



Otizm Spektrum Bozukluğunda Kullanılan Ölçeklerde Ebeveyn ve Öğretmenler Arasındaki Uyumun Farklı Tekniklere Göre Karşılaştırılması

Seher YALÇIN*, Hatice BAKKALOĞLU**, Meral Çilem ÖKCÜN AKÇAMUŞ***, Seyda DEMİR****, Zeynep BAHAP KUDRET*****

• Geliş Tarihi: 31.05.2020 • Kabul Tarihi: 04.01.2021 • Çevrimiçi Yayın Tarihi: 04.01.2021

Öz

Bu çalışmada, Otizm Spektrum Bozukluğu (OSB) olan çocuklarda bozukluğun derecesini ölçmede, bozukluğun göstergelerinden biri olan tekrarlayıcı davranışların ve sosyal iletişim-etkileşim davranışlarının değerlendirilmesinde alanyazında sıklıkla kullanılan dört ölçek için aynı çocukların hem ebeveynleri hem öğretmenleri tarafından yapılan puanlamalar arasındaki uyumu farklı tekniklere göre karşılaştırmak amaçlanmıştır. İlişkisel tarama modelinde olan bu çalışmanın katılımcıları, 2018-2019 eğitim-öğretim döneminde, Ankara ilinde Milli Eğitim Bakanlığına bağlı okullara ve rehabilitasyon merkezlerine devam eden, 3-23 yaş aralığındaki 300 ile 380 OSB tanılı bireyden ve bunların 292 ile 379 aralığındaki ebeveyni ile 298 ile 378 aralığındaki öğretmeninden oluşmaktadır. Araştırmada OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren dört farklı ölçek kullanılmıştır. Araştırmanın sonucunda farklı ölçekleri puanlayan ebeveyn ve öğretmenler için ölçeklere verilen puanlamaların kendi içinde iç tutarlılığının yüksek olduğu görülmüştür. Ayrıca bazı alt ölçeklerde, ebeveyn ve öğretmenlerin verdikleri puanlara ilişkin güvenilirlik katsayıları arasında öğretmenler lehine anlamlı farklılık bulunmuştur. Ebeveyn ve öğretmenlerin verdikleri puanlar arasındaki korelasyonların alt ölçeklere göre değiştiği (düşük, orta, yüksek düzeyde anlamlı ilişkiler ve anlamlı olmayan ilişki) tespit edilmiştir. Araştırmada elde edilen bulgular ışığında, sadece korelasyonlara dayalı olarak ebeveyn ve öğretmenler arasındaki uyum hakkında yorum yapmanın uygun olmadığı da görülmüştür.

Anahtar sözcükler: otizm spektrum bozukluğunda kullanılan ölçekler, ebeveyn ve öğretmenler arasındaki uyum

Atıf:

Yalçın, S., Bakkaloğlu, H., Akçamuş, M.Ç.Ö, Demir, Ş. ve Kudret, Z.B. (2021). Otizm spektrum bozukluğunda kullanılan ölçeklerde ebeveyn ve öğretmenler arasındaki uyumun farklı tekniklere göre karşılaştırılması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 52, 360-388 doi:10.9779.pauefd.745749.

* Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, ORCID ID: 0000-0003-0177-6727 yalcins@ankara.edu.tr

** Prof. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, ORCID ID: 0000-0002-3226-9077, hbakkaloglu@ankara.edu.tr

*** Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, ORCID ID: 0000-0003-3297-9711, meralcilem@gmail.com

**** Dr. Öğr. Üyesi, Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, ORCID ID: 0000-0002-7344-7853, demir.seyda@yahoo.com

***** Dr. Öğr. Üyesi, Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, ORCID ID: 0000-0001-5192-0892, zeynepbahap@gmail.com

Giriş

Eđitim-öđretim sürecinde, ölçme ve deđerlendirmenin en önemli amaçlarından biri bireylerin ele alınan özelliđe ne derece sahip olduđunun belirlenmesi ve bu belirlemeye göre bireyler hakkında verilen kararların isabetli olmasının sađlanmasıdır. Bu belirleme sıklıkla bir ölçme aracı kullanılarak yapılmaktadır. Bu nedenle kullanılan ölçme araçlarından elde edilen puanların geçerliđi ve güvenilirliđi önem taşımaktadır. Eđitimde ve psikolojide, ele alınan özellikler doğrudan ölçülemediđinden her ölçme işlemine bir miktar hata karışır. Bireyin bir ölçme işleminden elde edilen puanlarının teorik olarak gerçek puan ve hata puanından oluşması klasik test kuramının (KTK) temel varsayımıdır (Lord ve Novick, 1968). Yani hata, ölçülen özelliđin gerçek deđeri ile ölçme sonucunda elde edilen gözlenen deđeri arasındaki farktır (Atılđan, Kan ve Dođan, 2011). Hatalar; ölçülen bireyler, ölçmeyi yapan kişiler, ölçme aracı ve ölçmenin yapıldıđı ortamdan kaynaklanabilmektedir (Baykul 2000; Crocker ve Algina 1986).

KTK'da, güvenilirliđin hesaplanmasında farklı yaklaşımlar vardır. Bunlar, araştırmacının güvenilirliđi hangi anlamda ele aldıđına ve hata kaynađına göre deđişmektedir (Baykul, 2000). Bir testin belli zaman aralıđıyla iki kez uygulanmasıyla elde edilen güvenilirlik için test-tekrar-test yöntemi kullanılmakta, elde edilen güvenilirlik 'kararlılık' veya 'tutarlılık' anlamında deđerlendirilmektedir. Bu güvenilirlik kestiriminde iki uygulama arasında geçen sürenin, hata kaynađı olduđu düşünölmektedir. Ölçme aracında yer alan maddelerin ölçediđin tümünden elde edilen puanlarla olan tutarlılıđı, 'iç tutarlılık' güvenilirliđini ifade eder ve bu güvenilirlikte ölçme aracını oluşturan maddeler hata kaynađı olarak ele alınmaktadır. İki paralel formun kısa bir zaman aralıđıyla uygulanması sonucu elde edilen puanların 'tutarlılıđı'; eşdeđer (paralel) formlar yöntemi olarak adlandırılır ve bu yöntemde ölçme aracı hata kaynađı olarak görölmektedir. 'Puanlayıcılar arası güvenilirlik' ise, aynı özelliđe birden fazla puanlayıcının verdikleri puanlar arasındaki tutarlılıđa denilmekte ve bu güvenilirlik belirleme yönteminde puanlayıcılar, hata kaynađı olarak düşünölmektedir (Goodwin, 2001; Güler, 2008).

Puanlayıcılar arası güvenilirliđin olduđu önemli olduđu alanlardan biri kuşkusuz özel eđitimidir. Özel gereksinimli bireylerin farklı gelişimsel ve davranışsal özelliklerini deđerlendirmede sıklıkla kullanılan yöntemlerden biri (Murray, Ruble, Wills ve Molloy, 2009), aile ve öđretmen bildirimine dayalı (parents and teacher report scale) olan ölçeklerdir. Bu ölçekler, deđerlendirilmesi hedeflenen bireyle sıklıkla aynı ortamları paylaşan bireyler

tarafından doldurulan ölçekler ve kontrol listeleridir. Öğretmen ve aileler tarafından doldurulan, geçerlik ve güvenilirliği belirlenmiş ölçeklerle ve kontrol listeleriyle özel gereksinimli bireyin, farklı ortamlarda ortaya çıkan ya da çıkmayan davranışlarına ilişkin bilgi alınmaktadır. Bu ölçekler kullanılarak bireyler hakkında tarama, tanılama, müdahale süreçlerinde sıklıkla farklı puanlayıcılardan elde edilen bilgilere dayalı kararlar verilmektedir. Dolayısıyla bireyler hakkında verilen kararların geçerliği ve güvenilirliği, puanlayıcılara bağlı olmaktadır (Dinnebeil vd., 2013; Verhulst ve Akkerhuis, 1989). Sosyal iletişim ve etkileşim becerilerinde sınırlılıklar, tekrarlayıcı-sınırlayıcı davranışlar ve sınırlı ilgi alanlarıyla karakterize olan ve gelişimin erken dönemlerinde ortaya çıkan nöro-gelişimsel bir bozukluk olan otizm spektrum bozukluğu (OSB) olan çocuklar ve bireyler (APA, 2013), ebeveyn ve öğretmen bildirimine dayalı ölçeklerin sıkça kullanıldığı gruplardan biridir. Alanyazında, OSB olan çocuklar ve bireyler de dahil, özel gereksinimli çocuklara ve bireylere yönelik, tarama, tanılama ve müdahalelerin planlanması amacıyla yapılan değerlendirmelerde, çoklu kaynaktan veri toplanması (Raver, 2009; Vickerstaff, Heriot, Wong, Lopes, ve Dossetor, 2007; Voelker, Shore, Lee, ve Szuskiewicz, 2000), çoklu ölçme aracı kullanılması (Raver, 2009; Voelker ve diğ., 2000) ve çoklu değerlendirme yöntemlerinden yararlanılması önerilmektedir (Raver, 2009). Çocuğu ya da bireyi iyi tanıyan kişiler (ebeveynler ve öğretmenler gibi) arasındaki uyum; müdahale ekibi tarafından müdahalenin seçimi, ev ve okul arasındaki uygulamaların tutarlılığı, ailenin müdahaleden memnuniyeti gibi pek çok faktörü etkilemektedir (Dinnebeil vd., 2013; Gresham, Elliott, Cook, Vance ve Kettler, 2010; Tucker ve Schwartz, 2013). Bu nedenle, farklı veri toplama kaynaklarından elde edilen bilgiler arasında tutarsızlık olmasının yaratacağı sonuçlar göz önüne alındığında, çocuğu iyi tanıyan kişiler arasındaki uyum kritik bir çalışma alanı olarak görülmektedir (Kaat, Gadow ve Lecavalier, 2013).

Alanyazında aynı çocuk ya da birey hakkında bilgi sağlayan kişilerin uyum düzeylerine ilişkin tutarsız bulgular vardır. Bu araştırmalardan bazılarında ebeveynler ve öğretmenlerin doldurdukları ölçekler arasında güçlü bir tutarlılık olduğu bulunurken (ör. Constantino, Przybeck, Friesen ve Todd 2000; Constantino vd., 2003); bazılarında aynı çocuklar için ebeveyn ve öğretmenler tarafından doldurulan ölçek puanları arasında düşük düzeyde bir korelasyon olduğu (Murray vd., 2009) ya da anlamlı bir korelasyon olmadığı (Vickerstaff vd, 2007) bulunmuştur. Örneğin Constantino ve diğerleri (2000, 2003), Sosyal Yanıtlayıcılık Ölçeğini (Social Responsiveness Scale) kullandıkları araştırmalarında OSB olan çocukların sosyal davranışları konusunda anne-babalar ve öğretmenler arasında güçlü

bir tutarlılık olduğunu ve bu tutarlılığın iki yıllık bir izlemede de istikrar gösterdiğini bulmuşlardır. Buna karşın Kaat ve diğerleri (2013) yaptıkları çalışmada, ebeveynler ve öğretmenler arasında OSB ile ilgili bozukluklar ve semptomlar konusunda düşük düzeyde uyuma olduğunu tespit etmişlerdir. Vickerstaff ve diğerleri (2007) ise OSB olan çocuklar hakkında yaptıkları incelemede öğretmenler ve ebeveynler arasında anlamlı bir korelasyon olmadığını tespit etmişlerdir.

OSB olan çocuklarda, sosyal işlevlerde görülen yetersizliklerin bozukluğun önemli göstergelerinden biri olduğu bilinmektedir. Ancak, sosyal işlevselliği tanımlamak ve ölçmek zordur, ortama ve bağlama göre evde veya okulda ölçülmek istenen davranışta farklılıklar görülebilmektedir. Bu nedenle, ebeveynlerin ve öğretmenlerin sosyal işlevselliğin niteliği hakkında anlaşmazlıklar yaşamaları da şaşırtıcı değildir (Jepsen vd., 2012; Tasse & Lecavalier, 2000). Bu durumun sadece sosyal işlevsellikle ilgili davranışlar için değil aynı zamanda bozukluğun bir diğer göstergesi olan tekrarlayıcı davranışlar için de geçerli olabileceği düşünülmektedir. Çünkü, tekrarlayıcı davranışların sıklığında ve çeşitliliğinde de farklı bağlamlara göre farklılıklar görülebilmektedir (Joosten, Bundy, ve Einfeld, 2012; Kirby, Boyd, Williams, Faldowski, & Baranek, 2017; Lampi, Fitzpatrick, Romero, Amaral, ve Schmidt, 2018; Turner, 1999). Davranışın ortaya çıktığı bağlamın yanı sıra değerlendiricilerin ilgili davranışa bakış açıları ve bu davranışı algılamadaki farklılıklarının ve tutumlarının da puanlayıcılar arasında farklılıklara neden olabileceği düşünülmektedir (Azad, Reisinger, Xie, ve Mandell, 2016; Ryland, Hysing, Posserud, Gillberg, ve Lundervold, 2012). Değerlendiriciler arasında ortaya çıkan bu tür tutarsızlıklar, çocuklar hakkında çelişkili değerlendirme verilerini yorumlama ve müdahale hedeflerini belirleme süreçlerinde uzmanlar arasında anlaşma ve işbirliği konularında oldukça fazla zaman harcanmasına neden olmaktadır (Azad vd., 2016). Bu nedenle, puanlayıcılar arası güvenilirliğin sağlanması büyük önem taşımaktadır.

Puanlayıcılar arası güvenilirlik, aynı görev/kişi için iki ya da daha fazla puanlayıcı arasındaki uyumu ve tutarlılığı ifade eder. Puanlayıcılar arası güvenilirliği tespit etmede farklı yöntem ve yaklaşımlar vardır. Bunlardan en sık kullanılanları; korelasyon teknikleri, ortalamaların karşılaştırılması (t testi/varyans analizi), Cohen'in kappası (κ), ağırlıklandırılmış kappa (κ_w), Fleiss'in kappası, Kendall'ın uyum katsayısı (W) ve uyum yüzdesidir. Araştırmacıların bu yöntem ve yaklaşımları kullanması gereken durumlar; araştırmacının kaçındığı ölçme hatasına ve ölçümlerin elde edildiği ölçek türüne (sınıflama, sıralama, eşit aralıklı, eşit oranlı) bağlıdır (Goodwin, 2001).

Özel eğitim alanında öğretmen ya da ebeveyn gibi farklı puanlayıcılardan elde edilen verilere dayalı olarak hem bireylerin gelişimsel özelliklerinin belirlenmesi hem de müdahalenin planlanması süreçlerinde farklı kararlar verilebilmektedir. Bu süreçlerde bireyler hakkında verilen kararların geçerliği ve güvenilirliği puanlayıcılara bağlı olmaktadır. Bu nedenle, farklı veri toplama kaynaklarından elde edilen bilgiler arasında tutarsızlık olmasının yaratabileceği sonuçlar göz önüne alındığında, çocuğu iyi tanıyan kişiler arasındaki uyum kritik bir çalışma alanı olarak görülmektedir. Bu çalışmada, OSB olan çocuklarda bozukluğun derecesini ölçmede, bozukluğun göstergeleri arasında yer alan tekrarlayıcı davranışların ve sosyal iletişim-etkileşim davranışlarının değerlendirilmesinde alanyazında sıklıkla kullanılan dört ölçek için aynı çocukların hem ebeveynleri hem öğretmenleri tarafından yapılan puanlamalar arasındaki uyumun karşılaştırılması amaçlanmıştır. OSB olan bireylerde kullanılan bu ölçekler; i) Sosyal Yanıtlayıcılık Ölçeği-2-Okul Öncesi Türkçe Formu [SOYÖ-2-OTF], ii) Gilliam Otizm Derecelendirme Ölçeği-2-Türkçe Versiyonu (GOBDÖ-2-TV), iii) Sosyal İletişim Kontrol Listesi-Revize-Okulöncesi Türkçe Versiyonu (SİLKOL-R-OTV) ve iv) Tekrarlayıcı Davranışlar Ölçeği-Revize-Türkçe Versiyonu (TEDÖ-R-TV) ölçekleridir. Puanlayıcılar arasındaki uyumu belirlemede, korelasyon teknikleri, ortalamaların karşılaştırılması ve Cohen Kappa tekniği kullanılarak karşılaştırmalar yapılmıştır. Bu çalışmada OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini puanlayan ebeveyn ve öğretmenler açısından bu ölçeklere (SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV) ilişkin aşağıdaki sorulara yanıt aranmıştır:

1. Alt ölçekler ve tüm ölçek bazında ebeveyn ve öğretmenler için iç tutarlılık katsayıları nasıldır?
2. Ebeveyn ve öğretmenlerin puanlamaları arasındaki korelasyona dayalı tutarlılık, ebeveynler tüm grup ve eğitim düzeyine göre (lisans ve üstü ile altı olmak üzere) iki gruba ayrıldığında nasıldır?
3. Ebeveyn ve öğretmenlerin puanlamaları arasındaki tutarlılık, ortalamalar arası karşılaştırma tekniğine göre nasıldır?
4. SOYÖ-2-OTF ve GOBDÖ-2-TV ölçeklerini puanlayan ebeveyn ve öğretmenlerin çocukları tipik gelişen ve OSB olan çocuklar olarak sınıflamaları arasındaki tutarlılık Cohen kappa tekniğine göre nasıldır?

Yöntem

Araştırmanın Modeli

Bu çalışma, OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren farklı ölçeklerde ebeveyn ve öğretmenlerin aynı OSB olan çocuk için verdikleri tepkiler arasındaki uyumu farklı tekniklere göre belirlemeyi yani ebeveyn ve öğretmen tepkileri arasındaki uyumu belirlemeyi amaçladığından ilişkisel tarama modelinde bir araştırmadır (Karasar, 2000).

Çalışma Grubu

Bu çalışmanın katılımcıları, kolay erişilebilir örnekleme yaklaşımıyla belirlenmiştir. Çalışma grubu, 2018-2019 eğitim-öğretim döneminde, Ankara ilinde Milli Eğitim Bakanlığına bağlı okullara ve rehabilitasyon merkezlerine devam eden, SOYÖ-2-OTF için 2-7 yaşları arasında OSB tanısı olan 380 çocuğun 330 ebeveyni ve 378 öğretmeni (Bakkaloğlu, Kudret, Ökcün Akçamuş, Yalçın ve Demir, 2020); GOBDÖ-2-TV için 3-23 yaşları arasında OSB tanısı olan 380 bireyin 379 ebeveyni ve 329 öğretmeni (Bakkaloğlu vd., 2020); SİLKOL-R-OTV için 2-7 yaşları arasında OSB tanısı olan 365 çocuğun 315 ebeveyni ve 361 öğretmeni (İnan, Ökcün-Akçamuş, Bakkaloğlu ve Yalçın, 2020) ve TEDÖ-R-TV için 3-23 yaşları arası OSB tanısı olan 300 bireyin 292 ebeveyni ve 298 öğretmeninden (Ökcün-Akçamuş, Bakkaloğlu, Demir ve Bahap-Kudret, 2019) oluşmaktadır.

