

## Eğitim Alanında Ölçek Geliştirme Süreci: Bir İçerik Analizi Çalışması

### Scale Development Process in Educational Field: A Content Analysis Research

Melek Gülşah ŞAHİN<sup>a</sup>, Nağihan BOZTUNÇ ÖZTÜRK<sup>b</sup>

<sup>a</sup>Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara, Türkiye.

<sup>b</sup>Hacettepe Üniversitesi, Rektörlük, Ankara, Türkiye.

#### Özet

Bu çalışmada eğitim alanında geliştirilmiş ölçeklerin "ölçek geliştirme süreçlerinin" belirli ölçütler açısından incelenmesi amaçlanmıştır. Çalışma kapsamına 2010-2016 yılları arasında Türkiye'de eğitim alanında yayınlanmış olan 7 dergideki 72 makale dâhil edilmiştir. Makalelerin incelenmesinde "Ölçek Geliştirme Süreci Kontrol Formu" kullanılmıştır. Elde edilen bulgulara göre; makalelerde ölçülen yapının kuramsal temelleri tanımlanmakta ve ölçme aracının amacı belirtilmektedir. Ölçek maddelerin yazılma sürecinde kaynak tarama sıklıkla tercih edilirken neredeyse tüm çalışmalarda konu alanı uzmanlarından uzman görüşüne başvurulmuştur. Madde havuzu oluşturulurken olumsuz maddelere yer verildiği belirtilen makale sayısı az olmakla birlikte, hiçbir makalede kontrol maddelerine yer verilmesiyle ilgili bilgiye rastlanmamıştır. Yapı geçerliğinin belirlenmesinde AFA ve DFA genellikle beraber kullanılırken ölçüt geçerliği neredeyse hiç incelenmemiştir. Güvenirliğin kestirilmesinde ise tüm çalışmalarda iç tutarlılık anlamında güvenilirlik belirlenirken, katsayı olarak ise Cronbach  $\alpha$  katsayısı tercih edilmiştir.

**Anahtar Kelimeler:** Ölçek, ölçek geliştirme, güvenilirlik, geçerlik, içerik analizi

#### Abstract

In this study, it is aimed to examine "scale development process" developed in the field of education in terms of certain criteria. Within the scope of the study, 72 articles in 7 journals published in the field of education in Turkey between 2010-2016 were included. "Scale Development Process Control Form" was used in the analysis of the articles. According to findings; the theoretical basis of the structure measured in the articles is described and the purpose of the scale is specified. In the development of items, literature review is often preferred, but in almost all studies, expert opinion is sought from subject matter experts. While there was a small number of articles in which negative items were included in the creation of a item pool, there was no information about the inclusion of the control items in any article. While EFA and CFA are generally used together when the validity is determined, the criterion validity is hardly examined. In the estimation of reliability, internal consistency is always used in all studies and Cronbach  $\alpha$  coefficient is preferred for internal consistency calculation.

**Keywords:** Scale, scale development, reliability, validity, content analysis

## 1. Giriş

Eğitim ve psikolojide bireyler hakkında bilgi toplamada, ilgi, tutum, kaygı, motivasyon, kişilik vb. psikolojik özelliklerinin ölçülmesinde ölçeklerden yararlanılmaktadır. Ölçekler bireylerin doğrudan gözlenemeyen özelliklerinin ne ve nasıl olduğunu ortaya çıkarmaktadır (Erkuş, 2012). Literatür incelendiğinde araştırmacıların araştırmaları kapsamında var olan ölçekleri kullandıkları ya da ölçekleri oluşturdukları görülmektedir. Her iki çalışmada da öncelikle ölçülen değişkenin yapısı ve bu değişkenle ilişkili yapıların tanımlanması gerekmektedir (Cohen ve Sverdlık, 2010). Sonrasında tanımlanan yapı ile ilgili "mevcut ölçek" olup olmadığının belirlenmesi eğer var ise de "ne kadar duyarlı bir ölçek" olduğunun ve mevcut ölçeğin duyarlılığının yüksek olması durumunda ise bir uyarılma çalışmasının yapılmasının mı yoksa yeni bir ölçek oluşturmanın mı daha anlamlı olduğunun sorgulanması gerekmektedir (Tavşancıl, 2002). Hem ölçek geliştirmenin hem de uyarılmanın kendine özgü bir takım kuralları vardır ve uzmanlık gerektirmektedir (Erkuş, 2012; Hambleton ve Patsula, 1999).

Ölçek geliştirmeye karar veren araştırmacının sekiz aşamayı izlemesi beklenir (DeVellis, 2003). İlk olarak "ne ölçülmek istendiği" açıkça belirlenmelidir. Ölçülmek istenilen değişkenin ve ilişkili değişkenlerin kuramsal yapısı detaylı olarak ortaya konulmalıdır. İkinci aşamada madde havuzu oluşturulmalıdır. Üçüncü aşamada ölçme aracının formatına karar verilmesi yer almaktadır. Sırasıyla maddelerin uzmanlar tarafından gözden geçirilmesi, madde geçerliğinin sağlanması, ölçeğin uygulanması, maddelerin değerlendirilmesi, ölçeğin son halinin verilmesi aşamaları gelmektedir.

Madde havuzunun oluşturulmasında amaca uygun olarak farklı yöntemler kullanılabilir. Eğer doğrudan gözlenecek davranışlar ölçülecekse gözlem, eğer bir kuram temele alınacaksa literatür tarama ya da tutum gibi yapıların ölçülmesi söz konusu ise örneklem grubuna benzer bir grupta kompozisyon yazdırılması tercih edilebilir (Erkuş, 2012). Maddeler oluşturulurken, ölçülmek istenilen yapının olabilecek tüm alt boyutlarının gözlemlenmesi esastır (Tezbaşaran, 1997). Ölçekte kullanılması tasarlanan madde sayısının, olanak varsa üç-dört katının veya daha fazlasının hazırlanması gerekmektedir. Çünkü maddelerin bir grup üzerinde denenmesinden sonra amaca hizmet etmeyen yeterli güvenilirlik ve geçerlikte olmayan maddeler çıkabilir. Çok sayıda madde arasından istenen kapsayıcılıkta ve ayırt edicilikte madde seçme olanağı bu şekilde daha fazlalaşmaktadır (Tezbaşaran, 1997). Ölçme aracının formatına karar verilmesinde sonuçların nasıl yorumlanacağı önemlidir. Bu aşamada çeşitli ölçek türlerinden uygun olanın seçilmesi (Thurstone, Guttman, Likert vb.) ve cevap kategorilerinin belirlenmesi söz konusudur (DeVellis, 2003). Ölçek türlerinden en çok likert tipi ölçek kullanışlı olması ve dereceleme düzeyini artırma ile eşit aralık ölçeğinde ölçme sonuçları vermeleri nedeniyle sıklıkla tercih edilmektedir (Tezbaşaran, 1997).

