



## Otizm Spektrum Bozukluğu Gösteren Çocuklar için Eğitsel Değerlendirme ve Öğrenme Hızı Öngörüsü Kayıt Formlarının Psikometrik Niteliklerinin Belirlenmesi

### Determining Psychometric Properties of Educational Assessment and Prognosis of Learning Rate Record Forms for Children with Autism Spectrum Disorders

**Veysel Aksoy**, Yard. Doç. Dr., Anadolu Üniveristesi-Engelliler Araştırma Enstitüsü, [vaksoy@anadolu.edu.tr](mailto:vaksoy@anadolu.edu.tr)  
**İbrahim H. Diken**, Prof. Dr., Anadolu Üniveristesi-Engelliler Araştırma Enstitüsü, [ihdiken@anadolu.edu.tr](mailto:ihdiken@anadolu.edu.tr)

**ÖZ.** Bu araştırmanın amacı “Eğitsel Planlama için Otizm Tarama Aracının (Autism Screening Instrument for Educational Planning-3, ASIEP-3)” üçüncü güncellemesinde yer alan Eğitsel Değerlendirme Kayıt Formu (EGİDEKAF) ve Öğrenme Hızı Öngörüsü Kayıt Formu (ÖHÖKAF) informal değerlendirme araçlarının Türkçe formlarının psikometrik niteliklerini incelemektir. Ülkemizde Otizm Spektrum Bozukluğu (OSB) olan öğrencilerin eğitsel değerlendirmelerinde kullanılacak standardizasyonu yapılmış informal değerlendirme araçlarının eksikliği araştırmanın gerekçesini oluşturmaktadır. Araştırmanın çalışma grubunu üç özel özel eğitim ve rehabilitasyon merkezinde eğitim almakta olan öğrenciler oluşturmaktadır. EGİDEKAF’ın çalışma grubunda OSB ve Zihinsel Yetersizlik (ZY) tanılı, 35’i kız ve 91’i erkek olmak üzere toplam 126 öğrenci yer almıştır. ÖHÖKAF’ın çalışma grubunda ise OSB ve ZY tanılı, 34’ü kız ve 97’si erkek olmak üzere toplam 131 öğrenci yer almıştır. Performans alımına dayalı veri toplamayı gerektiren çalışmadan elde edilen veriler araçların geçerlik ve güvenirlik hesaplamalarının yapılması amacıyla analiz edilmiştir. Analiz bulguları her iki aracın da OSB olan çocukların eğitsel değerlendirmelerinde ve öğrenme hızlarının öngörülmesinde kullanılacak geçerli ve güvenilir araçlar olduğunu ortaya koymaktadır.

**Anahtar Kelimeler.** Otizm Spektrum Bozukluğu, İformel Değerlendirme, Eğitsel Değerlendirme, Öğrenme Hızı Öngörüsü

**ABSTRACT.** The purpose of this research is to study the psychometric properties of informal assessment tools, Educational Assessment Record Form (EARF) and Prognosis of Learning Rate Record Form (PLRRF), which appear in the third edition of Autism Screening Instrument for Educational Planning-3, ASIEP-3. The rationale for the research is the lack of standardised informal assessment tools to use for the educational assessment of children with Autism Spectrum Disorder (ASD) in Turkey. The study group consists of students, who receive education at three private institutions for special education and rehabilitation. A total of 126 students who are diagnosed with ASD and intellectual disability (ID) were included in the study group of EARF, 91 male and 35 female. And in the study group of PLRRF, a total of 131 students who are diagnosed with ID and ASD were included, 97 male and 34 female. The data that were obtained from the study requiring data collection based on performance feedback were analysed in order to assess the validity and reliability of the tools. The findings from the analysis revealed that both tools were valid and reliable to be used for educational assessments and prognosis of learning rate of children with ASD.

**Keywords.** Autism Spectrum Disorder, Informal Assessment, Educational Assessment, Prognosis of Learning Rate

#### SUMMARY

**Purpose and Significance:** The two tools whose psychometric properties were studied and adapted into within this study are two informal tools included in the third edition of the Autism Screening Instrument for Educational Planning (ASIEP-3), which was developed by Krug, Arick and Almond (2008). The ASIEP-3 tool set is an Autism Behaviour Checklist (ABC) that indicates children with a probability of ASD, consisting of four informal assessment record forms standardized through the outputs obtained from Intellectually Disabled (ID) and ASD individuals. Autism Behaviour Checklist (ABC), which is contained in the tool, is a norm referenced screening instrument. It is stated that the ABC could be used for screening and diagnosing specifically children suspected of having ASD. The other four tools that include the two tools in this set and

are studied for their psychometric properties as part of this study are standardised informal assessment tools. It is indicated that these tools are described as informal yet standardized tools, since validity and reliability analyses of them were performed, scored and they require standard procedure for their implementation, and interpreting their results (Krug and et al., 2008).

