

Matematik Eğitimi Lisansüstü Tezlerindeki Geçerlik ve Güvenirlik Çalışmalarının İncelenmesi¹

DOI: 10.26466/opus.735675

*

Sami Sezer Arbağ* - **Erhan Ertekin****

*Öğretmen, İstanbul İl Millî Eğitim Müdürlüğü Ölçme Değerlendirme Merkezi – İstanbul Fuat Sezgin Bilim ve Sanat Merkezi, İstanbul/Türkiye

E-Posta: samisezer.arbag@meb.gov.tr

ORCID: [0000-0002-9096-9952](https://orcid.org/0000-0002-9096-9952)

**Prof. Dr., Necmettin Erbakan Üniversitesi/Ahmet Keleşoğlu Eğitim Fakültesi Matematik ve Fen Bilimleri Eğitimi Bölümü, Matematik Eğitimi Anabilimdalı, Konya/Türkiye

E-Posta: ertekin75@gmail.com

ORCID: [0000-0002-6466-8996](https://orcid.org/0000-0002-6466-8996)

Öz

Bu araştırmanın amacı, matematik eğitimi alanında tamamlanan lisansüstü tezlerin geçerlik ve güvenirlilik çalışmalarını çağdaş standartlar yönünden incelemek ve hataları belirlemektir. Çalışma, matematik eğitiminde tamamlanmış doktora ve yüksek lisans tezlerinin geçerlik ve güvenirlilik çalışmalarının bazı kriterler yönünden incelenmesini hedeflediğinden nitel araştırmalardan durum çalışması olarak desenlenmiştir. Çalışmanın örneklemini 2000-2018 yılları arasında matematik eğitimi alanında yapılan ve YÖK Ulusal Tez Merkezinin internet sitesinde erişim engeli bulunmayan 92 doktora ve 343 yüksek lisans tezi olmak üzere toplam 435 lisansüstü tez oluşturmaktadır.

Araştırma sonucunda elde edilen bulgulara göre, incelenen tezlerin %65'inde rapor edilen güvenirlilik katsayısının 0.80'den düşük olduğu; %96'sında güvenirlilik katsayısının güven aralıklarının rapor edilmediği; %89'unda güvenirlilik kestirimine yönelik alınmış olan örneklem büyüklüğünün 400'den az olduğu belirlenmiştir. Bunun yanında en çok kullanılan güvenirlilik kestirim yöntemlerinin Kuder Richardson ve Cronbach alfa; geçerlik türlerinin kapsam ve yapı geçerliği; geçerlik kanıtlarının ise test içeriği ve iç yapı üzerine temellenmiş kanıt olduğu tespit edilmiştir. Bulgular tartışılarak olası çözüm ve öneriler ileri sürülmüştür.

Anahtar Kelimeler: Geçerlik, Güvenirlilik, Matematik Eğitimi

¹ Bu çalışma, Prof. Dr. Erhan Ertekin danışmanlığında, Sami Sezer Arbağ tarafından yapılan yüksek lisans tezinden üretilmiş ve 26-28 Eylül 2019'da gerçekleştirilen Uluslararası Türk Bilgisayar ve Matematik Eğitimi Sempozyumu'nda sözlü bildiri olarak sunulmuştur

Investigation of Validity and Reliability Studies in Mathematics Education Graduate Theses

*

Abstract

The aim of this research is to examine the validity and reliability studies of graduate theses completed in the field of mathematics education in terms of contemporary standards and to determine errors. The research was designed as a case study from qualitative research since it aimed to examine the validity and reliability studies of doctorate and master's theses completed in mathematics education in terms of some criteria. Sample of study was composed by a total of 435 postgraduate theses which include 92 doctorate theses and 343 postgraduate theses from the website of YÖK Center of International Thesis which were written between the years 2000-2018 in the field of mathematics education.

According to the findings obtained as a result of the research, the reliability coefficient reported in 65% of the examined theses is lower than 0.80; in 96% confidence interval of the reliability coefficients were not reported; in the sample size at 89% is less than 400. In additon, the most commonly used reliability estimation methods Kuder-Richardson and Cronbach alpha; scope and structure validity types; the validity evidence was found to be evidence based on test content and internal structure. Possible solutions and suggestions were proposed by discussing the findings.

Keywords: *Validity, Reliability, Math Education.*

Giriş

“Bilimsel bilginin üretilmesini ve toplumsal gelişmeyi sağlayan en önemli kurumlar üniversitelerdir” (Alkan, 2014). Üniversitelerin sorumluluğunda gerçekleştirilen lisansüstü eğitimler bilgi üretimine katkı sağlayan ve üniversitelerin araştırma yönünü destekleyen eğitimlerdir. Türkiye’deki üniversitelerde bazı lisans bölümlerinde bitirme projeleri veya mezuniyet projeleri yer alırken, yüksek lisans düzeyinde tezler veya bitirme projeleri, doktora düzeyinde ise, doktora tezleri şeklinde bilimsel çalışmalar yürütülmektedir. Üniversitelerde verilen eğitimin bir parçası olarak oluşturulan tezler, bilimin üretilmesine ve geliştirilmesine olumlu katkılar sağlamaktadır (Alkan, 2014; Benligiray, 2012).

Eğitim sistemine ilişkin uygulamaları şekillendirmede önemli bir rolü bulunan eğitim araştırmalarının sayısında son 10 yılda büyük bir artış olduğu gözlenmektedir. Yayınlanan bu çalışmalardan bir kısmı eğitim reformlarının temelini oluştururken, bir kısmı da literatürü gözden geçirme yoluyla önceki yapılan araştırma sonuçlarının güvenilirliğini test etmektedir (Odom, Brantlinger, Gersten, Horner, Thompson ve Harris, 2005; Onwuegbuzie ve Daniel, 2003). Eğitim veren bir lisansüstü programının eğilimlerini, izlediği seyri ve ilgi alanlarını ortaya koymada en önemli ve verimli yolun bu düzeyde yürütülen araştırmaların analiz edilmesi olduğu söylenebilir. Bu araştırmaların belirli zaman aralıklarında ve belirlenen çeşitli ölçütlere göre sistematik bir şekilde analiz edilmesi, alanın genel görüntüsünü ortaya çıkaracağından büyük önem arz etmektedir. Bu niceliksel artışa paralel olarak gelişen eğitim araştırmalarının nitelik olarak sorgulanması, araştırma yoluyla ulaşılan sonuçların gerçekleşmesi ve kullanılabilirliği çalışmaların kalitesinin ortaya konulabilmesi açısından büyük bir önem taşımaktadır. Söz konusu çalışmalar incelendiğinde, bazı yazarlar tarafından ortaya konan bulgularda yanlışların bulunabileceği, hatta gerçeğin, elde edilenin tam tersi olduğu da ifade edilmiştir. Bu durum, araştırmaların nitelikleri üzerine yapılan sentezlerin önemli ve gerekli bir çalışma olduğunu ortaya koymaktadır (Dunkin, 1996). Çalışmaların birçoğunda ölçme ve araştırma yöntem-bilimindeki belli başlı kavramların kullanımında bir takım hatalar göze çarpmaktadır. Hatalı kullanılan bu kavramların en önemlilerinden bazıları da geçerlik ve güvenirliliktir (Bademci, 2011a, 2017a, 2017b).

Amerika Eğitim Araştırma Derneği [American Educational Research Association (AERA)], Amerikan Psikoloji Derneği [American Psychological Association (APA)] ve Eğitimde Ölçme Ulusal Konseyi [National Council on Measurement in Education (NCME)] tarafından 1999 ve 2014 yıllarında yayınlanan ve “otorite” olarak da vurgulanan (Sireci, 2007) Eğitimsel ve Psikolojik Test Etme/Test Yapma Standartları [EPTS] (*Standards for Educational and Psychological Testing*) içinde geçerlik, “ölçümlerin kullanımlarının ve önerilen yorumların bir özelliği”; güvenirlilik ise, “testlerin kendilerinin değil, ölçümlerin bir özelliği” şeklinde kabul edilmiştir (AERA, APA ve NCME, 1999, 2014; Reynolds vd., 2009). EPTS'nin içeriğinde ölçüm güvenirliliği ve test ölçümlerinin yorumlarının geçerliğine ilişkin yönlendirici ilkeler, eğitim ve psikolojideki testlerin ve diğer ölçme araçlarının geliştirilmesi, uygulanması ve değerlendirilmesine yönelik ölçütler yer almaktadır. EPTS'de ve diğer bazı önemli kaynaklarda da açıkça görülebileceği üzere, geçerlik, ölçme araçlarının değil, ölçümlerin kullanımlarının bir özelliği; güvenirlilik ise, testlerin değil, ölçümlerin bir özelliğidir (Bademci, 2007, 2011a, 2017a, 2017b). Adı geçen husus bilimsel dayanaklarıyla ispatlanmış, dolayısıyla da “ölçüm yorum geçerliği” ve “ölçüm güvenirliliği” ifadelerinin daha uygun ifadeler olduğu ölçme yöntembilim dünyasında kabul görmüş ve geçerlik ve güvenirlilikle ilgili bir paradigma değişikliği ya da bir bilimsel devrim başlamıştır (Bademci, 2004, 2005a, 2005b, 2006a, 2006b, 2007, 2008, 2010, 2011a, 2017a, 2017b). Eldeki araştırma ile bu doğrultuda matematik eğitimi alanında yapılan yüksek lisans ve doktora tezlerindeki geçerlik ve güvenirlilik çalışmaları, geçerlik ve güvenirlilik kavramlarının kullanımı ve geçerlik-güvenirlilik çalışmalarında yapılan hatalar analiz edilmiştir. Böylece, geçerlik ve güvenirlilikle ilgili bu çalışma ile matematik eğitimine ilgi duyan yeni veya deneyimli araştırmacılara alandaki ölçme ve değerlendirmedeki yöntembilimsel durumla ilgili bilimsel tartışma ve sorgulama yapmalarına kaynak olunması hedeflenmiştir.

Alanyazın incelendiğinde lisansüstü tezlerin farklı açılardan incelendiği birçok araştırmayla karşılaşılabilir. Lisansüstü tezleri farklı bilim dallarında inceleyen araştırmalara (Altınkurt, 2007; Şimşek, Özdamar, Becit, Kılıçer, Akbulut ve Yıldırım, 2007; Kabaca ve Erdoğan, 2007; Tavşancıl, 2008; Korkmaz, 2010; Evrekli, İnel, Deniz ve Balım, 2011; Ergun ve Cilingir, 2013; Delice ve Ergene, 2015; Kutanis Özen, Özsoy, Karakiraz, Aras, Erol ve Uslu, 2015; Ulutaş, Üner, Turan Oluk, Yalçın Çelik ve Akkuş, 2015; Arslan, Geriş

ve Kulaksız, 2016; Kaya, Yazıcı, Deliveli ve Hoşgörür, 2016; İnce, Gül ve Bozyiğit, 2017; Şenyurt ve Özkan, 2017; Soycan ve Babacan, 2019) bakıldığında sadece güvenilirliğe odaklanılmış araştırmalar olsa da geçerlik ve güvenilirlik çalışmalarının birlikte ayrıntılı değerlendirildiği bir araştırmayla karşılaşmamıştır. Bu araştırmaların genelinde genel tez değerlendirmesinin içeriğinde geçerlik ve güvenilirlik çalışmalarına kısmen yer verildiği belirlenmiştir.

Lisansüstü tezlerde incelenen konuya uygun olan doğru yöntemlerin belirlenip kullanılması, böylelikle doğru ve bilimsel sonuçlara ulaşılması ve raporlanması son derece önemlidir. Bu çalışma matematik eğitimi alanında tamamlanmış tezlerdeki mevcut standartlar doğrultusunda geçerlik ve güvenilirlik kavramı ile ilgili hataların neler olduğunun ve bu hataların sıklıklarının tespit edilmesi, güvenilirlik ve geçerlik tespiti sürecinde yapılan hataların belirlenmesi ve benzer hataların tekrarının önüne geçilmesi açısından da önem arz etmektedir.