Maddelerin ölçülen yapıyı temsil etme gücü, anlaşılabilirliği, kapsam geçerliğinin değerlendirilmesi için uzman görüşlerinden yararlanılabilir (DeVellis, 2003; Erkuş, 2003; Yurdugül, 2005). Erkuş'a (2003) göre uzman görüşleri panel tartışmasıyla ve istatis-

tiksel yöntemler yoluyla alınabilir. Bu yöntemlerden biri veya ikisi birlikte uygulanabilir. Maddelerin uygunluğunun belirlenmesinden sonra ölçek formunun geçerliğinin ve güvenilirliğinin belirlenmesi amacıyla örneklem grubuna uygulanması söz konusudur. Örneklem grubunun evreni temsil etmesi önemlidir. Bu nedenle amaca uygun örneklem grubu ve sayısı belirlenmelidir. Nunually (1978) 300 kişilik bir örneklemin ölçek çalışmalarında yeterli olduğunu belirtmiştir. Comrey ve Lee (1992) ise 100 kişinin zayıf, 200 kişinin orta, 300 kişinin iyi, 500 kişinin çok iyi ve 1000 kişinin mükemmel olduğunu ifade etmiştir. Yine literatürde örneklem grubunun uygulanan madde sayısı ile ilişkili olarak belirlenmesine dair ölçütlerde söz konusudur. Örneklem sayısı madde sayısının en az beş katı (Bryman ve Cramer, 2001), 10 katı (Nunually, 1978), 15 katı (Gorush, 1983) olmasına dair görüşler söz konusudur (Delice ve Ergene, 2015). Çok az sayıda örneklem grubu kullanmak maddeler arası kovaryans örüntüsünün kararlı olmamasına neden olabilir yani maddenin iç tutarlılığı etkisi başka örneklem grupları için geçerli olmayabilir. Bir başka olası olumsuzluk ise örneklemin evreni temsil etmemesi olarak belirtilebilir (DeVellis, 2003). Örneklem grubuna ilişkin sayısının belirlenmesinden sonra araştırma kapsamındaki gruba ölçeğin uygulanması söz konusudur. Uygulamadan sonra maddelerin değerlendirilmesi yani maddelerin performansının ön incelenmesi söz konusudur. Bu aşamada madde toplam korelasyonları, madde varyansları, madde ortalamaları, faktör analizi, alfa katsayısı hesaplanabilir (DeVellis, 2003). Ölçeğin geçerlik ve güvenilirlik kanıtlarının sunulmasından sonra ölçeğe son hali verilebilir. Tüm bu bilgilerden yola çıkılarak ölçek geliştirilmesinin oldukça zahmetli bir iş olduğu belirtilebilir. Araştırmacıların geçerli ve güvenilir sonuçlar elde etmesi için ölçek geliştirme aşamalarını titizlikle uygulaması gerekmektedir.

Ulusal ve uluslararası literatür incelendiğinde; çok sayıda ölçeğin geliştirildiği görülmektedir. Geliştirilen ölçeklerin, ölçek geliştirme aşamalarını inceleyen çalışmalar da literatürde mevcuttur (Acar Güvendir ve Özer Özkan, 2015; Çüm ve Koç, 2013; Delice ve Ergene, 2015; Erkuş, 2016; Hinkin, 1995; Gül ve Sözbilir, 2015; Mor Dirlik, 2014; Slavec ve Drnovsek, 2012; Tavşancıl, Güler ve Ayan, 2014; Tosun ve Taşkesenligil, 2015; Worthington ve Whittaker, 2006). Ölçek geliştirme süreci, zahmetli bir süreç olmasının yanı sıra, ilgili alanlarda daha sonra yapılacak çalışmalar için genellenebilir ve hatalardan arınık sonuçlar elde edilmesini sağlama açısından oldukça önemli ve titizlikle takip edilmesi gereken bir süreçtir. Bu süreçte uygulanan adımların ve eğilimlerin ortaya çıkarılması, eğitim alanında tekrar gözden geçirilmesi gereken noktaların tespiti için önemli bir gösterge olabilir. Bu çalışmada ölçek geliştirme süreci ile ilgili ülkemizdeki çalışmalardan yola çıkarak mevcut durumun tespit edilmesi amaçlanmıştır. Bu amaçla ölçek geliştirme sürecinin aşamalarını içeren form geliştirilmiş ve belirlenen dergilerde yer alan likert tipi ölçek geliştirmeye ilişkin çalışmalar incelenmiştir. Ölçek geliştirme sürecinin incelendiği çalışmaların ülkemizde sınırlı olması bu çalışmanın önemini oluşturmaktadır. Ayrıca bu çalışmadan elde edilen bulgulardan yararlanılarak, ölçek geliştiricilere ve araştırmacılara sürecin iyileştirilmesine ve ölçek geliştirme çalışmalarının geçerliği ve güvenilirliğinin artırılmasına yönelik öneriler getirilecektir.

## 2. Yöntem

Bu çalışmada, Türkiye’de eğitim alanında yayınlanan SSCI ve/veya ULAKBİM ulusal veri tabanında indekslenen dergilerdeki ölçek geliştirme makaleleri içerik analizi yöntemi kullanılarak incelenmiştir. İçerik analizinde benzer verilerin belirli kavram ve temalar çerçevesinde bir araya getirilerek düzenlenmesi ve yorumlanması söz konusudur (Yıldırım ve Şimşek, 2013). İçerik analizi en genel anlamda; meta-analiz, meta sentez (tematik içerik analizi) ve betimsel içerik analizi olarak üç başlık altında toplanabilir (Çalık ve Sözbilir, 2014).

Bu çalışmada; eğitim alanındaki ölçek geliştirme süreçlerinde geçerli olan eğilimlerin ve sonuçların tanımlayıcı bir boyutta değerlendirilmesi söz konusu olduğundan araştırma betimsel içerik analizi olarak adlandırılabilir (Çalık ve Sözbilir, 2014).

