

Reliability Generalization Study of General Self-Efficacy Scale

Genel Öz Yeterlik Ölçeğinin Güvenirlik Genellemesi Çalışması

ABSTRACT

The aim of this study is to examine the meta-analytic reliability generalization of the General Self-Efficacy Scale (GSS) short form. It was also examined how different moderator variables affect the reliability coefficients. A set of criteria was determined to determine the studies to be included in the meta-analysis. 101 Cronbach's Alpha coefficients obtained from 87 studies selected according to the criteria were included in the meta-analysis. Analyses were conducted with the Random Effects Model using the Bonett Transform. Reliability generalization based on meta-analysis was used in the analysis of the data. The effect of moderator variables on the variability in reliability estimates as effect size was examined with Analog ANOVA. As a result of the research, it was determined that the mean alpha for general studies was .85. Moderator variable analyses were conducted with three categorical variables to explore the variability in Cronbach's Alpha values and the results were reported. According to the moderator variable analysis, it was determined that the studies conducted in the field of education in the groups where the scale was used differed significantly from the reliability values of the studies conducted in other fields. According to the results obtained from the research; it was determined that the Cronbach Alpha value of the scale was at a sufficient level and suggestions were made for other researchers in line with the findings.

Keywords: Reliability generalization, meta-analysis, self-efficacy, moderator

Öz

Bu çalışmanın amacı, Genel Öz Yeterlik Ölçeği (General Self-Efficacy Scale-GSS) kısa formunun meta-analitik güvenilirlik genellemesinin incelenmesidir. Ayrıca farklı moderatör değişkenlerin güvenilirlik katsayılarını nasıl etkilediği incelenmiştir. Meta-analize dâhil edilecek çalışmaların belirlenmesi için birtakım kriterler belirlenmiştir. Kriterlere göre seçilen 87 çalışmadan elde edilen 101 Cronbach Alfa katsayısından meta-analize dâhil edilmiştir. Bonett dönüşümü kullanılarak Rastgele etkiler modeli ile analizler yürütülmüştür. Verilerin analizinde meta-analize dayalı güvenilirlik genellemesi kullanılmıştır. Etki büyüklüğü olarak güvenilirlik tahminlerindeki değişkenlik üzerinde moderatör değişkenlerin etkisi Analog ANOVA ile incelenmiştir. Araştırma sonucunda genel çalışmalar için Cronbach Alfa katsayı ortalamasının ,85 olduğu belirlenmiştir. Cronbach Alfa değerlerindeki değişkenliği keşfetmek için üç kategorik değişken ile moderatör değişken analizleri yürütülmüştür ve sonuçlar raporlanmıştır. Moderatör değişken analizine göre ölçeğin kullanıldığı gruplarda eğitim alanında yapılan çalışmaların diğer alanda yapılan çalışmaların güvenilirlik değerlerine göre anlamlı şekilde farklılaştığı belirlenmiştir. Araştırmadan elde sonuçlara göre; ölçeğin Cronbach Alfa değerinin yeterli düzeyde olduğu belirlenmiş ve bulgular doğrultusunda diğer araştırmacılara yönelik önerilerde bulunulmuştur.

Anahtar Kelimeler: Güvenirlik genellemesi, meta-analiz, öz yeterlik, moderatör

Muammer
YÜKSEL¹



Milli Eğitim Bakanlığı, Manisa, Türkiye



Geliş Tarihi/Received 05.04.2025
Kabul Tarihi/Accepted 09.05.2025
Yayın Tarihi/Publication Date 31.08.2025

Sorumlu Yazar/Corresponding author:
Muammer YÜKSEL

E-mail: muammer.8845@gmail.com

Cite this article: Yüksel, M. (2025). Reliability Generalization Study of General Self-Efficacy Scale. *Journal of Psychometric Research*, 3(2), 128-135.



Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Giriş

Bandura'ya (1997) göre öz yeterlik; bireyin yapabileceklerine ve yapamayacaklarına ilişkin inancı olarak ifade edilmektedir. Snyder ve Lopez'e (2002) göre ise öz yeterlik, sadece bireyin kendi yetenek ya da yapabileceklerine olan inançları değil; problem çözme, karşılaşılan durumlar ile başa çıkabilmesi hakkında kendisine olan inançlarıdır. Ayrıca öz yeterlik, bireylerin erişmek istedikleri başarı kavramlarına ilişkin süreçleri düzenleme ve bu başarı düzeyine erişmek için yeteneklerine ilişkin inançları olarak da ifade edilebilir (Bandura, 1997). Öz yeterliğin temelinde yatan, bireyin kendisinden beklenen görevi gerçekleştirebilme durumuna yönelik sürece aradığı cevaplardır. Bireye ait olan öz yeterlik inancı bir işi yapma güdüsü ve bireyin başarılı olması ile de ilişkilidir. Birey kendini başarısız olarak gördüğünde, aynı performansı sergilemek için gerekli olan güdülenme düzeyi azalabilir (Dilekli, 2015). Bireyin hissettiği bu olumsuz duygu süreci kendisine ve yapabileceklerine yönelik inancını da etkiler. Ayrıca yeterlik düzeyleri kişilerin düşünme biçimlerini, duygusal tepkilerini ve problem çözme becerilerini de etkilemektedir. Öz yeterlik düzeyi düşük olan bireyler, karşılaşacakları olayların görüldüğünden de zor olduğunu düşünerek gözlerinde büyütürler ve karşılaştıkları problemlere çözüm bulmakta zorlanırlar (Schunk ve Greene, 2017).