Krug and et al. (2008) indicate that ASIEP-3 can be used for diagnosing children with ASD, placement and planning their educational programs and monitoring their development. Considering the existing diagnosing and screening tools in Turkey, it appears that we experience a lack of tools, which enable an objective comparison of educational assessments of individuals with ASD language-communication and social interaction performances of individuals with ASD to a sampling group. During post-diagnosis phases, the programs for students are planned through criterion-referenced assessments created by teachers, and student development is monitored along this program. These types of standardised informal tools allow for an objective measurement and assessment to be used during the phases of program development, program evaluation and monitoring for individuals with ASD. Thus, the Turkish adaptation of Educational Assessment Record Form (EARF) and Prognosis of Learning Rate Record Form in the scope of ASIEP-3 tools, and studying their psychometric properties were deemed important. In this regard, this study seeks answers to the following research questions:

Is the EARF a valid and reliable tool for assessing the levels of reciprocal language, productive language and body concept and speech imitation of the children with ASD? Is the PLLRF a valid and reliable tool for assessing the levels of learning rate of children with ASD?

#### **Methodology:**

**Research Design:** This study was conducted through a general screening model. In accordance with the model, in this study did both single and relational screenings. Through the tools, not only single type and quantity of variables were determined, but also the relationships between the specified dimensions of variables were studied through the tools properties whose psychometric properties were aimed to identify.

**Participants:** The participants of this study were the students with special needs and diagnosed with ASD and ID and teacher from three private institutions for special education and rehabilitation. In choosing the study group for the study, purposeful sampling from improbable sampling methods was used. A total of 33 teachers, 26 (79%) female and 7 (21%) male were included in the research during the phase of data gathering for the tools whose psychometric properties were studied. A total of 126 students who are diagnosed with ASD and intellectual disability (ID) were included in the study group of EARF, 91 male and 35 female. And in the study group of PLRRF, a total of 131 students who are diagnosed with ID and ASD were included, 97 male and 34 female.

**Data Collection Tool:** Since this is a study that determines the psychometric properties of assessment tools, validity and reliability results with regard to the scores of the assessment tools attained from the study were reported in the results section in detail. In this study, Turkish Early Language Development Test (TELDT) was chosen used in criterion referenced validity assessment. TELDT is the adaptation of reciprocal and productive language of children between the ages of 2 and 8; an assessment tool named as TELD-3 developed by Hresko, Reid and Hammill (1999).

**Data Analysis;** Analyses on authentic tool and additional analyses were done for the validity and reliability analyses of the tools. Analyses were performed through SPSS programme.

**Results: Validity-Reliability Results with regard to The Scores Achieved in the EARF;** Test-retest analysis was performed through the calculation of Pearson's correlation coefficient for the reliability analysis of the tool. R values calculated for subtests and total point were between .91 and .98. The Cronbach Alpha coefficient calculated for the internal consistency of the tool was  $\alpha = .87$ . Item-total correlation values of the four items that make up the total score of educational assessment were between .66 and .80. Pearson's correlation coefficients calculated for the subtests and total score for the Turkish version of TELD-3 and its criterion-referenced consistency were between .49 and .74. In

the construct validity of the tool, it was seen that it could distinguish children with ASD and ID from non-ASD and normal children, and its correlations with its subtests were significant.

*Validity-Reliability Results with regard to The Points Achieved in PLRRF:* For the inter-rater judgement, concurrence percentage of the records from two observers was viewed. Inter-rater reliability was found as %100. In the construct validity of the tool, it was seen that it could distinguish children with ASD and ID from non-ASD and normal children, and that its correlations with its subtests are significant

**Discussion and Conclusions:** Analyses by test-retest and item-total score correlation methods were conducted in order to test the reliability of EARF. In the analyses for validity dimension, on the other hand, inter-subtest correlation and reviewing subtest differences methods within the scope of criterion-referenced validity and construct validity were conducted. When interpreting the results, psychometric properties of Turkish version of the tool suggest that it is both reliable and valid.

In order to view reliability of the PLRRF tool, an analysis by inter-rater reliability was conducted. For the analyses for validity, on the other hand, inter-subtest correlation and reviewing the subtest differences methods within the scope of construct validity were conducted. When interpreting the results, psychometric properties of the Turkish version of the tool suggest that it is both reliable and valid.