Ebeveyn ve öğretmenlerin cinsiyet, yaş ve eğitim düzeyi gibi bazı demografik değişkenlere ilişkin betimleyici bilgileri de toplanmış ve takiben sunulmuştur. SOYÖ-2-OTF ve GOBDÖ-2-TV için ebeveynlerin yaş ortalaması 35.49 (ss=6.04) olup en düşük değer 21 en yüksek değer 61'dir. Öğretmenlerin ise cinsiyete göre dağılımına bakıldığında, katılımcıların 273 kadın (%71.84), 96 erkekten (%25.26) oluştuğu görülmüştür. SİLKOL-R-OTV için ebeveynlerin yaş ortalaması 35.40 (ss=6.01) olup en düşük değer 21 en yüksek değer 61'dir. SOYÖ-2-OTF, GOBDÖ-2-TV ve SİLKOL-R-OTV ölçeklerini yanıtlayan ebeveynlerin eğitim durumları incelendiğinde, katılımcıların 38'i ilkokul (%13.4), 34'ü ortaokul (%11.9), 91'i lise (%31.9), 93'ü üniversite (%32.6) ve 11'i lisansüstü (%3.9) mezundur. Onsekiz ebeveyn ise (%6.3) eğitim durumunu belirtmemiştir. Ebeveynlerin mesleklere göre dağılımı incelendiğinde 137'sinin yani yaklaşık yarısının (%48.07) ev hanımı olduğu ardından 17 ebeveynin (%5.96) ise öğretmen olduğu görülmüştür. Diğer pek çok meslek grubundan ebeveyn çalışmada yer almaktadır. SİLKOL-R-OTV için öğretmenlerin ise cinsiyete göre dağılımına bakıldığında, 209 kadın (%57.26) ve 70 erkekten

(%19.18) oluştuğu görülmüş ve bazı öğretmenlerin cinsiyete ilişkin bilgileri eksik olduğundan belirtilememiştir. TEDÖ-R-TV için ebeveynlerin yaş ortalaması 39.52 (ss=6.69) olup en düşük değer 24 en yüksek değer 66'dır. Ebeveynlerin eğitim durumu TEDÖ-R-TV için incelendiğinde, katılımcıların 40'ı ilkokul (%17.86), 14'i ortaokul (%6.25), 61'i lise (%27.23), 82'si üniversite (%36.61) ve 11'i lisansüstü (%4.91) mezunudur. 16 ebeveyn ise (%7.14) eğitim durumunu belirtmemiştir. TEDÖ-R-TV ebeveynlerin mesleklere göre dağılımı incelendiğinde, 118'inin yani yarıdan fazlasının (%52.68) ev hanımı olduğu ardından 22 ebeveynin (%9.82) ise öğretmen olduğu görülmüştür. Diğer pek çok meslek grubundan ebeveyn vardır. TEDÖ-R-TV için öğretmenlerin yaş ortalaması 31.36 (ss=8.41) olup en düşük değer 22 en yüksek değer 63'tür.

Veri Toplama Araçları

Sosyal Yanıtlayıcılık Ölçeği-2 (Social Responsiveness Scale-2)-Okul Öncesi Türkçe Formu (SOYÖ-2-OTF)

OSB ile ilgili semptomları ölçmek amacıyla Constantino ve Gruber (2012) tarafından geliştirilmiştir. Beş alt ölçek (sosyal farkındalık [8 madde], sosyal biliş [12 madde], sosyal iletişim [22 madde], sosyal motivasyon [11 madde], sınırlı ilgi alanları ve tekrarlayıcı davranışlar [12 madde]) olmak üzere toplam 65 maddeden oluşmaktadır. SOYÖ-2-OTF, DSM-5'teki OSB tanı kriterlerine (sosyal iletişim bozukluğu ve sınırlı ilgi/tekrarlayıcı davranışlar) uygun bir ölçektir. Ölçekteki maddeler dördümlü Likert derecelendirme ile puanlanmaktadır (0: Doğru değil, 1: Bazen doğru, 2: Sıklıkla doğru, 3: Hemen hemen her zaman doğru). Ölçekteki maddelerden on yedisi (3, 7, 11, 12, 15, 17, 21, 22, 26, 32, 38, 40, 43, 45, 48, 52, 55) tersine puanlanmakta ve ölçekten 65 ile 260 arasında puan alınmaktadır. Ölçekten yüksek puan alınması, OSB derecesinin yüksek olduğunu göstermektedir. Ölçeğin Türkçeye uyarlanması çalışması Bakkaloğlu ve diğerleri (2020) tarafından yapılmıştır. Beş faktörlü modelin uyum değerleri ebeveyn tepkilerine göre ($X^2=5302.684$, $Sd=2005$, $p<.05$, $X^2/Sd=2.64$, $RMSEA=0.071$, $CFI=0.78$) ve öğretmen tepkilerine göre ($X^2=6157.422$, $Sd=2005$, $p<.05$, $X^2/Sd=3.07$, $RMSEA=0.074$, $CFI=0.813$) incelendiğinde, modellerin veriye kabul edilebilir düzeyde uyum sağladığını göstermektedir. Ölçeğin iç tutarlılık katsayıları ebeveyn ve öğretmenler için sırasıyla .95 ile .96'dır. Alt ölçek güvenilirlik katsayıları ise ebeveyn için .59 ile .87 arasında, öğretmen için .69 ile .90 arasında değişmektedir (Bakkaloğlu ve diğ., 2020).

Gilliam Otizm Derecelendirme Ölçeği-2 (Gilliam Autism Rating Scale-2) Türkçe Versiyonu (GOBDÖ-2-TV)

Gilliam (2006) tarafından geliştirilen bu ölçek, 3-23 yaş aralığındaki bireyleri OSB belirtileri ya da tanı özellikleri açısından değerlendirmek amacıyla kullanılmaktadır. Ölçek, stereotipik davranışlar (14 madde), sosyal etkileşim (14 madde) ve iletişim (14 madde) olmak üzere üç alt ölçekten oluşmaktadır. Ölçekteki maddeler dörtlü derecelendirme ile puanlanmakta (0: Hiç gözlenmedi, 1: Nadiren gözlendi, 2: Bazen gözlendi, 3: Sıklıkla gözlendi) ve alınan toplam puanlar standart puanlara dönüştürülerek bir Otizm Bozukluk İndeksi (OBİ) puanı elde edilmektedir. Bu puanın 85 ve üstünde olması, OSB olma olasılığının yüksek, 70-84 puan arası orta ve 69 puan ve altında olması ise OSB olma olasılığının düşük olduğunu göstermektedir. GOBDÖ-2-TV'nin Türkçeye uyarlama ve geçerlik-güvenirlik çalışmaları Diken, Ardıç, Diken ve Gilliam (2012) tarafından yapılmıştır. Bu uyarlama çalışmasında 1191 katılımcıdan (%40.7'si ebeveyn, %59.3'ü öğretmen) toplanan verilerin tümü kullanılarak geçerlik güvenirlik analizleri yapılmış, öğretmen/uzman ve ebeveynler için ayrı ayrı analizler yapılmamıştır. Ölçeğin Türkçe versiyonunun doğrulayıcı faktör analizi sonuçları faktör yapısına ilişkin model uyumunun iyi düzeyde ($X^2=1730.08$, $Sd = 813$, $X^2/Sd=2.13$, $CFI=.89$, $RMSEA=0.071$) olduğunu göstermektedir. İç tutarlılık katsayıları ise .77 ve .85 arasında değişmektedir (Diken ve diğ., 2012).

Sosyal İletişim Kontrol Listesi-Revize (Social Communication Checklist-Revised) Okulöncesi Türkçe Versiyonu (SİLKOL-R-OTV)

Bu ölçek OSB olan çocuklarda sosyal iletişim ve etkileşim becerilerinin değerlendirilmesi amacıyla Ingersoll ve Dvortcsak (2010) tarafından geliştirilmiştir. Ölçek, sosyal katılıma yönelik 15, ifade edici dile yönelik 30, alıcı dile yönelik 8 ve taklit ile oyun becerilerine yönelik 17 madde olmak üzere toplam 70 madde içermektedir. Kontrol listesindeki maddeler “Nadiren ya da henüz değil (1 puan), Bazen ama sürekli değil (2 puan) ve Genellikle, zamanının en az %75'i (3 puan)” şeklinde puanlamakta ve 70 ile 210 arasında puan alınabilmektedir. Türkçeye uyarlaması İnan ve diğerleri (2020) tarafından gerçekleştirilen kontrol listesinin, doğrulayıcı faktör analizi sonucu uyum iyiliği değerlerinin hem ebeveyn formunda ($X^2=5121.967$, $Sd = 2339$, $p<.05$, $X^2/Sd=2.19$, $RMSEA: .068$; $CFI: .916$; $TLI: .913$) hem de öğretmen formunda ($X^2=5226.578$, $Sd = 2339$, $p<.05$, $X^2/Sd=2.23$, $RMSEA: .069$; $CFI: .954$, $TLI: .953$) iyi uyuma işaret ettiği görülmüş ve SİLKOL-R-OTV'nin dört faktörlü orijinal yapısının Türk kültüründe doğrulandığı sonucuna ulaşılmıştır. SİLKOL-R-OTV toplam puanları için Cronbach alfa iç tutarlılık katsayısı ebeveynler için .98 iken

öğretmenler için .99'dır. Ölçeğin güvenirlik analizlerinde hem ebeveyn hem de öğretmen formlarında tüm alt testler için iç tutarlılık güvenirlik katsayıları .80'in üzerindedir (İnan ve diğ., 2020).

Tekrarlayıcı Davranışlar Ölçeği-Revize (Repetitive Behavior Scale-Revised)-Türkçe Versiyonu (TEDÖ-R-TV)

Ölçek Bodfish, Symons, Parker ve Lewis (2000) tarafından tekrarlayıcı davranışları ve bu davranışların şiddetini değerlendirmek amacı ile deneysel olarak geliştirilmiş bir klinik derecelendirme ölçeğidir. Ölçek, stereotipik davranışlar (6 madde), kendine zarar verici davranışlar (8 madde), kompulsif davranışlar (8 madde), törensel davranışlar (6 madde), aynılık/tekdüzelik davranışları (11 madde), sınırlı ilgi alanı (4 madde) olmak üzere altı alt ölçekten ve toplam 43 maddeden oluşmaktadır. Ölçekte yer alan maddeler dörtlü derecelendirme kullanılarak (0: Davranış yok, 1: Hafif düzey, 2: Orta düzey, 3: Ağır düzey) puanlanmaktadır. Ölçekten alınan puan arttıkça tekrarlayıcı davranışların şiddetinin de arttığı kabul edilmektedir. Ölçeğin Türkçe uyarlama çalışması hem öğretmenler hem de ebeveynler tarafından doldurulan ölçekler için ayrı ayrı gerçekleştirilmiştir. Ebeveynler tarafından doldurulan ölçekler için geçerlik güvenirlik çalışması, Ökcün-Akçamuş ve diğerleri (2019) tarafından yapılmış ve bu araştırmanın sonucunda ölçeğin Türkçe Formunun altı faktörlü yapısının, özgün ölçeğin faktör yapısı ile uyumlu olduğu belirlenmiştir. Doğrulayıcı faktör analizinin sonuçları, uyum iyiliğine yönelik kestirimler göz önüne alındığında, altı faktörlü ölçek yapısının uyum indekslerinin iyi düzeyde ($X^2=16.7912$, $Sd=840$, $p<.05$, $X^2/Sd=2.00$, $RMSA=.059$, $SRMR=.067$, $NNFI=.94$, $CFI=.94$) olduğu görülmektedir. TEDÖ-R-TV ölçeğinin güvenilirliğine ilişkin analizler incelendiğinde, tüm alt ölçekler için ve toplam ölçek puanı için Cronbach alfa değerlerinin .73 ve üzerinde olduğu, toplam puan için Cronbach alfa katsayısının .94 olduğu bulunmuştur. Öğretmenler tarafından doldurulan ölçekler için geçerlik güvenirlik çalışması ise Demir ve diğerleri (2018) tarafından gerçekleştirilmiş, bu çalışmada da altı faktörlü ölçek yapısının uyum indekslerinin iyi düzeyde olduğu ($X^2=2415.92$, $Sd=838$, $p<.05$, $X^2/Sd=2.88$, $RMSA=.079$, $SRMR=.064$, $NNFI=.93$, $CFI=.94$), tüm alt ölçekler için Cronbach alfa değerlerinin .77 ve üzerinde olduğu, ölçek toplam puanı için ise bu değer .93 olduğu belirlenmiştir. Dolayısıyla analizler sonucunda, TEDÖ-R-TV'nin geçerli ve güvenilir bir araç olduğu saptanmıştır (Ökcün-Akçamuş ve diğ., 2019).

Verilerin Analizi

Çalışmanın ilk araştırma sorusu kapsamında OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren farklı ölçekleri (SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV) ebeveyn ve öğretmenlerin puanlamasına ilişkin iç tutarlılığın belirlenmesi amaçlanmıştır ve bu amaçla alt ölçekler ve tüm ölçek bazında Cronbach alfa iç tutarlılık katsayıları hesaplanmıştır. Cronbach alfa iç tutarlılık katsayısının 0.00 ile 0.40 arasında olması ölçeğin güvenilir olmadığını, 0.41 ile 0.60 aralığında olması düşük düzeyde güvenilir olduğunu, 0.61 ile 0.80 aralığında olması orta düzeyde güvenilir olduğunu, 0.81 ile 1.00 aralığında olması ise yüksek düzeyde güvenilir olduğunu göstermektedir (Alpar, 2010). İki grubun puanlamalarına ilişkin iç tutarlılık katsayıları arasındaki farkın manidarlığını belirlemek için güvenilirlik katsayıları Fischer Zr puanına dönüştürülmüştür. Fischer'in Zr değerine çevrilen katsayılar arasındaki farkın manidarlığı (1) numaralı eşitlik yardımıyla test edilmiştir (Akhun, 1984).

$$Z = \frac{Z_{r_1} - Z_{r_2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \quad (1)$$

Eşitlikte, "Z" güvenilirlik katsayıları farklarına ait Z istatistiği değerini, " Z_{r_1} " ebeveynlerin puanlamalarına ilişkin güvenilirlik katsayısının Fischer Zr değerine dönüştürülmüş değerini, " Z_{r_2} " öğretmenlerin puanlamalarına ilişkin güvenilirlik katsayısının Fischer Zr değerine dönüştürülmüş değerini, " n_1 " ebeveynlerin örneklem büyüklüğünü, " n_2 " öğretmenlerin örneklem büyüklüğünü göstermektedir.

İkinci araştırma sorusu kapsamında, ebeveyn ve öğretmenlerin ölçeklere ilişkin puanlamaları arasındaki tutarlılığın korelasyona dayalı olarak belirlenmesi amaçlanmıştır. Bu amaçla, ebeveyn ve öğretmenlerin ölçeklere ilişkin puanlamaları arasındaki ilişki incelenmeden önce ölçeklerden elde edilen puanların normal dağılım varsayımını karşılama durumu incelenmiştir. SOYÖ-2-OTF'nin ve SİLKOL-R-OTV'nin ebeveyn ve öğretmen tepkileri için alt ölçekler ve tüm ölçek toplam puanı için çarpıklık ve basıklık katsayıları incelenmiş, -1 ile +1 arasında değerler aldığı görülmüştür. Öğretmenler tarafından yanıtlanan GOBDÖ-2-TV'nin alt ölçekler ve tüm ölçek puanı için çarpıklık ve basıklık katsayıları incelenmiş, -1 ile +1 arasında değerler aldığı görülmüştür. GOBDÖ-2-TV'nin ebeveyn tepkileri için alt ölçekler ve tüm ölçek puanı için çarpıklık ve basıklık katsayıları incelenmiş, çarpıklık katsayılarının iletişim alt ölçeği için +2 değeri aldığı, basıklık

katsayısının ise stereotipik davranışlar alt ölçeğinde 4.747, iletişimde 6.914, sosyal etkileşimde 9.072 olduğu görülmüştür. Bu nedenle, GOBDÖ-2-TV'nin ebeveyn tepkileri için ayrıca ortalama, ortanca ve mod değerleri incelenmiş, bu değerlerin birbirine oldukça yakın olduğu görülmüş, dağılımın histogramı da incelendiğinde normal dağılım varsayımını karşıladığı sonucuna varılmıştır. Çarpıklık ve basıklık katsayılarının -3 ile +3 sınırları arasında olması ve ortalama, ortanca, mod değerlerinin birbirine yakınlığı nedeniyle normal dağılım varsayımının sağlandığı sonucuna ulaşılmıştır (Kalaycı, 2010). TEDÖ-R-TV'nin ebeveyn ve öğretmen tepkileri için alt ölçeklerin puanlarının ve ölçek toplam puanının çarpıklık ve basıklık katsayıları incelenmiş, çarpıklık katsayısının her iki form için de -2 ile +2 arasında olduğu görülmüştür. Basıklık katsayısının ise ebeveyn formunda KZVD alt ölçeğinde bu sınırları aştığı 4.010 değeri olduğu, öğretmen formunda ise KZVD alt ölçeği için 3.537, KD alt ölçeği için 3.425 olduğu diğer alt ölçeklerde tüm değerlerin -2 ile +2 arasında olduğu görülmüştür. KZVD alt ölçeği öğretmen formu için ortalama 2.79, ortanca 1 ve mod 0 olarak bulunmuştur. KZVD alt ölçeği ebeveyn formu için ortalama 2.13, ortanca 1 ve mod 0 olarak bulunmuştur. KD alt ölçeği öğretmen formu için ortalama 3.81, ortanca 3 ve mod 0 olarak bulunmuştur. Bu analizlerin sonucunda, KZVD ve KD alt ölçeklerinin normallik varsayımını sağlama durumu ile ilgili kuşkuya düşülmüş, TEDÖ-R-TV'nin tüm ölçek ve alt ölçekleri için hem ilişkili örneklem t testi hem de Wilcoxon İşaretli Sıralar testi yapılmıştır. Normallik varsayımı sağlanan ölçeklerde, ebeveynler hem tek bir grup olarak hem de eğitim düzeyine göre (lisans ve üstü ile lisans altı olmak üzere) iki gruba ayrıldığında, ebeveynler ile öğretmenlerin puanlamaları arasındaki ilişki Pearson Momentler Çarpımı Korelasyon (PMÇK) katsayısının hesaplanması ile incelenmiştir. TEDÖ-R-TV'nin normallik varsayımını sağlamama durumu nedeniyle Spearman Brown Korelasyon Katsayısı hesaplanmıştır. Korelasyon katsayısının mutlak değer olarak, 0.00 ile 0.30 arasında olması, düşük; 0.30 ile 0.70 arasında bir değer olması, orta ve 0.70 ile 1.00 arasında bir değer olması ise yüksek düzeyde ilişkiyi göstermektedir (Büyüköztürk, 2013).