Öğrencilerin öz yeterlik düzeyleri matematik problemi çözebileceklerine, kitap okuyabileceklerine, bisiklete binebileceklerine veya ayakkabılarını bağlayabileceklerine yönelik inançları ile ilgilidir. Öz yeterlik de, tıpkı kendi yeterlilik algılarının veya benlik kavramı inançlarının benzer inançları yansıtması gibi bireyin bu faaliyetleri yapıp yapamayacağına dair bazı yargıları içerir. Alanyazında öğrencilerin akademik başarı düzeyleri üzerindeki öz yeterlik etkisine yönelik çeşitli çalışmalar yapılmıştır (Koca ve Dadanlı, 2019; Polat, 2017; Yılmaz, Yiğit ve Kaşarçı, 2012). Ayrıca öz yeterlik öğrencilerin akademik erteleme, okulda kalma kararı gibi pek çok psikolojik ve öğrenme süreçleriyle de ilişkilidir (Caprara vd., 2008; Klassen vd., 2008). Öz yeterlik, bireyin amacını gerçekleştirmek için kendine duyduğu güven ve bu amacı gerçekleştirmek için yapılması gerekenleri yapabileceğine dair algısıdır. Bu algının derecesi bireyin göstereceği çabayı etkileyecektir ve bireyin en verimli performansı sergileyebilmesi için öz-yeterlik algısının yüksek olması önemlidir (Bandura, 1982). Öz yeterliğin bireyin başarı, duygusal durumu, iyi oluş hali üzerindeki etkileri göz önünde bulundurulduğunda; öz yeterlik düzeylerinin öğrencilerin öğrenme süreçlerindeki etkilerinin derinlemesine incelenmesi önemlidir.

Bireylerin öz yeterlik düzeylerinin belirlenmesi için araştırmacılar tarafından pek çok ölçme aracı geliştirilmiştir. Bu ölçme araçlarından biri de Jerusalem ve Schwarzer tarafından 1981 yılında geliştirilmiş olan "Genel Öz Yeterlilik Ölçeği" (General Perceived Self-Efficacy Scale - GSS)'dir. GSS ölçeği 10 maddeden oluşmaktadır. Ölçek ilk olarak Almanya'da yirmi yıl boyunca süren araştırmalar sonucunda geliştirilmiş ve sonrasında 28 dile uyarlanmıştır (Schwarzer ve Jerusalem, 1995). Ölçek dörtlü Likert tipinde geliştirilmiştir. Ölçekte yer alan maddelerden bireylerin alabilecekleri en düşük puan 10 ve en yüksek puan 40 olarak hesaplanmaktadır. Ölçekten alınan yüksek puanların bireylerin genel öz yeterlik düzeylerinin yüksek olduğunu yansıttığı varsayılmaktadır (Scholz vd., 2002). GSS ölçeğinin yüksek güvenilirliği, kararlılığı ve yapı geçerliliği daha önceki çalışmalarda doğrulanmıştır (Leganger vd., 2000; Schwarzer vd., 1999). Ölçeğin 28 ülke arasında yapısal olarak eşdeğer olduğu bulunmuştur ve yalnızca tek bir küresel boyut oluşturmaktadır (Scholz vd., 2002). Geliştirilen bu ölçek dünya çapında pek çok dile uyarlanmış ve üzerinde farklı ülkelerde pek çok araştırma gerçekleştirilmiştir (Brislin, 1970).

Ölçme sonuçlarının tekrarlanabilirliği olarak tanımlanan güvenilirlik (Crocker ve Algina, 1986), psikoloji ve eğitimde ölçme sonuçlarının yorumlanmasında kilit noktadır. Cronbach Alfa iç tutarlılık katsayısı Klasik Test Teorisine dayalı çalışmalarda sıklıkla kullanılan bir yöntemdir. Güvenirlik örneklemden örnekleme farklılık gösterebileceği için her çalışmada yeniden tahmin edilmesi ve raporlanması gerekir (Thompson ve Vacha-Haase, 2000). Klasik Test Teorisinde güvenilirlik hatayla ilişkilidir; hata ne kadar düşük olursa güvenilirlik o kadar yüksek olur. Güvenirlik Genelmesi (RG) ile bir ölçme aracının farklı çalışmalardan elde edilen güvenilirlik tahminleri bir araya getirilerek ortalama bir güvenilirlik katsayısı elde edilir (Vacha-Haase, 1998). Tekli çalışmalardan elde edilen bulguların istatistiksel olarak birleştirilmesi olan meta-analizin (Glass, 1976) bir uygulaması olan RG ile farklı çalışmalardan elde edilen güvenilirlik katsayıları birleştirilerek ortalama bir güvenilirlik katsayısı elde edilir. Başka bir açıdan güvenilirlik genelmesi, bir ölçüm aracının veya testin farklı bağlamlarda, gruplarda veya durumlarda ne kadar güvenilir olduğunu ifade eder. Yani, bir testin sonuçlarının, belirli bir örneklem veya koşul dışında da benzer sonuçlar verip vermediğini değerlendirmeyi amaçlar. Bu kavram, özellikle psikolojik testler ve ölçümler için önemlidir, çünkü bir testin sadece belirli bir grup üzerinde geçerli olması değil, daha geniş bir kitleye de uygulanabilir olması beklenir. Güvenirlik genelmesi, testin tutarlılığını ve geçerliliğini artırmak için kritik bir süreçtir. Bu yöntem güvenilirlik değerindeki

