

Tekrarlama Krizi ve Psikoterapi Etkililik Çalışmaları

The Replication Crisis and Psychotherapy Efficacy Studies

Kutlu Kağan Türkarlan¹ 



¹Araştırma Görevlisi, Orta Doğu Teknik Üniversitesi, Psikoloji, Ankara-Türkiye

ORCID: K.K.T. 0000-0002-2440-3977

Sorumlu yazar/Corresponding author:

Kutlu Kağan Türkarlan,
Orta Doğu Teknik Üniversitesi, Psikoloji,
Ankara-Türkiye

E-posta/E-mail:
kutlu.turkarlan@metu.edu.tr

Başvuru/Submitted: 29.06.2021

Revizyon Talebi/Revision Requested:
14.09.2021

Son Revizyon/Last Revision Received:
14.02.2022

Kabul/Accepted: 05.03.2022

Online Yayın/Published Online: 08.04.2022

Citation/Atf: Türkarlan, K.K. (2022).
Tekrarlama krizi ve psikoterapi etkililik
çalışmaları. *Psikoloji Çalışmaları - Studies in
Psychology*, 42(1): 1-41.
<https://doi.org/10.26650/SP2021-959170>

ÖZ

Son 10 yıldır psikoloji alanında kendini gösteren ve alanı derinden sarsan tekrarlama krizi, psikolojinin her alt alanını, araştırma bulgularını ve genel kabullerini yeniden değerlendirmesi için harekete geçirmiştir. Klinik psikoloji alanının geç de olsa tekrarlama krizi tartışmalarına katılmaya başladığı görülmektedir. Uzun yıllardır gerçekleştirilen psikoterapi etkililik (efficacy) çalışmaları, klinik psikoloji içindeki başlıca deneysel alanlardan birini oluşturmaktadır. Psikoterapi etkililik çalışmaları çeşitli psikolojik tedavilerin psikolojik rahatsızlıklar üzerindeki etkilerini değerlendirmek için gerçekleştirilmektedir. Tekrarlama krizi bağlamında ortaya çıkan bulgular, psikoterapilerin etkililiklerini değerlendirmek için sıkça kullanılan seçkisiz kontrollü çalışmaların ve bu çalışmaların sonuçlarının toplu olarak değerlendirilmesine yardımcı olan meta-analizlerin pek çok yöntemsel problem ve yanlılık içerdiğini ortaya çıkarmıştır. Seçkisiz kontrollü çalışmalar düşük örneklem sayısı, bağlılık yanlılığı, seçici sonuç raporlama, örnekleme dahil etme ya da dışarıda bırakma kriterlerinin katılığı, kontrol grubu olarak bekleme listesi kullanımı, seçkisizleştirme ve körleştirme sorunları, çalışmayı bırakan katılımcıların analizlere dahil edilmemesi ve psikoterapist etkilerinin ihmal edilmesi gibi problemler ve yanlılıklar içermektedir. Meta-analizler ile ilişkili başlıca sorunlar ise yayın yanlılığı ve düşük kaliteli çalışmaların meta-analizlere dahil edilmesidir. Tüm bu durumlar psikoterapilerin olumlu etkilerini olduğundan daha yüksek gösterirken, edinilen bulguların sağlıklı bir şekilde değerlendirilmesini engellemektedir. Uzun yıllardır gerçekleştirilen çalışmalar, psikoterapilerin çeşitli psikolojik rahatsızlıkların tedavisinde tercih edilebilecek etkili araçlar olduğunu göstermektedir. Ancak dikkat edilmesi gereken iki önemli husus vardır. Bunlardan ilki psikoterapilerin etkililiğini gösteren çalışmaların yöntemsel anlamda güncel bilimsel standartları yakalamakta zorlanmasıdır. İkinci önemli husus ise psikoterapilerin gerçek etkilerinin çalışmalarda bulunan etkilerden daha az olabileceğidir. Bu derleme makalesinin amacı, psikoterapi etkililik çalışmalarını ve meta-analizleri etkileyen yöntemsel problemlere ve onların çözümlerine değinerek klinik psikoloji alanında bu konuda ortaya çıkan farkındalığın artırılmasına ve gerçekleştirilecek çalışmaların yöntemsel olarak güçlendirilmesine yardımcı olmaktır.

Anahtar Kelimeler: Psikoterapi, tekrarlama krizi, seçkisiz kontrollü çalışma, meta-analiz, yanlılık, dodo kuşu yargısı

ABSTRACT

In the last ten years, the replication crisis, which has manifested itself and shaken the field of psychology profoundly, has led researchers in every subfield of psychology to reconsider research findings and general assumptions. Though late to this trend, clinical psychology researchers have started to participate in these discussions. Psychotherapy efficacy studies, carried out over many years, constitute a major experimental area of clinical psychology. Psychotherapy efficacy studies are conducted to evaluate the effects of psychological treatments on various mental disorders. The findings of the replication crisis indicate that randomized controlled trials, which are frequently used to examine the efficacy of psychotherapies and meta-analyses that help evaluate the results of randomized controlled trials, have several methodological problems and biases. Randomized controlled trials have problems and biases such as low sample size, allegiance bias, selective outcome reporting, strictness of inclusion or exclusion criteria, using wait-list as a control group, problems of randomization and blinding, exclusion of drop-out subjects from the analyses, and ignoring psychotherapists effects. The main methodological issues of meta-analyses are related to publication bias and the inclusion of low-quality studies in meta-analyses. Research over the years has demonstrated that psychotherapies are effective tools for treating various mental disorders. However, there are two critical points to note. First, the studies showing the efficacy of psychotherapies barely meet current scientific methodological standards. Second, the real-life effectiveness of psychotherapies may be lower than those found in clinical trials. This review article aims to increase awareness regarding replication crisis within the field of clinical psychology and strengthen the methodological approach of future studies by addressing the methodological problems of and suggesting solutions for psychotherapy efficacy studies/meta-analyses.

Keywords: Psychotherapy, replication crisis, randomized controlled trial, meta-analysis, bias, dodo bird verdict

EXTENDED ABSTRACT

Recent discussions in the field of psychology confirm the existence of the replication crisis and reveal that psychology has been facing a methodological crisis for several years (Maxwell et al., 2015; Open Science Collaboration, 2015). A major attempt (Open Science Collaboration, 2015) to replicate well-known cognitive and social psychology studies published in high-quality journals resulted in only a quarter of the studies being replicated. These unexpected results led researchers to question and review the research methods and practices used in the field of psychology (Nosek et al., 2015). Compared to other fields of psychology, clinical psychology was late to respond to these methodological discussions (Leichsenring et al., 2017; Sakaluk et al., 2019; Tackett et al., 2017). Clinical psychology research mainly includes studies on psychological evaluation, diagnosis of mental disorders, investigation of individual differences in psychopathologies, and treatment of mental disorders, meaning that the research utilizes both experimental and observational methods (Tackett et al., 2017). Psychotherapy efficacy studies are commonly conducted experimental studies in clinical psychology. Researchers use randomized controlled trials (RCTs) to examine the efficacy of a psychotherapy intervention for a specific mental disorder (Mulder et al., 2018; Wampold et al., 2011). Furthermore, the findings of RCTs are evaluated with meta-analyses, which can quantitatively summarize the results of numerous studies (Berkeljon & Baldwin, 2009; Leichsenring et al., 2017; Smith & Glass, 1977). With the

replication crisis, a need to review psychotherapy efficacy studies and meta-analyses in terms of methodological problems has arisen.

RCTs may have a number of methodological problems and biases, including (1) small sample size (Button et al., 2013; Sakaluk et al., 2019), (2) allegiance bias (Leichsenring et al., 2017; Leykin & DeRubeis, 2009), (3) selective outcome reporting (Bradley et al., 2017; Miguel et al. 2021; Shinohara et al., 2015), (4) the strictness of inclusion or exclusion criteria (Hoertel et al., 2014, 2015; Westen et al., 2004; von Wolff et al., 2014), (5) using wait-list as a control group (Cuijpers et al., 2016; Furukawa et al., 2014; Munder et al., 2019), (6) randomization and blinding problems (Cuijpers et al., 2010; Cuijpers & Cristea, 2016), (7) the exclusion of drop-out participants from the analyses (Cooper & Conklin, 2015; Cuijpers & Cristea, 2016; Fernandez et al., 2015), and (8) ignoring psychotherapists' effects (Del Re et al., 2012; Johns et al., 2019; Leichsenring et al., 2017; Owen et al., 2015). These problems can be prevented by applying a rigorous methodology. In terms of meta-analyses, publication bias (Driessen et al., 2015; Flint et al., 2015; Fonagy et al., 2017; Linardon et al., 2019; Niemeyer et al., 2012) and the inclusion of low-quality studies in meta-analyses (Coyne & Kok, 2014; Cuijpers & Cristea, 2016; Hengartner, 2018) deflate the validity of the results of meta-analyses. Therefore, a new publishing paradigm is needed to overcome publication bias (Bradley et al., 2017). Moreover, eliminating methodological problems and biases of RCTs by itself increases the quality of the studies analyzed in meta-analyses. Considering the problems related to psychotherapy efficacy studies and meta-analyses, it could be argued that methodological crisis in clinical psychology is responsible for the Dodo Bird Verdict (Lambert, 2013). Besides the problems related to RCTs and meta-analyses, the interpretation of the psychotherapy efficacy studies must consider four important issues: meanings of the effects size, the duration of the observed effect, response rate of the treatment, and transfer of the effect to real-world settings.

Discussion

Goldfried (2020) asserted that the field of psychotherapy is still an infant field, and it cannot be regarded as a *mature science*, which has a solid and established core as well as new developments. In contrast, the field of psychotherapy is a conflicting and competitive field in which different theoretical approaches strive to maintain their existence and prove the effectiveness of their methods (Budd & Hughes, 2009). Based on the literature, it could be asserted that psychotherapy is an effective treatment for various mental disorders (Lambert, 2013; Rubio-Aparicio et al., 2018). However, all these methodological problems and biases of the RCTs and meta-analyses indicate two important issues. First, the studies conducted to test the effects of psychotherapies barely meet current scientific standards

(Coyne & Kok, 2014; Hengartner, 2018; Leichsenring et al., 2017; Sakaluk et al., 2019). Second, the real-life effectiveness of psychotherapies is less than that found in previous studies (Schäfer & Schwarz, 2019; Wampold, 2019). For example, when specific biases were corrected, the efficacy of psychotherapies for depression was slightly above the accepted threshold (Cuijpers et al., 2018). The findings of RCTs and meta-analyses play a critical role in the selection of treatment options for mental disorders (Shean, 2016). This issue perhaps adds even more ethical responsibilities for clinical psychologists (Tackett et al., 2017). In conclusion, the findings suggest that clinical psychology researchers and clinicians must be aware of the effects of the replication crisis on the field and strive to bring their research closer to up-to-date scientific standards (Tackett et al., 2019).

Son yıllarda psikoloji alanında ortaya çıkan yöntemsel tartışmalar sonuç olarak *tekrarlama krizi* adı verilen bir olgunun varlığına işaret etmiş ve psikolojinin geçmişten beri devam eden yöntemsel bir kriz içinde olduğunu gözler önüne sermiştir (Maxwell ve ark., 2015; Open Science Collaboration, 2015). Open Science Collaboration'ın (2015) yüksek kalitedeki üç dergide yayınlanmış 100 kadar bilişsel ve sosyal psikoloji çalışmasını tekrarlama girişimi, çalışmalardan sadece %36'sının tekrarlanabildiğini ortaya çıkarmıştır. Bu çalışmadaki diğer ilginç bir bulgu ise tekrarlama çalışmalarında bulunan etki büyüklüklerinin, orijinal çalışmalarda bulunanlardan daha düşük olduğudur (Open Science Collaboration, 2015). Yıllardır varlığına kesin gözüyle bakılan, ders kitaplarına girmiş ve psikoloji alanının kendisi ile övünmesine sebep olan pek çok bulgunun yapılan tekrarlama çalışmaları ile desteklenememesi, araştırmacıları psikoloji alanındaki araştırma yöntemlerini ve uygulamalarını gözden geçirmek ve iyileştirmek için harekete geçirmiştir (Nosek ve ark., 2015). Psikolojinin diğer alanlarına kıyasla klinik psikoloji, alandaki yöntemsel kriz tartışmalarına katılmaya ve kendisini sorgulamaya yeni başlıyor gibi görünmektedir (Leichsenring ve ark., 2017; Sakaluk ve ark., 2019; Tackett ve ark., 2017). Klinik psikoloji araştırmaları psikolojik değerlendirme, psikolojik rahatsızların tanınması, psikopatolojiler ile ilgili kişisel farklılıkların araştırılması ve psikolojik rahatsızlıkların tedavisi gibi çeşitli konuları içermektedir. Deneysel çalışmaların gözlemsel çalışmalara göre tekrarlama krizinden daha yüksek oranda etkilendiği düşünüldüğünde (Tackett ve ark., 2017), ağırlıklı olarak deneysel olarak gerçekleştirilen psikoterapi etkililik çalışmalarının tekrarlama krizi ışığında yeniden değerlendirilmesi önemli bir ihtiyaç haline gelmiştir. Bu derleme makalesinin amacı psikoterapi etkililik çalışmalarının ve bu çalışmalar ile yapılan meta-analizlerin tekrarlama krizi ile nasıl ilişkili olduğunu göstermek ve aynı zamanda araştırmacıların ve klinisyenlerin bu konulardaki farkındalıklarını arttırmaya yardımcı olmaktır.

Kanıt Dayalı Terapiler

Bilindiği gibi psikoterapilerin etkililiği ile ilgili tartışmalar çok daha eski tarihlere dayanmaktadır (Eysenck, 1952, 1964; Rosenzweig, 1954; Strupp, 1963). Meta-analiz yönteminin psikoterapi etkililik çalışmalarının sonuçlarını değerlendirmek için kullanılmaya başlanması ile psikoterapilerin etkililiği ile ilgili soru işaretleri, psikoterapilerin lehine olacak şekilde ortadan kalkmaya başlamıştır (APA Presidential Task Force on Evidence-Based Practice, 2006; Lambert, 2013; Landman ve Dawes, 1982; Rubio-Aparicio ve ark., 2018; Sakaluk ve ark., 2019; Smith ve Glass, 1977). Yaklaşık 25 yıl önce

Amerikan Psikoloji Derneği Klinik Psikoloji Topluluğu (APDKPT), belli örneklemdeki etkililiği bilimsel çalışmalar ile gösterilmiş psikolojik tedavileri *ampirik olarak kanıtlanmış psikolojik tedaviler* şeklinde adlandırmış ve o dönemde ampirik kanıt olduğu bilinen psikoterapilerin bir listesini yayınlamıştır (Task Force on Promotion and Dissemination of Psychological Procedures Division of Clinical Psychology, 1995). Günümüzde *kanıta dayalı terapiler* olarak adlandırılan bu kavram ortaya çıktığı günden beri eğitim programlarının düzenlenmesi, sigortaların psikoterapi ücretlerini karşılaması ya da belli psikolojik rahatsızlıklara yönelik tedavi tavsiyeleri verilmesi gibi pek çok farklı bağlamda önemli bir karar kriteri olarak değerlendirilmektedir (Cuijpers ve Cristea, 2016; Shean 2016). Psikoterapi yöntemlerinin ampirik olarak desteklenmiş olup olmadıkları genelde, seçkisiz kontrollü çalışmalar (SKÇ) ve bu çalışmaların sonuçlarını toplu olarak değerlendiren meta-analizler ile ortaya konmaktadır (Chambless ve Hollon, 1998; Tolin ve ark., 2015).

APDKPT uzun yıllardır psikoterapilerin etkililiğine yönelik bulguları Chambless ve Hollon'un (1998) ortaya koydukları kriterlere göre üç farklı şekilde kategorilendirmekteydi (Sakaluk ve ark., 2019). Bir psikoterapi yönteminin *kuvvetli olarak desteklenmiş* sayılması için ilk olarak tedavi almama, plasebo tedavi alma ya da alternatif tedaviler alma durumlarına kıyasla üstünlük göstermesi gerekmektedir. Bu üstünlük aynı zamanda en az 2 ayrı araştırma grubu tarafından yeterli istatistiksel güce sahip çalışmalar ile ya da çok sayıda (dokuzdan fazla) tek vakalık deney çalışmaları ile desteklenmelidir. Sadece tek bir çalışma ya da az sayıda (üçten fazla) tek vakalık deney çalışması bir psikoterapinin etkililiğini destekliyorsa psikoterapi *makul destekli* olarak nitelendirilirken, psikoterapinin etkililiği ile ilgili olarak çelişkili bulguların olduğu durumlarda ise psikoterapi yöntemi *tartışmalı desteğe sahip* olarak değerlendirilmekteydi (Chambless ve Hollon, 1998; Sakaluk ve ark., 2019). Bu kriterlerin işaret ettiği gibi APDKPT'nin bir psikoterapi yönteminin kanıta dayalı olduğunu belirtmesi isteğe bağlı olarak değerlendirilebilecek kriterlere göre gerçekleşmektedir (Tackett ve ark., 2017). Çünkü bir psikoterapiyi desteklenmiş sınıfına atamak için onun etkili olduğunu ortaya koyan kaç çalışma yapılması gerektiğini gösteren herhangi bir somut ve istatistiksel kanıt yoktur. Örneğin, iki ayrı çalışmanın psikoterapinin etkililiğini gösterdiği ancak pek çok diğer çalışmanın psikoterapinin etkililiğini destekleyemediği durumlarda ne yapılacaktır? Chambless ve Hollon (1998) böyle bir durum oluştuğunda psikoterapinin etkililiği ile ilgili yapılan çalışmaların meta-analiz yöntemi kullanılarak değerlendirilmesini önermişlerdir.

