



Fen ve Matematik Eğitimi Alanında Gerçekleştirilen Ölçek Geliştirme Araştırmalarına Yönelik Tematik İçerik Analizi

Şeyda Gül ¹, Mustafa Sözbilir ²

Öz

Bu araştırmanın amacı, Türkiye’de fen ve matematik eğitimi alanında çalışan araştırmacılar tarafından yayımlanan ölçek geliştirme çalışmalarının içerik ve yöntemsel araçlar açısından durum değerlendirmeye yönelik bir analizini ortaya koymaktır. Bu amaçla, Türkiye’de eğitim bilimleri alanında yayım yapan altı temel akademik dergide 2000-2013 yılları arasında yayınlanan makaleler, araştırmaya dâhil olma ve araştırmadan hariç tutulma ölçütleri doğrultusunda taranmış ve toplam 22 makale araştırma kapsamına alınmıştır. Makalelerde ulaşılan bulgular, tematik içerik analizi (meta-sentez) şeklinde analiz ve sentez süreçlerinden geçirilerek çalışma gerçekleştirilmiştir. Bulgular, çalışmaların büyük çoğunluğunun tutumlar üzerinde yoğunlaştığını, daha çok matematik eğitimi alanında ölçeklerin geliştirildiğini, geçerlik analizi olarak en fazla yapı geçerliğinin incelendiğini ve bunlardan doğrulayıcı faktör analizinin sınırlı sayıda olduğunu, açıklayıcı faktör analizi yöntemlerine ait analizlerin genelde orta düzeyde yapıldığını ortaya koymuştur. Bununla birlikte güvenilirlik analizlerinde ise en fazla iç tutarlık yöntemlerinin tercih edildiği ortaya çıkmış ve ulaşılan sonuçlar doğrultusunda çeşitli önerilere yer verilmiştir.

Anahtar Kelimeler

Ölçek Geliştirme
Tematik içerik analizi/Meta-sentez
Fen ve Matematik Eğitimi

Makale Hakkında

Gönderim Tarihi: 19.10.2014
Kabul Tarihi: 06.03.2015
Elektronik Yayın Tarihi: 04.05.2015

DOI: 10.15390/EB.2015.4070

Giriş

Değişen dünyamızda toplumun ihtiyaçları hızla değişmekte ve bu ihtiyaçlara cevap verebilecek insan gücünün eğitilmesi için de özellikle eğitim konusunda yapılan bilimsel araştırmalar büyük önem taşımaktadır (Çiltaş, Güler ve Sözbilir, 2012). Bilimsel araştırmalar, sonuçları itibariyle tüm bilim alanlarındaki politika ve uygulamaları etkileyerek, uygulamalar için ampirik bir temel oluşturmakta, aynı zamanda mesleki faaliyetler için de uygulayıcılara rehber olarak kullanılabilmektedir. Eğitim alanındaki bilimsel araştırma sonuçları ise, ders kitapları, ansiklopediler ve el kitapları şeklinde yayınlanan çalışmalar açısından önemli bir içerik kaynağını oluşturur ve herhangi bir öğretim alanında üretilmiş bilgi olarak da kendisini gösterebilir (Karadağ, 2009).

Bilimsel araştırmaların genellikle toplumsal amaçları vardır. Araştırma, toplumu tanımaya, profilini betimlemeye, değişkenlerle ilgili ilişkileri ortaya çıkarmaya çalışan, bir sorunun ortaya çıkmasına yol açan etmenlerle ilgili neden, niçin, nasıl, ne zaman, nerede gibi sorulara cevap veren uğraşlar ve bilimsel bilgi üretme işlevidir (Özdamar, 2003, s:2). Bununla beraber bilimsel araştırmalar ancak veriler aracılığıyla yapılabilir (Ercan ve Kan, 2004). Bir bilimsel araştırmada verilerin elde

¹ Atatürk Üniversitesi, Kazım Karabekir Eğitim Fakültesi, Biyoloji Eğitimi Ana Bilim Dalı, Türkiye, seydagul@atauni.edu.tr

² Atatürk Üniversitesi, Kazım Karabekir Eğitim Fakültesi, Kimya Eğitimi Ana Bilim Dalı, Türkiye, sozibilir@atauni.edu.tr

edilebilmesi, elde edilen verilerin sayısallaştırılması, kaydedilmesi ve uygulama alanına konulmasında ise ölçekler önemli bir yere sahiptir (Karagöz ve Ekici, 2004).

Ölçek kavramı özünde ölçme sonuçlarının matematiksel niteliklerini göstermekle birlikte, eğitim ve psikoloji gibi davranış bilimlerinin birçok alanında, hedeflenen kişi veya kişiler, sistem, konu ya da içerik açısından bilgi toplamak amacıyla kullanılmaktadır (Yurdugül, 2005). Elbette bu bilgi toplama sürecinde kullanılan ölçeklerin/ölçme araçlarının seçiminde de dikkat edilmesi gereken bazı unsurlar vardır. Bu unsurlar arasında ölçme aracının araştırma konusuyla uyumlu olması, ölçme aracının güvenilirliği ve geçerliği önemli konuların başında gelmektedir. Zira bir bilimsel araştırmada, ölçek geliştirenlerin güvenilirlik ve geçerlik çalışmaları yapması ve ölçeği uygulayanların da uygulayacakları ölçeğin güvenilirlik ve geçerlik çalışmasının yapıp yapılmadığını sorgulaması, araştırma sonuçlarının doğruluğu açısından uygun veri sağlayacaktır (Ercan ve Kan, 2004). Bununla beraber, alan yazında önerilen geçerlik ve güvenilirlik önlemleri ile ilgili olarak araştırmacılar tarafından yapılan farklı sınıflamalar olmakla birlikte (Fraienkel ve Wallen, 2000; McMillan ve Schumacher, 2009; Topu, Baydaş, Turan ve Göktaş, 2013) ölçek geliştirme çalışmalarında önerilen geçerlik ve güvenilirlik önlemleri genel olarak aşağıdaki şekilde sıralanabilir:

Geçerlik Önlemleri;

- Uzman görüşünün alınması
- Veri toplama aracı hakkında bilgi verilmesi
- Veri toplama süreci hakkında bilgi verilmesi
- Veri analiz sürecinin ve değerlendirmenin kim/kimler tarafından yapıldığının açıklanması
- Varsayımlar ve sınırlılıkların betimlenmesi
- Örneklem özelliklerinin açıklanması
- Katılımcı gönüllülüğünün alınması
- Çalışmanın uygulama sürecinin betimlenmesi
- Madde sayısı

Güvenirlik Önlemleri;

- Uzman görüşünün alınması
- Güvenirlik hesaplamalarının belirtilmesi
- Değerlendiriciler arası güvenilirlik çalışmasının yapılması
- Katılımcı görüşünün (kontrolünün) alınması
- Veriler arasında tutarlılığın kontrol edilmesi
- Madde sayısı

Yukarıda ifade edilen geçerlik ve güvenilirlik önlemlerinin yanı sıra ölçek geliştirme sürecinde çok çeşitli yöntem ve tekniklere başvurulabilir. Bununla beraber; ölçme araçlarının geliştirilmesi temel olarak ölçme aracının tasarımı ve pilot uygulama olarak ifade edilen iki aşamada ele alınabilir. Her iki aşamaya özgü birbirinden farklı geçerlik kavramları söz konusudur. Özellikle ölçme aracının tasarımı aşamasında kapsam ve görünüş geçerliği ön plana çıkarken, uygulama aşamasında ise faktöriyel, ölçüt (yordama, uygunluk) ve yapı (yakınsama ve ıraksama) geçerlikleri ele alınmaktadır (Yurdugül ve Bayrak, 2012).

Genel anlamda değerlendirildiğinde, ölçek geliştirme çabaları uzun soluklu çalışmalardır ve ölçek geliştirecek araştırmacıların ölçülecek değişkene ilişkin alan bilgisi, ölçek geliştirme ve istatistik bilgi ve becerisinin yeterli düzeyde olması gerekmektedir. Ayrıca, geliştirilen ölçeklerde incelenen özellikle ilgili gelişmelere paralel olarak zaman içinde gerekli değişikliklerin yapılması; değişik örneklem ve zamanlarda ölçme aracının psikometrik niteliklerinin yeniden irdelenmesi de önemli bir husustur (Azaltun, 2008). Alan yazında ölçek geliştirmede kullanılan yöntem ve uygulamalar konusunda araştırmacıları bilgilendirmek ve ölçek geliştirme sürecini pekiştirmeye yönelik çalışmalara yer yer rastlanmaktadır (Azaltun, 2008; DiStefano, Zhu ve Mindrila, 2009; Ercan ve Kan, 2004; Erkuş, 2012; Yurdugül, 2005). Buna rağmen, Türkiye’de özellikle son yıllarda ölçek geliştirmeye yönelik olarak gerçekleştirilen çalışmalarda dikkate değer bir artış olduğu, bununla beraber akademisyenlerin üzerindeki yayın baskısının da itici gücüyle, ölçme ve değerlendirme alanından