dalgalanmaları ve varyansları açıklamak için en uygun yöntemdir (Sánchez-Meca vd., 2013). Alanyazın incelendiğinde, kullanılan pek çok ölçek için RG çalışmasının yapıldığı belirlenmiştir (Özdemir vd., 2020; Şen, 2022; Aslan vd., 2022; Dilek ve Gelbal, 2023; López-Ibáñez, López-Nicolás vd., 2024). Son yıllarda eğitim, sağlık, psikoloji ve sosyoloji gibi alanlarda farklı çalışmalarda kullanılan ölçeklerin güvenilirlik genellemesi çalışmalarının yapıldığı görülmüştür. Özellikle çalışmaların örneklemelerinin farklı büyüklüklerde olması, farklı ülkelerde yapılması, farklı düzeylerde uygulanması gibi durumların ölçeklerin örneklemelerden kestirilen güvenilirlik değerlerinin farklılaşmasına neden olabilir. Araştırmacılar bu açıdan kullanılan ölçekler üzerinde RG çalışması yapmakta ve sonuçlarını alanyazına sunarak sonraki çalışmalarda kullanılırken dikkat edilecek özellikleri sunmaktadırlar.

Öz yeterlik, eğitim başta olmak üzere psikoloji, sosyoloji, sağlık, yönetim gibi farklı alanlarda yapılan çalışmalarda sıkça ele alına bir psikolojik değişkendir. Öz yeterlik bireylerin yapabileceklerine olan inancı olarak ele alındığında farklı alanlardaki performans, yönelim ve duyuşsal düzeylerini etkileyen bir değişken haline gelmektedir. Bu açıdan bakıldığında öz yeterlik pek çok farklı alanda incelenen bir kavram olduğu için bu kavramı ölçmeyi amaçlayan bir ölçme aracı olarak ele alınan GSS üzerine yapılan bu çalışma, alana ölçeğin güvenilirliği ve farklı değişkenlere göre değişimi adına önemli katkılar sunması amaçlanmıştır.

Çalışma kapsamında ele alınan GSS ölçeğinin farklı alanlarda ve farklı örneklemeler üzerinde uygulandığı pek çok araştırmaya mevcuttur. Bu çalışma; farklı durumlarda ortaya çıkan güvenilirlik değerlerinin araştırmacılar tarafından yorumlama sürecinde karmaşaya neden olmaması, daha önce bu ölçeğe ait güvenilirlik genellemesi çalışmasının yapılmamış olması ve ileride yapılabilecek çalışmalar için araştırmacılara fikirler sunması açısından önem taşımaktadır.

Bu çalışma GSS'den elde edilen puanların güvenilirliğini genelleştirmeyi amaçlamaktadır. Bu ölçek, dünya genelinde bireylerin öz yeterliği üzerine yapılan çalışmalarda en yaygın kullanılan ölçeklerden biri olmuştur ve bu nedenle seçilmiştir. GSS ile birçok farklı çalışma yapılmış olmasına rağmen güvenilirliğin genelleştirilmesine odaklanan bir çalışmaya rastlanmamıştır. Bu nedenle bu çalışmada aşağıdaki sorulara yanıt aranacaktır:

- 1) GSS' den elde edilen puanların güvenilirlik düzeyi nedir?
- 2) GSS belirlenen moderatör değişkenlere göre güvenilirlik düzeyleri farklılık göstermekte midir?

