



Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi

Yıl: 2019, Cilt: 20, Sayı: 3, Sayfa No: 597-628

DOI: 10.21565/ozelegitimdergisi.421952

DERLEME

Gönderim Tarihi: 08.05.18

Kabul Tarihi: 13.03.19

Erken Görünüm: 23.03.19

Bilimsel-Dayanaklı Uygulamaları Belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” Yönergesinin Matematik Becerileri Öğretimi Örneğinde Ele Alınışı*

Orhan Aydın **
Anadolu Üniversitesi

Elif Tekin-İftar ***
Anadolu Üniversitesi

Salih Rakap ***
Ondokuz Mayıs Üniversitesi

Öz

Özel eğitim alanında bilimsel-dayanaklı uygulamalar her geçen gün uzmanlar, araştırmacılar ve uygulamacılar arasında daha fazla dikkate alınan bir konu haline gelmiştir. Bu çalışmada, Horner ve diğerleri tarafından 2005 yılında önerilen bilimsel-dayanaklı uygulamaları belirlemede “*Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri*” yönergesi ele alınmıştır. Tek-denekli deneysel araştırma desenlerine göre tasarlanmış bilimsel çalışmaların değerlendirilmesinde bu yönergenin; (a) nasıl ele alınarak yorumlanabileceği ve (b) yönergede yer alan göstergeler ışığında “kabul edilebilirlik” ölçütlerini karşılayan çalışmaların nasıl belirlenebileceği otizm spektrum bozukluğu olan bireylere matematik becerileri öğretimi konusunda yayımlanmış çalışmalardan örnekler verilerek açıklanması amaçlanmıştır. Bu çalışmanın, tek-denekli deneysel araştırma desenleriyle çalışmasını tasarlayacak olan araştırmacılara ve uygulamacılara nitelikli bir çalışma süreci yürütmeleri açısından yol gösterici bir kaynak olacağı düşünülmektedir. Ayrıca, bu çalışmanın tek-denekli deneysel araştırmaları değerlendirecek olan alan hakemlerine ve sistematik gözden geçirme ve/veya meta analiz yapacak olan araştırmacılara bilimsel dayanaklı uygulamaları belirleme ve değerlendirme konusunda niteliksel göstergelerin ele alınışı açısından rehberlik edeceği umulmaktadır.

Anahtar Sözcükler: Bilimsel araştırmalar, tek-denekli deneysel araştırmalar, bilimsel-dayanaklı uygulamalar, niteliksel göstergeler, işlevsel ilişki

Önerilen Atıf Şekli

Aydın, O., Tekin-İftar, E., & Rakap, S. (2019). Bilimsel dayanaklı uygulamaları belirlemede “tek-denekli deneysel araştırmaların niteliksel göstergeleri” yönergesinin matematik becerileri öğretimi örneğinde ele alınışı. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 20(3), 597-628. doi: 10.21565/ozelegitimdergisi.421952

*Çalışma birinci yazarın Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü Özel Eğitim Bölümünde tamamlanan yüksek lisans tez çalışmasından üretilmiştir. Çalışma Lisansüstü Tez Projeleri kapsamında Anadolu Üniversitesi Bilimsel Araştırma Projeleri tarafından desteklenmiştir (Proje No: 1609E619)

****Sorumlu Yazar:** Arş. Gör., E-posta: o_aydin@anadolu.edu.tr, <https://orcid.org/0000-0002-9695-2414>

***Prof. Dr., E-Posta: eltekin@anadolu.edu.tr, <https://orcid.org/0000-0001-5512-616X>

***Doç. Dr., E-posta: salih.rakap@omu.edu.tr, <https://orcid.org/0000-0001-7853-3825>

Eğitim ve davranış bilimlerinde bilimsel araştırmalar, çeşitli bakış açılarına göre sınıflandırılmaktadır. Bu sınıflandırmalardan biri, veri toplama ve analiz süreçlerine göre yapılmaktadır. *Yönteme* göre sınıflandırma olarak kavramsallaştırılan bu sınıflamada (a) betimsel araştırmalar, (b) bağıntısal araştırmalar, (c) nedensel-karşılaştırmalı araştırmalar ve (d) deneysel araştırmalar yer almaktadır. Betimsel araştırmalar, üzerinde çalışılan olgunun derinlemesine incelenmesine yönelik verilerin toplanarak analiz edildiği çalışmalardır. Bağıntısal araştırmalar, değişkenler arasındaki olası bağıntı (ilişki) ve bağıntı düzeyinin araştırıldığı araştırmalardır. Nedensel-karşılaştırmalı araştırmalar ise, önceden var olan (ya da gerçekleşmiş) bir durumun nedenlerini, bu nedenleri etkileyen olası değişkenleri ya da bu duruma bağlı olarak ortaya çıkan etkileri (sonuçları) belirlemeye yönelik tasarlanan araştırmalardır. Araştırmacı, nedensel-karşılaştırmalı araştırmalarda önceden ortaya çıkan durumun nedenlerinden ya da sonuçlarından yola çıkarak araştırmayı tasarlayabilir. Deneysel araştırmalara en yakın araştırma tasarımı olmasına rağmen bu araştırmalarda grupların önceden kendiliğinden ya da araştırmacıdan bağımsız olarak oluşması nedeniyle yansız atama yapılamaması ve araştırmacının etkisi araştırılan değişken üzerinde kontrol/uyarlama gücünün olmaması nedeniyle deneysel araştırmalardan farklılaşmaktadır. Deneysel araştırmalar ise araştırmacının araştırma sonuçlarını genellemeyi hedeflediği evrenden yansız atama yoluyla örneklem seçerek, bağımlı değişken üzerinde etkisini araştırdığı bağımsız değişkeni kontrol ederek/uyarlayarak değişkenler arasında neden-sonuç ilişkisini araştırmak üzere tasarlanan araştırmalardır. Deneysel araştırmalar başlığı altında araştırmacının katılımcılarının seçilme biçimine göre iki grup araştırmadan söz edilir. Bunlar gerçek/tam deneysel araştırmalar ve yarı deneysel araştırmalardır. Gerçek/tam deneysel araştırmalarda araştırmacının katılımcıları yansız olarak belirlenirken, yarı deneysel araştırmalarda katılımcılar yansız olarak belirlenmemektedir. Dolayısıyla, yansız atama yapıp yapılmadığı tam deneysel ve yarı deneysel araştırmalar arasındaki farkı belirleyen temel durumlardan birisidir. Eğitim ve davranış bilimlerinde bir uygulamanın etkili bir uygulama olup olmadığına ilişkin sorulara en güçlü yanıt veren, neden-sonuç ilişkisini güçlü bir şekilde ortaya koyan araştırmaların deneysel araştırmalar olduğuna ilişkin kabul edilmiş yaygın bir görüş vardır (Cook, Tankersley, Cook ve Landrum, 2008). Dolayısıyla, deneysel araştırmalar, söz konusu alanlarda bilgi üretiminde çok önemli role sahiptir. Ancak, deneysel araştırmalar arasında önemli bir yere sahip olan ve ifade edildiği bilgiyi en güçlü biçimde üretebilen grup deneysel araştırmaların çeşitli sınırlılıkları da söz konusudur. Örneğin, grup deneysel araştırmalarda bir bağımsız değişkenin bağımlı değişken üzerinde etkili olup olmadığına, bağımlı değişken açısından grup ortalama puanları üzerinden değerlendirilme yapılarak bakılmaktadır. Bu grup ortalama puanlarına rağmen bireysel farklılıklar nedeniyle katılımcıların bazıları söz konusu ortalama puanları elde edecek kadar uygulamadan etkilenmekte, bazı katılımcılar hiç etkilenmemekte ya da olumsuz yönde etkilenmektedir (Franklin, Allison ve Gorman, 2014). Ortalama puanlar üzerinden değerlendirme yapılması, özellikle bireysel farklılıkların önemli olduğu durumlarda yeterince hassas değerlendirme yapılmasını önlemekte/güçleştirmektedir (Barlow ve Hersen, 1984). Aynı zamanda grup deneysel araştırmalarda bağımsız değişken açısından birbirine benzer katılımcılar bulunması ve bu katılımcılardan iki denk (homojen) grup, deney ve kontrol grubu, oluşturulması gerekmektedir. Özellikle bireysel farklılıkların çok geniş olduğu gruplarda, örneğin özel gereksinimli çocuklarda ve bireylerde, yeterli sayıda katılımcıya ulaşarak birbirine denk gruplar oluşturmak mümkün olamayabilmektedir (Sheridan, 2014). Son olarak ise söz konusu özellikler açısından yeterli sayıda katılımcıya ulaşarak iki denk grup oluşturulabilse bile bu gruplardan kontrol grubunun belirli bir süre için uygulamadan mahrum bırakılması çoğu zaman etik ve eğitsel açıdan uygun ya da istendik bir durum değildir (Kırcaali-İftar, 2012). Tek-denekli deneysel araştırmaların belli başlı geliştirilme nedenleri arasında sıralanan bu sınırlılıkları kontrol edebilmek ya da elimine edebilmek yer almaktadır (Barlow ve Hersen, 1984; Franklin ve diğ., 2014; Kırcaali-İftar, 2012; Sheridan, 2014).

Tek-denekli araştırmalar, yarı deneysel araştırmalar başlığı altında ele alınmaktadır. Böyle sınıflandırılmasının temel nedeni bağımsız değişkenin kontrolünün tamamen araştırmacıda olması ve yansız atama yapılamamasından kaynaklanmaktadır (Tekin-İftar, 2012a). Ancak, tek-denekli deneysel araştırmaların katılımcıları hâlihazırda sorun davranışları olan ya da yeni davranışların kazandırılmasına gereksinim duyan katılımcılar arasından seçilmektedir. Dolayısıyla, araştırma katılımcıları yansız atama yoluyla belirlen(e)memektedir. Ancak, tek-denekli deneysel araştırmaları diğer yarı deneysel araştırma tasarımlarından

ayırarak nokta bu araştırmalarda katılımcı grupları üzerinde çalışılmaması ve her bir katılımcı ile yinelenen ölçümler yapılarak tekil değerlendirmelerin yapılmasıdır. Bu değerlendirmeler bazen tek bir katılımcı bazen de bir grup katılımcının (ör. bir grup öğrenci ya da sınıftaki tüm öğrenciler) tek bir katılımcı olarak ele alınmasıyla yapılmaktadır. Tek-denekli deneysel araştırmalar standart koşullar altında yinelenen ölçümler yapılarak sosyal açıdan önemli davranışlarda her bir katılımcının kendi kendisinin kontrolünü oluşturmasıyla davranış değiştirme ya da öğretim uygulamalarının etkililiğini tekil değerlendirmeler yaparak ortaya koyan araştırmalardır (Kennedy, 2005; Rilley-Tillman ve Burns, 2009; Tekin-İftar, 2012a).

Tek-denekli deneysel araştırmalarda, her bir katılımcı ya da katılımcı grubu kendi kontrolünü oluşturmada ve deneysel kontrol her bir katılımcı ya da katılımcı grubunda katılımcı içi ya da katılımcılar arası yinelenmeler yapılarak tek tek kurulmaktadır. Ancak, tek-denekli deneysel araştırmalarda deneysel kontrolün oluşması belli özelliklerin karşılanmasıyla elde edilebilmektedir. Bu özellikler; (a) standart koşullar, (b) yinelenen ölçümler, (c) başlama düzeyi evresi ve (d) uygulama evresidir (Tekin-İftar, 2012b). Diğer deyişle, standart koşullar sağlanarak başlama düzeyi ile uygulama arasında etki farkının gözlenmesi ve bu etki farkının en az üç durumda yinelenerek gözlenmesi ile deneysel kontrol sağlanabilmektedir (Horner ve Odom, 2014). *Standart koşullar*, tek-denekli deneysel araştırmalarda tasarlanan evrelerin her bir oturumda benzer koşullar altında düzenlenmesi olarak tanımlanır. Araştırmacı tasarladığı evrelerde bağımsız değişken dışındaki tüm değişkenleri benzer şekilde kullanır (Tekin-İftar, 2012a). Örneğin, araştırmacı doğru davranışları sözel pekiştirme yoluyla pekiştirmeyi planladıysa ya da katılımcının tepkide bulunması için yanıt aralığını beş saniye olarak belirlediyse tüm evrelerde bu özellikleri benzer şekilde kullanılmalıdır. Araştırmacı başlama düzeyi evresinde yanıt aralığını beş saniye, uygulama evresinde 10 saniye olarak belirlerse, uygulama evresindeki katılımcı tepkisinin bağımsız değişkenin uygulanmasından mı yoksa yanıt aralığının daha uzun süre olmasından mı ya da bu iki durumun birleşik etkisinden mi kaynaklandığını analiz etmek mümkün olmayabilir. *Yinelenen ölçümler*, bağımlı değişken için zaman içinde standart koşullar altında tekrar tekrar veri toplanması olarak tanımlanır. *Başlama düzeyi evresi*, grup deneysel araştırmalarda katılımcıların araştırmanın bağımlı değişkenlerine ilişkin düzeylerinin bir müdahale/uygulama sunulmadan belirlendiği ön-test değerlendirmesi görevini görür. *Başlama düzeyi evresi*, tek-denekli deneysel araştırmaların başlangıç noktası olup iki nedenle düzenlenir: (a) katılımcıların uygulama öncesinde bağımlı değişkeni hangi düzey, sıklık, oran veya yüzde ile sergilediklerini belirlemek ve (b) başlama düzeyi evresindeki katılımcı performansı ile uygulama evresi sırasında ya da sonrasındaki katılımcı performansı arasında karşılaştırma yaparak bağımsız değişkenin etkililiğini sınamak. Pek çok tek-denekli deneysel araştırma modelinde başlama düzeyi evresi düzenlemek deneysel kontrolün kurulması açısından bir zorunluluktur. Bazı modellerde zorunluluk olmamakla birlikte düzenlenmesi önerilmektedir. Başlama düzeyi evresi kararlı veri elde edinceye değin düzenlenmelidir. Bu bağlamda, başlama düzeyi evresinde en az beş oturum veri toplanması önerilmektedir. Uygulama evresi ise, başlama düzeyi evresinden hemen sonra düzenlenen bir ya da birden çok bağımsız değişkenin sistematik olarak uygulandığı evredir. Bu evrede bağımsız değişkenin işlevsel ve yinelenebilir biçimde tanımlanması gerekir (Barlow ve Hersen, 1984; Kazdin, 1982; Tawney ve Gast, 1984). Uygulama evresinin ne süreyle düzenleneceğine ilişkin genel bir kural olmamakla birlikte bağımlı değişkende ölçüt karşılanıp kararlı veri elde edilinceye değin sürdürülmesi önerilir. Başlama düzeyi evresinde ve uygulama evresinde bağımlı değişkenin tanımlanması ve özellikleri dikkate alınarak davranışsal kayıt tekniklerinden uygun olanıyla veri toplanır ve veriler grafiksel analiz yoluyla analiz edilir. Araştırmadan elde edilen sonuçların araştırmanın katılımcıları ile benzer özelliklerdeki diğer bireylerde diğer araştırmacılar ve uygulamacılar tarafından kullanılabilmesi için başlama düzeyi ve uygulama evrelerinin ve katılımcı özelliklerinin yinelenme yapmaya izin verecek nitelikte açık olarak tanımlanması gerekir (Alberto ve Troutman, 2009; Lane, Wolery, Reichow ve Rogers, 2007; Tekin-İftar, 2012b).