---

## GİRİŞ

Otizm Spektrum Bozukluğu (OSB), yaşamın ilk üç yılında kendisini gösteren, son yıllarda da görülme olasılığı giderek yaygınlaşan nöro-gelişimsel bir bozukluk olup, özel eğitim gerektiren bozukluk kategorisi olarak kabul edilmektedir. Çocukların dilsel, bilişsel ve sosyal gelişimlerini etkileyen nöro-gelişimsel bir bozukluk olan OSB'nin yaygınlığı diğer bozukluk türlerine göre son yirmi yıldır beklenmeyen bir şekilde artış göstermektedir (Allen, Robins, ve Decker, 2008). Dünyanın pek çok yerinde olduğu gibi Türkiye'de de kaç çocuğun OSB'den etkilendiği tam olarak bilinmemektedir. Her 160 ile 110 çocuktan birinde OSB görüldüğüne dair bulgular rapor edilmekle birlikte kesin rakamlar ifade edebilmek mümkün görünmemektedir (Glosser, 2011). Bu alandaki kayıtların daha uzun süreden beri tutulduğu ülkelerin başında gelen ABD'de Center of Disease-CDC adlı kuruluşun 2014'de yayımladığı verilere göre ülkedeki 8 yaşında ve okul çağında olan her 68 çocuktan biri OSB'den etkilenmektedir (CDC, 2014). Yaygınlığın artmasıyla birlikte genel olarak toplumda özel olarak da anne-babalarda, eğitim ve sağlık çalışanlarında ve araştırmacılarında bu bozukluk türüne olan ilginin artmakta olduğu belirtilmektedir (Dover ve Le Couteur, 2007).

OSB her çocukta farklı özellikler gösterse de OSB olan çocuklar genel olarak iletişim, sosyal etkileşim, tekrarlayan davranışlar ve sınırlı ilgiler olmak üzere üç temel alanda problemler yaşamaktadırlar. Bu alanlarda yaşanan problemlerin miktarının ve düzeyinin farklılık göstermesi OSB'nin belli bir süre şemsiye bir kavram halini almasında etkili olmuştur. Amerikan Psikiyatri Birliği (*American Psychiatric Association-APA*) tarafından yayınlanan Psikiyatride Hastalıkların Tanımlanması ve Sınıflandırılması Elkitabı, Yeniden Gözden Geçirilmiş dördüncü baskısında (DSM-IV-TR) (2001), bu şemsiyede beş alt bozukluk yer almaktadır. Bunlar Asperger Sendromu, Otizm, Rett Sendromu, Çocukluk Dezintegratif Bozukluğu, Başka Türü Adlandırılmayan Yaygın Gelişimsel Bozukluk (Atipik Otizmi de kapsar) şeklinde sıralanmaktadır. 2013 yılında kabul edilen ve Mayıs-2013 itibarıyla yayınlanan DSM-V'de ise bu şemsiye kavram yerini tek bir kavrama yani Otizm Spektrum Bozukluğu (OSB) kavramına bırakmıştır.

OSB olan çocukların değerlendirilmesi, çocukların gelişimi, alınması gereken önlemler ve yapılması gereken müdahaleler için karar verebilmek amacıyla farklı kaynaklardan, farklı değerlendirme yöntem ve araçları kullanılarak veri/bilgi toplama süreci olarak tanımlanmıştır (Sucuoğlu, 2012). OSB olan çocukların değerlendirilmesinde diğer özel gereksinimli çocuklarınkine benzer bir süreç işlemektedir. OSB olan çocukların bu değerlendirme süreçlerini Sucuoğlu (2012), tarama, tanılama ve gelişimsel veri toplama süreçleri olarak üç grupta incelenebileceğini ifade etmiştir. Tincani (2010), OSB tanısı konulması sürecinde tanının gelişimsel değerlendirmeye desteklenmesi gerektiğini belirtmekte ve çocukların OSB tanısı alması sürecinde tanı araçlarından