Yöntem

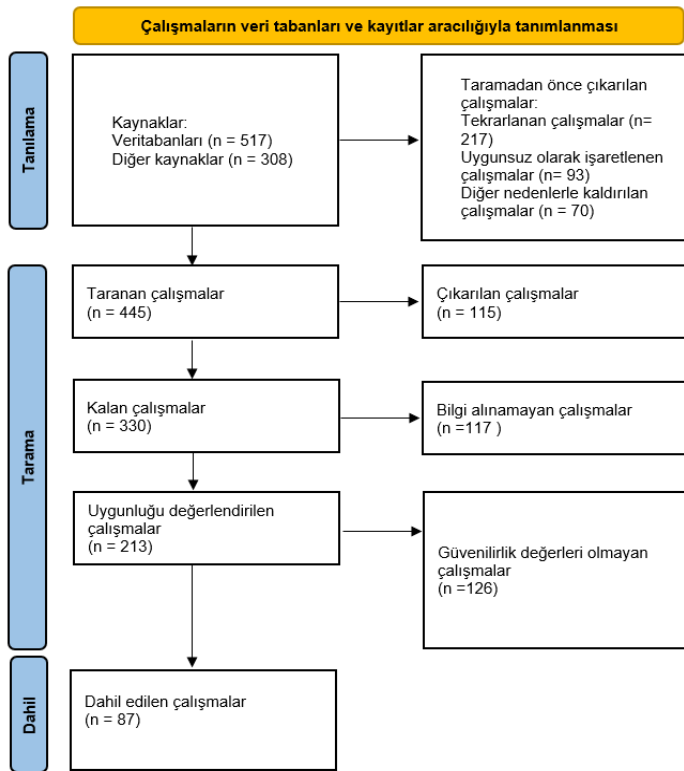
Araştırma Deseni

Araştırmada nicel araştırma yaklaşımı yürütülen meta analiz yöntemi kullanılmıştır. Meta-analiz, bağımsız bireysel çalışmalardan elde edilen bulguların istatistiksel bir kombinasyonunu sağlar (Glass, 1976). Güvenirlik genellemesi, bildirilen güvenilirlik katsayılarının ortalama bir tahminini elde etmek ve varyasyon kaynaklarını belirlemek için kullanılan istatistiksel bir yöntemdir (Vacha-Haase, 1998). Bu güvenilirlik genelleme çalışmasında bireysel çalışmaların araştırmaya dâhil edilme kriterleri incelenerek uygun olan çalışmalar dâhil edilmiştir.

Dâhil Edilme Kriterleri ve Verilerin Kodlanması

Genel Öz Yeterlik Ölçeğinin (Çapa vd., 2005) güvenilirliğine ilişkin meta-analize dâhil edilecek çalışmalar, Google Akademik, Ulusal Tez Merkezi, Ulakbim veri tabanlarında herhangi bir yıl kriteri olmaksızın taranmış ve 825 çalışmaya ulaşılmıştır. Çalışma, Türkçe "Genel Öz Yeterlik" anahtar kelimesi ve bu anahtar kelimenin İngilizce karşılığı olan "General Self-Efficacy" kelimeleri ile araştırılarak gerçekleştirilmiştir. Çalışmalar dâhil edilme kriterlerine göre analiz edilmiş olup, bu kriterler; Cronbach Alfa katsayısının bildirilmiş olması ve ölçeğin çalışmalarda madde sayısı ile kullanılmış olması ve çalışmanın dilidir. Son durumda, ölçeğin tamamı için bu kriterleri karşılayan 87 çalışma bulunmaktadır ve raporlanan Cronbach Alfa değeri 101'dir. Çalışma sayısı ile rapor edilen Cronbach Alfa değeri arasındaki fark, bazı çalışmaların birden fazla örnekleme birden fazla Cronbach Alfa katsayısı rapor etmesidir. Verilerin kodlanma sürecine ait sürecin PRISMA akış şeması Şekil 1'de verilmiştir.

Şekil 1
PRISMA Akış Şeması



Şekil 1’de çalışmaların kodlama sürecine ait akış verilmiştir. Olası moderatör değişkenler belirlenerek kodlama işlemi gerçekleştirilmiştir. Veriler iç tutarlılık katsayısı, yayın türü, uygulama grubu, örneklem büyüklüğü ve Likert tipine göre kodlanmıştır. Veriler daha sonra R programına aktarıldı.

Etki Büyüklüğünün Hesaplanması ve Diğer İstatistiksel İşlemler

Bu çalışmada GSS'in ortalama güvenilirlik katsayısının elde edilmesi amaçlanmaktadır. Bu amaçla tüm çalışmalarda bildirilen iç tutarlılık katsayısı olan Cronbach Alfa katsayısı kullanılmıştır. Güvenirlik genelleme çalışmalarında, değer 0'dan uzaklaştıkça çarpık bir dağılım gösterdiği için Cronbach Alfa değeri dönüştürülerek analizler yapılmaktadır (Field, 2005). Daha sonra etki büyüklüğü hesaplamasında hangi modelin seçileceğine karar verilmiştir. Bu modeller sabit etkiler modeli ve rastgele etkiler modelidir. Sabit etkiler modelinde etki büyüklüğünü etkileyen faktörlerin tüm bireysel çalışmalarda aynı olduğu varsayılmaktadır ancak eğitim gibi bir alanda etki büyüklüğünü etkileyen faktörlerin çok farklı olduğu bilinmektedir. Bu nedenle rastgele etkiler modelinin kullanılması gerekmektedir (Borenstein vd., 2009). Bu çalışmada rastgele etkiler modeli kullanılmıştır. Heterojenlik testi için Cochran'ın Q istatistiği ve bu

istatistiğin devamı sayılabilecek I^2 istatistiği kullanılmaktadır. Anlamlı bulunan bir Q istatistiği ve %75'in üzerinde bir I^2 istatistiği (Higgins vd., 2003) heterojenliğin kanıtıdır. Daha sonra karma etki modeli kullanılarak moderatör değişken analizleri yapılmıştır. Moderatör değişkenlerin istatistiksel anlamlılığını test etmek için ANOVA analizi ve Q-between (Q_B) istatistiği kullanılmıştır. Analizler R yazılımı ile yapılmıştır.