“Eğitim ve psikolojik test alanındaki en önemli ve temel kavramlardan biri güvenilirlik ve geçerliktir. Bu değere karşılık her iki kavramda en fazla yanlış anlaşılan ya da hatalı kullanılan kavramlar olmayı sürdürmektedir. Öncelikle lisansüstü programlar olmak üzere, yüksek lisans ve doktora programlarındaki ölçmeyle ilişkili konuların giderek azalması ya da ölçme konularıyla ilgili zayıf ve kalitesiz eğitim verilmesinin bu duruma sebep olduğu söylenebilir” (Bademci, 2007). Bu durumdan hareketle, Türkiye’de Matematik Eğitimi alanında tamamlanan lisansüstü tezlerdeki geçerlik ve güvenilirlik çalışmalarının incelemesini yapmak amacıyla planlanan bu araştırmada aşağıdaki sorulara yanıt aranmıştır.

1. Lisansüstü tezlerde güvenilirlik-geçerlik çalışması ve güvenilirlik doğrulturması yapılmı durumu nasıldır?
2. Lisansüstü tezlerde kullanılan güvenilirlik ve geçerlik ifadeleri nelerdir?
3. Güvenirlik çalışması yapılmış olan lisansüstü tezlerde kullanılan güvenilirlik kestirim yöntemi/yöntemleri rapor edilmiş midir? Rapor edilmiş ise hangi güvenilirlik kestirim yöntemi/yöntemleri kullanılmıştır?
4. Güvenirlik çalışması yapılan lisansüstü tezlerde güvenilirlik katsayısı rapor edilmiş midir? Rapor edilen güvenilirlik katsayısı yeterli düzeyde midir? Güvenirlik katsayısının güven aralıkları kestirilip rapor edilmiş midir?

5. Güvenirlik çalışması yapılmış lisansüstü tezlerde güvenirlik kestirimine yönelik alınmış örneklem büyüklüğünün rapor edilme ve yeterli olma durumu nasıldır?
6. Geçerlik çalışması yapılmış lisansüstü tezlerde hangi geçerlik türü/türleri kullanılmıştır?
7. Lisansüstü tezlerde geçerlik kanıtları rapor edilmiş midir? Geçerlik kanıtı rapor edilen lisansüstü tezlerde hangi geçerlik kanıtları kullanılmıştır?
8. Geçerlik çalışması yapılmış lisansüstü tezlerde, geçerlik kanıtı için hangi yöntem/yöntemler kullanılmıştır?

Yöntem

Araştırmanın Modeli

Bu araştırmada nitel araştırma yöntemlerinden durum çalışması kullanılmıştır. Durum çalışması, bir sistem hakkında genel bir bilgiye ulaşmak için tek veya çok sayıda sistem ögesinin gerçek bağlamı içinde ayrıntılı ve kapsamlı bir biçimde incelenmesi olarak ifade edilmektedir (Creswell, 2013). Bu araştırmada matematik eğitimi alanında tamamlanmış olan yüksek lisans ve doktora tezlerindeki geçerlik ve güvenirlik çalışmaları incelenen durumu oluşturmaktadır.

Veri Kaynağı

Nitel araştırma yöntemleri temel alınarak oluşturulan çalışmalarda araştırmacının amacı sonuçları tüm evrene genellemek değil ana problemi derinlemesine incelemektir. Bu yüzden araştırılan konunun daha iyi anlaşılması için araştırmacı amaçlı ya da isteyerek katılımcıları, çalışma grubunu belirler (Creswell, 2012). Bir araştırmada gözlenecek durum belirli niteliklere sahip olaylar, kişiler ve durumlardan oluşabilir. Bu durumda örneklem için belirlenen ölçütü karşılayan birimler (nesnelere, olaylar vb) örnekleme alınır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2008). Ölçüt örnekleme yöntemindeki temel amaç araştırmacı tarafından oluşturulan ya da daha önceden var olan ölçütleri karşılayan bütün durumların çalışılmasıdır (Yıldırım ve Şimşek, 2013).

Bu araştırmada Matematik Eğitimi alanında tamamlanmış yüksek lisans ve doktora tezlerindeki geçerlik ve güvenirlik ile ilgili mevcut durum ortaya

konulacaktır. Ancak alanda geçmişten günümüze yapılmış tüm tezlerin incelenmesi zaman, maliyet ve enerji açısından çok güç olduğundan araştırmanın örnekleminin oluşturulmasında amaçlı örnekleme yöntemlerinden ölçüt örnekleme yöntemi tercih edilmiştir. İncelenecek doküman seçimindeki üç ölçütten birincisi tezlerin 2000-2018 yılları arasında olması, ikincisi Matematik Eğitimi alanının alt anabilim dallarında tamamlanmış olması ve üçüncüsü nicel ve karma yöntem kullanılmış olmasıdır. Tarih kriteri belirlenirken EPTS'nin son sürümlerinin 1999 ve 2014 yılında yayımlanmış olması ve 2000 öncesinde yapılmış tezlerin dijital kopyalarına kolay ulaşım sağlanamaması durumları göz önünde bulundurulmuştur. Bu ölçütlere göre YÖK Ulusal Tez Merkezinin internet sitesinde matematik eğitimi alanında tamamlanmış ve erişim engeli bulunmayan 741 tez olduğu belirlenmiştir. Bu tezlerin 306 tanesi nitel araştırma ile yapılmış olduğu için araştırma kapsamından çıkarılmış, nicel ve karma araştırma ile yapılmış 92 doktora ve 343 yüksek lisans tezi olmak üzere toplam 435 tez araştırma kapsamına alınmıştır.

Veri Toplama Aracı ve Verilerin Toplanması

Araştırmada, nitel araştırma veri toplama yöntemlerinden doküman incelemesi kullanılmıştır. Doküman incelemesi, araştırılması hedeflenen olgu veya olgular hakkında bilgi içeren yazılı materyallerin analizini kapsar (Yıldırım ve Şimşek, 2013). Bu çalışma kapsamında ilk olarak akademik bir çalışmada geçerlik ve güvenilirlik çalışmalarının nasıl yapılması gerektiğine ilişkin yayınlar incelenmiştir. Ayrıca veri kaynakları tezler olduğu için tez yazım kılavuzları da gözden geçirilmiş ve sahip olması gereken özellikler belirlenmiş ve birinci araştırmacı tarafından tez incelemesinde kullanılmak üzere “Tez İnceleme Formu” oluşturulmuştur. Bu form 2 bölümden oluşmaktadır. Birinci bölümde teze ilişkin genel bilgiler, ikinci bölümde ise tezin yöntem bölümünde yer alan geçerlik ve güvenilirlik çalışmalarına ilişkin inceleme kriterleri yer almaktadır. Formun oluşturulmasında ilk olarak ilgili alanyazında daha önce başka araştırmacılar tarafından hazırlanılan formlar incelenmiştir. Ardından geçerlik ve güvenilirlik çalışmalarının standartlarına göre form oluşturulmuştur. “Tez İnceleme Formu” adıyla oluşturulan bu formun, çalışmanın amacına uygun olup olmadığını ortaya koymak amacıyla uzman görüşüne başvurulmuştur. Uzmanlardan alınan düzeltmeler

doğrultusunda, inceleme formu oluşturulmuştur. Bu formla araştırmacı tarafından 30 adet tez incelenmiş, eksiklikler belirlenmiş ve form güncellenmiştir. Formun bu haline ilişkin 3 farklı ölçme ve değerlendirme uzmanından tekrar görüş alınmış, forma son hâli verilmiştir.

Formun son hâli Tablo 1’de verilmiştir. Formda ölçütler ve sınırlandırılmış kodlamalar yer almaktadır.

Tablo1. Tez İnceleme Formu

Ölçütler	Kodlamalar
Tez Türü	(1) Doktora (2) Yüksek Lisans
Güvenirlik	
Güvenirlik Çalışması	(0) Yapılmamış (1)Yapılmış
Güvenirlik Kestirim Yöntemi	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Güvenirlik Kestirim Yöntemi	(1) Paralel Form (2) Test Tekrar Test (3) Yarıya Bölme (4) Kuder Richardson (5) Cronbach Alfa
Güvenirlik Doğuşturma	(0) Yapılmamış (1) Yapılmış
Güvenirlik İfadesi	(1) Testin güvenilirliği (2) Testin yanında aracın bir özelliği (3) Test güvenilirliği yanında ölçümlerin güvenilirliği (4) Farklı anlamlara gelen güvenilirlik ifadeleri (5) Güvenirlik ifadesine yer vermeyen
Güvenirlik Katsayısı	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Güvenirlik Katsayısı	(1) 0.80’den az (2) 0.80 ve 0.80’den fazla
Güvenirlik Katsayısının Güven Aralığı	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Güvenirlik Kestirimine Yönelik Alınmış Örneklem Büyüklüğü	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Güvenirlik Kestirimine Yönelik Alınmış Örneklem Büyüklüğü	(1) 400’den az (2) 400 ve 400’den fazla
Geçerlik	
Geçerlik Çalışması	(0) Yapılmamış (1) Yapılmış
Geçerlik Türü	(1) Kapsam geçerliği (2) Ölçüt ilişkili geçerlik(Yordayıcı, Eş zamanlı) (3)Yapı geçerliği
Geçerlik İfadesi	(1) Test geçerliği (2) Test ölçümlerinin geçerliği (3) Testin geçerliği yanında test ölçümlerinden yapılan yorumun özelliği (4) Farklı anlamlara gelen geçerlik ifadeleri (5) Geçerlik ifadesine yer vermeyen
Geçerlik Derecesi	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Geçerlik Derecesi	(1) Düşük (2) Orta (3) Yüksek
Geçerlik Kanıtı	(0) Rapor Edilmemiş (1) Rapor Edilmiş
Geçerlik Kanıt Kaynakları	(1) Test içeriği üzerine temellenmiş kanıt (2) Yanıt süreçleri üzerine temellenmiş kanıt (3) İç yapı üzerine temellenmiş kanıt (4) Diğer değişkenlerle ilişkiler üzerine temellenmiş kanıt (5) Test etmenin sonuçları üzerine temellenmiş kanıt
Geçerlik Kanıtına Yönelik Kullanılan Yöntem	(1) Uzman Görüşü (2) Korelasyon Analizi (3) Faktör Analizi (4) Test- Madde Analizi

Dokuman incelemesi sürecinde, bir araştırmacı öncelikle amacına yönelik mevcut kaynakları bulur, her bir kaynağı dikkatlice okur, gerekli bilgileri not alır ve aldığı notlardan yola çıkarak bazı değerlendirme işlemleri yapar (Çepni, 2014). Birinci araştırmacı YÖK Ulusal Tez Merkezinde yayımlanmış yüksek lisans ve doktora tezlerini bularak okumuş, “Tez İnceleme Formu” nu kullanarak araştırmaları geçerlik ve güvenilirlik açısından incelemiştir.

Araştırma sonuçlarının geçerliğinin teyidi için inandırıcılık ve aktarılabirliklerini incelemek gerekir (Lincoln ve Guba, 1985). İnanırıcılığı sağlamak için; araştırmada incelenen tezlerden bir kaçının yazarıyla görüşme sağlanmış güvenilirlik ve geçerlik çalışmaları ile ilgili görüşmeler yapılmış, bazı tezler farklı zamanlarda tekrar incelenmiş, elde edilen veriler özetlenmiş ve eleştirel bir gözle bakılarak araştırma sorularına yanıtlar açıklanmıştır.

Araştırma sonuçların güvenilirliğinin teyidi için tutarlık ve teyit edilebilirliklerini incelemek gerekir (Lincoln ve Guba, 1985). Tutarlığın sağlanması için ise kodlayıcılar arası puanlayıcı güvenilirliğini belirlemek amacıyla rastgele seçilen 30 tez ölçme değerlendirme uzmanları ve araştırmacı tarafından veri toplama aracı doğrultusunda oluşturulan forma göre ayrı ayrı puanlanmış ve puanlayıcılar arası tutarlığı belirlemek için Miles ve Huberman’ın (2015) önerdiği uzlaşma katsayısı hesaplanmıştır.

$$\text{Güvenirlik} = \text{Uzlaşma sayısı} / (\text{Uzlaşma sayısı} + \text{Uzlaşmama sayısı})$$

Elde edilen sonuç ile puanlayıcılar arası tutarlılık anlamındaki güvenilirliğin sağlandığı görülmüştür. 30 tezin üç farklı ölçme ve değerlendirme uzmanı tarafından incelenmesi sonucu hesaplanan güvenilirlik katsayısı %84 olarak bulunmuştur. Bu durumda, kodlamayı yapan uzmanlar arasında tutarlılık olduğu, diğer bir deyişle puanlayıcılar arası uzlaşma anlamındaki güvenilirliğin sağlandığı görülmüştür. Teyit edilebilirliğin sağlanması için ise araştırmada kullanılan inceleme formu ve Excel’de yapılan kodlamalar, gerekli olduğu takdirde incelenmeye sunulması için araştırmacı tarafından saklanmaktadır.