veya yakın uzak alanlardan pek çok akademisyenin, ne yazık ki bu engeli aşmada “en kolay” yol olarak ölçek geliştirme veya ölçek uyarlama çalışmalarını gördüğü ve bunun da pek çok hatayı ve eksikliği beraberinde getirdiği ileri sürülmektedir (Erkuş, 2007). Bunun yanı sıra belirli konularda yapılmış ve birbirinden bağımsız yürütülen bu çalışmalarda benzer veya farklı süreçlerin uygulandığı veya birbirinden farklı sonuçlara ulaşıldığı görülmektedir. Özellikle fen, fen-teknoloji ve matematik eğitimi alanında yapılan çalışmalara bakıldığında genel olarak tutum, öz-yeterlik, kaygı ve diğer (algı, görüş, inanç vb) konulara yönelik çok sayıda ölçek geliştirme çalışmasına rastlanmaktadır. Ayrıca, bu ölçek geliştirme çalışmaları genel olarak değerlendirildiğinde, söz konusu çalışmalarda araştırmacıların bireysel çabalarıyla çeşitli boyutların ortaya çıktığı yönünde bulguların elde edildiği görülmektedir. Dolayısıyla, bu bilgi yığını yorumlamak ve yeni çalışmalara yol açmak için, kapsayıcı ve güvenilir nitelikte üst çalışmaların yapılması büyük önem arz etmektedir (Akgöz, Ercan ve Kan, 2004). Ayrıca bu alanda yapılan makale, yüksek lisans veya doktora tez çalışmalarının incelendiği araştırmalarda en çok kullanılan veri toplama aracının ölçek olması bu ölçme aracını geliştirirken dikkat edilmesi gereken aşamaları daha önemli kılmaktadır (Acar-Güvendir ve Özer-Özkan (2015). Bu nedenle ölçek kavramının özellikle araştırmacılar tarafından doğru olarak algılanması, ancak bu konuda yapılan araştırmaların nasıl yapıldığının ve ne tür sonuçlar ortaya koyduğunun bilinmesiyle gerçekleşebilir. Bunu gerçekleştirmek için atılacak en önemli adımlardan biri, ölçek geliştirme sürecinde izlenen aşamaların belirlenmesi, var olan durumun ve eksikliklerin ortaya konulması ve böylece ölçek geliştirme konusunda yapılan araştırmaların ayrıntılı olarak irdelenmesi olabilir. Bu konu ile ilgili olarak alan yazında yurt dışında yapılan çalışmalar incelendiğinde, sınırlı sayıda yapılan bu çalışmaların çoğunlukla tek bir dergiye odaklandığı, sınırlı konuların ve istatistik tekniklerin incelendiği ve bulguların çoğunlukla nicel tekniklerle yorumlandığı görülmektedir (Dawis, 1987; Fassinger, 1987; Kahn, 2006; Tinsley ve Tinsley, 1987; Worthington ve Whittaker, 2006). Örneğin, Tinsley ve Tinsley (1987) ve Kahn (2006) bir dergideki yayınlarda sadece faktör analizine odaklanırken; Martens (2005) ve Quintana & Maxwell (1999) aynı dergide yapısal eşitlik modeline odaklanarak genel bir değerlendirme sunmaktadır. Bununla beraber Türkiye’de Çüm (2013) tarafından psikoloji ve eğitim bilimleri alanında yayımlanan makalelerin incelenmesine yönelik yapılan çalışma, Mor-Dirlik (2014) tarafından eğitim bilimleri alanında yapılmış 5 adet tezin incelendiği çalışma ile Acar-Güvendir ve Özer-Özkan (2015) tarafından seçilmiş üç dergideki eğitim alanında yapılan ölçek geliştirme/uyarlama çalışmalarının tarama yöntemi ile benzerlik ve farklılıklarının genel olarak incelendiği çalışmaya rastlanılmaktadır. Bunun yanında Tosun ve Taşkesenligil (2014) tarafından fen eğitimi alanında geliştirilen/adapte edilen ölçeklerden/başarı testlerinden oluşan toplam 62 makalenin bir doküman incelemesi yöntemiyle betimsel bir sunumunun yapıldığı çalışmaya da mevcuttur. Ancak Türkiye’de çalışma kapsamında belirlenen ana temalar çerçevesinde ölçek geliştirmeye yönelik yapılmış çalışmalardan yola çıkarak, ölçek geliştirme sürecinde izlenen tüm geçerlik ve güvenilirlik süreçlerini daha detaylı bir şekilde irdelleyen ve durumu nitel olarak sentezleyen bir araştırma bulunmamaktadır. Dolayısıyla bu konuda yapılacak bir tematik içerik analizi (meta-sentez) araştırması ölçek geliştirme ve uygulamaya yönelik araştırma yapmayı planlayan araştırmacıların var olan çalışmalardan haberdar olmaları ve mevcut durumu değerlendirebilmeleri sağlanarak ileride yapılacak çalışmalara da ışık tutacağı düşünülmektedir. Belirtilen gerekçe çerçevesinde bu çalışmada, ülkemizde eğitim bilimleri alanında yayın yapan altı temel dergide ölçek geliştirme alanında yayımlanan araştırma makaleleri tespit edilerek belli ölçütler açısından incelemek suretiyle mevcut durumu sentezlemek amaçlanmıştır. Bu amaç doğrultusunda aşağıda ifade edilen sorulara yanıt aranmaya çalışılmıştır. Buna göre;

Fen (biyoloji, fizik, kimya), fen-teknoloji ve matematik eğitimi alanında belirlenen ana temalar doğrultusunda;

- 1) Ölçek geliştirme çalışmaları daha çok hangi alanlarda yapılmıştır?
- 2) Ölçek geliştirme çalışmalarının aldığı atıf sayıları nedir?
- 3) Ölçek geliştirme çalışmalarında sıklıkla hangi tür örneklemeler ile çalışılmıştır?
- 4) Ölçek geliştirme çalışmalarında kullanılan örneklem büyüklükleri nasıldır?
- 5) Ölçek geliştirme çalışmalarında kullanılan geçerlik analiz türleri nelerdir?
- 6) Ölçek geliştirme çalışmalarında ele alınan yapı geçerliğine ait analizlerde hangi aşamalar sıklıkla tercih edilmektedir?
- 7) Ölçek geliştirme çalışmalarında kullanılan güvenilirlik analiz türleri nelerdir?

Yöntem

Bu araştırma, bir tematik içerik analizi (meta-sentez) çalışmasıdır. Tematik içerik analizi; aynı konu üzerinde yapılan birden fazla araştırmanın bulgularını, belirlenen tema veya şablonlar çerçevesinde ve ham verilerden farklı olarak eleştirel bir bakış açısıyla inceleyen, sentezleyen ve yorumlayan çalışmalardır (Au, 2007; Finfgeld, 2003; Walsh ve Downe, 2005). Bu özelliği ile tematik içerik analizi aynı zamanda, belli bir konunun farklı boyutlarını ele alan çalışmalardaki ortak ve benzer yönlerinin nitel olarak sentezlenmesi ve örneklendirilmesini sağladığından, daha fazla çalışmaya ulaşma açısından araştırmacılara, öğretmenlere ve karar alıcılara zengin bir başvuru kaynağı oluşturur (Çalık, Ayas ve Ebenezer, 2005; Çalık ve Sözbilir, 2014). Özetle ifade etmek gerekirse tematik içerik analizi çalışmaları belli bir alanda yapılmış araştırmaları nitel bir anlayışla ele alınıp, benzerlik ve farklılıkların karşılaştırmalı olarak ortaya konmasını içermektedir. Meta-analiz ve betimsel içerik analizleriyle kıyaslandığında, bu çalışmalarda genellikle incelemeye alınan araştırma sayısı (örneklem büyüklüğü) sınırlı olmaktadır (Çalık ve Sözbilir, 2014).

Araştırmanın Kapsamı ve Süreci

Verilerin Toplanması

Verilerin toplanması aşamasında öncelikle araştırma kapsamında Türkiye'deki altı akademik dergideki 2000-2013 yılları arasında yayınlanan makalelerin incelenmesine karar verilmiştir. Makalelerin yayımlandığı dergilerin seçiminde ülkemiz kaynaklı olup Thomson Reuters Web of Science© dizinindeki dergiler (Eğitim Araştırmaları Dergisi, Eğitim ve Bilim; Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Kuram ve Uygulamada Eğitim Bilimleri) ile fen eğitimi odaklı dergiler (International Journal of Environmental and Science Education, Türk Fen Eğitimi Dergisi) olmasına dikkat edilmiştir. Bu dergiler aynı zamanda ERIC dizininde de dizinlenen dergilerdir. Bunun temel nedeni, Acar-Güvendir ve Özer-Özkan (2015)'in de ifade ettiği gibi, söz konusu veri tabanlarında makaleler uluslararası standartlar doğrultusunda belirlenen ölçütlere göre bir değerlendirme sürecinden geçtiği varsayımı ve söz konusu dergilerdeki yayınların dikkate alınmasının daha güvenilir sonuçlar doğuracağı düşünülmektedir.

Makalelerin temini aşamasında, ücretsiz ve açık erişimli olan makalelere dergilerin kendi web sayfası üzerinden ulaşılırken, ücretli olan dergilere Atatürk Üniversitesi Kütüphanesi üzerinden erişilmiş ve tam metin olarak yayınlanan 200'ün üzerinde ölçek geliştirme/uyarlama konulu makaleye ulaşılmıştır. Bununla beraber, araştırmacılar tarafından ortaklaşa yapılan ön değerlendirmede makaleler "araştırmaya dâhil olma ve araştırmadan hariç tutulma" ölçütleri çerçevesinde yeniden incelenmiş ve böylece bu araştırmaya uygun olan toplam 22 makale değerlendirmeye alınmıştır. Araştırma kapsamında değerlendirmeye alınan makalelerin kodu ve künyesine ait detaylı bilgi Ek 1'de sunulmuştur.