Yayın Yanlılığı

Yayın yanlılığı meta-analiz sonuçlarını etkiler. Yayın yanlılığının olası nedenleri, kriterlerin belirli bir sonuca göre belirlenmesi, literatür taramasının yetersiz olması ve dâhil edilen çalışmaların bilimsel kanıtlarının yetersiz olması olabilir (Dinçer, 2021). Yayın yanlılığının belirlenmesi sonuçların güvenilirliğine katkı sağlayacaktır. Bu amaçla huni grafiği (Borenstein vd., 2009) ve Rosenthal'in güvenli N yöntemi (Rosenthal, 1979) yayın yanlılığının olup olmadığını belirlemek için kullanılmıştır.

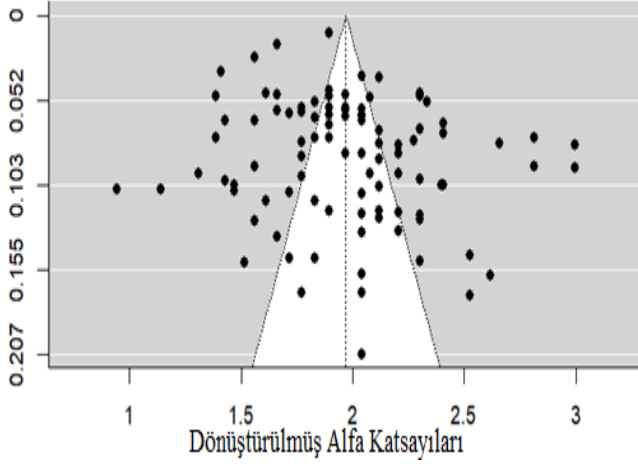
Bulgular

Gerçekleştirilen çalışmada öncelikle huni grafiği ve Rosenthal'in (1979) güvenli N değeri kullanarak yayın yanlılığını incelenmiştir. Araştırmadaki veriler için huni grafiği incelendiğinde, araştırmaya dâhil edilen çalışmaların dönüştürülmüş Cronbach Alfa güvenilirlik değerlerinin dağılımlarının simetrik yapıda olduğu belirlenmiştir. Huni grafiğinin incelenmesi öznellik içeren bir süreç olduğu için simetrikliğin bozulacağı araştırma sayısının kestirilmesine ilişkin güvenli N "fail-safe N" değerinin incelenmesi gerekmektedir (Davis ve Rothstein, 2006). Yayın yanlılığı olasılığı, güvenli N değerinin 1.620.230 olduğu belirlenmiş ve bu değer birleşik iki kuyruklu p değerinin .05'i aşması için 11.620.230 farklı çalışmayı bulup dâhil etmemiz gerektiğini göstermektedir. Bu değer meta-analiz sonuçlarını geçersiz kılmak için eklenmesi gereken yaklaşık çalışma sayısını göstermektedir. Belirlenen güvenli N değeri, istatistik değerinin anlamsız olması için araştırmaya eklenmesi gereken minimum ters yönlü çalışma sayısını ifade etmektedir. Ayrıca Rosenthal (1991), çalışma sayısına (k) göre araştırmadan elde edilecek sonuçların geçersiz olabilmesi için yayınlanamamış olan çalışma sayısını (N) hesapladığı $5k+10$ değerinin N değerinden küçük olması beklenir. Çalışmada hesaplanan güvenli N değeri, $5k+10$ 'dan (yani 515) daha büyüktür; bu da yayınlanmamış olan bu araştırmanın genel güvenilirlik tahmininin, mevcut meta-analizin birleşik katsayı alfasını değerini etkilemesinin muhtemel olmadığını göstermektedir. Bu sonuçlara göre, eklenmesi gereken çalışma sayısının oldukça yüksek olduğu

görülmüştür. Diğer yaklaşım olan huni grafiği asimetrisi Şekil 2’de görülmektedir.

Şekil 2

Güvenirlilik Değerlerine Ait Huni Grafiği



Araştırma kapsamında belirlenen bu değerler anlamlı çıkan p değerinin anlamsız çıkması için eklenmesi gereken en az çalışma sayısını ifade etmektedir. Yayın yanlılığı için gerçekleştirilen iki farklı yaklaşımla da bulunan değerlerin anlamsız çıkması için eklenmesi gereken çalışma sayısının çok fazla olması ve huni grafiğinin dağılımı göz önüne alındığında çalışmanın yayın yanlılığı için dayanıklı olduğu ve çalışmada yayın yanlılığı olmadığı şeklinde yorumlanabilir (Bakioğlu ve Göktaş, 2018). Ayrıca yayın yanlılığı incelemesine ilişkin gerçekleştirilen bir başka inceleme olarak Kendall Tau b değeri=.13 hesaplanmıştır. Belirlenen bu değer istatistiksel olarak anlamlı çıkmaması da çalışmada yayın yanlılığı bulunmadığına bir ek kanıt sunmaktadır.