Psikoterapilerin etkililiklerine yönelik bu değerlendirmelerin kısıtlılıklarını gidermek için Tolin ve arkadaşları (2015) daha kapsamlı ve bütüncül bir model önerisinde bulunmuştur. Bu modeldeki değerlendirme iki aşamalı olarak gerçekleştirilmektedir. Öncelikle yetkin kişilerden oluşan bir görev gücü belirli bir psikoterapinin etkililiği ile ilgili yapılandırılmış sistematik derleme çalışmalarını yanlılık risklerine de bakarak değerlendirecektir. Daha sonra ise bulunan kanıtların ne kadar güçlü olduğu fayda, maliyet, zarar ve tedavinin etki büyüklüğü gibi faktörler bağlamında uygulamaya dönük olacak şekilde tekrar gözden geçirilecek ve tedavinin ampirik olarak (çok kuvvetli, kuvvetli ya da zayıf) ne seviyede desteklenmiş olduğu belirlenecektir (Tolin ve ark., 2015). APDKPT birkaç sene önce internet sitesindeki kanıta dayalı terapiler listesini Tolin ve arkadaşlarının (2015) önerdiği yeni kriterlere göre tekrar gözden geçirmeye başlamıştır. Özetle bir psikoterapinin ampirik olarak desteklenmesi, onun etkili olduğunu gösteren SKÇ'lere ve bu çalışmaların ortalama etkisini gösteren meta-analizlere bakarak değerlendirilmektedir (Budd ve Hughes, 2009). SKÇ'ler temelde deneysel çalışmalardır ve yöntemsel olarak pek çok kusur içerebilirler. Benzer şekilde meta-analiz çalışmalarının sonuçları ise hem alanyazındaki SKÇ'lerin ne kadar kaliteli olduğundan hem de meta-analize hangi çalışmaların dahil edileceğinden etkilenmektedir. Uzun yıllardır yapılan araştırmaların sonuçlarının çeşitli yanlılıklar içerdiğine dair yapılan uyarılar ve tekrarlar krizi bağlamında oraya çıkan yeni tablo, psikoterapilerin etkililiğini göstermek için yapılan SKÇ ve meta-analizlerin pek çok yöntemsel kısıtlılık ve sorunlu araştırma uygulamasından mustarip olduğunu gözler önüne sermiştir (Bradley ve ark., 2017; Coyne ve Kok, 2014; Cuijpers ve Cristea, 2016; Hengartner, 2018; Luborsky ve ark., 1999; Munder ve ark., 2013; Sakaluk ve ark., 2019; Westen ve Morrison, 2001). Mevcut sorunların bir listesi ve çözüm önerileri Tablo 1'de görülebilir.

Tablo 1.*Seçkisiz Kontrollü Çalışmalar ve Meta-Analizlerdeki Yöntemsel Problemler*

Problemler	Çözüm Önerileri
Seçkisiz Kontrollü Çalışmalar	
<i>Düşük örneklem sayısı</i>	Örneklem sayısının istatistiksel güç analizi ile belirlenmesi, araştırma grupları iş birliği, katılımcıların çalışmada kalmasının sağlanması, eksik verilerin yordanması, daha basit hipotezler ve analizler
<i>Bağıllık yanlılığı</i>	Analiz ve puanlamaların bağımsız ve kör araştırmacılar tarafından gerçekleştirilmesi, farklı yaklaşımlardan araştırmacılarla iş birliği yapılması, çalışma kayıt protokollerinin güncellenmesi, araştırma yöntemleri ve etik eğitimi, bağıllığı değerlendiren/kontrol eden ölçüm araçlarının veya yöntemlerin kullanılması
<i>Seçici sonuç raporlama</i>	Ön kayıt, ön kayıt-yayın uyumunun kontrol edilmesi, fonların sadece ön kayıt-yayın tutarlılığı sağlayan araştırmacılara verilmesi, meslek topluluklarının uyumun sağlanması için adım atması, araştırmacıların yayın kara listesine alınması
<i>Örnekleme dahil etme ya da dışarıda bırakma kriterlerinin katılığı</i>	Bulguların pragmatik seçkisiz kontrollü çalışmalarla desteklenmesi
<i>Kontrol grubu olarak bekleme listesinin kullanımı</i>	Kontrol grubu olarak bilinen bir psikoterapi, ilaç tedavisi ya da hiç tedavi almamanın kullanılması
<i>Seçkisizleştirme ve körleştirme sorunları</i>	Örnekleme büyütme, seçkisizleştirme ve körleştirme süreçlerinin bağımsız araştırmacılar tarafından gerçekleştirilmesi, araştırma yöntemleri ve etik eğitimi
<i>Çalışmayı bırakan katılımcıların analizlere dahil edilmemesi</i>	Çalışmayı bırakan katılımcıların sonuç verilerinin yordanması, yarım bırakmaya sebep olacak etmenlerin tespit edilmesi ve önüne geçilmesi
<i>Psikoterapist etkilerinin ihmal edilmesi</i>	Psikoterapist seçim, eğitim ve süpervizyon süreçlerinin düzenlenmesi, psikoterapist etkisinin tedavi kılavuzlarında görünür hale getirilmesi, psikoterapistlerin vakalara seçkisiz olarak atanması, çok seviyeli modelleme yönteminin kullanılması
Meta-analizler	
<i>Yayın yanlılığı</i>	Zorunlu yayın hakkının olması, iki aşamalı değerlendirmeye geçilmesi, olumsuz sonuç dergilerinin kurulması, ön kayıt sistemine geçilmesi, gri alanyazının taranması, yayın yanlılığının etkisini kontrol eden istatistiksel yöntemlerin kullanılması
<i>Düşük kaliteli çalışmalar</i>	İki aşamalı değerlendirme ve ön kayıt sistemlerine geçilmesi, analizlere sadece yüksek kaliteli araştırmaları dahil etme, düşük kaliteli çalışmaların etkisinin kontrol edilmesi, SKÇ'lerdeki yöntemsel problemlerin önüne geçilmesi

Seçkisiz Kontrollü Çalışmalar ve Yöntemsel Problemleri

SKÇ'ler günümüzde psikoterapilerin etkililiğini değerlendirmek için en güvenilir yöntem olarak kabul edilmekte ve yaygın şekilde kullanılmaktadır (Freire, 2006; Mulder ve ark., 2018; Wampold ve ark., 2011). SKÇ'lerde katılımcılar seçkisiz olarak psikolojik bir tedavi verilen koşula ya da buna alternatif olacak başka bir tedavi ya da kontrol koşuluna atanırlar (Wampold ve ark., 2011). Bir psikoterapi yöntemi eğer etkili-

lik gösteriyor ise, psikoterapi alan gruptaki katılımcıların müdahale sonunda rahatsızlıklarını değerlendiren ölçüm araçlarından aldıkları ortalama puanın, diğer gruptaki katılımcıların ortalama puanından istatistiksel olarak anlamlı bir şekilde daha yüksek ya da daha düşük olması gerekir (Kendall, 2003). Örneğin, çalışma sonlandığında kaygı bozukluğu için bilişsel davranışçı terapi (BDT) almış ve kaygı bozukluğu tanısı olan katılımcıların kaygıyı değerlendiren ölçeklerden aldıkları puanların, bekleme sırasındaki aynı tanılı katılımcıların aldıkları puanlardan istatistiksel olarak anlamlı bir şekilde daha az olması beklenir. Eğer bu şekilde bir sonuç ortaya çıkıyorsa ise BDT'nin kaygı bozukluğu tanısı olan kişilerin kaygı belirtilerini azaltmak için etkililiği gösterilmiş psikolojik bir tedavi yöntemi olduğuna dair bulguların edinildiği söylenebilir.

Kendall'a (2003) göre kalitesi yüksek bir SKÇ'de katılımcıların hipotezlere uygun bir örneklemden seçilmesi; katılımcıların farklı koşullara seçkisiz bir şekilde atanması; farklı gruplara uygulanan müdahalelerin aktif tedavi bileşeni dışında olabildiğince aynı olması; çalışmayı yürüten araştırmacıların ve katılımcıların hangi koşullarda olduklarına kör olmaları ve tüm sürecin planlandığı gibi gittiğinin kontrolünün yapılması gerekmektedir. Böylece SKÇ'ler karıştırıcı etkileri olabildiğince azaltarak neden-sonuç ilişkilerinin daha sağlıklı kurulabilmesine imkân vermektedir (Wampold ve ark., 2011). SKÇ'lerin sayılan özellikleri göz önüne alındığında psikoterapilerin etkililiğini değerlendirmek için SKÇ'lerin kullanılmasının daha farklı bir bağlamda gerçekleştiği söylenebilir. Psikoterapi etkililik çalışmaları için SKÇ'ler planlanırken içsel geçerliliği arttırmak için çeşitli yöntemlere başvurulmaktadır. Philips ve Falkenström (2021) bu yöntemleri şu şekilde sıralamışlardır: İlk olarak SKÇ'ye katılma ihtimali olan katılımcılardan türdeş bir örneklem oluşturulmaya çalışılır. Yani katılımcıların benzer psikolojik rahatsızlıkları olmasına dikkat edilir. Örneğin bir katılımcının sadece depresyon rahatsızlığı varken diğer katılımcının hem depresyon rahatsızlığı hem de kişilik bozukluğu olması bu türdeşliği bozan bir etki ortaya çıkarmaktadır. İkinci olarak tedavi yöntemi, tedavi el kitabı olarak adlandırılan bir yönerge doğrultusunda gerçekleştirilir. Tedavi el kitabı tedavinin süreç boyunca nasıl uygulanacağını ve klinisyenin terapi süresi boyunca ne tür teknikler ve müdahalelerde bulunacağını adım adım somutlaştıran bir yönerge işlevi görmektedir. Üçüncü olarak tedaviyi uygulayacak terapistlerin test edilecek terapi yönteminin yeterli bir eğitimini almış olmaları gerekmektedir. Son olarak, terapistlerin test edilecek tedaviyi yeterli ve düzgün bir şekilde uyguluyor olduklarının değerlendirilmesi (tedavi bütünlüğü) gerekmektedir.

SKÇ'ler tıpta ilaçların etkililiğini değerlendirmek için geliştirilmiş kanıta dayalı tıp paradigmasının bir yansıması olarak psikoterapilerin etkililiğini değerlendirmek için uzun yıllardır kullanılmaktadır (Budd ve Hughes, 2009). Her ne kadar bir psikoterapinin etkililiğini değerlendirmek için altın standart olarak görülseler de SKÇ'lerin psikoterapilerin etkililiğini göstermek için uygun bir yöntem olup olmadığı ile ilgili tartışmalar hala devam etmektedir (Budd ve Hughes, 2009; Gelo ve ark., 2020; Persons ve Silberschatz, 1998; Philips ve Falkenström, 2021; Shean, 2012; Wampold ve ark., 2011; Westen ve ark., 2004). Persons ve Silberschatz (1998) el kitabına sıkı sıkıya uyularak yürütülen süreçlerin psikoterapide danışanın ihtiyaç ve durumuna göre gerçekleştirilebilecek esnemeleri engellediğine işaret etmektedir. Klinik uygulamada tedavi el kitabına sıkı sıkıya bağlı olmak her zaman mümkün olmamakla beraber bu tarz bir bağlılık danışanın mevcut haline göre tedavi yönteminde değişiklik yapılmasına engel oluşturmaktadır. Westen ve arkadaşları (2004) örneklemin türdeşleştirilmesinin ve katılımcıların kişiliklerinden ve kişilik patolojilerinden bağımsız olarak tedavi edilebileceklerinin varsayılmasının, SKÇ'lerin dış geçerliliğini düşürdüğüne dikkat çekmişlerdir. Yani etkili olduğu klinik çalışmalarla gösterilmiş psikolojik bir tedaviyi kullanan klinisyenler klinik uygulamada birden fazla eş tanısı ve kişilik bozukluğu olan danışanları (tedavisi görece daha güç kişiler), klinik çalışmalarda tek ya da az sayıda tanısı olan danışanlarda (tedavisi görece daha kolay) etkili olduğu bulunan tedavi yöntemlerini uygulayarak tedavi etmeye çalışmaktadırlar. Bu durum etkili bulunan psikoterapinin klinik uygulamada karşılaşılan danışanlarda daha düşük etki göstermesine neden olabilmektedir. Budd ve Hughes (2009) psikoterapilerde işe yarayan tedavi bileşenlerinin terapist-danışan etkileşimleri seviyesinde karşılıklı olarak çalıştığını ancak SKÇ'lerin hatalı olarak aktif bileşenlerin farklı psikoterapi tedavi paketleri seviyesinde çalıştığını varsaydığını vurgulamışlardır. Diğer bir deyişle SKÇ'ler danışanların ve terapist-danışan etkileşiminin psikoterapi sürecini nasıl etkilediğini göz ardı etmektedir. Wampold ve arkadaşları (2011) bir etkiyi bulmanın onu açıklamaktan daha kolay olduğunu hatırlatarak, SKÇ'lerin psikoterapide değişimin tam olarak nasıl gerçekleştiğini açıklamaya yardımcı olmakta zorlandıklarını belirtmektedir. SKÇ'lerde bulunan etki, test edilen tedavi paketinin tamamına aittir. Ancak belki de tedavi paketinin sadece bazı öğeleri (örn., BDT bilişsel ve davranışçı parçalardan oluşmaktadır) belli bir etkiyi ortaya çıkarmak için yeterli olmaktadır. Bu parçaların hangileri olduğunu öğrenmek için, bir psikoterapiyi işe yaradığı düşünülen parçalara ayırarak o spesifik parçanın işe yarayıp yaramadığını gösteren *parçalarına ayırma* çalışmalarının yapılması gerekmektedir. Shean'e (2012,

2016) göre SKÇ'ler genelde kalıcı ve anlamlı değişimlerin görece kısa sürede ortaya çıkacağını ve bu etkilerin ölçülebileceğini varsaymaktadır. Ancak her danışan belli bir psikolojik tedavi yöntemine sabit bir süre zarfında cevap vermiyor olabilir. Buna ek olarak seçkisizleştirilmenin tam anlamıyla işe yaraması için örneklem sayılarının yeterince büyük olması gerekmektedir (Shean, 2016). Düşük örnekleme sahip bir çalışmada seçkisiz atama yapılmış olsa bile karıştırıcı değişkenler farklı gruplara, kontrol edilebilecek şekilde dağılmamış olabilir. Gelo ve arkadaşları (2020) SKÇ'lerin nicel anlamda psikoterapi araştırmalarını tahakküm altında tutarak nitel araştırma yöntemlerini gölgede bırakmış olduklarını ve bu durumun bilimsel çoğulculuğu zayıflattığını belirtmişlerdir. Yaptıkları tarama çalışmasında taranan 9915 psikoterapi araştırmasının %74.67'sinin ampirik-nicel yöntemleri kullanan çalışmalar olduğunu raporlamışlardır (Gelo ve ark., 2020). Son olarak Philips ve Falkenström (2021) SKÇ'lerde çift taraflı körleme yapılmasının tam olarak mümkün olmadığını ve psikoterapinin iyileşmenin nedeni olduğuna yönelik direkt nedensellik bağlantısının kurulmasının problemliliğini belirtmişlerdir. SKÇ'lerdeki terapistler uyguladıkları terapiyi zaten bilmektedirler. Danışanlar ise uygulanan tedaviden yola çıkarak hangi koşulda olduklarını çıkarsayabilirler. Buna ek olarak SKÇ'ler direkt nedensellik kabulüne dayandıkları için insan zihni ve psikoterapi uygulamasını karmaşık bir sistem olarak görmek yerine, "X, Y'ye sebep oluyor" şeklindeki basit bir nedensellik bağlamında ele almakta ve böylece bilinçlilik ve özgür irade gibi süreci karmaşık hale getiren önemli noktaları atlamaktadırlar (Philips ve Falkenström, 2021). SKÇ'ler psikoterapi bağlamında tam anlamıyla deneysel çalışmalar olmasalar da deneysel çalışmaların başta seçkisiz atama ve kontrol grubu gibi temel önemli özelliklerine sahiptirler (Donnon, 2012; Kendall, 2003). Bu sebeple de çalışmaları tasarlayan, yürüten ve sonuçlarını analiz eden araştırmacıların bu süreçlerin yürütülmesi sırasında aldıkları kararlar ve gerçekleştirdikleri uygulamalardan etkilenirler. SKÇ'ler ile yapılan psikoterapi etkililik çalışmalarının açığa çıkardığı sonuçları aldatıcı bir şekilde etkileyen bu etmenlerden öne çıkanlar düşük örneklem sayısı, bağlılık yanlılığı, seçici sonuç raporlama, örnekleme dahil etme ya da dışarıda bırakma kriterlerinin katılığı, kontrol grubu olarak bekleme listesi kullanımı, seçkisizleştirme ve körleştirme problemleri, çalışmayı bırakan katılımcıların analizlere dahil edilmemesi ve psikoterapist etkilerinin ihmal edilmesidir.