### İçerik Analizine Tabi Tutulan Makaleler

Çalışmaya Türkiye’de yayınlanan SSCI ve/veya ULAKBİM ulusal veri tabanında indekslenen ve makalelerin tam metinlerine çevrimiçi ortamdan ulaşılabilen eğitim alanında en az 15 yıldır yayın yapan dergilerin yaklaşık yarısını temsil edecek şekilde toplamda 7 akademik dergi dâhil edilmiştir. Bu dergilerden 3 tanesi hem SSCI hem de ULAKBİM ulusal veri tabanında indekslendiği için doğrudan seçilmiştir (İncelemeye alınan dergilerden biri Ocak 2016 tarihi itibarıyla SSCI’de yer almamaktadır. Bu nedenle belirtilen tarihten sonra yayınlanmış ve çalışmaya alınmış makaleler SSCI’da yer almadığı göz önünde bulundurulmuştur). Sadece ULAKBİM’de indekslenen 4 dergi ise tesadüfi olarak seçilmiştir. Seçilen dergilerde yer alan ölçek geliştirme çalışmaları belirlenmiş ve kodlamalar verilmiştir. Bu çalışmaların %50’si tesadüfi yolla seçilerek araştırmaya dâhil edilmiştir.

İncelemeye alınacak ölçek geliştirme makaleleri belirlenirken makale başlığında “ölçek, ölçek geliştirme, geçerlik ve güvenilirlik çalışması” kelimeleri taranmıştır. Tarama sonucunda amacı sadece ölçek geliştirme olan makaleler çalışmaya dâhil edilmiş, hem ölçek geliştirme hem de geliştirilen ölçeği araştırma için kullanan makaleler çalışma dışında bırakılmıştır. Bu ölçütler göz önünde bulundurularak, toplamda 72 ölçek geliştirme makalesi içerik analizine tabi tutulmuştur. İnceleme sırasında 72 makaleden 69’unun likert tipi ölçek geliştirme sürecinin ele alındığı görülmüş ve çalışma bu makaleler üzerinden yürütülmüştür.

İncelenen makalelerin yer aldıkları dergilere bakıldığında %55,6’sının hem SSCI hem ULAKBİM, %44,4’ünün ise sadece ULAKBİM ulusal veri tabanında indekslendiği görülmüştür.

### Veri Toplama Aracı

Çalışmanın geçerliğini ve güvenilirliğini tehlikeye atmamak için incelenen makalelerin içerik analizi sürecinde Şahin, Boztunç Öztürk ve Kelecioğlu tarafından geliştirilen “*Ölçek Geliştirme Süreci Kontrol Formu*” kullanılmıştır. Bu formun kullanılma amacı

içerik analizi sürecinde standart ölçütlerin belirlenmesidir. Kontrol formu hazırlanırken ölçek geliştirme süreci aşamaları ve ölçek geliştirme sürecinde en sık karşılaşılan problem durumları ele alınmıştır. Hazırlanan formun kapsamını değerlendirmek için eğitimde ölçme ve değerlendirme alanında iki araştırma görevlisi ve üç öğretim elemanı toplam beş uzmanın görüşleri alınmış ve uzman görüşleri doğrultusunda gerekli güncellemeler yapılarak forma son hali verilmiştir.

Formun son hali Tablo 2’de verilmiştir. Tabloda formda yer alan ölçütler ve içerik analizinde kullanılan kodlamalar yer almaktadır.

**Tablo 2. Ölçek Geliştirme Süreci Kontrol Formu**

Ölçütler	Kodlama
No.	
Ölçeğin adı	
Yazarlar	
Basıldığı dergi	
Dergi SSCI’da yer alıyor mu?	(0) Hayır (1) Evet
Basıldığı yıl	
<b>Makalenin Giriş Bölümü</b>	
Ölçeğin amacının belirtilmesi	(0) Hayır (1) Evet
Ölçülen yapının kuramsal temellerinin tanımlanması	(0) Hayır (1) Evet
<b>Madde Havuzunun Hazırlanması</b>	
Madde Yazma Süreci	(1) Görüş alma (2) Kaynak tarama (3) Diğer
Olumsuz maddeler	(0) Yok (1) Var
Kontrol maddeleri	(0) Yok (1) Var
Uzman görüşü alınıp alınmadığı	(0) Hayır (1) Evet
Uzman sayısı	
Uzmanların nitelikleri	(1) Konu alanı (2) Ölçme ve değerlendirme (3) Dil (4) Diğer (5) Belirtilmemiş
Ön deneme uygulamasının yapılması	(0) Hayır (1) Evet
<b>Geliştirilen Form</b>	
Madde sayısı	
Cevap kategorisi sayısı	
Geliştirilen formun maddelerinin verilmesi	(0) Hayır (1) Evet
<b>Deneme Uygulamasının Örnekleme</b>	
Büyüklüğü	
Örneklemin seçimi	
<b>Analizler</b>	
Yapı geçerliği için yapılan çalışmalar	(1)AFA (2) DFA (3) Ölçüt gruplar (4) Yapılmamış
Hem AFA hem DFA yapılan çalışmalardaki gruplar	(1) Aynı (2) Farklı
Ölçüt geçerliği	(1) Uygunluk geçerlik (2) Farklılık geçerliği (3) Yapılmamış
Madde Analizi	(1) Alt-üst grup analizi (2) Madde test korelasyonu (3) Yapılmamış
Güvenirlilik	(1) İç tutarlılık (2) Test-tekrar test (3) Eşdeğer formlar (4) İki yarıya bölme (5) Yapılmamış

### Verilerin Analizi

Çalışma kapsamında yer alan 72 makale araştırmacılar tarafından dengeli bir şekilde paylaşılmış ve belirli bir süre içinde içerik analizinin tamamlanması hedeflenmiştir. Bireysel olarak içerik analizine başlamadan önce, araştırmacılar arasındaki tutarlılığı sağlamak adına öncelikle 72 makale arasından tesadüfi olarak seçilen 6 makale her iki araştırmacı tarafından da içerik analizine tabi tutulmuştur. Elde edilen veriler ile tutarlılık anlamında güvenilirliği kestirmek için Miles ve Huberman (1994) tarafından önerilen aşağıdaki formül kullanılmıştır.

Hesaplamalar sonucunda araştırmacılar arası tutarlılığın 0,81 olduğu görülmüştür. Bu sonuç yapılan değerlendirmenin güvenilir olduğuna ilişkin kanıt sağlamaktadır. Bunun yanı sıra araştırmacılar analiz sürecinde karşılaştıkları sorunlar ile ilgili görüş alışverişi yapıp farklı anlaşılmalari en aza indirmeye çalışmışlardır.