Tüm bilimsel araştırma tasarımlarında olduğu gibi, tek-denekli deneysel araştırmalarda da araştırmanın geçerlik ve güvenilirliğinin sağlanması elde edilen bulguların inandırıcılığı, anlamlılığı ve doğruluğu açısından son derece önemlidir. Bu durumun, özellikle, araştırmada ele alınan uygulamanın bilimsel-dayanaklı bir uygulama olup olmadığı konusunda karar verebilmek açısından önemi tüm bilimsel araştırma tasarımlarında çok büyüktür. Bilimsel-dayanaklı uygulama kavramı, önce tıp alanında ardından sosyal bilimlerde çeşitli alanlarda (ör. psikoloji, eğitim) bilim dallarının ideolojileri ve geleneklerinin zenginleşmesine hizmet etmiştir (Rakap, 2016; Reichow ve

Volkmar, 2011). Bilimsel-dayanaklı uygulamalar, özel eğitim alanında son 20 yılda araştırmacıların ve eğitim politikalarını belirleme sürecinde rol oynayan kurumların ve meslek örgütlerinin özel önem verdiği konuların başında gelmektedir. Bir uygulamanın bilimsel-dayanaklı olup olmadığının belirlenmesi konusundaki arayışlar sonucunda, bilimsel araştırmalarda bilimsel-dayanak sağlayabilecek araştırma ölçütlerinin neler olması gerektiğini belirleme konusunda araştırmacı ve meslek örgütleri çeşitli girişimler başlatarak farklı araştırma tasarımlarında aranabilecek niteliksel göstergelerin neler olabileceğini ileri sürmüşlerdir. Bilimsel-dayanaklı uygulamalar kavramına ilişkin henüz tam bir uzlaşma sağlanabilmiş olmasa da bu çabaların ürünü olarak özel eğitim alanında çeşitli yönerge (rubric) ve raporların yayımlandığı bilinmektedir. Örneğin, Özel Gereksinimli Çocuklar Konseyi (Council for Exceptional Children-CEC; www.cec.sped.org) bir Görev Grubu oluşturarak farklı araştırma tasarımlarında uygulamaların bilimsel-dayanaklarını belirlemek üzere kullanılacak niteliksel göstergelerin neler olabileceği konusunda bir çalışma yürütmüştür. Bu çalışmanın ürünleri nitel araştırmalarda (Brantlinger, Jimenez, Klingler, Pugach ve Richardson, 2005), bağıntısal araştırmalarda (Thompson, Diamond, McWilliam, Snyder ve Synder, 2005), grup ve yarı deneysel araştırmalarda (Gersten ve diğ., 2005) ve tek-denekli araştırmalarda (Horner ve diğ., 2005) çalışmaların karşılamaları gereken niteliksel göstergeleri ileri süren çalışmalar olmuştur. Ayrıca, Özel Gereksinimli Çocuklar Konseyi 2014 yılında “Özel Gereksinimli Çocuklar Konseyi Özel Eğitim Alanında Bilimsel-Dayanaklı Uygulamaların Standartları” başlıklı bir çalışma ile bir grup araştırmacının hem grup deneysel araştırmalar hem de tek-denekli deneysel araştırmalar için niteliksel göstergeleri belirleyerek araştırmaların metodolojik olarak güçlülük düzeylerini tanımlayan raporlar yayınlamıştır. Bu raporlara Özel Gereksinimli Çocuklar Konseyi tarafından hazırlanmış olan grup deneysel araştırmalar (Gersten ve diğ., 2005) ve tek-denekli deneysel araştırmalar (Horner ve diğ., 2005) niteliksel göstergeleri, What Work Clearinghouse tarafından yayınlanan rapor ve özel eğitim alanında araştırma metodolojileri konusunda çalışmalar yürüten 23 araştırmacı/uzmanla yürütülen delfi çalışmasında (uzman görüşlerinin analiz edildiği çalışma) ortaya çıkan görüşler kaynaklık etmiştir (Cook ve diğ., 2015). Genel özel eğitim alanında yayımlanan bu rapor ve yönergeler ek olarak, otizm alanında bilimsel-dayanaklı uygulamaları belirlemek üzere araştırmaların niteliksel ölçütlerini ve bu ölçütlere bağlı olarak metodolojik güçlülük düzeylerini belirlemeye yönelik (Reichow, Volkmar ve Cicchetti, 2008) ya da genel olarak eğitim alanında eğitimcilerle bilimsel-dayanaklı uygulamaları belirlemeye ve önermeye yönelik yönergeler (What Work Clearinghouse) de geliştirilmiştir (Rakap, 2016). Ancak, sıralanan bu çalışmaların araştırmacılar ve uygulamacılar tarafından tutarlı biçimde anlaşılması, değerlendirilmesi ve uygulanması bir sorun olabilmektedir. Benzer sorun bu çalışmayı tasarlayan araştırmacılar (ilk iki araştırmacı) tarafından da yaşanmıştır. Araştırmacılar otizm spektrum bozukluğu (OSB) olan çocuklara matematik becerilerinin öğretimi konulu bir sistematik gözden geçirme ve meta analiz çalışmasını yüksek lisans tez çalışması ve bilimsel araştırma projesi kapsamında çalışırken “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” yönergesini kullanmışlardır. Ancak, dâhil etme ölçütlerini karşılayan çalışmaların değerlendirilmesi aşamasında araştırmacılar yönergenin anlaşılması, değerlendirilmesi ve maddelerin çalışmalara uygulanması konularında sorunlar olabileceğini fark ederek bu konuda çalışacak olan araştırmacılara ve uzmanlara kaynaklık edebilmesi amacıyla üzerinde çalışılan projenin konusu olan OSB olan çocuklara matematik becerileri öğretimi örneğinde okumakta olduğunuz çalışmayı tasarlamışlardır. Çalışmanın amacı, Horner ve diğerleri tarafından 2005 yılında geliştirilen bilimsel-dayanaklı uygulamaları belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” yönergesinin bilimsel çalışmalarda nasıl ele alınabileceği, bu göstergeler ışığında “kabul edilebilirlik” ölçütlerini karşılayan çalışmaların nasıl belirlenebileceği ve çalışmalarda etkisi incelenen bir uygulamanın bilimsel-dayanaklı bir uygulama olarak değerlendirilmesi sürecinde hangi ölçütlerin dikkate alınabileceği gelişimsel yetersizliği olan bireylere matematik becerileri öğretimi konusunda yayımlanmış çalışmalardan örnekler verilerek açıklamaktır. İzleyen bölümde önce, bilimsel-dayanaklı uygulamaları belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” başlıklı yönerge tanıtılmış ardından bu yönergede yer alan maddelerin bir çalışmada nasıl yorumlanabileceği açıklanmış ve son olarak OSB olan bireylere matematik becerilerinin öğretimi konusunda gerçekleştirilmiş olan bir meta-analizi çalışmasına dâhil edilen çalışmalarda niteliksel göstergelerin tam olarak karşılandığı ideal örneklere ve tam olarak karşılanmadığı örneklere yer verilerek çalışmalardan alıntılar yapılmış ardından gerekçeleri açıklanmıştır.

Bilimsel-Dayanaklı Uygulamaları Belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” Yönergesi ve Göstergelerin Yorumlanması

Özel Gereksinimli Çocuklar Konseyi tarafından yayımlanan ve özel eğitim alanında saygınlık açısından ilk sıralarda yer alan *Exceptional Children* dergisi 2005 yılında, aslında geleneğinde olmayan bir biçimde tek bir konuyu irdeleyen bir grup çalışmadan oluşan özel bir sayı yayına hazırlamıştır. Bu sayıda, özel eğitim alanında araştırmalarda niteliksel göstergelerin geliştirilme sürecini ve etkili uygulamaların dayanaklarını ortaya koyan ilkeleri okuyucularla paylaşmak hedeflenmiştir. Bu sayıda yer alan çalışmalar, Özel Gereksinimli Çocuklar Konseyi Araştırma Birimi (CEC Divion of Research) ve Sam Odom tarafından yayına hazırlanarak dergiye sunulmuştur. Derginin editörü Steve Graham’a (2005) göre bu özel sayının hazırlanmasında üç temel amaç söz konusudur; (a) bilimsel-dayanaklı uygulamaların belirlenmesini, bilimsel-dayanaklı uygulamalara ilişkin bilginin paylaşılmasını ve bilimsel-dayanaklı uygulamaların alanda hizmet sunan uzmanlar tarafından kullanımının yaygınlaşabilmesini sağlayabilmek üzere değerlendirme ölçütlerini belirlemek, (b) eğitimin diğer alanlarına kıyasla özel eğitim alanında temel ve uygulamalı araştırmalara ilişkin yıllar içinde verilen özel önemi bu özel sayı ile daha da çarpıcı biçimde sergilemek ve (c) farklı araştırma tasarımlarında aranabilecek niteliksel göstergeleri yayımlayarak özellikle *Exceptional Children* dergisi için hizmet sunan alan hakemlerinin ve konuk hakemlerin makale değerlendirirken dikkate alabilecekleri göstergeleri sunmak. Bu çalışmalardan birisinde Horner ve diğerleri tarafından hazırlanmış ve tek-denekli deneysel araştırmaların bilimsel-dayanaklı uygulamaları belirlemede dikkate alınabilecek niteliksel göstergelerine yer vermiştir. Araştırmacılar bu göstergeleri; (a) katılımcılar ve ortamlar, (b) bağımlı değişken, (c) bağımsız değişken, (d) başlama düzeyi ve (e) geçerlik (iç geçerlik, dış geçerlik ve sosyal geçerlik) olmak üzere beş başlıkta ele almış ve bu başlıkları irdeleyen toplam 21 niteliksel göstergeye yer vermişlerdir. Tablo 1’de 21 niteliksel gösterge ile Horner ve diğerlerinin (2005) önerileri doğrultusunda bu göstergelerin nasıl değerlendirilebileceğine ilişkin açıklamalara yer verilmiştir.

Tablo 1

Tek-Denekli Deneysel Araştırmalarda Niteliksel Göstergeler ve Bu Göstergelerin İçeriklerine İlişkin Açıklamalar

| 1. Katılımcı ve Ortam | |
|---|---|
| Göstergeler | Açıklama |
| 1. <i>Katılımcılar yineleme yapmaya olanak sağlayacak açıklıkta tanımlanır.</i> | Katılımcı sayısı, yaşı/sınıf düzeyi, cinsiyeti ve tanısı belirtilmiş olmalıdır. Ayrıca katılımcının performansına ilişkin puanlama bilgileri sunulmalıdır. Puanlama bilgileri yoksa katılımcının yapabildikleri ve/veya yapamadıklarını betimleyen performans bilgileri ayrıntılı olarak sunulmalıdır. |
| 2. <i>Katılımcı seçme süreci yineleme yapmaya olanak sağlayacak açıklıkta tanımlanır.</i> | Katılımcının çalışmaya katılımı için sahip olması beklenen önkoşul özelliklerinin ve/veya katılımcının çalışma için seçilme ölçütlerinin açıklanması gerekmektedir. Ayrıca (bu önkoşullar/ölçütler sonucunda hedef beceriye ilişkin) katılımcının nasıl seçildiğine (Kim önerdi?, Kim tarafından seçildi?) ilişkin bilgilere yer verilmiş olması da yeterlidir. |
| 3. <i>Ortam yineleme yapmaya olanak sağlayacak açıklıkta tanımlanır.</i> | Çalışılan ortamın fiziksel özelliklerinin açıklanmış olması, ortamda bulunan başka bireyler varsa belirtilmesi, başka bireylere müdahale söz konusuysa açıklanması (ör. akran/öğretmen eğitimi) ve katılımcı-uygulamacının çalışma pozisyonunun açıklanmış olması gerekmektedir. Ayrıca, yalnızca okuldan söz ediliyorsa, okulun bulunduğu bölge/sosyo-ekonomik düzeyi, okuldaki öğrenci sayısı/ özel gereksinimli öğrenci sayısı veya özel eğitim sınıf sayısı gibi değişkenler açısından açıklanmış olması gerekmektedir. |

Tablo 1 (devamı)

| | | 2. Bağımlı Değişken | |
|--------------------|---|--|--|
| Göstergeler | | Açıklama | |
| 4. | <i>Bağımlı değişken işevuruk olarak kesin sınırları ile tanımlanır.</i> | Bağımlı değişken doğrudan ölçüme uygun olarak tanımlanmalı ve somut/nicel ifadelerle açıklanmış olmalıdır. Dolayısıyla, bağımlı değişkenin değerlendirilme süreci, tanımlama doğrultusunda, geçerli ve tutarlı ölçüm yapabilmeye ve tekrarlanabilmeye olanak sağlamalıdır. | |
| 5. | <i>Bağımlı değişken ölçülebilir olarak ifade edilir.</i> | Bağımlı değişken için doğru ve yanlış tepkiler açıklanmalıdır. Becerinin ölçümünde hangi davranış gerçekleşti olarak kabul edilmek üzere olumlu değerlendirildiğine (“+” verilecek) ve hangi davranış gerçekleşmedi olarak kabul edilmek üzere olumsuz değerlendirildiğine (“-” verilecek) ilişkin bilgi verilmiş olması veya gerektiğinde hedef beceri için beceri analizinin olması gerekmektedir. | |
| 6. | <i>Bağımlı değişkeni ölçme süreci kesin sınırları ile yinelenebilir olarak tanımlanır.</i> | Bağımlı değişkeni ölçme süreci diğer araştırmacılar tarafından tekrarlanabilecek şekilde açıklanmalıdır. | |
| 7. | <i>Bağımlı değişkene ilişkin tekrarlı ölçümler yapılır.</i> | Bağımlı değişken için uygulama öncesinde ve sürecinde performansın karşılaştırılmasına olanak verecek şekilde yinelenen ölçümler olmalıdır. | |
| 8. | <i>Gözlemciler arası güvenilirlik verisi raporlanır.</i> | Çalışma sürecinde bağımlı değişkeni ölçme güvenilirliğine ilişkin gözlemciler (kodlayıcılar) arası güvenilirlik (GAG) verisi toplanmış olmalıdır. Ayrıca bu verilerin minimum oranları GAG değeri için %80; Kappa değeri için %60 düzeyinde olması gerekmektedir. | |
| | | 3. Bağımsız Değişken | |
| Göstergeler | | Açıklama | |
| 9. | <i>Bağımsız değişken kesin sınırları ile yinelenebilir olarak tanımlanır.</i> | Uygulamanın nasıl gerçekleştirildiği (ör. 1:1 öğretim, haftada 3 kez) açıklanmış olmalı, uygulama sürecinde kullanılan araç-gereçlere (ör A4 çalışma kağıdı) gerektiğinde betimleme yaparak yer verilmeli ve uygulama süreci (ör. pekiştirme süreci) yeterli açıklıkta olmalıdır. | |
| 10. | <i>Bağımsız değişken sistematik olarak manipüle edilir.</i> | Bağımsız değişken yapılandırılmış bir şekilde uygulanmış olmalıdır. Uygulamacı bağımsız değişkeni her defasında katılımcıya planlandığı şekilde sunmuş ve/veya gerektiğinde geriye çekmiş olmalıdır. | |
| 11. | <i>Uygulama güvenilirliği tanımlanır.</i> | Bağımsız değişkene ilişkin uygulama güvenilirliği verisi toplanmış olmalıdır. | |
| | | 4. Başlama Düzeyi | |
| Göstergeler | | Açıklama | |
| 12. | <i>Başlama düzeyi evresi, uygulamadan önce, hedef davranışa ve hedef davranışın ileride göstereceği örüntüye ilişkin bilgi sağlamaktadır.</i> | Başlama düzeyinde uygulama yapılmadığından emin olunmalıdır. Başlama düzeyinde performansla ilişkin kararlı veri elde edilmiş olmalıdır. Başlama düzeyinde kabul edilebilir veri nokta sayısı beş olmalıdır. | |
| 13. | <i>Başlama düzeyi oturum planı kesin sınırları ile yinelenebilir olarak tanımlanır.</i> | Başlama düzeyi oturum planı yineleme yapmaya olanak verecek şekilde açıklanmalıdır. | |

Tablo 1 (devamı)

| | | 5. Geçerlik | |
|-----|---|-------------|---|
| | Göstergeler | | Açıklama |
| 14. | <i>DeneySEL etki en az üç farklı durumda gösterilir.</i> | | Araştırmada deneySEL etki en az üç kez gösterilmelidir. Bir diğer ifadeyle, araştırma tasarımı uygulamanın tahminde bulunma-doğrulama-yineleme boyutlarını katılımcı içerisinde veya katılımcılar arasında gösterir nitelikte olmalıdır. |
| 15. | <i>Tasarım iç geçerliğe ilişkin tehditleri kontrol eder.</i> | | Bir önceki madde ile ilişkili olarak deneySEL etki üç durumda (tahminde bulunma-doğrulama-yineleme) varsa o araştırma tasarımının iç geçerliği sağladığı söylenebilir. Bağımlı değişkendeki değişimin yalnızca bağımsız değişkenden kaynaklandığı uygulama süreci ve model üzerinde tahminde bulunma-doğrulama ve yineleme gerçekleştirilerek gösterilmelidir. |
| 16. | <i>Sonuçlar deneySEL kontrolün sağlandığına ilişkin bir örnek/kanıt sunmaktadır.</i> | | DeneySEL etkinin üç gösterimi (tahminde bulunma-doğrulama-yineleme) olmakla birlikte bu üç gösterimde hedef beceriye ilişkin performansın ölçütü karşılar düzeyde sonuçlanmış olduğu görülmelidir. |
| 17. | <i>DeneySEL etki yinelenmekte, dış geçerlik sağlanmaktadır.</i> | | Uygulama; farklı katılımcılarla, farklı ortamlarda, farklı davranışlarla veya farklı materyallerle yürütülmüş olmalıdır. Uygulamanın etkisi tekrar tekrar en az üç durum için (ör. üç farklı katılımcı, bir katılımcıda üç farklı davranış) sınanmış olmalıdır. Ayrıca çalışmalarda genelleme verisi toplanmışsa uygulamanın etkisinin yinelenmesi ve dış geçerliğin sağlanması oldukça artmış olacaktır. |
| 18. | <i>Bağımlı değişken sosyal olarak önemlidir.</i> | | Bağımlı değişken, kazandırılması hedeflenen beceri, sadece kavramsal bir teoriyi karşılamamalıdır. Aynı zamanda bireyin akademik veya günlük yaşamında (ör. Bireyselleştirilmiş Eğitim Programı'nda [BEP] olması, günlük yaşamında kullanması) önemli bir yer tutmalıdır. Ayrıca katılımcının öğretmeni, ailesi gibi yakın çevresinin becerinin önemine ilişkin görüş bildirmesi de o becerinin sosyal açıdan önemli olduğunun bir göstergesidir. |
| 19. | <i>Uygulamaya bağlı olarak bağımlı değişkende meydana gelen değişikliğin büyüklüğü sosyal olarak önemlidir.</i> | | Uygulama sonucunda becerinin ediniminin bireyin kendisi, öğretmenleri, ailesi veya akrabaları gibi yakın çevresi için yararlı olduğunun belirtilmesidir. Sosyal geçerlik verisi elde edilmiş çalışmalarda çoğunlukla bu değişkene ilişkin açıklamalar yer almaktadır. Ancak, sosyal geçerlik verisi sağlanmasa da araştırmalarda tartışma kısmında bu açıklamalar yer alabilmektedir. Burada dikkat edilmesi gereken seçilen beceri değil uygulama sonucunun sosyal önemidir. |
| 20. | <i>Bağımsız değişken ekonomiktir ve/veya uygulanması pratiktir.</i> | | Bağımsız değişkenin uygulamacı açısından uygulanabilir, kolay ve/veya ekonomik olması durumudur. Ayrıca, uygulamacıların deneySEL uygulama bittikten sonra da uygulamayı tercih edebileceklerini belirtmeleri bu değişkeni karşılar niteliktedir. |
| 21. | <i>Bağımsız değişken zamanla <u>tipik</u> (doğal) bağlamlarda/kişilerin varlığında uygulanmaktadır.</i> | | Bireyin yakın çevresindekiler (ör. öğretmenler, aile, akran) tarafından veya kendisi tarafından deneySEL uygulamanın uygulanmış olması veya katılımcıya doğal bağlamlarda (ör. okul ortamında, toplumsal bir ortamda) uygulamanın sunulması durumu söz konusu olmalıdır. |