Düşük Örneklem Sayısı

Psikoterapi etkililik çalışmaları yöntemsel olarak değerlendirildiğinde göze çarpan en önemli problemlerden ilki pek çok çalışmanın örneklem sayılarının ve bununla bağ-

lantılı olarak istatistiksel güçlerinin düşük olmasıdır (Sakaluk ve ark., 2019). İstatistiksel güç bir çalışmanın yokluk hipotezinin yanlış olması durumunda yokluk hipotezini doğru bir şekilde reddetme ihtimalidir (Button ve ark., 2013). Diğer bir deyişle istatistiksel güç, olan bir etkinin doğru bir şekilde tespit edilmesi ihtimali olarak tanımlanabilir. Bu sebeple bir çalışmanın istatistiksel gücü ne kadar yüksekse bir etkiyi doğru bir şekilde bulma ihtimali de o kadar yüksek olacaktır. İstatistiksel gücün alanyazındaki referans değeri .80 olarak kabul edilmektedir (Cohen, 1992). Bu değer altındaki istatistiksel güç değerleri düşük istatistiksel güç olarak tanımlanırken, bir istatistiksel gücün ne kadar düşük olduğu ile ilgili standart bir sınıflandırma yoktur. İstatistiksel güç temelde bir testin anlamlılık kriteri değerine, örneklem büyüklüğüne ve aranan etkinin popülasyondaki tahmini büyüklüğüne bağlı olarak değişmektedir (Rossi, 2013). Anlamlılık kriteri genelde .05 olarak kabul edildiği ve aranan etkinin popülasyondaki büyüklüğü araştırmacıların kontrolü dışındaki bir değişken olduğu için örneklem büyüklüğü istatistiksel güç belirlenirken daha önemli bir değişken haline gelmektedir. Örneğin en basit şekilde bağımsız iki grubun kıyaslandığı bir t-testinin, çift taraflı olarak orta büyüklükteki bir etkiyi ($d = .50$) .80 istatistiksel güç ile tespit edebilmesi için toplam 128 kişilik bir örneklem büyüklüğüne sahip olması gerekmektedir (Faul ve ark., 2007). Diğer değişkenler sabit tutulurken örneklem büyüklüğü 64 kişiye düşürülürse çalışmanın ulaşabildiği istatistiksel güç değeri de .50'ye düşecektir. İstatistiksel güç sadece örneklem büyüklüğüne bağlı olmasa da uygulamada onunla güçlü bir ilişkisi olduğu söylenebilir.

Kabul edilen .80 değerinin altında kalan düşük istatistiksel güç genelde tip 2 hata (β) ile ilişkilendirilmesine rağmen araştırmacılar düşük istatistiksel gücün aynı zamanda tip 1 hatayı (α) arttırdığına işaret etmişlerdir (Button ve ark., 2013; Christley, 2010; Dumas-Mallet ve ark., 2017). Bir testin doğru sonucu göstermesi ihtimali istatistiksel güç ($1 - \beta$) ve tip 1 hata değeri ile ilişkili olan olumlu yordama değerine ($[(1 - \beta) \times R] / [(1 - \beta) \times R + \alpha]$) dayanmaktadır (Button ve ark., 2013; Christley, 2010). Aranan etkinin incelenen etkiler içinde yok olmama ihtimali (R) sabit tutulduğunda, düşük istatistiksel güç tip 1 hata oranını arttırmaktadır. Yani düşük istatistiksel gücü olan çalışmalar hatalı pozitif sonuçlar vererek olmayan etkileri varmış gibi gösterebilirler. Psikoterapi araştırmaları bağlamında düşünüldüğünde bu hatalı pozitif sonuç, bir psikoterapinin uygulandığı rahatsızlıkta etkili olduğunu gösteren sonuçtur. Diğer bir deyişle bu psikoterapi o rahatsızlık için aslında etkili bir yöntem değilken yanıltıcı bir şekilde etkili gibi gözükemektedir. Düşük istatistiksel güç ile ilgili diğer olumsuz bir sonuç ise bulunan etkinin,

örneklem büyüklüğünün az olması ve ancak büyük bir etki bulunduğunda anlamlı gözükmesi sebebiyle gerçek dünyadaki etkinin şişirilmiş bir hali olmasıdır (Button ve ark., 2013; Cuijpers ve ark., 2010; Gelman ve Carlin, 2014). Örneğin, klinik çalışmada etki büyüklüğü (tedavinin sonuç değişkeni üzerindeki etkisinin büyüklüğünü gösteren istatistiksel değer) .80 bulunan bir psikoterapinin gerçek dünyadaki etki büyüklüğü .50 olabilir. Özetle düşük örneklem sayısına ve düşük istatistiksel güce sahip çalışmaların ortaya çıkardığı sonuçlar gerek psikoterapinin etkili olması gerekse de etkisinin büyüklüğü açısından yeterince güvenilir değildir.

Sakaluk ve arkadaşları (2019) APDKPT'nin internet sitesinde kanıta dayalı oldukları belirtilen 79 farklı psikoterapi yöntemini destekleyen çalışmaları, çeşitli yöntemsel ve istatistiksel değerler açısından analiz etmişlerdir. Gerçekleştirilen psikoterapi etkililik çalışmalarının ortalama istatistiksel güç değeri .58 ($SS = .21$) olarak bulunurken, ortalama tekrarlamaya indeksi .52 ($SS = .24$) olarak bulunmuştur. Daha spesifik olarak psikolojik rahatsızlıklar ve onlara özgü psikoterapilerin değerlerine bakıldığında ise bazı psikoterapilerin ortalama istatistiksel güçlerinin standartların altında kaldığı görülmektedir. Örneğin, kuvvetli olarak desteklenmiş kategorisinde olan panik bozukluk için BDT çalışmalarının ortalama istatistiksel gücü .53 iken, depresyon için bilişsel terapi çalışmalarının ulaştığı ortalama istatistiksel güç .57 olarak bulunmuştur (Sakaluk ve ark., 2019). Bu sonuçlar, APDKPT'nin kanıta dayalı psikoterapi olarak nitelendirdiği pek çok psikoterapi yönteminin mevcut bilimsel standartları sağlayamayan çalışmalar ile destekleniyor olduğunu göstermektedir.

Düşük Örneklem Sayısına Yönelik Öneriler

Düşük örneklem sayısı problemini düzeltmeye yönelik en başta gelen öneri, araştırmaların örneklem sayılarının araştırmanın tasarlanması aşamasında istatistiksel güç analizi yapılarak belirlenmesidir (Button ve ark., 2013). Günümüzde bu amaçla kullanılacak pek çok bilgisayar yazılımı bulunmaktadır. Araştırmacılar bu yazılımları kullanarak araştırmalarının en az .80 istatistiksel güce sahip olması için hangi büyüklükte bir örnekleme sahip olması gerektiğini kolayca hesaplayabilirler. Bu hesaplama yapılırken dikkat edilmesi gereken önemli bir nokta, popülasyondaki tahmini etki büyüklüğünün (alanyazındaki düşük istatistiksel güce sahip çalışmalar sebebiyle) gerçek etki büyüklüğünün şişirilmiş bir hali olduğudur. Yani genelde meta-analizlerden yola çıkılarak tahmin edilen popülasyondaki etki büyüklüğünün gerçekte görülenden daha düşük bir değer olduğu varsayılmalıdır (Ferguson, 2009). Bu durum yeterli istatistiksel güce ulaşmak için gereken örneklem sayısının daha büyük bir değere denk geleceği anlamına

gelmektedir.

Pek çok özel popülasyonda (örn., az görülen psikolojik rahatsızlıklar) yeterli katılımcı sayısına ulaşmak ciddi bir zorluk yaratmaktadır. Böyle durumlarda farklı araştırma gruplarının bağımsız araştırmalar yerine güçlerini birleştirdikleri ortak çalışmalar yapmalarıyla tek başlarına ulaşamayacakları katılımcı sayılarına ulaşmaları sağlanabilir. Hansen ve Collins (1994) örneklem büyüklüğünün sabit tutulduğu durumlarda istatistiksel gücün azalmaması için katılımcıların çalışmada kalmasının sağlanmasını veya çalışmayı bırakan katılımcıların eksik verilerinin istatistiksel yöntemler kullanılarak tahmin edilmesini önermişlerdir. Son olarak, araştırma hipotezlerinin ve onları test etmek için yapılan analizlerin karmaşık hale gelmesi istatistiksel gücü düşüren bir etki göstermektedir. Bu sebeple araştırmacılar analizleri, etkileşim etkileri yerine ana etkileri hedefleyecek şekilde gerçekleştirerek ve araştırma tasarımlarını olabildiğince basit tutarak düşük örneklem sayısının istatistiksel gücü azaltıcı etkisini hafifletebilirler (Abraham ve Russell, 2008).

Bağlılık Yanlılığı

Psikoterapi araştırmaları bağlamında bağlılık yanlılığı, psikoterapi etkililik çalışmasının yürütücüleri arasında test edilen psikoterapi yöntemi ile yakından ilgisi ve bağlı bulunan kişilerin bulunmasının sonuçları psikoterapinin lehine olacak şekilde etkilemesidir (Luborsky ve ark., 1999; Thase, 1999). Leichsenring ve arkadaşları (2017) bağlılık yanlılığının sadece araştırmacılık değil psikoterapistlik, süpervizörlük, dergi hakemliği ve editörlük gibi konularda da görülebileceğini belirtmişlerdir. Leykin ve DeRubeis (2009) bağlılık yanlılığının potansiyel kaynaklarını araştırmacının eğitiminin yetersizliği, araştırmacının araştırılan terapi yönteminin etkililiği konusunda aşırı hevesli olması, bulunan olumsuz sonuçların yayınlanmaması ve araştırmacının araştırmada uzmanlığı olmayan bir terapi yöntemini kullanması sebebiyle o terapinin tedavi bütünlüğünü düşürmesi olarak listelemişlerdir. Bağlılık yanlılığı etkisini seçici sonuç değişkeni kullanımı ya da rakip psikoterapi yönteminin daha zayıf şekilde uygulanması (tedavi bütünlüğünün bozulması) gibi durumlar üzerinden gösterebilmektedir (Leichsenring ve ark., 2017; Leykin ve DeRubeis, 2009; Luborsky ve ark., 1999; Shedler, 2018). Örneğin, iki farklı psikoterapi yaklaşımının etkilerinin kıyaslandığı bir çalışmada, araştırmacılar kendi bağlı oldukları psikoterapi yöntemini uygulayacak psikoterapistlere, rakip psikoterapiyi uygulayacak psikoterapistlerden daha iyi bir eğitim veya süpervizyon verebilirler ve böylece çalışma sonucunda bulunan fark aslında psikoterapistlere verilen yanlı eğitimlerden kaynaklı olabilir (Leykin ve DeRubeis, 2009; Shedler, 2018). Bu du-

rumda aslında iki eşit terapi yönteminin kıyaslanması yerine bir terapi yöntemi, zayıflatılmış diğer bir terapi yöntemi ile kıyaslanmaktadır. Alanyazında bu zayıflatılmış terapi yöntemine *başarısız olması amaçlanmış tedavi* denmektedir (Leichsenring ve ark., 2017). Munder ve arkadaşları (2013) psikoterapi sonuç araştırmalarına dair 30 meta-analiz çalışmasını inceledikleri meta-meta-analiz çalışmasında araştırmacı bağıllığı ile sonuç değişkeni arasında pozitif ve orta güçte ($r = .26$) bir ilişki bulmuşlardır. Bu ilişki-sel etki büyüklüğünün d-etki büyüklüğüne çevirildiği durumda ise bağıllık yanlılığı içeren bir çalışmalı ile içermeyen çalışma arasında, bağıllık yanlılığı lehinde .54'lük bir etki büyüklüğü farkı ortaya çıkmaktadır.

Bağıllık Yanlılığına Yönelik Öneriler

Leichsenring ve arkadaşları (2017) bağıllık yanlılığının önüne geçmek için analizlerin bağımsız araştırmacılar tarafından kör şekilde gerçekleştirilmesi, başka teorik yaklaşımlardan araştırmacılarla iş birliği yapılması ve çalışma kayıt protokollerinin bağıllık yanlılığını tespit edebilecek şekilde güncellenmesi gibi önerilerde bulunmuşlardır. Bağıllık yanlılığının potansiyel kaynakları düşünüldüğünde araştırmacıların araştırma yöntemleri ve etiği açısından iyi bir şekilde yetiştirilmelerinin önemli olduğu görülmektedir. Luborsky ve arkadaşları (1999) çalışmalarda araştırmacıların bağıllık yanlılığını değerlendiren ölçeklerin kullanılmasını ve böylece yanlılığın sonuç değişkeni üzerindeki etkisinin kontrol edilmesini önermektedirler. Thase (1999) başka teorik yaklaşımlardan araştırmacılar ile iş birliği yapılması önerisinin yeterince uygulanabilir olmadığını belirtmiştir. Öte yandan diğer teorik yaklaşımlardan araştırmacıların, araştırma sürecinin tamamına dahil olmaksızın, ana araştırmacıların bağlı olmadığı psikoterapi yönteminin nasıl uygulanacağını tasarlanması ve tedavi bütünlüğünün nasıl sağlanacağını belirlemesi ile ilgili prosedürlere katkı sağlaması bağıllık yanlılığının kontrol altına alınmasına yardımcı olabilir (Luborsky ve ark., 1999). Son olarak meta-analizler alanyazındaki bağıllık yanlılığının etkisinin büyüklüğünün tahmin edilmesi için kullanılabilir (Thase, 1999). Ancak bu kontrol geç bir aşamada yapıldığı için tedavi bütünlüğünde bağıllık yanlılığından kaynaklı olarak ortaya çıkan problemleri tam anlamıyla engellemeyecektir.

Seçici Sonuç Raporlama

Seçici sonuç raporlama, araştırmacıların çalışmalarındaki çeşitli sonuç değişkenlerinden sadece olumlu ya da anlamlı sonuçlar gösterenleri raporladıkları bir yanlılıktır (Bradley ve ark., 2017). Örneğin, bir çalışmada katılımcıların depresyon seviyelerini değerlendirmek için Beck Depresyon Ölçeği, Hastane Anksiyete Depresyon Ölçeği ve

Hamilton Depresyon Derecelendirme Ölçeği kullandığında, bunlardan sadece istatistiksel olarak anlamlı sonuç veren ölçeklerin sonuç değişkeni olarak raporlanması seçici sonuç raporlama olarak değerlendirilmektedir. Shinohara ve arkadaşları (2015), psikiyatri alanında depresyon tedavilerini (BDT ve antidepresan tedavisi) değerlendirmek için yapılan 170 çalışmayı ön kayıt (çalışmanın yöntemsel özellikleri ve prosedürlerinin çalışma başlanmadan önce herkese açık bir platforma kayıt edilmesi) ve seçici sonuç raporlama oranları açısından incelemiştir. Çalışmaların sadece %54.1'nin ön kayıt içerdiğini ve bu çalışmalardan sadece %18.8'inde ön kayıt ve raporlanan birincil sonuç değişkenlerinin uyumlu olduğunu bulmuşlardır (Shinohara ve ark., 2015). Benzer şekilde Bradley ve arkadaşları (2017), 112 seçkisiz kontrollü psikoterapi çalışmasını inceledikleri analizlerinde, çalışmalardan sadece 13 tanesinin ön kaydının olduğunu ve birincil sonuç ölçümlerinin net bir şekilde belirtildiğini bulmuşlardır. Bu 13 çalışmadan ancak yedi tanesinde seçici sonuç raporlama gerçekleştirildiğini ve dört çalışmada bu değişikliklerin klinik olarak anlamlı sonuçlar doğuracak yönde yapıldığını göstermişlerdir (Bradley ve ark., 2017). Araştırmacıların seçici sonuç raporlamayı ön kaydı olan çalışmalarda bile gerçekleştirmeleri dikkat edilmesi gereken önemli bir noktadır. Miguel ve arkadaşları (2021), 2005 yılından beri yapılan ve psikoterapinin yetişkin depresyonu üzerindeki etkisini değerlendiren 185 klinik çalışmayı inceledikleri meta-analizde, etki büyüklüğünün seçici sonuç raporlama yapılan çalışmalar dahil edildiğinde .81, bu çalışmalar çıkarıldığında ise .54 olduğunu bulmuşlardır. Yani bu yanlışlık, bulunan etkilerin olduğundan daha şişirilmiş görünmesine ve böylece tedavi anlamında yapılacak tercihlerin yanlış yönlendirilmesine sebep olabilmektedir (Kirkham ve ark., 2010).