Araştırmacılar tarafından öncelikle güvenirlilik katsayıları üzerinde dönüşümlerin uygulanıp uygulanmayacağına ve uygulanacak ise hangi dönüşümün tercih edileceğine karar verilmelidir. Bugüne kadarki çoğu RG çalışması, bazı

araştırmacılar (Bachner ve O'Rourke, 2007) analiz etmiştir (Şen ve Yıldırım, 2020; Thompson ve Vacha-Haase, 2000) tarafından önerildiği gibi dönüştürülmemiş alfa katsayıları üzerinden gerçekleştirilmiştir. Diğer bir açıdan elde edilen güvenirlilik değerlerinin dağılımını normalleştirmek veya varyansları stabilize etmek için de katsayılar dönüştürülerek kullanılmaktadır (Sawilowsky, 2000). Başlangıçta Fisher z dönüşümleri kullanılsa da Cronbach Alfa katsayıları için Fisher z'den daha iyi olan dönüşümler (Hakstian-Whalen ve Bonett dönüşümleri gibi) mevcuttur. Bunlardan Hakstian-Whalen dönüşümü, güvenirlilik katsayılarının dağılımını normalleştirmemizi sağlar. Bonett (2002) tarafından önerilen alfa katsayılarının dağılımını normalleştirmemizi ve varyanslarını sabitlememizi sağladığından, daha iyi bir alfa katsayı dönüşümünü sunmuştur. Bu dönüşümler katsayıların standartlaştırılmasını ve güvenirlilik değerlerinin karşılaştırılmasını sağlamaktadır. Bu çalışmada çarpık Cronbach Alfa değerlerini normalleştiren ve varyans değerlerini sabitleyen Bonett dönüşümü kullanılmıştır.

Ölçekten elde edilen toplam puanlar için Bonett (2002) dönüşüm yöntemi kullanılarak hesaplanan ortalama güvenirlilik katsayısının ,86 (%95: ,85-,87) olduğu ve bu değer istatistiksel olarak anlamlı olduğu belirlenmiştir ($p<,05$). Bu ortalama değer ölçeğin başlangıç değerine yakındır (Çapa vd., 2005). Q istatistiği de anlamlıydı ($Q_{100} = 2644.238$, $p<,05$) ve I^2 istatistiği 97,53 idi. Bu değerler heterojenliğin yüksek olduğunu göstermektedir.

Moderatör Analizi

Analizler sonucunda heterojenliğin nedenini belirlemek amacıyla moderatör analizleri yapılmıştır. Belirlenen moderatör değişkenler üç tanedir: uygulanan grup, Likert tipi (4-kategori ve diğer), uygulanan alan (eğitim ve diğer). Moderatör değişken analizleri Tablo 1’de gösterilmektedir.

Tablo 1

Ölçeğe Ait Moderatör Değişkenlere Yönelik Analiz Sonuçları

| Değişken | Kategori | k | α | %95 _{alt} | %95 _{üst} | Q_B | Sd | p |
|----------------|-----------------|-----|----------|--------------------|--------------------|-------|----|------|
| Seçenek Sayısı | 4 | 87 | ,87 | ,84 | ,89 | 1,95 | 1 | ,16 |
| | Diğer (5 ve 7) | 14 | ,85 | ,84 | ,86 | | | |
| Uygulanan Grup | Öğrenci Olmayan | 45 | ,86 | ,78 | ,94 | 4,58 | 1 | ,10 |
| | Öğrenci | 49 | ,84 | ,83 | ,85 | | | |
| Uygulanan Alan | Eğitim | 49 | ,84 | ,82 | ,85 | 7,42 | 1 | ,01* |
| | Diğer | 52 | ,86 | ,85 | ,88 | | | |

Tablo 1 incelendiğinde GSS ölçeğinin uygulandığı alana (Eğitim ve diğer) göre güvenilirlik değerlerinin için anlamlı farklılık gösterdiği, ancak ölçeğin kullanıldığı formdaki seçenek sayısı (4'lü Likert ve diğer) ve uygulanan grubun özellikleri (Yetişkin ve öğrenci) açısından anlamlı bir farklılık göstermediği belirlenmiştir. Anlamlı farklılığın kaynağının belirlenmesi için yapılan analiz sonucunda anlamlı farklılığın eğitim alanından kaynaklandığı belirlenmiştir.

Tartışma, Sonuç ve Öneriler

Güvenirlik ölçekle ilişkili değildir, örneklemden örneğe farklı değerler verebilir, bu nedenle her çalışmada raporlanmalıdır (Crocker ve Algina, 2008). Çalışmalarda güvenilirlik katsayısının raporlanmaması sonuçların güvenilirliğini şüpheye sokmaktadır. Araştırmacılar bulgularının güvenilir olduğunu kanıtlamaktan sorumludur, bu nedenle daha dikkatli olmalıdırlar. Gerçekleştirilen bu çalışmada ölçeğin tümü için güvenilirlik katsayısı .85 olarak belirlenmiştir. Burada raporlanan değer Cronbach Alfa değeridir ve bu değer .70'in üzerinde kabul edilebilir bir değerdir (Nunnally, 1978).