Çalışma kapsamına giren makalelerin içerik analizi tamamlandıktan sonrasında veriler SPSS ortamında tek bir dosyada birleştirilmiştir. Sonlandırılan veri setinde hatalı kodlamaların olup olmadığı incelenmiştir. Kodlamasında hata olduğu düşünülen bir makale diğer araştırmacı tarafından tekrar içerik analizine alınmıştır. Gerekli düzeltmeler yapıldıktan sonra ölçek geliştirme kontrol formunun genel başlıkları altında elde edilen verilerin frekans ve yüzde analizleri yapılmış ve tablolaştırılmıştır.

### 3. Bulgular ve Yorumlar

Araştırmanın bulguları; araştırmanın tanıtımı, maddelerin hazırlanması, geliştirilen form, ölçeğin uygulanması, geliştirilen form ve analiz alt başlıkları altında verilmiştir.

#### Araştırmanın Tanıtımı

Araştırmanın tanıtımı kapsamında incelenen ölçütlere bakıldığında; makalelerin % 83,6'sında ölçülen yapının kuramsal temellerinin tanımlandığı ve %100'ünde ölçme aracının amacının belirtildiği görülmüştür. Geliştirilen ölçek türleri incelendiğinde %95,9'unun likert tipi ölçek olduğu görülmektedir. Bunların yanı sıra %2,7'si açık uçlu sorulardan oluşan bir ölçek iken %1,4 ise ankettir. Çalışmanın amacı likert tipi ölçeklerin geliştirme sürecinin incelenmesi olduğu için bundan sonraki aşamalarda analizler bu 69 ölçek üzerinden yürütülmüştür.

#### Maddelerin Hazırlanması

Maddelerin hazırlanma aşamasında; madde yazma süreci, yazılan maddelerde olumsuz maddeler ve kontrol maddelerine yer verilip verilmediği, uzman görüşü alınıp alınmadığı, uzmanların nitelikleri ve ön deneme uygulamasının yapılma durumu incelenmiştir. Madde yazma sürecine ait bulgular Tablo 4'te verilmiştir.

**Tablo 4. Madde Yazma Sürecinde Başvurulan Kaynakların Dağılımı**

Madde Yazma Süreci	Frekans (f)	Yüzdeler (%)
Görüş alma	2	2,9
Kaynak tarama	39	56,5
Görüş alma - Kaynak tarama	21	30,4
Görüş alma - Diğer	2	2,9
Görüş alma - Kaynak tarama - Diğer	2	2,9
Diğer	2	2,9
Belirtilmemiş	1	1,5
Toplam	69	100

Tablo 4'e bakıldığında; geliştirilen ölçeklerin madde havuzunun hazırlanması aşamasında, çalışmaların %2,9'unda sadece görüş alındığı ve %56,5'inde sadece kaynak taraması yapıldığı görülmektedir. Bunların yanı sıra çalışmaların %30,4'ünde hem görüş alınmış hem kaynak taraması yapılmıştır. Elde edilen verilere bakıldığında, madde havuzunu hazırlamak için kaynak tarama yönteminin neredeyse tüm çalışmalarda kullanıldığı görülmektedir.

Maddeler hazırlanırken incelenen çalışmaların %31,9'unda olumsuz maddelerin yer aldığı belirtilirken, çalışmaların hiçbirinde kontrol maddesine yer verildiğine değinilmemiştir.

Hazırlanan ölçek maddelerinin uzmanlar tarafından incelenmesi durumuna bakıldığında neredeyse tüm çalışmalarda (%98,6) uzman görüşlerine yer verildiği görülmüştür. Uzman görüşüne yer verilen çalışmaların %23,53'ünde görüş alınan uzmanların sayısı belirtilmezken, diğer çalışmalarda uzman sayısı 2-31 ranjında değişim göstermektedir. Uzman sayısı dağılımının detayına bakıldığında; görüş alınan uzman sayısı çalışmaların %26,92'sinde 5'in altında, %51,92'sinde 5-10 aralığında, %15,38'inde 10-20 arasında ve %5,77'sinde 20 ve üstünde olacak şekilde değişim görülmüştür.

Tablo 6'da görüş alınan uzmanların niteliklerine yer verilmektedir.

**Tablo 6. Görüş Alınan Uzmanların Nitelikleri**

Uzman Nitelikleri	Frekans (f)	Yüzdeler (%)
Konu Alanı	50	72,46
Eğitimde Ölçme ve Değerlendirme	20	28,99
Dil	24	34,78
Diğer	18	26,09
Belirtilmemiş	10	14,49

Tablo 6'ya bakıldığında; çalışmaların %72,46'sında konu alanı uzmanı, %28,57'sinde eğitimde ölçme ve değerlendirme uzmanı, %34,78'inde Türkçe dil uzmanı yer aldığı görülmektedir. Bunların dışında çalışmaların %26,09'unda diğer uzmanlık alanları da yer almıştır. Ancak çalışmaların %14,49'unda uzmanların nitelikleri belirtilmemiştir. Hem konu alanı hem ölçme ve değerlendirme uzmanı hem de dil uzmanından görüş alındığı çalışmalara bakıldığında; %12 olduğu görülmüştür.

Çalışmalarda ön deneme uygulamasının yapılma yüzdesi incelendiğinde; %75,4'ünde ön deneme uygulamasının yapılmadığı görülmüştür.

#### Geliştirilen Form

Geliştirilen formların bölümünde tercih edilen cevap kategorisi sayısı ve formun ya da formda yer alan maddelerin çalışmalarda

yer verilip verilmemesi incelenmiştir. Geliştirilen ölçeklerin cevap kategorisi sayılarının dağılımına bakıldığında; geliştirilen ölçeklerin neredeyse tamamı (%91,3) 5 cevap kategorisine sahiptir. 5 kategorili ölçeklerin dışında kalan ölçeklere bakıldığında; 3 ve 4 kategorili ölçeklerin ilköğretim düzeyindeki öğrencilere yönelik, 7 kategorili ölçeklerin ise üniversite düzeyindeki öğrenciler için geliştirildiği görülmüştür. Geliştirilen ölçeklere ya da ölçeklerin maddelerine makalelerde yer verilip verilmemesine bakıldığında çalışmaların %46,4'ünde ne ölçeklere ne de ölçek maddelerine yer verilmediği görülmüştür.

### Ölçeğin Uygulanması

Ölçeğin uygulanması sürecinde; deneme grubunun seçim yöntemi, grubun büyüklüğü ve madde başına katılımcı sayısı incelenmiştir. Uygulamanın yapıldığı grupların seçilme yöntemi Tablo 8'de verilmiştir.