alınan puanların çoğu kez yeterli olmayacağını, tanının kesinleştirilmesi için çocuğun gelişimsel olarak özellikle dil, sosyal iletişim, bilişsel-akademik alanlarda da değerlendirilmesi gerektiğini ifade etmektedir. Bahsi geçen değerlendirmelere ek olarak OSB tanı belirtisi olan parmak ucunda yürüme, kollarıyla kanat çırpma, elleri ile sürekli alkış yapma gibi stereotipik davranışlar da gözlem ve görüşme yöntemleri kullanılarak değerlendirilebilir. Özel eğitim hizmetleri açısından değerlendirmenin nihai amacı öğrencilerin akademik, davranışsal ya da fiziksel özelliklerini belirlemek ve bu özelliklere uygun yasal ve eğitsel kararlar alabilmek amacıyla veri toplama süreci olduğu söylenebilir (Gürsel, 2012). Avcıoğlu (2011), tanılama sonucunda özel eğitim hizmetleri için uygunluğu belirlenen öğrencilerin uygun yasal ve eğitimsel kararları almak, eğitsel müdahale programlarını planlamak ve daha sonra da müdahalelerin etkililiğini ölçümlemek amacıyla kapsamlı eğitsel değerlendirmelerden geçirilmesi gerektiğini belirtmektedir. Gelişimsel veri toplama süreci eğitsel değerlendirme yapılmasını gerektiren süreçtir. Peterson ve Meier (1988), eğitsel değerlendirme ile çocuğun farklı gelişimsel alanlarda işlevde bulunma düzeyinin belirlenerek çocuk için geliştirilecek bireyselleştirilmiş eğitim programında (BEP) kullanılacak performans düzeyinin ortaya konulduğunu ifade etmişlerdir. Bu sayede çocuğun güçlü ve zayıf yanları ve hangi alanlarda ne tür bir müdahaleye gereksinim olduğu belirlenmektedir. Eğitsel değerlendirme kapsamında OSB olan çocukların ağırlıklı olarak iletişim ve dil becerileri, sosyal ve uyumsal davranışları ile bilişsel ve akademik becerilerinin değerlendirmeleri yapılmakta ve bu süreçte formel ve informal değerlendirme araçları kullanılabilir (Sucuoğlu, 2012).

Bilindiği gibi formel değerlendirmeler norm referanslı ve standardize edilmiş değerlendirmelerdir. OSB'nin taranması ve tanınması sürecinde dünyada ve ülkemizde kullanılan araçlar bu özellikleri taşımaktadır. Klinik bulgu elde etmek ve değerlendirme yapmak için kullanılan ölçme araçlarının bu özellikleri taşıması beklenmektedir. Ancak bir çocukta OSB'nin bulunup bulunmadığını ortaya koymak her zaman çok kesin bir biçimde mümkün değildir. Bu nedenle değerlendirmelerin özellikle dil, sosyal iletişim, bilişsel-akademik ve davranışsal alanlarda yapılacak değerlendirmelerle desteklenmesi hem tanıyı desteklemeye hem de çocuk için geliştirilecek müdahale programını planlamaya katkı sağlayacaktır.

Bu çalışma kapsamında Türkçeye kazandırılarak psikometrik özellikleri incelenmiş olan iki araç Krug, Arick ve Almond (2008) tarafından geliştirilmiş olan Eğitsel Planlama için Otizm Tarama Aracının (*Autism Screening Instrument for Educational Planning-3, ASIEP-3*) üçüncü güncellemesinde yer alan iki informal araçtır. ASIEP-3, uygulanan çocuklarda OSB olma olasılığını ortaya koyan bir otizm davranış kontrol listesi (Autism Behavior Checklist- ABC) ve OSB olan ve zihinsel yetersizliği (ZY) olan çocuklarından elde edilen verilerle standardize edilmiş dört informal değerlendirme kayıt formunun yer aldığı bir araç setidir. Araçta yer alan Otizm Davranış Kontrol Listesi (ODKL) norm referanslı standart bir tarama aracıdır. Özellikle OSB şüphesi olan çocukların taranması ve tanınmasında kullanılabilmesi ifade edilmiştir. Sette yer alan ve bu çalışma kapsamında psikometrik özellikleri incelenen iki aracın da içinde yer aldığı diğer dört araç ise standardize edilmiş informal değerlendirme araçlarıdır. Araçların standardizasyonu yapılmış olmakla birlikte geniş bir örnekleme norm referansları oluşturulmamıştır. Bunun yerine tanıli bir örneklem grubuyla karşılaştırma olanağı sağlayan örneklem referansları mevcuttur. Araçların geçerlik ve güvenilirlik analizlerinin yapılmış olması ve ayrıca uygulanması, puanlanması ve sonuçlarının yorumlanmasının standart prosedürler gerektirmesi nedeniyle informal ancak standardize edilmiş araçlar olarak tanımlandığı belirtilmektedir (Krug ve diğ., 2008).