Tek-denekli deneySEL araştırma yöntemleri kullanılarak tasarlanmış bir araştırmanın Horner ve diğerlerinin (2005) geliştirmiş olduğu bilimsel-dayanaklı uygulamaları belirlemede “Tek-Denekli DeneySEL Araştırmaların Niteliksel Göstergeleri” açısından “kabul edilebilir” bir çalışma olarak değerlendirilebilmesi için o araştırmada sıralanan beş özelliğe ilişkin bilgi verilmiş olması gerekmektedir: (a) uygulamanın işlevsel tanımı, (b) bulguların işlevsel tanımı, (c) uygulama güvenilirliği, (d) uygulama ile elde edilen bulgular arasında işlevsel ilişki ve (e) deneySEL kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması. Bu özelliklerin tek-denekli deneySEL araştırma yöntemlerini kullanan bir çalışmada bulunup bulunmadığını belirlemek amacıyla 21

niteliksel göstergeden hangilerinin dikkate alınacağına ilişkin alanyazında görüş birliği yoktur. Bu çalışma kapsamında Tablo 2’de, OSB olan bireylere yönelik yürütülmüş matematik beceri öğretim çalışmaları incelenerek ve Horner ve diğerlerinin (2005) önerileri dikkate alınarak pratik bir değerlendirme süreci oluşturulmuştur. Böylece, bir çalışmanın “*kabul edilebilir*” olarak değerlendirilmesi için sıralanan beş özelliğe ilişkin söz konusu 21 niteliksel göstergeden altı göstergenin “*kabul edilebilirlik*” özelliğini karşılamak üzere kullanılabilmesi ileri sürülmüştür (Aydın, 2017). İlk dört “*kabul edilebilirlik*” ölçütü için Horner ve diğerlerinin önerileri doğrultusunda ilgili niteliksel göstergeler belirlenerek değerlendirme süreci oluşturulmuş ve benimsenmiştir. Ancak, “*deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması*” ölçütü için Horner ve diğerleri (2005) “*uygulamaya ilişkin hakemli dergilerce yayımlanmış, niteliksel göstergeleri minimum düzeyde karşılayan ve deneysel kontrolü sağlanmış en az beş çalışmanın yürütülmüş olması, araştırmaların en az üç farklı bölgede, üç farklı araştırmacı grubuyla yürütülmüş olması ve bu çalışmaların en az 20 katılımcıyla yürütülmüş olması*” gerektiğinden söz etmişlerdir. Bu çalışma kapsamında örnek değerlendirmeye alınan matematik beceri öğretim çalışmalarında tek bir uygulama incelenmediğinden ve alanyazında bu konuda sınırlı sayıda çalışma yürütüldüğünden söz konusu ölçütü karşılama konusunda çalışmalar değerlendirilmemiştir. Bu ölçütün karşılanmasında çalışmaların kendi içerisinde değerlendirilmesinin uygun olduğuna karar verilerek niteliksel ölçütlerden “*etki yinelenmiş, dış geçerlik sağlanmıştır*” göstergesinin “*kabul edilebilirlik*” açısından ölçüt olarak ele alınması benimsenmiştir. Ancak değerlendirme yapılan konuda çok sayıda çalışmanın olması durumunda deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmesi “*uygulamaya ilişkin en az beş çalışmanın yürütülmüş olması, araştırmaların en az üç farklı bölgede, üç farklı araştırmacı grubuyla yürütülmüş olması ve toplamda en az 20 katılımcıyla yürütülmüş olması*” ile değerlendirilmelidir.

Tablo 2’de görüldüğü gibi bir çalışmanın “*kabul edilebilir*” olarak değerlendirilmesinde “*uygulamanın işlevsel tanımı*” ölçütü için niteliksel göstergelerden (n = 21, Bk. Tablo-4) “*bağımsız değişken kesin sınırları ile yinelenebilir olarak tanımlanmış*” göstergesi (madde 9); “*bulguların işlevsel tanımı*” ölçütü için “*deneysel etkinin üç kanıtı/gösterimi*” ve “*tasarım iç geçerliğe ilişkin tehditleri kontrol etmiş*” göstergeleri (madde 14 ve 15); “*uygulama güvenilirliği*” ölçütü için “*uygulama güvenilirliği tanımlanmış*” göstergesi (madde 11); “*uygulama ile elde edilen bulgular arasında işlevsel ilişki*” ölçütü için “*sonuçlar deneysel kontrolün sağlandığına ilişkin bir örnek/kanıt sunmuş*” göstergesi (madde 16); “*deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması*” ölçütü için “*etki yinelenmiş, dış geçerlik sağlanmış*” göstergesi (madde 17) kullanılmıştır. Bu bağlamda, bir çalışmanın “*kabul edilebilir*” olarak değerlendirilmesinde 21 niteliksel göstergeden altı göstergenin (9., 11., 14., 15., 16. ve 17. maddeler) olumlu olarak kodlanması önerilmektedir.

Horner ve diğerleri (2005) tarafından geliştirilen bu niteliksel göstergelerin ve kabul edilebilirlik ölçütlerinin nasıl ele alınabileceği ve yorumlanabileceğine ilişkin alan yazında farklı yaklaşımların olduğu görülmektedir (ör. Chard, Ketterlin-Geller, Baker, Doabler ve Apichatabutra, 2009; Mayton, Wheeler, Menendez ve Zang, 2010; Rogers ve Graham, 2008). Bu çalışma kapsamında, Horner ve diğerleri (2005) tarafından geliştirilen niteliksel göstergelerin ve kabul edilebilirlik ölçütlerinin nasıl kullanılabilmesi ve yorumlanacağı OSB olan bireylere yönelik yürütülmüş matematik beceri öğretim çalışmaları kullanılarak ele alınmıştır.

Tablo 2

Tek-Denekli Deneysel Araştırmalarda “Kabul Edilebilirlik” Ölçütlerinin Belirlenmesi

| Kabul Edilebilirlik Ölçütleri | Horner ve Diğerlerinin (2005) Görüşleri | Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri Açısından Değerlendirilmesi |
|--|--|---|
| <i>Uygulamanın işlevsel tanımı</i> | Uygulamanın yürütülme sürecine ilişkin, başka araştırmacılar tarafından da uygulama sürecinin tekrarlanmasına olanak verebilecek nitelikte, detaylı bilgilere yer verilmesi gerekmektedir. | “Bağımsız değişken kesin sınırları ile yinelenebilir olarak tanımlanmış” göstergesi bu ölçütü karşılar niteliktedir (Bk. Tablo 4, Madde 9). |
| <i>Bulguların işlevsel tanımı</i> | Uygulama sürecinde kullanılan özel durumların tanımlanması, uygulamayı uygulayan araştırmacıların yetkinliği, katılımcı özelliklerinin tanımlanması ve bağımlı değişkene ilişkin çıktılarının uygulamadan kaynaklı olması özelliklerini içermelidir. | “Deneysel etkinin üç kanıtı/gösterimi” ve “tasarım iç geçerliğe ilişkin tehditleri kontrol etmiş” göstergeleri bu ölçütü karşılar niteliktedir (Bk. Tablo 4, Madde 14 ve 15). |
| <i>Uygulama güvenilirliği</i> | Uygulama güvenilir bir şekilde uygulanmış olmalıdır. | “Uygulama güvenilirliği tanımlanmış” göstergesi bu ölçütü karşılar niteliktedir (Bk. Tablo 4, Madde 11). |
| <i>Uygulama ile elde edilen bulgular arasında işlevsel ilişki</i> | Uygulama ile bağımlı değişkende meydana gelen değişim arasında işlevsel bir ilişkinin olması gerekmektedir. | “Sonuçlar deneysel kontrolün sağlandığına ilişkin bir örnek/kanıt sunmuş” göstergesi bu ölçütü karşılar niteliktedir (Bk. Tablo 4, Madde 16). |
| <i>Deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmesi</i> | Uygulamaya ilişkin en az beş çalışmanın yürütülmüş olması gerekir. Bu araştırmaların ise, en az üç farklı bölgede, üç farklı araştırmacı grubuyla ve en az 20 katılımcıyla yürütülmüş olması gerekmektedir. | “Etki yinelenmiş, dış geçerlik sağlanmış” göstergesi bu ölçütü karşılar niteliktedir (Bk. Tablo 4, Madde 17) |

Bilimsel-Dayanaklı Uygulamaları Belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” Yönergesinin Örnek Kullanımı

Tablo 3’te okuyucuya fikir vermesi bakımından matematik becerilerinin öğretimi ile ilgili alan yazında yer alan çalışmalardan makale kapsamında değerlendirilecek olanların çeşitli özelliklerine ilişkin bilgiler sunulmuştur. Bu çalışmaların niteliksel göstergeler açısından incelenmesine ve değerlendirilmesine geçilmeden önce 5 kategori altında 21 niteliksel göstergenin tamamını içeren bir kodlama formu (Bk. Tablo 4) ve her bir göstergenin nasıl puanlanacağına ilişkin açıklamalar geliştirilmiştir (Bk. Tablo 1).

Niteliksel göstergelerin kodlanması sırasında, forma çalışmaların künyesi yazıldıktan sonra çalışmanın her bir göstergesi karşılayıp karşılamadığı, çalışmanın söz konusu göstergesi karşılaması durumunda “Evet” anlamında “E”, karşılamaması durumunda “Hayır” anlamında “H” harfi kullanılarak ilgili sütunda belirtilmiştir. Tablo 4’te her bir çalışmanın niteliksel göstergelerin hangilerini karşıladığı ve hangilerini karşılamadığı gösterilmektedir. Araştırmacılar benzer şekilde çalışmaları değerlendirecekleri zaman bu kodlamaları birbirlerinden bağımsız en az iki kodlayıcı tarafından kodlanmasını sağlamalı ve kodlayıcılar arasında tutarlık analizi hesaplamalarını yapmalıdırlar. Kodlamacılar arasında görüş ayrılığının olması durumunda ise birbirlerini ikna ederek uzlaşma sağlamalıdırlar. Aşağıda, Tablo 4’te yer alan kodlamalar ışığında, 5 kategori altında ele alınan 21 niteliksel göstergenin her biri için göstergesi karşılayan ve karşılamayan çalışmalara örnekler sunulmuş, bir çalışmada ilgili göstergenin neden karşılandığı ya da karşılanmadığı makaleden elde edilen alıntılar ışığında gerekçeleriyle açıklanmıştır.

Tablo 3

Örnek Değerlendirme Kapsamında İncelenen Çalışmalar ve Betimleyici Bilgileri

| Çalışmalar | Katılımcı Özellikleri Sayı, Tanı/Tanı Aracı, Cinsiyet/Yaş veya Sınıf | Ortam ve Öğretim Düzenlemesi | Bağımlı Değişken | Bağımsız Değişken | Araştırma Modeli | Uygulamacı Bilgisi | Güvenirlilik GAG/UG | Genelleme/ İzleme/ Sosyal Geçerlik |
|-------------------------------------|--|--|--|---|---|----------------------------------|------------------------|--|
| Adcock ve Cuvo, 2009 | 3, YGB(1)-AS(1)-OSB(1),--, E(3)/7-8-10 | Terapi odası/Grup öğretimi | Çıkarma/çarpma/toplama/madeni paraları tanıma/zamanı söyleme | Harmanlanmış öğretim uygulaması | Çoklu yoklama modeli | Araştırmacı | +/- | -/+/- |
| Akmanoğlu ve Batu, 2004 | 3, Otizm(3)/--, E(2)-K(1)/12-17-6 | Özel eğitim sınıfı/ Üniversite bünyesindeki birim sınıfı/1:1 öğretim | Rakamları gösterme becerisi | Eşzamanlı ipucuyla öğretim | Çoklu yoklama modeli | Araştırmacı | +/+ | +/+/+ |
| Cihak ve Grim, 2008 | 4, Otizm(4)-IQ ₁ :50, IQ ₂ :47, IQ ₃ :45, IQ ₄ :35 /CARS-WISC IV, E(2)-K(2)/ 16-17-15-16 | Kaynaştırma sınıfı, okul kitap mağazası/1:1 öğretim | Satın alma becerisi | Sayma stratejisi | Çoklu yoklama modeli | Kaynaştırma sınıf öğretmeni | +/+ | +/+/- |
| Collins, Hager ve Galloway, 2011 | 1, Otizm-IQ=47/ UNIT, E/14 | Özel eğitim sınıfı/1:1 öğretim | Satış vergisi hesaplama | Sabit bekleme süreli öğretim | Çoklu yoklama modeli | Özel eğitim ve yardımcı öğretmen | +/+ | +/-/- |
| Kellems ve diğ., 2016 | 3, Otizm(3)-IQ ₁ :53-IQ ₂ :77-IQ ₃ :78/ WJ-Cog-WISC-III, E(2)-K(1)/19-20-21 | Bireysel sınıf odası/1:1 öğretim | Birim fiyat kıyaslama/bahşiş hesaplama/kişi sayısına göre tarif miktarını ayarlama | Video ipucu uygulaması | Çoklu yoklama modeli | Araştırmacı | +/+ | -/+/+ |
| Morrison ve Rosales-Luiz, 1997 | 1, Otizm-IQ:36/BSID, E/5 | Ev/1:1 öğretim | Sayma becerisi | Tercih edilen nesnelere yapılan öğretim | Dönüşümlü uygulamalar modeli | Araştırmacı | +/- | -/-/- |
| Levingston, Neef ve Cihon, 2009 | 1, Otizm/--, E/10 | Genel eğitim sınıfı/1:1 öğretim | Çarpma ve bölme sözel problemleri | İşlem öncesi becerilerin edinimi | Çoklu başlama düzeyi modeli | Araştırmacı | +/+ | +/-/- |
| Root, Browder, Saunders ve Lo, 2017 | 3, OSB(3)-IQ ₁ :58-IQ ₂ :46-IQ ₃ :62/ LIPS R-Stanford Binet-CARS-GARS II, E(3)/7-9-11 | Bireysel sınıf odası/1:1 öğretim | Karşılaştırma sözel problemlerini çözme | Şema yaklaşımı öğretimi | Çoklu yoklama ve dönüşümlü uygulamalar modeli | Özel eğitim öğretmeni | +/+ | -/+/+ |
| Bouxsein ve Fisher, 2007 | 1, AS/--, E-19 | Terapi odası/1:1 öğretim | Kolay-orta-zor sorulara akıcı, orta düzeye doğru cevap verme | Ayrımlı pekiştirme | Tersine çevirme modeli | Terapist | +/- | +/-/- |
| Waters ve Boon, 2011 | 2, Otizm(1)-AS(1)-IQ ₁ :64-IQ ₂ :64/ WISC-III, E(2)/15-16(1 ay) | Özel eğitim sınıfı/1:1 öğretim | Ondalık para değerlerini çıkarma | Nokta belirleme tekniği | Çoklu yoklama modeli | Özel eğitim öğretmeni | +/+ | +/+/+ |

E: Erkek, K: Kadın; AS: Asperger Sendromu, OSB: Otizm Spektrum Bozukluğu, YGB: Yaygın Gelişimsel Bozukluk; IQ: Intelligence Quotient; BSID: Bayley Scales of Infant Development, CARS: Childhood Autism Rating Scale, GARS: Gilliam Autism Rating Scale, LIPS-R: Leiter International Performance Scale-Revised, UNIT: Universal Nonverbal Test of Intelligence; WJ-Cog: Woodcock Johnson Tests of Cognitive Abilities, WISC: Wechsler Intelligence Scale for Children; GAG: Gözlemciler Arası Güvenirlilik, UG: Uygulama Güvenirliliği.