Seçici Sonuç Raporlamaya Yönelik Öneriler

Seçici sonuç raporlamanın önüne geçmek için başvuru birincil yöntem ön kayıt sistemidir (Button ve ark., 2013; Tackett ve ark., 2017). Araştırmacılar araştırmalarını gerçekleştirmeden önce hipotezlerini, çalışmanın prosedürünü ve örneklemini, kullanılacak ölçüm araçlarını ve verilere uygulanacak analizleri belirleyerek ön kayıt platformlarına kaydederler. Ancak Bradley ve arkadaşlarının (2017) ortaya koyduğu gibi ön kayda rağmen yayınlarda seçici sonuç raporlamanın gerçekleştiği görülmektedir. Bu sebeple Bradley ve arkadaşları (2017) dergi editörlerinin, çalışmanın yayına kabulünden önce makalenin ön kayıt ile uyumlu olup olmadığını kontrol etmesi, fon sağlayan ajansların sadece ön kayıt ile uyumlu şekilde yapılan çalışmalarını desteklemesi ve meslek topluluklarının ön kayıt protokollerine bağlılığı cesaretlendirecek hamleler atması

gerektiğini belirtmişlerdir. Ön kayda rağmen gerçekleştirilen seçici sonuç raporlama durumu etik bir problem olarak değerlendirilebilir. Çünkü araştırmacılar ön kayıt yaparak bir anlamda kayıtlara bağlı kalacaklarına yönelik bir söz vermektedirler. Bunun yaşandığı durumlarda araştırmacıların bir süreliğine yayın kara listelerine alınmaları, bu sorunun önüne geçmeye yardımcı olabilir.

Örnekleme Dahil Etme ya da Dışarıda Bırakma Kriterlerinin Katılığı

SKÇ'lerde örneklemin homojenliğinin sağlanması ve böylece iç geçerliliğin artırılması için, çalışmaya başvuran kişiler dışarıda bırakma kriterlerine göre değerlendirilir ve bu kriterlere uymayan katılımcılar çalışmaya dahil edilmez (Shean, 2016). Örneğin depresyon için bir psikoterapinin etkililiğini değerlendiren bir çalışmanın örneklemini oluşturulurken, başvuran kişilerin içinden depresyon tanısının yanında kişilik bozukluğu tanısı olan kişiler çıkarılmaktadır. Böylece çalışmanın son örneklemini ağırlıklı olarak ya da sadece depresyon tanısı olan katılımcılardan oluşmaktadır. Bu durum çalışmaya dahil edilen katılımcıların gerçek hayatta kliniklere başvuran kişilerden yaş, cinsiyet, etnik kimlik, sosyo-ekonomik statü gibi demografik özellikler ya da eş tanılar ve hastalığın şiddeti gibi pek çok açıdan farklı olmalarına neden olmaktadır (Fensterheim ve Raw, 1996; Hotopf, 2002; Persons ve Silberschatz, 1998). Klinik uygulamada tedavi almak için gelen danışanların çoğu zaman birden fazla psikolojik rahatsızlık tanısı olduğu bilinmektedir (Westen ve ark., 2004). Aynı zamanda SKÇ'lerde test edilen psikoterapiler ile klinikte uygulanan psikoterapiler arasında da seans sıklığı, psikoterapi süresi, el kitabına katı şekilde uyulması ve psikoterapistlerin becerileri gibi alanlarda ciddi farklar olduğu belirtilmektedir (Addis ve Krasnow, 2000; Goldfried ve Wolfe, 1998; Persons, 1991). Pek çok araştırmacı bu sebeplerle SKÇ bulgularının dış geçerlilik açısından problemliliğini iddia etmektedir (Hotopf, 2002; Morrison ve ark., 2003; Mulder ve ark., 2018; Persons, 1991; Shean, 2012). SKÇ'lerin dışarıda bırakma kriterlerinin, genelde hastalığın şiddetini de etkileyen etmenleri (örn., eş hastalıklar, kişilik bozuklukları, madde kullanımı, intihar riski) içermeleri sebebiyle hem sonuçların dış geçerliliğini azalttığı hem de türdeş katılımcılar ile bulunan etkilerin gerçek hayattaki etkilerden daha büyükmüş gibi gözükmesine neden olduğu belirtilmektedir (Shedler, 2018; Westen ve Morrison, 2001). Westen ve arkadaşlarının (2004) bulgularına göre gerçek dünyadaki hastaların yaklaşık %40-70'i psikoterapi etkililik çalışmalarının dışarıda bırakma kriterlerine takılmaktadır. Hoertel ve arkadaşları (2014), 34653 kişiyi kapsayan Birleşik Devletler örnekleminde gerçekleştirdikleri incelemede, sosyal anksiyete bozukluğu tanısı

olan her 10 kişiden altısının etkililik çalışmalarının dahil etme kriterlerine uymadığını rapor etmişlerdir. Aynı durum sınır kişilik bozukluğu hastaları için 10 kişiden beşinin uymadığı şeklindedir (Hoertel ve ark., 2015). Gerçek hayattaki kronik depresyon hastalarının SKÇ'lere dahil edilme kriterlerini sağlayıp sağlamadıklarını inceleyen von Wolff ve arkadaşları (2014), hastaların sadece %25'inin çalışmaların dahil etme kriterlerini sağlayabildiğini bulmuşlardır.

Örnekleme Dahil Etme ya da Dışarıda Bırakma Kriterlerinin Katılığına Yönelik Öneriler

SKÇ'lerde iç geçerliliği attırmak için örneklemin türdeşleştirilmesi durumunun ortaya çıkardığı dış geçerlilik sorununu azaltmak için pragmatik seçkisiz kontrollü çalışmaların (PSKÇ) kullanılabilceği belirtilmiştir (Hotopf, 2002; Lambert, 2013). PSKÇ'ler ile örneklemin dışında bırakma kriterleri en aza indirilir, katılımcılar klinik uygulamada tedaviye başvuran kişilerden seçilir, psikoterapistler danışanların ihtiyacına göre müdahalelerini esnetebilirler ve sonuç ölçümleri belirtilerden ziyade gerçek hayatta anlamı olan işlevsellik ile ilişkili (örn., işe dönme, hastane ziyaretinin sıklığı ve intihar oranları) ölçümleri içerir (Hotopf, 2002). Ancak PSKÇ'ler SKÇ'lere göre daha yüksek dışsal geçerliliğe sahip olsa da pek çok karıştırıcı etkinin işin içine girmesi sebebiyle iç geçerlilik açısından SKÇ'lerin gerisinde kalmaktadırlar.

Kontrol Grubu Olarak Bekleme Listesi Kullanımı

Kontrol grubu olarak bekleme listesinde olan kişilerin kullanılması psikoterapi etkililik çalışmalarında sık rastlanan bir uygulamadır. Cuijpers ve arkadaşları (2016), BDT'nin depresyon ve çeşitli kaygı bozukluklarındaki etkisini değerlendiren 144 çalışmayı meta-analiz yöntemi ile incelemiştir. Kaygı bozukluğu üzerine yapılan çalışmaların yaklaşık %80'inde, depresyon çalışmalarının ise %44.4'ünde kontrol grubu olarak işe yaradığı bilinen alternatif bir müdahale yerine bekleme listesi kullanıldığı ve bekleme listesinin etkisi kontrol edildiğinde bulunan etki büyüklüklerinin düştüğü gözlemlenmiştir (Cuijpers ve ark., 2016). Bekleme listesi kullanımı çalışmalarda gerçekte olduğundan daha büyük bir etki büyüklüğü bulunmasının nedenlerinden birisi olarak değerlendirilmektedir (Coyne ve Kok, 2014). Çünkü bekleme listesinde bulunmak no-sebo etkisi sebebiyle katılımcıların daha kötüye gitmesine (Furukawa ve ark., 2014) ya da bekleme süresi bitiminde tedavi alacaklarını düşünmelerinden dolayı problemleri için bir şeyler yapmaktan geri durmalarına neden olabilmektedir (Cuijpers ve ark., 2019). Sonuç olarak bekleme listesi grubunun ortalama sonuç puanları, kullanılabilcek

alternatif geçerli bir tedaviye göre daha düşük kalkmakta böylece incelenen psikoterapinin etkililiği olduğundan daha yüksek gözükülebilmektedir. Furukawa ve arkadaşlarının (2014) 49 SKÇ'yi dahil ettikleri ağ meta-analizine göre kontrol grubu olarak bekleme listesi kullanımı BDT'nin çalışmalardaki etki büyüklüğünü herhangi bir tedavi almama koşulunun kullanımından bile daha büyük göstermektedir. Öte yandan Munder ve arkadaşları (2019) bekleme listesinin bunun tam tersi bir etki gösterebileceğini, bekleme listesindeki katılımcıların sadece çalışmaya dahil edilmekten dolayı bile iyileşme yönünde ilerleyebileceklerini belirtmiş; Furukawa ve arkadaşlarının (2014) bekleme listesi ile nosebo etkisini ilişkilendirdiği meta-analizdeki kısıtlılıklara dikkat çekmiştir.

Kontrol Grubu Olarak Bekleme Listesi Kullanımına Yönelik Öneriler

Bu problemin önüne geçmek için araştırmacılar kontrol koşulu için gerçek tedavi olarak nitelendirilen ve alanda kabul gören başka bir psikoterapiyi ya da ilaç tedavisini kullanabilirler. Hiç tedavi almama koşulunun bekleme listesi koşulundan daha gerçekçi sonuçlar doğurduğuna dair bulgular, tedavi alama koşulununun bekleme listesinin bir alternatifi olabileceğine işaret etmektedir. Öte yandan Rifkin'e göre (2007) asıl önemli soru, bir psikoterapinin etkili olup olmadığından ziyade hali hazırda uygulanan bir tedaviye (örn., ilaç tedavisi) kıyasla etkili olup olmadığıdır. Kontrol koşulu olarak işe yaradığı kabul edilen bir tedavinin uygulanması bekleme listesi ya da tedavi almama koşullarına göre daha maliyetli olsa da bekleme listesindeki katılımcıları tedavisiz bırakmayacağı için etik anlamda daha doğru bir tercih olacaktır.

Seçkisizleştirme ve Körleştirme Sorunları

SKÇ'lerde katılımcıların farklı koşullara atanmaları seçkisiz bir şekilde gerçekleştirilir ve araştırmacılar araştırma süreçlerine olabildiğince kör bir şekilde dahil olmaya çalışırlar. Seçkisizleştirmenin gruplar arasındaki farkları kontrol edebilmesi için örneklem olabildiğince büyük olması gerekmektedir (Hsu, 1992; Shean, 2016). Çünkü düşük örneklem sayısına sahip bir çalışmada seçkisizleştirme yapılsa bile gruplar şans eseri uygun şekilde dağılmamış olabilir (Leichsenring ve ark., 2017; Wampold ve ark., 2011). Aynı zamanda körleme süreçlerinin sıkı tutulmadığı koşullarda seçkisizleştirme açısından problem yaratacak sonuçlar doğabilir. Örneğin, araştırmacılar farkında olmadan katılımcıları daha çok fayda göreceklarını düşündükleri koşullara atayabilirler (Cuijpers ve Cristea, 2016). Körleştirme ile ilgili diğer bir problem ise klinik sonuç değerlendirmelerinin kör değerlendiriciler tarafından yapılmaması sonucunda ortaya çıkmaktadır. Değerlendirici hangi koşulda olduğunu bildiği ve belli bir psikoterapiyi

alan kişinin sonuç puanlarını daha yüksek ya da düşük verme yanlılığı gösterebilir. Örneğin körlemede bir problem yaşandıysa ve çalışmanın sonuç ölçümü olarak klinisyen tarafından puanlanan Hamilton Depresyon Değerlendirme Ölçeği kullanılıyorsa, araştırmacı etkili olmasını beklediği psikoterapi grubundaki katılımcının depresyon puanını daha düşük değerlendirme yanlılığı gösterebilir. Cuijpers ve arkadaşları (2010) psikoterapilerin yetişkinlerdeki depresyon üzerindeki etkilerini inceledikleri meta-analizlerinde atamanın düzgün bir şekilde gizlendiği ve sonuç değişkenlerinin değerlendirilmesinin kör şekilde gerçekleştirildiği çalışmalarda bunların yapılmadığı çalışmalara oranla daha düşük etki büyüklükleri bulunduğunu raporlamışlardır.

Seçkisizleştirme ve Körleştirme Sorunlarına Yönelik Öneriler

Seçkisizleştirmenin istenilen etkiyi gösterebilmesi için örneklem sayılarının büyü-tülmesi gerekmektedir (Hsu, 1992; Wampold ve ark., 2011). Seçkisizleştirme ve körleştirmede yaşanan diğer sorunlar bu prosedürleri araştırmacılardan bağımsız ve süreçlere kör olan kişilerin yapması ile aşılabilecek problemlerdir (Miller ve Stewart, 2011). Aynı zamanda araştırmacılara araştırma yöntemleri ve etiği gibi konularda yeterli eğitim verilmesi bu sorunların önüne geçmeye yardımcı olacaktır.

Çalışmayı Bırakan Katılımcıların Analizlere Dahil Edilmemesi

Psikoterapi çalışmalarında bazı katılımcıların çalışma tamamlanmadan çalışmayı bırakmalarına sık bir şekilde rastlanmaktadır (Wierzbicki ve Pekarik, 1993). Cuijpers ve Cristea (2016) çalışmayı bırakan katılımcıların genelde tedaviye cevap vermeyen ya da yan etki deneyimleyen katılımcılar olduğunu ve bu kişilerin analizlere dahil edilmemelerinin, sonuçları psikoterapi müdahalesinin lehinde etkileyebileceğini belirtmişlerdir. Çünkü psikoterapiden daha az fayda görebilecek katılımcıların sonuç analizlerine dahil edilmemesi, grubun ortalama puanını yükselterek psikoterapinin etkisini artırdığından daha büyük gösterebilir. Fernandez ve arkadaşları (2015) 115 BDT çalışmasını dahil ederek gerçekleştirdikleri meta-analizde, tedavi öncesi ve tedavi sürecinde çalışmayı bırakan katılımcıların oranlarını sırası ile %15.9 ve %26.2 olarak bulmuşlardır. Benzer şekilde Cooper ve Conklin de (2015) depresyon için psikoterapinin etkililiğini inceleyen 54 çalışmayı dahil ettikleri analizlerinde, çalışmayı bırakan katılımcı oranını %17.5 olarak rapor etmişlerdir. Bu bulgular çalışmayı bırakma yüzdelerinin, çalışmaların sonuçlarını anlamlı olarak etkileyebilecek büyüklükte olduğunu göstermektedir.

Çalışmayı Bırakan Katılımcıların Analizlere Dahil Edilmemesine Yönelik Öneriler

Araştırmacılar çalışmayı bırakan katılımcıların sonuç puanlarının olmaması sebebiyle bu puanları tahmin etmeye yarayan istatistiksel yöntemlerin kullanılmasını önermektedirler (Crameri ve ark., 2015; Cuijpers ve Cristea, 2016). Ayrıca, araştırmalar tasarlanırken katılımcıların çalışmayı yarım bırakmasına sebep olabilecek potansiyel etmenler düşünülerek bunların önüne geçmek için çeşitli çözümler geliştirilebilir. Bu etmenlerin neler olduğunu öğrenmek için çalışmayı bırakan katılımcılar ile daha sonradan iletişime geçilerek detaylı bilgiler alınabilir.

Psikoterapist Etkilerinin İhmal Edilmesi

Bazı araştırmacılar bir psikoterapinin belli bir rahatsızlık üzerindeki etkisini incelerken psikoterapistlerin bu sürece olan etkilerinin de değerlendirilmesi gerektiğini vurgulamışlardır (Del Re ve ark., 2012; Johns ve ark., 2019; Leichsenring ve ark., 2017; Owen ve ark., 2015). Tracey ve arkadaşları (2014) yetkin ve etkili terapistlerin anlamlı örüntüleri daha iyi tespit edebildiklerini, edindikleri bilgileri hızlı ve doğru şekilde işleyebildiklerini, uzun ve kısa süreli hafızalarının daha iyi olduğunu, bir problemi daha derin anlamıyla ele alabildiklerini, problemleri anlamak ve analiz etmek için daha fazla vakit ayırdıklarını ve kendini gözlem becerilerinin daha gelişmiş olduğunu belirtmiştir. Her psikoterapistin belli bir psikoterapiyi aynı şekilde ve etkinlikte uygulayabileceğini varsaymak psikoterapistlerin kişilik özellikleri ve yeteneklerinin psikoterapi süreci üzerindeki etkisinin göz ardı edilmesine neden olmaktadır. Tedavi el kitapları psikoterapinin uygulanmasını olabildiğince standardize hale getirirse de psikoterapi teknikleri etkilerini psikoterapistin uygulaması üzerinden göstermektedir. Owen ve arkadaşlarına (2015) göre psikoterapistlerin sonuçlar üzerindeki etkileri, psikoterapistlerin farklı müdahale koşullarına seçkisiz bir şekilde atanmamaları ve danışan seviyesinde değerlendirilen verilerin etki büyüklüklerine ait standart hataların bağımsızlık varsayımını ihlal etmeleri sebebiyle yanlı olmaları gibi noktalarda belirgin hale gelmektedir. Araştırmacılar bu etkileri göstermek için gerçekleştirdikleri simülasyon çalışmasında psikoterapist etkisinin düşük tutulduğu durumda çalışmaların %20'sinde psikoterapinin etkisi açısından anlamlı sonuç bulunamadığını, psikoterapist etkisi yüksek tutulduğunda ise çalışmaların %80'inde anlamlı sonuç bulunamadığını belirtmişlerdir (Owen ve ark., 2015). Çok seviyeli modelleme kullanarak 19 çalışmayı analiz eden Johns ve arkadaşları (2019), psikoterapist etkisinin %2 ile %29 arasında değiştiğini, ağırlıklı ortalamanın tüm çalışmalar için %5 iken SKÇ'ler için %8.2 olduğunu raporlamışlardır. Diğer yön-

temsel problemlere benzer şekilde psikoterapist etkileri de tip 1 hata oranını arttırmakta ve tespit edilen etki büyüklüklerinin şişmesine neden olmaktadır (Leichsenring ve ark., 2017).