Makalelerin Araştırmaya Dâhil Olma ve Araştırmadan Hariç Tutulmasında Dikkate Alınan Ölçütler

- **Seçilen dergilerde yayınlanması:** Araştırma kapsamında değerlendirilen makalelerin sadece seçilmiş olan dergilerde yayınlanmasına dikkat edilmiştir.
- **Belirlenen yıllarda yayınlanmış olması:** İncelenen dergilerdeki makalelerin 2000-2013 yılları arasında yayınlanmış olmasına dikkat edilmiştir.
- **Makalelerin ölçek geliştirmeye yönelik olması:** Bu tematik içerik analizi araştırmasında sadece ölçek geliştirme çalışmaları incelenmiştir. Bu nedenle test geliştirme veya ölçek uyarlama çalışmaları araştırmaya dâhil edilmemiştir. Ayrıca ölçek yerine anket geliştirmeye yönelik çalışmalar analiz dışında bırakılmıştır.
- **Araştırma alanına uygun olması:** Bu araştırma kapsamında incelenen ölçek geliştirme çalışmalarından sadece biyoloji, fizik, kimya, fen-teknoloji ve matematik eğitimine yönelik konularda yayınlananlar analizlere dâhil edilmiştir.
- **Örneklemin uygun olması:** Tematik içerik analizine dâhil edilen makalelerdeki örneklemin sadece Türkiye sınırlarından seçilmiş olmasına dikkat edilmiştir.

Tematik İçerik Analizine Dâhil Edilen Çalışmalarda Ana Temaları Oluşturan Konular ve Kodlama Yöntemi

Araştırma kapsamında tematik içerik analizi yapılacak makalelerin kodlanmasında ve temaların belirlenmesinde her iki araştırmacı birlikte çalışmıştır. Buna göre araştırmaya dâhil edilen makaleler araştırmacılar tarafından tek tek incelenerek ana temalar ve alt temalar belirlenmiştir. Belirlenen ana ve alt temalar detaylı bir biçimde tüm yönleriyle ele alınıp kodlanarak tablo haline getirilmiştir (Tablo 1, Tablo 2, Tablo 3).

Tablo 1. Makalelere Ait Kodlar ve Ana Temalar

Ana Tema Kodları	Temalar
TÖGÇ	Tutum, ilgi vb. yönelik ölçek geliştirme çalışmaları
ÖÖGÇ	Öz-yeterliğe yönelik ölçek geliştirme çalışmaları
KÖGÇ	Kaygı, endişe vb. yönelik ölçek geliştirme çalışmaları
DÖGÇ	Diğer konularda yapılan ölçek geliştirme çalışmaları

Araştırmada ana ve alt temalar belirlendikten sonra, araştırmanın güvenilirliğini sağlamak için her iki araştırmacı, belirlenen ana temaların her birinden rastgele seçmiş oldukları birer makaleyi bireysel olarak incelemiştir. Yapılan bireysel incelemeler sonucunda veriler karşılaştırılmış ve büyük çoğunluğunda araştırmacıların fikir birliği içinde olduğu görülmüştür. Ayrıca yapılan karşılaştırmalarda herhangi bir tutarsızlık görüldüğünde söz konusu makaleler araştırmacılar tarafından birlikte incelenmiş ve gerekli düzeltmeler yapılarak tutarsızlıklar giderilmeye çalışılmıştır.

Araştırmada güvenirliliğin yanı sıra geçerliliğin sağlanmasına da dikkat edilmiştir. Bu amaçla Sandelowski ve Barroso'nun (2007), meta-sentezde geçerliliği sağlamak için tanımladıkları üç tür geçerlikten hareketle, bu araştırma aşağıdaki geçerlik ölçütleri çerçevesinde yürütülmüştür (akt: Aküzüm ve Özmen, 2013):

1. Tanımlayıcı geçerlik: Verilerin doğruluğunu gerçeklere dayanarak tanımlamayan bir geçerlik türüdür. Bu, çalışmada kullanılan her bir rapordan elde edilen anlamlı ve doğru tanımlamalardır.
2. Yorumlayıcı geçerlik: Bakış açılarıyla ilgili araştırmacıların anlayışlarının tam ve doğru temsil edilmesini sağlamaktadır.
3. Kuramsal geçerlik: Bulguların yorumlanmasında araştırmacının güvenirliliğine başvurmaktadır. Bu, bilgileri birleştirmede, verileri yorumlamak için kullanılan yöntemle bağlı olmak anlamına gelmektedir.

Bulgular

Araştırmada tematik içerik analizine dâhil edilen makaleler belirlenen ana temalar altında konu alanları ve yayımlandıkları yıllara göre sıralanarak biyoloji eğitimi makaleleri “B1, B2...” şeklinde, fizik eğitimi makaleleri “F1, F2...” şeklinde, kimya eğitimi makaleleri “K1, K2...” şeklinde, fen ve teknoloji eğitimi makaleleri “FT1, FT2...” şeklinde ve matematik eğitimi makaleleri “M1, M2...” olacak şekilde kodlanmıştır. Dolayısıyla yapılan analizlerde bu kod sıraları dikkate alınmıştır. Daha sonra araştırma kapsamında belirlenen her bir ana tema için konu alanı, yayın yılı, atıf sayısı, örneklem türü, örneklem büyüklüğü, geçerlik analizleri ve güvenilirlik analizlerine ait bulgular tablolar halinde sırasıyla sunulmuştur (Tablo 2). Bunun yanı sıra, ölçek geliştirme çalışmalarında sıklıkla başvurulan yapı geçerliği aşamalarının belirlenmesinde Worthington ve Whittaker (2006) tarafından belirlenen tasnif sisteminden yararlanılmıştır (Tablo 3). Buna göre yapılan analizlerden elde edilen bulgular aşağıda sunulmuştur.

Tablo 2 incelendiğinde TÖĞÇ ana temasına ait 13 makalenin olduğu ve söz konusu makaleler içinde B1 ve B2 kodlu makalelerin en fazla atıf aldığı anlaşılmaktadır. Diğer makalelerin ise oldukça az atıf aldığı veya hiç almadığı (FT4) belirlenmiştir. ÖÖĞÇ ana temasına bakıldığında, bu ana temaya ait 3 makalenin olduğu ve bu makaleler içinde M3 kodlu makalenin oldukça fazla atıf aldığı, M5 kodlu makalenin ise hiç atıf almadığı anlaşılmaktadır. Benzer şekilde KÖĞÇ ana temasına ait bulgular, bu konuda sadece 2 makalenin olduğunu ve bu makalelerden birinin 3 atıf aldığı diğerinin ise hiç atıf almadığı göstermektedir. Son olarak DÖĞÇ başlıklı ana temaya bakıldığında, bu ana temaya ait 4 makalenin olduğu, söz konusu makalelerden birinin (M7) fazla atıf aldığı diğerlerinin ise oldukça az atıf aldığı veya hiç almadığı (M9) belirlenmiştir.

Tematik içerik analizine dâhil edilen makaleler örneklem türü açısından incelendiğinde; çalışmaların büyük çoğunluğunun ortaöğretim ve lisans öğrencileri ile yürütüldüğü, yine az bir kısmında eğitimciler ve ilköğretim öğrencilerine yer verildiği anlaşılmaktadır. Bunun yanında örneklem büyüklüğü açısından elde edilen bulgular, makalelerin yarısına yakın bir kısmında 301-500 arasındaki örneklerle çalışıldığını ve geriye kalanların da yarısının 300’ün altındaki kişilerden oluştuğunu ortaya koymaktadır.

Geçerlik analizleri ile ilgili olarak; makalelerin tamamında yapı geçerliğinin incelendiği (Tablo 2) ve bunların hepsinde faktör analizinin kullanıldığı (Tablo 3), ayrıca 8 tanesinde ayırt edicilik geçerliğine (discriminant validity) yer verildiği, yarısından fazlasında ise kapsam geçerliğinin yapıldığı görülmektedir. Ancak ölçüt geçerliğine sadece 1 makalede bakılmakla beraber makalelerin yarısından fazlasında görünüş geçerliğinin yapıp yapılmadığı net olarak anlaşılamamaktadır. Güvenirlik analizlerine ait bulgular incelendiğinde; makalelerin tamamında iç tutarlılığa bakıldığı, yaklaşık üçte birinde ise testi yarılama yönteminin tercih edildiği görülmektedir. Bunun yanında test tekrar test yönteminin sadece 2 makalede kullanıldığı, eşdeğer formların ise hiç kullanılmadığı dikkati çekmektedir.

