmıştır.

Ayre, C., & Scally A. J. (2014). Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development, 47* (1), 79–86. doı: 10.1177/0748175613513808.

Basham, A., & Sedlacek, W. E. (2009). Validity. In American Counseling Association (Ed.), *The ACA encyclopedia of counseling* (p. 557). Alexandria, VA: American Counseling Association.

Batdı, V. (2013). İşbirlikli öğrenmenin yabancı dil öğretimindeki önemine ilişkin öğretmen görüşleri. *Eğitim ve Öğretim Araştırmaları Dergisi, 2* (1), s. 158-165.

Brains, C., Willnat, L,, Manheim, J., & Rich, R. (2011). *Empirical Political Analysis* 8th edition. Boston, MA: Longman. p.76.

Brinkman. W.-P. (2009). *Design of a questionnaire ınstrument. handbook of mobile technology research methods.* ISBN 978-1-60692-767-0. pp. 31-57. Netherlands: Nova Publisher.

Büyüköztürk, Ş. (2005). Anket geliştirme. *Türk Eğitim Bilimleri Dergisi, 3* (2), 133-151.

Cohen, R. J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Cronbach, J. L., & Meehl, P. E. (1951). Construct validity in psychological tests. *Psychological Bulletin, 52* (4), 281-302.

Davis L.L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5,* 194-197.

Demiralp, D., & Kazu, H. (2012). İlköğretim birinci kademe programlarının öğrencilerin yansıtıcı düşünmelerini geliştirmedeki katkısına yönelik öğretmen görüşleri. *Pegem Eğitim ve Öğretim Dergisi, 2* (2), 29-38.

Ercan, İ., & Kan, İ. (2004). Ölçeklerde güvenirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi, 30* (3), 211-216.

Frank-Stromberg, M., & Olsen, S. J. (2004). *Instruments for clinical health-care research*. London, England: Jones & Bartlett.

James, L. R., Demaree, R. G., & Wolf, G. (1993). *R*wg: An assessment of within-group interrater agreement. *Journal of Applied Psychology*, *78*, 306–309.

Kalaycı, Ş. (2016). *SPSS uygulamalı çok değişkenli istatistik teknikleri* (7th ed.). Ankara: Asil Yayın Dağıtım, 426 s.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, *28*(4), 563-575.

Leech, N.L., Barrett, K.C., & Morgan, G.A. (2005). *SPSS for Intermediate Statistics*: *Use and interpretation* (2th ed.). Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, London, 240.

Lindell, M. K., & Brandt, C. J. (1999). Assessing interrater agreement on the job relevance of a test:

A comparison of the *CVI*, *T*, *rwg(j)*, and *r*wg(j)* indexes. *Journal of Applied Psychology*, *84*, 640–647.

Lindell, M. K., Brandt, C. J., & Whitney, D. J. (1999). A revised index of interrater agreement for multi-item ratings of a single target. *Applied Psychological Measurement*, *23*, 127–135.

McGartland Rubio, D., Berg-Weger, M., Tebb, S.S., Lee, E.S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research , 27*(2), 94-104.

McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (7th ed.). New York, NY: Pearson.

Otrar, M., & Argın F. S. (2015). Öğrencilerin sosyal medyaya ilişkin tutumlarını belirlemeye yönelik bir ölçek geliştirme çalışması. *Eğitim ve Öğretim Araştırmaları Dergisi, 4* (1), s. 391-403

Öngöz, S. (2011). Elektronik ders kitabı değerlendirme formunun geliştirilmesi: geçerlik ve güvenirlik çalışması. *11th Internatıonal Educatıonal Technology Conference, IETC* - May 25-27, 2011 Istanbul, Turkey, Proceedings Book (Volume II), s.1481-1485.

Özbek, R., Kahyaoğlu, M., & Özgen, N. (2007). Öğretmen adaylarının öğretmenlik mesleğine yönelik görüşleri. *Sosyal Bilimler Dergisi, 9*(2), s.221-232.