Verilerin Analizi

Araştırma kapsamında matematik eğitimi alanındaki lisansüstü tezler için oluşturulan formula ulaşılan veriler kaydedilmiştir. Elde edilen veriler Excel

programı kullanılarak frekans ve yüzde değerleri ile çözümlenmiştir. Sonuçlar yüzde ve frekans tabloları ile betimsel olarak analiz edilmiştir.

Aktarılabirliğin sağlanması için; verileri ayrıntılı bir biçimde betimlemek için elde edilen bulguların yazılmasında tablo, grafik ve şekiller kullanılmış, ayrıntılı şekilde betimlemeler yapıp yorumlanmış, araştırmada matematik eğitimi alanında tamamlanmış tezler incelenmiş, alt ana bilim dalları ile çeşitlilik sağlanmaya çalışılarak bazı ölçütler kullanılarak asıl örneklem belirlenmiştir.

Bulgular

Bu bölümde matematik eğitimi alanında tamamlanmış olan lisansüstü tezlerindeki geçerlik ve güvenirlik çalışmalarının incelenmesi sonucunda elde edilen bulgulara yer verilmiştir.

Lisansüstü tezlerde, güvenirlik-geçerlik çalışması ve güvenirlik doğrulturması yapıma durumuna ilişkin dağılımlar Tablo-2'de sunulmuştur.

Tablo 2. Lisansüstü tezlerin güvenirlik-geçerlik çalışması ve güvenirlik doğrulturması yapıma durumuna ilişkin dağılım

Durum		n	%	n	%	n	%
		Doktora		Yüksek Lisans		Toplam	
Güvenirlik Çalışması	Yapılmış	86	19.77	265	60.92	351	80.69
	Yapılmamış	6	1.38	78	17.93	84	19.31
Geçerlik Çalışması	Yapılmış	89	20.46	285	65.52	374	85.98
	Yapılmamış	3	0.69	58	13.33	61	14.02
Geçerlik ve Güvenirlik Çalışması	Sadece Güvenirlik Çalışması Yapılmış	6	1.38	72	16.55	78	17.93
	Sadece Geçerlik Çalışması Yapılmış	3	0.69	42	9.66	45	10.34
	Hem Güvenirlik Hem Geçerlik Çalışması Yapılmış	83	19.08	207	47.59	290	66.67
	Güvenirlik ve Geçerlik Çalışması Yapılmamış	0	0.00	22	5.06	22	5.06
Güvenirlik Doğulturması	Yapılmış	14	3.22	81	18.62	95	21.84
	Yapılmamış	78	17.93	262	60.23	340	78.16

Genel anlamda lisansüstü tezler incelendiğinde; 351'i (%80.69) için güvenirlik çalışmasının yapıldığı; 84'ü (%19.31) için ise yapılmadığı tespit edilmiştir. Buna göre lisansüstü tezlerde neredeyse her 5 tezdten 1'inde güvenirlik çalışması yapılmadığı söylenebilir. Özel olarak doktora tezleri incelendi-

ğinde; 92 tezin 86'sı (%19.77) için güvenilirlik çalışması yapılırken, 6'sı (%1.38) için yapılmamıştır. Yüksek lisans tezlerinde ise 343 tezin 265'i (%60.92) için güvenilirlik çalışması yapıldığı, 78'i (%17.93) için yapılmadığı görülmektedir. Geçerlik bağlamında ise toplamda 435 tezin 374'ü (%85.98) için geçerlik çalışmasının yapıldığı, 61'i (%14.02) için yapılmadığı tespit edilmiştir. Doktora tezleri incelendiğinde 92 tezin 89'u (%20.46) için geçerlik çalışması yapılırken, 3'ü (%0.69) için yapılmamıştır. Yüksek lisans tezlerinde ise 343 tezin 285'i (%65.52) için geçerlik çalışması yapıldığı, 58'i (%13.33) için yapılmadığı görülmektedir.

Lisansüstü tezlerin geçerlik ve güvenilirlik çalışmaları birlikte incelendiğinde, toplam 435 tezin 290'ı (%66.67) için hem güvenilirlik hem de geçerlik çalışmasının yapıldığı, 78'i (%17.93) için geçerlik çalışması yapılmaksızın sadece güvenilirlik çalışmasının yapıldığı, 45'i (%10.34) için ise güvenilirlik çalışması yapılmaksızın sadece geçerlik çalışması yapıldığı tespit edilmiştir. 22 tezde (%5.06) hem güvenilirlik ve hem geçerlik çalışması yapılmadığı ve bu tezlerin tümünün yüksek lisans tezleri olduğu dikkat çekmektedir. Matematik eğitimi alanında tamamlanmış her 20 tezin 1'inde hem güvenilirlik ve hem geçerlik çalışması yapılmamıştır. Doktora tezleri incelendiğinde 92 tezin 83'ü (%19.08) için hem güvenilirlik hem de geçerlik çalışması, 6'sı (%1.38) için geçerlik çalışması yapılmaksızın sadece güvenilirlik çalışması, 3'ü (%0.69) için ise güvenilirlik çalışması yapılmaksızın sadece geçerlik çalışmasının yapıldığı belirlenmiştir. Yüksek lisans tezlerinde ise 343 tezin 207'si (%47.59) için hem güvenilirlik hem geçerlik çalışması, 72'si (%16.55) için geçerlik çalışması yapılmaksızın sadece güvenilirlik çalışması, 42'si (%9.66) için ise güvenilirlik çalışması yapılmaksızın sadece geçerlik çalışmasının yapıldığı görülmektedir. 22'sinde (%5.06) ise ne geçerlik ne de güvenilirlik çalışması yapılmıştır.

Lisansüstü tez yazarlarının kullanmış oldukları veri toplama araçları için aracın geliştiricisinin örneklemeden elde ettiği güvenilirlik kanıtlarını kendi tezinde rapor etme durumu incelendiğinde, lisansüstü tezlerin 95'i (%21.05) için güvenilirlik doğrulturmasının yapıldığı bir başka ifadeyle daha önceden geliştirilen aracın uygulanmasından elde edilen güvenilirlik analizlerinin değiştirilmeden rapor edildiği ve 340'ı (%78.16) için ise güvenilirlik doğrulturmasının yapılmadığı tespit edilmiştir. Doktora tezleri incelendiğinde 92 tezin 14'ü (%3.22) için güvenilirlik doğrulturması yapılırken, 78'i (%17.93) için ise yapılmamıştır. Yüksek lisans tezlerinde ise 343 tezin 81'i için (%18.62)

güvenirlik doğuşturmamasının yapıldığı, 262'si (%60.23) için ise yapılmadığı tespit edilmiştir.

Ayrıca araştırmamız bağlamında lisansüstü tezlerde kullanılan güvenilirlik ve geçerliğe ilişkin ifadeler incelenmiş ve ifadelere ilişkin dağılımlar Tablo-3'te sunulmuştur.

Tablo 3. Lisansüstü tezlerdeki güvenilirlik ve geçerlik ifadelerine ilişkin dağılım

İfadeler	n	%	n	%	n	%	
	Doktora		Yüksek Lisans		Toplam		
Güvenirlik İfadeleri	"Testin Güvenirliği"	53	12.18	169	38.85	222	51.03
	"Testin Güvenirliği" yanında "Aracın Bir Özelliği"	15	3.45	62	14.25	77	17.7
	"Test Güvenirliği" yanında "Ölçümlerin Güvenirliği"	17	3.91	38	8.74	55	12.64
	Farklı Anlamalara Gelen Güvenirlik İfadeleri	2	0.46	13	2.99	15	3.45
	Güvenirlik İfadesine Yer Vermeyen	5	1.15	61	14.02	66	15.17
	"Testin Geçerliği"	74	17.01	242	55.63	316	72.64
	"Test Ölçümlerinin Geçerliği"	2	0.46	24	5.52	26	5.98
Geçerlik İfadeleri	"Testin Geçerliği" yanında "Test Ölçümlerinden Yapılan Yorumun Özelliği"	12	2.76	5	1.15	17	3.91
	Farklı Anlamalara Gelen Geçerlik İfadeleri	1	0.23	20	4.60	21	4.83
	Geçerlik İfadelerine Yer Vermeyen	3	0.69	52	11.95	55	12.64

Tablo 3 incelendiğinde, lisansüstü tezlerin 222'si (%51.03) için "Testin güvenilirliği", 77'si (%17.7) için "Testin Güvenirliği" yanında "Aracın Bir Özelliği", 55'i (%12.64) için "Test Güvenirliği" yanında "Ölçümlerin Güvenirliği" ifadesinin kullanıldığı tespit edilmiştir. Güvenirlik ifadesine yer vermeyen tez sayısının 66 (%15,17) olduğu görülmektedir. Ayrıca 15 tezde (%3.45) farklı anlamlara gelen güvenilirlik ifadeleri ile karşılaşmıştır. Doktora tezleri incelendiğinde 92 tezin 53'ü (%12.18) için "Testin güvenilirliği", 17'si (%3.91) için "Test Güvenirliği" yanında "Ölçümlerin Güvenirliği", 15'i (%3.45) için "Testin Güvenirliği" yanında "Aracın Bir Özelliği" ifadeleri kullanılmıştır. 5 doktora tezinde (%1.15) güvenilirlik ifadesine yer verilmemiş olup, 2 tezde (%0.46) ise farklı anlamlara gelen güvenilirlik ifadeleri kullanılmıştır. Yüksek lisans tezlerinde ise 343 tezin 169'u (%38.85) için "Testin güvenilirliği", 62'si (%14.25) için "Testin Güvenirliği" yanında "Aracın Bir Özelliği"

ği”, 38’i (%8.74) için “Test Güvenirliği” yanında “Ölçümlerin Güvenirliği” ifadeleri kullanılmıştır. 61 yüksek lisans tezinde (%14.02) güvenilirlik ifadesine yer verilmemiş olup, 13 tezde (%2.99) ise farklı anlamlara gelen güvenilirlik ifadeleri kullanılmıştır.

Toplamda 435 tezin 316’sı (%72.64) için ise “Testin geçerliği”, 26’sı (%5.98) için “Test Ölçümlerinin Geçerliği”, 17’si (%3.91) için “Testin Geçerliği” yanında “Test Ölçümlerinden Yapılan Yorumun Özelliği” ifadesinin kullanıldığı tespit edilmiştir. Geçerlik ifadesine yer vermeyen tez sayısının 55 (%12.64) olduğu görülmektedir. Ayrıca 21 tezde (%4.83) farklı anlamlara gelen geçerlik ifadeleri ile karşılaşmıştır. Doktora tezleri incelendiğinde 92 tezin 74’ü (%17.01) için “Testin geçerliği”, 2’si (%0.46) için “Test Ölçümlerinin Geçerliği”, 12’si (%2.76) için “Testin Geçerliği” yanında “Test Ölçümlerinden Yapılan Yorumun Özelliği” ifadeleri kullanılmıştır. 3 doktora tezinde (%0.69) geçerlik ifadesine yer verilmemiş olup, 1 tezde (%0.23) ise farklı anlamlara gelen geçerlik ifadeleri kullanılmıştır. Yüksek lisans tezlerinde ise 343 tezin 242’si (%55.63) için “Testin geçerliği”, 24’ü (%5.52) için “Test Ölçümlerinin Geçerliği”, 5’i (%1.15) için “Testin Geçerliği” yanında “Test Ölçümlerinden Yapılan Yorumun Özelliği” ifadeleri kullanılmıştır. 52 yüksek lisans tezinde (%11.95) geçerlik ifadesine yer verilmemiş olup, 20 yüksek lisans tezinde (%4.60) ise farklı anlamlara gelen geçerlik ifadeleri kullanılmıştır.

Lisansüstü tezlerde güvenilirlik kestirim yöntemi, güvenilirlik katsayısı ve güvenilirlik kestirimine yönelik örneklem büyüklüğüne ilişkin dağılımlar Tablo-4’te sunulmuştur.

Güvenirlik çalışması yapılmış 351 lisansüstü tezin 295’i (%84.05) için güvenilirlik kestirim yönteminin rapor edildiği; 56’sı (%15.95) için ise rapor edilmediği tespit edilmiştir. Doktora tezleri incelendiğinde 86 tezin 81’i (%23.08) için güvenilirlik kestirimi rapor edilirken, 5’i (%1.42) için rapor edilmemiştir. Yüksek lisans tezlerinde ise 265 tezin 214’ü (%60.97) için güvenilirlik kestirim yönteminin rapor edildiği, 51’i (%14.53) için rapor edilmediği görülmektedir.