Tablo 4

Bilimsel-Dayanaklı Uygulamaları Belirlemede “Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri” Yönergesi Açısından Çalışmaların Örnek Değerlendirilmesi

| Göstergeler | Adcock ve Cuvo (2009) | Akmanoğlu ve Batu (2004) | Cihak ve Grim (2008) | Collins ve diğerleri, (2011) | Kellems ve diğerleri, (2016) | Levingston ve diğerleri, (2009) | Morrison ve Rosales-Luiz (1997) | Root ve diğerleri, (2017) | Tiger ve diğerleri, (2007) | Waters ve Boon (2011) |
|---|-----------------------|--------------------------|----------------------|------------------------------|------------------------------|---------------------------------|---------------------------------|---------------------------|----------------------------|-----------------------|
| Katılımcılar ve Ortamlar | | | | | | | | | | |
| 1. Yeterli tanımlanmıştır. | E | E | E | E | E | E | E | E | E | E |
| 2. Seçilme süreci yeterli tanımlanmıştır. | H | E | E | E | H | E | H | E | E | E |
| 3. Ortam yeterli tanımlanmıştır. | E | E | H | E | H | H | H | E | E | E |
| Bağımlı Değişken | | | | | | | | | | |
| 4. İşevuruk olarak kesin sınırları ile tanımlanmıştır. | E | E | E | E | E | E | E | E | E | E |
| 5. Ölçülebilirdir. | H | E | E | E | E | E | E | E | E | E |
| 6. Ölçme kesin sınırları ile yinelenabilir olarak tanımlanmıştır. | H | E | E | E | E | E | E | E | E | E |
| 7. Tekrarlı ölçümler yapılmıştır. | E | E | E | E | E | E | E | E | E | E |
| 8. Gözlemciler arası güvenilirlik verisi raporlanmıştır. | E | E | E | E | E | E | E | E | E | E |
| Bağımsız Değişken | | | | | | | | | | |
| 9. Kesin sınırları ile yinelenabilir olarak tanımlanmıştır. | E | E | E | E | E | E | E | E | E | E |
| 10. Sistematik olarak manipule edilmiştir. | E | E | E | E | E | E | E | E | E | E |
| 11. Uygulama güvenilirliği tanımlanmıştır. | H | E | E | E | E | E | H | E | H | E |
| Başlama Düzeyi | | | | | | | | | | |
| 12. Evre, uygulamadan önce, örüntüye (pattern) kanıt sağlamıştır. | E | E | E | E | E | E | H | E | E | E |
| 13. Kesin sınırları ile yinelenebilir olarak tanımlanmıştır. | E | E | E | E | E | H | H | E | H | H |
| Geçerlik | | | | | | | | | | |
| 14. Deneysel etkinin üç kanıtı/gösterimi vardır. | E | E | E | E | E | H | H | E | E | E |
| 15. Tasarım iç geçerliğe ilişkin tehditleri kontrol etmiştir. | E | E | E | E | E | H | H | E | H | E |
| 16. Sonuçlar deneysel kontrolün sağlandığına ilişkin bir örnek/kanıt sunmuştur. | E | E | E | H | E | H | H | E | H | E |
| 17. Etki yinelenmiş, dış geçerlik sağlanmıştır. | E | E | E | E | E | E | H | E | H | E |
| 18. Bağımlı değişken sosyal olarak önemlidir. | E | E | E | E | E | E | H | E | E | E |
| 19. Uygulamaya bağlı olarak bağımlı değişkende meydana gelen değişikliğin büyüklüğü sosyal olarak önemlidir | H | E | H | H | E | H | H | E | H | E |
| 20. Bağımsız değişken uygun maliyetli ve/veya pratiktir. | H | E | H | H | E | H | H | H | H | E |
| 21. Bağımsız değişken zamanla tipik (doğal) bağlamlarda/kişilerin varlığında uygulanmıştır. | H | H | E | E | E | E | H | E | H | E |
| Karşılanan göstergeler/toplam göstergeler | 14/21 | 20/21 | 18/21 | 18/21 | 19/21 | 14/21 | 08/21 | 20/21 | 13/21 | 20/21 |

Katılımcılar ve Ortamlar. Katılımcı ve ortam kategorisi kapsamında yer alan üç niteliksel göstergenin tamamını karşıladı olarak kodlanan (“E” ile kodlanan) çalışma (Akmanoğlu ve Batu, 2004) ile bu üç niteliksel göstergenin ikisini karşılamadı olarak kodlanan (“H” ile kodlanan) çalışma (Kellems ve diğ., 2016) kapsamında ilişkili göstergeler ile ilgili sunulan bilgiler Tablo 5’te görülmektedir.

Tablo 5

Katılımcı ve Ortam Açısından Çalışmalarda Sunulan Bilgiler

| Gösterge | Akmanoğlu ve Batu, 2004 | | Kellems ve diğerleri, 2016 | |
|---|-------------------------|--|----------------------------|--|
| | E/H | Alıntılar | E/H | Alıntılar |
| 1. Yeterli tanımlanmıştır. | E | <ul style="list-style-type: none"> “Katılımcılar yaşları 6, 12 ve 17 olan 2 erkek, 1 kız olmak üzere üç otizmlili öğrencidir.” “Erkek öğrenciler aynı zamanda zihinsel yetersizlik tanısına sahiptir. Hiçbir öğrencinin temel tepki ipuçlarına ilişkin öncesinde sistematik aldığı bir eğitim yoktur.” “Erkek öğrenciler özel bir özel eğitim okuluna giderken, kız öğrenci bir üniversitede gelişimsel yetersizliği olan çocukların eğitim gördüğü merkeze gitmektedir.” “Ömer, bağımsız olarak, elbisesini giyip çıkarabilme ve tuvaletini yapabilme gibi temel öz bakım becerilerine sahiptir...” “Barış kendisine söylendiğinde bazı ince motor becerileri (ör. kalem tutma, şekli boyama), kaba motor becerileri (ör. top tutma, grup oyunu oynama) ve öz bakım becerilerini (ör. yemek yeme) bağımsız olarak yapabilmektedir...” “...Serap basit yönergeleri anlayabilmekte, iki kelimeli cümleler kurabilmekte ve ekolali konuşmaktadır.” | E | <ul style="list-style-type: none"> “Çalışmaya dokuz yetişkin katılımcı katılmıştır. Yaşları 18-21 arasında değişmekte ve otizm, zihinsel yetersizlik, öğrenme güçlüğü ve diğer sağlık bozuklukları dâhil olmak üzere çeşitli yetersizliğe sahiptirler.” Katılımcıların her birine ilişkin yaş, cinsiyet, zeka testinden aldıkları puanlar, matematik performans düzeyleri, uyumsal davranış değerlendirme puanları çalışma içerisinde sunulan tablo da yer almaktadır (Bk. Kellems ve diğ., 2016, s. 34). “Katılımcıların tamamı IDEA kapsamında özel eğitim hizmetlerinden yararlanmaktadırlar.” |
| 2. Seçilme süreci yeterli tanımlanmıştır. | E | <ul style="list-style-type: none"> “Katılımcılar ilk yazar tarafından şu ön koşul becerilere dayalı olarak seçilmiştir: (a) etkinliğe dikkatini en az 10 dk. verme (pilot çalışma süresince bir oturumun 10 dk. sürdüğü belirlendiği için), (b) basit sözel yönergeleri başarılı bir şekilde takip edebilme (ör. bana ... ver, bana ... göster) ve (c) verilen üç nesneden adı söylenen nesneyi seçme” “Ayrıca, numaraları adlandırma yeni bir beceridir ve katılımcılar tarafından henüz öğrenilmemiştir.” | H | <ul style="list-style-type: none"> “İnsan katılımcılar için gerekli olan etik kurul onayı ilgili tüm kuruluşlardan alınmıştır.” Katılımcıların kim tarafından seçildiği, nasıl belirlendiği bilgileri çalışmada yer almamaktadır. |
| 3. Ortam yeterli tanımlanmıştır. | E | <ul style="list-style-type: none"> “Ömer ve Barış için tüm oturumlar kendi okullarındaki bir sınıfta yürütülmüştür. Serap için tüm oturumlar üniversite eğitim merkezinde bir sınıfta yürütülmüştür.” “Her iki sınıfta da bir masa ve iki sandalye yer almaktadır... Bir video kamera oturumları kaydetmek için kullanılmıştır.” “Öğrenci ve eğitimci birbirleriyle yüz yüze gelecek şekilde masada oturmuşlardır...” | H | <ul style="list-style-type: none"> “Çalışma batıda 68,000 kayıtlı öğrencinin olduğu bir banliyö okul bölgesinde yürütülmüştür.” “Tüm oturumlar, okuldaki aynı odada düzenlenmiştir. Çalışma süresince oda başka bir aktivite için kullanılmamıştır.” “Katılımcılar gerekli materyallerin yerleştirildiği bir masaya oturmuşlardır.” “İki araştırma asistanı katılımcılarla çalıştıkları süresince davranış kontrol listesi kullanarak veri toplamışlardır.” |

Her iki çalışmada da “katılımcılar yeterli tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Çünkü iki çalışmada da çalışmaya katılan katılımcı sayısı, katılımcıların yaşları, cinsiyetleri ve tanıları belirtilmiştir. Her bir katılımcının performanslarına ilişkin bilgiler sunulmuştur. Ayrıca, Kellems ve diğerlerinin (2016) yürüttükleri çalışmada katılımcıları tanılamada kullanılan testler ve aldıkları puanları belirten bilgiler de mevcuttur. Akmanoğlu ve Batu’nun (2004) yürüttükleri çalışmada, katılımcıların belirlenen önkoşul becerileri sergileyip sergileyememelerine göre araştırmacı tarafından seçilmelerinden dolayı ve çalışma kapsamında hedeflenen becerilerin katılımcılar tarafından öğrenilecek yeni beceriler olmasının belirtilmesinden dolayı “seçilme süreci yeterli tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Buna karşın, Kellems ve diğerlerinin (2016) yürüttükleri çalışmada katılımcı özelliklerinden söz edilmesine karşın çalışmaya katılım ile ilgili önkoşul beceriler açıkça belirtilmemesinden ve katılımcıların kim tarafından seçildiği, nasıl belirlendiği gibi bilgilerin yer almamasından dolayı ilgili gösterge “H” olarak kodlanmıştır. “Ortam yeterli tanımlanmıştır” göstergesi için Akmanoğlu ve Batu’nun yürüttükleri çalışmada her bir katılımcı için çalışmanın yürütüleceği ortamın neresi olduğuna dair bilgi sunulmuştur. Ortamda bulunan mobilya (ör. masa, sandalye) ve araç-gereçler (ör. video kamera kaydedici) açıklanmıştır. Ortamda kimlerin yer aldığı ve katılımcı-uygulamacı çalışma pozisyonu belirtilmiştir. Bu nedenle ilgili gösterge Akmanoğlu ve Batu’nun çalışması için “E” olarak kodlanmıştır. Ancak, Kellems ve diğerlerinin yürüttükleri çalışmada okula ilişkin ayrıntılı açıklamaya yer verilmemiş, oturumların yürütüleceği ortam hakkında bilgi verilmesine karşın ortam özelliklerine yönelik (ör. araç-gereçler, mobilyalar) detaylı bir bilgi sunulmamış ve katılımcı-uygulamacı çalışma pozisyonu (oturma pozisyonları) açıklanmamıştır. Bu nedenle ilgili gösterge bu çalışma için “H” olarak kodlanmıştır.

Bağımlı Değişken. Bağımlı değişken ölçütü kapsamında yer alan beş niteliksel göstergenin tamamı karşıladı olarak kodlanan (“E” olarak kodlanan) çalışmaya (Collins ve diğ., 2011) ve bu beş niteliksel göstergenin üçü karşıladı olarak kodlanan (“E” olarak kodlanan) ikisi ise karşılamadı olarak kodlanan (“H” olarak kodlanan) çalışmaya (Adcock ve Cuvo, 2009) ilişkin sunulan bilgiler Tablo 6’da görülmektedir.

Tablo 6
Bağımlı Değişken Açısından Çalışmalarda Sunulan Bilgiler

| Bağımlı Değişken | Collins ve diğerleri, 2011 | | Adcock ve Cuvo, 2009 | |
|--|----------------------------|---|----------------------|---|
| | E/H | Alıntılar | E/H | Alıntılar |
| 4. İşevuruk olarak kesin sınırları ile tanımlanmıştır. | E | <ul style="list-style-type: none"> • Jason, Morgan ve Rena için “<i>Gazete reklamlarında yer alan ürünler için satış vergisi ücretinin hesap makinasıyla hesaplanması</i>” bağımlı değişken olarak ifade edilmiştir. • Ölçülebilirlik kısmında bağımlı değişken (satış vergisi hesaplama) somut ve nicel ifadelerle açıklanmıştır. | E | <ul style="list-style-type: none"> • Herbir katılımcı için edinim ve kalıcılık bağımlı değişkenleri (ör. “<i>iki basamaklı sayıdan tek basamaklı sayının çıkarılması, çeyrek kalalı saati söyleme becerisi</i>” gibi) açıklanmıştır (Bk. Tablo 1, Adcock ve Cuvo, 2009, s. 322). |
| 5. Ölçülebilirdir. | E | <ul style="list-style-type: none"> • “<i>Jason için beceri analizi ürünün fiyatını gir, X bas, 6’ya bas, % bas, kâğıda miktarı yaz, temizle düğmesine bas, 1. hesabı gir, + bas, vergi miktarını gir, = bas ve toplam miktarı kâğıda yaz.</i>” • “<i>Öğretmen Rena ve Morgan yazma bilmediği için beceri analizini onlara yönelik uyarlamıştır...</i>” • “<i>Bu amaçların herbiri için ölçüt üç oturum boyunca %100 doğrulukta sergilemektedir.</i>” | H | <ul style="list-style-type: none"> • Bağımlı değişkenlere ilişkin bir beceri analizi yoktur. • Çalışma bağımlı değişkeni ölçüm (measurement, data collection) sürecine ilişkin bir başlık veya bilgi içermemektedir. |

Tablo 6 (devamı)

| Bağımlı Değişken | Collins ve diğerleri, 2011 | | Adcock ve Cuvo, 2009 | |
|---|----------------------------|---|----------------------|---|
| | E/H | Alıntılar | E/H | Alıntılar |
| 6. Ölçme kesin sınırları ile yinelenabilir olarak tanımlanmıştır. | E | <ul style="list-style-type: none"> • “Öğretim süresi boyunca eğitmen; ipuculu ve ipucusuz tüm doğru tepkileri +, ipuculu ve ipucusuz tüm yanlış tepkileri -, ipucunun ardından tepki verilmediğinde 0 olarak kaydetmiştir.” • “Herbir doğru tepki için eğitmen pekiştireç (praise) sunmuş ve her bir yanlış tepki ya da tepkide bulunmama için eğitmen doğru tepkiye ilişkin model olmuştur.” | H | <ul style="list-style-type: none"> • Ölçüm süreci açıklanmamıştır. • “... Çocuk jestsel ipucuyla doğru tepkide bulunursa sonraki edinim becerisine geçilmiştir. Çocuk yanlış tepki verirse jestsel ipucu ard arda üç deneme yeniden verilmiştir...” • “Edinim becerisi sözlü bir tepki gerektiriyorsa ve hata yapıldıysa deneme tam sözlü bir ipucuyla sunulmuştur...” |
| 7. Tekrarlı ölçümler yapılmıştır. | E | <ul style="list-style-type: none"> • Başlama düzeyi, uygulama ve izleme evrelerinde performans karşılaştırılması yapmaya olanak verecek düzeyde bağımlı değişkenin yinelenen ölçümleri grafikler üzerinde görülmektedir (Bk. Collins ve diğ., 2011, s. 30, 31 ve 32). | E | <ul style="list-style-type: none"> • Evreler arasında bağımlı değişkene yönelik performans karşılaştırılması yapmaya olanak verecek düzeyde yinelenen ölçümler grafikler üzerinde görülmektedir (Bk. Adcock ve Cuvo, 2009, s. 325 ve 326). |
| 8. Gözlemciler arası güvenilirlik verisi rapor edilmiştir. | E | <ul style="list-style-type: none"> • “Bağımlı değişkene ilişkin araştırma sürecinde gözlemciler arası güvenilirlik verisi %99.1’dir (ranj: %81.8-%100).” | E | <ul style="list-style-type: none"> • “Güvenirlilik verisi oturumların %33’ünden elde edilmiştir.” • “Edinim becerilerinin tamamı için gözlemciler arası güvenilirlik verisi ortalama %98’dir... Kalıcılık becerileri için gözlemciler arası güvenilirlik verisi %100’dür.” |

Collins ve diğerleri (2011) ve Adcock ve Cuvo (2009) tarafından yürütülen çalışmalarda bağımlı değişken “işevruk olarak tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Çünkü her iki çalışmada da, bağımlı değişkenler somut ve nicel ifadelerle açıklanmıştır. Collins ve diğerlerinin (2011) yürüttükleri çalışmada bağımlı değişkenin gerçekleştirilmesine yönelik bir beceri analizi oluşturulmuş ve katılımcının bu beceri basamaklarının her birini üç oturum boyunca %100 doğrulukta sergiler olması ölçüt olarak belirtildiği için bağımlı değişkenin “ölçülebilir” olmasına ilişkin gösterge “E” olarak kodlanmıştır. Buna karşın, Adcock ve Cuvo (2009) tarafından yürütülen çalışmada, bağımlı değişkene ilişkin bir beceri analizinin olmayışı, doğru ve yanlış tepkilerden söz edilmesine karşın bağımlı değişkenleri ölçme sürecinde (ör. çıkarma işlemi, çarpma işlemi) olumlu ve olumsuz tepkilerin nasıl değerlendirildiği yinelenabilir açıklıkta betimlenmemesinden dolayı ilgili gösterge “H” olarak kodlanmıştır. Collins ve diğerlerinin yürüttükleri çalışmada “ölçme kesin sınırları ile yinelenebilir açıklıkta tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Çünkü çalışmada katılımcıların beceri analizinde yer alan basamakları sergileyip sergilememelerine göre (başlama düzeyinde, uygulamada ve yoklamada) uygulamacının nasıl değerlendireceğine ilişkin detaylı bilgi sunulmuştur. Ayrıca, çalışmada uygulamacının söz konusu evrelerde katılımcının kaç saniye içerisinde verdiği doğru tepkilere “+” verileceği, yanlış tepkilere “-” verileceği veya tepkide bulunmayınca nasıl değerlendireceği ve ölçme sürecinde pekiştireç kullanma şekli yineleme yapmaya olanak verecek şekilde açıklanmıştır. Ancak, Adcock ve Cuvo’nun yürüttükleri çalışmada ise doğru ve yanlış tepkiler tam olarak belirtilmediği için, diğer deyişle bir beceri analizi oluşturulmadığı için, uygulama sürecinde becerilerin değerlendirilme süreci yineleme yapabilecek nitelikte açık değildir. Ayrıca, uygulama sürecine ilişkin detaylı bilgi sunulmuştur ancak bu bağımsız değişkenin detaylı açıklanması ile ilişkili bir durum olup bağımlı değişkenin nasıl ölçümlendiğini açıklamamaktadır. Tüm bu nedenlerden dolayı, ilgili gösterge bu çalışma için “H”

olarak kodlanmıştır. Her iki çalışmada da, bağımlı değişkenin yinelenen ölçümlerinin evreler arasında yapılmasından dolayı “tekrarlı ölçümler yapılmıştır” göstergesi “E” olarak kodlanmıştır. Bununla birlikte, iki çalışmada da bağımlı değişkeni ölçmeye ilişkin gözlemciler arası güvenilirlik verisi elde edilmiş ve rapor edilmiştir. Dolayısıyla her iki çalışma için de “gözlemciler arası güvenilirlik verisi rapor edilmiştir” göstergesi “E” olarak kodlanmıştır.