Psikoterapist Etkilerinin İhmal Edilmesi Yönelik Öneriler

Johns ve arkadaşları (2019) bu sorunun etkilerinin azaltılmasına yönelik bazı tavsiyelerde bulunmuşlardır. İlk olarak, psikoterapistlerin eğitimleri için yapılan seçmelerde ağırlıklı olarak akademik başarı kriterine odaklanması yerine, rol yapma gibi aktif yöntemler seçim sürecine dahil edilebilir. İkinci olarak, çalışmalarda daha etkili terapistlerin daha zor vakalar ile eşleştirilmesiyle bir denge kurulması sağlanabilir. Üçüncü olarak, psikoterapist etkilerini değerlendirmeye yarayan yöntemlerin kullanımı süpervizyon süreçlerinde verilen geri bildirimlerin güvenilirliğini arttırabilir. Dördüncü olarak, etkili psikoterapistler süpervizyon görevleri üslenebilir ya da seanslarının kayıtlarını akranları ile paylaşabilirler. Son olarak psikolojik tedavi kılavuzlarına psikoterapistlerin bireysel etkilerine dikkat çeken bilgilerin eklenmesi danışanlara psikoterapistleri ile kuracakları terapi ilişkisinin psikoterapinin sonucunu etkileyen önemli bir faktör olduğunu hatırlatabilir (Johns ve ark., 2019). Psikoterapist etkisini tamamen ortadan kaldırmaya da azaltacak diğer bir yöntem psikoterapistlerin vakalara seçkisiz bir şekilde atanması olabilir (Owen ve ark., 2015). Son olarak alternatif bir araştırma yöntemi olarak çok seviyeli modelleme analizlerine başvurulabilir (Wampold ve ark., 2011). Bu yöntem psikoterapist etkisinin tedavi sonucu üzerindeki etkisinin ölçülmesine olanak sağlamaktadır.

Meta-Analizler ve Yöntemsel Problemleri

Geçtiğimiz yüzyılda yapılan bilimsel araştırmaların sayısının beklenmedik seviyede artmasıyla, araştırmacılar çok sayıda biriken bulgunun nasıl sentezleneceği üzerine kafa yormaya başlamışlardır (O’rourke, 2007). Bu ihtiyaç sonucu 1970’li yıllarda ortaya çıkan meta-analiz yöntemi, belli sayıda çalışmanın sonuçlarını nicel olarak bir araya getirerek sonuçların özetini sunmaya yarayan istatistiksel bir yöntemdir (Berkeljon ve Baldwin, 2009; Glass, 1982; Smith ve Glass, 1977). Meta-analiz bir yandan çok sayıda araştırma bulgusunun derli toplu bir sentezini ortaya koysa da diğer yandan bazı kısıtlılıklar içerebilmektedir. İlk olarak, meta-analiz yöntemini kullanmak ciddi bir çaba ve istatistiksel uzmanlık gerektirmektedir (Shelby ve Vaske, 2008). Çok sayıda yayın farklı veritabanları üzerinden sistematik olarak tarananak meta-analize uygunluk açısından değerlendirilmekte, uygun bulunan çalışmalardan gerekli veriler çıkarılarak meta-analize dahil edilmektedir (Berkeljon ve Baldwin, 2009). Ancak günümüzde analiz sürecini

kolaylaştıran bilgisayar programlarının ortaya çıkması bu konuda ciddi bir kolaylık yaratmıştır. İkinci olarak, sadece nicel sonuçların bir araya getirilmesi araştırmalardaki nitel sonuçların göz ardı edilmesine sebep olabilmektedir (Berkeljon ve Baldwin, 2009). Meta-analiz yöntemi kabaca nicel sonuçların ortalamasının alınmasına benzediği için nitel sonuçların bir araya getirilip değerlendirilmesine olanak sağlamamaktadır. Üçüncü olarak, farklı yöntemlere ve yöntemsel kalitelere sahip çalışmaların bir araya getirilmesi bir yanlılık ve hata kaynağı olabilmektedir (Berkeljon ve Baldwin, 2009; Shelby ve Vaske, 2008). Ancak meta-analiz, çalışmaların yöntem ve kalitelerindeki farkların bir değişken olarak analizlere dahil edilmesine ve etkilerinin belirlenebilmesine olanak sağlamaktadır (Shelby ve Vaske, 2008). Ayrıca meta-analizlere sadece belli bir yöntem ve yüksek kaliteye sahip çalışmalar dahil edilerek bu sorun aşılabilmektedir (Berkeljon ve Baldwin, 2009). Son olarak, bir konuda sadece yayınlanmış çalışmaların meta-analizlere dahil edilmesi, sonuçları olduğundan daha büyük gösteren temel bir probleme sebep olmaktadır. Çünkü alanyazında sonuçları olumlu olan çalışmaların daha çok yayınlandığına yönelik bir yanlılık olduğu bilinmektedir (Fanelli, 2011). Bu potansiyel problem çeşitli istatistiksel düzeltme yöntemlerinin kullanımı ya da yayınlanmamış çalışmaların taranması ile aşılabilmektedir (Berkeljon ve Baldwin, 2009; Shelby ve Vaske, 2008; Vevea ve Woods, 2005). Günümüzde meta-analiz, artıları ve eksileri ile beraber çok sayıda çalışmanın bulgularını bir araya getirmede standart olarak kabul edilen bir yöntem haline gelmiştir.

Yapılan ilk psikoterapi etkililik çalışmasından bu zamana psikoterapilerin etkililiğini değerlendiren çalışmaların sayısındaki ciddi artış, çalışmaların etkilerini toplu olarak değerlendirebilen meta-analiz yönteminin daha sık kullanılmasına yol açmıştır (Budd ve Hughes, 2009). Fakat yapılan meta-analizlerin artışı da birbiri ile çelişen pek çok meta-analiz bulgusunun (Cristea ve ark., 2017; Friborg ve Johnsen, 2017; Johnsen ve Friborg, 2015; Leichsenring ve Rabung, 2011; Smit ve ark., 2012; Tolin, 2010) ve meta-analiz çalışmalarının güvenilirliği ile ilgili bazı soru işaretlerinin ortaya çıkmasına sebep olmuştur (Coyne ve Kok, 2014; Hengartner, 2018). Bir meta-analiz çalışmasının ortaya koyduğu sonuçlar, büyük ölçüde analize hangi çalışmaların dahil edildiğinden ve bu çalışmaların hangi kriterlere göre analiz edildiğinden etkilenmektedir (González-Blanch ve Carral-Fernández, 2017). Psikoterapi etkililik çalışmaları bağlamında meta-analiz bulgularını olumsuz bir şekilde etkileyen başlıca etmenlere örnek olarak yayın yanlılığı ve düşük kaliteli çalışmalar verilebilir.

Yayın Yanlılığı

Bir meta-analize dahil edilecek çalışmalar, pek çok farklı kaynağın taranması sonucu ulaşılan çalışmaların arasından dahil etme kriterlerini sağlayan çalışmaların seçilmesi ile belirlenir (Berkeljon ve Baldwin, 2009). *Yayın yanlılığı* kavramı, alanyazındaki yayınlanmış çalışmaların, çoğunlukla olumlu sonuçlar gösteren çalışmalardan oluştuğunu ifade etmektedir (Sohn, 1996). Yapılan analizler bilimsel dergilerin olumlu sonuçlar ortaya koyan çalışmaları, olumsuz sonuçlar bulan çalışmalara kıyasla daha çok yayınladığını açığa çıkarmıştır (Easterbrook ve ark., 1991; Fanelli, 2011). Olumsuz sonuç bulduğu için yayınlanmayan çalışmaların ortaya çıkardığı durum alanyazında *çekmece-de kalan çalışmalar* olarak adlandırılmıştır (Coyne ve Kok, 2014). Eğer bir meta-analize dahil edilen çalışmalar, ağırlıklı olarak yayınlanan çalışmalardan oluşuyorsa meta-analiz bulduğu etki büyüklüğünün gerçekte olan etki büyüklüğünden daha büyük olması beklenir (Fanelli, 2011; Hopewell ve ark. 2005).

Yayın yanlılığının bulimiya nervoza, şizofreni ve sınır kişilik bozukluğu için geliştirilen psikoterapilerin etkililiğini değerlendiren çalışmalarda mevcut olduğu belirtilmektedir (Fonagy ve ark., 2017; Linardon ve ark., 2019; Niemeyer ve ark., 2012). Flint ve arkadaşlarının (2015) psikoterapinin depresyon üzerindeki etkisini inceleyen 149 çalışmayı değerlendirdikleri meta-analiz çalışmasında, çalışmaların etki büyüklüğünü $d = .55$ ve ortalama istatistiksel güçlerini ise $.49$ olarak bulmuşlardır. Sadece BDT'leri içeren 92 çalışmanın etki büyüklüğü ise $d = .58$ olarak bulunmuştur. Bulgulardan yola çıkarak yapılan hesaplamalar 104 çalışmanın istatistiksel olarak anlamlı ve olumlu sonuç vermesi gerektiğini gösterirken, mevcut durumda bu sayının 123 olduğu bulunmuştur (Flint ve ark., 2015). Yayın yanlılığına yönelik yapılan düzeltmeler meta-analizlerde bulunan ilk etki büyüklüklerini ciddi anlamda düşürebilmektedir. Driessen ve arkadaşları (2015) Amerika Birleşik Devletleri Milli Sağlık Enstitüsü tarafından fonlanan ve psikolojik tedavilerin depresyondaki etkililiğini inceleyen 55 çalışmanın, %23.6'sının yayınlanmadığını rapor etmişlerdir. Bu yayınlanmayan çalışmalara yönelik yapılan düzeltme, meta-analizde bulunan etki büyüklüğünün $g = .52$ 'den, $g = .39$ 'a düştüğünü göstermiştir. Dahası, araştırmacılar bulunan orta seviyedeki etkinin, seçici sonuç raporlama durumu kontrol edildiğinde daha da düşeceğini belirtmişlerdir (Driessen ve ark., 2015).

Yayın Yanlılığına Yönelik Öneriler

Yayın yanlılığının önüne geçebilmek için yöntemsel olarak güçlü ancak olumsuz sonuç bulan çalışmaların alanyazındaki temsil gücünün artırılması gerekmektedir. Araş-

tirmacılar bunu sağlamak için başvurulacak pek çok yol olduğunu belirtmektedir. Carroll ve arkadaşları (2017) alanda ortaya atılmış dokuz yoldan hangilerinin uygulanabilir olduğunu dergi editörlerine ve akademisyenlere sordukları çalışmalarında, en uygulanabilir çarelerin zorunlu yayın (sonuçlardan bağımsız olarak çalışmanın yayınlanması hakkı), iki aşamalı değerlendirme (değerlendirme için yayınevine önce giriş ve yöntem bölümlerinin gönderilmesi ve kabul durumunu takiben bulgular ve sonuç bölümlerinin gönderilmesi), olumsuz sonuç dergileri (olumsuz sonuç bulan çalışmaları yayınlayan dergilerin kurulması) ve araştırma kaydı (ön kayıt) olduğunu rapor etmişlerdir. Yani, yayın yanlılığının önüne geçmek için öncelikli olarak mevcut yayın yapma ve dergicilik paradigmasının değiştirilmesi gerekmektedir. Meta-analize dahil edilecek çalışmalar özelindeki diğer bir çare, kar amacı güden yayıncıların kontrolünde olmayan ve *gri alanyazın* adı verilen yayınlanmamış çalışmaların da bulunduğu veritabanlarının taranmasıdır. Gri alanyazın içinde lisanüstü tezlerini, konferans sunumlarını, klinik çalışmaları, araştırma kayıtlarını ve yayınlanmamış makaleleri barındırmaktadır (Rothstein ve Hopewell, 2009). Gri alanyazındaki çalışmaların genelde olumsuz sonuç bulan çalışmaları içerdiği düşünüldüğünde, bu çalışmaların meta-analize dahil edilmesi yayın yanlılığı etkisini azaltacak bir etki gösterecektir. Son olarak gri alanyazının taranmasından sonra bile yayın yanlılığı etkisi olduğu tespit edilmişse yayın yanlılığının meta-analiz sonuçları üzerindeki etkisinin kontrol etmek için *kırp ve doldur* ya da *ağırlık işlevi* modeli gibi istatistik yöntemler kullanılabilir (Vevea ve Woods, 2005).

Düşük Kaliteli Çalışmalar

Meta-analizleri etkileyen diğer başlıca problem düşük örneklem sayısına sahip, iyi tasarlanmamış ve yüksek yanlılık riski içeren çalışmaların meta-analiz hesaplamalarına dahil edilmesidir. Bu tarz çalışmalar, istatistiksel yetersizliklerinden dolayı gerçekte olandan daha büyük etki büyüklükleri ortaya çıkararak meta-analizlerin ortaya koyduğu etki büyüklüklerinin daha büyük gözükmesine neden olmaktadır (Coyne ve Kok, 2014; Cuijpers ve Cristea, 2016; Hengartner, 2018). Cuijpers ve arkadaşları (2016) BDT'nin çeşitli psikolojik rahatsızlıklardaki etkisini değerlendiren 144 çalışmayı dahil ettikleri meta-analizde, çalışmalardan sadece %17.4'ünün yüksek kaliteli sayılabilecek kriterleri sağladığını belirtmişlerdir. Daha güncel bir depresyon meta-analiz çalışmasında Cuijpers ve arkadaşları (2018) batı ülkelerinde yapılan çalışmaların sadece %20'sinin yanlılık riski açısından düşük seviyede olduğunu raporlamışlardır. Batı ülkelerinde yapılan 325 çalışmada, tüm psikoterapilerin genel etkisi $g = .63$ bulunurken, çeşitli yanlılıklar kontrol edildiğinde bu oranın $g = .31$ 'e düştüğü belirtilmiştir (Cuijpers ve ark., 2018).

Düşük Kaliteli Çalışmalara Yönelik Öneriler

Düşük kalitedeki çalışmaların alanyazındaki temsil gücünü azaltmak için daha önce de bahsedilen iki aşamalı değerlendirme ya da ön kayıt gibi çarelere başvurulabilir (Carroll ve ark., 2017; Leichsenring ve ark., 2017). Aslında düşük kalitedeki çalışmaların meta-analiz sonuçları üzerindeki etkisini azaltmanın en kolay yolu, bu çalışmaları meta-analize dahil etmemektir. Bu yol meta-analize dahil edilecek çalışma sayısını azalttığı için genelde göz ardı edilebilmektedir. Yine de meta-analize dahil edilen çalışmaları kalitelerine göre sınıflandırarak önce tüm çalışmalar, sonra da düşük kalitedeki çalışmalar çıkarıldıktan sonra kalan çalışmaların nasıl bir etki büyüklüğü ortaya çıkardığını görebilmek mümkündür. Çalışmalarının yüksek kalitede olmasını sağlamak için araştırmacıların araştırma yöntemleri ve etiği açısından iyi bir eğitim almış olmaları gerekmektedir. Araştırmacıların SKÇ'lerin yöntemsel problemleri bölümünde dile getirilen sorunların önüne geçecek şekilde tasarlanması ve yürütülmesi çalışmaların kalitesini yükselten bir etki gösterecektir.

Psikoterapi Çalışmalarının Sonuçlarının Yorumlanması

SKÇ'ler ve meta-analizlerin yorumlanması ve bulguların doğru bir şekilde değerlendirilmesi için daha önce bahsedilen yanlılıkların dışında dikkat edilmesi gereken başka önemli noktalar da bulunmaktadır. Bunlar etki büyüklüğünün anlamı, tedavinin faydasının uzun dönemde korunup korunmadığı, tedaviye yanıt oranları ve bulunan etkilerin gerçek dünyaya transferidir. Psikoterapi etkililik çalışmaları bağlamında etki büyüklüğü, müdahale grubu ve kontrol grubunun sonuç puanları arasındaki farkı gösteren istatistiksel bir değerdir (Kelley ve Preacher, 2012). Yani etki büyüklüğü bir psikoterapinin belli bir rahatsızlığı tedavi etmede kontrol olarak kullanılan seçenekten ne derece daha etkili olduğunu gösterir. Etki büyüklüğünün alanyazındaki en sık gösterimlerinden biri Cohen'in d 'sidir (Peng ve Chen, 2014). Bir etki büyüklüğünün Cohen'in d 'si olarak .50 bulunması gruptan birinin ortalama puanının diğer grubun ortalama puanından .50 standart sapma daha yüksek olduğu anlamına gelir. Bu farkın klinik uygulamadaki etkilerini netleştirebilmek için sonuç ölçüm değişkeni açısından kaç birime denk geldiğine dikkat edilmelidir. Örneğin, puan aralığı 0 ile 63 olan Beck Depresyon Ölçeği'ndeki 2 ya da 3 puanlık bir düşüş istatistiksel olarak anlamlı bir sonuç verse de bu durum klinik uygulama bağlamında anlamlı bir sonuç doğurmayabilir (Shedler, 2018). Wampold "Büyük Psikoterapi Tartışması" kitabında (2001) psikoterapi çalışmalarındaki etki büyüklüğünün istatistiksel olarak nasıl yorumlanacağına yönelik kullanışlı bir tablo sun-

maktadır. Örneğin, Arnberg ve Öst (2014) depresif belirtileri olan çocuklar için BDT'nin etkisini inceledikleri meta-analizde, BDT'nin etkisini $d = .66$ (yayın yanlılığı kontrol edildiğinde $d = 0.44$) olarak bulmuşlardır. Wampold'un tablosundaki "0.7" değerine bakıp yaklaşık olarak bir değerlendirme yaparsak bu etki büyüklüğü psikoterapi alan çocukların %75.8'nin psikoterapi almayan çocuklardan daha iyi bir durumda olacağı; sonuç değişkenindeki değişimin %10.9'unun BDT'den kaynaklı olduğu; BDT almayan çocukların tedavi başarısı oranının .33 olduğu ve BDT almanın bu başarı oranını .67'ye çıkaracağı gibi anlamlara gelmektedir (Wampold, 2001).