Güvenirlik çalışması yapılmış ve güvenilirlik kestirim yöntemi rapor edilmiş toplam 295 lisansüstü tezin 101’i (%34.24) için sadece Cronbach Alfa yöntemi, 45’i (%15.25) için Cronbach Alfa ve Kuder-Richardson yöntemleri, 38’i (%12.88) için sadece Kuder Richardson yöntemi, 33’ü (%11.19) için yarıya bölme ve Cronbach Alfa yöntemi, 25’i (%8.47) için Cronbach Alfa, Kuder-

Richardson ve yarıya bölme yöntemleri, 18'i (%6.10) için yarıya bölme ve Kuder-Richardson yöntemleri kullanılmıştır. Sadece paralel form ve sadece test-tekrar test yöntemi kullanılan teze rastlanmamıştır. Sadece 1 yöntem kullanan tezlerin sayısının 145 (%49.16), 2 yöntemi birlikte kullanan tezlerin sayısının 118 (%40), 3 yöntem kullanan tezlerin sayısının ise 32 (%10.84) olduğu görülmektedir.

Tablo 4. Lisansüstü tezlerdeki güvenirlilik kestirim yöntemi, güvenirlilik katsayısı ve güvenirlilik kestirimine yönelik alınmış olan örneklem büyüklüğüne ilişkin dağılım

Durum		n	%	n	%	n	%	
		Doktora		Yüksek Lisans		Toplam		
Güvenirlilik Kestirim Yöntemi	Rapor Edilmiş	81	23.08	214	60.97	295	84.05	
	Rapor Edilmemiş	5	1.42	51	14.53	56	15.95	
Güvenirlilik Katsayısı	Rapor Edilmiş	81	23.08	234	66.67	315	89.74	
	Rapor Edilmemiş	5	1.42	31	8.83	36	10.26	
Güvenirlilik Katsayısının Güven Aralığı	Rapor Edilmiş	4	1.27	8	2.54	12	3.81	
	Rapor Edilmemiş	77	24.44	226	71.75	303	96.19	
Güvenirlilik Kestirimine Yönelik Alınmış Örneklem Büyüklüğü	Rapor Edilmiş	84	23.93	247	70.37	331	94.30	
	Rapor Edilmemiş	2	0.57	18	5.13	20	5.70	
Yöntem ve Düzey								
Güvenirlilik Kestirim Yöntemi	Sadece Paralel Form	0	0.00	0	0.00	0	0.00	
	Sadece Test Tekrar Test	0	0.00	0	0.00	0	0.00	
	Sadece Kuder Richardson	6	2.03	32	10.85	38	12.88	
	Sadece Cronbach Alfa	17	5.76	84	28.47	101	34.24	
	Sadece Yarıya Bölme	1	0.34	5	1.69	6	2.03	
	Cronbach Alfa ve Kuder Richardson	14	4.75	31	10.51	45	15.25	
	Yarıya Bölme ve Kuder Richardson	6	2.03	12	4.07	18	6.10	
	Test Tekrar Test ve Cronbach Alfa	3	1.02	7	2.37	10	3.39	
	Yarıya Bölme ve Cronbach Alfa	12	4.07	21	7.12	33	11.19	
	Kuder Richardson ve Test Tekrar Test	4	1.36	8	2.71	12	4.07	
	Cronbach Alfa, Kuder Richardson ve Test Tekrar Test	7	2.37	0	0	7	2.37	
	Cronbach Alfa, Kuder Richardson ve Yarıya Bölme	11	3.73	14	4.75	25	8.47	
	Güvenirlilik Katsayısı	0.80'den az	19	6.03	92	29.21	111	35.24
		0.80 ve 0.80'den fazla	62	19.68	142	45.08	204	64.76
	Güvenirlilik Kestirimine Yönelik Alınmış Örneklem Büyüklüğü	400'den az	61	18.43	234	70.69	295	89.12
400 ve 400'den fazla		23	6.95	13	3.93	36	10.88	

Güvenirlilik çalışması yapılmış 351 lisansüstü tezin 315'i (%89.74) için güvenirlilik katsayısının rapor edildiği; 36'sı (%10.26) için ise rapor edilmediği tespit edilmiştir. Doktora tezleri incelendiğinde 86 tezin 81'i (%23.08) için güvenirlilik katsayısı rapor edilirken, 5'i (%1.42) için rapor edilmemiştir. Yüksek lisans tezlerinde ise 265 tezin 234'ü (%66.67) için güvenirlilik katsayısının rapor edildiği, 31'i (%8.83) için ise rapor edilmediği görülmektedir. Araştırma kapsamına alınan lisansüstü tezlerdeki her 10 tezden 1'inde güvenirlilik katsayısı rapor edilmemiştir.

Güvenirlilik çalışması yapılmış ve güvenirlilik katsayısı rapor edilmiş toplam 315 lisansüstü tezin 111'i (%35.24) için güvenirlilik katsayısının 0.80'den az ve 204'ü (%64.76) için 0.80 ve 0.80'den fazla olduğu tespit edilmiştir. Doktora tezleri incelendiğinde 81 tezin 19'u (%6.03) için güvenirlilik katsayısı 0.80'den az iken, 62'si (%19.68) için 0.80 ve 0.80 üzerindedir. Yüksek lisans tezlerinde ise 234 tezin 92'si (%29.21) için güvenirlilik katsayısının 0.80'den az olduğu, 142'si (%45.08) için ise 0.80 ve 0.80 üzerinde olduğu belirlenmiştir.

Güvenirlilik çalışması yapılmış ve güvenirlilik katsayısı rapor edilmiş toplam 315 lisansüstü tezin 12'si (%3.81) için güvenirlilik kat sayısının güven aralıkları rapor edilirken, 303'ü (%96.19) için rapor edilmemiştir. Doktora tezleri incelendiğinde 81 tezin 4'ü (%1.27) için güvenirlilik katsayısının güven aralıkları rapor edilmiş, 77'si (%24.44) için ise rapor edilmemiştir. Yüksek lisans tezlerinde ise 234 tezin 8'i (%2.54) için güvenirlilik katsayısının güven aralıkları rapor edilmiş, 226'sı (%71.75) için rapor edilmemiştir.

Güvenirlilik çalışması yapılmış 351 lisansüstü tezin 331'i (%94.30) için örneklem büyüklüğünün rapor edildiği; 20'si (%5.70) için rapor edilmediği tespit edilmiştir. Doktora tezleri incelendiğinde 86 tezin 84'ü (%23.93) için örneklem büyüklüğü rapor edilirken, 2'si (%0.57) için rapor edilmemiştir. Yüksek lisans tezlerinde ise, 265 tezin 247'si (%70.37) için örneklem büyüklüğü rapor edilirken, 18'i (%5.13) için rapor edilmemiştir. Buna göre; her 20 lisansüstü tezin 1'inde güvenirlilik kestirimine yönelik alınmış olan örneklem büyüklüğünün rapor edilmediği ifade edilebilir.

Güvenirlilik çalışması yapılmış ve güvenirlilik kestirimine yönelik alınmış olan örneklem büyüklüğü rapor edilmiş toplam 331 lisansüstü tezin 295'i (%89.12) için örneklem büyüklüğünün 400'den az, 36'sı (%10.88) için ise 400 ve üzerinde olduğu görülmektedir. Doktora tezleri incelendiğinde; 84 tezin 61'i (%18.43) için örneklem büyüklüğünün 400'den az, 23'ü (%6.95) için ise

400 ve üzeri olduğu tespit edilmiştir. Yüksek lisans tezlerinde ise 247 tezin 234'ü (%70.69) için örneklem büyüklüğünün 400'den az, 13'ü (%3.93) için ise 400 ve üzerinde olduğu görülmektedir. Buna göre; araştırma kapsamına alınan lisansüstü tezlerin yaklaşık %11'inde örneklem büyüklüğünün yeterli düzeyde olduğu, yaklaşık %89'unda ise yetersiz olduğu (400'den az) söylenebilir.

Lisansüstü tezlerde geçerlik türü, derecesi, kanıtı, kanıt kaynakları ve geçerlik kanıtına yönelik kullanılan yöntemlere ilişkin dağılımlar Tablo-5'te sunulmuştur.

Tablo 5. Lisansüstü tezlerdeki geçerlik türü, derecesi, kanıtı, kanıt kaynakları ve geçerlik kanıtına yönelik kullanılan yöntemlere ilişkin dağılım

Durum		n		%			
		Doktora	Yüksek Lisans	Toplam			
Geçerlik Derecesi	Rapor Edilmiş	0	0	0	0		
	Rapor Edilmemiş	89	23.80	285	76.20		
Geçerlik Kanıtı	Rapor Edilmiş	89	23.80	285	76.20		
	Rapor Edilmemiş	0	0	0	0		
Tür, Kanıt ve Yöntem							
Geçerlik Türü	Sadece Kapsam Geçerliği	26	6.95	143	38.24	169	45.19
	Sadece Yapı Geçerliği	13	3.48	31	8.29	44	11.76
	Kapsam Geçerliği ve Ölçüt İlişkili Geçerlik	5	1.34	11	2.94	16	4.28
	Kapsam Geçerliği ve Yapı Geçerliği	37	9.89	92	24.60	129	34.49
	Kapsam Geçerliği, Ölçüt İlişkili Geçerlik ve Yapı Geçerliği	8	2.14	8	2.14	16	4.28
Geçerlik Kanıtı	Sadece "Test İçeriği Üzerine Temellenmiş Kanıt"	24	6.42	138	36.90	162	43.32
	Sadece "Yanıt Süreçleri Üzerine Temellenmiş Kanıt"	2	0.53	0	0	2	0.53
	Sadece "İç Yapı Üzerine Temellenmiş Kanıt"	8	2.14	15	4.01	23	6.15
	Sadece "Diğer Değişkenlerle İlişkiler Üzerine Temellenmiş Kanıt"	4	1.07	6	1.6	10	2.67
	Sadece "Test Etme Sonuçları Üzerine Temellenmiş Kanıt"	0	0.00	0	0.00	0	0.00
	"Test İçeriği Üzerine Temellenmiş Kanıt" ve "Yanıt Süreçleri Üzerine Temellenmiş Kanıt"	2	0.53	19	5.08	21	5.61
	"Test İçeriği Üzerine Temellenmiş Kanıt" ve "İç Yapı Üzerine Temellenmiş Kanıt"	35	9.36	79	21.12	114	30.48
	"Test İçeriği Üzerine Temellenmiş Kanıt", "İç Yapı Üzerine Temellenmiş Kanıt", "Diğer Değişkenlerle İlişkiler Üzerine Temellenmiş Kanıt"	10	2.67	18	4.81	28	7.49
	"Test İçeriği Üzerine Temellenmiş Kanıt", "Yanıt Süreçleri Üzerine Temellenmiş Kanıt" ve "İç Yapı Üzerine Temellenmiş Kanıt"	4	1.07	10	2.67	14	3.74
	Sadece "Uzman Görüşü"	23	6.15	135	36.10	158	42.25
	Sadece "Korelasyon Analizi"	8	2.14	9	2.41	17	4.55
Sadece "Faktör Analizi"	5	1.34	7	1.87	12	3.21	
Sadece "Test Madde Analizi"	3	0.80	10	2.67	13	3.48	
Geçerlik Kanıt Yöntemi	"Uzman Görüşü" ve "Test Madde Analizi"	23	6.15	64	17.11	87	23.26
	"Uzman Görüşü" ve "Faktör Analizi"	13	3.48	20	5.35	33	8.82
	"Uzman Görüşü" ve "Korelasyon Analizi"	4	1.07	21	5.61	25	6.68
	"Uzman Görüşü", "Test Madde Analizi" ve "Faktör Analizi"	10	2.67	19	5.08	29	7.75

Geçerlik çalışması yapılmış bütün lisansüstü tezlerde geçerlik kanıtları rapor edilmiştir. Geçerlik kanıtı rapor edilmeyen herhangi bir teze rastlanmamıştır. Sadece uzman görüşü alınsa bile “geçerlik kanıtı rapor edilmiş” olarak değerlendirildiği göz önünde bulundurulmalıdır.

Geçerlik çalışması yapılmış hiçbir lisansüstü tezde geçerlik derecesi ifade edilip yorumlanmamıştır.