**Tablo 8. Deneme Uygulamasının Yapıldığı Grubun Seçimi**

Uygulamanın Yapıldığı Gruplar	Frekans (f)	Yüzdeler (%)
Örneklem	17	24,65
Çalışma grubu	34	49,28
Belirtilmemiş	18	26,09
Toplam	69	100

Tablo 8'e göre; geliştirilen ölçeklerin uygulandığı gruplar çalışmaların %49,28'inde çalışma grubu olarak adlandırılmıştır. Çalışmaların %24,65'inde örneklem üstünde uygulamalar yapılırken, %26,09'unda ise uygulama yapılan grubun nasıl seçildiği belirtilmemiştir. Örneklem seçilerek yapılan bazı çalışmalarda örnekleme yöntemi belirtilmezken, örnekleme yöntemi belirtilen çalışmalarda tabakalı, basite seçkisiz, küme, amaçsal, uygun, gelişigüzel ve ulaşılabilir örnekleme yöntemlerinin tercih edildiği görülmüştür.

Uygulama yapılan grupların büyüklükleri ise 61 - 1542 arasında değişkenlik göstermiştir. Grup büyüklüklerin kategorize edilmiş dağılımları Tablo 9'da verilmiştir.

**Tablo 9. Grup Büyüklüklerinin Dağılımı**

Grup Büyüklükleri	Frekans (f)	Yüzdeler (%)
100'ün altı	1	1,45
100-199	10	14,49
200-299	15	21,74
300-499	28	40,58
500-999	12	17,39
1000 ve üstü	3	4,35
Toplam	69	100

Tablo 9'a göre; incelenen ölçek geliştirme çalışmalarının %40,58'inin uygulandığı grup büyüklüğü 300-499 arasında olduğu, %21,74'ünün de 200-299 arasında olduğu görülmektedir. 500-999 arasında grup büyüklüğüne sahip çalışmalar incelenen makalelerin %17,39'unu oluştururken, %4,35'i ise 1000 ve üstünü oluşturmaktadır.

Madde başına katılımcı sayısının dağılımı Tablo 10'da verilmiştir.

**Tablo 10. Madde Başına Katılımcı Sayısının Dağılımı**

Madde Başına Katılımcı Oranı	Frekans (f)	Yüzdeler (%)
5'in altı	19	27,54
5-9	27	39,13
10-14	10	14,49
15-19	9	13,04
20 ve üstü	4	5,80
Toplam	69	100

Tablo 10'a bakıldığında; madde başına düşen katılımcı sayısının 5-9 arasında daha yoğun olduğu (%39,13) görülmektedir. 5'in altında %27,54, 10-14 arasında %14,49, 15-19 arasında %13,04 ve 20 ve üstünde %5,08 oranında madde başına katılımcı bulunmaktadır.

### Analiz

Analiz bölümünde geliştirilen ölçeklerin geçerlik ve güvenilirliklerinin kestirilmesi için kullanılan yöntemler incelenmiştir. Geçerliğin belirlenmesi sürecinde öncelikle yapı geçerliğinin belirlenmesinde kullanılan yöntemler incelenmiştir. Elde edilen bulgular Tablo 11'de verilmiştir.

**Tablo 11. Yapı Geçerliliğinin Belirlenmesinde Kullanılan Yöntemlerin Dağılımı**

Yapı Geçerliliği Belirleme Yöntemleri	Frekans (f)	Yüzdeler (%)
Açımlayıcı Faktör Analizi	29	42,0
Doğrulamalı Faktör Analizi	4	5,8
Açımlayıcı ve Doğrulamalı Faktör Analizi	32	46,4
Açımlayıcı Faktör Analizi ve Ölçüt Gruplar	2	2,9
Açımlayıcı Faktör Analizi, Doğrulamalı Faktör Analizi ve Ölçüt Gruplar	2	2,9
Toplam	69	100

Tablo 11'e göre; incelenen ölçek geliştirme çalışmalarının %46,4'ünde Açımlayıcı Faktör Analizi (AFA) ve Doğrulamalı Faktör Analizinin (DFA) birlikte kullanıldığı görülmektedir. Sadece AFA kullanılan çalışmalar incelenen ölçek geliştirme çalışmalarının %42,0'sini oluştururken, sadece DFA kullanılan çalışmalar ise %5,8'lik bir değere sahiptir. AFA, DFA ve ölçüt grupların birlikte uygulandığı çalışmalar da incelenen çalışmaların %2,9'unu oluşturmaktadır. Buna göre; yapı geçerliliği belirleme yöntemlerinden AFA ve DFA'nın birlikte oldukça sık kullanıldığı sonucu elde edilmiştir. AFA ve DFA'nın birlikte kullanıldığı çalışmalara bakıldığında, bu makalelerin %61,11'inde aynı grubun kullanıldığı görülmektedir. %38,89'unda ise farklı grupların verileri ile AFA ve DFA yapılmıştır.

Geçerlik kapsamında yer alan bir diğer geçerlik türü olan ölçüt geçerliliğinin belirlenmesinde kullanılan yöntemler Tablo 12'de verilmiştir.

**Tablo 12. Ölçüt Geçerliliğinin Belirlenmesinde Kullanılan Yöntemlerin Dağılımı**

Güvenirlilik Kestirme Yöntemleri	Frekans (f)	Yüzdeler (%)
Uygunluk	6	8,7
Yapılmamış	63	91,3
Toplam	69	100

Tablo 12'ye göre; incelenen çalışmaların %91,3'ünde ölçüt geçerliliği sınamamıştır. Ölçüt geçerliliğinin yapıldığı çalışmalarda ise uygunluk geçerliliği tercih edilmiştir. Elde edilen bulgulara göre; ölçek geliştirme sürecinde ölçüt geçerliliğinin çoğunlukla incelenmediği görülmüştür.

İncelenen çalışmalarda, madde analizi belirlenmesinde kullanılan yöntemler Tablo 13'te verilmiştir.

**Tablo 13. Madde Analizinde Kullanılan Yöntemlerin Dağılımı**

Madde Analizleri	Frekans (f)	Yüzdeler (%)
Alt-Üst Grup Analizi	5	7,2
Madde-Test Korelasyonu	26	37,7
Alt-Üst Grup Analizi ve Madde-Test Korelasyonu	19	27,5
Yapılmamış	19	27,5
Toplam	69	100

Tablo 13'e göre; incelenen çalışmaların %37,7'sinde yalnızca madde-test korelasyonu incelenmiştir. Araştırmaların yaklaşık olarak dörtte birlik bir kısmında (%27,5) ise madde analizlerinin yapılmadığı görülmektedir. Alt-üst grup analizi ve madde-test korelasyonun birlikte incelendiği çalışmalar da makalelerin %27,5'ini oluşturmaktadır.