Psikoterapilerin ortaya çıkardığı kazanımların ne kadar süre geçerli olduğunu değerlendirmek için katılımcılardan tedavi bitiminden belli süreler (örn. üç, altı, 12 ay) sonra tekrar sonuç ölçümleri alınmaktadır. Van der Zwerde ve arkadaşları (2019) 30 tane uykusuzluk için BDT çalışmasının uzun süreli etkilerini inceledikleri meta-analiz çalışmasında, BDT'nin uykusuzluk şiddeti üzerindeki etkilerini çalışma bitiminden üç ay sonra $g = .64$, altı ay sonra $g = .40$ ve 12 ay sonra $g = .25$ olarak bulmuşlardır. Bu örnekte de görüldüğü gibi psikoterapilerin etkisi uzun vadede azalabilmektedir. Kazanımları hızlıca sönen bir psikoterapi, tedavi alan kişinin tekrar benzer sorunlar yaşamasına ve yeniden tedavi arayışına girmesine neden olabilir. Örneğin, Wojnarowski ve arkadaşları (2019) majör depresyon bozukluğu için BDT'lerdeki nüksetme oranını ortalama %33.4 olarak bulmuşlardır. Bu yüzden tedavi kazanımlarının sürekliliği ve nüksetme oranı belli bir psikoterapinin tedavi seçeneği olarak değerlendirilmesinde önemli bir rol oynayabilir.

SKÇ'lere başvuran tüm kişilerin çalışmalara dahil edilmediği düşünüldüğünde çalışmalara dahil edilen katılımcıların ne kadarlık bir kısmının iyileşiyor olduğu önemli bir soru olarak karşımıza çıkar (Leichsenring et al., 2018). Westen ve Bradley (2005) majör depresyon bozukluğu, kaygı bozukluğu, bulimiya nervoza ve travma sonrası stres bozukluğu gibi rahatsızlıklarda psikoterapilerin etkisini inceleyen altı meta-analiz çalışmasının sonuçlarını değerlendirmiş ve çalışmaya katılan katılımcılardan ancak yarısının iyileşme gösterdiğini bildirmişlerdir. Benzer şekilde Springer ve arkadaşları (2018) çeşitli kaygı bozukluklarında BDT'nin etkisini inceleyen 100 çalışmayı dahil ettikleri meta-analizlerinde, iyileşme oranlarını tüm rahatsızlıklar dahil edildiğinde %51 olarak rapor etmişlerdir. İyileşme oranları her rahatsızlığa ve uygulanan psikoterapiye göre değişiyor olsa bile çalışmaya dahil edilme oranları ile beraber değerlendirildiğinde SKÇ'lerde iyileşme gösterdiği belirtilen katılımcıların gerçek dünyadaki hastaların az bir

kısmına denk geldiği açıkça görülmekte ve bu durum ciddi bir dış geçerlilik sorunu yaratmaktadır (Shedler, 2018).

İç geçerliliği yükseltmeye odaklanarak gerçekleştirilen SKÇ'lerde elde edilen etkilerin klinik pratiğe transfer edilemediği cevaplanması gereken diğer önemli bir sorudur (Shean, 2016; Wampold ve ark., 2011). Gerçek dünya geçerliliğinin değerlendirilmesi için daha çok dış geçerliliğe odaklanan PSKÇ'lere başvurulmaktadır (Hotopf, 2002). Psikoterapi araştırmalarında ortaya çıkan önemli bir bulgu, SKÇ'lerde etkili bulunan psikoterapi yöntemlerinin klinik uygulamada de işe yaradığı ancak bulunan etki büyüklüklerinin SKÇ'lerdeki etki büyüklerinden daha küçük olduğudur (Craigie ve Nathan, 2009; Lambert, 2013; Wampold, 2019; Westbrook ve Kirk, 2005). Bu aslında beklenmedik bir bulgu değildir. Çünkü klinik uygulamada bir psikoterapinin uygulanmasında tedavi bütünlüğüne ne kadar uyulduğu bilinmemektedir. Buna ek olarak klinik uygulamada karşılaşılan danışanların tanılarının ve problemlerinin SKÇ'lere dahil edilen katılımcılarınkinden daha karmaşık bir tablo göstermesi bulunan etkinin büyüklüğünü düşürüyor olabilir (Craigie ve Nathan, 2009; Persons ve ark., 2006). Morrison ve arkadaşları (2003) psikoterapistler ile yaptıkları bir çalışmada, psikoterapi çalışmalarında etkili olduğu gösterilen süre kısıtlı psikoterapilerin klinik uygulamada ne kadar süre ile uygulandığını ve tedavinin uzunluğu ile danışanın ek tanıları arasındaki ilişkiyi incelemişlerdir. Klinik uygulamada BDT'nin ortalama olarak depresyon için 33 seans, panik bozukluk için 29 seans, diğer kaygı bozuklukları için 44 seans sürdüğü bulunmuştur. Ayrıca ek psikopatoloji ya da kişilik bozukluğu tanıları ile tedavi süresi arasında pozitif ilişkiler olduğu rapor edilmiştir (Morrison ve ark., 2003). Bu yüzden genelde 12-16 hafta süren klinik çalışmalarda bulunan etkinin aynı şekilde ve aynı süre zarfında klinik uygulamada da bulunabileceğini varsaymak yanlış olacaktır.

Dodo Kuşu Yargısı

Geçmiş yıllarda psikoterapilerin birbirlerine kıyasla ne kadar etkili olduklarını değerlendirmek için yapılan meta-analiz çalışmaları çelişen sonuçlar ortaya çıkarmıştır. Bu yüzden bazı araştırmacılar farklı psikoterapilerin etkililik anlamında aşağı yukarı aynı sonuçlar verdiğini iddia etmiş ve bu durumu "Alice Harikalar Diyarında" kitabına referans vererek *Dodo Kuşu Yargısı* olarak isimlendirmişlerdir (Baardseth ve ark., 2013; Luborsky ve ark., 2002; Luborsky ve ark., 1975; Smith ve Glass, 1977; Wampold ve ark., 1997). Lambert'e (2013) göre psikoterapiler arasında anlamlı farklar bulunamamasına bazı açıklamalar getirilebilir. Bunlardan ilki, farklı psikoterapiler farklı süreçler ile

benzer sonuçlara ulaşıyor olabilir. İkincisi, psikoterapiler arasındaki fark geçmişteki araştırma yöntemleri ile tespit edilemiyor olabilir. Üçüncüsü ise farklı psikoterapiler iyileşmeye sebep olan ortak etmenler içeriyor olabilir (Lambert, 2013). Öte yandan psikoterapilerin etkililik açısından anlamlı farklılıklar gösterdiğini belirten araştırmacılar da vardır (de Felice ve ark., 2019; Lilienfeld, 2014; Marcus ve ark., 2014; Shapiro ve Shapiro, 1982; Tolin, 2010). Lehide ve aleyhindeki kanıtlarla beraber *Dodo Kuşu Yargısı* hala çözüme ulaşmamış bir problem olarak klinik psikoloji alanındaki varlığını sürdürmektedir (Budd ve Hughes, 2009; Cuijpers ve ark., 2019). *Dodo Kuşu Yargısı* fenomeninin doğru olduğunu kabul eden bazı araştırmacılar psikoterapilerin terapötik ilişki, empati veya beklenti gibi iyileşmeye sebep olan ortak etmenler içermesini, psikoterapiler arasında bulunamayan farkların bir açıklaması olarak kabul etmişlerdir (Laska ve ark., 2014; Nahum ve ark., 2019; Wampold, 2015). Tekrarlama krizi bağlamında düşünüldüğünde, Lambert'in (2013) bahsettiği ikinci ihtimalin yani psikoterapilerin arasındaki farkın yöntemsel kısıtlılıklar (düşük kaliteli SKÇ'ler, onlar üzerinden gerçekleştirilen meta-analizler, alternatif araştırma yöntemlerinin kullanımında eksiklikler) sebebiyle tespit edilememesinin ciddi bir şekilde tekrar değerlendirilmesi gereken bir husus olduğu söylenebilir (Cuijpers ve ark., 2019). Yöntemsel esneklik ve sakıncalı araştırma uygulamalarının kullanımı ile herhangi bir psikoterapinin etkili olduğuna dair sonuçlar rahatlıkla bulunabilmektedir (Cuijpers ve Cristea, 2016). Etkili olduğu belirtilen psikoterapilerden hangilerinin yüksek kaliteli ve aynı zamanda, gerçek hayat etkilerini de değerlendiren çalışmalar ile tekrar sınıandıktan sonra hala, etkili olarak nitelendirilebileceği cevaplanması gereken önemli bir sorudur. Klinik psikoloji alanında farklı psikoterapiler arasında kayda değer farklar olmadığına yönelik bir algı olsa da, González-Blanch ve Carral-Fernández (2017) bu durumun yüzlerce psikoterapiden şimdiye kadar test edilebilmiş çok az bir kısmı için geçerli olduğunun unutulmaması gerektiğini vurgulamıştır.

Tartışma

Goldfried (2020), psikoterapi alanın yaklaşık olarak 100 yaşında olmasına rağmen emekleme dönemini yaşadığını ve hala sosyologların yeni gelişmelerin yanında bazı sağlam ve temel kabullerin olduğunu belirttikleri *olgun bilim* kategorisine giremediğini vurgulamaktadır. Psikoterapi alanı temel kabul ve ortaklaşmaların olduğu bir alandan ziyade, farklı teorik yaklaşımların varlıklarını sürdürmek ve kendi yöntemlerinin daha etkili olduğunu göstermek için uğraştıkları, çatışmalı ve rekabet içeren bir alandır (Budd

ve Hughes, 2009). Bu derleme makalesinde de görüldüğü gibi psikoterapilerin etkililiğini değerlendirmek için şimdiye kadar yapılmış SKÇ'ler ve meta-analizler pek çok yöntemsel problem ve yanlılık içermeleri sebebiyle elde edilen bulguların sağlıklı bir şekilde değerlendirilmesine izin vermemektedir. Meta-analizler bazen herkesin kazandığını gösterirken (Luborsky ve ark., 2002), bazen de herkesin kaybettiğine işaret etmektedir (Cuijpers ve ark., 2018). Bu açıdan bakıldığında aslında klinik psikolojinin tekrarlama krizinin merkezindeki alanlardan birisi olduğunu söylemek yanlış olmayacaktır (Tackett ve ark., 2017; 2019). Öte yandan psikoterapilerin etkililiğini değerlendiren SKÇ'lerin sayısı göz önünde bulundurulduğunda, psikoterapi alanında gerçekleştirilen tekrarlama çalışmalarının sayısının (istatistiksel güçleri düşük olsa bile) psikolojinin diğer alanlarına kıyasla çok daha fazla olduğu da göz çarpmaktadır.

Rubio-Aparicio ve arkadaşları (2018) 1285 çalışmayı içeren 54 meta-analiz çalışmasının sonuçlarını tekrar meta-analize sokarak psikolojik tedavilerin genel etki büyüklüğünü .75 olarak bulmuşlardır. Bu sebeple psikoterapinin çeşitli psikoloji rahatsızlıkları iyileştirmek için etkili bir tedavi yöntemi olduğu rahatlıkla dile getirebilir (Lambert, 2013; Rubio-Aparicio ve ark., 2018). Ancak bahsedilen tüm bu yöntemsel problem ve yanlılıklar dikkat edilmesi gereken iki nokta olduğuna işaret etmektedir. Bunlardan ilki, psikoterapilerin etkilerini göstermek için yapılan çalışmaların, güncel bilimsel standartları yakalamakta zorlandığı ve tekrarlama anlamında ciddi problemler içerdiğidir (Coyne ve Kok, 2014; Hengartner, 2018; Leichsenring ve ark., 2017; Sakaluk ve ark., 2019). İkincisi ise psikoterapilerin etki gücünün, geçmişte yapılan çalışmalarda hesaplanandan daha az olduğudur (Schäfer ve Schwarz, 2019). Örneğin, bazı yanlılık düzeltmeleri yapıldığında depresyon için psikoterapilerin etkililiğinin kabul edilen en alt sınırın biraz üstünde olduğuna dair sonuçlar çıkabilmektedir (Cuijpers ve ark., 2018). Cuijpers ve arkadaşları (2019) bekleme listesinin geçerli bir kontrol grubu sayılması durumunda bile psikoterapilerin etkililiğini düşüren diğer yanlılıkların, psikoterapi etkililik araştırmaları için ciddi problemler yaratmaya devam ettiğini belirtmişlerdir.

Alanyazında biriken tüm bilgilere rağmen psikoterapilerin etkililiklerinin yüksek istatistiksel güce sahip, iyi tasarlanmış, sakıncalı araştırma uygulamalarını içermeyen daha pek çok etkililik çalışması ile test edilmesi gerekmektedir. Öyle gözükmektedir ki araştırmacılar SKÇ'leri ve meta-analizleri ilerleyen zamanlarda da psikoterapilerin etkililiğini değerlendirmede birincil araçlar olarak kullanmaya devam edeceklerdir. Araştırmacılar ve klinisyenler bu alandaki eksikliklerin giderilmesi ve psikoterapi etkililik

çalışmalarının kalitelerinin artırılması için pek çok öneride bulunmaktadırlar. Bunlardan başlıcaları tedavilerin tanılardan ziyade sıkıntı veren deneyimleri ve problemleri davranışları azaltmayı hedeflemesi (Budd ve Hughes, 2009; Goldfried ve Wolfe, 1998); araştırma uygulamalarının, SKÇ'lerin ve meta-analizlerin kalitesinin artırılması için ciddi önlemler alınması (Coyne ve Kok, 2014; Leichsenring ve ark., 2017; Sakaluk ve ark., 2019; Tackett ve ark., 2017; 2019); araştırmacı-klinisyen iş birliğinin ve klinisyenlerin araştırma bulgularını değerlendirme yeteneklerinin artırılması (Goldfried ve Wolfe, 1998; González-Blanch ve Carral-Fernández, 2017) ve psikoterapilerin etkililiklerini gösteren çalışmalardan ziyade değişim mekanizmasını açıklayan çalışmaların yapılmasıdır (González-Blanch ve Carral-Fernández, 2017; Philips ve Falkenström, 2021; Westen ve ark., 2004). Aynı zamanda psikoterapilerin etkililiğinin değerlendirilmesi için SKÇ'lere alternatif ve ek olarak fenomenolojik yöntemlerin (Shean, 2016), natüralist/pragmatik çalışmaların (Lambert, 2013), panel verisi ve kohort dizaynlarının (Philips ve Falkenström, 2021) ve çok seviyeli istatistiksel modellerin (Wampold ve ark., 2011) kullanılabileceğine dikkat çekilmektedir.