Geçerlik çalışması yapılmış toplam 374 lisansüstü tezin 169’u (%45.19) için sadece kapsam geçerliği, 129’u (%34.49) için kapsam geçerliği ve yapı geçerliği, 44’ü (%11.76) için sadece yapı geçerliği, 16’sı (%4.28) için kapsam geçerliği-ölçüt ilişkili geçerlik ve yapı geçerliği türleri kullanılmıştır. Doktora tezleri incelendiğinde 89 tezin 26’sı (%6.95) için sadece kapsam geçerliği, 37’si (%9.89) için kapsam geçerliği ve yapı geçerliği, 13’ü (%3.48) için sadece yapı geçerliği, 8’i (%2.14) için kapsam geçerliği-ölçüt ilişkili geçerlik ve yapı geçerliği, 5’i (%1.34) için ise kapsam geçerliği ve ölçüt ilişkili geçerlik türleri kullanıldığı tespit edilmiştir. Yüksek lisans tezlerinde ise 285 tezin 143’ü (%38.24) için sadece kapsam geçerliği, 92’si (%24.60) için kapsam geçerliği ve yapı geçerliği, 31’i (%8.29) için sadece yapı geçerliği, 11’i (%2.94) için ise kapsam geçerliği ve ölçüt ilişkili geçerlik, 8’i (%2.14) için ise kapsam geçerliği-ölçüt ilişkili geçerlik ve yapı geçerliği türlerinin kullanıldığı tespit edilmiştir. Lisansüstü tezlerinde en çok kullanılan geçerlik türlerinin kapsam ve yapı geçerliği olduğu görülmektedir. Lisansüstü tezlerin neredeyse yarısında sadece kapsam geçerliği kullanılmıştır.

Geçerlik çalışması yapılmış ve geçerlik kanıtı rapor edilmiş toplam 374 lisansüstü tezin 162’si (%43.32) için sadece test içeriği üzerine temellenmiş kanıt, 114’ü (%30.48) için test içeriği üzerine temellenmiş kanıt ve iç yapı üzerine temellenmiş kanıt kaynağı kullanıldığı tespit edilmiştir. 42 (%11.23) tezde 3 kanıt kaynağı, 135 (%36.09) tezde 2 kanıt kaynağı birden kullanılmıştır. 197 (%52.60) tezde ise sadece bir kanıt kaynağı kullanılmıştır. En çok kullanılan kaynağın test içeriği üzerine temellenmiş kanıt, en az kullanılan kaynağın ise yanıt süreçleri üzerine temellenmiş kanıt olduğu görülmektedir. Test etmenin sonuçları üzerine temellenmiş kanıt kaynağının kullanıldığı lisansüstü tezle karşılaşılmamıştır. Tezlerin yarısından fazlasında sadece tek bir kanıt kaynağı kullanılmış olması düşündürücüdür.

Geçerlik çalışması yapılmış ve geçerlik kanıt yöntemi rapor edilmiş 374 lisansüstü tezin 158’i (%42.25) için sadece uzman görüşü, 87’si (%23.26) için uzman görüşü ve test madde analizi, 33’ü (%8.82) için uzman görüşü ve

faktör analizi, 29'u (%7.75) için uzman görüşü-test madde analizi ve faktör analizi kullanıldığı tespit edilmiştir. 29 (%7.75) tezde 3 kanıt yöntemi, 145 (%37.76) tezde 2 kanıt yöntemi birden kullanılmıştır. 200 (%53.49) tezde ise sadece bir kanıt yöntemi kullanılmıştır. Doktora tezleri incelendiğinde 89 tezin 23'er tanesi için (%6.15) sadece uzman görüşü ve uzman görüşü-test madde analizi kullanılmıştır. Yüksek lisans tezleri incelendiğinde 285 tezin 135'inde (%36.10) sadece uzman görüşü, 64'ünde (%17.11) uzman görüşü ve test madde analizi kullanılmıştır. En çok kullanılan yöntemin uzman görüşü, en az kullanılan yöntemin ise korelasyon analizi yöntemi olduğu görülmektedir. Tezlerin yarısından fazlasında sadece tek bir geçerlik kanıtı yöntemi kullanıldığı söylenebilir.

Tartışma ve Sonuç

Araştırma kapsamına alınan 435 tezin %81'inde (351 tez) güvenirlik çalışması rapor edilmiş, kalan %19'unda (84 tez) ise güvenirlik çalışmalarının rapor edilmediği tespit edilmiştir. Korkmaz (2010) çalışmasında araştırmamız sonuçlarına benzer şekilde her 5 tezin 1'inde; Kutanis vd. (2015) incelenen tezlerin %26'sında; Ulutaş vd. (2015) inceledikleri makalelerde kullanılan ölçeklerin %61'inde güvenirlik çalışması yapılmadığını belirlemişlerdir.

Araştırma sonucunda matematik eğitimi alanında tamamlanmış 435 lisansüstü tezin %86'sında (374 tez) geçerlik çalışması rapor edilmiştir. Ancak %14'ü (62 tez) gibi büyük bir oranda geçerlik çalışması rapor edilmemiştir. Ergun ve Cilingir (2013) çalışmalarında da inceledikleri tezlerin yaklaşık %28'inde geçerlik çalışmasının yapılmadığını belirtmişlerdir. Kutanis vd. (2015) çalışmalarında inceledikleri tezlerin %57'sinde; Ulutaş vd. (2015) incelenen makalelerde kullanılan ölçeklerin %60'ında geçerlik çalışmasının yapılmadığını tespit etmişlerdir. Geçerlik çalışması yapılmama oranı gerek çalışmamız gerekse diğer çalışmalar için benzerlik göstermese de lisansüstü tezlerde önemli sayılabilecek oranda geçerlik çalışmasının yapılmadığını söylemek mümkündür.

Lisansüstü tezlerde güvenirlik ve geçerlik çalışması rapor edilme durumu birlikte değerlendirildiğinde, 435 tezin %67'sinde hem güvenirlik hem geçerlik çalışması, %18'inde geçerlik çalışması yapılmaksızın sadece güvenirlik çalışması, %10'unda güvenirlik çalışması yapılmaksızın sadece geçerlik çalışması yapılmıştır. Geri kalan %5'inde ise ne güvenirlik ne de geçerlik

çalışmasına yer verilmiştir. Kabaca ve Erdoğan'nun (2007) çalışmalarında araştırma sonuçlarına benzer şekilde neredeyse incelenen her 4 tezin 1'inde geçerlik-güvenirlik çalışması; her 10 tezin 3'ünde ise hem geçerlik hem de güvenilirlik çalışması yapılmadığı belirlenmiştir. Altinkurt (2007) çalışmasında, ölçme aracı kullanılan tezlerin yarısına yakınında ölçme aracına ilişkin geçerlik ve güvenilirlik kanıtlarının sunulmadığını tespit etmiştir. Tez yazarları yaptıkları çalışmaların benzerlerini incelediklerinde güvenilirlik ve geçerlik çalışmaları ne düzeyde yapıldıysa bir benzerini kendi çalışmalarında uyguladıkları düşünülmektedir. Buradan hareketle lisansüstü öğrencilerin ölçme yöntemini daha önce tamamlanan araştırmalardan öğrenme eğiliminde oldukları; ölçme yöntemindeki gelişmeleri takip etmedikleri; geçerlik ve güvenilirlik çalışmalarının önemini fark edemedikleri söylenebilir.

Diğer taraftan matematik eğitimi alanında tamamlanmış lisansüstü tezlerin %10'unda güvenilirlik çalışması yapılmaksızın sadece geçerlik çalışmasının yapılmasının sebebi "ölçümlerden elde edilen yorumlar geçerli olduğunda ölçümler güvenilir olur" düşüncesinden kaynaklanmış olabilir.

Araştırmamız sonuçlarına göre tezlerin %5'inde hem güvenilirlik hem de geçerlik çalışmasına rastlanmamıştır. Bu çalışmaların tamamının yüksek lisans tezleri olduğu tespit edilmiştir. Benzer şekilde Evrekli vd. (2011) çalışmalarında incelenen tezlerin yaklaşık %3'ünde geçerlik ve güvenilirlik çalışmalarının rapor edilmediğini belirtmişlerdir. Bu durum yüksek lisans çalışmalarının lisansüstü çalışmaların ilk aşaması olması sebebiyle araştırmacıların güvenilirlik ve geçerlik incelemelerini tam olarak kavrayamadıklarına işaret sayılabilir. Bir başka ifadeyle, araştırmacıların lisans ve lisansüstü programlarda yeterince test, anket ya da ölçek hazırlama süreci için teorik bilgiye sahip olmaması veya bilgiye sahip olsa bile uygulamaya dönük bilgisinin olmaması sebebiyle güvenilirlik ve geçerlik çalışmalarına gereken önemi vermemiş olduklarını düşündürmektedir.

Araştırma kapsamına alınan 435 tezin yaklaşık %22'sinde (95 tez) güvenilirlik doğrulturması yapıldığı tespit edilmiştir. Araştırmada kullanılan testi uygulamadan testin güvenilirliğinden bahsetmek veya bir başkasının hesapladığı bir teste ait ölçümlerin güvenilirlik katsayısını hatalı bir kabul ile testin kendi özelliği gibi kabul edip ve yine yanlış bir biçimde ve bilinçsizce bir kabul ile ölçümlerin ve onlardan elde edilen güvenilirlik katsayısının değişmeyeceğini kabul ederek uygulama yapmaksızın bir test hakkında güvenilir-

lik kararına varılmış olması güvenirlilik doğuşturmalarının yapıldığı anlamına gelmektedir (Bademci, 2007). Güvenirlilik yalnızca testin kendisinin bir fonksiyonu değil, örneklemin de bir fonksiyonudur (Aktaran: Bademci, 2007). Her 5 tezin 1'inde test ölçümlerinin güvenilir olduğu göz ardı edilerek daha önce yapılmış çalışmalardaki güvenirlilik çalışmalarından elde edilen sonuçlar hiç bir güvenilirlik çalışması yapılmadan rapor edilmiştir. Araştırma bulgularına benzer şekilde Ergun ve Cilingir'in (2013) çalışmalarında incelenen tezlerin %17'sinde güvenirliliği ve geçerliği belirlenmiş ölçme aracını hiç bir geçerlik ve güvenirlilik çalışması yapmadan kendi örneklemi üzerine uyguladıklarını belirlemişlerdir. Ulutaş vd. (2015) çalışmalarında ise tezlerin %14'ünde güvenirlilik değeri için daha önce alan yazındaki değerlerin uygulama yapmadan verildiğini, yani güvenirlilik doğuşturma yapıldığını belirlemişlerdir. Araştırma sonuçlarından hareketle daha önceki araştırmalarda rapor edilmiş güvenirlilik ve geçerlik kanıtlarının kendi araştırmalarında değiştirmeden kullandığı söylenebilir. Ayrıca "Testler güvenilir değildir, güvenilir olan ölçümlerdir" düşüncesini göz önünde bulundurduğumuzda başka bir örneklem grubuna uygulanmış ölçeğin ya da testin güvenirlilik sonuçlarının doğrudan kullanılması uygun değildir. Çünkü ölçekten ya da testten elden ettiğimiz ölçümlerin güvenirlilik değerleri gruptan gruba göre değişebilir. Matematik eğitimi alanında tamamlanmış lisansüstü her 10 tezin 2'sinde güvenirlilik doğuşturmalarının yapıldığı görülmektedir. Bu durum lisansüstü öğrencilerin ölçme ve değerlendirmedeki yenilikleri takip etmesinden kaynaklanıyor olabilir. Tezlerde güvenirlilik doğuşturma yapılmayıp, hangi grup için bir test yapılacaksa testi uygulayıp test ölçümlerini yorumlayarak güvenirlilik kanıtları sunulmalıdır.

Araştırmamız bağlamında güvenirlilik çalışması yapılmış tezlerdeki güvenirlilik kestirim yöntemi/yöntemleri tezlerin %84'ünde rapor edilmiştir. Bir araştırmada kullanılmış olan bütün yöntemler ayrıntıları ile rapor edilmesi gerekmektedir. Crocker ve Algina (1986), test geliştiricisi ya da uygulayıcısı belirli koşullar altında yapılan testin ölçüm güvenirliliğini incelemekle kalmayıp, bu bilgiyi potansiyel kullanıcılara aktarmalıdır ifadesi ile benzer bir vurgu yapmıştır (Korkmaz, 2010: 49). Bademci (2010, 2011a) araştırmalarda ölçüm güvenirlilik kestirim yönteminin mutlaka rapor edilmesi gerektiğinin önemini vurgulamıştır. Ancak bu araştırmada 5'i doktora, 51'i yüksek lisans tezi olmak üzere toplam 56 tezde (%16) ölçüm güvenirliliğine yönelik kestirim yönteminin rapor edilmediği tespit edilmiştir.