Tablo 3’te yapı geçerliğine ilişkin bulgular detaylı incelendiğinde, makalelerin tamamında faktör analiz türlerinden açımlayıcı faktör analizinin (AFA) kullanıldığı, çok az bir kısmında doğrulayıcı faktör analizine (DFA) yer verildiği görülmektedir. Açımlayıcı faktör analizi ile ilgili olarak elde edilen bulgulara bakıldığında ise verilerin faktör analizi yapmaya uygunluğunun belirlenmesinde Küresel Bartlett testi ve Kaiser-Meyer-Olkin istatistiğinin en fazla tercih edildiği, rotasyon metodu olarak ise dik döndürmenin (Orthogonal) sıklıkla yapıldığı anlaşılmaktadır. Öte yandan madde eleme/tutma ölçütlerine ait bulgular incelendiğinde; sadece bir makalede faktör yük değerlerine bakılmadığı, yarısından fazlasında çapraz/binişik yüklerin de dikkate alındığı, ayrıca makalelerin çoğunluğunda madde analizinin sıklıkla tercih edildiği görülmektedir. Bunun yanında ortak varyanslara ise sadece 4 makalede yer verildiği anlaşılmaktadır. Bununla beraber faktör eleme/tutma ölçütlerine bakıldığında; makalelerin yarısından fazlasında özdeğerler, yamaç eğim grafiği ve faktörlerin açıkladığı minimum varyans değerleri dikkate alınırken çok sınırlı sayıdaki makalede faktör başına düşen madde sayısına yer verildiği anlaşılmaktadır.

Table 3. İçerik analizine dâhil edilen çalışmalarındaki yapı geçerliği aşamaları

Ana temalar	Makalenin kodu	AFA	DFA	Korelasyon matrisinin faktör analizi yapmaya uygunluğu				Rotasyon metotları				Madde eleme/tutma ölçütleri				Faktör eleme/tutma ölçütleri					
				Madde başına düşen katılımcı sayısı	Küresel Bartlett testi	KMO	Diğer	Yapılmanış/Belirlenmiş	Dik döndürme	Eğik döndürme	Yapılmanış/Belirlenmiş	Yükler	Binişik yükler	Ortak varyanslar	Madde analizi	Özdeğerler	Yamaç eğim grafiği	Faktörlerin açıkladığı minimum varyans	Faktör başına düşen madde sayısı		
TÖĞÇ	B1	✓	-	-	✓	✓	-	-	-	✓	-	-	-	✓	✓	✓	-	-			
	B2	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	-	-			
	B3	✓	✓	-	✓	✓	-	-	-	✓	✓	-	-	✓	✓	✓	✓	✓			
	B4	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	-	-			
	F1	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	-	-	-			
	F2	✓	-	✓	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓			
	K1	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	-	-	-			
	FT1	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓			
	FT2	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓			
	FT3	✓	✓	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓			
FT4	✓	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓				
M1	✓	-	-	-	-	-	-	-	✓	-	-	-	✓	✓	✓	-	-	-			
M2	✓	-	-	-	✓	✓	-	-	-	-	-	-	✓	✓	✓	✓	✓	✓			
Toplam	13	✓	2	2	12	12	2	1	11	1	1	1	13	6	1	9	11	12	10	2	
M3	✓	-	-	-	-	-	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	
M4	✓	-	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	
M5	✓	-	-	-	-	-	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	
Toplam	3	✓	1	0	1	1	1	2	1	1	1	2	2	2	1	2	1	1	1	1	0
FT5	✓	-	-	✓	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
M6	✓	-	-	✓	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
Toplam	2	0	0	2	2	2	2	0	2	0	0	0	2	0	2	2	2	2	1	1	0
K2	✓	-	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
M7	✓	-	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
M8	✓	-	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
M9	✓	-	-	-	✓	✓	-	-	✓	-	-	-	✓	✓	✓	✓	✓	✓	✓	✓	✓
Toplam	4	2	2	4	4	4	0	0	2	1	1	4	4	4	3	4	3	4	2	2	1
Genel Toplam	22	5	4	18	19	18	5	3	16	3	3	21	12	17	18	14	17	17	14	14	3

Tartışma, Sonuç ve Öneriler

Bu araştırma Türkiye’de ölçek geliştirme/uyarlama yönelik çok sayıda çalışmanın yapıldığını, ancak sadece ölçek geliştirmeye yönelik olarak biyoloji, fizik, kimya, fen-teknoloji ve matematik eğitimi alanlarında yapılan ölçek geliştirme çalışmalarının büyük çoğunluğunun matematik eğitimi alanında yapıldığı görülmektedir. Matematik eğitimi alanında yapılan ölçek geliştirme çalışmalarının fen alanlarına göre daha fazla olmasının sebeplerine dair kabul edilebilir bir açıklama geliştirmenin zor olmasına rağmen, belki ölçek geliştirmeye yönelik analizlerin ileri düzeyde istatistik bilgisi gerektirmesi ve matematik eğitimi araştırmacılarının alanları gereği istatistik işlemler için gerekli temel bilgi ve becerilere sahip olmaları etkili olmuş olabilir.

İncelenen makaleler ana temalara göre değerlendirildiğinde, makalelerin yarısından fazlasının tutum ölçeği geliştirmeye yönelik olduğu, öz-yeterlik, kaygı düzeyi ve diğer konulara yönelik makalelerin ise sınırlı sayıda kaldığı görülmektedir. Benzer durum atıf sayıları için de geçerlidir. Tosun ve Taşkesenligil (2014) tarafından Türkiye’de fen eğitimi alanında geliştirilen/adapte edilen ölçeklerin doküman analizinin yapıldığı çalışmada da benzer bulgulara ulaşılmıştır. Birey davranışlarının önemli ve kritik bir yordayıcısı olarak görülen tutumlar bilişsel, duyuşsal ve davranışsal boyutları içermesi (Anderson, 1988) nedeniyle kaygı, öz-yeterlik vb. diğer psikolojik yapılardan daha kapsamlı olarak düşünülebilir. Fen ve matematik eğitimi alanında yapılan içerik analiz çalışmaları (Çiltaş vd., 2012; Sozibilir, Kutu ve Yasar, 2012) öğrenme, öğretme temel konu alanları dışında en yaygın olarak araştırılan değişkenler arasında tutumun yer aldığını göstermektedir. Yaygın olarak incelenen bir değişken olması da değişik dersler kapsamında birçok tutum ölçeğinin geliştirilmesini gerekli kılmaktadır. Dolayısıyla bu durum ölçek geliştirme çalışmalarında tutumlara daha fazla yer verilmesine sebep olarak görülebilir.

Çalışmada atıf sayıları açısından TÖGÇ’na yönelik bulgular; B1 ve B2 kodlu makalelerin en fazla atıf aldığını ortaya koymuştur. B1 kodlu makalenin laboratuvara yönelik öğretmen tutumlarını belirlemeye yönelik olduğu düşünüldüğünde, özellikle fen öğretim sürecinin temel bileşenlerinden olan öğretmenler ve laboratuvar dersine yönelik tutumların belirlenmesi konusunda çok fazla çalışmada kullanılması doğaldır. Ayrıca söz konusu makalenin yayım yılının oldukça eski bir tarihte olması da çok sayıda atıf almasının nedenleri arasında sayılabilir. Benzer durum ÖÖGÇ ana temasına ait M3 ile DÖGÇ ana temasına ait M7 kodlu makaleler için de geçerlidir. Yanı sıra B2 kodlu makalenin çevreye yönelik olması çok fazla atıf almasına neden olabilir. Zira Gul ve Sozibilir (2015) tarafından yapılan çalışmalarda da gerek Türkiye’de gerekse dünyadaki seçilmiş önemli dergilerde yayınlanan biyoloji eğitimi araştırma makalelerinde en fazla çevre eğitimi konusuna yer verildiği, bunun nedeni olarak ise bu konunun çoğu zaman disiplinler arası olarak ele alınması nedeniyle biyoloji eğitimi alanı dışında farklı araştırmacılar tarafından da çalışılabilirliği gösterilmiştir. KÖGÇ’na yönelik bulgulara bakıldığında FT5 kodlu makalenin çok yeni bir tarihte yayınlanmasına rağmen bu ana tema altında toplanan diğer makaleden daha fazla atıf alması dikkat çekici bir durumdur. Bu durum söz konusu makalenin öğretmen, öğrenci ve velilere yönelik ölçekleri içermesi nedeniyle daha geniş bir örnekleme hitap etmesi ve bu nedenle daha fazla çalışmada kullanılabilir olmasına atfedilebilir.

Çalışmanın bulguları örneklem grubu açısından genel olarak değerlendirildiğinde; tematik içerik analizine dâhil edilen makalelerin büyük çoğunluğunun ortaöğretim ve lisans öğrencileri ile yürütüldüğünü ortaya koymaktadır. Tosun ve Taşkesenligil (2014) tarafından da elde edilen benzer bulgular araştırmacıların örnekleme ulaşma kolaylığından kaynaklandığını düşündürebilir. Zira söz konusu makalelerin özellikle lisans öğrencilerinin çoğunlukla öğretmen adayları ile yürütüldüğü düşünüldüğünde bu bulgu araştırmacıların örnekleme ulaşma kolaylığından bu grubu tercih ettikleri şeklinde yorumlanabilir. Alan yazında araştırma makalelerinin içerik analizine yönelik yapılan birçok çalışmada da aynı bulgular ortaya konularak benzer yorumlar ileri sürülmüştür (Çiltaş vd., 2012, Gul ve Sozibilir, 2015). Öte yandan örneklem çeşitliliğinin sınırlı sayıda olması, öğretim sürecinin temel bileşenlerinin (öğrenci, öğretmen, veli, yönetici vb.) genel anlamda göz ardı edildiğini düşündürebilir. Ayrıca, tutum, özyeterlik, kaygı vb. özelliklerin sadece bireyin kendisinden değil, çevresindeki diğer öge ve kişilerden de etkilenebileceği düşünülürse (Deveci, Çalmaz ve Açık, 2012; Gençtürk ve Memiş,

2010), ileride yapılacak çalışmalarda kullanılan örneklemin geniş bir tabana yayılması mevcut sorunların tespiti ve çözümüne yönelik daha güvenilir ve doğru sonuçlar elde edilmesine katkı sağlayabilir.