Bağımsız Değişken. Bağımsız değişken ölçütü kapsamında yer alan üç niteliksel göstergenin tamamı karşıladı olarak kodlanan (“E” olarak kodlanan) çalışmaya (Root ve diğ., 2017) ve bu üç niteliksel göstergenin ikisi karşıladı olarak kodlanan (“E” olarak kodlanan) biri ise karşılamadı olarak kodlanan (“H” olarak kodlanan) çalışmaya (Tiger ve diğ., 2007) ilişkin bilgiler Tablo 7’de görülmektedir.

Tablo 7

Bağımsız Değişken Açısından Çalışmalarda Sunulan Bilgiler

| Bağımsız Değişken | Root ve diğerleri, 2017 | | Tiger ve diğerleri, 2007 | |
|---|-------------------------|--|--------------------------|---|
| | E/H | Alıntılar | E/H | Alıntılar |
| 9. Kesin sınırları ile yinelenebilir olarak tanımlanmıştır. | E | <ul style="list-style-type: none"> Bağımsız değişken (şema yaklaşımıyla yapılan öğretim) görsel bir platformda tanıtılmıştır (Bk. Root ve diğ., 2017, s. 4). “İlk yazar uygulamaların tamamını herbir katılımcıya birebir uygulamıştır.” <ul style="list-style-type: none"> “Uygulama oturumları 10-15 dk sürmüştür...” “Herbir uygulama oturumundaki derseler ipucunun giderek artırılması stratejisinin içine gömülen doğrudan öğretim formatında gerçekleştirilmiştir.” <ul style="list-style-type: none"> “... İpucunun giderek artırılması hiyerarşisi üç düzeyden (sözel, tam sözel ve model ipucu) oluşmaktadır.” “... Model ipucunda eğitmen basamakta gerekli doğru tepkiye tam model olmuştur...” “Çalışmada kullanılan materyaller; (a) bir problem çözme matı, (b) laminant kaplı kâğıt formatında ya da iPad 3’te gösterilen bir grafik organizatörü (şema diyagramı), (c) laminant ile kaplı öğrenci öz-egitim kâğıdı ve (d) karşılaştırma türü sözel problemler içermektedir.” | E | <ul style="list-style-type: none"> “Ayrımlı pekiştirme (kısa sürede cevap verme) uygulamasında oturumlarda kolay veya zor problemler 5 sn.’den daha az sürede cevaplanırsa (doğruya yanlışa bakılmaksızın) kısa bir sözlü pekiştirme ve video izlemeye dönüştürülebilir bir sembol pekiştireç sunulmuştur.” “Doğru cevabı ayrımlı pekiştirme uygulama oturumlarında orta düzeyde sorulara doğru cevap verilmesi (cevaplama süresine bakılmaksızın) kısa sürede cevap verme uygulamasında olduğu gibi pekiştirilmiştir.” “Kolay problemler tek ve çift basamaklı çarpma ve toplama işlemlerinden ve üslü işlemlerden oluşmaktadır...Orta düzey zorluktaki problemler, uzun bölme işlemi gerektiren işlemlerden... Zor problemler, geometri, trigonometri ve integral problemlerinden oluşmaktadır.” |
| 10. Sistematik olarak manipüle edilmiştir. | E | <ul style="list-style-type: none"> “Üç öğretim oturumu boyunca üç sözel problemin çözümünde, eğitmen ipucunun giderek artırılması uygulamasını katılımcılar gereksinim duydukça sunmuştur...” “... Genellikle, sözel ipucu katılımcılara sonraki adımın ne olduğunu görmeleri için kendine yönerge verme listesine bakmalarında yöneltilmiştir...” | E | <ul style="list-style-type: none"> “Her bir oturumda üç problem türü tesadüfi olarak sunulmuştur. Her deneme terapistin matematik probleminin olduğu kartı katılımcıya vermesiyle başlamıştır...” Ayrımlı pekiştirme uygulaması kolay, orta ve zor düzey sorular için madde 9’da belirtildiği şekli ile sunulmuştur. |
| 11. Uygulama güvenilirliği tanımlanmıştır. | E | <ul style="list-style-type: none"> “Uygulama güvenilirliği oranı %98 - %100 arasında %99 elde edilmiştir.” | H | <ul style="list-style-type: none"> Uygulama güvenilirliğine ilişkin veri elde edilmemiş ve araştırmada rapor edilmemiştir. |

Bağımsız değişken “kesin sınırları ile yinelenabilir olarak tanımlanmıştır” göstergesi örnek değerlendirilen her iki çalışmada da “E” olarak kodlanmıştır. Çünkü her iki çalışmada da uygulama süreci detaylandırılmış ayrıca; tepkiler, öğretim düzenlemesi, uygulama sürecindeki pekiştirme durumları gibi içerikler yinelenebilir açıklıkta betimlenmiştir. Bağımsız değişken “sistemik olarak manipüle edilmiştir” göstergesi, her iki çalışmada da bağımsız değişkenin uygulanması belli bir sistemik içerisinde planlanmış ve planlandığı şekli uygulandığı için “E” olarak kodlanmıştır. Root ve diğerlerinin (2017) çalışmasında uygulama güvenilirliğine ilişkin bir kontrol listesi oluşturulmuş (11 maddelik) ve bağımsız bir kodlayıcı uygulama sürecinin bir kısmını izleyerek söz konusu maddeleri kodlamış ve güvenilirlik verisi elde edilmiştir. Bu nedenle, bu çalışma için “uygulama güvenilirliği tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Buna karşın, Tiger ve diğerlerinin (2007) çalışmasında uygulama güvenilirliğine ilişkin bir veri elde edilmediğinden ilgili gösterge “H” olarak kodlanmıştır.

Başlama Düzeyi. Başlama düzeyi ölçütü kapsamında yer alan iki niteliksel göstergenin tamamı karşıladı olarak kodlanan (“E” olarak kodlanan) çalışmaya (Cihak ve Grim, 2008) ile bu iki niteliksel göstergenin tamamı karşılamadı olarak kodlanan (“H” olarak kodlanan) çalışmaya (Morrison ve Rosales-Luiz, 1997) ilişkin sunulan bilgiler Tablo 8’de görülmektedir.

Tablo 8

Başlama Düzeyi Açısından Çalışmalarda Sunulan Bilgiler

| Başlama Düzeyi | Cihak ve Grim, 2008 | | Morrison ve Rosales-Luiz, 1997 | |
|---|---------------------|--|--------------------------------|---|
| | E/H | Alıntılar | E/H | Alıntılar |
| 12. Evre, uygulamadan önce, örüntüye (pattern) kanıt sağlamıştır. | E | <ul style="list-style-type: none"> “... <i>Ek bir yardım, öğretim ya da dönüt başlama düzeyi boyunca sunulmamıştır.</i>” Grafiksel verilerde (Bk. Cihak ve Grim, 2008, s. 722 ve 723) başlama düzeyinde en az beş oturumda kararlı veri elde edildiği görülmektedir “<i>Öğrencilere üç farklı ücret aralığı setinin (5.00-9.99\$, 10.00-14.99\$ ve 15.00-20.00\$) herbirinden rastlantısal 10 ücret sunulmuştur. Öğretmen öğrencinin dikkatini çekmiş ve ardından flash kartta yazan ücreti gösterirken sözlü olarak ücreti belirtmiştir. Öğretmen öğrencinin tepkide bulunması için (ücreti denkleştirmek için bağımsız olarak para sayması) 10 sn. fırsat sunmuş ve daha sonra bağımsız ya da yanlış tepkileri kaydetmiştir.</i>” | H | <ul style="list-style-type: none"> Başlama düzeyi oturumları düzenlenmemiştir. |
| 13. Kesin sınırları ile yinelenebilir olarak tanımlanmıştır. | E | <ul style="list-style-type: none"> “... <i>Ek bir yardım, öğretim ya da dönüt başlama düzeyi boyunca sunulmamıştır.</i>” Grafiksel verilerde (Bk. Cihak ve Grim, 2008, s. 722 ve 723) başlama düzeyinde en az beş oturumda kararlı veri elde edildiği görülmektedir “<i>Öğrencilere üç farklı ücret aralığı setinin (5.00-9.99\$, 10.00-14.99\$ ve 15.00-20.00\$) herbirinden rastlantısal 10 ücret sunulmuştur. Öğretmen öğrencinin dikkatini çekmiş ve ardından flash kartta yazan ücreti gösterirken sözlü olarak ücreti belirtmiştir. Öğretmen öğrencinin tepkide bulunması için (ücreti denkleştirmek için bağımsız olarak para sayması) 10 sn. fırsat sunmuş ve daha sonra bağımsız ya da yanlış tepkileri kaydetmiştir.</i>” | H | <ul style="list-style-type: none"> Başlama düzeyi oturumları düzenlenmemiştir. |

Not: Başlama düzeyinde kabul edilebilir veri nokta sayısı beş olmalıdır (Horner ve diğ., 2005). Ancak, OSB olan bireylere yönelik yürütülmüş matematik beceri öğretim çalışmaları incelendiğinde, çok az çalışmada bu ölçüt karşılandığından üç ve üzeri veri sayısı (12. madde için) kabul edilebilir olarak değerlendirilmiştir.

Başlama düzeyi kapsamında “evre uygulamadan önce örüntüye kanıt sağlamıştır” göstergesi Cihak ve Grim’in (2008) yürüttükleri çalışma için “E” olarak kodlanmıştır çünkü söz konusu çalışmada başlama düzeyinde ipucu sunulmamış, öğretim yapılmamış veya herhangi bir dönüt verilmemiştir. Araştırma modeli incelendiğinde başlama düzeyinde kararlı verilerin elde edildiği görülmektedir. Buna karşın, Morrison ve Rosales-Luiz’in (1997) yürüttükleri çalışmada başlama düzeyi oturumları düzenlenmediği için ilgili gösterge “H” olarak kodlanmıştır. Bununla birlikte, Cihak ve Grim’in yürüttükleri çalışmada başlama düzeyinde herhangi bir öğretim sağlanmamakla birlikte neler yapıldığı (ör. öğrenciye sunulan ürünlere ilişkin bilgiler, yanıt aralığı, deneme sayısı) okuyucu tarafından tekrarlanabilir şekilde açıklandığı için başlama düzeyi “kesin sınırları ile yinelenebilir olarak

tanımlanmıştır” göstergesi “E” olarak kodlanmıştır. Buna karşın, Morrison ve Rosales-Luiz tarafından yürütülen çalışmada başlama düzeyi oturumları düzenlenmediği için ilgili gösterge “H” olarak kodlanmıştır.

Geçerlik. Geçerlik ölçütü kapsamında yer alan sekiz niteliksel göstergenin tamamı karşıladı olarak kodlanan (“E” olarak kodlanan) çalışmaya (Waters ve Boon, 2011) ve bu sekiz niteliksel göstergenin ikisi karşıladı olarak kodlanan (“E” olarak kodlanan) altısı ise karşılamadı olarak kodlanan (“H” olarak kodlanan) çalışmaya (Levingston ve diğ., 2009) ilişkin bilgiler Tablo 9’da görülmektedir.

Tablo 9

Geçerlik Açısından Çalışmalarda Sunulan Bilgiler

| Geçerlik | Waters ve Boon, 2011 | | Levingston ve diğerleri, 2009 | |
|---|----------------------|---|-------------------------------|--|
| | E/H | Alıntılar | E/H | Alıntılar |
| 14. Deneysel etkinin üç kanıtı/ gösterimi vardır. | E | <ul style="list-style-type: none"> Deneysel etki üç farklı durumda, tahminde bulunma-doğrulama-yineleme basamakları için, (katılımcılar arasında) gerçekleştirilmiştir (Bk. Şekil 3, Waters ve Boon, 2011, s. 551) | H | <ul style="list-style-type: none"> İşlem öncesi becerilerin ediniminin problem çözme üzerindeki etkisi iki katılımcı üzerinde incelenmiştir (Bk. Şekil 1, Levingston ve diğ., 2009, s. 365). |
| 15. Tasarım iç geçerliğe ilişkin tehditleri kontrol etmiştir. | E | <ul style="list-style-type: none"> “Çalışmada katılımcılar arası çoklu yoklama modeli kullanılmıştır...” Uygulama sürecinde “Nokta Belirleme Tekniği” kullanılarak bir öğretim süreci yürütülmüştür. Grafiksel veriler “Nokta belirleme Tekniği” ile yapılan uygulama sürecine ilişkindir (Bk. Şekil 3, Waters ve Boon, 2011, s. 551). | H | <ul style="list-style-type: none"> “İşlem öncesi becerilerin edinim sırası iki katılımcıda farklılaşmaktadır.” Bu durum problem çözme becerilerini geliştirmede işlem öncesi becerilerinden hangisi ya da hangilerinin etkili olduğuna dair kestirimi güçleştirmektedir (Bk. Şekil 1, Levingston ve diğ., 2009, s. 365). |
| 16. Sonuçlar deneysel kontrolün sağlandığına ilişkin bir örnek/kanıt sunmuştur. | E | <ul style="list-style-type: none"> Grafiksel veriler incelendiğinde her üç katılımcının performansında da deneysel etkiyi gösteren düzey ve eğilim değişikliği görülmektedir (Bk. Şekil 3, Waters ve Boon, 2011, s. 551). | H | <ul style="list-style-type: none"> Deneysel etki yalnızca iki katılımcı için görülmektedir. (Bk. Şekil 1, Levingston ve diğ., 2009, s. 365). İşlem öncesi becerilerden birinin ya da bir kaçının problem çözme becerisi üzerinde etkili olup olmadığını kestirim yapmak mevcut araştırma tasarımıyla güçleştirmektedir. |
| 17. Etki yinelenmiş, dış geçerlik sağlanmıştır. | E | <ul style="list-style-type: none"> Uygulama üç farklı katılımcıyla (Tent, Michael ve Alex) yürütülmüş ve deneysel etki katılımcılar arasında ve katılımcılar içerisinde yinelenerek grafiksel veri üzerinde gösterilmiştir (Bk. Şekil 3, Waters ve Boon, 2011, s. 551). “Öğrencilerin Nokta Belirleme Stratejisini genelleyip genellemediklerini belirlemek için aynı matematik becerileri için farklı uyarılar (farklı sorular) kullanılmıştır.” | H | <ul style="list-style-type: none"> İşlem öncesi becerilerin ediniminin birbirleriyle etkileşimi söz konusudur (Bk. Şekil 1, Levingston ve diğ., 2009, s. 365). Etki yinelenmesi görülmektedir. Ancak bu etki sabit bir değişkenden daha çok birden fazla değişkene aittir. |

Tablo 9 (devamı)

| Geçerlik | Waters ve Boon, 2011 | | Levingston ve diğerleri, 2009 | |
|--|----------------------|--|-------------------------------|--|
| | E/H | Alıntılar | E/H | Alıntılar |
| 18. Bağımlı değişken sosyal olarak önemlidir. | E | <ul style="list-style-type: none"> Katılımcıların tamamında hedeflenen beceri, devam ettikleri okul programlarında yer alan bireyselleştirilmiş eğitim programları (BEP) kapsamında amaçlanmıştır. | E | <ul style="list-style-type: none"> "... Her iki katılımcı da sözel problemleri çözmekte zorlanmaktaydılar ve aileleri ve sınıf öğretmenleri çalışmaya katılmaları için söz konusu katılımcıları önermiştir." "... Üçüncü ve dördüncü sınıf matematik müfredat programından elde edilerek... problemler oluşturuldu." |
| 19. Uygulamaya bağlı olarak bağımlı değişkende meydana gelen değişikliğin büyüklüğü sosyal olarak önemlidir. | E | <ul style="list-style-type: none"> "10 maddelik bir sosyal geçerlik anketi <i>Nokta Belirleme Tekniği</i>'nin matematik öğretiminde sosyal geçerliğini belirlemek üzere öğrenciler, aileleri ve öğretmenlere yönelik kullanılmıştır." "Öğrenciler, öğretmenler ve aileleri <i>Nokta Belirleme Tekniği</i> uygulamasının genellikle yararlı olduğuna dair fikir beyan ettikleri anket üzerinde görülmüştür." Öğretmenler "<i>Nokta Belirleme Tekniği</i>" uygulamasını kullanmayı öğretmenin kolay olduğunu ve öğrencilerin kısa sürede öğrendiklerini ifade etmiştir. | H | <ul style="list-style-type: none"> Sosyal geçerlik verisi yoktur. Bağımlı değişkende meydana gelen değişimin öğrencinin yaşamına ne gibi katkı sunduğuna dair bir açıklama yoktur. |
| 20. Bağımsız değişken uygun maliyetli ve/veya pratiktir. | E | <ul style="list-style-type: none"> Öğrenciler "<i>Nokta Belirleme Tekniği</i>" uygulamasının kullanımının kolay ve eğlenceli olduğunu ve ekranlarına önerebileceklerini ifade etmiştir. | H | <ul style="list-style-type: none"> Bağımsız değişkenin kullanım kolaylığı ya da maliyetine ilişkin çalışmada bir açıklama yoktur. |
| 21. Bağımsız değişken zamanla <u>tipik</u> (doğal) bağlamlarda/kişilerin varlığında uygulanmıştır. | E | <ul style="list-style-type: none"> "<i>Veri toplama, öğretim ve uygulama oturumları aynı (öğrencilerin devam ettikleri) özel eğitim matematik sınıfında yürütülmüştür.</i>" Özel eğitim öğretmeni kendi öğrencilerine kendi sınıflarında 1:1 öğretim sunmuştur. | E | <ul style="list-style-type: none"> "<i>Öğretim oturumları bireysel olarak bir başka öğretmen tarafından (ilk yazar) katılımcının sınıfının arka kısmında bir masada gerçekleştirilmiştir.</i>" |