Güvenirliliğin kestirilmesinde kullanılan yöntemler incelenmiş ve elde edilen bulgular Tablo 14'te verilmiştir.

**Tablo 14. Güvenirliliğin Kestirilmesinde Kullanılan Yöntemlerin Dağılımı**

Güvenirlilik Kestirme Yöntemleri	Frekans (f)	Yüzdeler (%)
İç Tutarlılık	45	65,2
İç Tutarlılık ve Test Tekrar Test	9	13,0
İç Tutarlılık ve Eşdeğer Formlar	1	1,4
İç Tutarlılık ve İki Yarıya Bölme	8	11,6
İç Tutarlılık, Test Tekrar Test ve İki Yarıya Bölme	4	5,8
Yapılmamış	2	2,9
Toplam	69	100

Tablo 14'e göre; incelenen çalışmalarda güvenirliliğin kestirilmesinde çoğunlukla (%65,2) iç tutarlılık yönteminin tek başına kullanılması tercih edilmiştir. İkinci sırada ise İç tutarlılık ve test tekrar test yöntemlerinin birlikte kullanıldığı çalışmalar da %13,0'lük bir yüzdeyi oluşturmaktadır. İç tutarlılık ve iki yarıya bölme yöntemlerinin birlikte kullanıldığı çalışmalar da incelenen çalışmaların %11,6'sını oluşturmaktadır. Sonuç olarak ölçek geliştirme çalışmalarında iç tutarlılık yönteminin çoğunlukla tercih edildiği ve güvenirliliğin kestirilmesinde ise çalışmaların tamamında Cronbach  $\alpha$  katsayısı hesaplanmış olduğu belirlenmiştir.

#### 4. Tartışma

Bu çalışmada eğitim alanındaki ölçek geliştirme süreçlerinin incelenmesi amaçlanmıştır. Çalışma kapsamında 72 ölçek geliştirme makalesi ele alınmış ve likert tipi ölçek geliştirme sürecini anlatan 69 makale çeşitli ölçütlere dikkat edilmiştir. Araştırma kapsamında incelenen makalelerin yarısından fazlası hem SSCI hem ULAKBİM ulusal veri tabanında indekslenmektedir. Makalelerin çoğunda ölçülen yapının kuramsal temelleri tanımlanmış ve tümünde ölçme aracının amacı belirtilmiştir. Bu sonuç ölçek geliştirme süreçlerini incelemiş farklı çalışmaların sonuçları ile benzerlik teşkil etmektedir (Delice ve Ergene, 2015; Çüm ve Koç, 2013).

Madde havuzunun oluşturulmasında kaynak taramanın önemli bir yeri olduğu görülmektedir. Bazı makalelerde ise hem kaynak taraması yapılmış hem de görüş alınmıştır. Acar Güvendir ve Özer Özkan (2015)'in çalışmalarında da ölçek geliştirme sürecinde madde havuzu oluşturulurken literatürden sıklıkla yararlanıldığı belirtilmiştir.

Hazırlanan ölçek maddeleri uzman görüşüne tabii tutulması açısından incelendiğinde neredeyse tüm çalışmalarda uzman görüşüne başvurulduğu görülmüştür. Literatürde de incelenen çalışmaların çoğunda uzman görüşüne yer verildiği görülmektedir (Acar Güvendir ve Özer Özkan, 2015; Çüm ve Koç, 2013). Uzmanların niteliklerine bakıldığında; çalışmaların yaklaşık dörtte üçünde konu alanı uzmanlarına mutlaka yer verildiği, dil ve ölçme-değerlendirme uzmanları ile daha az çalışıldığı sonucu elde edilmiştir. Bu durum ölçek geliştirme sürecinde disiplinler arası çalışmaların yaygın olmadığını bir göstergesi olabilir. Benzer bulgulara farklı çalışmalarda da rastlanmaktadır (Acar Güvendir ve Özer Özkan, 2015). Ancak ölçek maddelerinin, konu alanını temsiliyetinin irdelenmesinin yanı sıra psikometrik olarak doğruluklarının da irdelenmesi geliştirilen ölçeğe önemli katkılar sağlayacaktır. Bu nedenle konu alanı uzmanları ile birlikte özellikle ölçme değerlendirme uzmanlarından ve dil uzmanlarından da uzman görüşünü istenilmelidir.

Deneme uygulanmasında genellikle çalışma grubu tercih edilirken, az da olsa bazı çalışmalarda ise basit tesadüfi örnekleme yöntemi kullanılarak örneklem seçilmiştir. Erkuş (2012), ölçek geliştirme çalışmalarında ölçülen özelliğin ranjının örnekleme oranının önemli olduğunu ve katılımın gönüllülük esasına dayanması gerektiğini vurgulamaktadır. Örneklemin seçkisiz olması gönüllülük durumuna ters düşeceğinden, deneme uygulamasın için en uygun örnekleme yönteminin genellikle amaçlı örnekleme olduğunu belirtmektedir. İncelenen ölçek geliştirme çalışmalarının yaklaşık %3'ünde amaçsal örnekleme kullanılmıştır. Uygulanan grupların büyüklüklerine bakıldığında çalışmaların yarısından fazlasında 300 ve üstünde grup büyüklüklerinin kullanıldığı görülmüştür. Bunun yanı sıra; ölçek maddesi başına katılımcı sayısına bakıldığında 5.00-9.99 oranının daha yoğun olarak kullanıldığı görülmüştür. Literatürde farklı örneklem büyüklükleri ölçütlerine rastlanmasıyla birlikte esas önemli olanın doğru örneklemin seçilmesidir (Erkuş, 2016; Gül ve Sözbilir, 2015).