Öte yandan elde edilen bulgular, incelenen makalelerdeki örneklem büyüklüğünün genel anlamda yarıya yakınının 301-500 arası, geriye kalanların yarısının ise 300 ve altındaki kişilerle gerçekleştirildiği yönündedir. Örneklem sayısının genellikle 301-500 arasında tutulması geçerlik ve güvenilirlik açısından olumlu olmakla birlikte daha güvenilir sonuçlar elde etmek adına örneklem sayısının daha da artırılması gerekli görülmektedir. Zira Comrey ve Lee (1992), örneklem büyüklüğü için gerekli sayının 100= zayıf, 200= orta, 300=iyi, 500= çok iyi ve 1000= mükemmel olarak nitelendirirken; Aleamoni (1976) alt sınırın 400 olması gerektiğini ifade etmiştir. Ayrıca bu çalışmada AFA analizlerinde de faktör başına düşen kişi sayısı ölçütüne yönelik katılımın düşük düzeyde olması ve Tosun ve Taşkesenligil (2014) tarafından ölçek geliştirme çalışmalarında kullanılan örneklem büyüklüğünün genel olarak 101-200 kişi arasında değiştiğine yönelik elde edilen bulgular, Türkiye’de ölçek geliştirme çalışmalarında örneklem büyüklüğü açısından güvenilirlik konusunda bir zayıflığın olduğunu düşündürebilir. Nitekim Delice (2010), örneklem büyüklüğünün çok küçük seçilmesi halinde testin gücünün azalabileceği düşüncesinden hareketle bir güvenilirlik analizinde örneklem büyüklüğünün oldukça önemli bir faktör olduğunu vurgulayarak benzer görüşlerini ortaya koymuştur. Diğer taraftan örnekleme artırmayı güvenilirlik sorununu çözen bir durum gibi düşünmek te her zaman için doğru bir yol olarak görülmemektedir. Zira örneklem büyüklüğü kadar örneklemin ne derece doğru bir örneklem olduğu ve verilen formları ne derece ciddiyetle doldurduğu da göz ardı edilmemesi gereken önemli bir husustur. Yalnız güvenilirlikle ilgili bu husu hakkında ölçek geliştirme çalışmalarında herhangi bir açıklamaya rastlanılmamıştır.

Geliştirilen ölçeklerin geçerlik analizlerine yönelik yapılan araştırma sonuçları incelendiğinde; ölçüt geçerliğinin sadece bir makalede yapıldığı görülmektedir. Bu bulgu çalışma açısından beklenen bir durumdur. Zira ölçüt geçerliği doğası gereği genellikle eğitim alanında testlerin geliştirilme sürecinde daha çok tercih edilen bir geçerlik türüdür. Bu bulguya ek olarak araştırma sonuçları, 2000 yılından bu yana yayımlanan makalelerde kapsam geçerliğinin test etme yöntemi konusunda bazı sorunların devam ettiğini ve günümüzde de güncelliğini kaybetmediğini göstermektedir. Nitekim, Slavec ve Drnovsek (2012) tarafında yapılan bir araştırmada içerik analizi yapılan çalışmaların çok az bir kısmında kapsam geçerliğine yer verildiği ortaya çıkmıştır. Bu çalışmada incelenen makalelerin yaklaşık yarısında kapsam geçerliğine ilişkin analiz ve bulgulara kısaca yer verilirken, geriye kalanların büyük çoğunluğunda kapsam geçerliğinin araştırıldığına dair net bir bilgiye rastlanılmamıştır. Aynı durum görünüş geçerliğine ait bulgular için de söz konusudur. İncelenen çalışmalarda görünüş ve kapsam geçerliğine ait net bir bulguya ulaşılmamasının nedeni; başlangıçta ölçek maddelerinin hazırlığı aşamasında bazı makalelerde doğrudan kapsam geçerliği için uzman görüşü alınmasına rağmen bazılarında ise dil, ifadelerin anlaşılabilirliği, ilgili davranışı ölçüp ölçmediği vb. amaçlar için uzman görüşüne başvurulduğuna yönelik ifadelerden kaynaklanmaktadır. Brinkmann (2009)’ın ifade ettiği gibi, görünüş geçerliğine benzer şekilde kapsam geçerliği de bir konsensüs konusudur ve bu nedenle kapsam geçerliği için uzmanların yapının tüm yönleriyle ele alınarak oluşturulduğunu kabul etmeleri gerekmektedir. Bu açıdan bakıldığında, araştırmacıların her iki geçerlik türü için de doğrudan uzman görüşüne başvurdukları düşünülmektedir. Buna paralel olarak Tavşancıl (2002), görünüş geçerliğinin genelde kapsam geçerliği içinde değerlendirildiğini ileri sürerek bu geçerlik türünde geçerlik düzeyinin sayısal değerlerle değil, uzman görüşü ile belirlendiğini ifade etmiştir. Öte yandan, bir bütün olarak testin ve testteki her bir maddenin maksada ne derece hizmet ettiği (Tekin, 1982, s.45) şeklinde tanımlanan kapsam geçerliği, çalışmanın genellikle ilk aşamalarında mutlaka yapılması gereken temel basamaklardan birini oluştururken; ölçeğin ölçmek istediği özelliği ölçüyor gözükmesi (Ercan ve Kan, 2004) şeklinde tanımlanan görünüş geçerliği ise ölçek geliştirildikten sonra sınırlanır. Dolayısıyla incelenen makalelerin bir kısmında ölçeğin görünüş geçerliğinin başlangıç aşamalarında yapılması, bazılarında ise görünüş geçerliğine ait yapılan işlemlerin kapsam geçerliği şeklinde tanımlanması araştırmacıların her iki geçerlik türünü tam olarak tanımlayamadıklarını hatta birbirine karıştırdıklarını düşündürmektedir. Bunun yanında, kapsam

geçerliği Ercan ve Kan (2004) tarafından ifade edildiği gibi mantıki yoldan ve istatistiki yoldan olmak üzere iki türlü incelenmektedir. Ancak tematik içerik analizine dâhil edilen makalelerin hiç birinde kapsam geçerlik oranı veya kapsam geçerlik indeksi gibi istatistiki yollara başvurulmadığı, sadece uzman görüşüne başvurularak kapsam geçerliğinin mantıki yoldan incelendiği görülmektedir. Dolayısıyla bu bulgular, geçerlik analizleri konusunda araştırmacıların kendilerini biraz daha geliştirmeleri gerektiğini göstermektedir.

Diğer taraftan geçerlik analizleri ile ilgili olarak tematik içerik analizine dâhil edilen makalelerin tamamında yapı geçerliğinin incelendiği görülmektedir. Bunun yanında yapı geçerliğinin sağlanmasında birçok yöntem önerilmekle birlikte en sık kullanılanlar faktör analizi, ayırt edicilik/ıraksaklık geçerliği (discriminant validity) ve yakınsaklık (convergent validity) geçerliğidir (Çokluk, Şekercioğlu ve Büyüköztürk, 2014). Çalışmada yapı geçerliği ile ilgili olarak elde edilen bulgulara bakıldığında; sadece 8 makalede ayırt edicilik geçerliğine başvurulduğu, yakınsaklığa ilişkin geçerliğin ise hiç kullanılmadığı görülmektedir. Churchill (1979)'e göre bir ölçme aracında yapının geçerli olabilmesi için, o yapıyı oluşturan değişkenler ile diğer yapıları ölçen değişkenler arasında düşük korelasyon gösteren bir ayırt edicilik geçerliğinin (discriminant validity) sağlanması gerekmektedir. Bu açıdan bakıldığında araştırmacıların ölçek geliştirme çalışmalarında ayırt edicilik geçerliğini göz ardı ettikleri görülmektedir. Araştırmacıların bu iki geçerlik türünü sıklıkla tercih etmemeleri, bu geçerlik türleri konusunda yeterli bilgiye sahip olmamalarından kaynaklanabileceği gibi, yapı geçerliği olarak faktör analizinin daha popüler olması ve bu analizi yapmayı yeterli görmelerinde kaynaklanabilir. Nitekim çalışmada elde edilen bulgulara bakıldığında; incelenen makalelerin tamamında yapı geçerliğini araştırmak için en sık kullanılan yöntemlerden biri olan faktör analizinin kullanıldığı anlaşılmaktadır. Erkuş (2012), psikolojik ölçek geliştirme sürecinde, faktör analizinin mutlaka başvurulması gereken bir istatistiksel işlem olduğunu vurgulamaktadır. Bu açıdan bakıldığında incelenen makalelerde faktör analizinin tercih edilmesi, çalışma açısından sevindiricidir. Faktör analizi ile ilgili olarak ise makalelerin tamamında açılımlayıcı faktör analizine (AFA) başvurulduğu, sadece 5 makalede doğrulayıcı faktör analizinin (DFA) yapıldığı görülmektedir. Hinkin (1995), Tosun ve Taşkesenligil (2014) tarafından yapılan çalışmalarda da AFA'nin sıklıkla tercih edildiği yönünde benzer bulgular ortaya konulmuştur. Bilindiği gibi AFA var olan psikolojik yapıyı anlamaya, DFA ise bu yapıyı test etmeye yönelik olup her iki tür de ölçek geliştirme sürecinde önemli olup birbirini tamamlamaktadırlar (Erkuş, 2012). Bununla beraber, tipik faktör analitik tekniklerinin zayıflıklarından biri, ortaya çıkan faktör yapısının uyum iyiliğini ortaya koymadaki yetersizliğidir (Long, 1983). Bu nedenle temel yapıyı belirlemek için analizlere AFA ile başlanmalı, daha sonra belirlenen bu yapıyı değerlendirmek için farklı bir örneklem üzerinden DFA yoluyla analizlere devam edilmesi gerekmektedir (Cabrera-Nguyen, 2010; Worthington ve Whittaker, 2006). Dolayısıyla, incelenen makalelerin büyük çoğunluğunda DFA'ya yer verilmemesi araştırmacıların bu konudaki bilgi ve beceri eksikliğinden kaynaklanabilir. Ayrıca, AFA analizlerinin SPSS gibi araştırmacılar tarafından yaygın olarak kullanılan istatistik programlarında yapılabilmesine rağmen DFA analizlerinin yapılabilmesi için LISREL, AMOS gibi daha az bilinen programlara ihtiyaç duyulmaktadır. Nitekim Hinkin (1995) tarafından yapılan çalışmada da ölçek geliştirmeye yönelik analizlerde LISREL programı kullanımının SPSS programına göre çok daha az tercih edildiği ortaya çıkmıştır. Buradan hareketle, araştırmacıların bu programların kullanımı konusundaki yetersizliklerinden dolayı DFA analizlerini çoğunlukla ihmal ettikleri söylenebilir.