Geçerlik kategorisi altında "deneysel etkinin üç kanıtı/gösterimi vardır" göstergesi üç farklı katılımcı arasında bağımsız değişkenin etkisi sınıdığı için Waters ve Boon (2011) tarafından yürütülen çalışmada "E" olarak kodlanmıştır. Buna karşın, Levingston ve diğerlerinin (2009) yürüttükleri çalışmada uygulama (işlem öncesi beceriler) iki katılımcıya öğretim içeriğinin sıralanmasında farklılaştırma yapılarak sunulmuş ve problem çözme becerilerindeki etkisi incelenmiştir. Bu nedenle ilgili gösterge bu çalışma için "H" olarak kodlanmıştır. "Tasarım iç geçerliğe ilişkin tehditleri kontrol etmiştir" göstergesi Waters ve Boon'un yürüttükleri çalışmada, araştırma modeli (deneysel etkinin üç farklı durumda gösterimi) ve öğretim süreci incelendiğinde, deneysel süreçte, katılımcıların performansında (ondalıklı para değerlerini çıkarma) meydana gelen değişimin yalnızca

uygulamadan (nokta belirleme tekniği) kaynaklandığı güçlü bir biçimde görüldüğü için ilgili gösterge “E” olarak kodlanmışken, Levingston ve diğerlerinin yürüttükleri çalışmada ise, deneysel etki (işlem öncesi becerilerin edinimi) tahminde bulunma-doğrulama-yineleme basamaklarında güçlü biçimde görülmediği için (ör. üçüncü katılımcıyla) ve işlem öncesi becerilerinin ediniminin birbirini etkileme olasılığı nedeniyle araştırma tasarımının iç geçerliği kontrol edebildiğini söylemek mümkün olmamış ve bu nedenle de ilgili gösterge “H” olarak kodlanmıştır. “Sonuçlar deneysel kontrolün sağlandığına ilişkin bir örnek/kanıt sunmuştur” göstergesi açısından çalışmalar incelendiğinde ise Waters ve Boon’un yürüttükleri çalışmada deneysel etki her üç durumda da (katılımcılar arasında) ilerleme kaydettirdiği ve kontrolün sağlandığına ilişkin güçlü bir kanı oluşturduğu için ilgili gösterge bu çalışma için “E” olarak kodlanırken, Levingston ve diğerlerinin yürüttükleri çalışmada deneysel etki iki katılımcı ile gösterildiğinden ve bu gösterim sürecinde de işlem öncesi becerilerinden tamamının mı yoksa birkaçının mı problem çözme becerisi üzerinde etkili olup olmadığının kestirimini yapmak mevcut araştırma tasarımıyla güçleştiği için bu çalışma için “H” olarak kodlanmıştır. Waters ve Boon’un yürüttükleri çalışmada “etki yinelenmiş, dış geçerlik sağlanmıştır” göstergesi “E” olarak kodlanmıştır. Çünkü katılımcı içerisinde ve/veya katılımcılar arasında yinelenen ölçümler yapılmış ve uyaran genellemesi yapılarak tekrarlı ölçümler devam etmiştir. Ancak, Levingston ve diğerlerinin yürüttükleri çalışmada etki yinelenmesi söz konusu olsa bile bu yinelenmenin sabit bir değişkenden değil birden fazla değişkenden kaynaklı olduğu ve bu nedenle de deneysel etkinin ayırt edilemediği gerekçesiyle “H” olarak kodlanmıştır. Bir diğer gösterge olan “bağımlı değişken sosyal olarak önemlidir” değişkeni ise her iki çalışma için de “E” olarak kodlanmıştır. Çünkü iki çalışmada da, bağımlı değişkenin öğrencilerin BEP’lerinde yer aldığı ve akademik olarak gereksinim duyduğu becerileri öğretmeye yönelik olarak seçildiği ve söz konusu becerilerin aile, öğretmen gibi çocuğu birinci derecede etkileyen kişiler tarafından önerildiği görülmektedir. Waters ve Boon’un yürüttükleri çalışmada, sosyal geçerlik verisinin elde edilip rapor edilmesi ve öğrenciler, öğretmenler ve aileleri becerinin ediniminin (para hesaplama becerileri) öğrencilerin bağımsız yaşam becerilerinin desteklenmesi açısından önemli ve gerekli bir beceri olduğunu belirtmelerinden kaynaklı olarak “uygulamaya bağlı olarak bağımlı değişkende meydana gelen değişikliğin büyüklüğü sosyal olarak önemlidir” göstergesi “E” olarak kodlanmıştır. Buna karşın Levingston ve diğerlerinin yürüttükleri çalışmada, sosyal geçerlik verisinin elde edilmediği ve öğrencinin yaşamına sunduğu katkı betimlenmediği için ilgili gösterge “H” olarak kodlanmıştır. Waters ve Boon’un yürüttükleri çalışmada öğretmen ve öğrenci görüşleri uygulamanın (nokta belirleme tekniği) kolay ve eğlenceli olduğu yönündedir. Bu nedenle “bağımsız değişken uygun maliyetli ve/veya pratiktir” göstergesi “E” olarak kodlanmıştır. Buna karşın, ilgili göstergeyi karşılayacak bir açıklama Levingston ve diğerlerinin yürüttükleri çalışmada olmamasından dolayı “H” olarak kodlanmıştır. Geçerlik kategorisindeki son madde olan “bağımsız değişken zamanla tipik (doğal) bağlamlarda/kişilerin varlığında uygulanmıştır” göstergesi için çalışmaların ikisi de incelendiğinde özel eğitim ya da bir sınıf öğretmenin uygulamacı olması ve sınıf ortamında uygulamayı gerçekleştirmeleri nedeniyle ilgili gösterge iki çalışma için de “E” olarak kodlanmıştır.

Sonuç olarak; bilimsel-dayanaklı uygulamaların belirlenmesinde “*Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri*” yönergesi kapsamında beş grupta yer alan niteliksel göstergelerden tamamını karşılayan çalışmalar olmakla birlikte, niteliksel göstergelerin tamamını, çoğunluğunu veya bir kısmını karşılamayan çalışmalar da bulunmaktadır. Bu durumun niteliksel özellikleri açısından araştırmaları sorgulanabilir kıldığı söylenebilir. Ayrıca, araştırmaların Horner ve diğerleri tarafından önerilen (2005) “*kabul edilebilirlik*” ölçütlerini karşılayıp karşılamamaları açısından değerlendirilmeleri ile bilimsel-dayanaklı olup olmadıklarına yönelik ileri analiz süreçleri kapsamında (ör. meta analizi) ele alınması gerekmektedir. Niteliksel göstergeler kapsamında ele alınan 10 çalışmanın “*kabul edilebilirlik*” ölçütlerine ilişkin yapılan kodlamalar Tablo 10’da yer almaktadır.

Tablo 10
 “Kabul Edilebilirlik” Ölçütlerine İlişkin Çalışmaların Örnek Değerlendirilmesi

| Çalışmalar | “Kabul Edilebilirlik” Ölçütlerine Göre Değerlendirme | | | | | Deneyel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması (Bk. Tablo 4, Madde 17) |
|--------------------------------|--|--|--|--|---|--|
| | Uygulamanın İşlevsel Tanımı (Bk. Tablo 4, Madde 9) | Uygulama güvenilirliği (Bk. Tablo 4, Madde 11) | Bulguların İşlevsel Tanımı (Bk. Tablo 4, Madde 14) | Bulguların İşlevsel Tanımı (Bk. Tablo 4, Madde 15) | Uygulama ile elde edilen bulgular arasında işlevsel ilişki (Bk. Tablo 4, Madde, 16) | |
| Adcock ve Cuvo, 2009 | E | H | E | E | E | E |
| *Akmanoğlu ve Batu, 2004 | E | E | E | E | E | E |
| *Cihak ve Grim, 2008 | E | E | E | E | E | E |
| Collins ve diğerleri, 2011 | E | E | E | E | H | E |
| *Kellems ve diğerleri, 2016 | E | E | E | E | E | E |
| Morrison ve Rosales-Luiz, 1997 | E | H | H | H | H | H |
| Levingston ve diğerleri, 2009 | E | E | H | H | H | E |
| *Root ve diğerleri, 2017 | E | E | E | E | E | E |
| Tiger ve diğerleri, 2007 | E | H | E | H | H | H |
| *Waters ve Boon, 2011 | E | E | E | E | E | E |

*“Kabul edilebilirlik” ölçütü olarak belirlenen göstergelerin tamamında “E” olarak kodlanan çalışmalar.

“Kabul edilebilirlik” ölçütlerine ilişkin değerlendirilen 10 çalışmadan beşinin (Akmanoğlu ve Batu, 2004; Cihak ve Grim, 2008; Kellems ve diğ., 2016; Root ve diğ., 2017 ve Waters ve Boon, 2011) Horner ve diğerlerinin (2005) bir çalışmanın “kabul edilebilir” düzeyde değerlendirilmesi için önerdiği beş özelliği karşıladığı yani bu özellikler kapsamında ele alınan altı maddenin tamamında “E” olarak kodlandığı görülmektedir. 10 çalışmadan beşi ise (Adcock ve Cuvo, 2009; Collins ve diğ., 2011; Levingston ve diğ., 2009; Morrison ve Rosales-Luiz, 1997; Tiger ve diğ., 2011) “kabul edilebilirlik” ölçütlerine ilişkin altı niteliksel göstergeden en az birini karşılamadığı için “kabul edilebilir” düzeyde değerlendirilmemiştir. Sonuç olarak, deneysel sürece etki edebilecek değişkenlerin araştırma sürecinde niteliksel açıdan yeterli olması araştırmanın “kabul edilebilir” düzeyde değerlendirilmesinde etken olduğu görülmektedir. Dolayısıyla bir çalışma niteliksel özellikler açısından yeterli olmasının yanısıra, özellikle, deneysel süreç açısından yeterli niteliksel özelliklere sahip olması gerekmektedir.

Tartışma, Sonuç ve Öneriler

Bu çalışmada, Horner ve diğerleri (2005) tarafından tek-denekli deneysel araştırmaların niteliğinin incelenmesi ve yürütülecek tek-denekli deneysel çalışmalara rehberlik etmesi amacıyla geliştirilen ve özellikle

uluslararası alanyazında sıklıkla kullanılan niteliksel göstergelerin tek-denekli deneysel araştırma çalışmaları bağlamında nasıl kullanılacağı açıklanmıştır. Bu bağlamda, niteliksel göstergeler kullanılarak “*kabul edilebilirlik*” ölçütlerini karşılayan çalışmaların nasıl belirleneceği gelişimsel yetersizliği olan bireylere matematik becerileri öğretimi konusundaki yürütülen bir alanyazın taramasında yer alan çalışmalar kullanılarak anlatılmıştır. Bir niteliksel göstergeyi karşılayan ve karşılamayan çalışmalara ilgili makalelerde sunulan bilgiler ışığında örnekler sunulmuş ve çalışmanın göstergeleri neden karşıladığı ya da karşılamadığı ayrıntılı bir şekilde anlatılmıştır. Tek denekli deneysel araştırmaları rapor eden çalışmaların niteliksel göstergeler bağlamında değerlendirilmesi çalışmaların niteliğinin belirlenmesi için meta analizler kapsamında sıklıkla kullanılan bir uygulamadır. Bu çalışmanın ülkemizde tek-denekli deneysel araştırmaların meta-analizlerini gerçekleştiren araştırmacılara, gerçekleştirdikleri meta-analizler kapsamında yer alan çalışmaları, niteliksel göstergeler bağlamında nasıl değerlendirecekleri konusunda rehberlik edeceği düşünülmektedir. Ayrıca, bu çalışmanın tek denekli deneysel araştırmalar planlayan ve uygulayan araştırmacılara niteliksel göstergeler bağlamında nelere dikkat etmeleri gerektiği konusunda destek sunacağı öngörülmektedir. Aşağıda bu çalışma kapsamında elde edilen bulgular kısaca tartışılmış ve ilgili alanyazın ışığında araştırmacılara bazı öneriler sunulmuştur.

Eğitim alanında bilimsel-dayanaklı uygulamaların kullanılmasına yönelik hareket Amerika Birleşik Devletleri’nde 2001 yılında çıkarılan Hiçbir Çocuk Geride Kalmasın (No Child Left Behind) ve 2004 yılında çıkarılan Özel Gereksinimli Bireylerin Eğitimi Yasası (Individuals with Disabilities Education Act) ile birlikte hız kazanmıştır (Rakap, 2016). Bu bağlamda 2000’li yılların ortalarından itibaren araştırmacıların ve eğitim ile ilgili organizasyonlar ile çeşitli kuruluşların genel ve özel eğitim alanında bilimsel dayanağı olan ve öğrenci çıktıları üzerinde olumlu etkileri olan uygulamaları belirlemeye yönelik çalışmalara hız verdiği görülmektedir. Bu çalışma kapsamında ele alınan ve Horner ve diğerleri (2005) tarafından geliştirilen “*Tek-Denekli Deneysel Araştırmaların Niteliksel Göstergeleri*” de bu çabaların bir sonucu olarak ortaya çıkmıştır. Bu çalışma bağlamında incelenen 10 çalışmadan 6’sının 21 nitelik ölçütünden en az 18’ini karşıladığı görülürken, 2 çalışmanın 14, bir çalışmanın 13 ve 1 çalışmanın da 8 göstergeyi karşıladığı belirlenmiştir. Çalışmaların yayımlandıkları yılın güncelliği ile karşıladıkları niteliksel gösterge sayısı arasında bir pozitif korelasyon ilişkisinin ($r = 0.65$) olması, her geçen yıl yayınlanan çalışmalarda daha fazla niteliksel göstergenin karşılandığına işaret etmektedir. Ancak, tek-denekli deneysel araştırmaların meta-analizi üzerinde çalışmalar yapan araştırmacıların özen göstermesi gereken konulardan birisi, özellikle 2005 yılından önce yayınlanan makalelerin bu ölçütlerden bazılarını karşılamıyor olmasının doğrudan makalede rapor edilen çalışmanın niteliksiz olduğu anlamına gelmediğidir. Çoğu zaman dergilerin sahip olduğu sayfa ya da kelime sayısı sınırı nedeniyle ve tek denekli deneysel araştırmaları rapor eden makalelerde ele alınması gereken özelliklerle ilgili açık bir rehberin olmaması sebebiyle, 2005 yılı öncesinde yayınlanan makalelerde, makalede rapor edilen çalışma yürütülürken niteliksel göstergelerin hepsi ya da çoğu dikkate alınmış olmasına karşın, ölçütleri karşılayacak düzeyde bilgi rapor edilmemiş olması oldukça muhtemeldir.

Bu çalışmanın alanyazına önemli katkılarından biri Horner ve diğerleri (2005) tarafından ele alınan “*kabul edilebilirlik*” ölçütlerini karşılama bağlamında değerlendirilebilecek niteliksel göstergelerin belirlenmesidir. Bu bağlamda bu çalışmada, beş kabul edilebilirlik ölçütünün (uygulamanın işlevsel tanımı, bulguların işlevsel tanımı, uygulama güvenilirliği, uygulama ile elde edilen bulgular arasında işlevsel ilişki ve deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması) değerlendirilmesinde altı niteliksel göstergenin (Bk. Tablo 10) kullanılması önerilmektedir. Bilimsel-dayanaklı uygulamaların belirlenmesine yönelik meta-analizlere bir çalışmanın dâhil edilebilmesi için o çalışmanın Horner ve diğerleri tarafından geliştirilen 21 niteliksel göstergenin tamamını karşılaması beklenirken, göstergelerin tamamını karşılamayan çalışmaların en azından “*kabul edilebilirlik*” ölçütleri bağlamında ele alınan altı niteliksel göstergeyi karşılaması önerilmektedir. Daha öncede açıklandığı gibi, bu çalışma kapsamında incelenen çalışmaların farklı öğretim uygulamalarını kullanmaları ve dolayısıyla tek bir uygulama için yeterli sayıda araştırma olmaması nedeniyle “deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması” ölçütü niteliksel göstergelerden “etki yinelenmiş dış geçerlik sağlanmıştır” göstergesi ile değerlendirilmiştir. Bu doğrultuda bir uygulamanın bilimsel dayanağının olup olmadığının incelendiği meta-analiz

çalışmalarında, bireysel çalışmaların 21 niteliksel göstergesi ve ilk dört “*kabul edilebilirlik*” ölçütü karşılayıp karşılamadığının belirlenmesi ve daha sonra göstergeleri ve ölçütleri karşılayan çalışmaların beşinci ölçüt olan “deneysel kontrolün yeterli sayıda çalışma, araştırmacı ve katılımcıyla gerçekleştirilmiş olması” ölçütünü karşılayıp karşılamadığının belirlenmesi gerekmektedir. Bir uygulama ile ilgili niteliksel ölçütlerin tamamını veya kabul edilebilirlik ölçütlerinin ilk dördünü karşılayan, en az üç farklı bölgede, üç farklı araştırmacı grubuyla yürütülmüş, toplamda en az 20 katılımcının yer aldığı en az beş tek-denekli deneysel araştırma yürütüldüğünde o uygulama bilimsel-dayanaklı uygulamalar bağlamında ele alınabilecektir.