Krug ve diğerleri (2008), ASIEP-3'ün OSB olan çocukların tanınması, yerleştirilmesi, eğitsel programlarının planlanması ve gelişimlerinin izlenmesi amacıyla kullanılabilmesini ifade etmektedirler. Ülkemizde var olan tanılama ve tarama araçları incelendiğinde OSB olan çocukların eğitsel performanslarını bir örneklem grubuyla nesnel olarak karşılaştırmaya olanak sağlayan araçların bulunmadığı görülmektedir. Tanı sonrası süreçlerde öğretmen yapımı ölçüt bağımlı değerlendirmeler ya da programa dayalı değerlendirmeler yoluyla öğrencilerin programı planlanmakta ve öğrenci gelişimi bu program doğrultusunda izlenmektedir. Bu türden standardize edilmiş informal araçların varlığında OSB olan çocuklar için program geliştirme, programı değerlendirme ve izleme süreçlerinde kullanılacak nesnel bir ölçme ve değerlendirme olanağı ortaya çıkmaktadır. Bu nedenlerle ASIEP-3'te yer alan Eğitsel Değerlendirme Kayıt Formu-EGİDEKAF (*Educational Assessment Record Form*) ve Öğrenme Hızı Öngörüsü Kayıt Formu-ÖHÖKAF

(*Prognosis of Learning Rate Record Form*) araçlarının Türkçeye kazandırılarak psikometrik niteliklerinin incelenmesi önemli görülmüştür. Bu amaçla gerçekleştirilen çalışmada aşağıdaki araştırma sorularına yanıt aranmıştır.

EGİDEKAF, OSB olan çocukların alıcı dil, ifade edici dil, beden kavramı ve konuşma taklidi düzeylerini ölçmede güvenilir ve geçerli bir araç mıdır? ÖHÖKAF, OSB olan çocukların öğrenme hızı düzeylerini ölçmede güvenilir ve geçerli bir araç mıdır?

## YÖNTEM

### Araştırma Deseni

Bu araştırma, tarama modellerinden genel tarama modeli ile gerçekleştirilmiştir. Genel tarama modelleri, çok sayıda elemandan oluşan bir evrende, evren hakkında genel yargıya varmak amacı ile evrenin tümü ya da evrenden alınan bir örneklem üzerinde yapılabilmektedir (Karasar, 2008). Model gereği hem tekil hem de ilişkisel taramaların yapıldığı bu çalışmada psikometrik niteliklerinin belirlenmesi amaçlanan araçlar yoluyla değişkenlerin hem tek tek tür ve miktarları belirlenmiş hem de değişkenlerin belirli boyutları arasındaki ilişkiler incelenmiştir.

### Çalışma Grubu

Bu araştırmanın katılımcıları üç Özel Özel Eğitim ve Rehabilitasyon Merkezine devam etmekte olan OSB ve ZY tanısı almış özel gereksinimli öğrenciler ve aynı kurumlarda görev yapmakta olan öğretmenlerdir. Psikometrik nitelikleri belirlenmesi amaçlanan araçların OSB tanısı almış çocukların değerlendirilmesinde kullanılacak olmasına rağmen çalışma grubunda OSB tanısı olan çocukların yanında ZY tanısı olan çocuklardan da veri toplanmasının nedenleri şu şekilde açıklanabilir. İlk olarak özgün çalışmada da her iki gruptan veri toplanmıştır. Psikometrik niteliklerin belirlenmesi çalışmasında da özgün çalışmaya bağlı kalınması amaçlandığından benzer bir uygulama gerçekleştirilmiştir. İkinci olarak aracın geçerliliğini sınamak amacıyla yapılan ayırt edicilik analizlerinde sahip oldukları özellikler nedeniyle OSB'ye en yakın grup olan orta ve ağır düzeyde ZY tanısı olan çocuklardan veri toplanması ayırt ediciliğe daha gerçekçi kanıtlar sağlamaktadır. Üçüncü olarak gerek OSB olan gerekse de ZY tanısı olan çocukların eğitsel değerlendirmesinde özel eğitim uygulamaları tarafından benzer yöntem ve prosedürler kullanılmaktadır.

Araştırmanın çalışma grubu seçilirken olasılıklı olmayan örnekleme yöntemlerinden amaçlı örnekleme tekniği kullanılmıştır. Araştırma deseni gereği ortalama olarak 2-20 yaş arası OSB ve ZY tanısı olan çocuklara belirli bir sayının üzerinde ulaşmak gerekliliği vardır. Ayrıca araştırma uygulaması bu çocuklarla çalışmakta olan öğretmenlerin uygulamacı olmasını gerektirdiğinden öğrenci tanı, yaş ve sayı ölçütlerine öğretmenlerin gönüllülüğü ölçütü de eklenmiştir. Araştırmanın uygulama yöntemi her bir öğrencinin yarı yapılandırılmış ortamlarda, belirli bir süreyle performansının alınması ve bunun kaydedilmesini gerektiren özel olarak düzenlenmiş ortam, materyal ve zaman gerekleri olan etkileşime dayalı uygulamalardır. Bu durum özellikle kamu okullarında uygulamayı güçleştiren bir etkidir. Bu nedenle örneklem seçimine öğrenci özellikleri ve öğretmenlerin gönüllülüğünün yanı sıra erişilebilirlik ve uygulanabilirlik olarak tanımlayabileceğimiz değişkenler de katılmaktadır. Tüm bu faktörler nedeniyle araştırmanın örnekleme seçilirken yönetici ve öğretmenleri araştırmaya katılmaya gönüllü olan özel özel eğitim ve rehabilitasyon merkezleri uygulama alanı olarak seçilmiştir. Bu merkezlere devam etmekte olan öğrencilerden ölçütleri karşılayanlar da örnekleme dâhil edilmişlerdir. Örnekleme, OSB tanısı olan öğrencilerin yüksek işlevli ya da atipik otizm tanılı olanları yer almamıştır. ZY grubunda ise orta-ağır ZY tanılı çocuklar örnekleme dâhil edilmiştir.