Geliştirilen ölçeklerin yapı geçerliğinin belirlenmesinde AFA ve DFA'nin beraber kullanılması sıklıkla tercih edilmiştir. Bunun yanı sıra sadece AFA da çalışmaların %42,0'sinde kullanılmıştır. Literatüre bakıldığında ise; ölçek geliştirme çalışmalarında yalnızca AFA daha çok tercih edilirken, hem AFA hem de DFA'nin birlikte kullanıldığı çalışmaların daha az olduğu görülmüştür (Gül ve Sözbilir, 2015; Tosun ve Taşkesenligil, 2015). AFA ve DFA'nin birlikte kullanıldığı çalışmalara bakıldığında bu çalışmaların yarısından fazlasında analizlerin aynı veriler kullanılarak yapıldığı, farklı veriler kullanılarak yapılan bazı çalışmalarda ise örneklemin ikiye bölünerek verilerin elde edildiği gözlenmiştir. Literatüre bakıldığında ise, ölçülen yapının belirlenmesi için analizlere AFA ile başlanıp sonrasında belirlenen yapının yeni bir örneklem kullanılarak DFA ile analizlerin devam etmesi gerektiği görülmektedir (Cabrera-Nyugen, 2010; Henson ve Roberts, 2006; Worthington ve Whittaker, 2006). Ayrıca Henson ve Roberts (2006), AFA ve DFA'nin aynı örnekleme yapılması durumunun çok bilgi vermeyeceğini ve yanıltıcı olabileceğini belirtirken; Erkuş (2016) da, örneklemin yapay olarak ikiye bölünüp yarısına AFA diğer yarısına DFA yapılmasının kesinlikle doğru olmadığını belirtmektedir.

Ölçüt geçerliği ise çok az çalışmada incelenmiştir. Benzer çalışmalarda da ölçüt geçerliğinin genellikle incelenmediği görülmektedir (Acar Güvendir ve Özer Özkan, 2015; Gül ve Sözbilir, 2015). Madde analizleri incelendiğinde; madde-test korelasyonu daha sık kullanılırken alt-üst grup analizi ve madde-test korelasyonunun beraber kullanılması da çalışmaların yaklaşık dörtte birinde yer aldığı görülmüştür.

Güvenirlilik belirleme yöntemleri incelendiğinde, daha çok tek uygulamaya dayalı yöntemlerin tercih edildiği görülmüştür. Ölçeklerin tamamında iç tutarlılık anlamında güvenirlilik kestirimi yapılırken, az bir kısmında ise kararlılık anlamındaki güvenirlilik de incelenmiştir. Birden fazla uygulamaya dayalı güvenirlilik kestirme yöntemlerinin kullanımı hem maliyet hem de zaman açısından araştırmacılara ekstra bir yük getirdiği için daha az tercih edildiği düşünülmektedir. İç tutarlılık anlamında güvenirliliğin kestirilmesine detaylı olarak bakıldığında; çalışmaların tümünde Cronbach  $\alpha$  katsayısının hesaplandığı, bunun yanı sıra bazı çalışmalarda ise Spearman Brown iki yarı güvenirliliklerinin de hesaplandığı görülmüştür. İç tutarlılık anlamında güvenirlilik, testin maddeleri arasındaki ilişkilere dayanarak elde edildiği için bir çalışmada aynı amaca hizmet eden farklı güvenirlilik kestirim yöntemlerinin birlikte kullanılmasından ziyade, farklı amaca hizmet eden güvenirlilik kestirim yöntemlerinin bir arada kullanılması daha uygun olabilir. Ulusal literatür incelendiğinde de; ölçek geliştirme sürecinde çalışmaların çoğunda (Delice ve Ergene, 2015) ya da tamamında (Acar Güvendir ve Özer Özkan, 2015; Gül ve Sözbilir, 2015)'in Cronbach  $\alpha$  katsayısının raporlandığı görülmektedir. Cronbach  $\alpha$ 'nın daha çok tercih edilmesinin sebebi olarak bu katsayının hesaplanması için hazır programların mevcut olması düşünülmektedir.

Elde edilen bulgulara genel olarak bakıldığında; ölçek geliştirme süreciyle ilgili sıkıntıların halen devam etmekte olduğu ifade edilebilir. Bu durum ölçek geliştirilmesinin ne kadar önemli ve bir o kadar da zahmetli bir süreç olduğunun henüz anlaşılmadığı sonucunu doğurmaktadır. Doğru bir ölçek geliştirme sürecine sahip olmadan geliştirilen ölçeklerin barındırdığı hatalar yeni yapılacak

araştırmaların da hatalı sonuçlar içermesine sebep olacaktır. Bu nedenledir ki; ölçek geliştirme konusunun üzerinde önemle durulmalıdır.

Ölçek geliştirme sürecinin daha doğru yönetilebilmesi için neler yapılabileceği konusuna gelince akla birkaç öneri gelmektedir. Bunlar; ölçek geliştirme ile ilgili ulusal bağlamda kurum ya da kuruluşların oluşturulması, ölçek geliştirmenin yapılacağı araştırmaya ölçme-değerlendirme uzmanlarının da dâhil edilmesi, ölçme-değerlendirme uzmanlarından oluşan bir grup tarafından ölçek geliştirme çalışmasını yapmak isteyen kişilere danışmanlık verilmesi, lisansüstü programlarda ölçek geliştirme dersinin seçmeli ders olarak sunulması, ölçme-değerlendirme uzmanlarınca ölçek geliştirme süreci ile ilgili standartları detaylı bir şekilde açıklayan doküman oluşturulması ve bu dokümanın dergi hakemleri ile paylaşılması, dergi yönetiminin ölçek geliştirme sürecinin nasıl olması gerektiği ile ilgili hakem kurulunda yer alan ölçme ve değerlendirme uzmanlarıyla ortak bir karar alması, dergi editörünün ölçek geliştirme makalelerine konu alanı uzmanlarının yanı sıra ölçme değerlendirme uzmanlarını da mutlaka hakem olarak ataması olarak sıralanabilir.