Üçüncü araştırma sorusu kapsamında ebeveyn ve öğretmenlerin verdiği puanların ortalamaları arasında anlamlı bir fark olup olmadığı normallik varsayımı sağlandığından İlişkili Örneklem t Testi ile kıyaslanmıştır. Puanlayıcılar arasında puanlamada katılık veya cömertlik açısından manidar fark varken, PMÇK katsayısı yüksek pozitif olabilmektedir. Korelasyonun yüksekliği, puanların birlikte değişim gösterdiğini ifade etmektedir ancak tek başına PMÇK katsayısının dikkate alınmaması, ortalamaların da kıyaslanması önerilmektedir (Goodwin, 2001).

Dördüncü araştırma sorusu kapsamında, ebeveyn ve öğretmen puanlayıcıları arasındaki uyum Cohen kappa katsayısı ile hesaplanmıştır. Cohen kappa istatistiği, iki kategorik değişken dizisi arasındaki (sadece iki bağımsız değerlendiricinin değerlendirmeleri arası) uyumun sorgulandığı bir tekniktir (Can, 2016). Kappa katsayısının .20'den küçük olması yetersiz uyuma, .21 ile .40 arası düşük düzeyde uyuma, .41 ile .60 arası orta düzeyde uyuma, .61 ile .80 arası iyi düzeyde uyuma, .81 ile 1 arası ise çok iyi düzeyde uyuma işaret etmektedir (Can, 2016). SOYÖ-2-OTF'de toplam puan ölçeğin orijinalinde belirtildiği gibi T puanına dönüştürülmüş, T puanları 59 ve altı olanlar (tipik gelişen çocuklar) ile 60 ve üstü (hafif, orta ve yüksek riskli çocuklar) olmak üzere iki kategorili hale dönüştürülmüştür (Constantino & Gruber, 2012). GOBDÖ-2-TV'den elde edilen OBİ puanlarına göre katılımcı çocuklar orta ve ağır ile hafif grup olarak iki kategoriye dönüştürülmüştür. SİLKOL-R-OTV ve TEDÖ-R-TV'ye ilişkin kesme puanları olmadığından o ölçekler için puanlayıcılar arası uyum Cohen kappa istatistiğine göre hesaplanamamıştır.

Bulgular

İç Tutarlılık Katsayısına Göre Ebeveyn ve Öğretmen Puanlamaları Arasındaki Tutarlılık

OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenler için alt ölçekler ve tüm ölçek bazında iç tutarlılık katsayılarını belirlemek için hesaplanan Cronbach alfa katsayıları Tablo 1'de sunulmuştur.

Tablo 1'e göre, SOYÖ-2-OTF'nin ebeveyn tepkilerine göre iç tutarlılık katsayıları .610 ile .950 arasında; öğretmen tepkilerine göre iç tutarlılık katsayıları .692 ile .962 arasında; GOBDÖ-2-TV'nin ebeveyn tepkilerine göre iç tutarlılık katsayıları .832 ile .941 arasında; öğretmen tepkilerine göre iç tutarlılık katsayıları .846 ile .942 arasında; SİLKOL-R-OTV ebeveyn tepkilerine göre iç tutarlılık katsayıları .930 ile .981 arasında; öğretmen tepkilerine göre iç tutarlılık katsayıları .929 ile .982 arasında; TEDÖ-R-TV'nin ebeveyn tepkilerine göre iç tutarlılık katsayıları .686 ile .929 arasında; öğretmen tepkilerine göre iç tutarlılık katsayıları .775 ile .938 arasında değişmektedir. Bu farklılıkların manidar olup olmadığını test etmek için iç tutarlılık katsayılarına Fisher Z dönüşümü yapılarak katsayılar arasındaki farkın manidarlığı test edilmiştir.

Tablo 1. *SOYÖ-2-OTF*, *GOBDÖ-2-TV*, *SİLKOL-R-OTV* ve *TEDÖ-R-TV* ölçeklerine ilişkin ebeveyn ve öğretmen puanlamaları için cronbach alfa katsayıları

Ölçek	Alt Ölçek	Ebeveyn		Öğretmen		Z
		N	α	N	α	
SOYÖ-2-OTV	SF	313	.610	355	.692	-1,845
	SB	307	.724	346	.804	-2,502*
	Sİ	287	.881	330	.906	-1,607
	SM	323	.783	354	.824	4,368*
	SİTD	310	.809	349	.811	-0,075
	Toplam	258	.950	286	.962	-1,677
GOBDÖ-2-TV	İ	316	.877	368	.882	-0,312
	SE	267	.832	323	.846	-0,610
	SD	309	.909	362	.917	-0,621
	Toplam	251	.941	310	.942	-0,105
SİLKOL-R-OTV	SK	300	.946	354	.929	1,824
	İED	280	.966	336	.965	0,177
	AD	315	.938	361	.939	-0,108
	TO	288	.930	294	.936	-0,553
	Toplam	235	.981	241	.982	-0,304
TEDÖ-R-TV	SD	283	.738	294	.791	-1,529
	KZVD	281	.732	291	.795	-1,816
	KD	275	.772	293	.806	-1,082
	TD	287	.727	257	.783	-1,563
	ATD	276	.868	284	.870	-0,098
	SİA	287	.686	298	.775	-2,298*
	Toplam	248	.929	240	.938	-0,784

Not: *SOYÖ-2-OTF* (Sosyal Yanıtlayıcılık Ölçeği-2-Okul Öncesi Türkçe Formu): SF: Sosyal Farkındalık, SB: Sosyal Biliş, Sİ: Sosyal İletişim, SM: Sosyal Motivasyon, SİTD: Sınırlı İlgi Alanları ve Tekrarlayıcı Davranışlar.

GOBDÖ-2-TV (Gilliam Otizm Derecelendirme Ölçeği-2-Türkçe Versiyonu): İ: İletişim, SE: Sosyal Etkileşim, SD: Stereotipik Davranışlar.

SİLKOL-R-OTV (Sosyal İletişim Kontrol Listesi-Revize-Okulöncesi Türkçe Versiyonu): SK: Sosyal Katılım, İED: İfade Edici Dil, AD: Alıcı Dil, TO: Taklit/Oyun.

TEDÖ-R-TV (Tekrarlayıcı Davranışlar Ölçeği-Revize-Türkçe Versiyonu): SD: Stereotipik Davranışlar, KZVD: Kendine Zarar Verici Davranışlar, KD: Kompulsif Davranışlar, TD: Törenselleşmiş Davranışlar, ATD: Aynılık-Tekdüzelik Davranışları, SİA: Sınırlı İlgi Alanı.

*p<.05

Tablo 1’de görüldüğü gibi, SOYÖ-2-OTF’nin SB ve SM alt ölçeklerinde ve TEDÖ-R-TV’nin SİA alt ölçeğinde ebeveyn ve öğretmenlerin verdikleri puanlara ilişkin güvenilirlik katsayıları arasındaki fark manidardır ($p<.05$). Üç alt ölçekte de öğretmenlerin verdikleri puanlamalara ilişkin güvenilirlik katsayılarının ebeveynlerden manidar olarak daha yüksek olduğu tespit edilmiştir.

İlişki Katsayılarına Göre Ebeveyn ve Öğretmenlerin Puanlamaları Arasındaki Tutarlılık

OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerine ilişkin ebeveyn ve öğretmenlerin puanlamaları arasındaki ilişkiyi tespit etmek için ebeveyn ve öğretmenlerin tüm alt ölçek ve ölçek toplam puanları arasındaki ilişki incelenmiş ve sonuçlar Tablo 2’de sunulmuştur. Ayrıca Tablo 2’de ebeveynler eğitim düzeyine göre (lisans ve üstü ile lise ve altı olmak üzere) iki gruba ayrıldığında, öğretmenler ile ebeveynlerin puanları arasındaki ilişkiler de sunulmuştur.

Tablo 2’de görüldüğü gibi, eğitim düzeyinden bağımsız olarak ebeveynler tek grup olarak alındığında, SOYÖ-2-OTF’de ebeveyn ve öğretmen arasındaki korelasyonlar incelendiğinde, Sosyal Farkındalık alt ölçeğinde anlamlı ilişki yokken diğer boyutlarda korelasyonlar .304 (sosyal biliş) ile .577 (SOYÖ-2-OTF tüm ölçek) arasında değişmekte ve bu boyutlarda orta düzeyde anlamlı ilişkiler olduğu görülmektedir. GOBDÖ-2-TV’de ebeveynler ve öğretmenlerin puanlamaları arasındaki korelasyonlar incelendiğinde, korelasyon katsayılarının .489 (iletişim) ile .954 (sosyal etkileşim) arasında değiştiği, bu değerlerin orta ve yüksek düzeyde anlamlı olduğu görülmektedir. SİLKOL-R-OTV’de ebeveyn ve öğretmen arasındaki korelasyonlar incelendiğinde, korelasyon katsayılarının .472 (sosyal katılım) ile .644 (ifade edici dil) arasında değiştiği ve ilişkilerin orta düzeyde anlamlı olduğu görülmektedir. TEDÖ-R-TV’de ebeveyn ve öğretmenlerin puanları arasındaki korelasyon katsayılarının da .279 (aynılık-tekdüzelik davranışları) ile .574 (kendine zarar verici davranışlar) arasında değiştiği ve bu değerlerin düşük ve orta düzeyde anlamlı olduğu görülmektedir. Ebeveynlerin eğitim düzeylerine göre bulgular incelendiğinde (Tablo 2), SOYÖ-2-OTV’de, lise ve altında eğitim düzeyine sahip ebeveynlerin puanları ile öğretmenlerin puanları arasında SF alt ölçeği dışında, .286 ile .459 arasında anlamlı korelasyon katsayıları elde edilmiştir.

Tablo 2. Ebeveynlerin eğitim düzeyine göre öğretmenler ile ebeveynlerin SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçek puanları arasındaki ilişkiler

Ölçek	Alt Ölçek	Eğitim düzeyi-1*		Eğitim düzeyi-2*		Tümü	
		N	r	N	r	N	r
SOYÖ-2-OTF	SF	139	-.013	96	.047	250	.056
	SB	131	.286**	89	.334**	236	.304**
	Sİ	113	.382**	82	.402**	207	.392**
	SM	133	.402**	96	.559**	245	.458**
	SİTD	133	.459**	89	.582**	237	.496**
	Toplam	163	.426**	104	.505**	261	.577**
GOBDÖ-2-TV	İ	131	.437**	74	.552**	219	.489**
	SE	149	.965**	99	.955**	266	.954**
	S	140	.376**	84	.403**	280	.563**
	Toplam	162	.754**	104	.775**	203	.821**
SİLKOL-R-OTV	SK	162	.538**	104	.382**	283	.472**
	İED	162	.682**	104	.645**	284	.644**
	AD	161	.515**	103	.458**	282	.497**
	TO	163	.565**	104	.470**	285	.507**
	Toplam	163	.632**	104	.575**	285	.597**
TEDÖ-R-TV	SD	122	.400**	77	.451**	211	.406**
	KZVD	119	.477**	77	.638**	208	.574**
	KD	118	.440**	76	.552**	205	.434**
	TD	101	.291**	77	.589**	186	.354**
	ATD	109	.392**	78	.345**	196	.279**
	SİA	124	.498**	80	.483**	216	.416**
	Toplam	81	.503**	66	.485**	155	.397**

Not: SOYÖ-2-OTF (Sosyal Yanıtlayıcılık Ölçeği-2-Okul Öncesi Türkçe Formu): SF: Sosyal Farkındalık, SB: Sosyal Biliş, Sİ: Sosyal İletişim, SM: Sosyal Motivasyon, SİTD: Sınırlı İlgi Alanları ve Tekrarlayıcı Davranışlar.

GOBDÖ-2-TV (Gilliam Otizm Derecelendirme Ölçeği-2-Türkçe Versiyonu): İ: İletişim, SE: Sosyal Etkileşim, SD: Stereotipik Davranışlar.

SİLKOL-R-OTV (Sosyal İletişim Kontrol Listesi-Revize-Okulöncesi Türkçe Versiyonu): SK: Sosyal Katılım, İED: İfade Edici Dil, AD: Alıcı Dil, TO: Taklit/Oyun.

TEDÖ-R-TV (Tekrarlayıcı Davranışlar Ölçeği-Revize-Türkçe Versiyonu): SD: Stereotipik Davranışlar, KZVD: Kendine Zarar Verici Davranışlar, KD: Kompulsif Davranışlar, TD: Törenselleşmiş Davranışlar, ATD: Aynılık-Tekdüzelik Davranışları, SİA: Sınırlı İlgi Alanı.

*Eğitim düzeyi-1: Lise ve altı, Eğitim düzeyi-2: Üniversite ve üstü.

**p<.01

Eğitim düzeyi üniversite ve üstünde olan ebeveynlerin puanları ile öğretmenlerin puanları arasındaki korelasyonların ise lise ve altında eğitim düzeyine sahip olanlar

arasındaki korelasyonlardan ve toplam ölçek puanı dışında, tüm alt ölçeklerde eğitim düzeyi ayrımı yapılmadığı durumdan da daha yüksek olduğu görülmektedir. Benzer bir bulgu TEDÖ-R-TV'nin SD, KZVD ve TD alt ölçekleri için de geçerlidir. KD alt ölçeğinde ise eğitim düzeyi üniversite ve üstünde olan ebeveynlerin puanları ile öğretmenlerin puanları arasındaki korelasyonların lise ve altında eğitim düzeyine sahip olanlar arasındaki korelasyonlardan daha yüksek olduğu ancak tüm bireylerde bu durumun olmadığı tespit edilmiştir. Ancak GOBDÖ-2-TV'de eğitim düzeyine göre ebeveynler gruplara ayrılrsa da her iki grubun öğretmenlerin puanlamaları ile korelasyonlarının birbirine yakın olduğu görülmektedir. SİLKOL-R-OTV'de ise bu bulguların tersine lise ve altında eğitim düzeyine sahip ebeveynlerin puanları ile öğretmenlerin puanları arasındaki korelasyonların tüm alt ölçeklerde ve toplam ölçek puanında eğitim düzeyi ayrımı yapılmadığı durumdan da daha yüksek olduğu görülmektedir.

Ortalamalar Arası Karşılaştırmaya Göre Ebeveyn ve Öğretmenlerin Puanlamaları Arasındaki Tutarlılık

SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenler arasındaki tutarlılığı, ortalamalar arası karşılaştırmaya dayalı olarak incelemek için İlişkili Örneklem t Testi yapılmış, elde edilen sonuçlar Tablo 3'te sunulmuştur.

Tablo 3. SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenlerin ortalamalarının karşılaştırılmasına ilişkin t-testi sonuçları

Ölçek	Alt Ölçek	Puanlayıcı	\bar{x}	N	ss	t	sd	p
SOYÖ-2-OTF	SF	Ebeveyn	12.26	250	4.77	-	249	.000
		Öğretmen	18.56	250	3.47	17.343		
	SB	Ebeveyn	19.19	236	6.33	-	235	.000
		Öğretmen	25.90	236	4.62	15.588		
	Sİ	Ebeveyn	31.11	207	12.77	-	206	.000
		Öğretmen	45.90	207	8.95	17.178		
	SM	Ebeveyn	16.42	245	6.69	-	244	.000
		Öğretmen	23.98	245	5.08	18.815		
	SİTD	Ebeveyn	13.89	237	6.83	-	236	.000
		Öğretmen	25.74	237	6.66	26.925		
	Toplam	Ebeveyn	92.21	285	33.76	-	284	.000
		Öğretmen	138.44	285	24.56	24.803		
GOBDÖ-2-TV	Ebeveyn	8.94	218	2.83	-	217	.000	
	Öğretmen	14.10	218	8.46	10.156			

SE	Ebeveyn	8.10	265	3.31	-	264	.000		
	Öğretmen	15.24	265	9.97	17.000				
SD	Ebeveyn	7.82	279	3.51	-	278	.000		
	Öğretmen	12.34	279	8.89	10.096				
Toplam	Ebeveyn	23.43	284	8.93	-	283	.000		
	Öğretmen	37.97	284	21.18	15.894				
SİLKOL-R-OTV	SK	Ebeveyn	42.80	283	13.30	.398	282	.691	
		Öğretmen	42.48	283	12.84				
	İED	Ebeveyn	76.89	284	24.87	2.369	283	.018	
		Öğretmen	73.83	284	26.51				
	AD	Ebeveyn	22.49	282	7.53	-1.059	281	.291	
		Öğretmen	22.95	282	7.04				
	TO	Ebeveyn	43.93	285	13.32	1.482	284	.139	
		Öğretmen	42.75	285	13.71				
	Toplam	Ebeveyn	185.47	285	53.13	1.225	284	.222	
		Öğretmen	181.93	285	55.52				
	TEDÖ-R-TV	SD	Ebeveyn	4.33	211	3.52	-3.109	210	.002
			Öğretmen	5.22	211	4.07			
KZVD		Ebeveyn	2.05	208	2.99	-2.656	207	.009	
		Öğretmen	2.66	208	3.95				
KD		Ebeveyn	3.93	205	3.88	-.414	204	.680	
		Öğretmen	4.06	205	4.61				
TD		Ebeveyn	3.44	186	3.13	-1.310	185	.192	
		Öğretmen	3.82	186	3.80				
ATD		Ebeveyn	6.37	196	5.83	-.830	195	.407	
		Öğretmen	6.82	196	6.60				
SİA		Ebeveyn	2.86	216	2.58	-1.270	215	.205	
		Öğretmen	3.12	216	3.03				
Toplam		Ebeveyn	22.78	155	17.25	-2.217	154	.028	
		Öğretmen	26.64	155	21.64				

Not: SOYÖ-2-OTF (Sosyal Yanıtlayıcılık Ölçeği-2-Okul Öncesi Türkçe Formu): SF: Sosyal Farkındalık, SB: Sosyal Biliş, Sİ: Sosyal İletişim, SM: Sosyal Motivasyon, SİTD: Sınırlı İlgi Alanları ve Tekrarlayıcı Davranışlar.

GOBDÖ-2-TV (Gilliam Otizm Derecelendirme Ölçeği-2-Türkçe Versiyonu): İ: İletişim, SE: Sosyal Etkileşim, SD: Stereotipik Davranışlar.

SİLKOL-R-OTV (Sosyal İletişim Kontrol Listesi-Revize-Okulöncesi Türkçe Versiyonu): SK: Sosyal Katılım, İED: İfade Edici Dil, AD: Alıcı Dil, TO: Taklit/Oyun.

TEDÖ-R-TV (Tekrarlayıcı Davranışlar Ölçeği-Revize-Türkçe Versiyonu): SD: Stereotipik Davranışlar, KZVD: Kendine Zarar Verici Davranışlar, KD: Kompulsif Davranışlar, TD: Törenselleşmiş Davranışlar, ATD: Aynılık-Tekdüzelik Davranışları, SİA: Sınırlı İlgi Alanı.