Ölçeğin bütünü için heterojenlik göz önüne alındığında I² değeri 75'in üzerinde ve yeterli düzeydedir (Higgins vd., 2003). Bu durum ölçeğin örneklemden örnekleme farklı sonuçlar verdiğini ve genellenemeyeceğini göstermektedir. Bu nedenle heterojenliğe neden olabilecek moderatör değişkenler belirlenmiş ve moderatör değişken analizleri yapılmıştır. Bu bağlamda üç kategorik değişken belirlenmiştir. Belirlenen moderatör değişkenler; kullanılan likert tipi (4 ve diğer), uygulama grubu (yetişkin-öğrenci) ve uygulanan alandır (eğitim ve diğer). Ölçeğin bütünü açısından bu çalışma için belirlenen likert tipi ve uygulanan grup moderatör değişkenleri arasında istatistiksel olarak fark yoktur. Ancak uygulanan alan moderatör değişkenine göre istatistiksel olarak anlamlı farklılık vardır. Bu anlamlılık eğitim alanında uygulan ölçekler lehinedir. Bu farklılığın, öz yeterlik kavramının eğitim araştırmalarında daha sık ve sistematik biçimde kullanılmasıyla ilişkili olabileceği düşünülmektedir.

Sosyal bilimler çalışmalarında kullanılan ölçeklerden elde edilen ölçümlerin güvenilirliği, ölçüm sonuçlarına dayalı olarak yapılacak yorumlar için önemlidir. Düşük güvenilirlik değerleri ile yapılan yorumların güvenilirliği de düşük olacaktır. Bu çalışma sonucunda GSS'nin ölçeğinden elde edilen iç tutarlılık katsayısının yüksek olduğu görülmüştür.

Bu çalışmada GSS'nin güvenilirlik genellemesi belirlenen moderatör değişkenlerle yapılmıştır. Gelecek araştırmalarda, farklı moderatör değişkenlerin dâhil

edilmesiyle güvenilirliğin bağlamsal varyasyonları daha ayrıntılı biçimde incelenebilir. Bu çalışmada meta-regresyon yapılmamıştır; araştırmacılar belirleyecekleri değişkenlerle meta-regresyon yapabilirler. Çalışma Bonett (2002) dönüşümü kullanılarak yapılmıştır. Gelecek çalışmalarda araştırmacıların farklı örneklem grupları, farklı güvenilirlik katsayısı dönüşümleri yapılarak, farklı kültürel bağlamlar ya da farklı istatistiksel yöntemlere odaklanarak araştırma tasarımları yapmaları önerilmektedir.

Etik Komite Onayı: Bu araştırmada erişime açık olan veriler yayınlardan elde edilen veriler kullanıldığı için Etik Kurul Onayı alınmasına gerek yoktur.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar, çıkar çatışması olmadığını beyan etmiştir.

Finansal Destek: Yazar, bu çalışma için finansal destek almadığını beyan etmiştir.

Ethics Committee Approval: Since this research uses data obtained from openly accessible publications, it is not necessary to obtain Ethics Committee Approval.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflicts of interest to declare.

Financial Disclosure: The author declared that this study has received no financial support.

Kaynaklar

- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37(2), 122-147. <https://doi.org/10.1037/0003-066X.37.2.122>
- Bandura, A. (1997). *Self-efficacy; The exercise of control*. Freeman.
- Bakioğlu, A., & Gökaş, E. (2018). An educational policy making method: Meta-analysis. *Medeniyet Eğitim Araştırmaları Dergisi*, 1(2), 35-54.
- Bachner, Y. G., & O'rourke, N. (2007). Reliability generalization of responses by care providers to the Zarit Burden Interview. *Aging & Mental Health*, 11(6), 678-685. <https://doi.org/10.1080/13607860701529965>
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185-216. <https://doi.org/10.1177/135910457000100301>