Kestirim yöntemi/yöntemleri rapor edilmiş tezlerin %34'ünde (101 tez) sadece Cronbach Alfa, %15'inde (45 tez) Cronbach Alfa ve Kuder-Richardson, %13'ünde (38 tez) sadece Kuder-Richardson, %11'inde (33 tez) yarıya bölme ve Cronbach Alfa yöntemleri kullanılmıştır. Araştırmada en çok kullanılan yöntemlerin Cronbach Alfa ve Kuder-Richardson, en az kullanılan yöntemlerin ise paralel form, test-tekrar test ve yarıya bölme yöntemleri olduğu belirlenmiştir. Araştırmamız sonuçlarına benzer şekilde Kabaca ve Erdoğan'ın (2007), Korkmaz'ın (2010), Delice ve Ergene'nin (2015), Kutanis vd. 'nin (2015) ve Ulutaş vd. 'nin (2015) çalışmalarında da en çok kullanılan yöntemin Cronbach Alfa olduğu tespit edilmiştir. Yine benzer şekilde Soycan ve Babacan'ın (2019) çalışmalarında incelenen tezlerde güvenilirlik kanıtı için yarıya bölme, Kuder-Richardson ve Cronbach Alfa yöntemlerinin kullanıldığı, Şenyurt ve Özkan'ın (2017) çalışmalarında ise en çok kullanılan güvenilirlik kestirim yönteminin Cronbach Alfa ve Kuder-Richardson olduğu belirlenmiştir.

Cronbach Alfa ve Kuder-Richardson yöntemlerinin çok kullanılmış olması tek uygulama gerektiren yöntemlerden olması ve istatistik programlarından rahatça hesaplanabilmesinden kaynaklanmış olabileceği düşünülmektedir. Ayrıca nicel araştırmalarda genellikle 1-0 ile puanlanan testler ve likert tipi gibi çoklu puanlamaya sahip ölçeklerin kullanılması bu yöntemlerin ağırlıklı olarak kullanılmış olmasının nedeni olarak gösterilebilir. Paralel formlar yönteminin az kullanılmasının sebebi de eşit güçlükte ve yapıda iki test formunun oluşturulmasının daha fazla zaman alması olabilir. Test-tekrar test yöntemi de iki uygulama gerektirdiği için az tercih edilmiş olabilir. Yarıya bölme yöntemi tek uygulama gerektiren bir yöntem olmasına rağmen düzeltme formülünün kullanılması gerekliliğinden kaynaklı Cronbach Alfa ve Kuder-Richardson yöntemlerine nazaran daha az kullanılmış olduğu söylenebilir.

Ayrıca araştırmamız sonucunda güvenilirlik kestirimine yönelik kullanılan yöntemler birlikte değerlendirildiğinde, sadece 1 yöntem kullanan tezlerin sayısının 145 (%49.16), 2 yöntemi birlikte kullanılan tezlerin sayısının 118 (%40), 3 yöntem kullanan tezlerin sayısının ise 32 (%10.84) olduğu belirlenmiştir.

Güvenirlik, sınavı alanların belirli bir grubu için bir test üzerindeki ölçümlerin bir özelliği olarak tanımlanmıştır (Crocker ve Algina, 1986). Bir başka ifadeyle güvenilirlik, test sonuçlarının bir özelliğidir. "Test güvenilir-

dir” veya “testin güvenirliliği” ya da “aracın güvenirliliği” gibi ifade biçimleri doğru değildir. Çünkü güvenirlik, “testlerin değil ölçümlerin bir özelliğidir.” ve ifade edilmesi gereken de ölçüm güvenirliliğidir (Bademci, 2007). “Testin güvenirliliği” ifadesi, araştırmaya dahil edilmiş tezlerin %51’inde kullanılmıştır. Kavram yanlılığı çağırışımı olarak belirtilen bu durum, incelenen tezlerin yarısından fazlasında görülmektedir. Korkmaz’ın (2010) çalışmasında ise incelenen tezlerin %79’unda aynı durumun olduğu belirlenmiştir. Otorite olarak kabul edilen (Sireci, 2007) American Educational Research Association, American Psychological Association ve National Council on Measurement in Education [AERA, APA ve NCME] tarafından 1999 ve 2014 yıllarında yayınlanan Eğitimsel ve Psikolojik Test Etme/Test Yapma Standartları’nda (*Standards for Educational and Psychological Testing* [EPTS]) “testin güvenirliliği” ifadesinin “kabul edilemez” olduğu açıkça ifade edilmesine ve aradan yıllar geçmesine rağmen, matematik eğitimi alanında tamamlanmış ve araştırma kapsamına alınmış, lisansüstü tezlerin %51’inde “testin güvenirliliği” ifadesinin görülmüş olması, matematik eğitimi alanındaki lisansüstü tezlerin bu anlamda yetersiz olduğunu göstermektedir. Ayrıca tezlerin %18’inde “Testin Güvenirliliği” yanında “Aracın Bir Özelliği”, %13’ünde “Test Güvenirliliği” yanında “Ölçümlerin Güvenirliliği” ifadesinin kullanıldığı tespit edilmiştir. Tezlerin %15’inde güvenirlik ifadesine yer verilmemiş, %3’ünde ise farklı anlamlara gelen güvenirlik ifadeleri kullanılmıştır. Güvenirlik ifadelerindeki sonuçlardan hareketle ölçme ve yontembilimindeki çağdaş standartları lisansüstü öğrencilerin takip etmedikleri söylenebilir.

Güvenirlik çalışması yapmış tezlerin %90’ında (315 tez) güvenirlik katsayısı rapor edilmiştir. Çağdaş standartlara göre güvenirlik katsayısı rapor edilen tezlerin %35’inde katsayı yeterli düzeydeyken (0.80 ve 0.80’den fazla), %65’inde yeterli düzeyde (0.80’den az) değildir. Ancak, güvenirlik katsayılarının, güvenirlik çalışması için alınan örneklem büyüklüğü ile birlikte değerlendirilmesi, yani her ikisinin de istenilen düzeyde olması gerekliliği unutulmamalıdır (Bademci, 2007). Araştırma bulgularına benzer şekilde Şimşek vd. (2007) çalışmalarında 64 doktora tezinden yalnızca birkaçında yüksek düzeyde güvenirlik katsayısının rapor edildiği; çoğu çalışmada ise orta düzeyde güvenirlik katsayısının yeterli görüldüğü belirlenmiştir. Korkmaz’ın (2010) çalışmasında ise incelenen tezlerin %63’ünde güvenirlik kestiriminde kullanılan güvenirlik katsayısının 0.80’den düşük olduğu belir-

lenmiştir. Araştırma bulgularının aksine Delice ve Ergene'nin (2015) çalışmasında incelenen makalelerden güvenilirlik katsayısı rapor edilmiş makalelerin yaklaşık %85'inde güvenilirlik katsayısının 0.80'den küçük olmadığı tespit edilmiştir.

Güvenirlik katsayısı rapor edilen tezlerin %96'sında (303 tez) güvenilirlik kat sayısının güven aralıkları rapor edilmemiştir. EPTS (1999, 2014) güvenilirlik katsayısının güven aralıklarının rapor edilmesi gerektiğini ortaya koymasına rağmen araştırmalarda bu durum göz ardı edilmiştir. Sadece tezlerin %4'ünde güvenilirlik kat sayısının güven aralıkları rapor edilmiştir. Güven aralığının dar ya da geniş olması ölçüm güvenilirlik katsayısının isabetli olma durumunu ortaya koyacağı için önemlidir. Araştırmamız sonuçlarına göre araştırmacıların neredeyse tamamına yakını güvenilirlik katsayısı hesaplamasında alınan güven aralığını raporlamadığı için güvenilirlik katsayılarının daha isabetli ve duyarlı olup olmadığı konusunda yorum yapılamadığı söylenebilir. Temel araştırma amaçları için, EPTS (1999, 2014) güvenilirlik katsayısının güven aralıklarının rapor edilmesi gerektiğini ortaya koymuştur. Bademci (2011b) çalışmasında örneklem büyüklüğünün artırılması ile, örneğin güvenilirlik katsayısı 0.90 için güven aralığı genişliğinin daralacağını, böylece daha büyük örneklem için daha duyarlı, daha güvenilir kestirime ulaşabileceğini kanıtlamıştır. Dolayısıyla, örneklem büyüklüğü, güvenilirlik katsayısı ve güven aralığını bir bütün olarak düşünmek gerekmektedir. Güven aralığının %99 veya %95 olarak ifade edilmiş olması test ölçümlerinin güvenilirliği hakkında daha net bir güvenilirlik kanıtı koyulmasını sağlar. Araştırma kapsamına alınan lisansüstü tezlerde bu kanıtı sunup güvenilirlik katsayısının güven aralıklarını rapor eden çok az tezle karşılaşmış olması dikkat çekicidir.

Güvenirlik kestirimi için alınan örneklem büyüklüğü tezlerin %94'ünde (331 tez) rapor edilmiş olup, %6'sında (20 tez) güvenilirlik kestirimine yönelik alınmış örneklem büyüklüğü rapor edilmemiştir. 331 tezin 295'inde (%89) örneklem büyüklüğü 400'den az olarak rapor edilmiştir. Benzer şekilde Korkmaz'ın (2010) çalışmasında incelediği tezlerin %76'sında güvenilirlik kestirimine yönelik alınmış olan örneklem büyüklüğünün 400'den az olduğu belirlenmiştir. Araştırmamız bulgularına benzer şekilde Ulutaş vd. (2015) çalışmalarında incelenen makalelerin %1'inde örneklem büyüklüğünün rapor edilmediği; rapor edilenlerin %87'sinde örneklem büyüklüğünün 200 ve altında; %11'inde örneklem büyüklüğünün 400 ve üzerinde olduğunu

belirlemişlerdir. Ölçüm güvenirliliğinin gruptan gruba göre değiştiği göz önünde bulundurulduğunda, örneklem büyüklüğü ölçüm güvenirliliğinin duyarlı kestirimi konusunda çok önemli bir faktördür (Bademci, 2007). Güvenirlik çalışmaları için örneklem büyüklüğü kestiriminde kullanılan örneklemdeki kişi sayısı için farklı görüşler mevcut olsa da EPTS’de ölçüm güvenirliliğinin kestirildiği hallerde büyük örneklem kullanılması, güvenirliliğin daha duyarlı kestirimi için önemli ve yararlı olduğunu, örneklem büyüklüğünün en az 400 denek olması gerektiğini vurgulamıştır. EPTS göz önüne alındığında lisansüstü tezlerin yaklaşık %90’ının çağdaş standartlara güvenirlik kestirimine yönelik alınmış olan örneklem büyüklüğü açısından uygun olmadığı söylenebilir. Başka bir ifadeyle, örneklem büyüklüğü rapor edilmiş her 10 tezin sadece 1’inin çağdaş standartlara güvenirlik kestiriminde kullanılan örneklem büyüklüğü açısından daha uygun olduğu ifade edilebilir.