Tablo 3'te görüldüğü gibi, SOYÖ-2-OTF ile GOBDÖ-2-TV ölçeklerinde tüm alt boyutlarda ebeveyn ve öğretmenlerin verdikleri puanların ortalamaları arasındaki farklar anlamlıdır. Her iki ölçekte de öğretmenlerin çocuklar/bireyler için verdikleri ortalama puanlar ebeveynlerden anlamlı olarak daha yüksektir. SİLKOL-R-OTV ölçeğinde sadece *İfade Edici Dil* alt ölçeğinde ebeveyn ve öğretmenlerin puan ortalamaları arasındaki fark anlamlı bulunurken ($t=2.369, p<.05$), diğer alt ölçeklerde anlamlı bir farklılık yoktur. *İfade Edici Dil* alt ölçeğinde ebeveynlerin çocukları için verdikleri puanların ortalaması ($\bar{x}=76.89$) öğretmenlerden ($\bar{x}=73.83$) anlamlı olarak daha yüksektir. TEDÖ-R-TV için ebeveyn ve öğretmenlerden elde edilen puanların ortalamaları arasında *Stereotipik Davranışlar* ve *Kendine Zarar Verici Davranışlar* alt ölçeklerinde ve tüm ölçek toplam puanında istatistiksel olarak anlamlı farklılıklar varken, diğer alt ölçeklerde anlamlı farklılık bulunmamıştır. Her iki alt ölçekte ve ölçeğin tümünden elde edilen puan ortalamaları incelendiğinde, öğretmenlerin öğrenciler için verdikleri puan ortalamaları ebeveynlerinkinden anlamlı olarak daha yüksektir.

TEDÖ-R-TV için ayrıca Wilcoxon İşaretli Sıralar Testi de uygulanmıştır. Sonuçlar, ebeveyn ve öğretmenlerden elde edilen puanların ortalamaları arasında *Stereotipik Davranışlar* ve *Kendine Zarar Verici Davranışlar* alt ölçeklerinde istatistiksel olarak manidar farklılıklar olduğunu göstermektedir ($z=-3.035, p<.05$; $z=-2.123, p<.05$). Diğer alt ölçeklerde ve tüm ölçek toplam puanında ebeveyn ve öğretmenlerin ortalama puanları arasındaki fark manidar değildir, alt ölçeklerin sırasına göre sonuçların $z=-.201, p>.05$; $z=-1.262, p>.05$; $z=-.816, p>.05$; $z=-1.082, p>.05$; $z=-1.820, p>.05$ olduğu görülmüştür. İlişkili Örneklem t Testi ile sonuçlar kıyaslandığında, sadece tüm ölçek toplam puanında farklılık olduğu görülmektedir. Ölçeğin tamamından elde edilen puanlar normallik varsayımını karşıladığından İlişkili Örneklem t Testi sonuçları tartışılmıştır.

Ebeveyn ve Öğretmenlerin Çocukları Tipik Gelişen ve OSB Olan Çocuk Olarak Sınıflamaları Arasındaki Tutarlılık

OSB'de kullanılan ölçekleri (SOYÖ-2-OTF ve GOBDÖ-2-TV) puanlayan ebeveyn ve öğretmenlerin çocukları tipik gelişen ve OSB olan çocuk olarak sınıflamaları arasındaki tutarlılık Cohen kappa tekniğine göre incelenmiş, sonuçlar Tablo 4'te sunulmuştur.

Tablo 4. Ebeveyn ve öğretmenlerin sınıflamaları arasındaki tutarlılığa ilişkin cohen kappa sonuçları

Ölçekler	Kappa	p
SOYÖ-2-OTF	.243	.000
GOBDÖ-2-TV	.560	.000

Not: SOYÖ-2-OTF (Sosyal Yanıtlayıcılık Ölçeği-2-Okul Öncesi Türkçe Formu), GOBDÖ-2-TV (Gilliam Otizm Derecelendirme Ölçeği-2-Türkçe Versiyonu)

SOYÖ-2-OTF için ebeveyn ve öğretmenlerin çocukları, OSB riski taşımayan ile OSB açısından hafif, orta ve yüksek riskli çocuklar olarak sınıflandırmaları arasındaki tutarlılık anlamlıdır ancak düşük düzeyde uyuma işaret etmektedir. GOBDÖ-2-TV’de ise ebeveyn ve öğretmenlerin çocukları OSB olma olasılığı orta ve yüksek ile düşük olarak sınıflandırmaları arasındaki tutarlılık anlamlı ve orta düzeyde uyuma işaret etmektedir.

Sonuç, Tartışma ve Öneriler

Bu çalışmada, OSB’ye özgü davranışsal ve gelişimsel özellikleri değerlendiren farklı ölçeklerde ebeveyn ve öğretmenlerin OSB olan aynı birey için verdikleri tepkiler arasındaki uyumu farklı tekniklere göre belirlemek amaçlanmıştır. Bu kapsamda, ilk olarak OSB olan bireylerin bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenler için alt ölçekler ve tüm ölçek bazında iç tutarlılık katsayıları hesaplanmıştır. SOYÖ-2-OTF ve TEDÖ-R-TV ölçeklerinde bazı alt ölçeklerin orta düzeyde güvenilir olduğu ancak diğer ölçeklerin ve alt ölçeklerin yüksek düzeyde güvenilir olduğu sonucuna ulaşılmıştır. Bu durum, puanlayıcıların yaptıkları puanlamaların kendi içinde iç tutarlılığının yüksek olduğu şeklinde yorumlanabilir. Ayrıca bazı alt ölçeklerde orta düzeyde iç tutarlılık güvenilirlik katsayısı elde edilmesi, alt ölçeği oluşturan madde sayısının diğer alt ölçeklere göre daha az olmasından da kaynaklanabilir. Örneğin, SOYÖ-2-OTF’nin en az maddeye sahip olan alt ölçeği “Sosyal Katılım” iken, bu ölçek içinde en düşük güvenilirlik katsayılarının da bu alt ölçeğe ait olduğu görülmektedir. Alanyazında da, güvenilirliğin madde sayısından etkilendiği ifade edilmektedir (Baykul, 2000; Crocker ve Algina, 1986). Ek olarak, ebeveynlere göre öğretmenlerin puanlamalarına ilişkin Cronbach alfa katsayılarının SİLKOL-R-OTV’nin SK ve İED alt ölçekleri hariç diğer tüm ölçek ve alt ölçeklerde daha yüksek olduğu tespit edilmiştir. SOYÖ-2-OTF’nin SB ve SM alt ölçeklerinde ve TEDÖ-R-TV’nin SİA alt ölçeğinde ebeveyn ve öğretmenlerin

puanlamalarına ilişkin güvenilirlik katsayıları arasındaki fark istatistiksel olarak anlamlı bulunmuştur. Üç alt ölçekte de öğretmenlerin puanlamalarına ilişkin güvenilirlik katsayılarının ebeveynlerinkinden anlamlı olarak daha yüksek olduğu tespit edilmiştir. Murray ve diğerleri de (2009), OSB olan çocukların sosyal beceriler ve davranış ölçümleri üzerine ebeveyn ve öğretmenlerin değerlendirmelerinin karşılaştırıldığı bir çalışmada, öğretmenlerin ebeveynlerden daha tutarlı değerlendirme yaptıklarını bulmuşlardır. Bu bağlamda, yapılan çalışmadan elde edilen sonuçlar, OSB olan çocukların tekrarlayıcı davranışları ve sosyal iletişim ile etkileşim becerileri konusunda öğretmenlerin daha nesnel ve gerçekçi değerlendirme yapmalarından kaynaklanabileceği gibi, ilgili alt ölçeklerde ölçülen davranışların daha nesnel puanlamayı gerektirmesinden de kaynaklanabilir. Ayrıca, ebeveynlerin çocuklarına olan duygusal bağları da nesnel puanlama yapmalarına engel olabilmektedir (Azad ve diğ., 2016). Ek olarak, ebeveynlerin tipik olarak OSB olan az sayıdaki çocuğu gözlemleme fırsatı olduğundan ölçekleri yanıtlarken tipik gelişen çocuklarla kıyaslayarak yanıt verdikleri düşünülmektedir. Öğretmenler ise farklı derecelerde OSB olan daha fazla çocuğu gözlemleme fırsatına sahiptir (Azad ve diğ., 2016; Ryland ve diğ., 2012). Bu durumun da elde edilen sonuçları etkilemiş olabileceği düşünülmektedir.

OSB olan çocukların bozuklukla ilişkili davranışsal ve gelişimsel özelliklerini değerlendiren SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenlerin puanlamaları arasındaki ilişkiyi tespit etmek için ebeveyn ve öğretmenlerin tüm alt ölçek ve ölçek toplam puanları arasındaki ilişki incelenmiştir. SOYÖ-2-OTF’de iki grup arasındaki korelasyonlar incelendiğinde, Sosyal Farkındalık alt ölçeği dışındaki diğer boyutlarda orta düzeyde anlamlı ilişkiler olduğu görülmektedir. GOBDÖ-2-TV ve SİLKOL-R-OTV ölçeklerinde ebeveyn ve öğretmenlerin puanlamaları arasında orta ve yüksek düzeyde anlamlı korelasyonlar elde edilmiştir. TEDÖ-R-TV’de ise ebeveyn ve öğretmenlerin puanlamaları arasında düşük ve orta düzeyde anlamlı ilişki olduğu görülmektedir. Bazı alt ölçeklerde korelasyon katsayılarının düşük olmasına rağmen anlamlı çıkmasının nedeni örneklem büyüklüğü olabilir. Korelasyon katsayısı örneklem büyüklüğünden etkilenmektedir (Bates, Zhang, Dufek, ve Chen, 1996). Başka bir çalışmada da OSB olan çocukların sosyal beceriler ve davranış ölçümleri üzerine ebeveynler ve öğretmenler arasında düşük düzeyde fakat anlamlı ilişki olduğu görülmüştür (Murray ve diğ., 2009). Bu bağlamda, şimdiki çalışmadan elde edilen bulgulara göre özellikle SOYÖ-2-OTF ve TEDÖ-R-TV ölçeklerinde sadece öğretmen ya da ebeveynler tarafından yapılan puanlamaların çocuğun durumunu gerçekçi biçimde yansıtmasının kuşkulu olduğu sonucuna

varılabilir. Alanyazında da öznel, içselleştirilmiş sorunlara ilişkin ölçeklerde ebeveynler ile öğretmenlerin puanlamaları arasındaki uyumun daha düşük olduğu, daha açık, gözlenebilir davranışlarla ilgili maddeleri içeren ölçeklerde ise iki grup arasında düşük-orta düzeyde uyum olduğu görülmüştür (Jepsen, 2012). Vickerstaff ve diğerleri (2007) de OSB olan çocuklar hakkında yaptıkları incelemede benzer sonuçlara ulaşmıştır. Öğretmenler ve ebeveynler arasındaki tutarlılığın, nesnel (daha açık, gözlenebilir) davranışları ölçen alt ölçeklerde daha yüksek olduğu görülmüştür. Yapılan bu çalışmada da bazı ölçeklerde (SOYÖ-2-OTF'nin SF alt ölçeği ile TEDÖ-R-TV'nin ATD alt ölçeği) ebeveynler ile öğretmenlerin puanlamaları arasında düşük düzeyde ilişki bulunmasının olası bir nedeni, öğretmenin sık sık olarak algıladığı bir davranış ebeveynler tarafından daha seyrek algılanabilir ya da tam tersi olabilir. Bu durum, öğretmenlerin OSB tanımlı farklı öğrenciler ile karşılaşırken ebeveynlerin sadece kendi çocuklarını gözlemliyor olmalarından da kaynaklanabilir. Ayrıca, GOBDÖ-2-TV ve SİLKOL-R-OTV ölçeklerinde ebeveyn ve öğretmenler arasındaki korelasyon düzeylerinin orta ve yüksek düzeyde anlamlı olması sadece puanların birlikte değişim gösterdiğinin kanıtıdır. Yani, sadece korelasyon değerleri kullanılarak ebeveyn ve öğretmenlerin aynı birey için benzer puanlamalar yaptığı şeklinde yorumlanmamalıdır. Böyle bir yorum için ebeveyn ve öğretmenler arasındaki korelasyon ile birlikte her iki grubun ortalamalarının da incelenmesi gerekir.

Ebeveynler eğitim düzeyine göre (lisans ve üstü ile lise ve altı olmak üzere) iki gruba ayrıldığında, öğretmenler ile ebeveynlerin puanlamaları arasındaki ilişkilerin SOYÖ-2-OTF ve TEDÖ-R-TV ölçekleri için eğitim düzeyi üniversite ve üstünde olan ebeveynlerin puanları ile öğretmenlerin puanları arasındaki korelasyonların tüm alt ölçeklerde eğitim düzeyi ayrımı yapılmadığı durumdan da daha yüksek olduğu görülmektedir. Sadece SOYÖ-2-OTF tüm ölçek toplam puanının korelasyonu eğitim düzeyi ayrımı yapılmadığında daha yüksektir. Ancak GOBDÖ-2-TV'de eğitim düzeyine göre gruplara ayrılrsa da korelasyonların birbirine yakın olduğu görülmektedir. SİLKOL-R-OTV'de ise diğer ölçeklerden farklı olarak eğitim düzeyi lise ve altında olan ebeveynlerin puanları ile öğretmenlerin puanları arasındaki korelasyonların tüm alt ölçeklerde ve toplam ölçek puanında eğitim düzeyi ayrımı yapılmadığı durumdan da daha yüksek olduğu görülmektedir. SOYÖ-2-OTF ve TEDÖ-R-TV ölçeklerinde, eğitim düzeyi yüksek ebeveynlerin öğretmenler ile daha tutarlı olduğu tespit edilmiştir. Bu ölçeklerin ifadeleri daha karmaşık, anlaşılması zor olabilir, alt eğitim düzeyindekiler yeterince anlamadığı için de doğru tepkiler vermemiş olabilirler. Bu bağlamda, ebeveynlerden veri toplanacağı zaman ölçek ifadelerinin

eğitim düzeyi düşük olan ebeveynler tarafından doğru anlaşıldığından emin olmak için ön görüşmeler yapılması önerilmektedir.

SOYÖ-2-OTF, GOBDÖ-2-TV, SİLKOL-R-OTV ve TEDÖ-R-TV ölçeklerini puanlayan ebeveyn ve öğretmenler arasındaki tutarlılığı, ortalamalar arası karşılaştırmaya dayalı olarak belirlemek amacıyla İlişkili Örneklemeler t testi yapılmıştır. SOYÖ-2-OTF ile GOBDÖ-2-TV ölçeklerinde tüm alt boyutlarda ebeveyn ve öğretmenlerin puanlamaları arasında anlamlı farklılık olduğu, öğretmenlerin öğrenciler için verdikleri puanların ebeveynlerden daha yüksek olduğu belirlenmiştir. GOBDÖ-2-TV ölçeğinde ebeveyn ve öğretmenlerin puanlamaları arasında orta ve yüksek düzeyde anlamlı korelasyonlar elde edilmekle birlikte, puanlayıcılar arasında puanlamada katılık/cömertlik bakımından anlamlı farklılıklar olduğu görülmüştür. Bu ölçeklerin her ikisi de OSB'nin derecesine ilişkin bilgi vermektedir (Constantino & Gruber, 2012; Gilliam, 2006). Dolayısıyla bu araştırmanın bulgularına göre ebeveynlerin çocuklarının OSB derecesini daha hafif puanlama eğiliminde olmalarına karşın, öğretmenlerin OSB derecesini daha şiddetli olarak puanlama eğiliminde oldukları söylenebilir. Bu bağlamda, sadece korelasyonlara göre ebeveyn ve öğretmenler arasındaki uyum hakkında yorum yapmanın uygun olmadığı da tespit edilmiştir. SİLKOL-R-OTV ölçeğinde sadece İfade Edici Dil alt ölçeğinde ebeveyn ve öğretmenlerin tepki ortalamaları arasındaki farklar manidarken diğer alt ölçeklerde manidar bir farklılık yoktur. İfade Edici Dil alt ölçeğinde ebeveynlerin çocukları için verdikleri ortalama puanlar öğretmenlerden manidar olarak daha yüksektir. Bu durum, Vickerstaff ve diğerlerinin (2007) de belirttiği gibi, ebeveyn ve öğretmenlerin çocukların sosyal davranışları hakkındaki gözlemlerinin çevresel etmenler nedeniyle farklı olmasından da kaynaklanabilir. TEDÖ-R-TV için ebeveyn ve öğretmenlerden elde edilen puanların ortalamaları arasında Stereotipik Davranışlar ve Kendine Zarar Verici Davranışlar alt ölçeklerinde ve tüm ölçek puanında istatistiksel olarak manidar farklılıklar varken diğer alt ölçeklerde ebeveyn ve öğretmenlerin ortalama puanları arasındaki fark manidar değildir. Her iki alt ölçekte ve ölçeğin tümünden elde edilen puan ortalamaları incelendiğinde, öğretmenlerin öğrenciler için verdikleri ortalama puanlar ebeveynlerden manidar olarak daha yüksektir. Alanyazında da (Jepsen, 2012; Pekin, 2005; Vickerstaff ve diğerleri, 2007) alt ölçeğin ölçtüğü davranışa göre puanlamalar arasında farklılıklar görüldüğü tespit edilmiştir. Örneğin, Otizm Sosyal Beceriler Profili ölçeğinin beş puanlayıcı tarafından puanlanması ve arasındaki uyum düzeyinin farklı tekniklere göre belirlendiği bir çalışmada, puanlayıcıların katılık veya cömertlik düzeylerinin farklı olduğu ve puanlayıcılar arası tutarsızlığın olduğu tespit

edilmiştir (Pekin, 2015). Bu araştırmada hem TEDÖ-R-TV Stereotipik Davranışlar ve Kendine Zarar Verici Davranışlar Alt Ölçeği, hem de SOYÖ-2-OTF ile GOBDÖ-2-TV alt ölçeklerinde öğretmenlerin puanlarının ebeveyn puanlarından anlamlı olarak yüksek olması, ebeveynlerin spektrum derecesine yönelik daha düşük puan verme eğiliminde olmaları ya da öğretmenlerin daha yüksek puan verme eğiliminde olmaları ile de açıklanabilir.