Rubio, D. M, Berg-Weger, M., Tebb, S., Lee, S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. S*ocial Work Research (ProQuest Psychology Journals), 27* (2), s.94-104.

Shuttleworth M. (2016). Content validity. Retrieved from https://explorable.com/content-validity.html.

Tinsley, H. E. A., & Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology*, *22*, 358–376.

Veneziano L. & Hooper J. (1997). A method for quantifying content validity of health-related questionnaires. *American Journal of Health Behavior, 21* (1), s. 67-70.

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*, 197–210. doi:10.1177/0748175612440286.

Yıldırım A, & Şimşek H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri* (9. Basım). Ankara: Seckin Yayıncılık, 446 s.

Yurdugül, H. (2005). Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması, *XIV. Ulusal Eğitim Bilimleri Kongresi*, Pamukkale Üniversitesi Eğitim Fakültesi, 28–30 Eylül Denizli, 16.06.2016 tarihinde http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf adresinden alınmıştır.

Extended Summary

1. Introduction

In education and psychology, it has a great importance to determine the views and attitudes in an appropriate way (Özbek, Kahya and Özgen, 2007; Yurdugül, 2005). This situation puts forward the studies of scale development. However, the top priority aspect that should be taken into consideration while developing a scale, is the validity of the measurement tool (Brains, Willnat, Manheim ve Rich, 2011). Likewise, the validity is an important criterion for the credibility and generalizability of a scientific study (Basham ve Sedlacek, 2009; Yıldırım ve Şimşek, 2013). In the studies of validity, one should elaborate to ensure a consistent relation between the scale desired to be developed and the properties desired to be evaluated. In such a study which is called as content validity, it is revealed that in which extent each and every item serves to the purpose. By this means, instead of the irrelevant statements that are off-topic, a place will be given to the statements that have a strong representational ability while developing a scale (Brinkman, 2009; Basham and Sedlacek, 2009; Ayre and Scally, 2014; Frank-Stromberg and Olsen, 2004; Wilson et al., 2012). In the studies of content validity which is also called as logical, rational or context validity (Shuttleworth, 2016), a preliminary study must be carried out in which a sufficient number of expert views are taken in order to determine that into what extent a scale item covers the property desired to be measured (Allen and Yen, 2002; Basham and Sedlacek, 2009; Lawshe, 1975; Ercan and Kan, 2004; Brinkman, 2009; McMillan and Schumacher, 2010; Büyüköztürk, 2005; Otrar, 2015; Rubio et al., 2003). Because, in this type of preliminary studies, in the case that the number of experts are kept in an adequate number (between 5 and 40), the validity of the scale which is going to be developed, will also be high (Ayre and Scally, 2014; Lawshe, 1975; Veneziano and Hooper, 1997; Wilson et al., 2012).

Different techniques for the determination of the content validity are developed (Lawshe, 1975; Davis, 1992; Tinsley & Weiss, 1975; James, Demaree, and Wolf's, 1993; Lindell, Brandt, and Whitney's, 1999; Lindell & Brandt,1999; Cohen, 1960). The most widely used one is the technique developed by Lawshe (1975) (Veneziano and Hooper 1997; Wilson et al., 2012; Ayre and Scally 2014). When the technique developed by Lawshe is compared with the other alternatives of it, it comes to the front with its features such as being simple and handy (Wilson et al., 2012).

The aim of the present study is to present a road map in which the answer of the question of how the content validity studies can be carried out is given with introducing content validity ratio (CVR) and Content Validity Index (CVI) through a sample study.

2. Method

Content analysis is done for the sample study. And the sample study is a qualitative one in which the hypothetic form is acquired with the hypothetic processes. However, the obtained qualitative data were converted to quantitative data to analyze and test the compatibility among the expert opinions more easily.