Araştırmada psikometrik nitelikleri belirleme çalışması yapılan araçların veri toplama sürecine 26'sı (%79) kadın ve 7'si (%21) erkek olmak üzere toplam 33 öğretmen katılmıştır. Öğretmenlerin yaşları 22 ila 57 arasında değişmekte olup yaş ortalamaları 33.8, yaşlarının standart sapması 10.72 ve ranjı 35'dir. Öğretmenlerin mesleki deneyimleri 1 ila 26 yıl arasında değişmekte olup mesleki deneyim ortalamaları 7.03, standart sapması 5.77 ve ranjı 25'dir. Çalışma kapsamında verilerin elde edildiği öğrencilere ilişkin bilgiler Tablo 1'de verilmiştir.

**Tablo 1. Araştırmaya Katılan Özel Gereksinimli Çocuklara İlişkin Bilgiler.**

	Cinsiyet	Yaş*						
		n	%	min.	max.	X	ss	Ranj
EGİDEKAF	OSB-Kız	22	30	2;2	22;7	9;4	4;11	20;5
	OSB-Erkek	53	70	2;1	17;7	9;4	4;1	15;6
	Toplam	75	100	2;1	22;7	9;4	4;4	20;6
	ZY-Kız	13	25	5;7	12;9	10;8	1;10	7;2
	ZY-Erkek	38	75	2;6	17;10	10;2	4;4	15;4
	Toplam	51	100	2;6	17;10	10;3	3;10	15;4
	Genel-Kız	35	28	2;1	22;7	9;9	4;2	20;6
	Genel-Erkek	91	72	2;1	17;10	9;8	4;3	15;9
	Genel-Toplam	126	100	2;1	22;7	9;9	4;2	20;6
	ÖHÖKAF	OSB-Kız	20	27	3;1	22;7	8;1	4;8
OSB-Erkek		55	73	2;1	17;7	9;5	4;5	15;6
Toplam		75	100	2;1	22;7	9	4;6	20;6
ZY-Kız		14	25	2;2	17;10	10;7	3;6	15;8
ZY-Erkek		42	75	2;6	17;10	8;10	4;2	15;4
Toplam		56	100	2;2	17;10	9;4	4;1	15;8
Genel-Kız		34	26	2;2	22;7	9;2	4;4	20;5
Genel-Erkek		97	74	2;1	17;10	9;2	4;4	15;9
Genel-Toplam		131	100	2;1	22;7	9;2	4;4	20;6

\*Yaşlar yıl ve ay olarak verilmiş olup “;” işaretinden önce olan değer yılı sonraki değer ayı göstermektedir (Örn; 12;7, 12 yaş 7 ay olarak okunur.). Yaş verileri ay cinsinden toplanmış olup analizler yapıldıktan sonra yıl ve ay olarak dönüştürülerek tabloda sunulmuştur.

### Veri Toplama Araçları

Bu araştırma ölçme araçlarının psikometrik niteliklerini belirleme çalışması olduğundan araştırmadan elde edilen ölçme araçlarının puanlarına ilişkin geçerlik ve güvenilirlik bulguları araştırmının bulgular bölümünde detaylı olarak rapor edilmiştir. Bu nedenle bu bölümde aracın özgün formunun geliştirilmesinde rapor edilen geçerlik güvenilirlik bulguları ve izleyen bölümde işlem süreci rapor edilmiştir.