OSB olan bireylerin gelişimsel ve davranışsal özelliklerini değerlendirmede kullanılan ölçekleri (SOYÖ-2-OTF ve GOBDÖ-2-TV) puanlayan ebeveyn ve öğretmenlerin çocukları OSB olan ve olmayan şeklinde sınıflamaları arasındaki tutarlılığı Cohen kappa tekniğine göre karşılaştırmak için ebeveyn ve öğretmenlerin puanları arasındaki uyum incelenmiştir. SOYÖ-2-OTF için ebeveyn ve öğretmenlerin çocukları, OSB riski olmayan ile OSB açısından hafif, orta ve yüksek riskli çocuklar olarak gruplama davranışları arasındaki tutarlılık anlamlı ancak düşük düzeyde uyuma işaret etmektedir. GOBDÖ-2-TV'de ise ebeveyn ve öğretmenlerin çocukları OSB olma olasılığı orta ve yüksek ile düşük olarak gruplama davranışları arasındaki tutarlılık anlamlı ve orta düzeyde uyuma işaret etmektedir. Alanyazında yapılan çalışmalarda, çocuğun OSB semptomlarının ciddiyetinin de ebeveyn ve öğretmen değerlendirmeleri arasındaki uyumu etkileyebileceği ifade edilmektedir (Azad ve diğ., 2016; Hundert, Morrison, Mahoney, Mundy, ve Vernon 1997; Voelker ve diğ., 2000). Bu çalışmada, ciddi OSB'ye sahip çocuk sayısı fazla olmadığından semptom ciddiyetinin puanlayıcılar arası uyuma etkisi incelenememiştir. Bu bağlamda, gelecek çalışmalarda çocuklar semptom ciddiyetine göre gruplandırıldıktan sonra ebeveyn ve öğretmen değerlendirmeleri arasındaki uyum incelenebilir.

Çalışmadan elde edilen sonuçlara genel olarak bakıldığında, bazı alt ölçeklerde, öğretmenlerin ebeveynlerden daha tutarlı (iç tutarlılık) değerlendirme yaptıkları, sadece öğretmen ya da ebeveynler tarafından yapılan puanlamaların çocuğun durumunu gerçekçi biçimde yansıtmasının kuşkuğu olduğu, sadece korelasyon değerleri kullanılarak ebeveyn ve öğretmenlerin aynı birey için benzer puanlamalar yaptığı şeklinde yorumlanmaması gerektiği tespit edilmiştir. Ayrıca bazı alt ölçeklerde eğitim düzeyi yüksek ebeveynlerin puanlamalarının öğretmenler ile daha tutarlı olduğu, puanlayıcıların (ebeveynler ve öğretmenler) katılık veya cömertlik düzeylerinin farklı olduğu ve bazı alt ölçeklerde puanlayıcılar arası tutarsızlığın olduğu saptanmıştır.

OSB olan çocukların değerlendirilmesi sürecinde çoklu bireyden alınan bilgilerin (örneğin, aileler, terapistler, öğretmenler) kullanılması önerilmektedir. Bu çalışmanın sonucunda elde edilen bulgular da bu durumu desteklemektedir. Ölçeğin ölçtüğü özellik,

ölçek maddelerinin anlaşılabilirliği gibi özellikler de ölçeklerin değerlendirilme sürecinde dikkate alınmalıdır. Ayrıca, dikkatsizlik, kişisel yanlılık, hareleme etkisi, merkeze kayma etkisi, genelleme hatası gibi durumlar da puanlayıcılar arası uyumu etkileyen diğer faktörlerdir (Çıkrıkçı, 2019; Özgüven, 2007). Araştırmacıların bu etkileri de göz önünde bulundurarak puanlayıcılar arası tutarlılığı yorumlamaları ve uygulama öncesi gerektiğinde ebeveynlere bilgilendirme yapılması önerilir.

Etik Kurul İzin Bilgisi: *Bu araştırma, Ankara Üniversitesi Etik kurulunun 04/03/2019 tarihli 77 ve 78 sayılı kararları ile alınan izinlerle yürütülmüştür.*

Yazar Çıkar Çatışması Bilgisi: *Yazarlar herhangi bir çıkar çatışması beyan etmemektedir.*

Yazar Katkısı: *Seher YALÇIN: araştırma deseni, veri toplama, verilerin analizi, makale yazımı; Hatice BAKKALOĐLU: konuyu bulma, araştırma deseni, makale yazımı; Meral Çilem ÖKCÜN AKÇAMUŞ: veri toplama, makale yazımı; Seyda DEMİR: veri toplama, makale yazımı; Zeynep BAHAP KUDRET: veri toplama, makale yazımı.*

Kaynakça

- Akhun, İ. (1984). İki korelasyon katsayısı arasındaki farkın manidarlığının test edilmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 17(1), 1-7. doi: 10.1501/Egifak_0000001034
- Alpar, R. (2010). *Uygulamalı istatistik ve geçerlik, güvenirlik*. Ankara: Detay Yayıncılık.
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders*. Arlington: American Psychiatric Publishing.
- Atılğan, H., Kan, A., ve Doğan, N. (2011). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Azad, G. F., Reisinger, E., Xie, M., & Mandell, D. S. (2016). Parent and teacher concordance on the Social Responsiveness Scale for children with autism. *School Mental Health*, 8, 368-376. doi: 10.1007/s12310-015-9168-6
- Bakkaloğlu, H., Ökcün-Akçamuş, M. Ç., Demir, Ş., Bahap-Kudret, Z., ve Yalçın, S., (2019). *Sosyal Yanıtlayıcılık Ölçeği-2-Okulöncesi Türkçe Formu'nun Türk kültürüne uyarlanması: Geçerlik ve güvenirlik çalışması*. Yayınlanmamış araştırma raporu.
- Bates, B. T., Zhang, S., Dufek, J. S., & Chen, F. C. (1996). The effects of sample size and variability on the correlation coefficient. *Medicine and Science in Sports and Exercise*, 28(3), 386-391. doi: 10.1097/00005768-199603000-00015
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders*, 30(3), 237-243. doi: 10.1023/a:1005596502855
- Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı* (18. baskı). Ankara: Pegem Akademi.
- Can, A. (2016). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Ankara: Pegem Akademi.
- Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., & Brophy, S. L. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview- revised. *Journal of Autism and Developmental Disorders*, 33(4), 427-433. doi: 10.1023/a:1025014929212

- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale-second edition (SRS-2)*. Torrance: Western Psychological Services.
- Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*, 2-11. doi: 10.1097/00004703-200002000-00002
- Crocker, L. M., & Algina, L. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Çıkrıkçı, N. (2019). Ölçmede güvenirlik. R. N., Çıkrıkçı (Ed.), *Eğitimde ölçme ve değerlendirme içinde (77-100)*. Ankara: Anı Yayıncılık.
- Demir, Ş., Ökcün-Akçamuş, M. Ç., Bahap Kudret, Z., & Bakkaloğlu, H. (2018, 11-13 Ekim). *OSB olan çocuklar için öğretmenler tarafından doldurulan Tekrarlayıcı Davranışlar Ölçeği-R-Türkçe Versiyonunun psikometrik özelliklerinin incelenmesi (Sözlü Bildiri)*. 28. Ulusal Özel Eğitim Kongresi, Eskişehir Osmangazi Üniversitesi, Eskişehir, Türkiye
- Diken, İ. H., Ardıç, A., Diken, Ö., ve Gilliam, J. E. (2012). Gilliam Otistik Bozukluk Derecelendirme Ölçeği-2-Türkçe Versiyonu'nun (GOBDÖ-2-TV) geçerlik ve güvenirliğinin araştırılması: Türkiye standardizasyon çalışması. *Eğitim ve Bilim, 37*(166), 318-328.
- Dinnebeil, L. A., Sawyer, B. E., Logan, J., Dynia, J. M., Cancio, E., & Justice, L. M. (2013). Influences on the congruence between parents' and teachers' ratings of young children's social skills and problem behaviors. *Early Childhood Research Quarterly, 28*(1), 144-152. <https://doi.org/10.1111/j.1469-7610.1989.tb00772.x>
- Gilliam, J. E. (2006). *Gilliam autism rating scale*. Austin, TX.: PRO-ED.
- Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Psychological Education and Exercises Science, 5*(1), 13-34. doi: 10.1207/S15327841MPEE0501_2
- Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010). Cross-informant agreement for ratings for social skill and problem behavior ratings: An investigation of the Social Skills Improvement System Rating Scales. *Psychological Assessment, 22*(1), 157.

- Güler, N. (2008). *Klasik test kuramı genellenabilirlik kuramı ve rasch modeli üzerine bir araştırma* (Doktora Tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara. <https://tez.yok.gov.tr/UlusalTezMerkezi/> adresinden alınmıştır.
- Hundert, J., Morrison, L., Mahoney, W., Mundy, F., & Vernon, M. L. (1997). Parent and teacher assessments of the developmental status of children with severe, mild/moderate, or no developmental disabilities. *Topics in Early Childhood Special Education*, 17, 419-434. doi: 10.1177/027112149701700404
- İnan, B., Ökcün-Akçamuş, M. Ç., Bakkaloğlu, H., ve Yalçın, S. (2020). Turkish adaptation and validity-reliability study of the revised Social Communication Checklist for children with autism spectrum disorder. *Journal of Early Childhood Studies*, 4(4), 172-196. doi: [10.24130/eccd-jecs.1967202041193](https://doi.org/10.24130/eccd-jecs.1967202041193)
- Jepsen, M. I., Gray, K. M., & Taffe, J. R. (2012). Agreement in multiinformant assessment of behaviour and emotional problems and social functioning in adolescents with Autistic and Asperger's Disorder. *Research in Autism Spectrum Disorders*, 6, 1091-1098. doi: 10.1016/j.rasd.2012.02.008
- Joosten, A. V., Bundy, A. C., & Einfeld, S. L. (2012). Context influences the motivation for stereotypic and repetitive behaviour in children diagnosed with intellectual disability with and without autism. *Journal of Applied Research in Intellectual Disabilities*, 25(3), 262-271. doi: 10.1111/j.1468-3148.2011.00663.x
- Kaat, A. J., Gadow, K. D., & Lecavalier, L. (2013). Psychiatric symptom impairment in children with autism spectrum disorders. *Journal of Abnormal Child Psychology*, 41, 959-969. doi: 10.1007/s10802-013-9739-7
- Kalaycı, Ş. (2010). *SPSS uygulamalı çok değişkenli istatistik teknikleri* (5. baskı). Ankara: Asil Yayın Dağıtım.
- Karasar, N. (2000). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın ve Dağıtım.
- Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2017). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 21(2), 142-154. doi: 10.1177/1362361316632710
- Lampi, A., Fitzpatrick, P., Romero, V., Amaral, J., & Schmidt, R. C. (2018). Understanding the influence of social and motor context on the co-occurring frequency of restricted and repetitive behaviors in autism. *Journal of Autism and Developmental Disorders*, 50, 1479-1496. doi: [10.1007/s10803-018-3698-3](https://doi.org/10.1007/s10803-018-3698-3)

- S.Yalçın, H.Bakkaloğlu, M.C.Ö. Akçamuş, Ş.Demir ve Z.B.Kudret / *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 52, 360-388, 2021
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison Wesley, Reading, MA.
- Mirenda, P., Smith, I., Vaillancourt, T., Georgiades, S., Duku, E., Szatmari, P., et al. (2010). Validating the repetitive behavior scale-revised in young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 40(12), 1521-1530. doi: 10.1007/s10803-010-1012-0
- Murray, D. S., Ruble, L. A., Willis, H., & Molloy, C. A. (2009). Parent and teacher report of social skills in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, 40(2), 109-115. doi: 10.1044/0161-1461
- Ökcün-Akçamuş, M., Bakkaloğlu, H., Demir, Ş., ve Bahap-Kudret, Z. (2019). Otizm spektrum bozukluğunda Tekrarlayıcı Davranışlar Ölçeği-Revize Türkçe Sürümünün geçerlilik ve güvenilirlik çalışması. *Anadolu Psikiyatri Dergisi*, 20(Ek sayı1.1), 65-72. doi: 10.5455/apd.42649
- Özgüven, İ. E. (2007). *Psikolojik testler*. Ankara: PDREM Yayınları.
- Pekin, Z. (2015). *Otizm sosyal beceriler profili ölçeğinde puanlayıcılar arası güvenirlığın klasik test kuramı ve genellenebilirlik kuramına göre karşılaştırılması* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Raver, S. A. (2009). *Early childhood-special education-0-8 years: Strategies for positive outcomes*. Pearson Education, Inc.
- Ryland, H. K., Hysing, M., Posserud, M. B., Gillberg, C., & Lundervold, A. J. (2012). Autism spectrum symptoms in children with neurological disorders. *Child and Adolescent Psychiatry and Mental Health*, 6(34), 1-9. doi: 10.1186/1753-2000-6-34
- Tasse, M., & Lecavalier, L. (2000). Comparing parent and teacher ratings of social competence and problem behaviors. *American Journal on Mental Retardation*, 105, 252-259. doi: 10.1352/0895-8017
- Tucker, V., & Schwartz, I. (2013). Parents' perspectives of collaboration with school professionals: Barriers and facilitators to successful partnerships in planning for students with ASD. *School Mental Health*, 5(1), 3-14. doi: 10.1007/s12310-012-9102-0

- Turner, M. (1999). Annotation: Repetitive behaviour in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry*, 40(6), 839-849. doi: [10.1111/1469-7610.00502](https://doi.org/10.1111/1469-7610.00502)
- Vickerstaff, S., Heriot, S., Wong, M., Lopes, A., & Dossetor, D. (2007). Intellectual ability, self-perceived social competence, and depressive symptomatology in children with high-functioning autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 37, 1647-1664. doi: 10.1007/s10803-006-0292-x.
- Verhulst, F. C., & Akkerhuis, G. W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4–12. *Journal of Child Psychology and Psychiatry*, 30(1), 123-136. <https://doi.org/10.1111/j.1469-7610.1989.tb00772.x>
- Voelker, S. L., Shore, D. L., Lee, C. H., & Szuskiewicz, T. A. (2000). Congruence in parent and teacher ratings of adaptive behavior of low-functioning children. *Journal of Developmental and Physical Disabilities*, 12(4), 367-376. doi: 10.1023/A:1009436230984
- Wainer, A. L., Berger, N. I., & Ingersoll, B. R. (2017). Brief report: The preliminary psychometric properties of the Social Communication Checklist. *Journal of Autism and Developmental Disorders*, 47(4), 1231-1238. doi: 10.1007/s10803-016-3026-8



Comparison of the Consistency between Parents and Teachers in Scales Used in Autism Spectrum Disorder According to Different Techniques

Seher YALÇIN*, Hatice BAKKALOĞLU**, Meral Çilem ÖKCÜN AKÇAMUŞ***, Seyda DEMİR****, Zeynep BAHAP KUDRET*****

• **Received:** 31.05.2020 • **Accepted:** 04.01.2021 • **Online First:** 04.01.2021

Abstract

In this study, it was aimed to compare the consistency between the scores of both the parents and teachers of the same children according to different techniques for the four scales commonly used in the literature in measuring the degree of the disorder in children with autism spectrum disorder (ASD), and in the evaluation of repetitive behaviors and social communication-interaction behaviors, which are the leading indicators of the disorder. Participants of this study, which is a relational survey model, are the individuals diagnosed with ASD, between the ages of 3-23 (and N between 300 and 380), and attending schools and rehabilitation centers affiliated with the Ministry of National Education in the 2018-2019 academic year, as well as their parents (N between 292 and 379) and teachers (N between 298 and 378). Four different scales were used to evaluate the behavioral and developmental characteristics of children with ASD in the study. As a result of the research, it was seen that the scores obtained from parents and teachers, who completed different scales, were high in internal consistency within groups. Moreover, in some subscales, the difference between the reliability coefficients regarding the scores obtained from parents and teachers was statistically significant in favor of teachers. It was determined that the correlations regarding the scores obtained from the parents and teachers differed according to the subscales (low, medium, high level, and non-significant relationship). In this context, it was also found that it is not appropriate to interpret the consistency between parents and teachers based only on correlations.

Keywords: scales used in autism spectrum disorder, the consistency between parents and teachers

Cited:

Yalçın, S., Bakkaloğlu, H., Akçamuş, M.Ç.O., Demir, Ş., & Kudret, Z.B. (2021). Comparison of the consistency between parents and teachers in scales used in autism spectrum disorder according to different techniques. *Pamukkale University Journal of Education*, 52, 360-388. doi:10.9779.pauefd.745749.

* Assoc. Prof. Dr., Ankara University, Faculty of Educational Sciences, ORCID ID: 0000-0003-0177-6727
yalcins@ankara.edu.tr

** Prof. Dr., Ankara University, Faculty of Educational Sciences, ORCID ID: 0000-0002-3226-9077,
hbakkaloglu@ankara.edu.tr

*** Assoc. Prof. Dr., Ankara University, Faculty of Educational Sciences, ORCID ID: 0000-0003-3297-9711,
meralcilem@gmail.com

**** Asst. Prof. Dr., Ankara University, Faculty of Educational Sciences, ORCID ID: 0000-0002-7344-7853,
demir.seyda@yahoo.com

***** Asst. Prof. Dr., Ankara University, Faculty of Educational Sciences, ORCID ID: 0000-0001-5192-0892,
zeynepbahap@gmail.com

Introduction

One of the most important purposes of measurement and assessment in the education and training process is to determine to what extent individuals have the attribute addressed and ensure that individuals' decisions are accurate according to this determination. This determination is often made by using a measurement tool. Therefore, the validity and reliability of the scores obtained from the measurement tools used are important. In education and psychology, a certain amount of error is involved in every measurement process since the characteristics addressed cannot be directly measured. The basic assumption of the classical test theory (CTT) is that the individual's scores obtained from a measurement process are theoretically composed of true scores and error scores (Lord & Novick, 1968). In other words, the error is the difference between the true value of the measured feature and the observed value obtained as a result of the measurement (Atılğan, Kan, & Doğan, 2011). Errors may arise from the individuals measured, the people who perform the measurement, the measurement tool, and the environment in which the measurement is made (Baykul 2000; Crocker and Algina 1986).

In CTT, there are different approaches to calculating reliability. These vary according to the meaning of the researcher's reliability and the source of error (Baykul, 2000). The test-retest method is used for the reliability obtained by applying a test twice within a certain time interval. The obtained reliability is considered in terms of "stability" or "consistency." In this reliability estimation, the time between two applications is considered to be a source of error. The consistency of the items in the measurement tool with the whole scale's scores expresses the "internal consistency" reliability. In this reliability, the items that make up the measurement tool are considered as the source of error. The "consistency" of the scores obtained as a result of the application of two parallel forms with a short time interval is called the equivalent (parallel) forms method. In this method, the measuring tool is seen as the source of error. 'Inter-rater reliability' is called the consistency between the scores given by more than one rater to the same feature, and the raters are considered as a source of error in this method (Goodwin, 2001; Güler, 2008).

One of the areas where reliability between raters is very important is undoubtedly special education. One of the frequently used methods to assess the different developmental and behavioral characteristics of individuals with special needs (Murray, Ruble, Wills, & Molloy, 2009) is the parent and teacher report scales based on family and teacher ratings.