Horner ve diğerleri (2005) tarafından tek denekli deneysel araştırmaların niteliğini değerlendirmek üzere geliştirilen göstergeler tek denekli deneysel araştırmaların yöntemsel niteliklerini incelemeye yönelik geliştirilen ilk ölçütlerdendir. Bu çalışmayı takiben, özel gereksinimli çocukların eğitimi ile ilişkili birçok kurum ve kuruluş (ör. Özel Gereksinimli Çocuklar Konseyi, What Work Clearinghouse, Amerikan Ulusal Otizm Merkezi) tek-denekli deneysel araştırmaların niteliğini belirlemeye yardımcı olmak amacıyla benzer ölçütler geliştirmiştir. Tek-denekli deneysel araştırmalar yürüten ve bu alanda bilimsel-dayanaklı uygulamaların belirlenmesine yönelik meta-analiz çalışmaları gerçekleştiren araştırmacıların bu kurum ve kuruluşlar tarafından geliştirilen ölçütlere aşina olması bir uygulama ile ilgili en güvenilir şekilde elde edilen kanıtların belirlenmesi bağlamında oldukça önemlidir. Dahası, araştırmacıların bir kurum ya da kuruluş tarafından geliştirilen niteliksel ölçütlere göre bilimsel-dayanaklı olduğu belirlenen bir uygulamanın başka bir ölçüt seti ile değerlendirildiğinde farklı sonuçlar ortaya çıkabileceğinin farkında olmaları ve bu nedenle inceledikleri uygulamanın hangi ölçütlere göre bilimsel-dayanaklı olduğunu belirtmeleri gerekmektedir (Cook ve diğ., 2015).

Araştırma ile uygulama arasındaki açıklığın kapatılması ve etkili olduğu birçok deneysel araştırma ile ortaya konan uygulamaların özel gereksinimli çocuklara öğretim yaparken kullanılması ve bu vesileyle özel gereksinimli öğrencilerin kendi gerçek potansiyellerine ulaşmak için eğitim ortamlarında söz konusu etkili uygulamalara maruz kalması oldukça önemlidir (Cook ve diğ., 2015). Bu makale kapsamında ele alınan niteliksel göstergeler ile diğer kurum ve kuruluşlar tarafından geliştirilen niteliksel ölçütlerin temel amacı bilimsel-dayanaklı uygulamaların belirlenmesi yoluyla özel gereksinimli çocukların gelişimsel ve öğrenmeye yönelik çıktılarının geliştirilmesidir. Tek-denekli deneysel araştırmalar bağımlı ve bağımsız değişkenler arasındaki nedensel ya da işlevsel ilişkilerin belirlenmesinde kullanılan güçlü yöntemlerdir ve bu yöntemler kullanılarak yürütülen çalışmaların bilimsel-dayanaklı uygulamaları belirlemeye yönelik sistematik ve meta-analizler kapsamında değerlendirilmesi oldukça önemlidir (Horner ve diğ., 2005). Bu bağlamda tek-denekli deneysel araştırma desenlerini kullanarak çalışmalar yürüten araştırmacıların; (a) tek-denekli deneysel araştırma desenlerinin yöntem ile ilgili anahtar özelliklerine hâkim olmaları, (b) niteliksel göstergeleri ve metodolojik açıdan niteliksel ölçütlerini karşılayan yüksek nitelikli çalışmalar gerçekleştirmeleri, (c) bu çalışmaları niteliksel göstergeleri ve ölçütlerini göz önünde bulundurarak raporlaştırmaları, (d) çalışmanın raporlaştırılması bağlamında sayfa sınırı ile ilgili kaygıların olduğu durumlarda niteliksel göstergelerin karşılanıp karşılanmadığını kısaca belirtmeleri ya da tablolaştırmaları, (e) periyodik aralıklarla sistematik bir şekilde bilimsel-dayanaklı uygulamaları belirlemeye yönelik çalışmalar gerçekleştirmeleri ve (f) farklı kurum ya da kuruluşlar tarafından geliştirilen ölçütleri karşılaştırmaları olarak inceleyerek diğer araştırmacıları ve uygulayıcıları bilgilendirmeleri önerilmektedir.

Bu çalışmanın iki sınırlılığı bulunmaktadır. Birincisi, niteliksel göstergelerin tek-denekli deneysel araştırma çalışmaları bağlamında nasıl kullanılacağı, gelişimsel yetersizliği olan bireylere matematik becerileri eğitimi konusundaki yürütülen bir alanyazın taramasında yer alan çalışmalar kullanılarak anlatılmıştır. Niteliksel göstergelerin nasıl kullanılacağına farklı konulara odaklanan çalışmalar bağlamında açıklanması ve örneklerin çoğaltılması, özellikle genç araştırmacıların tek-denekli deneysel araştırma çalışmalarının niteliğinin değerlendirilmesinde niteliksel göstergeleri doğru bir şekilde kullanmalarını destekleyecektir. İkinci sınırlılık ise, bu çalışma kapsamında örnek olarak kullanılan çalışmalardan hiçbirinin altı niteliksel gösterge (katılımcılar ve ortamlar yeterli tanımlanmıştır; bağımlı değişken işlev olarak kesin sınırları ile tanımlanmıştır; bağımlı değişkenin tekrarlı ölçümleri yapılmıştır; bağımlı değişken için gözlemcilerarası güvenilirlik verisi toplanmıştır; bağımsız değişken kesin sınırları ile yinelenabilir olarak tanımlanmıştır; bağımsız değişken sistematik olarak manipüle edilmiştir) bağlamında “Hayır” olarak puanlanmamış olması ve dolayısıyla ilgili niteliksel göstergeleri karşılamayan çalışmalara örnek verilememiş olmasıdır. Yukarıda belirtilen altı gösterge incelendiğinde bu göstergeleri karşılamayan tek-denekli deneysel araştırmaların yayınlanma olasılığının çok düşük olduğu ya da hiç olmadığı görülmektedir. İleriki çalışmalarda, yayınlanmamış alanyazın incelenerek niteliksel göstergeleri karşılamayan çalışmalara örnekler sunulabilir.

Kaynaklar

*Örnek değerlendirilen çalışmalar

- *Adcock, J., & Cuvo, A. J. (2009). Enhancing learning for children with autism spectrum disorders in regular education by instructional modifications. *Research in Autism Spectrum Disorders*, 3(2), 319-328. doi:10.1016/j.rasd.2008.07.004
- *Akmanoğlu, N., & Batu, S. (2004). Teaching pointing to individuals with autism using simultaneous prompting. *Education and Training in Developmental Disabilities*, 39(4), 326-336.
- Alberto, P. A., & Troutman, A. C. (2009). *Applied behavior analysis for teachers* (8 th ed.). New Jersey: Pearson.
- Aydın, O. (2017). *Otizm spektrum bozukluğu olan bireylere matematik becerilerinin öğretimi: Tek-denkli araştırmalarda betimsel ve meta analiz [Teaching math skills to individuals with autism spectrum disorder: A descriptive and meta analysis in single subject research designs]* (Yayımlanmamış Yüksek Lisans Tezi) [Unpublished master's thesis]. Anadolu Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir [Anadolu University, Educational Science Institute, Eskişehir].
- Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon.
- Brantlinger, E., Jimenez, R., Klinger, J., Pugach, M., & Richardson, V. (2005). Qualitative studies in special education. *Exceptional Children*, 71(2), 195-207.
- Chard, D. J., Ketterlin-Geller, L. R., Baker, S. K., Doabler, C., & Apichatabutra, C. (2009). Repeated reading interventions for students with learning disabilities: Status of the evidence. *Exceptional Children*, 75(3), 263-281.
- *Cihak, D. F., & Grim, J. (2008). Teaching students with autism spectrum disorder and moderate intellectual disabilities to use counting-on strategies to enhance independent purchasing skills. *Research in Autism Spectrum Disorders*, 2(4), 716-727. doi:10.1016/j.rasd.2008.02.006
- *Collins, B. C., Hager, K. L., & Galloway, C. C. (2011). Addition of functional content during core content instruction with students with moderate disabilities. *Education and Training in Autism and Developmental Disabilities*, 46(1), 22-39.
- Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D. W. (2015). CEC's standards for classifying the evidence base of practices in special education. *Remedial and Special Education*, 36(4), 220-234. doi:10.1177/0741932514557271
- Cook, B. G., Tankersley, M., Cook, L., & Landrum, T. J. (2008). Evidence-based practices in special education: Some practical considerations. *Intervention in School and Clinic*, 44(2), 69-75. doi:10.1177/1053451208321452
- Council for Exceptional Children-CEC. (2018). <http://www.cec.sped.org>
- Franklin, R. D., Allison, B. D., & Gorman, B. S. (2014). Introduction. In R. D. Franklin, D. B. Allison, & B. S. Gorman (Eds.), *Design and analysis of single-case research* (pp. 41-65). New York: Psychology Press.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, 71(2), 149-164.
- Graham, S. (2005). Preview. *Exceptional Children*, 71(2), 135.

- Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based in special education. *Exceptional Children, 71*(2), 165-180.
- Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 27-51). Washington, DC: APA.
- Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- *Kellems, R. O., Frandsen, K., Hansen, B., Gabrielsen, T., Clarke, B., Simons, K., & Clements, K. (2016). Teaching multi-step math skills to adults with disabilities via video prompting. *Research in Developmental Disabilities, 58*, 31-44. doi:10.1016/j.ridd.2016.08.013
- Kennedy, C. H. (2005). *Single-case designs for educational research*. Boston, MA: Pearson/Allyn Bacon.
- Kırcaali-İftar, G. (2012). Tek-denekli araştırmaların tarihçesi [History of single case research]. E. Tekin-İftar (Ed.), *Eğitim ve davranış bilimlerinde tek-denekli araştırmalar içinde* (ss. 1-14). Ankara: Türk Psikologlar Derneği Yayını [Ankara: Turkish Psychologists Association Publication].
- Lane, K., Wolery, M., Reichow, R., & Rogers, L. (2007). Describing baseline conditions: Suggestions for study reports. *Journal of Behavioral Education, 16*(3), 224-234.
- *Levingston, H. B., Neef, N. A., & Cihon, T. M. (2009). The effects of teaching precurent behaviors on children's solution of multiplication and division word problems. *Journal of Applied Behavior Analysis, 42*(2), 361-367. doi: 10.1901/jaba.2009.42-361
- Mayton, M. R., Wheeler, J. J., Menendez, A. L., & Zhang, J. (2010). An analysis of evidence-based practices in the education and treatment of learners with autism spectrum disorders. *Education and Training in Autism and Developmental Disabilities, 45*(4), 539-551.
- *Morrison, K., & Rosales-Ruiz, J. (1997). The effect of object preferences on task performance and stereotype in a child with autism. *Research in Developmental Disabilities, 18*(2), 127-137. doi:10.1016/S0891-4222(96)00046-7
- Rakap, S. (2016). Özel eğitimde bilimsel dayanaklı uygulamalar [Evidence based interventions in special education]. V. Aksoy (Ed.), *Özel eğitim içinde* (ss. 181-211). Ankara: Pegem Akademi [Ankara: Pegem Academy].
- Reichow, B. & Volkmar, F. R. (2011). Evidence-based practices in autism: Where we started. In B. Reichow, P. Doehring, D. V. Cicchetti, & F. R. Volkmar (Eds.), *Evidence-based practices and treatments for children with autism* (pp. 3-24). New York: Springer.
- Reichow, B., Volkmar, F. R., & Cicchetti, D. V. (2008). Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disabilities, 38*(7), 1311-1319. doi:10.1007/s10803-007-0517-7
- Riley-Tillman, T. C., & Burns, M. K. (2009). *Evaluating educational interventions: Single-case design for measuring response to intervention*. New York: The Guildford Press.
- Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879-906. doi:10.1037/0022-0663.100.4.879
- *Root, J. R., Browder, D. M., Saunders, A. F., & Lo, Y. Y. (2017). Schema-based instruction with concrete and virtual manipulatives to teach problem solving to students with autism. *Remedial and Special Education, 38*(1), 42-52. doi:10.1177/0741932516643592

- Sheridan, S. M. (2014). Single-case designs and large-n studies: The best of both worlds. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 299-308). Washington, DC: APA.
- Tawney, J. W., & Gast, D. L. (1984). *Single subject research in special education*. Columbus, OH: Charles E. Merrill.
- Tekin-İftar, E. (2012a). Tek-denekli arařtırmalar ve temel kavramlar [Single case research and basic conceptions]. E. Tekin-İftar (Ed.), *Eđitim ve davranıř bilimlerinde tek-denekli arařtırmalar* içinde (ss. 15-39). Ankara: Türk Psikologlar Derneđi Yayını [Ankara: Turkish Psychologists Association Publication].
- Tekin-İftar, E. (2012b). Bilimsel arařtırmalarda geđerlik [Validity in scientific research]. E. Tekin-İftar (Ed.), *Eđitim ve davranıř bilimlerinde tek-denekli arařtırmalar* içinde (ss. 15-39). Ankara: Türk Psikologlar Derneđi Yayını [Ankara: Turkish Psychologists Association Publication].
- *Tiger, J. H., Bouxsein, K. J., & Fisher, W. W. (2007). Treating excessively slow responding of a young man with Asperger Syndrome using differential reinforcement of short response latencies. *Journal of Applied Behavior Analysis*, 40(3), 559-563. doi:10.1901/jaba.2007.40-559
- Thompson, B., Diamond, K. E., McWilliam, R., Snyder, P., & Snyder, S. W. (2005). Evaluating the quality of evidence from correlational research for evidence-based practice. *Exceptional Children*, 71(2), 181-194.
- *Waters, H. E., & Boon, R. T. (2011). Teaching money computation skills to high school students with mild intellectual disabilities via the TouchMath program: A multi-sensory approach. *Education and Training in Autism and Developmental Disabilities*, 46(4), 544-555.
- What Works Clearinghouse. (2014). *Procedures and standards handbook* (Version 3.0). Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf



Ankara University Faculty of Educational Sciences Journal of Special Education

Year: 2019, Volume: 20, No: 3, Page No: 597-628

DOI: 10.21565/ozelegitimdergisi.421952

REVIEW

Received Date: 08.05.18

Accepted Date: 13.03.19

OnlineFirst: 23.03.19

Explanation of “Quality Indicators of Single-Subject Experimental Research Studies” Rubric Using Studies Focused on Teaching Mathematics Skills *

Orhan Aydın **
Anadolu University

Elif Tekin-Iftar ***
Anadolu University

Salih Rakap ****
Ondokuz Mayıs University

Abstract

Evidence-based practices have become a subject of further consideration among experts, researchers and practitioners in the field of special education in recent years. In this study, we described (a) how quality indicators developed by Horner and colleagues (2005) for single-subject experimental research designs could be used to determine quality of single-subject experimental research studies using studies focused on teaching mathematical skills to individuals with developmental disabilities and (b) how these indicators could be used to determine whether a study meets acceptability standards developed by Horner et al. For this purpose, we provided examples of studies that met and failed to meet the quality indicators and explained in detail why the study met or did not meet certain criteria. We expect that this study will guide researchers and practitioners who conduct or evaluate studies and meta-analyses of studies employing single-subject experimental research designs in using quality indicators.

Key Words: Single-subject experimental research, evidence-based practices, quality indicators, acceptability standards, functional relationship.

Recommended Citation

Aydın, O., Tekin-Iftar, E., & Rakap, S. (2019). Explanation of “quality indicators of single-subject experimental research studies” rubric using studies focused on teaching mathematics skills. *Ankara University Faculty of Educational Sciences Journal of Special Education* 20(3), 597-628. doi: 10.21565/ozelegitimdergisi.421952

*The study was produced from the master thesis of the first author, which was completed at Anadolu University Educational Sciences Institute, Special Education Department.

****Corresponding Author:** Res. Assist., E-mail: o_aydin@anadolu.edu.tr, <https://orcid.org/0000-0002-9695-2414>

***Prof., E-mail: eltekin@anadolu.edu.tr, <https://orcid.org/0000-0001-5512-616X>

****Asst. Prof., E-mail: salih.rakap@omu.edu.tr, <https://orcid.org/0000-0001-7853-3825>

Identification and implementation of evidence-based practices in classrooms have been among the important topics in which educational researchers, professional organizations, and institutions have paid special attention over the last 20 years. As a result of the research for determining whether a practice is evidence-based, researchers and professional associations have initiated various initiatives to determine what should be the research criteria that can provide a scientific basis in scientific research. This work has resulted in the development of quality indicators for different research designs. Although a consensus has yet to be reached on the concept of evidence-based practices, various guidelines (rubrics) and reports have been published in the field of special education as a product of these efforts. For example, the Council for Exceptional Children formed a Task Force to study what might be the quality indicators that could be used to determine the scientific basis of practice in different research designs. As a result of these efforts, quality indicators that a study has to meet in order to provide evidence for a practice have been developed for group or quasi-experimental researcher designs (Gersten et al., 2005), single-subject experimental research designs (Horner et al., 2005), qualitative (Brantlinger, Jimenez, Klinger, Pugach and Richardson, 2005) and correlational (Thompson, Diamond, McWilliam, Snyder and Synder, 2005) researcher designs. In addition, with a study titled "Standards for Evidence-Based Practices in Special Education" the Council for Exceptional Children developed quality indicators and design standards for both group experimental studies and single-subject experimental research in 2014. In addition to these guidelines developed and published in the field of general special education, specific guidelines and criteria have been developed to determine evidence-based practices in the area of autism spectrum disorder (Reichow, Volkmar and Cicchetti, 2008) or general education (Rakap, 2016). Although these guidelines and criteria have been developed to guide the field of special education in determining evidence-based practices, correct and consistent understanding, evaluation and implementation of these guidelines and criteria can be problematic for researchers and practitioners. Using the literature on teaching mathematical skills to individuals with developmental disabilities as an example, the purpose of the present study, therefore, was to examine how scientific research studies could address the "Quality Indicators of Single-Subject Experimental Research Studies" developed by Horner and colleagues (2005), how studies that meet "*acceptability*" standards using these indicators are determined, and what criteria can be taken into account in the process of being considered as an evidence-based practice. In the following section, we first introduced "Quality Indicators of Single-Subject Experimental Research Studies" and then explained how each quality indicator could be addressed in a study. Finally, using the literature on teaching mathematical skills to individuals with developmental disabilities, we provided examples of studies meeting or not meeting quality indicators and explained why a particular study met or did not meet each quality indicator.