Tüm artı ve eksilere rağmen SKÇ'ler ve meta-analizler psikoterapi araştırmalarında kullanılması gereken önemli araçlardır. Bu derleme makalesinin göstermeyi amaçladığı temel nokta tekrarlamaya krizi bağlamında psikoterapi araştırmalarında ortaya çıkan asıl büyük problemin, SKÇ'lere ve meta-analizlere özgül içsel problemlerden ziyade onların kaliteli bir şekilde uygulanmasında yaşanan sorunlardan kaynaklı olduğudur. SKÇ'lere ve meta-analizlere özgün içsel problemler, bu araçlara alternatif olarak önerilen diğer araştırma yöntemlerinin psikoterapi araştırma süreçlerine dahil edilmesi ile aşılabılır. Örneğin SKÇ'lerin düşük dış geçerliliği problemi ile bağlantılı olarak gerçekleştirilen PSKÇ'ler bir psikoterapinin yönteminin gerçek dünyadaki etkililiğinin gösterilmesine yardımcı olabilir (Lambert, 2013). Değişkenler arasındaki etkilerin seanslar boyunca nasıl değiştiğinin gözlemlenmesine imkân sağlayan *Çapraz gecikmeli panel veri modelleri* uzun vadeli süreçlerdeki karıştırıcı etkilerin kontrol edilmesine olanak tanımaktadır (Philips ve Falkenström, 2021). Çok seviyeli modelleme yöntemlerinin kullanılması danışan, psikoterapist ve psikoterapi seviyelerindeki değişkenlerin sonuç üzerindeki etkilerinin belirlenmesine katkıda bulunabilir (Wampold ve ark., 2011). Bu sebeplerle bir psikoterapinin kanıtı dayalı olarak nitelendirilmesinde SKÇ bulgularının yanında, alternatif araştırma araçlarından edinilen bulguların değerlendirilmesinin önemli bir yeri olmalıdır (Tolin ve ark., 2015). Öte yandan SKÇ'ler yapılırken ortaya çıkan ve tekrarlamaya

krizi tartışmalarının da işaret ettiği çalışmaların kalitesini düşüren yöntemsel sorunlar PSKÇ'lerin yürütülmesi sırasında da gerçekleşirse tek başına alternatif bir araştırma yöntemi kullanmanın uygulamada bir faydası olmayacaktır. Bu yüzden tekrarlama krizi bağlamında psikoterapi alanında ortaya çıkan sorunların önüne geçebilmek için öncelikli olarak SKÇ'lerin ve meta-analizlerin yöntemsel kalitelerinin artırılması ve bu yöntemlerin onların içsel problemlerine çözüm sunan alternatif araştırma yöntemleri ile beraber kullanılmaları gerekmektedir. Diğer bir deyişle eldeki araçları olumsuz olarak değerlendirmeden önce, o araçların kullanma kılavuzuna uygun olarak kullanılıp kullanılmadığı incelenmelidir. Bu derleme makalesi kapsamında ortaya konan bulgular araçların kullanma kılavuzuna uygun şekilde kullanılmadığına işaret etmektedir. Bu uygun kullanmama hali her zaman şartların imkansızlığından kaynaklanmadığı için ona neden olan diğer etmenlerin açığa çıkarılması için çaba sarf edilmesi gerekmektedir. Örneğin akademinin *yayınla ya da yok ol* kültürü ve kapitalist ilkelerle ilerlemeye çalışması mevcut yöntemsel krizin altında yatan en büyük nedenlerden bazıları olarak görülmektedir (Kiai, 2019; Münch, 2014; Stricker ve Günther, 2019).

Klinik psikoloji araştırma ve uygulama etkileşiminin yüksek olduğu bir alandır. Psikoterapi etkililik çalışmalarından elde edilen bulgular, gerçek hayatta psikolojik rahatsızlıkların tedavisine yönelik seçeneklerin oluşturulmasında ve bu bağlamda verilecek kararlarda kritik bir rol oynamaktadır. Bu durum belki de klinik psikologlara, psikolojinin diğer alanlarındaki psikologlara kıyasla daha fazla etik sorumluluk yüklemektedir (Tackett ve ark., 2017). Klinik psikolojinin bilimsanı-uygulayıcı modeline (Jones ve Mehr, 2007) daha sıkı sarılması ve bu modelin uygulanmasında yaşanan sorunları ortadan kaldırmak için çaba sarf etmesi uzun vadede yaşanan yöntemsel krizin üstesinden gelinmesine yardımcı olabilir. Ortaya çıkan mevcut tablo, klinik psikoloji araştırmacılarını ve klinisyenlerini tekrarlama krizinin alana olan etkilerini fark etmeye ve araştırmalarını güncel bilimsel standartlara yaklaştırmak için çaba sarf etmeye yönlendirmelidir (Tackett ve ark., 2019). Son olarak tekrarlama krizi ile ilgili tartışmaların başlaması ile beraber ortaya çıkan açık bilim/açık veri hareketleri ve ön kayıt sisteminin yaygınlaşması gibi gelişmeler hem psikoloji hem de klinik psikoloji ve psikoterapi alanları için umut verici gelişmeler olarak değerlendirilmelidir.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

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

Peer-review: Externally peer-reviewed.

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

Grant Support: The author declared that this study has received no financial support.

Kaynakça/References

- Abraham, W. T., ve Russell, D. W. (2008). Statistical power analysis in psychological research. *Social and Personality Psychology Compass*, 2(1), 283-301. <https://doi.org/10.1111/j.1751-9004.2007.00052.x>
- Addis, M. E. ve Krasnow, A. D. (2000). A national survey of practicing psychologists' attitudes toward psychotherapy treatment manuals. *Journal of Consulting and Clinical Psychology*, 68(2), 331-339. <https://doi.org/10.1037/0022-006X.68.2.331>
- APA Presidential Task Force on Evidence-Based Practice. (2006). Evidence-based practice in psychology. *The American Psychologist*, 61(4), 271-285. <https://doi.org/10.1037/0003-066X.61.4.271>
- Arnberg, A. ve Öst, L.-G. (2014). CBT for children with depressive symptoms: A meta-analysis. *Cognitive Behaviour Therapy*, 43(4), 275-288. <https://doi.org/10.1080/16506073.2014.947316>
- Baardseth, T. P., Goldberg, S. B., Pace, B. T., Wislocki, A. P., Frost, N. D., Siddiqui, J. R., Lindemann, A. M., Kivlighan, D. M. III, Laska, K. M., Del Re, A. C., Minami, T., & Wampold, B. E. (2013). Cognitive-behavioral therapy versus other therapies: Redux. *Clinical Psychology Review*, 33(3), 395-405. <https://doi.org/10.1016/j.cpr.2013.01.004>
- Berkeljon, A., & Baldwin, S. A. (2009). An introduction to meta-analysis for psychotherapy outcome research. *Psychotherapy Research*, 19(4-5), 511-518. <https://doi.org/10.1080/10503300802621172>
- Bradley, H. A., Rucklidge, J. J. ve Mulder, R. T. (2017). A systematic review of trial registration and selective outcome reporting in psychotherapy randomized controlled trials. *Acta Psychiatrica Scandinavica*, 135(1), 65-77. <https://doi.org/10.1111/acps.12647>
- Budd, R. ve Hughes, I. (2009). The Dodo Bird Verdict—controversial, inevitable and important: A commentary on 30 years of meta-analyses. *Clinical Psychology & Psychotherapy*, 16(6), 510-522. <https://doi.org/10.1002/cpp.648>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J. ve Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365-376. <https://doi.org/10.1038/nrn3475>
- Carroll, H. A., Toumpakari, Z., Johnson, L. ve Betts, J. A. (2017). The perceived feasibility of methods to reduce publication bias. *PLoS One*, 12(10), e0186472. <https://doi.org/10.1371/journal.pone.0186472>
- Chambless, D. L. ve Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 66(1), 7-18. <https://doi.org/10.1037/0022-006X.66.1.7>
- Christley, R. M. (2010). Power and Error: Increased Risk of False Positive Results in Underpowered Studies. *The Open Epidemiology Journal*, 3(1), 16-19. <http://dx.doi.org/10.2174/1874297101003010016>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037//0033-2909.112.1.155>
- Cooper, A. A. ve Conklin, L. R. (2015). Drop out from individual psychotherapy for major depression: A meta-analysis of randomized clinical trials. *Clinical Psychology Review*, 40, 57-65. <https://doi.org/10.1016/j.cpr.2015.05.001>
- Coyne, J. C. ve Kok, R. N. (2014). Salvaging psychotherapy research: A manifesto. *Journal of Evidence-Based Psychotherapies*, 14(2), 105-124.

- Craigie, M. A. ve Nathan, P. (2009). A nonrandomized effectiveness comparison of broad-spectrum group CBT to individual CBT for depressed outpatients in a community mental health setting. *Behavior therapy*, 40(3), 302-314. <https://doi.org/10.1016/j.beth.2008.08.002>
- Cramer, A., von Wyl, A., Koemeda, M., Schulthess, P. ve Tschuschke, V. (2015). Sensitivity analysis in multiple imputation in effectiveness studies of psychotherapy. *Frontiers in Psychology*, 6, 1042. <https://doi.org/10.3389/fpsyg.2015.01042>
- Cristea, I. A., Stefan, S., Karyotaki, E., David, D., Hollon, S. D. ve Cuijpers, P. (2017). The effects of cognitive behavioral therapy are not systematically falling: A revision of Johnsen and Friborg (2015). *Psychological Bulletin*, 143(3), 326-340. <https://doi.org/10.1037/bul0000062>
- Cuijpers, P. ve Cristea, I. A. (2016). How to prove that your therapy is effective, even when it is not: A guideline. *Epidemiology and Psychiatric Sciences*, 25(5), 428-435. <https://doi.org/10.1017/S2045796015000864>
- Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M. ve Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*, 15(3), 245-258. <https://doi.org/10.1002/wps.20346>
- Cuijpers, P., Karyotaki, E., Reijnders, M. ve Ebert, D. D. (2018). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences*, 28(1), 21-30. <https://doi.org/10.1017/S2045796018000057>
- Cuijpers, P., Karyotaki, E., Reijnders, M. ve Ebert, D. D. (2019). Is psychotherapy effective? Pretending everything is fine will not help the field forward. *Epidemiology and Psychiatric Sciences*, 28(3), 356-357. <https://doi.org/10.1017/S204579601800080X>
- Cuijpers, P., Reijnders, M. ve Huibers, M. J. H. (2019). The Role of Common Factors in Psychotherapy Outcomes. *Annual Review of Clinical Psychology*, 15(1), 207-231. <https://doi.org/10.1146/annurev-clinpsy-050718-095424>
- Cuijpers, P., van Straten, A., Bohlmeijer, E., Hollon, S. D., & Andersson, G. (2010). The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychological Medicine*, 40(2), 211-223. <https://doi.org/10.1017/S0033291709006114>
- de Felice, G., Giuliani, A., Halfon, S., Andreassi, S., Paoloni, G. ve Orsucci, F. F. (2019). The misleading Dodo Bird verdict. How much of the outcome variance is explained by common and specific factors? *New Ideas in Psychology*, 54, 50-55. <https://doi.org/10.1016/j.newideapsych.2019.01.006>
- Del Re, A. C., Flückiger, C., Horvath, A. O., Symonds, D. ve Wampold, B. E. (2012). Therapist effects in the therapeutic alliance-outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review*, 32(7), 642-649. <https://doi.org/10.1016/j.cpr.2012.07.002>
- Donnon, T. (2012). Experimental or RCT research designs: A crisis of nomenclature in medical education. *Canadian Medical Education Journal*, 3(2), 82-84.
- Driessen, E., Hollon, S. D., Bockting, C. L. H., Cuijpers, P. ve Turner, E. H. (2015). Does Publication Bias Inflate the Apparent Efficacy of Psychological Treatment for Major Depressive Disorder? A Systematic Review and Meta-Analysis of US National Institutes of Health-Funded Trials. *PLoS One*, 10(9), e0137864. <https://doi.org/10.1371/journal.pone.0137864>
- Dumas-Mallet, E., Button, K. S., Boraud, T., Gonon, F. ve Munafò, M. R. (2017). Low statistical power in biomedical science: A review of three human research domains. *Royal Society Open Science*, 4(2), 160254. <https://doi.org/10.1098/rsos.160254>
- Easterbrook, P. J., Gopalan, R., Berlin, J. A. ve Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 337(8746), 867-872. [https://doi.org/10.1016/0140-6736\(91\)90201-Y](https://doi.org/10.1016/0140-6736(91)90201-Y)
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. *Journal of Consulting Psychology*,

- 16(5), 319-324. <https://doi.org/10.1037/h0063633>
- Eysenck, H. J. (1964). The outcome problem in psychotherapy: A reply. *Psychotherapy: Theory, Research & Practice*, 1(3), 97-100. <https://doi.org/10.1037/h0088591>
- Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891-904. <https://doi.org/10.1007/s11192-011-0494-7>
- Faul, F., Erdfelder, E., Lang, A.-G., ve Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. <https://doi.org/10.3758/BF03193146>
- Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532-538. <https://doi.org/10.1037/a0015808>
- Fensterheim, H. ve Raw, S. D. (1996). Psychotherapy Research Is Not Psychotherapy Practice. *Clinical Psychology: Science and Practice*, 3(2), 168-171. <https://doi.org/10.1111/j.1468-2850.1996.tb00067.x>
- Fernandez, E., Salem, D., Swift, J. K. ve Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology*, 83(6), 1108-1122. <https://doi.org/10.1037/ccp0000044>
- Flint, J., Cuijpers, P., Horder, J., Koole, S. L. ve Munafò, M. R. (2015). Is there an excess of significant findings in published studies of psychotherapy for depression? *Psychological Medicine*, 45(2), 439-446. <https://doi.org/10.1017/S0033291714001421>
- Fonagy, P., Luyten, P. ve Bateman, A. (2017). Treating Borderline Personality Disorder With Psychotherapy: Where Do We Go From Here? *JAMA Psychiatry*, 74(4), 316-317. <https://doi.org/10.1001/jamapsychiatry.2016.4302>
- Freire, E. S. (2006). Randomized Controlled Clinical Trial in Psychotherapy Research: An Epistemological Controversy. *Journal of Humanistic Psychology*, 46(3), 323-335. <https://doi.org/10.1177/0022167806286276>
- Friborg, O. ve Johnsen, T. J. (2017). The effect of cognitive-behavioral therapy as an antidepressive treatment is falling: Reply to Ljótsson et al. (2017) and Cristea et al. (2017). *Psychological Bulletin*, 143(3), 341-345. <https://doi.org/10.1037/bul0000090>
- Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., Chen, P., Hunot, V., & Churchill, R. (2014). Waiting list may be a placebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, 130(3), 181-192. <https://doi.org/10.1111/acps.12275>
- Gelman, A. ve Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, 9(6), 641-651. <https://doi.org/10.1177/1745691614551642>
- Gelo, O. C. G., Lagetto, G., Dinoi, C., Belfiore, E., Lombi, E., Blasi, S., Aria, M., & Ciavolino, E. (2020). Which methodological practice(s) for psychotherapy science? A systematic review and a proposal. *Integrative Psychological & Behavioral Science*, 54(1), 215-248. <https://doi.org/10.1007/s12124-019-09494-3>
- Glass, G. V. (1982). Meta-analysis: An Approach to the Synthesis of Research Results. *Journal of research in science teaching*, 19(2), 93-112.
- Goldfried, M. R. (2020). The field of psychotherapy: Over 100 years old and still an infant science. *Clinical Psychology in Europe*, 2(1), e2753. <https://doi.org/10.32872/cpe.v2i1.2753>
- Goldfried, M. R. ve Wolfe, B. E. (1998). Toward a more clinically valid approach to therapy research. *Journal of Consulting and Clinical Psychology*, 66(1), 143-150. <https://doi.org/10.1037/0022->

006X.66.1.143

- González-Blanch, C. ve Carral-Fernández, L. (2017). Cage up Dodo, please! The tale of all psychotherapies being equally effective. *Papeles del Psicólogo*, 38(2), 94-106. <https://doi.org/10.23923/pap.psicol2017.2828>
- Hansen, W. B. ve Collins, L. M. (1994). Seven ways to increase power without increasing N. *NIDA Research Monograph*, 142, 184-195.
- Hengartner, M. P. (2018). Raising awareness for the replication crisis in clinical psychology by focusing on inconsistencies in psychotherapy research: How much can we rely on published findings from efficacy trials? *Frontiers in Psychology*, 9, Article 256. <https://doi.org/10.3389/fpsyg.2018.00256>
- Hoertel, N., de Maricourt, P., Katz, J., Doukhan, R., Lavaud, P., Peyre, H. ve Limosin, F. (2014). Are participants in pharmacological and psychotherapy treatment trials for social anxiety disorder representative of patients in real-life settings? *Journal of Clinical Psychopharmacology*, 34(6), 697-703. <https://doi.org/10.1097/JCP.0000000000000204>
- Hoertel, N., López, S., Wang, S., González-Pinto, A., Limosin, F. ve Blanco, C. (2015). Generalizability of pharmacological and psychotherapy clinical trial results for borderline personality disorder to community samples. *Personality Disorders: Theory, Research, and Treatment*, 6(1), 81-87. <https://doi.org/10.1037/per0000091>
- Hopewell, S., Clarke, M. ve Mallett, S. (2005). Grey Literature and Systematic Reviews. H. R. Rothstein & A. J. Sutton (Ed.), *Publication Bias in Meta-Analysis* içinde (s. 49-72). John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470870168.ch4>
- Hotopf, M. (2002). The pragmatic randomised controlled trial. *Advances in Psychiatric Treatment*, 8(5), 326-333. doi:10.1192/apt.8.5.326
- Hsu, L. M. (1992). Random sampling, randomization, and equivalence of contrasted groups in psychotherapy outcome research. A. E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* içinde (s. 91-105). American Psychological Association. <https://doi.org/10.1037/10109-018>
- Johns, R. G., Barkham, M., Kellett, S. ve Saxon, D. (2019). A systematic review of therapist effects: A critical narrative update and refinement to Baldwin and Imel's (2013) review. *Clinical Psychology Review*, 67, 78-93. <https://doi.org/10.1016/j.cpr.2018.08.004>
- Johnsen, T. J. ve Friborg, O. (2015). The effects of cognitive behavioral therapy as an anti-depressive treatment is falling: A meta-analysis. *Psychological Bulletin*, 141(4), 747-768. <https://doi.org/10.1037/bul0000015>
- Jones, J. L. ve Mehr, S. L. (2007). Foundations and assumptions of the scientist-practitioner model. *American Behavioral Scientist*, 50(6), 766-771. <https://doi.org/10.1177/0002764206296454>
- Kelley, K. ve Preacher, K. J. (2012). On effect size. *Psychological Methods*, 17(2), 137-152. <https://doi.org/10.1037/a0028086>
- Kendall, J. (2003). Designing a research project: Randomised controlled trials and their principles. *Emergency Medicine Journal: EMJ*, 20(2), 164-168. <https://dx.doi.org/10.1136%2Femj.20.2.164>
- Kiai, A. (2019). To protect credibility in science, banish "publish or perish". *Nature Human Behaviour*, 3(10), 1017-1018. <https://doi.org/10.1038/s41562-019-0741-0>
- Kirkham, J. J., Dwan, K. M., Altman, D. G., Gamble, C., Dodd, S., Smyth, R. ve Williamson, P. R. (2010). The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*, 340, c365. <https://doi.org/10.1136/bmj.c365>
- Lambert, M. J. (2013). The Efficacy and Effectiveness of Psychotherapy. M. J. Lambert (Ed.), *Bergin and Garfields Handbook of Psychotherapy and Behavior Change* (6. Bs) içinde (s. 169-218). John

Wiley & Sons, Ltd.