## 5. Kaynakça

- Acar Güvendir, M. ve Özer Özkan, Y. (2015) Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 52, 23-33
- Cabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results in the journal of the society for social work and research. *Journal of the Society for Social Work and Research*, 1(2), 99-103.
- Cohen, R.J. ve Swerdik, M.E. (2010). *Psychological testing and assessment*. Boston: McGraw-Hill Companies.
- Comrey, A.L. ve Lee, H.L.(1992). *A first course in factor analysis*, Hillsdale, New Jersey: Erlbaum.
- Çalık, M. ve Sözbilir, M. (2014). İçerik analizinin parametreleri. *Eğitim ve Bilim*, 39(174), 33-38. doi:10.15390/EB.2014.3412
- Çüm, S. ve Koç, N. (2013). Türkiye'de psikoloji ve eğitim bilimleri dergilerinde yayımlanan ölçek geliştirme ve uyarlama çalışmalarının incelenmesi. *Eğitim Bilimleri ve Uygulama*, 12 (24), 115-135.
- DeVellis, R. F. (2003). *Scale development theory and applications* (Second Edition). SAGE Publication, Inc.
- Delice, A. ve Ergene, Ö. (2015). Ölçek geliştirme ve uyarlama çalışmalarının incelenmesi: Matematik eğitimi makaleleri örneği. *Karaelmas Journal of Educational Sciences*, 3, 60-75.
- Erkuş, A. (2007). Ölçek geliştirme ve uyarlama çalışmalarında karşılaşılan sorunlar. *Türk Psikoloji Bülteni*, 13(40) 17.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme-1: Temel kavramlar ve işlemler*. Ankara: Pegem Akademi.
- Erkuş, A. (2016). Ölçek geliştirme ve uyarlama çalışmalarındaki sorunlar ile yazım ve değerlendirilmesi. *Pegem Atıf İndeksi*, 0, 1221-1234. doi:10.14527/9786053183563.075
- Hambleton, R.K. ve Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1(1), 1-30.
- Henson, R. ve Roberts, J. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66, 393-416.
- Hinkin, T. R. (1995). A review of scale development practices in the study in organizations. *Journal of Management*, 21(5), 967-988.
- Gül Ş. ve Sözbilir, M. (2015). Fen ve matematik eğitimi alanında gerçekleştirilen ölçek geliştirme araştırmalarına yönelik tematik içerik analizi. *Eğitim ve Bilim*, 40(178), 85-102. doi: 10.15390/EB.2015.4070
- Miles, M. B. ve Huberman, A. M. (1994). *An expanded sourcebook qualitative data analysis*. London: SAGE Publications.
- Mor Dirlik, E. (2014). Ölçek geliştirme konulu doktora tezlerinin test ve ölçek geliştirme standartlarına uygunluğunun incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5 (2), 62-78.
- Nunnally, J.C. (1978). *Psychometric theory*. NewYork: McGraw-Hill Companies.
- Slavec, A. ve Dmosek, M. (2012). A perspective on scale development in entrepreneurship research. *Economic and Business Review*, 14(1), 39-62.
- Tavşancıl, E., Güler, G. ve Ayan, C. (2014). 2002-2012 yılları arasında Türkiye'de Geliştirilen Bazı Tutum Ölçeği Geliştirme Çalışmalarının Ölçek Geliştirme Süreci Açısından İncelenmesi. IV. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi (Uluslararası Katılımlı) 9-13 Haziran, Hacettepe Üniversitesi, Ankara.
- Tezbaşaran, A. (2008). *Likert tipi ölçek hazırlama kılavuzu*. Ankara: Türk Psikologlar Derneği Yayınları
- Tosun, C. ve Taşkesenligil, Y. (2015). The instruments used in science education in Turkey: a descriptive content analysis. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 15(2), 364- 383.
- Worthington, R. L. ve Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838.
- Yıldırım, A. ve Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin Yayıncılık.

## Extended Abstract

*The process of developing a scale is a process that is very important and required to be followed carefully since the scale can be generalized for further studies in related fields and it provides obtaining error-free results as well as it is a hard process. In this study, articles on developing scale published in the field of education in Turkey in SSCI and/or ULAKBIM (Turkish National Academic Network and Information Center) national database were examined by using content analysis method. 72 articles in 10 journals published in the field of education in Turkey between 2010 and 2016 were included in the study. "Scale Development Process Control Form" was used in the examination of the articles. The opinions of five assessment and evaluation experts were taken to evaluate the content of the form, and the form was finalized by necessary updates according to the expert opinions. The form includes the criteria related to the scale*



development process determined under the titles of "The Introductory Chapter of the Article, Preparing Item Pool, Developed Form, The Group Performed the Test Practice and Analysis." Before the analysis, the coefficient of consistence between the researchers was obtained as 0,81, and it was determined that the interrater reliability related to the evaluations was sufficient.

When the criteria in the introductory chapters of the articles, it was viewed that theoretical fundamentals of the measured construction were defined in 83,6% of the articles, and the purpose of the assessment instrument was stated in all of them. It was seen that literature review method (56,5%) was preferred most and conducting questionnaire method (1,4%) least during the process of item writing in the criteria examined under the title of Preparing Item Pool. The most preferred number of the experts in the studies was changing between 5 and 10 (51,9%). When the qualifications of the experts whose opinions were taken were viewed, it was determined that domain experts (50%), language experts (34,8%) and assessment and evaluation experts in education (28,9%) were preferred respectively.

The distribution of the response categories of the developed scale in the section of the developed form were analysed, and it was observed that 5 categories (91,3%) were selected at most. When it was examined if the developed scales or the items of the scales were included in the articles, it was seen that in 52,2% of the articles included neither scales nor scale items. In the chapter of "Implementation of the Scale," the selection of the group on which the test practice was performed was examined, and it was seen that the group was named as 'study group' in most of the studies (49,3%), and sampling methods were used in 24,7% of the studies. While in some studies conducted by selecting sample, sampling method was not stated, and it was seen that stratified, simple random, cluster, teleological, convenience, accidental and availability sampling methods were used in the studies conducted by selecting sample. In addition, when the size of the group that the test practice was performed, it was found that 300 – 499 range was selected most (40,6%). When the range of the participant number for each item was viewed, it was seen that 5 – 9 participants were preferred most (30,1%) and 20 and over were preferred least (5,8%).

In the Analysis chapter, the methods used to determine the validity and reliability of the scales were analyzed. It was seen that in 46,4% of the scales, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used together and in 42% only Exploratory Factor Analysis was used. It was determined that while item-test correlation was preferred (37,7%) most and sub-superior group analysis was preferred least (7,2%), not any item analysis was used in 27,5% of the studies. When reliability prediction methods were viewed, it was seen that internal consistency method was used most (65,2%) and internal consistency and equivalence forms methods were preferred together least (1,4%). It was also determined that in all studies in which internal consistency method was preferred Cronbach  $\alpha$  coefficient was calculated.

When the findings obtained from the study are viewed generally, it can be said that troubles about scale development is still continuing. This situation shows that it hasn't been understood how scale development is important and also that it is a hard process. Errors in the scales developed without correct scale development processes will cause further studies to have wrong results. Therefore the scale development is needed to be emphasized in a vital way. To manage the scale development process, it can be suggested that a group of measurement and evaluation specialist should provide consultancy service to the researchers willing to conduct scale development studies, a document explaining the standards related to the scale development process in detail should be developed by measurement and evaluation specialist and it should be shared with journal reviewers.