These assessment tools are scales and checklists filled by individuals who often share the same environment with the target individual. Information about the behaviors of individuals with special needs that occur or do not occur in different environments is obtained with scales and checklists of which validity and reliability are determined, and teachers and families fill that. By using these scales, decisions about individuals are often made based on information obtained from different raters during screening, diagnosis, and intervention processes. Therefore, the validity and reliability of individuals' decisions depend on the raters (Dinnebeil et al., 2013; Verhulst & Akkerhuis, 1989). Children and individuals with autism spectrum disorder (ASD), a neurodevelopmental disorder characterized by limitations in social communication and interaction skills, repetitive behaviors, and limited interests (APA, 2013), are among the groups that parent and teacher report scales are used frequently. In the literature, it is recommended to collect data from multiple sources in assessment for screening, diagnosing, and intervention planning for children and individuals with special needs, including children and individuals with ASD as well (Raver, 2009; Vickerstaff, Heriot, Wong, Lopes, & Dossetor, 2007; Voelker, Shore, Lee, & Szuszkiewicz, 2000), it is also suggested to use multiple assessment tools (Raver, 2009; Voelker et al., 2000) and multiple assessment methods (Raver, 2009). Consistency between people who know the child or individual well (such as parents and teachers); affects many factors such as the choice of the services and supports by the intervention team, the consistency of the practices between home and school, and the satisfaction of the family with the interventions (Dinnebeil et al., 2013; Gresham, Elliott, Cook, Vance, & Kettler, 2010; Tucker & Schwartz, 2013). Therefore, considering the consequences of inconsistency between the information obtained from different data collection sources, the consistency between people who know the child well is seen as a critical area of study (Kaat, Gadow, & Lecavalier, 2013).

In the literature, there are inconsistent findings regarding the consistency levels of individuals who provide information about the same child or individual. While some of these studies have found a strong consistency between the scales completed by parents and teachers (e.g., Constantino, Przybeck, Friesen, & Todd 2000; Constantino et al., 2003); in some, it was found that there is a low level of correlation (Murray et al., 2009) or no significant correlation (Vickerstaff et al., 2007) between the scale scores filled in by parents and teachers for the same children. For example, Constantino et al. (2000, 2003), using the Social Responsiveness Scale, found that there was a strong consistency between parents and

teachers regarding the social behaviors of children with ASD, and this consistency was consistent over a two-year follow-up. On the other hand, Kaat et al. (2013) found a low level of agreement between parents and teachers in terms of disorders and symptoms related to ASD. Vickerstaff et al. (2007) found no significant correlation between teachers and parents of children with ASD.

It is known that impairments in social functions are an important indicator of the disorder in children with ASD. However, social functioning is difficult to define and measure, and depending on the setting and context, differences in desired behavior at home or school can be seen. Therefore, it is not surprising that parents and teachers have disagreements about the quality of social functioning (Jepsen et al., 2012; Tasse & Lecavalier, 2000). It is thought that this situation may be valid not only for behaviors related to social functioning but also for repetitive behaviors, another indicator of ASD. Because the frequency and variety of repetitive behaviors can also differ according to different contexts (Joosten, Bundy, & Einfeld, 2012; Kirby, Boyd, Williams, Faldowski, & Baranek, 2017; Lampi, Fitzpatrick, Romero, Amaral, & Schmidt, 2018; Turner, 1999), in addition to the context in which the behavior occurs, it is thought that the perspectives of the raters, as well as the differences in perceptions regarding and attitudes towards the related behavior, may cause differences among the raters (Azad, Reisinger, Xie, & Mandell, 2016; Ryland, Hysing, Posserud, Gillberg, & Lundervold, 2012). Such inconsistencies that arise among raters cause a lot of time to be spent on agreement and cooperation between experts in interpreting conflicting data collected during the evaluation of children and identifying intervention goals (Azad et al., 2016). Therefore, ensuring inter-rater reliability is crucial.

Inter-rater reliability refers to the coherence and consistency between two or more raters for the same task/person. There are different methods and approaches to determine the reliability between raters. The most frequently used of these are correlation techniques, comparing the means (t-test /variance analysis), Cohen's kappa (κ), Weighted kappa (κ_w), Fleiss' kappa, Kendall coefficient of concordance (W), and compliance percentage. For the situations where researchers should use these methods and approaches, the choice must depend on the measurement error that the researcher wants to avoid and the type of scale from which the measurements are obtained (nominal, ordinal, interval, or ratio) (Goodwin, 2001).

In special education, different decisions can be made in the processes of determining the developmental characteristics of individuals and planning the intervention based on the data obtained from different raters such as teachers or parents. The validity and reliability of the decisions made about individuals in these processes depend on the raters. Therefore, considering the consequences of inconsistency between the information obtained from different data collection sources, the consistency between people who know the child well is seen as a critical field of study. In this study, it was aimed to compare the consistency between the scores obtained from both parents and teachers of the same children for four scales frequently used in the literature to measure the degree of the disorder, to assess the repetitive behaviors, and social communication-interaction behaviors among the indicators of the disorder. These scales used in individuals with ASD are as follows i) Social Responsiveness Scale-2-Preschool Turkish Form [SRS-2-PTF], ii) Gilliam Autism Rating Scale-2-Turkish Version (GARS-2-TV), iii) Social Communication Checklist-Revised-Preschool Turkish Version (SCC-R-PTV) and iv) Repetitive Behavior Scale-Revised-Turkish Version (RBS-R-TV). Correlation techniques, comparison of means, and Cohen's Kappa technique were used to determine raters' agreement. In this study, answers were sought for the following questions regarding these scales (SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV) for the parents and teachers who scored the behavioral and developmental characteristics of children with ASD:

1. What are the internal consistency coefficients for parents and teachers based on subscales' scores and the whole scale scores?
2. How is the consistency based on the correlation between parents' and teachers' scores, according to parents when they are considered a single group, and according to the two groups separated by education level (undergraduate and higher and lower)?
3. How is the consistency between parents' and teachers' scores according to the comparison of means technique?
4. What is the consistency in classifying children as typically developing and children with ASD according to Cohen's kappa technique by the parents and teachers who scored the SRS-2-PTF and GARS-2-TV?

Method

Research Model

This study was conducted in the correlational survey model (Karasar, 2000) since it was aimed to identify the consistency between the ratings of parents and teachers for the same children with ASD based on scores of different scales that assess the behavioral and developmental characteristics of children with ASD utilizing different techniques, in other words since it was aimed to determine the level of coherence between parent and teacher scores.

Participants

Participants of this study were determined using a convenience sampling method. Participants of the study consisted of parents and teachers of children who attended schools and rehabilitation centers affiliated with the Ministry of National Education in Ankara during the 2018-2019 academic year. Participants consist of 330 parents and 378 teachers of 380 children with ASD between the ages of 2-7 for SRS-2-PTF (Bakkaloğlu, Kudret, Ökcün Akçamuş, Yalçın, & Demir, 2020); 379 parents and 329 teachers of 380 individuals with ASD between the ages of 3-23 for GARS-2-TV (Bakkaloğlu et al., 2020); 315 parents and 361 teachers of 365 children with ASD between the ages of 2-7 for SCC-R-PTV (İnan, Ökcün-Akçamuş, Bakkaloğlu, & Yalçın, 2020) and 292 parents and 298 teachers of 300 individuals with ASD between the ages 3-23 for RBS-R-TV (Ökcün-Akçamuş, Bakkaloğlu, Demir and Bahap-Kudret, 2019).

Descriptive statistics of parents and teachers on demographic variables such as gender, age, and educational level were also collected and presented subsequently. The parents' mean age is 35.49 (sd = 6.04) for SRS-2-PTF and GARS-2-TV, and the lowest value is 21, and the highest value is 61. When the distribution of teachers according to gender was examined, it was seen that the participants consisted of 273 women (71.84%) and 96 men (25.26%). For SCC-R-PTV, the mean age of the parents is 35.40 (sd = 6.01), and the lowest value is 21, and the highest value is 61. When the educational levels of the parents who answered the SRS-2-PTF, GARS-2-TV, and SCC-R-PTV scales were examined, 38 of the participants (13.4%) graduated from primary school, 34 of them (11.9%) graduated from secondary school, 91 of them (31.9%) graduated from high school,

93 of them (32.6%) had a bachelor's degree level, and 11 of them (3.9%) had a postgraduate degree. According to professions, when the distribution of parents was examined, it was seen that 137 of them, that is approximately half of the sample (48.07%), were housewives, and then 17 parents (5.96%) were teachers. Parents from many other professions were involved in the study. When the distribution of teachers by gender for SCC-R-PTV was examined, it was seen that 209 of the teachers were women (57.26%) and 70 were men (19.18%), and some teachers' knowledge on gender could not be stated because of missing information. The mean of age of parents for RBS-R-TV is 39.52 (sd = 6.69), and the lowest value is 24, and the highest value is 66. When the educational level of the parents for RBS-R-TV was examined, it was seen that 40 of the participants (17.86%) graduated from primary school, 14 of them (6.25%) graduated from secondary school, 61 of them (27.23%) graduated from high school, 82 of them (36.61%) had a bachelor's degree, and 11 of them (4.91%) had a postgraduate degree. A total of 16 parents (7.14%) did not indicate their educational level. When the distribution of parents for RBS-R-TV according to professions was examined, it was seen that 118 of them (52.68%) were housewives, and 22 parents (9.82%) were teachers. There were parents from many other professions. For RBS-R-TV, the mean of teachers' age is 31.36 (sd = 8.41), and the lowest value is 22, and the highest value is 63.

Data Collection Tools

Social Responsiveness Scale-2- Preschool Turkish Form (SRS-2-PTF)

It was developed by Constantino and Gruber (2012) to assess the symptoms related to ASD. It consists of 65 items, including five subscales (social awareness [8 items], social cognition [12 items], social communication [22 items], social motivation [1 item], and restricted interests and repetitive behaviors [12 items]). SRS-2-PTF is a scale suitable for ASD diagnosis criteria (social interaction and communication disorder, and restricted interest/repetitive behaviors) in DSM-5. The scale items are scored with a four-point Likert scale (0: not true, 1: sometimes true, 2: often true, 3: almost always true). Seventeen items in the scale (3, 7, 11, 12, 15, 17, 21, 22, 26, 32, 38, 40, 43, 45, 48, 52, 55) are scored in reverse, and scores between 65 and 260 can be obtained from the scale. Higher scores on the scale indicate a higher ASD level. The adaptation of the scale to Turkish was conducted by Bakkaloğlu et al. (2020). The fit values of the five-factor model are based on parental responses ($X^2 = 5302.684$, $df = 2005$, $p < .05$, $X^2/df = 2.64$, $RMSEA = 0.071$, $CFI = 0.78$)

and teacher responses ($X^2 = 6157.422$, $df = 2005$, $p < .05$, $X^2/df = 3.07$, $RMSEA = 0.074$, $CFI = 0.813$) showing that the models fit the data at an acceptable level. The internal consistency coefficients of the scale are .95 and .96 for parents and teachers, respectively. The subscale reliability coefficients range from .59 to .87 for the parent and from .69 to .90 for the teacher (Bakkaloğlu et al., 2020).

Gilliam Autism Rating Scale-2-Türkçe Version (GARS-2-TV)

Developed by Gilliam (2006), this scale is used to assess individuals between the ages of 3-23 regarding ASD symptoms or diagnostic features. The scale consists of three subscales: stereotyped behaviors (14 items), social interaction (14 items), and communication (14 items). The items in the scale are scored with four-point ratings (0: never observed, 1: seldom observed, 2: sometimes observed, 3: frequently observed), and an Autism Index (AI) score is obtained by converting the total scores into standard scores. If this score is 85 and above, having ASD is very likely; the average between 70-84 points indicates that the probability of having ASD is possibly and 69 points and below indicates that the probability of having ASD is unlikely. Adaptation and validity-reliability studies of GARS-2 into Turkish were conducted by Diken, Ardıç, Diken, and Gilliam (2012). In this adaptation study, validity and reliability analyses were made using all of the data collected from 1191 participants (40.7% parents, 59.3% teachers/experts), and separate analyses were not made for teachers/experts and parents. Confirmatory factor analysis' results of the Turkish version of the scale show that the model fit for the factor structure is at a good level ($X^2 = 1730.08$, $df = 813$, $X^2/df = 2.13$, $CFI = .89$, $RMSEA = 0.071$). The internal consistency coefficients vary between .77 and .85 (Diken et al., 2012).

Social Communication Checklist-Revised-Preschool-Turkish Version (SCC-R-PTV)

This scale was developed by Ingersoll and Dvortcsak (2010) to assess the social communication and interaction skills of children with ASD. The scale includes a total of 70 items; 15 for social participation, 30 for expressive language, 8 for receptive language, and 17 for imitation and play skills. The items in the checklist are scored as "not observed (0 points); rarely or not yet (1 point), sometimes but not consistently (2 points); and usually, at least 75% of the time (3 points)," and scores between 70 and 210 can be obtained. The Turkish adaptation of the checklist was performed by İnan et al. (2020), and as a result of the confirmatory factor analysis, the goodness of fit values of the checklist indicate good fit in both the parent form ($X^2 = 5121.967$, $df = 2339$, $p < .05$, $X^2/df = 2.19$, $RMSEA: .068$; $CFI: .916$; $TLI: .913$) and teacher form ($X^2 = 5226.578$, $df = 2339$, $p < .05$, $X^2/df = 2.23$, $RMSEA:$

.069; CFI: .954, TLI: .953), and it was concluded that the original structure of SCC-R-PTV with four factors was confirmed in Turkish culture. Cronbach's alpha internal consistency coefficient for SCC-R-PTV total scores is .98 and .99 for parents and teachers, respectively. In the checklist's reliability analysis, the internal consistency reliability coefficients for all subtests in both parent and teacher forms are over 80 (İnan et al., 2020).

Repetitive Behavior Scale-Revised-Turkish Version (RBS-R-TV)

The scale is a clinical rating scale developed experimentally by Bodfish, Symons, Parker, and Lewis (2000) to assess repetitive behaviors and the severity of these behaviors. RBS-R-TV consists of six subscales and 43 items: stereotyped behaviors (6 items), self-injurious behaviors (8 items), compulsive behaviors (8 items), ritualistic behaviors (6 items), sameness behaviors (11 items), and restricted behaviors (4 items). The scale score items are scored using the four-point scale (0: behavior does not occur, 1: mild level, 2: moderate level, 3: severe level). It is accepted that as the score obtained from the scale increases, the severity of repetitive behaviors also increases. The Turkish adaptation study of the scale was carried out separately for scales filled in by both teachers and parents. The validity and reliability study for the scales filled by the parents was carried out by Ökcün-Akçamuş et al. (2019). As a result of this study, it was determined that the six-factor structure of the Turkish form of the scale was compatible with the factor structure of the original scale. The results of the confirmatory factor analysis showed that the goodness of fit values of the six-factor scale structure were at a good level ($X^2 = 16.7912$, $df = 840$, $p < .05$, $X^2/df = 2.00$, $RMSEA = .059$, $SRMR = .$, $NNFI = .94$, $CFI = .94$). When the reliability analysis of the RBS-R-TV scale was examined, it was found that the Cronbach alpha values for all subscales were .73 and above, and the Cronbach alpha coefficient for the total score was .94. The validity and reliability study for the scales filled by teachers was carried out by Demir et al. (2018). In this study, the fit indices of the six-factor scale structure were found to be at a good level ($X^2 = 2415.92$, $df = 838$, $p < .05$, $X^2/df = 2.88$, $RMSEA = .079$, $SRMR = .064$, $NNFI = .93$, $CFI = .94$), Cronbach alpha values for all subscales were found to be .77 and above, and this value was .93 for the total score of the scale. Therefore, as a result of the analysis, it was determined that RBS-R-TV is a valid and reliable tool (Ökcün-Akçamuş et al., 2019).

Data Analysis

Within the scope of the first research question of the study, to identify the internal consistency regarding the scoring of parents and teachers on different scales (SRS-2-PTF,

GARS-2-TV, SCC-R-PTV, and RBS-R-TV) that assess the behavioral and developmental characteristics of children with ASD, Cronbach alpha internal consistency coefficients were calculated based on the scores obtained from the subscales and the whole scale. Cronbach's alpha internal consistency coefficient between 0.00 and 0.40 indicates that a scale is not reliable, between 0.41 and 0.60 indicates low reliability, between 0.61 and 0.80 indicates moderate reliability. A range between 0.81 and 1.00 indicates high reliability (Alpar, 2010). To determine the significance of the difference between the internal consistency coefficients for the scores of the two groups, the reliability coefficients were converted into Fischer Zr scores. The significance of the difference between the coefficients converted into Fischer's Zr value was tested with the help of equation (1) (Akhun, 1984).

$$Z = \frac{Z_{r_1} - Z_{r_2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

In the equation (1), "Z" is the Z statistic value belonging to the differences in reliability coefficients, " Z_{r_1} " is the converted value of the reliability coefficient for the scores of the parents to Fischer Zr value, " Z_{r_2} " is the converted value of the reliability coefficient for the teachers' scores to Fischer Zr value, " n_1 " is the sample size of the parents, " n_2 " shows the sample size of the teachers.

Within the scope of the second research question, it was aimed to determine the consistency between the scores of parents and teachers regarding the scales based on correlation. For this purpose, before examining the relationship between parents' and teachers' scores on the scales, the condition of providing the normal distribution assumption of the scores obtained from the scales was examined. The skewness and kurtosis coefficients for the subscales for parent and teacher responses of the SRS-2-PTF and SCC-R-PTV and the total score of these scales were examined, and it was found that these values were between -1 and +1. The skewness and kurtosis coefficients of GARS-2-TV, which the teachers answered, were examined for the subscales and the total scale score, and it was found that these values were also between -1 and +1. The skewness and kurtosis coefficients of parental responses for the subscales of GARS-2-TV and the total scale score were examined, and the skewness coefficients were found to be +2 for the communication subscale, while the kurtosis coefficient was 4.747 in the stereotyped behaviors subscale,

6.914 in the communication and 9.072 in the social interaction. Therefore, the mean, median, and mode values of GARS-2-TV were also examined for parental responses, and these values were found to be very close to each other. When the histogram of the distribution was examined, it was concluded that the normal distribution assumption was met. Since the skewness and kurtosis coefficients are between the limits of -3 and +3, and the mean, median, and mode values are close to each other, it was concluded that the normal distribution assumption was achieved (Kalaycı, 2010). The skewness and kurtosis coefficients of the subscales' scores and the total score of the scale for the parent and teacher responses of RBS-R-TV were examined, and it was seen that the skewness coefficients were between -2 and +2 for both forms. The kurtosis coefficient was 4.010 for the self-injurious behaviors subscale in the parental form; in the teacher form, the kurtosis coefficients were 3.537 for the self-injurious behaviors subscale, 3.425 for the ritualistic behaviors subscale, and all the values in the other subscales were between -2 and +2. The mean score of the self-injurious behaviors subscale for the teacher form was 2.79, the median was 1, and the mode was 0. The mean, median, and mode of the self-injurious behaviors subscale were 2.13, 1, and 0, respectively, for the parent form. The mean of the ritualistic behaviors subscale was 3.81, the median was 3, and the mode was 0 for the teacher form. As a result of these analyses, doubts were raised about the assumption of normality of the self-injurious behaviors and ritualistic behaviors subscales, and both the paired-samples t-test and the Wilcoxon signed-ranks test were used for all scale scores and subscale scores of the RBS-R-TV. In the scales that provided the normality assumption, when parents were considered as a single group and divided into two groups according to their educational level (undergraduate and higher, and below undergraduate), the relationship between parents' and teachers' scores was analyzed by the Pearson product-moment correlation (PPMC) coefficient. The Spearman-Brown formula was calculated due to the failure to provide the assumption of normality in RBS-R-TV. As an absolute value, the correlation coefficient between 0.00 and 0.30 indicates a low level of correlation; a value between 0.30 and 0.70 indicates a moderate correlation, and a value between 0.70 and 1.00 indicates a high level of correlation (Büyüköztürk, 2013).