Considering the studies that are done on the determination of the content validity (Davis, 1992; Veneziano ve Hooper 1997; Lawshe, 1975; Wilson et al., 2012; Ayre ve Scally 2014; Tinsley & Weiss, 1975; James, Demaree, and Wolf's, 1993; Lindell, Brandt, and Whitney's, 1999; Lindell & Brandt,1999; Cohen, 1960) , a road map is presented through a sample study in which "modified the Lawshe's technique" is used for the content validity measurements. In the study, the technique developed by Lawshe (1975) and the critical content ratio determined by Ayre and Scally (2014) are taken into consideration.

26 experts are joined to the sample study. As a data collection tool an 'expert form' in which 66 items are cross-examined, is used. In order to be able to measure the content validity ratios of the items in the sample form, 3, 2 and 1 points are given for the options "Proper", "Should be corrected", and "Should be removed" respectively.

In addition, experts are asked to write their views for the each item that they marked the option "should be removed" and "should be corrected". For this purpose open ended questions that are "If your answer is 'should be corrected', what will your suggestion about what form it should be?" and "If your answer is 'should be removed', why?", are added just below the options of "should be removed" and "should be corrected".

In the sample study, qualitative data obtained from experts views are converted to quantitative data by calculating content validity ratio (CVR) and content validity index (CVI). In this process firstly the content validity ratio (CVR) and then the content validity index (CVI) are calculated. The calculations are performed by using Microsoft Excel 2013.

According to CVR; if the half of the experts express their view as 'Proper' about the item in the scale, CVR will be equal to 0(CVR=0); and if more than half of the experts express their view as 'Proper', CVR will be greater than 0(CVR>0); and if less than half of the experts express their view as 'Proper', CVR will be less than 0(CVR<0). If CVR is equal to 0 or less than 0, there is no content validity of the item which has value like that. Thereby, those items are eliminated directly (Lawshe, 1975; Wilson et al., 2012; Ayre ve Scally 2014).

CVRcritical values are used to determine the minimum number of experts needed to approve so that an item can be determined to be necessary or unnecessary. However CVRcritical values differ in respect to the number of experts. Ayre and Scally (2014) pointed out that CVRcritical values will change if the number of experts that are going to join to the study increases or decreases by even only 1 person. When the arguments put forward by Lawshe (1975), Wilson et al. (2012) and Ayre and Scally (2014) on their studies of CVRcritical are all evaluated, it is decided to predicate the CVRcritical values revealed by Ayre and Scally (2014) upon.

The CVR is a statistic item that is used in rejection or retention of certain items. After the items are identified as to be included in the scale with determination of CVR, the CVI is calculated for the whole test. In this case, the CVI value is obtained by calculating mean of the CVR values of items which will retain in the scale (Lawshe, 1975).

3. Findings, Discussion and Results

In the light of views of 26 experts engaging in the study, the content validity ratios are computed for each and every item in the expert form and the two items with zero and negative values are directly excluded from the scale. Based on the 'critical content validity ratio' derived by Ayre and scally (2014), it was decided whether the items with values above zero will retain in scale or not. Thus, 6 items which have values of critical content ratio below 0,385 which is determined for 26 experts were excluded from the scale. For the 58 items retained in the sample scale, content validity index (0,691) is computed. Accordingly, among the values computed in the sample study, CVI (0,691) is greater than CVR (0,385). So, it can be said that the content validity of the items (58 items) retained in the scale which is going to be developed, is statistically significant. As a result, a scale which has a content validity is successfully acquired.

Via this sample study, content validity ratio (CVR) and content validity index (CVI) are introduced and a road map is provided so that the content analysis can be done easily and conveniently in the scale development studies.

As a result, by following the road map pointed out in this study, a useful scale, which has sufficiency in coverage and provides advantages such as being able to control the number of item factors, offering preliminary item analysis and increasing the construct validity of a construction which is intended to be measured by any scale, can be developed.

Araştırma makalesi: Yeşilyurt, S. ve Çapraz, C. (2018). Ölçek geliştirme çalışmalarında kullanılan kapsam geçerliği için bir yol haritası. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi. 20*(1), 251-264.