Çalışmada AFA'ne yönelik elde edilen bulgular detaylı olarak incelendiğinde; korelasyon matrisinin faktör analizi yapmaya uygunluğu konusunda incelenen makalelerin sadece 3 tanesinde herhangi bir bilgiye rastlanmamıştır. Bununla birlikte makalelerin büyük bir çoğunluğunda Küresel Bartlett testi ve Kaiser-Meyer-Olkin (KMO) testinin birlikte kullanıldığı, ayrıca sadece 4 makalede bu testlere ek olarak madde başına düşen katılımcı sayısının da dikkate alındığı ortaya çıkmıştır. Worthington ve Whittaker (2006)'in yaptığı çalışmada da ortaya konulan benzer bulgular araştırmacıların genel olarak faktör analizinin başlangıç aşamasındaki faktörleşme konusunda öne sürülen ölçütlere kısmen dikkat ettiğini göstermektedir. Bununla beraber (Erkuş, 2012, s.99) faktör analizinin temelde maddeler arası korelasyona dayandığını ve korelasyonun da örneklemdaki gözlem

sayısına çok duyarlı olması nedeniyle faktör analizinin örneklem büyüklüğünden çok etkilendiğini ifade etmektedir. Bu durum korelasyon matrisinin faktör analizine uygunluğunun belirlenmesinde örneklemin de dikkate alınmasının bir gereklilik olduğunu göstermektedir. Ancak incelenen makalelerin çok az bir kısmında madde başına düşen katılımcı sayısının dikkate alınması dikkat çekici bir durumdur. Bu durum araştırmacıların örneklem yeterliliğinin test edilmesinde kullandıkları K.M.O. testini çalışmanın bu aşaması için yeterli gördüklerini düşündürebilir. Bununla beraber bu çalışma kapsamında, ölçek geliştirme çalışmalarında faktörleşme ölçütlerinin tamamının dikkate alınmasının ölçek geliştirme sürecinin sonraki aşamaları için daha doğru sonuçlar ortaya koyması açısından faydalı olacağı düşünülmektedir.

Yapı geçerliği aşamalarında kullanılan rotasyon metotları incelendiğinde, makalelerin çoğunluğunda döndürme işlemine başvurulduğu görülmektedir. Döndürme işlemi, faktörlerin daha kolay yorumlanabilmelerine olanak sağladığından, çalışmaların büyük çoğunluğunda bu yöntemle başvurulması doğaldır. Öte yandan döndürme işlemi genel olarak dik (orthogonal) ve eğik (oblique) olarak iki şekilde uygulanmaktadır. Bu çalışma kapsamında elde edilen bulgular incelendiğinde, makalelerin büyük çoğunluğunda dik döndürme (özellikle varimax) işleminin kullanıldığı görülmektedir. Büyüköztürk (2002), her iki döndürme sonuçlarının hemen hemen her zaman benzer sonuçlar ürettiğini belirterek, uygulamaların tamamına yakınında yorumlamada kolaylık sağladığından dik döndürmenin daha fazla tercih edildiği ifade etmektedir. Bu açıdan bakıldığında, incelenen makalelerde dik döndürmenin sıklıkla kullanılması beklenen bir durumdur.

AFA aşamalarında madde eleme/tutma ölçütleri ile ilgili olarak Worthington ve Whittaker (2006), faktör analizi sürecinde madde eleminin bu sürecin bir parçası olarak oldukça yaygın kullanıldığını ifade etmiş ve araştırmacıların bu süreçte faktör yükleri ile çapraz/binişik yük değerlerini çoğunlukla dikkate aldığını vurgulamıştır. Nitekim bu çalışmada da, Worthington ve Whittaker (2006)'in çalışmasına benzer şekilde incelenen makalelerin çok büyük bir kısmında faktör yükleri ve çapraz yüklerin dikkate alındığı görülmüştür. Bununla beraber ortak varyanslara sınırlı sayıda çalışmada yer verildiği ortaya çıkmıştır. Faktör yükleri, maddelerin ilgili faktörle korelasyonunu gösterir, bu bakımdan ölçek geliştirme sürecinde maddelerin hangi faktör altında yer aldığını belirlemede bu değerler dikkate alınır (Erkuş, 2012, s:98). Dolayısıyla araştırmacıların çalışmalarında bu iki özelliği göz ardı etmemeleri olumlu bir durum olarak karşılanabilir. Bununla beraber, uygulamada az sayıda madde için 0.30'a kadar indirilebilse de Büyüköztürk (2002) madde eleme için alt sınırın 0.45 veya yüksek olmasının seçim için iyi bir ölçüt olduğunu ileri sürmektedir. Ayrıca, son faktör sayısını belirlemeden önce maddeleri silmek, aslında tutulan faktör sayısını azaltabilir. Diğer taraftan, faktör çözümüne anlamlı bir katkı sağlamayan maddeleri gereksiz yere tutmak, belirlenecek son faktör sayısına karar vermeyi daha da zorlaştıracaktır (Worthington ve Whittaker, 2006). Bu açıdan bakıldığında, bu çalışmaya benzer nitelikte ileride yapılabilecek çalışmalarda araştırmacıların maddeleri elerken belirledikleri alt sınırların irdelenmesinin faydalı olacağı düşünülmektedir. Çalışmada yukarıda ifade edilenlerin yanı sıra, madde eleme/tutma ölçütü açısından makalelerin yarısından fazlasında madde analizinin yapıldığı görülmektedir. Madde analizinin Likert ölçeğinde kullanılma nedeni, Likert ölçekleme tekniğinin en önemli sayıtlısı olan "tek boyutluluğu" sağlamak içindir (Tavşancıl, 2002). Diğer taraftan Erkuş (2012), günümüzde artık faktör analizi tekniklerine sadece "yapı geçerliği" için başvurmanın anlamlı olmadığını ve faktör analizinde madde analizinin mutlaka yapılması gereken bir işlem olduğunu vurgulamaktadır. Dolayısıyla, incelenen makalelerin büyük çoğunluğunda madde analizinin yapılması sevindirici olmakla birlikte bu oran yeterli değildir. Dolayısıyla, ileride yapılacak herhangi bir ölçek geliştirme çalışmasında madde analizine mutlaka yer verilmesi gerektiği önerilmektedir.

İncelenen makalelerde AFA ile ilgili olarak faktör eleme/tutma ölçütlerine de bakılmıştır. Bulgular incelendiğinde, makalelerin yarısından fazlasında özdeğer ve yamaç eğim grafiğine yer verildiği görülmektedir. Makalelerde özdeğer ve yamaç eğim grafiklerinin sıklıkla kullanılmasının doğru bir tercih olduğunu destekler nitelikte Hinkin (1998), ölçeğin yapı geçerliği çalışmalarında teoriksel ayrımı desteklemede birden büyük özdeğerler ve yamaç eğim grafiğinin kullanılması

gerektiğini ifade etmiştir. Bununla birlikte makalelerin yarısından fazlasında faktör eleme/tutma ölçütü olarak açıklanan varyans oranına da bakıldığı anlaşılmaktadır. Çokluk vd (2014), faktör analizi sonucunda faktörlerin her bir değişken üzerinde yol açtıkları ortak varyans olarak tanımladıkları ortak faktör varyansı sonuçlarına bakarak madde çıkarma kararının verilmemesi gerektiğini, böyle bir problemin diğer çözümlerinde (faktör yükü, özdeğer vb.) gözlenmesi gerektiğini ifade etmektedir. Bu açıdan bakıldığında incelenen makalelerde faktör eleme açısından yukarıda ifade edilen diğer özelliklerin dikkate alınması araştırmacıların doğru bir seçim yaptığını göstermektedir. Öte yandan faktör eleme/tutma ölçütlerine bakıldığında, faktör başına düşen madde sayısının göz ardı edildiği anlaşılmaktadır. Bu bulgu, ilgili makalelerde faktör başına yeterli sayıda madde düşmesi nedeniyle araştırmacıların bu konuda bilgi vermeye gerek duymamasından da kaynaklanabilir. Bununla beraber Erkuş (2012), faktör sayısını belirlemede, sadece özdeğerler ve yamaç eğim grafiğinin birbirine göre farkı vs bağlı kalmanın doğru olmadığını; aynı zamanda olabildiğince her türlü faktör analitik tekniklere, ortaya çıkan olası alt yapıların birbirleriyle ilişkilerine vb kavramsal yapıyı da dikkate alarak başvurmak gerektiğini ifade etmektedir. Bu nedenle, ölçek geliştirme çalışmalarında faktör eleme veya faktör sayısına karar verme noktasında, olası tüm ölçütlerin dikkate alınarak analizlerin yürütülmesi önemli ve gerekli görülmektedir.