Within the scope of the third research question, whether there is a meaningful difference between the mean of the scores obtained from the parents and teachers was compared with the paired-samples t-test since the assumption of normality was provided. If there is a significant difference between raters in terms of rigidity or generosity in scoring,

the PPMC coefficient can be highly positive. The high correlation indicates that the scores vary together, but it is recommended not to consider the PPMC coefficient alone, but to compare the means (Goodwin, 2001).

Within the fourth research question's scope, the consistency between parent and teacher ratings was calculated by Cohen's kappa coefficient. Cohen's kappa statistics is a technique in which the compatibility between two sets of categorical variables (only when between the ratings of two independent raters) is questioned (Can, 2016). A Kappa coefficient less than .20 indicates poor fit; between .21 and .40 indicates a low level of agreement, between .41 and .60, indicates a moderate level of agreement, between .61 and .80, indicates a good level of agreement, and between .81 and 1 indicates a very good level of agreement (Can, 2016). The total score in the SRS-2-PTF was converted into T score as stated in the original of the scale, and the T scores were converted into two categories as those with 59 and below (typically developing children) and 60 and above (mild, moderate, and high-risk children (Constantino & Gruber, 2012)). According to the ADI scores obtained from GARS-2-TV, the children participating in the study were divided into two categories as having a moderate or high risk for ASD and mild risk for ASD. Since there were no cut-off points for SCC-R-PTV and RBS-R-TV, the inter-rater agreement for those scales could not be calculated according to Cohen's kappa statistics.

Results

Consistency between Parent and Teacher Scores According to Internal Consistency Coefficient

Cronbach alpha coefficients, calculated to determine the internal consistency coefficients based on subscales and the whole scale for parents and teachers who score SRS-2-PTF, GARS-2-TV, SCC-R-PTV and RBS-R-TV scales evaluating the behavioral and developmental characteristics of children with ASD are presented in Table 1.

According to Table 1, the internal consistency coefficients of SRS-2-PTF according to parent responses range from .610 to .950, and according to teachers' responses, the internal consistency coefficients range from .692 to .962. The internal consistency coefficients of GARS-2-TV according to parent responses range from .832 to .941, and according to teachers' responses, internal consistency coefficients range from .846 to .942. The internal consistency coefficients of SCC-R-PTV according to parent responses range from .930 to .981 and according to teacher responses, internal consistency coefficients range from .929 to .982; RBS-R-TV's internal consistency coefficients according to parent

responses range from .686 to .929 and according to teachers' responses internal consistency coefficients range from .775 to .938. To test whether these differences are significant or not, the significance of the difference between the coefficients was tested by making Fisher Z transform the internal consistency coefficients.

Table 1. Cronbach's alpha coefficients for parent and teacher scoring of SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales

Scale	Subscale	Parent		Teacher		Z
		N	α	N	α	
SRS-2-PTV	SA	313	.610	355	.692	-1,845
	SC	307	.724	346	.804	-2,502*
	SCOM	287	.881	330	.906	-1,607
	SM	323	.783	354	.824	4,368*
	RIRB	310	.809	349	.811	-0,075
	Total	258	.950	286	.962	-1,677
GARS-2-TV	C	316	.877	368	.882	-0,312
	SI	267	.832	323	.846	-0,610
	SB	309	.909	362	.917	-0,621
	Total	251	.941	310	.942	-0,105
SCC-R-PTV	SP	300	.946	354	.929	1,824
	EL	280	.966	336	.965	0,177
	RL	315	.938	361	.939	-0,108
	IP	288	.930	294	.936	-0,553
	Total	235	.981	241	.982	-0,304
RBS-R-TV	SB	283	.738	294	.791	-1,529
	SIB	281	.732	291	.795	-1,816
	CB	275	.772	293	.806	-1,082
	RB	287	.727	257	.783	-1,563
	SUB	276	.868	284	.870	-0,098
	RIB	287	.686	298	.775	-2,298*
	Total	248	.929	240	.938	-0,784

Note: SRS-2-PTF (Social Responsiveness Scale-2-Preschool Turkish Form): SA: Social Awareness, SC: Social Cognition, SCOM: Social Communication, SM: Social Motivation, RIRB: Restricted Interests and Repetitive Behaviors.

GARS-2-TV (Gilliam Autism Rating Scale-2-Turkish Version): C: Communication, SI: Social Interaction, SB: Stereotypical Behaviors.

SCC-R-PTV (Social Communication Checklist-Revised Preschool Turkish Version): SP: Social Participation, EL: Expressive Language, RL: Receptive Language, IP: Imitation / Play

RBS-R-TV (Repetitive Behaviors Scale-Revised-Turkish Version): SB: Stereotyped Behavior, SIB: Self-injurious Behavior, CB: Compulsive Behaviors, RB: Ritualistic Behavior, SUB: Sameness-Uniformity Behaviors, RIB: Restricted Interest Behavior.

* $p < .05$

As seen in Table 1, the differences between the reliability coefficients of the scores obtained from parents and teachers in the SC and SM subscales of the SRS-2-PTF and the

RIB subscale of the RBS-R-TV is significant ($p < .05$). It was determined that the reliability coefficients of the scores obtained from the teachers in all three subscales were significantly higher than the parents.

Consistency between Parents' and Teachers' Scores According to Correlation Coefficients

To determine the correlation between parents and teachers' scores on SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales that evaluate the behavioral and developmental characteristics of children with ASD, the correlation between all subscales score and scale total scores of parents and teachers were examined. The results are presented in Table 2. Also, in Table 2, when parents are divided into two groups according to their educational level (a) undergraduate and graduate b) high school and lower), the correlations between teachers' and parents' scores are also presented.

Table 2. *The correlations between teachers and parents' SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scale scores according to the educational level of the parents*

Scale	Subscale	Educational Level-1*		Educational Level -2*		Total	
		N	r	N	r	N	r
SRS-2-PTF	SA	139	-.013	96	.047	250	.056
	SC	131	.286**	89	.334**	236	.304**
	SCOM	113	.382**	82	.402**	207	.392**
	SM	133	.402**	96	.559**	245	.458**
	RIRB	133	.459**	89	.582**	237	.496**
	Total	163	.426**	104	.505**	261	.577**
GARS-2-TV	C	131	.437**	74	.552**	219	.489**
	SI	149	.965**	99	.955**	266	.954**
	SB	140	.376**	84	.403**	280	.563**
	Total	162	.754**	104	.775**	203	.821**
SCC-R-PTV	SP	162	.538**	104	.382**	283	.472**
	EL	162	.682**	104	.645**	284	.644**
	RL	161	.515**	103	.458**	282	.497**
	IP	163	.565**	104	.470**	285	.507**
	Total	163	.632**	104	.575**	285	.597**
RBS-R-TV	SB	122	.400**	77	.451**	211	.406**
	SIB	119	.477**	77	.638**	208	.574**
	CB	118	.440**	76	.552**	205	.434**
	RB	101	.291**	77	.589**	186	.354**
	SUB	109	.392**	78	.345**	196	.279**
	RIB	124	.498**	80	.483**	216	.416**
	Total	81	.503**	66	.485**	155	.397**

Note: SRS-2-PTF (Social Responsiveness Scale-2-Preschool Turkish Form): SA: Social Awareness, SC: Social Cognition, SCOM: Social Communication, SM: Social Motivation, RIRB: Restricted Interests and Repetitive Behaviors.

GARS-2-TV (Gilliam Autism Rating Scale-2-Turkish Version): C: Communication, SI: Social Interaction, SB: Stereotypical Behaviors.

SCC-R-PTV (Social Communication Checklist-Revised Preschool Turkish Version): SP: Social Participation, EL: Expressive Language, RL: Receptive Language, IP: Imitation / Play

RBS-R-TV (Repetitive Behaviors Scale-Revised-Turkish Version): SB: Stereotyped Behavior, SIB: Self-injurious Behavior, CB: Compulsive Behaviors, RB: Ritualistic Behavior, SUB: Sameness-Uniformity Behaviors, RIB: Restricted Interest Behavior.

*Educational level-1: High school and lower, Educational level-2: undergraduate and graduate

** $p < .01$

As seen in Table 2, when the parents are taken as a single group regardless of their educational level, and the correlations between parents and teachers in SRS-2-PTF are examined, it is observed that while there is no significant relationship in the Social Awareness subscale, the correlations in other dimensions ranged from .304 (social cognition) to .577 (SRS-2-PTF full scale). There are moderately significant relationships in these dimensions. When the correlations between the scores in GARS-2-TV obtained from parents and teachers are examined, it is seen that the correlation coefficients range from .489 (communication) to .954 (social interaction), and these values are moderately and highly significant. When the correlations between parents and teachers in SCC-R-PTV are examined, it is seen that the correlation coefficients range from .472 (social participation) to .644 (expressive language), and the relationships are moderately significant. It is seen that the correlation coefficients between the scores obtained from parents and teachers in the RBS-R-TV range from .279 (sameness-uniformity behaviors) to .574 (self-injurious behaviors), and these values are low and moderately significant.

When the findings are examined according to the educational levels of the parents (Table 2), significant correlation coefficients between .286 and .459 are obtained between the scores of the parents with high school and lower educational level as well as the scores of the teachers in the SRS-2-PTF except for SA subscale. It is observed that the correlations between the scores of parents with university or higher educational level and the scores of teachers are higher than the correlations between those with high school and lower educational level, and higher than the situation in all subscales, except the full-scale score, when there is no discrimination according to educational level. A similar finding is also valid for SB, SIB, and RB subscales of RBS-R-TV. In the CB subscale, it is determined that

the correlations between the scores of parents with a university or higher educational level and the scores of teachers are higher than the correlations between those with high school and lower educational levels. However, it is determined that this situation does not exist in all individuals. However, in GARS-2-TV, although parents are divided into groups according to educational level, it is seen that the correlations of both groups with the scores of the teachers are close to each other. On the contrary, to these findings in SCC-R PTV, it is seen that the correlations between the scores of parents with high school and lower educational level and the scores of the teachers are higher in all subscales and full-scale scores than in the case of no discrimination according to educational level.

Consistency between Parents' and Teachers' Scoring According to Comparison between Means

In order to examine the consistency between the parents and teachers who scored the SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales based on the comparison between the means, the Paired Samples t-test was conducted, and the obtained results are presented in Table 3.

Table 3. *t*-test results regarding the comparison of means of parents' and teachers' scoring the SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV

Scale	Subscale	Scorer	\bar{x}	N	sd	t	df	p
SRS-2-PTF	SA	Parent	12.26	250	4.77	-	249	.000
		Teacher	18.56	250	3.47	17.343		
	SC	Parent	19.19	236	6.33	-	235	.000
		Teacher	25.90	236	4.62	15.588		
	SCOM	Parent	31.11	207	12.77	-	206	.000
		Teacher	45.90	207	8.95	17.178		
	SM	Parent	16.42	245	6.69	-	244	.000
		Teacher	23.98	245	5.08	18.815		
	RIRB	Parent	13.89	237	6.83	-	236	.000
		Teacher	25.74	237	6.66	26.925		
	Total	Parent	92.21	285	33.76	-	284	.000
		Teacher	138.44	285	24.56	24.803		
GARS-2-TV	C	Parent	8.94	218	2.83	-	217	.000
		Teacher	14.10	218	8.46	10.156		
	SI	Parent	8.10	265	3.31	-	264	.000
		Teacher	15.24	265	9.97	17.000		
	SB	Parent	7.82	279	3.51	-	278	.000
		Teacher	12.34	279	8.89	10.096		
	Total	Parent	23.43	284	8.93	-	283	.000

		Teacher	37.97	284	21.18	15.894		
SCC-R-PTV	SP	Parent	42.80	283	13.30	.398	282	.691
		Teacher	42.48	283	12.84			
	EL	Parent	76.89	284	24.87	2.369	283	.018
		Teacher	73.83	284	26.51			
	RL	Parent	22.49	282	7.53	-1.059	281	.291
		Teacher	22.95	282	7.04			
	IP	Parent	43.93	285	13.32	1.482	284	.139
		Teacher	42.75	285	13.71			
	Total	Parent	185.47	285	53.13	1.225	284	.222
		Teacher	181.93	285	55.52			
RBS-R-TV	SB	Parent	4.33	211	3.52	-3.109	210	.002
		Teacher	5.22	211	4.07			
	SIB	Parent	2.05	208	2.99	-2.656	207	.009
		Teacher	2.66	208	3.95			
	CB	Parent	3.93	205	3.88	-.414	204	.680
		Teacher	4.06	205	4.61			
	RB	Parent	3.44	186	3.13	-1.310	185	.192
		Teacher	3.82	186	3.80			
	SUB	Parent	6.37	196	5.83	-.830	195	.407
		Teacher	6.82	196	6.60			
	RIB	Parent	2.86	216	2.58	-1.270	215	.205
		Teacher	3.12	216	3.03			
	Total	Parent	22.78	155	17.25	-2.217	154	.028
		Teacher	26.64	155	21.64			

Note: SRS-2-PTF (Social Responsiveness Scale-2-Preschool Turkish Form): SA: Social Awareness, SC: Social Cognition, SCOM: Social Communication, SM: Social Motivation, RIRB: Restricted Interests, and Repetitive Behaviors.

GARS-2-TV (Gilliam Autism Rating Scale-2-Turkish Version): C: Communication, SI: Social Interaction, SB: Stereotypical Behaviors.

SCC-R-PTV (Social Communication Checklist-Revised Preschool Turkish Version): SP: Social Participation, EL: Expressive Language, RL: Receptive Language, IP: Imitation / Play

RBS-R-TV (Repetitive Behaviors Scale-Revised-Turkish Version): SB: Stereotyped Behavior, SIB: Self-injurious Behavior, CB: Compulsive Behaviors, RB: Ritualistic Behavior, SUB: Sameness-Uniformity Behaviors, RIB: Restricted Interest Behavior.

As seen in Table 3, the differences between the means of scores given by the parents and teachers in all sub-dimensions of SRS-2-PTF and GARS-2-TV scales are significant. In both scales, the mean scores given by teachers for children/individuals are significantly higher than parents. In the SCC-R-PTV scale, while the difference between the mean scores of parents and teachers are significant ($t = 2.369$, $p < .05$) in the *Expressive Language*

subscale, there is no significant difference in the other subscales. In the Expressive Language subscale, the mean of the parents' scores ($\bar{x}=76.89$) for their children is significantly higher than teachers' ($\bar{x}=73.83$). While there are statistically significant differences between the averages of the scores obtained from parents and teachers for RBS-R-TV in the Stereotyped Behaviors and Self-injurious Behavior subscales and total score of the full scale, no significant difference is found in the other subscales. When the mean scores obtained from both the full scale and the subscales are examined, the teachers' mean scores for the students are significantly higher than those of the parents.

Wilcoxon signed-ranks test is also applied for RBS-R-TV. The results show that statistically significant differences in the Stereotyped Behaviors and Self-injurious Behavior subscales between the mean scores of the parents and teachers ($z=-3.035, p<.05$; $z=-2.123, p<.05$). The difference between the parent's and teachers' mean scores in the other subscales and the total score is not significant. According to the order of the subscales, it is seen that the results are $z=-.201, p>.05$, $z=-1.262, p>.05$; $z=-.816, p>.05$; $z=-1.082, p>.05$; $z=-1.820, p>.05$, respectively. When the results are compared with the paired Samples t-Test, it is seen that the only difference is in the total score of the full scale. Since the scores obtained from the full scale meet the normality assumption, the results of the paired samples t-test are discussed.

Consistency between Parents 'and Teachers' Classification of Children as Typically Developing or with ASD

The consistency between the parents and teachers who scored the scales used in ASD (SRS-2-PTF and GARS-2-TV), classifying their children as typically developing or with ASD is examined according to the Cohen kappa technique, and the results are presented in Table 4.

Table 4. *Cohen kappa results regarding consistency between parents' and teachers' classifications*

Scales	Kappa	p
SRS-2-PTF	.243	.000
GARS-2-TV	.560	.000

Note: SRS-2-PTF (Social Responsive Scale-2-Preschool Turkish Form), GARS-2-TV (Gilliam Autism Rating Scale-2-Turkish Version)

For SRS-2-PTF, the consistency between parents' and teachers' classification of children as having no ASD risk and children with mild, moderate, and high risk for ASD is significant but indicates a fair agreement. On the other hand, in GARS-2-TV, the

consistency between parents' and teachers' classification of children as having moderate possibility and having the high and low possibility of having ASD is significant and indicates a moderate level of agreement.

Discussion

This study aims to determine the consistency between the responses of parents and teachers for the same individual with ASD, using different scales that evaluate behavioral and developmental characteristics specific to ASD, utilizing different techniques. In this context, firstly, the internal consistency coefficients are calculated on the basis of subscales and full scale for parents and teachers who scored SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales that evaluate the behavioral and developmental characteristics of individuals with ASD. It is concluded that some subscales in SRS-2-PTF and RBS-R-TV scales are moderately reliable, but other scales and subscales are highly reliable. This situation can be interpreted as the internal consistency of the scorings made by the raters is high. Also, obtaining a moderate internal consistency reliability coefficient in some subscales may be due to the lower number of items in the subscale than other subscales. For example, while the subscale of SRS-2-PTF with the least number of items is "Social Participation," it is seen that the lowest reliability coefficients of this scale belong to this subscale. The literature states that reliability is affected by the number of items (Baykul, 2000; Crocker & Algina, 1986). In addition, Cronbach alpha coefficients related to teachers' scores are found to be higher than parents in all scales and subscales except SP and EL subscales of SCC-R-PTV. The difference between the reliability coefficients of the scores of the parents and teachers in the SC and SM subscales of the SRS-2-PTF and the RIB subscale of RBS-R-TV is found statistically significant. It is determined that the reliability coefficients for the teachers' scores in all three subscales are significantly higher than those of the parents. Murray et al. (2009), in a study comparing parents' and teachers' evaluations on social skills and behavioral measures for children with ASD, found that teachers made more consistent evaluations than parents. In this context, the results obtained from the study may be due to teachers' making more objective and realistic evaluations about the repetitive behaviors of children with ASD and their social communication and interaction skills. The behaviors measured in the relevant subscales require more objective scoring. Besides, parents' emotional bond with their children can prevent them from objective scoring (Azad et al., 2016). Additionally, it is thought that since parents have the opportunity to observe a small number of children who typically have ASD, they compare their children with typically developing children when completing scales. On the other hand, teachers have the

opportunity to observe more children with different degrees of ASD (Azad et al., 2016; Ryland et al., 2012). It is thought that this situation may have affected the obtained results.