- Bonett, D. G. (2002). Sample size requirements for testing and estimating coefficient alpha. *Journal of Educational and Behavioral Statistics*, 27(4), 335-340. <https://doi.org/10.3102/10769986027004335>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*, 100(3), 525-534. <https://doi.org/10.1037/0022-0663.100.3.525>
- Crocker, L., & Algina, J. (1986). *Introduction classical and modern test theory*. CBS College Publishing Company.
- Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Cengage Learning.
- Çapa, Y., Çakıroğlu, J. ve Sarıkaya, H. (2005), Öğretmen yeterlik duygusu ölçeğinin Türkçe versiyonunun geliştirilmesi ve doğrulanması, *Eğitim ve Bilim*, 30(137), 74-81. <https://doi.org/10.15390/ES.2005.668>
- Davis, A.L., & Rothstein, H.R. (2006). The effects of the perceived behavioral integrity of managers on employee attitudes: A meta-analysis. *J Bus Ethics* 67, 407-419. <https://doi.org/10.1007/s10551-006-9034-4>
- Dilek, H., & Gelbal, S. (2023). Reliability generalization meta analysis of the teachers' sense of efficacy scale. *Recep Tayyip Erdoğan Üniversitesi Eğitim Fakültesi Dergisi (REFAD)*, 3(1), 1-27.
- Dilekli, Y. (2015). *Öğretmenlerin düşünmeyi öğretmeye yönelik yaptıkları sınıf içi uygulamaları, öz-yeterlik düzeyleri ve öğretim stilleri arasındaki ilişki*. (Yayınlanmamış Doktora Tezi). Balıkesir Üniversitesi, Balıkesir, Türkiye.
- Dinçer, S. (2021). *Eğitim bilimlerinde uygulamalı meta analiz* (3. bs.). Ankara: Pegem A.
- Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods*, 10(4), 444-467. <https://doi.org/10.1037/1082-989X.10.4.444>
- Glass, G. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8. <https://doi.org/10.3102/0013189X005010003>
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Education and Debate*, 6(327), 557-560.
- Jerusalem, M., & Schwarzer, R. (1995). General self-efficacy scale English version (Gse-R). *APA PsycTests*. <https://doi.org/10.1037/t18916-000>
- Klassen R. M., Krawchuk L. L., & Rajani S. (2008). Academic procrastination of undergraduates: low self-efficacy to self-regulate predicts higher levels of procrastination. *Contemporary Educational Psychology*, 33(4), 915-931. <https://doi.org/10.1016/j.cedpsych.2007.07.001>
- Koca, F., & Dadandı, İ. (2019). Akademik öz-yeterlik ile akademik başarı arasındaki ilişkide sınav kaygısı ve akademik motivasyonun aracı rolü. *Elementary Education Online*, 18(1), 241-252. <https://doi.org/10.17051/ilkonline.2019.527207>
- Leganger, A., Kraft, P., & Røysamb, E. (2000). Perceived self-efficacy in health behaviour research: Conceptualisation, measurement and correlates. *Psychology & Health*, 15(1), 51-69. <https://doi.org/10.1080/08870440008400288>
- López-Ibáñez, C., López-Nicolás, R., Blázquez-Rincón, D. M., & Sánchez-Meca, J. (2024). Reliability generalization meta-analysis: Comparing different statistical methods. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues*, 43(20), 18275–18293. <https://doi.org/10.1007/s12144-023-05604-y>
- Nunnally, J. C. (1978) *Psychometric theory* (2nd ed.). McGraw-Hill.
- Özdemir, V., Yıldırım, Y., & Tan, Ş. (2020). A meta-analytic reliability generalization study of the Oxford Happiness Scale in Turkish sample. *Journal of Measurement and Evaluation in Education and Psychology*, 11(4), 374-404. <https://doi.org/10.21031/epod.766266>
- Polat, E. (2017). Kaygı düzeyi ve akademik öz-yeterlik inancının akademik başarı ile ilişkisi. *Yönetim Ve Ekonomi Araştırmaları Dergisi*, 15(1), 1-22. <https://doi.org/10.11611/yead.312024>
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86(3), 638-641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Rosenthal, R. (1991). Quality-weighting of studies in meta-analytic research. *Psychotherapy Research*, 1(1), 25-28. <https://doi.org/10.1080/10503309112331334031>
- Sawilowsky, S. S. (2000). Psychometrics versus datametrics: Comment on Vacha-Haase's "reliability generalization" method and some EPM editorial policies. *Educational and Psychological Measurement*, 60(2), 157-173. <https://doi.org/10.1177/00131640021970439>

- Sánchez-Meca, J., López-López, J. A., & López-Pina, J. A. (2013). Some recommended statistical analytic practices when reliability generalization studies are conducted. *British Journal of Mathematical and Statistical Psychology*, *66*(3), 402-425. <https://doi.org/10.1111/j.2044-8317.2012.02057.x>
- Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, *18*(3), 242-251. <https://doi.org/10.1027/1015-5759.18.3.242>
- Schunk, D. H., & Greene, J. A. (2017). *Handbook of self-regulation of learning and performance*. Routledge Publishing.
- Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived general self-efficacy on the Internet: Data collection in cyberspace. *Anxiety, Stress & Coping: An International Journal*, *12*(2), 145-161. <https://doi.org/10.1080/10615809908248327>
- Snyder, C. R., & Lopez, S. J. (2002). *Handbook of positive psychology*. Oxford University Press.
- Şen, S. & Yıldırım, İ. (2020). *CMA ile meta-analiz uygulamaları*. Anı Yayıncılık.
- Şen, S. (2022). A Reliability generalization meta-analysis of runco ideational behavior scale, *Creativity Research Journal*, *34*(2), 178-194, <https://doi.org/10.1080/10400419.2021.1960719>
- Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, *60*(2), 174-195. <https://doi.org/10.1177/00131640021970448>
- Vacha-Haase, T. (1998). Reliability generalization: Exploring variance in measurement error affecting score reliability across studies. *Educational and Psychological Measurement*, *58*(1), 6-20. <https://doi.org/10.1177/0013164498058001002>
- Yılmaz, E., Yiğit, R., & Kaşarç, İ. (2012). İlköğretim öğrencilerinin özyeterlilik düzeylerinin akademik başarı ve bazı değişkinler açısından incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, *1*(23), 371-388.