Geçerlik çalışması yapılmış olan tezlerin %45’inde (169 tez) sadece kapsam geçerliği, %35’inde (129 tez) hem kapsam hem de yapı geçerliği, %12’sinde (44 tez) sadece yapı geçerliği, %5’inde (16 tez) hem kapsam hem de ölçüt ilişkili geçerlik, yine %5’inde (16 tez) hem kapsam hem yapı hem de ölçüt ilişkili geçerlik türleri kanıtları sunulmuştur. Tezlerde sadece belirtke tabloları ya da uzman görüşü ile yalnız kapsam geçerliğinin kanıtlarının sunulmuş olması matematik eğitimi alanında tamamlanmış tezlerin geçerlik kanıtları anlamında yetersiz olduğu şeklinde yorumlanabilir. En çok kullanılan geçerlik türü kapsam geçerliği ve yapı geçerliğidir. Araştırmamız bulgularının aksine Kabaca ve Erdoğan’ın (2007) çalışmasında incelenen tezlerde yapı geçerliği ve ölçüt ilişkili geçerlik türlerine rastlanmadığı belirtilmiştir. Ergun ve Cilingir (2013) ise incelenen tezlerin %21’inde kapsam geçerliği, %9’unda yapı geçerliği ve %2’sinde ölçüt ilişkili geçerlik türü kullanıldığını tespit etmişlerdir. Kutanis vd. (2015) çalışmalarında incelenen tezlerin %5’inde kapsam geçerliği, %10’unda yapı geçerliği türü kullanıldığını tespit etmişlerdir. Ulutaş vd. (2015) çalışmalarında ise geçerlik çalışması yapılan tezlerin %45’inde kapsam geçerliği, %10’unda yapı geçerliği türü kullanıldığını; araştırmamız bulgularının aksine tezlerin %51’inde geçerlik türünden bahsedilmediğini tespit etmişlerdir. Geçerlik test yapma standartlarında kapsam geçerliği, ölçüt ilişkili geçerlik ve yapı geçerliği şeklindeki bölünmüş geleneksel bakış açısı eleştirilmiş, bunun yerine geçerlik, çeşitli geçerlik kanıtlarına dayalı bütüncül bir kavram olarak ele alınmıştır (Ba-

demci, 2017a). Bu bağlamda lisansüstü tezlerin yaklaşık %39’unda 2 geçerlik türü, yaklaşık % 4’ünde 3 geçerlik türü birden kullanıldığı göz önünde bulundurulduğunda lisansüstü tezlerin % 43’ünün diğerlerine nazaran geçerlik türündeki kanıt sunumu açısından çağdaş standartlara daha uygun olduğu söylenebilir.

“Testin geçerliği” ifadesi, araştırmaya dâhil edilmiş tezlerin %76’sında kullanılmıştır. Kavram yanılıgısı çağrışımı olarak belirtilen bu durum, incelenen tezlerin dörtte üçünde görülmektedir. Tezlerin %5’inde “Test Ölçümlerinin Geçerliği”, %2’sinde “Testin Geçerliği” yanında “Test Ölçümlerinden Yapılan Yorumun Özelliği” ifadesinin kullanıldığı tespit edilmiştir. Tezlerin %13’ünde geçerlik ifadesine yer verilmemiş, %4’ünde ise farklı anlamlara gelen geçerlik ifadeleri kullanılmıştır. Geçerlik, bir testin ya da ölçme aracının kendisinin bir özelliği değildir. “Test geçerlidir” veya “testin geçerliği” ,“aracın geçerliği” ya da “test ölçümlerinin geçerliği” veya “ölçümler geçerlidir” gibi ifade biçimleri doğru değildir. Çünkü geçerlik, test ölçümlerine bağımlı yorumların bir özelliğidir. Daha doğru ifade biçimi “test ölçümlerinden yapılmış yorumlar geçerlidir” veya “ölçümlerden yapılmış kullanım ve yorumlar geçerlidir” şeklinde olmalıdır (Bademci, 2011a, 2017a). Test ölçümlerinin geçerliği, ölçümlerin geçerliği ve testin geçerliği ifadeleri hatalı olduğu düşünüldüğünde tezlerin %81’inde yani her 5 tezin 4’ünde geçerlik hakkında hatalı ifadeler olduğu söylenebilir.

Geçerlik çalışması yapılmış ve geçerlik kanıtı rapor edilmiş tezlerin %43’ünde sadece test içeriği üzerine temellenmiş kanıt, %30’unda test içeriği ve iç yapı üzerine temellenmiş kanıt, %7’sinde test içeriği üzerine temellenmiş kanıt, iç yapı üzerine temellenmiş kanıt ve diğer değişkenlerle ilişkiler üzerine temellenmiş kanıt, %6’sında ise sadece iç yapı üzerine temellenmiş kanıt kaynakları kullanılmıştır. 1999 ve 2014 test yapma standartlarında, test ölçümlerinin önerilen yorumunun geçerliğinin değerlendirilmesinde test içeriği üzerine temellenmiş kanıt, yanıt süreçleri üzerine temellenmiş kanıt, iç yapı üzerine temellenmiş kanıt, diğer değişkenlerle ilişkiler üzerine temellenmiş kanıt ve test etmenin sonuçları üzerine temellenmiş kanıt kaynaklarının geçerlik kanıtı olarak sunulmasının gerekliliğinden bahsedilmiştir (AERA, APA ve NCME, 2014). Bu çerçevede, geçerlik kanıtlarından birden fazlasını kullanmayı önermiş olsa da bu durum kural olarak söz konusu değildir. Ancak belli kullanımlar doğrultusunda önerilen yorum için birden çok kanıt kaynağı kullanmak gerekmektedir (Bademci, 2017b). Ma-

tematik eğitimi alanında tamamlanmış tezlerin yaklaşık %11'inde 3 kanıt kaynağı, yaklaşık %36'sında ise 2 kanıt kaynağının birden kullanılmış olmasından dolayı tezlerin yaklaşık %47'sinin çağdaş standartlara daha uygun olduğu söylenebilir. EPTS'de geçerlik kanıtlarının 5'inin de rapor edilmesi önerilmesine rağmen araştırmamız kapsamında incelenen tezlerin hiç birinde 5 kanıt kaynağının birden kullanılmadığı tespit edilmiştir. Lisansüstü tezlerin yaklaşık %53'ünde ise sadece bir kanıt kaynağı kullanılmıştır. Tezlerin çoğunluğunda test içeriği üzerine temellenmiş kanıt kaynağının kullanılmasının sebebinin araştırmacıların uzman görüşü almasının ve uzlaşma çalışmalarını yapmasının daha hızlı ve kolay olmasından kaynaklanmış olabileceği düşünülmektedir.

Çağdaş standartlarda test ölçümlerinin önerilen bir yorumu uygun kanıtla desteklendiği takdirde yüksek geçerliğe, yeterli gerekçeye dayanmışsa veya haklı çıkmamışsa düşük geçerliğe sahip olduğu; geçerliğin varlık ya da yokluk şeklinde değil, derece ile ifade edilmesi belirtilmiştir (Bademci, 2017a). Ancak incelediğimiz matematik eğitimi lisansüstü tezlerinde geçerlik derece olarak yorumlanmamıştır.

Geçerlik çalışması yapılmış ve geçerlik kanıtı rapor edilmiş olan tezlerin %42'sinde sadece uzman görüşü, %23'ünde hem uzman görüşü hem de test-madde analizi, %9'unda uzman görüşü ve faktör analizi, %8'inde uzman görüşü, test-madde analizi ve faktör analizi yöntemleri birlikte kullanılmıştır. Uzman görüşüne bir çok tezde başvurulmuş olması çalışmalarda daha hızlı ve kolay bir dönüt almak için tercih edilmiş olabilir, ancak geçerlik kanıtı için sadece uzman görüşü alınmış olması yeterli değildir. Araştırma sonuçlarına benzer şekilde Kabaca ve Erdoğan (2007) çalışmalarında incelenen doktora tezlerinin geçerlik kanıtlarının sadece uzman görüşü ile sunulduğunu belirlemişlerdir. Kutanis vd. (2015) çalışmalarında incelenen tezlerin %78'inde faktör analizi, %4'ünde uzman görüşü kullanıldığını belirlemişlerdir. Geçerlik çalışmalarında faktör analizinin ve uzman görüşünün diğer yöntemlere göre daha çok kullanılmasının sebebi bu yöntemlerin diğer yöntemlere göre geçerlik çalışmalarında daha ön planda olduğu gösterilebilir.

Öneriler

Araştırmadan elde edilen sonuçlar doğrultusunda, araştırmacılara, akademik kadrolarda bulunanlara ve lisans-lisansüstü öğrencilere yönelik olarak, bazı öneriler aşağıda sunulmuştur.

- Mevcut ve eski standartlar yerine, EPTS'nin 1999 ve 2014'te önerdiği çağdaş ve yeni standartlar Türk eğitim ve bilimi içerisinde hızla yerleştirilmeli ve böylelikle, Türk eğitimi ve bilimin ilerlemesine ve kalitesinin artmasına katkıda bulunulmalıdır.
- Geçerlik ve güvenilirlikteki kavram yanılığısı çağrışımı bilimsel düşüncenin önünde önemli bir engel olarak görülmektedir. Kavram yanılığarı, çalışmaların bilimsel niteliğini derinden etkilemektedir. Dolayısıyla ölçme ve değerlendirme derslerinde geçerlik ve güvenilirlik konusunda kavram yanılığlarının giderilmesine yönelik önlemler alınmalıdır.
- Danışmanlar öğrencilerini ölçme ve ölçek geliştirme süreci ile ilgili çalışmalarını incelemeye veya ölçme ve ölçek geliştirme sürecini uygulamaya yönlendirerek daha çok anlamlandırmalarını sağlayabilir.
- Araştırmalarda kullanılan ölçme araçlarının geçerlik ve güvenilirlik kanıtlarının sunulup sunulmadığını, enstitülerde bulunan tez inceleme biriminin daha dikkatli biçimde kontrol etmesi için çalışmalar yapılabilir. Bundan sonra yapılacak olan araştırmalara yönelik öneriler ise aşağıda yer almaktadır.
- Bu araştırma matematik eğitimi alanında yapılan tezler üzerinde yürütülmüştür. Üniversite sayısı genişletilerek ya da alan değişikliği (fen bilimleri eğitimi, temel eğitim vb.) ile benzer çalışmalar yapılabilir.
- Bu araştırmada yüksek lisans ve doktora tezleri ele alınmıştır; benzer çalışmalar eğitim bilimleri alanında veya matematik eğitimi alanlarında yazılmış olan makaleler üzerinde de yapılabilir.
- Benzer bir araştırma yıl aralığı genişletilerek gerçekleştirilebilir.

EXTENDED ABSTRACT

**Investigation of Validity and Reliability Studies in
Mathematics Education Graduate Theses**

*

Sami Sezer Arbağ – Erhan Ertekin

Ministry of National Education, Necmettin Erbakan University

This study is on the use of the concept of validity and reliability of post-graduate theses completed in the field of mathematics education and contemporary standards of validity and reliability. With this research, the validity and reliability of doctoral and master theses completed in the field of mathematics education were examined in accordance with contemporary standards, around errors. By adopting a qualitative research approach, it is aimed to prevent the repetition of the same mistakes in future scientific researches in this research. Focusing on validity and reliability studies of doctoral and master's theses, especially completed in mathematics education, made it necessary to conduct the research with a case study design. The sample of the study is a total of 435 graduate theses, 92 doctoral theses and 343 master's theses, written in the field of mathematics education in the 2000-2018 research and have no access barriers on the website of YÖK National Thesis Center. The analysis and analysis of the doctoral and master's theses were collected in this research with the "Thesis Examination Form" according to the research under the categories for the research. Descriptive and categorical analyzes were examined in another Excel program, which was scored and obtained by different evaluation experts, expressing different scoring and reliability expressing reliability. Credibility and transferability for validity of research results; For the reliability of the results, the consistency and verifiability evidence was examined.

Theses, differences in the use of validity and reliability concepts, measurement reliability estimation methods, reliability constructions, measurement reliability coefficients, sample sizes used in reliability studies, confidence intervals of measurement reliability coefficients, validity and reliability studies, validity type used, validity proofs, validity proof. It has been evaluated in terms of compliance with contemporary standards in validity

and reliability, taking into account various dimensions such as the methods used.

According to the results of the research, it was seen that 19% of the theses did not report reliability and 14% of them did not report validity. When the reliability and validity study reporting situations are examined together, both reliability and validity studies were not conducted in 67% of the theses, only a reliability study was not conducted in 18%, and only a validity study was not done in 10%. In the remaining 5%, neither reliability nor validity studies were included.

The reliability estimation method/methods in the theses whose reliability study was conducted was reported in the majority of the theses (84%). Estimation method/methods were reported for 34% of theses, using only Cronbach Alfa, 15% by Cronbach Alfa and Kuder Richardson, 13% only by Kuder Richardson, 11% by halving and Cronbach Alpha methods. It has been observed that the most used methods are Cronbach Alfa and Kuder Richardson.

It has been determined that 22% of the theses are built for reliability, in other words, the values in the literature are reported without a reliability study of a scale whose reliability has been determined in the literature.