Bu araştırma kapsamında yukarıda ifade edilenlere ek olarak, tematik içerik analizine dâhil edilen makalelerde kullanılan güvenilirlik analiz yöntemleri de incelenmiştir. Buna göre elde edilen bulgular, makalelerin tamamında güvenilirlik analizleri için iç tutarlığa (C. Alpha) bakıldığını ortaya koymaktadır. Bununla beraber bulgular incelendiğinde, makalelerin yaklaşık üçte birinde testi yarılama yöntemlerinin tercih edildiği, ayrıca az sayıda da olsa test-tekrar test yönteminin kullanıldığı anlaşılmaktadır. Eşdeğer formların ise hiç kullanılmadığı görülmektedir. Bilindiği gibi iç tutarlık ve testi yarılama yöntemleri tek uygulamaya dayalı güvenilirlik yöntemleri iken; test-tekrar test ve eşdeğer formlar iki uygulamaya dayalı güvenilirlik yöntemleri olup bir ölçme aracının eş değer iki formunun hazırlanmasını veya aynı formun iki defa uygulamasını gerektirmektedir. Buradan anlaşılacağı üzere iki uygulamaya dayalı yöntemler daha fazla zaman, emek, masraf vb. gerektirir. Bu durum araştırmacıların güvenilirlik olarak tek uygulamaya dayalı yöntemleri daha fazla tercih etmelerine sebep olmuş olabilir. Bunun yanında güvenilirlik analizlerinde tüm makalelerde iç tutarlık yönteminin kullanılması doğru bir tercih olarak düşünülebilir. Zira Tezbaşaran (1996), Likert tipi ölçeklerin yapısıyla ilgili temel sayıtlardan birinin, ölçekte bulunan her bir maddenin ölçülen özellikle monotonik bir ilişki içinde olması yani her bir maddenin aynı özelliği ölçmesi nedeniyle Likert tipi ölçeklerde öncelikle iç tutarlığın (Cronbach α) sınanması gerektiğini belirtmektedir.

Yukarıda ifade edilen sonuçlardan yola çıkarak şu önerilere yer verilebilir:

- Tematik içerik analizinden elde edilen bulgular, araştırmacıların çoğunlukla tutum ölçeği geliştirmeye odaklandıklarını göstermektedir. Bu nedenle belirlenen ihtiyaçlar dâhilinde, öz-yeterlik, kaygı, algı vb. konularda daha fazla çalışmaya ihtiyaç duyulmaktadır.
- Elde edilen bulgular ölçek geliştirme çalışmalarının daha çok matematik eğitimi alanında yapıldığını ortaya koymaktadır. Bu nedenle biyoloji, fizik ve kimya eğitimi araştırmacılarının da kendi alanlarında daha fazla ölçek geliştirme çalışması yapmaya yönlendirilmeleri faydalı olacaktır.
- Ölçek geliştirme çalışmaları belirli bir hedef kitleye yönelik olmaktan ziyade farklı örneklem gruplarını da içerek şekilde geniş bir tabana yayılmalıdır.
- Çalışmanın bulguları ölçek geliştirme çalışmalarının genellikle orta düzeyde bir örneklem sayısını içerdiğini göstermektedir. Bu nedenle geliştirilen ölçeklerin daha güvenilir olabilmesi için daha büyük örneklemle çalışmaya ihtiyaç vardır.
- Araştırmacıların farklı geçerlik ve güvenilirlik yöntemleri konularında daha detaylı bilgi edinmesi sağlanmalıdır.

- Arařtırmacıların DFA analizlerini yapabilmeleri için SPSS dıřında LISREL gibi farklı istatistik programlarının kullanımı konusunda lisansüstü dersler açılmalı veya hizmet ii eđitim verilmelidir.
- Faktör ve madde eleme/tutma ölçütlerinin tamamı dikkate alınmalı ve buna göre bulgular bir bütün olarak deđerlendirilerek analizler yürütülmelidir.
- Bu alıřmada tematik ierik analizine dâhil edilen makalelerde DFA'nin oldukça sınırlı sayıda yürütüldüğü ortaya ıkmıř, bu nedenle DFA'ne ait detaylı incelemeler yapılmamıřtır. Bu nedenle ileride yapılabilecek benzer alıřmalarda DFA ile ilgili daha detaylı incelemeler yapılması önerilmektedir.
- Son olarak arařtırmaya dâhil olma ve hari tutulma ölçütleri dikkate alındığında, bu alıřmanın sadece fen, fen ve teknoloji ve matematik eđitimi alanında yapılmıř olması, belirli dizinlerde taranan dergilerin dikkate alınması nedeniyle özellikle YÖK veri tabanı ile ULAKBİM veri tabanında yayınlanan makale ve tezlerin arařtırma kapsamı dıřında tutulması, ölek uyarlama alıřmalarının arařtırma kapsamı dıřında tutularak sadece ölek geliştirme alıřmalarının incelenmesi gibi durumlar bu arařtırmanın sınırlılıkları olarak düşünölmektedir. Dolayısıyla ileride yapılacak benzer nitelikteki alıřmalarda bu sınırlılıkların da dikkate alınarak alıřma kapsamının daha da genişletilmesi önerilmektedir.

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*, 14(52), 23-33.
- Akgöz, S., Ercan, İ. ve Kan, İ. (2004). Meta-analizi. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30(2), 107-112.
- Aküzüm, C. ve Özmen, F. (2013). Eğitim denetmenlerinin rollerini gerçekleştirme yeterlikleri bir meta-sentez çalışması. *EKEV Akademi Dergisi*, 17(56), 97-120.
- Aleamoni, L. M. (1976). The relation of sample size to the number of variables in using factor analysis techniques. *Educational and Psychological Measurement*, 36, 879-883.
- Anderson, L. W. (1988). Attitudes and their measurement. J. P. Keeves, (Ed.), *Educational research, methodology and measurement: An international handbook* içinde (ss.421-426). New York: Pergamon Press.
- Azaltun, M. (2008). VI. araştırma yöntemleri semineri-ölçme ve ölçek geliştirme. *Anatolia: Turizm Araştırmaları Dergisi*, 19(1), 104-111.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36, 258-267.
- Brinkman, W. P. (2009). Design of a questionnaire instrument. S. Love (Edt). *Handbook of mobile technology research methods* içinde (ss.31-53). London: Nova Publisher.
- Büyüköztürk, Ş. (2002). Faktör analizi: Temel kavramlar ve ölçek geliştirmede kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*, 32, 470-483.
- 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.
- Churchill, G. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64-73.
- Comrey, A. L. ve Lee, H. L. (1992). *A first course in factor analysis*, Hillsdale, New Jersey: Erlbaum.
- Çalık, M., Ayas, A. ve Ebenezer, J.V. (2005). A review of solution chemistry studies: Insights into students' conceptions. *Journal of Science Education and Technology*, 14(1), 29-50.
- Çalık, M. ve Sözbilir, M. (2014). İçerik analizinin parametreleri. *Eğitim ve Bilim*, 39(174), 33-38.
- Çiltaş, A., Güler, G. ve Sözbilir, M. (2012). Türkiye'de matematik eğitimi araştırmaları: İçerik analizi çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(1), 515-580.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları* (3. Baskı). Pegem yayıncılık: Ankara.
- 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.
- Dawis, R. V. (1987). Scale construction. *Journal of Counseling Psychology*, 34(4), 481-489.
- Delice, A. (2010). Nicel araştırmalarda örneklem sorunu. *Kuram ve Uygulamada Eğitim Bilimleri*, 10(4), 1969-2018.
- Deveci, S. E., Çalmaz, A. ve Açık, Y. (2012). Doğu Anadolu'da yeni açılan bir üniversitenin öğrencilerinde kaygı düzeylerinin sağlık, sosyal ve demografik faktörler ile ilişkisi. *Dicle Tıp Dergisi*, 39(2), 189-196.
- DiStefano, C., Zhu, M. ve Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment Research & Evaluation*, 14(20), 1-10.
- Ercan, İ. ve Kan, İ. (2004). Ölçeklerde güvenilirlik ve geçerlik. *Uludağ Üniversitesi Tıp Fakültesi Dergisi*, 30(3), 211-216.