In order to determine the relationship between the scores of parents and teachers who scored the SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales that evaluate behavioral and developmental characteristics of children with ASD associated with the disorder, the relationship between all subscales and scale total scores of parents and teachers is examined. When the correlations between the two groups in SRS-2-PTF are examined, it is seen that there are moderately significant relationships in all dimensions except the Social Awareness subscale. In GARS-2-TV and SCC-R-PTV scales, moderate and high levels of significant correlations are obtained between parents' and teachers' ratings. In RBS-R-TV, it is seen that there is a low and moderate level of a significant relationship between the parents' and teachers' ratings. The correlation coefficients in some subscales are low, but significant may be due to the sample size. The correlation coefficient is affected by the sample size (Bates, Zhang, Dufek, & Chen, 1996). In another study, it was observed that there was a low but significant relationship between parents' and teachers' measurements on the social skills and behaviors of children with ASD (Murray et al., 2009). In this context, according to the findings obtained from the current study, it can be concluded that it is doubtful that the ratings made by only the teacher or the parent in the SRS-2-PTF and RBS-R-TV scales reflect the situation of the child in a realistic way. In the literature, it was observed that the consistency between the scores of parents and teachers in the scales about subjective and internalized problems was lower, and there was a low-moderate level of consistency between two groups in the scales containing more explicit and observable behaviors (Jepsen, 2012). Vickerstaff et al. (2007) reached similar results in their examination of children with ASD. It was observed that the consistency between teachers and parents was higher in subscales measuring objective (more explicit, observable) behaviors. In this study, a possible reason for the low level of correlation between parents and teachers' scores in some scales (SA subscale of SRS-2-PTF and SUB subscale of RBS-R-TV) might be since behavior the teacher regards as frequently occurring is regarded as less frequently accruing by the parent or vice versa. This may also be attributed to the fact that while teachers encounter different students with ASD, parents observe only their children. Furthermore, the fact that the correlation levels between parents and teachers in the GARS-2-TV and SCC-R-PTV scales are moderately and highly significant is only evidence indicating that the scores covary. In other words, using only correlation coefficients, it should not be interpreted as that parents and teachers make similar ratings for the same

individual. For such an interpretation, it is necessary to examine both groups' means together with the correlation between parents and teachers.

When the parents are divided into two groups according to their educational level (undergraduate and higher, high school and lower), and the correlations between the scores obtained from teachers and parents for all subscales of SRS-2-PTF and RBS-R-TV scales are examined, it is seen that the correlations between the scores of teachers and the parents who are at least holding a bachelor's degree are higher than when there is no distinction between educational level is made. Only the correlation of SRS-2-PTF full-scale score is higher when the educational level is not differentiated for the whole group. However, although GARS-2-TV scores are divided into groups according to educational level, it is seen that the correlations are close to each other. In SCC-R-PTV, unlike other scales, it is observed that the correlations between the scores of the parents whose educational level is high school and lower and the scores of the teachers are higher in all subscales and full-scale scores than in the case of no distinction made between educational levels. In SRS-2-PTF and RBS-R-TV scales, it is found that scores of parents with higher educational levels are more consistent with scores of teachers. The expressions used in these scales can be more complex and difficult to understand; parents with lower educational levels may not have given correct responses because they might not have understood the items well enough. In this context, for future research, when data is collected from parents, preliminary interviews should be conducted to make sure that the expressions of the scales are understood correctly by parents with lower educational levels.

In order to determine the consistency between parents and teachers who scored SRS-2-PTF, GARS-2-TV, SCC-R-PTV, and RBS-R-TV scales, based on the comparison between means, a Paired Samples t-test was conducted. It is determined that there is a significant difference between the scores obtained from parents and teachers in all sub-dimensions of SRS-2-PTF and GARS-2-TV scales, and the scores given by teachers for students are higher than parents. In the GARS-2-TV scale, it is observed that although there are moderate and high-level significant correlations between the scores obtained from parents and teachers, there are significant differences between raters in terms of rigidity/generosity in scoring. Both of these scales provide information on the degree of ASD (Constantino & Gruber, 2012; Gilliam, 2006). Therefore, according to the findings of this study, it can be suggested that parents tend to score their children's degree of ASD mildly, while teachers tend to score the degree of ASD more severely. In this context, it is determined that it is not appropriate to comment on parents' and teachers' consistency based

only on correlations. Only in the Expressive Language subscale of SCC-R-PTV, while the differences between means of the parents' and teachers' responses are significant, there is no significant difference in the other subscales. In the Expressive Language subscale, parents' mean scores for their children are significantly higher than teachers'. As Vickerstaff et al. (2007) also stated, this situation might be due to differences in parents' and teachers' observations about children's social behaviors because of environmental factors. For RBS-R-TV, there are statistically significant differences between the means of the scores obtained from parents and teachers in the Stereotypical Behaviors and Self- Injurious Behavior subscales. In a full-scale score, the difference between parents' and teachers' mean scores in other subscales is not significant. When the mean scores obtained from both subscales and full scale are examined, the average scores the teachers give for the students are significantly higher than the parents. In the literature (Jepsen, 2012; Pekin, 2005; Vickerstaff et al., 2007), it was determined that there are differences between the scores according to the behaviors measured by the subscales. For example, in a study in which the Autism Social Skills Profile scale was scored by five raters and the level of consistency was determined according to different techniques, it was found that the rigidity or generosity levels of the raters were different, and there was an inconsistency between the raters (Pekin, 2015). In this study, significantly higher teachers' scores on both the RBS-R-TV Stereotypical Behaviors and Self- Injurious Behavior Sub-Scale, and the SRS-2-PTF and GARS-2-TV subscales than the parent scores, can also be explained by the tendency of parents to score lower on the degree of the spectrum or the tendency of teachers to score higher.

To compare the consistency between the classification of children as with and without ASD by parents and teachers who scored the scales (SRS-2-PTF and GARS-2-TV) that are used to evaluate the developmental and behavioral characteristics of individuals with ASD, according to the Cohen kappa technique, the consistency between the scores obtained from parents and teachers is examined. For SRS-2-PTF, the consistency between parents' and teachers' grouping of children without ASD risk and children at mild, moderate, and high risk for ASD indicates significant but low consistency. In GARS-2-TV, the consistency between parents' and teachers' grouping of the moderate and high probability of having ASD with a low probability of having ASD indicates significant and moderate consistency. The studies conducted in the literature suggest that the severity of the child's ASD symptoms may also affect the consistency between parent and teacher evaluations (Azad et al., 2016; Hundert, Morrison, Mahoney, Mundy, & Vernon 1997; Voelker et al., 2000). In the current study, the effect of ASD severity on consistency is not examined since the

number of children with severe ASD is low. In this context, the consistency between parent and teacher evaluations can be examined after children are grouped by symptom severity in future studies.

Considering the results obtained from the study in general, it is determined that teachers evaluate some subscales more consistently (internal consistency) than parents, that it is doubtful whether the ratings made by the teachers or the parents alone reflect the situation of the child in a realistic way and that it should not be interpreted that parents and teachers make similar scores for the same individual by using only correlation values. In addition, it is determined that in some subscales, the scores of parents with higher education levels are more consistent with teachers, that raters (parents and teachers) have different levels of rigidity or generosity. That inconsistency between raters is evident in some subscales.

It is recommended to use information from multiple individuals (e.g., families, therapists, teachers) in the evaluation process of children with ASD. The obtained findings from this study also support this situation. Features such as the property measured in the scale and the comprehensibility of the scale items should also be taken into account in the scales' evaluation process. Also, situations such as carelessness, personal bias, moiré effect, shifting towards the center effect, and generalization error are other factors that affect the consistency between raters (Çıkrıkçı, 2019; Özgüven, 2007). Researchers are recommended to interpret the consistency between raters considering these effects and inform parents before implementing it.

Ethical Approval: *This research was conducted with the permission obtained with the Ankara University Ethics Committee's decisions dated 04/03/2019, numbered 77 and 78.*

Conflict Interest: *The authors declare no conflict of interest.*

Authors Contributions: *Seher YALÇIN: research design, data collection, data analysis, and writing the manuscript; Hatice BAKKALOĞLU: finding the topic, research design, and writing the manuscript; Meral Çilem ÖKCÜN AKÇAMUŞ: data collection and writing the manuscript; Seyda DEMİR: data collection and writing the manuscript; Zeynep BAHAP KUDRET: data collection and writing the manuscript.*

References

- Akhun, İ. (1984). İki korelasyon katsayısı arasındaki farkın manidarlığının test edilmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 17(1), 1-7. doi: 10.1501/Egifak_0000001034
- Alpar, R. (2010). *Uygulamalı istatistik ve geçerlik, güvenilirlik*. Ankara: Detay Yayıncılık.
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders*. Arlington: American Psychiatric Publishing.
- Atılğan, H., Kan, A., ve Doğan, N. (2011). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Azad, G. F., Reisinger, E., Xie, M., & Mandell, D. S. (2016). Parent and teacher concordance on the Social Responsiveness Scale for children with autism. *School Mental Health*, 8, 368-376. doi: 10.1007/s12310-015-9168-6
- Bakkaloğlu, H., Ökcün-Akçamuş, M. Ç., Demir, Ş., Bahap-Kudret, Z., ve Yalçın, S., (2019). *Sosyal Yanıtlayıcılık Ölçeği-2-Okulöncesi Türkçe Formu'nun Türk kültürüne uyarlanması: Geçerlik ve güvenilirlik çalışması*. Yayınlanmamış araştırma raporu.
- Bates, B. T., Zhang, S., Dufek, J. S., & Chen, F. C. (1996). The effects of sample size and variability on the correlation coefficient. *Medicine and Science in Sports and Exercise*, 28(3), 386-391. doi: 10.1097/00005768-199603000-00015
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları.
- Bodfish, J. W., Symons, F. J., Parker, D. E., & Lewis, M. H. (2000). Varieties of repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders*, 30(3), 237-243. doi: 10.1023/a:1005596502855
- Büyüköztürk, Ş. (2013). *Sosyal bilimler için veri analizi el kitabı* (18. baskı). Ankara: Pegem Akademi.
- Can, A. (2016). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Ankara: Pegem Akademi.
- Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., & Brophy, S. L. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview- revised. *Journal of Autism and Developmental Disorders*, 33(4), 427-433. doi: 10.1023/a:1025014929212

- Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale-second edition (SRS-2)*. Torrance: Western Psychological Services.
- Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*, 2-11. doi: 10.1097/00004703-200002000-00002
- Crocker, L. M., & Algina, L. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Çıkrıkçı, N. (2019). Ölçmede güvenilirlik. R. N., Çıkrıkçı (Ed.), *Eğitimde ölçme ve değerlendirme içinde (77-100)*. Ankara: Anı Yayıncılık.
- Demir, Ş., Ökcün-Akçamuş, M. Ç., Bahap Kudret, Z., & Bakkaloğlu, H. (2018, 11-13 Ekim). *OSB olan çocuklar için öğretmenler tarafından doldurulan Tekrarlayıcı Davranışlar Ölçeği-R-Türkçe Versiyonunun psikometrik özelliklerinin incelenmesi (Sözlü Bildiri)*. 28. Ulusal Özel Eğitim Kongresi, Eskişehir Osmangazi Üniversitesi, Eskişehir, Türkiye
- Diken, İ. H., Ardıç, A., Diken, Ö., ve Gilliam, J. E. (2012). Gilliam Otistik Bozukluk Derecelendirme Ölçeği-2-Türkçe Versiyonu'nun (GOBDÖ-2-TV) geçerlik ve güvenilirliğinin araştırılması: Türkiye standardizasyon çalışması. *Eğitim ve Bilim, 37*(166), 318-328.
- Dinnebeil, L. A., Sawyer, B. E., Logan, J., Dynia, J. M., Cancio, E., & Justice, L. M. (2013). Influences on the congruence between parents' and teachers' ratings of young children's social skills and problem behaviors. *Early Childhood Research Quarterly, 28*(1), 144-152. <https://doi.org/10.1111/j.1469-7610.1989.tb00772.x>
- Gilliam, J. E. (2006). *Gilliam autism rating scale*. Austin, TX.: PRO-ED.
- Goodwin, L. D. (2001). Interrater agreement and reliability. *Measurement in Psychological Education and Exercises Science, 5*(1), 13-34. doi: 10.1207/S15327841MPEE0501_2
- Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010). Cross-informant agreement for ratings for social skill and problem behavior ratings: An investigation of the Social Skills Improvement System Rating Scales. *Psychological Assessment, 22*(1), 157.
- Güler, N. (2008). *Klasik test kuramı genellenebilirlik kuramı ve rasch modeli üzerine bir araştırma* (Doktora Tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara. <https://tez.yok.gov.tr/UlusalTezMerkezi/> adresinden alınmıştır.

Hundert, J., Morrison, L., Mahoney, W., Mundy, F., & Vernon, M. L. (1997). Parent and teacher assessments of the developmental status of children with severe, mild/moderate, or no developmental disabilities. *Topics in Early Childhood Special Education*, 17, 419-434. doi: 10.1177/027112149701700404

İnan, B., Ökcün-Akçamuş, M. Ç., Bakkaloğlu, H., ve Yalçın, S. (2020). Turkish adaptation and validity-reliability study of the revised Social Communication Checklist for children with autism spectrum disorder. *Journal of Early Childhood Studies*, 4(4), 172-196. doi: [10.24130/eccd-jecs.1967202041193](https://doi.org/10.24130/eccd-jecs.1967202041193)

Jepsen, M. I., Gray, K. M., & Taffe, J. R. (2012). Agreement in multiinformant assessment of behaviour and emotional problems and social functioning in adolescents with Autistic and Asperger's Disorder. *Research in Autism Spectrum Disorders*, 6, 1091-1098. doi: 10.1016/j.rasd.2012.02.008

Joosten, A. V., Bundy, A. C., & Einfeld, S. L. (2012). Context influences the motivation for stereotypic and repetitive behaviour in children diagnosed with intellectual disability with and without autism. *Journal of Applied Research in Intellectual Disabilities*, 25(3), 262-271. doi: 10.1111/j.1468-3148.2011.00663.x

Kaat, A. J., Gadow, K. D., & Lecavalier, L. (2013). Psychiatric symptom impairment in children with autism spectrum disorders. *Journal of Abnormal Child Psychology*, 41, 959-969. doi: 10.1007/s10802-013-9739-7

Kalaycı, Ş. (2010). *SPSS uygulamalı çok değişkenli istatistik teknikleri* (5. baskı). Ankara: Asil Yayın Dağıtım.

Karasar, N. (2000). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın ve Dağıtım.

Kirby, A. V., Boyd, B. A., Williams, K. L., Faldowski, R. A., & Baranek, G. T. (2017). Sensory and repetitive behaviors among children with autism spectrum disorder at home. *Autism*, 21(2), 142-154. doi: 10.1177/1362361316632710

Lampi, A., Fitzpatrick, P., Romero, V., Amaral, J., & Schmidt, R. C. (2018). Understanding the influence of social and motor context on the co-occurring frequency of restricted and repetitive behaviors in autism. *Journal of Autism and Developmental Disorders*, 50, 1479-1496. doi: [10.1007/s10803-018-3698-3](https://doi.org/10.1007/s10803-018-3698-3)

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison Wesley, Reading, MA.

Mirenda, P., Smith, I., Vaillancourt, T., Georgiades, S., Duku, E., Szatmari, P., et al. (2010). Validating the repetitive behavior scale-revised in young children with autism

- spectrum disorder. *Journal of Autism and Developmental Disorders*, 40(12), 1521-1530. doi: 10.1007/s10803-010-1012-0
- Murray, D. S., Ruble, L. A., Willis, H., & Molloy, C. A. (2009). Parent and teacher report of social skills in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, 40(2), 109-115. doi: 10.1044/0161-1461
- Ökcün-Akçamuş, M., Bakkaloğlu, H., Demir, Ş., ve Bahap-Kudret, Z. (2019). Otizm spektrum bozukluğunda Tekrarlayıcı Davranışlar Ölçeği-Revize Türkçe Sürümünün geçerlilik ve güvenilirlik çalışması. *Anadolu Psikiyatri Dergisi*, 20(Ek sayı.1), 65-72. doi: 10.5455/apd.42649
- Özgüven, İ. E. (2007). *Psikolojik testler*. Ankara: PDREM Yayınları.
- Pekin, Z. (2015). *Otizm sosyal beceriler profili ölçeğinde puanlayıcılar arası güvenirliliğin klasik test kuramı ve genellenebilirlik kuramına göre karşılaştırılması* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Raver, S. A. (2009). *Early childhood-special education-0-8 years: Strategies for positive outcomes*. Pearson Education, Inc.
- Ryland, H. K., Hysing, M., Posserud, M. B., Gillberg, C., & Lundervold, A. J. (2012). Autism spectrum symptoms in children with neurological disorders. *Child and Adolescent Psychiatry and Mental Health*, 6(34), 1-9. doi: 10.1186/1753-2000-6-34
- Tasse, M., & Lecavalier, L. (2000). Comparing parent and teacher ratings of social competence and problem behaviors. *American Journal on Mental Retardation*, 105, 252-259. doi: 10.1352/0895-8017
- Tucker, V., & Schwartz, I. (2013). Parents' perspectives of collaboration with school professionals: Barriers and facilitators to successful partnerships in planning for students with ASD. *School Mental Health*, 5(1), 3-14. doi: 10.1007/s12310-012-9102-0
- Turner, M. (1999). Annotation: Repetitive behaviour in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry*, 40(6), 839-849. doi: [10.1111/1469-7610.00502](https://doi.org/10.1111/1469-7610.00502)
- Vickerstaff, S., Heriot, S., Wong, M., Lopes, A., & Dossetor, D. (2007). Intellectual ability, self-perceived social competence, and depressive symptomatology in children with

- high-functioning autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 37, 1647-1664. doi: 10.1007/s10803-006-0292-x.
- Verhulst, F. C., & Akkerhuis, G. W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4–12. *Journal of Child Psychology and Psychiatry*, 30(1), 123-136. <https://doi.org/10.1111/j.1469-7610.1989.tb00772.x>
- Voelker, S. L., Shore, D. L., Lee, C. H., & Szuszkiewicz, T. A. (2000). Congruence in parent and teacher ratings of adaptive behavior of low-functioning children. *Journal of Developmental and Physical Disabilities*, 12(4), 367-376. doi: 10.1023/A:1009436230984
- Wainer, A. L., Berger, N. I., & Ingersoll, B. R. (2017). Brief report: The preliminary psychometric properties of the Social Communication Checklist. *Journal of Autism and Developmental Disorders*, 47(4), 1231-1238. doi: 10.1007/s10803-016-3026-8