*Eğitsel Değerlendirme Kayıt Formu (EGİDEKAF):* EGİDEKAF'ın uygulanabilmesi için öğrencinin belirli giriş-düzeyi davranışlarına sahip olması gerekmektedir. Bu davranışlar yerinde oturarak bekleme, belirli bir nesneye bakma gibi davranışlardır. Ayrıca başarılı bir değerlendirme için çocuğun değerlendirilmesini engelleyecek davranış problemlerinin olmaması gerekmektedir. Bu değerlendirmenin bulguları çocuklarla çalışmaya başlamak için öneriler sağlamaktadır. Eğitsel değerlendirmenin amacı, çocuğun işlevsellik seviyesini beş alanda incelemektir: (1) Oturma davranışı, (2) alıcı dil, (3) ifade edici dil, (4) beden kavramı ve (5) konuşma taklidi. Eğitsel Değerlendirmeyi yapmak için gereken süre 10-20 dakika arasında değişmektedir.

*Çocuğun oturma davranışı* önemli etkisinden dolayı, diğer dört alanla eşit düzeyde ağırlığa sahiptir (en yüksek puan 12). Öğrencilerin, kaçmak için çaba göstermeden sandalyelerinde oturma davranışı sergilemeleri testin sonraki evrelerinin uygulanabilmesi açısından esastır. Bu davranışı sergileyen öğrenciler 12 puan almaktadırlar.

*Alıcı dil alanı* 12 maddeden oluşmaktadır. Her bir madde 1 veya 0 olarak puanlanmaktadır. Puanlamada 1 "evet davranış ortaya çıktı" ve 0 "hayır davranış ortaya çıkmadı" şeklinde puanlanmaktadır. Bu alandaki 12 madde, öğrencinin anlama ve farklı işitsel ve sözel uyarıcılara tepki verme yeterliliğinin incelenmesini sağlamaktadır.

*İfade edici dil alanında* 12 madde yer almaktadır. Bu maddeler, çocuğun yakın çevresiyle ilgili sorulara ilişkin ifade edici tepkilerini (örneğin, sözel ifadeler, işaretler ve jestler) ortaya koymak amacıyla tasarlanmıştır.

*Beden kavramı alanındaki* 12 madde, motor taklit becerilerini ve bedeni tanımlayıcı bilgileri değerlendirmek amacıyla kullanılmaktadır. Bu alanda çocuğun öğretim yoluyla ya da çevreden dolaylı olarak edindiği kendine ilişkin bilgilerinin düzeyini incelenmek amaçlanmaktadır.

*Konuşma taklidi alanında* da 12 madde yer almaktadır. Bu maddeler öğrencinin farklı türlerdeki seslendirmeleri telaffuz etme yeterliğini değerlendirmektedir. Bu maddelerden elde edilen bilgiler eğitsel programlamaya başlamak için kullanılmaktadır.

Uygulamada uygulayıcı ve çocuk karşılıklı oturarak eğitsel performans alma işlemi gerçekleştirilmektedir. Araç setinin içerisinde oyuncak araba, bardak, oyun küpü, yaklaşık 9 cm uzunluk ve genişlikte renkli kare kâğıtlar (kırmızı, yeşil ve sarı), plastik küçük bir kaşık ve selofan parçası bulunmaktadır. Uygulama için ayrıca kurabiye ve meyve suyu gibi birincil pekiştireçler ve çocuğun sevdiği bir oyuncak kullanılmaktadır. Çocuğa uyarılar yönergede belirtilen sırada sunulmakta ve çocuğun doğru tepkileri için "1" yanlış tepkileri için "0" verilerek değerlendirme gerçekleştirilmektedir.

EGİDEKAF'ın özgün standardizasyon çalışmasında 80 uygulamacı ve 132 katılımcıdan elde edilen verilerle standardizasyon işlemi yapılmıştır. Katılımcılar 41 OSB olan çocuk ve 91 ZY tanısı olan çocuktan oluşmaktadır. Çalışmada örneklemin yaş aralığı ve herhangi bir ek tanılarının olup olmadığı rapor edilmemiştir. Türkçe formun psikometrik niteliklerinin belirlenmesi çalışmasında ASIEP-3'te yer alan diğer informel araçların örneklemindeki yaş aralığı esas alınarak çalışma grubu oluşturulmuştur.

Güvenirlilik boyutunda değerlendiriciler arası güvenirlilik ve test-tekrar test güvenirliliğine ilişkin analizler yapılmıştır. Test-tekrar test güvenirliliği 11 OSB olan öğrenci ve uygulayıcısıyla gerçekleştirilmiştir. Beş alt boyutun test-tekrar test güvenirliliğine ilişkin korelasyon katsayıları şu şekildedir. Oturması alt boyutunda 1.00, alıcı dil alt boyutunda .93, ifade edici dil alt boyutunda .95, beden kavramı alt boyutunda .99, konuşma taklidi alt boyutunda .97 ve eğitsel değerlendirme toplam puanında .99' dur. Değerlendiriciler arası güvenirlilik katsayısı 432 veri noktasında 9 denek ve 18 gözlemci için %100 bulunduğu belirtilmiştir.