- Erkuş, A. (2007). Ölçek geliştirme ve uyarılama çalışmalarında karşılaşılan sorunlar. *Türk Psikoloji Bülteni*, 13(40), 17-25.
- Erkuş, A. (2012). *Psikolojide ölçek ve ölçek geliştirme-I: Temel kavramlar ve işlemler* (1. Baskı). Ankara: Pagem yayıncılık.
- Fassinger, R. (1987). Use of structural equation modeling in counseling psychology research. *Journal of Counseling Psychology*, 34(4), 425-436.
- Finfgeld, D. L. (2003). Metasynthesis: The state of the art-so far. *Qualitative Health Research*, 13(7), 893-904.
- Fraenkel, J. R. ve Wallen, N. E. (2000). *How to design & evaluate research in education* (4. baskı). London: McGraw Hill.
- Gençtürk, A. ve Memiş, A. (2010). İlköğretim okulu öğretmenlerinin öz-yeterlik algıları ve iş doyumlarının demografik faktörler açısından incelenmesi. *İlköğretim Online*, 9(3), 1037-1054.
- Gül, Ş. ve Sözbilir, M. (2015). Biology education research trends in Turkey: 1997-2012. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(1), 93-109. doi: 10.12973/eurasia.2015.1309a.
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967-988.
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 1(1), 104-121. doi: 10.1177/109442819800100106
- Kahn, J. H. (2006). Factor analysis in counseling psychology research, training, and practice: Principles, advances, and applications. *The Counseling Psychologist*, 34(5), 684-718.
- Karadağ, E. (2009). Eğitim bilimleri alanında yapılmış doktora tezlerinin tematik açıdan incelenmesi. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, 10(3), 75-87.
- Karagöz, Y. ve Ekici, S. (2004). Sosyal bilimlerde yapılan uygulamalı araştırmalarda kullanılan istatistiksel teknikler ve ölçekler. *Cumhuriyet Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 5(1), 25-43.
- Long, J. S. (1983). *Confirmatory factor analysis*. Beverly Hills, CA: Sage.
- Martens, M. P. (2005). The use of structural equation modeling in counseling psychology research. *The Counseling Psychologist*, 33(3), 269-298.
- McMillan, J.H. ve Schumacher, S. (2009). *Research in education: Evidence-based inquiry* (7th ed.). London: Pearson.
- 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.
- Özdamar, K. (2003). *Modern bilimsel araştırma yöntemleri*. Eskişehir: Kaan Kitapevi.
- Quintana, S. M. ve Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist*, 27(4), 485-527.
- Slavec, A. ve Drnovsek, M. (2012). A perspective on scale development in entrepreneurship research. *Economic and Business Review*, 14(1), 39-62.
- Sozbilir, M., Kutu, H. ve Yasar, M.D. (2012). Science education research in Turkey: A content analysis of selected features of papers published. D. Jorde ve J. Dillon (Eds), *Science education research and practice in Europe: Retrospective and prospective* içinde(ss.341-374). Rotterdam: Sense Publishers.
- Tavşancıl, E. (2002). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel yayınevi.
- Tekin H. (1982). *Eğitimde ölçme ve değerlendirme* (3. Baskı). Ankara: Mars Matbaası.
- Tezbaşaran A. A. (1996) *Likert tipi ölçek geliştirme kılavuzu*. Ankara: Türk Psikologlar Derneği Yayınları.
- Tinsley, H. E. A. ve Tinsley, D. J. (1987). Use of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, 34(4), 414-424.
- Topu, F. B., Baydaş, Ö., Turan, Z. ve Göktaş, Y. (2013). Öğretim teknolojisi araştırmalarında geçerlik ve güvenilirlik önlemleri. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 42(1), 110-126.

- Tosun, C. ve Taşkesenligil, Y. (2014, Eylül). *Türkiye’de fen eğitimi alanında geliştirilen/adapte edilen ölçeklerin ve başarı testlerinin doküman analizi*. XI. Ulusal Fen bilimleri ve Matematik Eğitimi Kongresi’nde sunulan sözlü bildiri, Adana.
- Walsh, D. ve Downe, S. (2005). Meta-synthesis method for qualitative research: A literature review. *Journal of Advanced Nursing*, 50(2), 204-211.
- 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.
- Yurdugül, H. (2005). *Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması*. XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi.
- Yurdugül, H. ve Bayrak, F. (2012). Ölçek geliştirme çalışmalarında kapsam geçerlik ölçüleri: Kapsam geçerlik indeksi ve Kappa istatistiğinin karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Özel Sayı 2*, 264-271.

Ek 1. Araştırmaya Dâhil Edilen Makalelerin Listesi

Makalenin kodu	Makalenin Künyesi
B1	Ekici, G. (2002). Biyoloji öğretmenlerinin laboratuvar dersine yönelik tutum ölçeği (BÇLDYTÖ). <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 22, 62-66
B2	Uzun, N. ve Sağlam, N. (2006). Ortaöğretim öğrencileri için çevresel tutum ölçeği geliştirme ve geçerliliği. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 30, 240-250
B3	Arıca, O. T. ve Ilgaz, G. (2007). Açımlayıcı ve doğrulayıcı faktör analizi ile biyoloji dersi tutum ölçeğinin yapı geçerliliğinin incelenmesi. <i>Eurasian Journal of Educational Research</i> , 28, 1-8
B4	Darçın, E. S. ve Güven, T. (2008). Development of an attitude measure oriented to biotechnology for the pre-service science teachers. <i>Türk Fen Eğitimi Dergisi</i> , 5(3), 72-81
F1	Tekbıyık, A. ve Akdeniz, A. R. (2010). Ortaöğretim öğrencilerine yönelik güncel fizik tutum ölçeği: Geliştirilmesi, geçerlik ve güvenilirliği. <i>Türk Fen Eğitimi Dergisi</i> , 7(4), 134-144
F2	Taşlıdere, E. ve Eryılmaz, A. (2012). Basit elektrik devreleri konusuna yönelik tutum ölçeği geliştirilmesi ve öğrencilerin tutumlarının değerlendirilmesi. <i>Türk Fen Eğitimi Dergisi</i> , 9(1), 31-46
K1	Koçak, C. ve Önen, A. S. (2012). Günlük yaşam kimyası tutum ölçeği geliştirme çalışması. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 43, 318-329
K2	Feyzioğlu, B., Demirdağ, B., Akyıldız, M. ve Altun, E. (2012). Kimya öğretmenlerinin laboratuvar uygulamalarına yönelik algıları ölçeği geliştirilmesi. <i>Türk Fen Eğitimi Dergisi</i> , 9(4), 44-63
FT1	Afacan, Ö. ve Aydoğdu, M. (2006). The science technology society (STS) course attitude scale. <i>International Journal of Environmental and Science Education</i> , 2(1), 189-201
FT2	Yaşar, Ş. ve Anagün, Ş. S. (2009). Reliability and validity studies of the science and technology course scientific attitude scale. <i>Türk Fen Eğitimi Dergisi</i> , 6(4), 43-54
FT3	Evrekli, E., İnel, D., Balım, A., G. ve Kesercioğlu, T. (2009). Fen öğretmen adaylarına yönelik yapılandırmacı yaklaşım tutum ölçeği: Geçerlilik ve güvenilirlik çalışması. <i>Türk Fen Eğitimi Dergisi</i> , 6(2), 134-148
FT4	Kağıtçı, B. ve Kurbanoglu, N. İ. (2013). Fen ve teknoloji dersine yönelik kaygı ölçeğinin geliştirilmesi: Güvenirlik ve geçerlik çalışması. <i>Türk Fen Eğitimi Dergisi</i> , 10(3), 95-107
FT5	Deveci, İ. ve Önder, İ. (2013). Fen ve teknoloji derslerinde verilen ödevlere yönelik öğretmen, öğrenci ve veli ölçeklerini geliştirme çalışması. <i>Türk Fen Eğitimi Dergisi</i> , 10(3), 159-184
M1	Çağırğan-Gülten, D. ve Derelioğlu, Y. (2006). Öğretmen adaylarının matematik öğrenmeyi öğretmeye ilişkin tutumlarını incelemeye yönelik bir ölçek geliştirme çalışması. <i>Eurasian Journal of Educational Research</i> , 24, 103-111
M2	Turanlı, N., Karakaş-Türker, N. ve Keçeli, V. (2008). Matematik alan derslerine yönelik tutum ölçeği geliştirilmesi. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 34, 254-262
M3	Işıksal, M. ve Aşkar, P. (2003). İlköğretim öğrencileri için matematik ve bilgisayar öz-yeterlik algısı ölçekleri. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 25, 109-118
M4	Cantürk-Günhan, B. ve Başer, N. (2007). Geometriye yönelik öz-yeterlik ölçeğinin geliştirilmesi. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 33, 68-76
M5	Özyürek, R. (2010). Matematik yetkinlik beklentisi bilgilendirici kaynaklar ölçeği'nin geçerlik ve güvenilirlik çalışmaları. <i>Kuram ve Uygulamada Eğitim Bilimleri</i> , 10(1), 419-447
M6	Deniz, L. ve Üldaş, İ. (2008). Validity and reliability study of the mathematics anxiety scale involving teachers and prospective teachers. <i>Eurasian Journal of Educational Research</i> , 30, 49-62
M7	Çalıkloğlu-Bali, G. (2002). Matematik öğretiminde dil ölçeği. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 23, 57-61
M8	Kayhan-Altay, M. ve Umay, A. (2013). İlköğretim ikinci kademe öğrencilerine yönelik sayı duyusu ölçeği'nin geliştirilmesi. <i>Eğitim ve Bilim</i> , 38(167), 241-255
M9	Arslan, O., Işıksal-Bostan, M. ve Şahin, E. (2013). Origaminin matematik eğitiminde kullanılmasına yönelik inanç ölçeği geliştirilmesi. <i>Hacettepe Üniversitesi Eğitim Fakültesi Dergisi</i> , 28(2), 44-57