Quality Indicators of Single-Subject Experimental Research Studies to Determine Evidence-Based Practices and Interpretation of the Indicators

Exceptional Children journal, published by the Council for Exceptional Children and ranked first among special education journals, prepared a special issue in 2005, consisting of a group of studies focused on a single topic. In this special issue, the quality indicators for different research designs used in the area of special education and the process of developing these indicators were shared with the readers. The work in this issue has been prepared by the Division of Research of Council for Exceptional Children and Sam Odom. According to Steve Graham (2005), the editor of Exceptional Children, there were three main objectives in the preparation of this special issue: (a) establishing criteria for the identification of evidence-based practices, sharing information on evidence-based practices, and promoting the use of evidence-based practices in the field; (b) showing the importance given and special attention paid to basic and applied research in special education over the years more clearly; and (c) presenting quality indications for different research designs used in special education to guide reviewers in evaluating manuscripts submitted to the journal. One of the studies published in this special issue was prepared by Horner and colleagues and focused on quality indicators for studies using single-subject experimental research designs. Horner et al. (2005) developed 21 quality indicators for single-subject experimental research under five area: (a) participants and environments, (b) dependent variable, (c) independent variable, (d) baseline phase, and (e) validity (internal validity, external validity, and social validity). Table 1 presents a summary of 21 quality indicators and how Horner and colleagues (2005) suggested that these indicators be assessed.

In order for a study to be considered as “*acceptable*” based on the quality indicators for single-subject experimental research developed by Horner and colleagues (2005), the study has to present sufficient information for the following five standards: a) functional description of the implementation, (b) functional description of findings, (c) implementation fidelity, (d) functional relationship between implementation and findings, and (e) sufficient number of studies showing experimental control with sufficient number of participants implemented by sufficient number of researchers/research groups. There is no consensus in the literature on which of 21 quality indicators should be taken into account to determine whether these five standards are present in a study using a single-subject experimental research design. In the present study, a practical evaluation process was developed (and presented at Table 2) for this purpose by examining studies focused on teaching mathematical skills to individuals with developmental disabilities and considering suggestions provided by Horner and colleagues (2005). In this respect, six quality indicators (among 21 indicators) that can be used to determine whether five acceptability standards are present in a study are identified (Aydin, 2017). For the first four acceptability standards, associated quality indicators were identified based on the suggestions of Horner and colleagues. Horner et al. (2005) stated that for the last acceptability standard of “*experimental control being demonstrated across a sufficient range of studies, researchers, and participants,*” it is necessary that “*a minimum of five single-subject studies that meet minimally acceptable methodological criteria and document experimental control have been published in peer-reviewed journal, the studies are conducted by at least three different researchers across at least three different geographical locations, and studies include at least 20 participants.*” Within the context of the present study, we were not able to evaluate this acceptability standard as sample articles examined did not focus on a single practice to teach mathematical skills to individuals with developmental disabilities. For this reason, for a study to meet this acceptability standard, we decided that it had to meet the indicator associated with external validity, “*experimental effects are replicated across participants, settings, or materials to establish external validity.*” However, when there is a sufficient number of studies focused on a single practice, this standard should be used as described by Horner and colleagues (2005).

As shown in Table 2, for a study to meet acceptability standard of “*practice is operationally defined,*” it has to meet the quality indicator 9, *independent variable is described with replicable precision*; to meet the acceptability standard of “*the context and outcomes associated with a practice are clearly defined,*” the study has to meet the quality indicator 14 (i.e., *the design provides at least three demonstrations of experimental effect at three different points in time*) and indicator 15 (i.e., *the design controls for common threats to internal validity*); to meet the acceptability standard of “*the practice is implemented with documented fidelity,*” the study has to meet the quality indicator 11, *overt measurement of the fidelity of implementation for the independent variable is highly desirable*; to meet the acceptability standard of “*the practice is functionally related to change in valued outcomes,*” the study has to meet indicator 16, *the results document a pattern that demonstrates experimental control*; and to meet the acceptability standard of “*experimental control is demonstrated across a sufficient range of studies, researchers, and participants to allow confidence in the effect,*” the study has to meet the quality indicator 17, *experimental effects are replicated across participants, settings, or materials to establish external validity*. In this context, it is suggested that 6 of the 21 quality indicators (9th, 11th, 14th, 15th, 16th, and 17th indicators) should be coded as “Yes” for a study to be considered as “acceptable.”

Prior to examination and evaluation of the studies focused on teaching math skills, a coding form (see Table 4) containing all 21 quality indicators under 5 categories (see Table 4) and explanations of how to score each indicator were developed (see Table 1). Following the development of the form, studies focused on teaching mathematical skills to individuals with developmental delay (see Table 3 for more information about these studies) were examined, evaluated, and coded in view of the quality indicators.

During coding, after identifying information about a study is coded on the form, the letter “Y” for “Yes” is used when a study meets the corresponding quality indicator and the letter “N” for “No” is used when a study does not meet the respective quality indicator. Table 4 shows quality indicators each study met or did not meet. When researchers evaluate studies, they must ensure that these codes are encoded independently by at least two encoders

and that they perform inter-coder reliability analysis between coders. When a disagreement occurs between coders, they must come to an agreement by discussing why they coded the respective indicator in the way they coded.

In the following section, based on the coding provided in Table 4, we present examples for studies meeting and not meeting 21 quality indicators under five categories and explain why a certain study met or did not meet the quality indicators.

Participants and Settings. Akmanoglu and Batu (2004) study met all three quality indicators under *participants and settings* category, while Kellems et al. (2016) met two of the three quality indicators under this category. Table 5 shows information presented in each article that was used when evaluating whether the study met the respective quality indicators. In both studies, "*participants are described with sufficient detail*" indicator was coded as Yes. Because in both studies, the number of participants, the age, gender and diagnosis of the participants were described. Moreover, both studies provided information concerning the performance level of each participant. As Akmanoglu and Batu (2004) provided detailed information about prerequisite skills each participant had to demonstrate to participate in the study and aimed to teach new skills to participants during the study, the quality indicator of "*the process for selecting participants is described with replicable precision*" was coded as Yes. On the other hand, although Kellems et al. (2016) described the characteristics of participants, they did not provide information about prerequisite skills needed to participate in the study, and how participants were identified and selected. Therefore, the quality indicator of "*the process for selecting participants is described with replicable precision*" was coded as No for this study. In terms of the quality indicator of "*critical features of the physical setting are described with sufficient precision to allow replication*," Akmanoğlu and Batu (2004) provided detailed information about the settings in which the study was conducted for each participant. The furniture (e.g., table, chair, etc.) and equipment (e.g., camcorder) present in the setting was described. Moreover, individuals present in the setting while the study was being conducted and participant-practitioner positioning during the implementation was described. For this reason, this indicator was coded as Yes for Akmanoglu and Batu (2004). On the other hand, although Kellems et al. (2016) provided information about the settings in which the study was conducted, they did not provide information about the school in which the study took place, nor had they described characteristics of the study settings (e.g., materials, furniture). Moreover, the authors did not describe participant-practitioner positioning during the implementation. For this reason, this indicator was coded as "No" for Kellems et al. (2016).

Dependent Variable. Collins, Hager, and Galloway (2011) met all five quality indicators under the *dependent variable* category, while Adcock and Cuvo (2009) met three of the five quality indicators under this category. Table 6 shows information presented in each article that was used when evaluating whether the study met the respective quality indicators. Both Collins et al. (2011) and Adcock and Cuvo (2009) met the quality indicator of "*dependent variables are described with operational precision*," as the dependent variable in both studies were explained with concrete and quantitative statements. Collins et al. (2011) created a skill analysis of the dependent variable to be performed and determined acquisition criteria of 100% accuracy on each step of the skill for three sessions. Thus, the indicator of "*each dependent variable is measured with a procedure that generates a quantifiable index*" was coded as Yes for this study. However, in the study conducted by Adcock and Cuvo (2009), the lack of a skill analysis of the dependent variable as well as the lack of information about how correct and incorrect responding would be evaluated resulted in this indicator being coded as No for this study. In Collins et al. (2011), the quality indicator of "*measurement of the dependent variable is valid and described with replicable precision*" is coded Yes, because the authors described how the instructor would evaluate whether participant demonstrated each step of the skill analysis during baseline, implementation, and probe sessions. However, in Adcock and Cuvo (2009), the definitions of correct and incorrect responses were not comprehensive and detailed. Therefore, the measurement of dependent variable was not described in a replicable fashion. For this reason, this indicator was coded as No for Adcock and Cuvo (2009). Because both studies involved repeated measurement of the dependent variable across the study phases and reported interobserver agreement (IOA) data, the quality indicators of "*dependent variables are measured repeatedly over time*" and "*data are collected on the*

reliability or IOA associated with each dependent variable, and IOA levels meet minimal standards” were coded Yes for both studies.

Independent Variable. Root, Browder, Saunders and Lo (2017) met all three quality indicators under the *independent variable* category, while Tiger et al. (2007) met two of the three quality indicators under this category. Table 7 shows information presented in each article that was used when evaluating whether the study met the respective quality indicators. The quality indicator of *“independent variable is described with replicable precision”* was coded Yes for both studies because the authors described implementation process in detail in a replicable fashion. In addition, the quality indicator of *“independent variable is systematically manipulated and under the control of the experimenter”* was coded Yes for both studies as in each study independent variable was systematically manipulated by the researchers based on the research design used. Root et al. (2017) developed a checklist (11 items) for measuring fidelity of implementation of the independent variable. An independent coder observed implementation of the independent variable and used the checklist to determine implementation fidelity. For this reason, the quality indicator of *“overt measurement of the fidelity of implementation for the independent variable is highly desirable”* was coded Yes for this study. In contrast, Tiger et al. (2007) did not obtain data on the implementation fidelity of the independent variable; as a result, the relevant indicator was coded as No.

Baseline. Cihak and Grim (2008) study met both quality indicators under *baseline* category, while Morrison and Rosales-Luiz (1997) did not meet any of the two quality indicators under this category. Table 8 shows information presented in each article that was used when evaluating whether the study met the respective quality indicators. The quality indicators of *“the majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur”* and *“baseline conditions are described with replicable precision”* were coded Yes for Cihak and Grim (2008) as the researchers collected sufficient and stable data to predict future performance of the participants, no intervention was provided during the baseline phase, and baseline phase was described in sufficient detail to allow replication of the procedures used. On the other hand, because Morrison and Rosales-Luiz (1997) did not collect baseline data, two indicators under this category were coded as No for their study.

Validity. Waters and Boon (2011) met all eight quality indicators under *validity* category, while Levingston, Neef and Cihon (2009) met two of the eight quality indicators under this category. Table 9 shows information presented in each article that was used when evaluating whether the study met the respective quality indicators. The quality indicator of *“the design provides at least three demonstrations of experimental effect at three different points in time”* was coded as Yes for Waters and Boon study as the research design employed in the study allowed three demonstrations of experimental effect across three participants or behaviors. However, this indicator was coded No for Levingston et al. study because there were only two participants in the study and only one behavior (problem solution) was studied with each participant. Moreover, the next two quality indicators of *“the design controls for common threats to internal validity”* and *“the results document a pattern that demonstrates experimental control”* were also coded Yes for Waters and Boon’s study as the study used research design that permitted elimination of rival hypothesis and the researcher were able to show that change in the dependent variables occurred only when the intervention were implemented. However, these two indicators were coded No for Levingston et al. study as experimental effect was not observed depending on implementation of the intervention. In addition, quality indicators of *“experimental effects are replicated across participants, settings, or materials to establish external validity”* was coded Yes for Waters and Boon (2011) as experimental effect was observed three conditions (across three participants/behaviors) and coded No for Levingston et al. study as the experimental effect was replicated with only two participants. *“The dependent variable is socially important”* were coded as Yes for both studies because behaviors identified as targets for the interventions were selected from participants’ IEPs and were considered to be important in terms of participants’ academic development. Moreover, parents and teachers of study participants stated that their children needed to learn behaviors selected as target for the interventions implemented. The quality indicator of *“the magnitude of change in the dependent variable*

resulting from the intervention is socially important" was coded Yes for Waters and Boon (2011) study as the researchers reported social validity data collected from students, teachers, and parents of participating students and No for Levingston et al. (2009) as social validity data were not reported. Similarly, the quality indicator of *"implementation of the independent variable is practical and cost-effective"* was coded Yes for Waters and Boon (2011) study and No for Levingston et al. (2009) because the former study reported that students and teachers found the intervention easy to implement and practical and the latter did not report any information about the practicality and cost-effectiveness of the intervention implemented. The last quality indicator, *"social validity is enhanced by the implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts"* was coded Yes for both studies because the studies were implemented within the classroom environment by a special education teacher or a classroom teacher.

As described above, the literature examined in this study did not include any studies that met all 21 quality indicators listed under five categories. One can claim that this finding makes this body of literature questionable in terms of the quality characteristics. In addition to evaluating studies based on these 21 quality indicators, researchers conducting meta-analyses should evaluate studies based on the acceptability standards offered by Horner et al. (2005). Table 10 shows the coding of the acceptability standards for the 10 studies examined in this study. Across 10 studies, 5 met the acceptability standards (Akmanoğlu and Batu, 2004; Cihak and Grim, 2008; Kellems et al., 2016; Root et al., 2017; Waters and Boon, 2011). The rest of the studies did not meet at least one of the acceptability standards.

Discussion, Results and Implications

In this study, we described how quality indicators developed by Horner and colleagues (2005) for single-subject experimental research designs could be used to determine quality of single-subject experimental research studies using studies focused on teaching mathematical skills to individuals with developmental disabilities and how these indicators could be used to determine whether a study meets acceptability standards developed by Horner et al. (2005). For this purpose, we provided examples of studies that met and failed to meet the quality indicators and explained in detail why the study met or did not meet certain criteria. We expect that this study will guide researchers who are conducting studies or meta-analyses of studies employing single-subject experimental research designs. In the following section, we briefly discussed main findings of the present study and provided suggestions in the light of the extant literature.

The movement towards the use of evidence-based practices in education has gained momentum in the United States with the passage of No Child Left Behind in 2001 and the Individuals with Disabilities Education Act of 2004 (Rakap, 2016). In this context, beginning in the mid-2000s, researchers and various organizations related to education or special education, started to work on determining evidence-based practices that have positive impacts on student outcomes. Quality indicators (Horner et al., 2005) investigated in the present study were developed as a result of these efforts. Of the 10 studies examined in this study, 6 were found to meet at least 18 of 21 quality indicators, while 1 study each met 8, 13, 14 and 16 quality indicators. A positive correlation between the publication years of the studies and the number of indicators they meet ($r = 0.6548$) suggests that studies published more recently meet more quality indicators. However, one of the issues that should be considered by researchers working on the meta-analysis of single-subject experimental research studies is that although these criteria have been developed based on the literature focused on single-subject experimental research, the fact that the articles published before 2005 failing to meet some of these criteria does not indicate that the work reported on the articles are low-quality. Due to page or word limitations journals set and unclear guidelines about what features of single-subject experimental research studies should be addressed in the articles, it is possible that although researchers of the articles published prior 2005 considered all quality feature when designing and implementing their studies, they did not report all of them when publishing the results of their studies.

One of the important contributors of the present study to the literature is the identification of quality indicators that can be used to evaluate acceptability standards developed by Horner and colleagues (2005). In this

context, we proposed to use 6 of the 21 quality indicators to evaluate five acceptability standards. For a study employing a single-subject experimental research design to be included in a meta-analysis to determine evidence-base for a practice, the study is expected to meet all 21 quality indicators developed by Horner and colleagues. Studies that do not meet all 21 quality indicators should at least meet the 6 indicators to be considered as acceptable. When there are at least five single-subject experimental research studies investigating a particular practice that meet all quality indicators and the first four acceptability standard, and the studies are conducted in at least three different geographical regions by three different researchers or research groups and included at least 20 participants, the practice can be considered as evidence-based.

Indicators developed by Horner et al. (2005) to assess the quality of single-subject experimental research studies are among the first criteria developed to examine the methodological quality of single-subject experimental research. Following this work, many institutions and organizations associated with the education of children with special needs (e.g., the Council for Exceptional Children, the What Works Clearinghouse, and National Autism Center) have developed similar criteria to help determine the quality of single-subject experimental research studies. Thus, it is very important for researchers and organizations conducting meta-analysis of single-subject experimental research studies to become familiar with these different criteria or standards in order to determine the most reliable evidence for a practice. Furthermore, researchers should be aware of the fact that a practice judged to be evidence-based by a set of quality indicators developed by an institution or organization could be found as non-evidence-based by another set of criteria developed by a different organization (Cook et al., 2015).

Closing the gap between research and practice and using evidence-based practices in the classrooms are particularly important for children with special needs who need to be exposed to effective practices to reach their potential (Cook et al., 2015). The purpose of the quality indicators explained in this study and others is to support development and learning of children with special needs by identifying evidence-based practices. Single-subject experimental research designs are powerful methods used to determine causal or functional relationships between dependent and independent variables, and it is important to include studies employing single-subject experimental research designs in the meta-analyses conducted to determine evidence-based practices (Horner et al., 2005). In this context, researchers conducting studies using single-subject experimental research designs should (a) know key features of single-subject experimental research designs, (b) implement high-quality studies that meet qualitative indicators and methodological quality criteria, (c) consider quality indicators when reporting their findings, (d) systematically conduct meta-analyses to determine evidence-based practices, and (e) inform researchers and practitioners about the quality indicators by comparing and contrasting criteria developed by different organizations.