It was determined that 51% of the theses used "Reliability of the Test", 18% "Reliability of the Test" as well as "A Feature of the Vehicle", 13% "Test Reliability" as well as "Reliability of Measurements". In addition, 15% did not include reliability statements, and 3% used reliability statements with different meanings. It is not the tests that are reliable, but the measurements obtained from the tests. According to this situation, it was seen that there was a misconception connotation in the sense of reliability statement in more than half of the theses. When the validity statements were examined, the expression "Validity of the test" was used in 76% of the theses included in the study. In 5% of the theses, it was determined that "Validity of Test Measurements", "Test Validity" in 2% of the theses, and "Feature of the Interpretation from Test Measurements" were found. Validity statement was not included in 13% of the theses, and validity expressions with different meanings were used in 4%. Considering that it is the characteristic of the interpretation made from the test measurements, not the valid tests, it has been observed that there is a misconception connotation in the sense of validity statement in 81% of the theses.

The reliability coefficient has been reported in 90% of the theses whose reliability study has been done. While the reliability coefficient is sufficient (more than 0.80 and 0.80) in 35% of these theses whose reliability coefficient is reported, it is not sufficient (less than 0.80) in 65%. In 96% of the theses whose reliability coefficient was reported, the confidence intervals of the reliability coefficient were not reported. Confidence intervals of the reliability coefficient were reported only in 4% of the theses. The sample size taken for the reliability estimation was reported in 94% of the theses, it was observed that the sample size was reported as less than 400 (insufficient) in 89%, and the sample size as 400 and more (sufficient) in 11%. Considering contemporary standards, it can be stated that 65% of theses are insufficient in terms of reliability coefficient and 89% of theses are insufficient in terms of sample size.

In 45% of theses whose validity study was made, only content validity, 34% had both content and construct validity, 12% had only construct validity, 5% had both scope and criterion-related validity, and again 5% had both evidence of both the structure and criterion-related validity types were presented. It has been determined that the most used types of validity are content and construct validity.

In 43% of theses whose validity study was conducted and proof of validity was reported, only evidence based on test content was used, in 30% test content and sources of evidence based on internal structure were used. Considering that 5 of the validity proofs in EPTS standards suggest to be reported, 5 sources of evidence were not used in any of the theses completed in the field of mathematics education. In 11% of the theses, 3 sources of evidence were used and in 36% 2 sources of evidence were used at once. In 53% of the theses, only one source of evidence was used.

Of the theses whose validity study was conducted and proof of validity was reported, 42% had only expert opinion, 23% both expert opinion and test-item analysis, 9% expert opinion and factor analysis, 8% expert opinion, test-Item analysis and factor analysis methods were used together. It was seen that the most used methods to present the validity proof were expert opinion and test-item analysis.

Kaynakça / References

- AERA, APA ve NCME [American Educational Research Association, American Psychological Association ve National Council on Measurement in Education] (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- AERA, APA ve NCME [American Educational Research Association, American Psychological Association ve National Council on Measurement in Education] (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Alkan, G. (2014). Türkiye’de muhasebe alanında yapılan lisansüstü tez çalışmaları üzerine bir araştırma (1984-2012). *Muhasebe ve Finansman Dergisi*, 61, 41-52.
- Altınkurt, L. (2007). Sanat eğitimi ile ilgili lisansüstü tezlerin incelenmesi. *Milli Eğitim Dergisi*, 173, 105-113.
- Arslan, A., Geriş, A. ve Kulaksız, T. (2016). Eğitsel Yazılım Değerlendirme Formlarının ve Ölçeklerinin İncelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 7(2), 272-284.
- Bademci, V. (2004). ‘Testin güvenilirliği’ veya ‘Test güvenilirlidir’ diye ifade etmek doğru değildir. *Türk Eğitim Bilimleri Dergisi*, 2(3), 367-372.
- Bademci, V. (2005a). Araştırmalarda ölçme ile ilgili bazı büyük hataları düzeltmek ve bir reformu başlatmak: Güvenirlik, testlerin bir özelliği değildir. *Eğitim Fakültelerinde Yeniden Yapılandırmanın Sonuçları ve Öğretmen Yetiştirme Sempozyumu*, 22-23-24 Eylül, Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara.
- Bademci, V. (2005b). Testler güvenilir değildir: Ölçüm güvenilirliğine yeterli dikkat ve güvenilirlik çalışmaları için örneklem büyüklüğü. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, 17, 33-45.
- Bademci, V. (2006a). Güvenirliği doğru anlamak ve bazı klişeleri yıkmak: Bilinenlerin aksine, Cronbach’ın alfa katsayısı, negatif ve -1’den küçük olabilir. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 7(12), 3-26.
- Bademci, V. (2006b). Tartışmayı sonlandırmak: Cronbach’ın alfa katsayısı, iki değerli [0,1] ölçümlenmiş maddeler ile kullanılabilir. *Kazım Karabekir Eğitim Fakültesi Dergisi*, 13, 438-446.
- Bademci, V. (2007). *Ölçme ve araştırma yöntem bilminde paradigma değişikliği: Testler güvenilir değildir*, Ankara: Yenyap Yayınları.
- Bademci, V. (2008). Araştırmalarda ölçme ile ilgili bazı büyük hataları düzeltmek ve eğitimde yeniden yapılanmayı sürdürmek: Güvenirlik, testlerin bir özelliği değildir. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, 22, 50-69.

- Bademci, V. (2010). Türk eğitim ve biliminde paradigma değişikliği: Testler veya ölçekler güvenilir ve geçerli değildir. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dekanlığı, 26 Nisan, G.Ü. Gazi Eğitim Fakültesi Resim-İş Eğitimi Anabilim Dalı Konferans Salonu, Ankara.*
- Bademci, V. (2011a). Türk eğitim ve biliminde paradigma değişikliği: Testler veya ölçekler güvenilir ve geçerli değildir. *Dicle Üniversitesi Eğitim Fakültesi Dergisi, 16, 116-132.*
- Bademci, V. (2011b). Kuder-Richardson 20, Cronbach Alfa, Hoyt'un varyans analizi, genellenirlik kuramı ve ölçüm güvenirliliği üzerine bir çalışma. *Dicle Üniversitesi Eğitim Fakültesi Dergisi, 17, 173-193.*
- Bademci, V. (2017a). Ölçme ve araştırma yöntembiliminde çağdaş gelişmeler ve yeni standartlar 1: Geçerlik, ölçümlerin kullanımlarının ve önerilen yorumlarının bir özelliğidir. *Eğitim ve Toplum Araştırmaları Dergisi, 4(1), 63-80.*
- Bademci, V. (2017b). Ölçme ve araştırma yöntembiliminde çağdaş gelişmeler ve yeni standartlar 2: Geçerlikte üçleme (kapsam, ölçüt ilişkili ve yapı geçerlikleri) öğretisinin reddi ve geçerlik kanıtının kaynakları. *Eğitim ve Toplum Araştırmaları Dergisi, 4(1), 81-97.*
- Benligiray, S. (2012). Araştırmacıların bankacılık alanına bilimsel katkıları: Bankacılık konusunda yapılan lisansüstü tezlere ilişkin bir araştırma. *Süleyman Demirel Üniversitesi Vizyoner Dergisi, 3(6), 26-46.*
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., ve Demirel, F. (2008). *Bilimsel Araştırma Yöntemleri* (17. Baskı). Ankara: Pegem Yayınları.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson Education.
- Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Crocker, L., ve Algina, J. (1986). *Introduction to classical and modern test theory*, Ford Worth: Holt, Rinehart and Winston, Inc.
- Çepni, S. (2014). *Araştırma ve proje çalışmalarına giriş*. (7. Baskı). Trabzon: Celepler Matbaacılık.
- Delice, A., ve Ergene, Ö. (2015). Ölçek geliştirme ve uyarlama çalışmalarının incelenmesi: Matematik eğitimi makaleleri örneği. *Karaelmas Eğitim Bilimleri Dergisi, 3, 60-75.*
- Dunkin, M. J. (1996). Types of errors in synthesizing research in education. *Review of Educational Research, 66(2), 87-97.*

- Ergun, M. ve Cilingir, F. (2013). İlköğretim Bölümünde yapılan lisansüstü tezlerin incelenmesi: Ondokuz Mayıs Üniversitesi örneği. Sözel bildiri, *VI. Ulusal Lisansüstü Eğitim Sempozyumu*, Sakarya Üniversitesi, Sakarya.
- Evrekli, E., İnel, D., Deniz, H., ve Balım, A. G. (2011). Fen Eğitimi alanındaki lisansüstü tezlerdeki yöntemsel ve istatistiksel sorunlar. *İlköğretim Online*, 10(1), 206-218.
- İnce, M., Gül, H. ve Bozyiğit, S. (2017). Türkiye’de turizm pazarlaması konusunda yazılan lisansüstü tezlerin içerik analizi yöntemi ile incelenmesi. *MANAS Sosyal Araştırmalar Dergisi*, 6(5), 113-130.
- Kabaca, T., ve Erdoğan, Y. (2007). Fen bilimleri, bilgisayar ve matematik eğitimi alanlarındaki tez çalışmalarının istatistiksel açıdan incelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 22 (2), 54-63.
- Kaya, Ç., Yazıcı, A. Ş., Delivelı, K., ve Hoşgörür, V. (2016). Türkiye’de eğitim denetimi alanında yapılan lisansüstü çalışmalarının değerlendirilmesi. *Adnan Menderes Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 7(1), 38-51.
- Korkmaz, A. (2010). *Vahit Bademci’nin paradigma değişikliği üzerine bir araştırma: Testler değil, ölçümler güvenilirdir*. Yayımlanmamış yüksek lisans tezi. Sosyal Bilimler Enstitüsü, Zonguldak.
- Kutanis Özen, R., Özsoy, E., Karakiraz, A., Aras, M., Erol, E., ve Uslu, O. (2015). Örgüt kültürü çalışmalarının yöntem ve kapsam bakımından incelenmesi: Lisansüstü tezler üzerinden bir inceleme. *Yönetim ve Ekonomi Araştırmaları Dergisi*, 13(1), 123-142.
- Lincoln, Y. S., ve Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Miles, M. B., ve Huberman, A. M. (2015). *Nitel veri analizi* (S. Akbaba-Altun ve A. Ersoy Çev. Ed.). Ankara: Pegem Akademi.
- Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., ve Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Council for Exiceptional Children*, 71(2), 137-148.
- Onwuegbuzie, A. J., ve Daniel L. G. (2003). Typology of analytical and interpretational errors in quantitative and qualitative educational research. *Current Issues in Education*, 6(2).
- Reynolds, C. R., Livingston, R. B., ve Willson, V. (2009). *Measurement and assessment in education*. Upper Saddle River, New Jersey: Pearson Education International.
- Sireci, S. G. (2007) On validity theory and test validation, *Educational Research*, 36(8), 477-481.

- Soycan, M., ve Babacan, E. (2019). Müziksel işitme, okuma ve yazma ile ilgili geliştirilmiş ölçme araçlarının incelenmesi: İçerik analizi çalışması. *Elektronik Sosyal Bilimler Dergisi*,18(69), 343-353.
- Şenyurt, S., ve Özer Özkan, Y. (2017). Eğitimde ölçme ve değerlendirme alanında yapılan yüksek lisans tezlerinin tematik ve metodolojik açıdan incelenmesi. *İlköğretim Online*,16(2), 628-653.
- Şimşek, A., Özdamar, N., Becit, G., Kılıçer, K., Akbulut, Y., ve Yıldırım, Y. (2007, Mayıs). Türkiye'deki eğitim teknolojisi araştırmalarında güncel eğilimler. Sözel bildiri, *I. Uluslararası Bilgisayar ve Öğretim Teknolojileri Sempozyumu*, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale.
- Tavşancıl, E. (2008). Lisansüstü tez çalışmalarının yöntem bölümünün içerik çözümlenmesi. Yazılı bildiri, *I. Ulusal Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi* Ankara Üniversitesi, Ankara.
- Ulutaş, B., Üner, S., Turan Oluk, N., Yalçın Çelik, A., ve Akkuş, H. (2015). Türkiye'deki Kimya eğitimi makalelerinin incelenmesi: 2000-2013. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 16(2),141-160.
- Yıldırım, A. ve Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri*. (9. Baskı). Ankara: SeçkinYayımları.

Kaynakça Bilgisi / Citation Information

Arbağ, S. S. ve Ertekin, E. (2020). Matematik eğitimi lisansüstü tezlerindeki geçerlik ve güvenirlik çalışmalarının incelenmesi. *OPUS–Uluslararası Toplum Araştırmaları Dergisi*, 16(31), 3924-3957. DOI: 10.26466/opus.735675