- Landman, J. T. ve Dawes, R. M. (1982). Psychotherapy outcome: Smith and Glass' conclusions stand up under scrutiny. *American Psychologist*, 37(5), 504-516. <https://doi.org/10.1037/0003-066X.37.5.504>
- Laska, K. M., Gurman, A. S. ve Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy*, 51(4), 467-481. <https://doi.org/10.1037/a0034332>
- Leichsenring, F., Abbass, A., Hilsenroth, M. J., Leweke, F., Luyten, P., Keefe, J. R., Midgley, N., Rabung, S., Salzer, S. ve Steinert, C. (2017). Biases in research: Risk factors for non-replicability in psychotherapy and pharmacotherapy research. *Psychological Medicine*, 47(6), 1000-1011. <https://doi.org/10.1017/S003329171600324X>
- Leichsenring, F., Abbass, A., Hilsenroth, M. J., Luyten, P., Munder, T., Rabung, S. ve Steinert, C. (2018). "Gold standards," plurality and monocultures: The need for diversity in psychotherapy. *Frontiers in Psychiatry*, 9, 159. <https://doi.org/10.3389/fpsy.2018.00159>
- Leichsenring, F. ve Rabung, S. (2011). Long-term psychodynamic psychotherapy in complex mental disorders: Update of a meta-analysis. *The British Journal of Psychiatry*, 199(1), 15-22. <https://doi.org/10.1192/bjp.bp.110.082776>
- Leykin, Y. ve DeRubeis, R. J. (2009). Allegiance in Psychotherapy Outcome Research: Separating Association From Bias. *Clinical Psychology: Science and Practice*, 16(1), 54-65. <https://doi.org/10.1111/j.1468-2850.2009.01143.x>
- Lilienfeld, S. O. (2014). The Dodo Bird verdict: Status in 2014. *The Behavior Therapist*, 37(4), 91-95.
- Linardon, J., Kothe, E. J. ve Fuller-Tyszkiewicz, M. (2019). Efficacy of psychotherapy for bulimia nervosa and binge-eating disorder on self-esteem improvement: Meta-analysis. *European Eating Disorders Review*, 27(2), 109-123. <https://doi.org/10.1002/erv.2662>
- Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., Halperin, G., Bishop, M., Berman, J. S., & Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, 6(1), 95-106. <https://doi.org/10.1093/clipsy.6.1.95>
- Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T. P., Berman, J. S., Levitt, J. T., Seligman, D. A., & Krause, E. D. (2002). The dodo bird verdict is alive and well--mostly. *Clinical Psychology: Science and Practice*, 9(1), 2-12. <https://doi.org/10.1093/clipsy.9.1.2>
- Luborsky, L., Singer, B. ve Luborsky, L. (1975). Comparative studies of psychotherapies: Is it true that "everyone has won and all must have prizes"? *Archives of General Psychiatry*, 32(8), 995-1008. <https://doi.org/10.1001/archpsyc.1975.01760260059004>
- Marcus, D. K., O'Connell, D., Norris, A. L. ve Sawaqdeh, A. (2014). Is the Dodo bird endangered in the 21st century? A meta-analysis of treatment comparison studies. *Clinical Psychology Review*, 34(7), 519-530. <https://doi.org/10.1016/j.cpr.2014.08.001>
- Maxwell, S. E., Lau, M. Y. ve Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, 70(6), 487-498. <https://doi.org/10.1037/a0039400>
- Miguel, C., Karyotaki, E., Cuijpers, P. ve Cristea, I. A. (2021). Selective outcome reporting and the effectiveness of psychotherapies for depression. *World Psychiatry*, 20(3), 444-445. <https://dx.doi.org/10.1002%2Fwps.20900>
- Miller, L. E., ve Stewart, M. E. (2011). The blind leading the blind: Use and misuse of blinding in randomized controlled trials. *Contemporary Clinical Trials*, 32(2), 240-243. <https://doi.org/10.1002%2Fwps.20900>

- org/10.1016/j.cct.2010.11.004
- Morrison, K. H., Bradley, R. ve Westen, D. (2003). The external validity of controlled clinical trials of psychotherapy for depression and anxiety: A naturalistic study. *Psychology and Psychotherapy: Theory, Research and Practice*, 76(2), 109-132. <https://doi.org/10.1348/147608303765951168>
- Mulder, R., Singh, A. B., Hamilton, A., Das, P., Outhred, T., Morris, G., Bassett, D., Baune, B. T., Berk, M., Boyce, P., Lyndon, B., Parker, G., & Malhi, G. S. (2018). The limitations of using randomised controlled trials as a basis for developing treatment guidelines. *Evidence-based mental health*, 21(1), 4–6. <https://doi.org/10.1136/eb-2017-102701>
- Munder, T., Brüttsch, O., Leonhart, R., Gerger, H. ve Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review*, 33(4), 501-511. <https://doi.org/10.1016/j.cpr.2013.02.002>
- Munder, T., Flückiger, C., Leichsenring, F., Abbass, A. A., Hilsenroth, M. J., Luyten, P., Rabung, S., Steinert, C., & Wampold, B. E. (2019). Is psychotherapy effective? A re-analysis of treatments for depression. *Epidemiology and psychiatric sciences*, 28(3), 268–274. <https://doi.org/10.1017/S2045796018000355>
- Münch, R. (2014). *Academic capitalism: Universities in the global struggle for excellence*. Routledge.
- Nahum, D., Alfonso, C. A. ve Sönmez, E. (2019). Common Factors in Psychotherapy. A. Javed ve K. N. Fountoulakis (Ed.), *Advances in Psychiatry* içinde (s. 471-481). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-70554-5_29
- Niemeyer, H., Musch, J. ve Pietrowsky, R. (2012). Publication bias in meta-analyses of the efficacy of psychotherapeutic interventions for schizophrenia. *Schizophrenia Research*, 138(2), 103-112. <https://doi.org/10.1016/j.schres.2012.03.023>
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., Ishiyama, J., ... Yarkoni, T. (2015). SCIENTIFIC STANDARDS. Promoting an open research culture. *Science (New York, N.Y.)*, 348(6242), 1422–1425. <https://doi.org/10.1126/science.aab2374>
- Open Science Collaboration. (2015). PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*, 349, Article aac4716. <https://doi.org/10.1126/science.aac4716>
- O’rouke, K. (2007). An historical perspective on meta-analysis: Dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine*, 100(12), 579-582. <https://doi.org/10.1177/0141076807100012020>
- Owen, J., Drinane, J. M., Idigo, K. C. ve Valentine, J. C. (2015). Psychotherapist effects in meta-analyses: How accurate are treatment effects? *Psychotherapy*, 52(3), 321-328. <https://doi.org/10.1037/pst0000014>
- Peng, C.-Y. J. ve Chen, L.-T. (2014). Beyond Cohen’s d: Alternative Effect Size Measures for Between-Subject Designs. *The Journal of Experimental Education*, 82(1), 22-50. <https://doi.org/10.1080/00220973.2012.745471>
- Persons, J. B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: A proposed remedy. *American Psychologist*, 46(2), 99-106. <https://doi.org/10.1037/0003-066X.46.2.99>
- Persons, J. B., Roberts, N. A., Zalecki, C. A. ve Brechwald, W. A. (2006). Naturalistic outcome of case formulation-driven cognitive-behavior therapy for anxious depressed outpatients. *Behaviour Research and Therapy*, 44(7), 1041-1051. <https://doi.org/10.1016/j.brat.2005.08.005>
- Persons, J. B. ve Silberschatz, G. (1998). Are results of randomized controlled trials useful to

- psychotherapists? *Journal of Consulting and Clinical Psychology*, 66(1), 126-135. <https://doi.org/10.1037/0022-006X.66.1.126>
- Philips, B., & Falkenström, F. (2021). What Research Evidence Is Valid for Psychotherapy Research?. *Frontiers in psychiatry*, 11, 625380. <https://doi.org/10.3389/fpsy.2020.625380>
- Rifkin, A. (2007). Randomized controlled trials and psychotherapy research [Editorial]. *The American Journal of Psychiatry*, 164(1), 7–8. <https://doi.org/10.1176/appi.ajp.164.1.7>
- Rosenzweig, S. (1954). A transvaluation of psychotherapy: A reply to Hans Eysenck. *The Journal of Abnormal and Social Psychology*, 49(2), 298-304. <https://doi.org/10.1037/h0061172>
- Rossi, J. S. (2013). Statistical power analysis. J. A. Schinka, W. F. Velicer ve I. B. Weiner (Ed.), *Handbook of psychology: Research methods in psychology* içinde (s. 71–108). John Wiley & Sons, Inc..
- Rothstein, H. R. ve Hopewell, S. (2009). Grey literature. H. Cooper, L. V. Hedges ve J. C. Valentine (Ed.), *The handbook of research synthesis and meta-analysis* içinde (s. 103–125). Russell Sage Foundation.
- Rubio-Aparicio, M., Marín-Martínez, F., Sánchez-Meca, J. ve López-López, J. A. (2018). A methodological review of meta-analyses of the effectiveness of clinical psychology treatments. *Behavior Research Methods*, 50(5), 2057-2073. <https://doi.org/10.3758/s13428-017-0973-8>
- Sakaluk, J. K., Williams, A. J., Kilshaw, R. E. ve Rhyner, K. T. (2019). Evaluating the evidential value of empirically supported psychological treatments (ESTs): A meta-scientific review. *Journal of Abnormal Psychology*, 128(6), 500-509. <https://doi.org/10.1037/abn0000421>
- Schäfer, T. ve Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00813>
- Shapiro, D. A. ve Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, 92(3), 581-604. <https://doi.org/10.1037/0033-2909.92.3.581>
- Shean, G. (2012). Some limitations on the external validity of psychotherapy efficacy studies and suggestions for future research. *American Journal of Psychotherapy*, 66(3), 227-242. <https://doi.org/10.1176/appi.psychotherapy.2012.66.3.227>
- Shean, G. (2016). Psychotherapy Outcome Research: Issues and Questions. *Psychodynamic Psychiatry*, 44(1), 1-24. <https://doi.org/10.1521/pdps.2016.44.1.1>
- Shedler, J. (2018). Where Is the Evidence for “Evidence-Based” Therapy? *Psychiatric Clinics of North America*, 41(2), 319-329. <https://doi.org/10.1016/j.psc.2018.02.001>
- Shelby, L. B. ve Vaske, J. J. (2008). Understanding meta-analysis: A review of the methodological literature. *Leisure Sciences*, 30(2), 96-110. <https://doi.org/10.1080/01490400701881366>
- Shinohara, K., Tajika, A., Imai, H., Takeshima, N., Hayasaka, Y. ve Furukawa, T. A. (2015). Protocol registration and selective outcome reporting in recent psychiatry trials: New antidepressants and cognitive behavioural therapies. *Acta Psychiatrica Scandinavica*, 132(6), 489-498. <https://doi.org/10.1111/acps.12502>
- Smit, Y., Huibers, M. J. H., Ioannidis, J. P. A., van Dyck, R., van Tilburg, W. ve Arntz, A. (2012). The effectiveness of long-term psychoanalytic psychotherapy—A meta-analysis of randomized controlled trials. *Clinical Psychology Review*, 32(2), 81-92. <https://doi.org/10.1016/j.cpr.2011.11.003>
- Smith, M. L. ve Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(9), 752-760. <https://doi.org/10.1037/0003-066X.32.9.752>

- Sohn, D. (1996). Publication bias and the evaluation of psychotherapy efficacy in reviews of the research literature. *Clinical Psychology Review*, 16(2), 147-156. [https://doi.org/10.1016/0272-7358\(96\)00005-0](https://doi.org/10.1016/0272-7358(96)00005-0)
- Springer, K. S., Levy, H. C. ve Tolin, D. F. (2018). Remission in CBT for adult anxiety disorders: A meta-analysis. *Clinical Psychology Review*, 61, 1-8. <https://doi.org/10.1016/j.cpr.2018.03.002>
- Stricker, J. ve Günther, A. (2019). Scientific misconduct in psychology: A systematic review of prevalence estimates and new empirical data. *Zeitschrift für Psychologie*, 227(1), 53-63. <http://dx.doi.org/10.1027/2151-2604/a000356>
- Strupp, H. H. (1963). The outcome problem in psychotherapy revisited. *Psychotherapy: Theory, Research & Practice*, 1(1), 1-13. <https://doi.org/10.1037/h0088565>
- Tackett, J. L., Brandes, C. M., King, K. M. ve Markon, K. E. (2019). Psychology's replication crisis and clinical psychological science. *Annual Review of Clinical Psychology*, 15, 579-604. <https://doi.org/10.1146/annurev-clinpsy-050718-095710>
- Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., Oltmanns, T. F., & Shrout, P. E. (2017). It's time to broaden the replicability conversation: Thoughts for and from clinical psychological science. *Perspectives on Psychological Science*, 12(5), 742-756. <https://doi.org/10.1177/1745691617690042>
- Task Force on Promotion and Dissemination of Psychological Procedures Division of Clinical Psychology. (1995). Training in and dissemination of empirically-validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48(1), 3-23.
- Thase, M. E. (1999). What is the investigator allegiance effect and what should we do about it? *Clinical Psychology: Science and Practice*, 6(1), 113-115. <https://doi.org/10.1093/clipsy.6.1.113>
- Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W. ve Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist*, 69(3), 218-229. <https://doi.org/10.1037/a0035099>
- Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies?: A meta-analytic review. *Clinical Psychology Review*, 30(6), 710-720. <https://doi.org/10.1016/j.cpr.2010.05.003>
- Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D. ve Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology: Science and Practice*, 22(4), 317-338. <https://doi.org/10.1111/cpsp.12122>
- van der Zweerde, T., Bisdounis, L., Kyle, S. D., Lancee, J. ve van Straten, A. (2019). Cognitive behavioral therapy for insomnia: A meta-analysis of long-term effects in controlled studies. *Sleep Medicine Reviews*, 48, 101208. <https://doi.org/10.1016/j.smr.2019.08.002>
- Vevea, J. L. ve Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods*, 10(4), 428-443.
- von Wolff, A., Jansen, M., Hölzel, L. P., Westphal, A., Härter, M. ve Kriston, L. (2014). Generalizability of findings from efficacy trials for chronic depression: An analysis of eligibility criteria. *Psychiatric Services*, 65(7), 897-904. <https://doi.org/10.1176/appi.ps.201300309>
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Routledge.
- Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14(3), 270-277. <https://doi.org/10.1002/wps.20238>
- Wampold, B. E. (2019). The basics of psychotherapy: An introduction to theory and practice (2. Bs.). *American Psychological Association*. <https://doi.org/10.1037/0000117-000>
- Wampold, B. E., Hollon, S. D. ve Hill, C. E. (2011). Unresolved questions and future directions in psychotherapy research. *History of psychotherapy: Continuity and change* (2. Bs) içinde (s. 333-

- 356). American Psychological Association. <https://doi.org/10.1037/12353-011>
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K. ve Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin*, 122(3), 203-215. <https://doi.org/10.1037/0033-2909.122.3.203>
- Weinberger, J. (1995). Common factors aren't so common: The common factors dilemma. *Clinical Psychology: Science and Practice*, 2(1), 45–69. <https://doi.org/10.1111/j.1468-2850.1995.tb00024.x>
- Westbrook, D. ve Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour Research and Therapy*, 43(10), 1243-1261. <https://doi.org/10.1016/j.brat.2004.09.006>
- Westen, D. ve Bradley, R. (2005). Empirically supported complexity: Rethinking evidence-based practice in psychotherapy. *Current Directions in Psychological Science*, 14(5), 266-271. <https://doi.org/10.1111/j.0963-7214.2005.00378.x>
- Westen, D. ve Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. *Journal of Consulting and Clinical Psychology*, 69(6), 875-899. <https://doi.org/10.1037/0022-006X.69.6.875>
- Westen, D., Novotny, C. M. ve Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, 130(4), 631-663. <https://doi.org/10.1037/0033-2909.130.4.631>
- Wierzbicki, M. ve Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology: Research and Practice*, 24(2), 190-195. <https://doi.org/10.1037/0735-7028.24.2.190>
- Wojnarowski, C., Firth, N., Finegan, M. ve Delgadillo, J. (2019). Predictors of depression relapse and recurrence after cognitive behavioural therapy: A systematic review and meta-analysis. *Behavioural and Cognitive Psychotherapy*, 47(5), 514-529. <https://doi.org/10.1017/S1352465819000080>