Geçerlik boyutunda literatür ve uzman görüşüne dayalı kapsam geçerliğinin yanı sıra yapı geçerliği ve ölçüt bağımlı geçerlik analizleri gerçekleştirilmiştir. EGİDEKAF'a ilişkin yapı geçerliği kapsamında yapılan analizlerde OSB olan çocukların OSB olan olmayan ZY tanısı olan çocuklara göre daha düşük alıcı dil becerisine, ( $F(1, 130) = 18.3, p < .01$ ), düşük ifade edici dil becerisine ( $F(1, 130) = 10.5, p < .01$ ), düşük beden kavramı becerisine ( $F(1, 130) = 9.0, p < .01$ ) ve düşük konuşma taklidi becerisine ( $F(1, 130) = 7.63, p < .01$ ) sahip oldukları rapor edilmiştir. OSB olan grubun ayrıca anlamlı düzeyde düşük eğitsel değerlendirme toplam puanına sahip oldukları bildirilmiştir ( $F(1, 130) = 16.5, p < .01$ ). Çocuğun oturması alt boyutunda puan anlamlı düzeyde farklılaşmamıştır ( $p < .05$ ). Ölçüt bağımlı geçerlik amacıyla EGİDEKAF ile Hedrick, Prather ve Tobin (1975) tarafından geliştirilmiş olan İletişim Gelişimi Sıralı Envanteri (*Sequenced Inventory of Communication Development/SCID*) dil puanları arasındaki korelasyonların incelenmesi yapılmıştır. EGİDEKAF ile SCID puanı arasında anlamlı düzeyde olumlu ilişki bulunduğu rapor edilmiştir ( $r = .75, p < .01$ ). Bir diğer anlamlı ilişki Eğitsel Değerlendirme ve Öğrenme Hızı Öngörüsü kayıt formunun A basamağının ölçüt için yanıtları alt testiyle bulunmuştur ( $r = -.78, p < .01$ ), (Krug ve diğ., 2008).

*Öğrenme Hızı Öngörüsü Kayıt Formu (ÖHÖKAF)*: ÖHÖKAF aracı öğrencilerin iki basamaklı, siyah-beyaz sıralı öğrenme etkinliğini öğrenmek için gerekli olan tepkileri edinme düzeylerini değerlendirmektedir. Etkinlik, herhangi bir öğrenci için yeni olması amacıyla tasarlanmıştır ve her hangi bir önkoşul öğrenmeye gereksinim yoktur. Öğrenci bu etkinliği, doğru tepkilerinin birincil pekiştireçlerle pekiştirilmesi yoluyla öğrenmektedir. Yanlış tepkilerde hata düzeltmeleri ipucu verilmesi yoluyla yapılmaktadır. ÖHÖKAF'da ağır düzeyde yetersizliği olan çocuklarla çalışan birçok öğretmen tarafından kullanılmakta olan doğrudan öğretim-ayrık denemeler formatı kullanılmaktadır. Bu araç, fiziksel olarak plastik bir fişi alabilen veya en azından fişi almaya çabalayan çocuklarda uygulanabilmektir. Ellere şekil vermektan uyarana karşı aşırı seçicilikle ilgili son teste kadar çeşitli öğretme aşamaları vardır.

Bir çocuğun öğrenme hızı, bu araç için gerekli süreyi de ortaya koyacaktır. Bir oturum, ölçüt karşılanmadıkça, 48 denemeden oluşur ve öğrenme etkinliğiyle ilgili ek adımlar tamamlanmadıkça bir okul gününde yaklaşık olarak 15 dakika sürmektedir. Çocuğun öğrenme hızına dayalı olarak 15 dakikalık ek oturumlarla isteğe bağlı öğrenme etkinlikleri verilebilmektedir ancak bu tarama için gerekli değildir. ÖHÖKAF'yi uygulayabilmek için gereken malzemeler, tepki tepsisi, bir beyaz kare fiş, bir beyaz daire fiş, bir siyah daire fiş, yiyecek pekiştireçleri, kalem ve yapışkan banttır.

Öğrenme hızı etkinliğindeki amaç, çocuğa, tepsiye her zaman ilk olarak daire şeklindeki siyah plastik fişi (daire şeklinde siyah fiş ve kare şeklinde beyaz fiş kullanılacaktır) koymasını öğretmektir. Bu, belirli sözel yönergeler veya fiziksel ipuçları vermeksizin çocuğun doğru seçimi yapmasını pekiştirerek yapılmaktadır. Uygulayıcı “*Fişi tepsiye koy*” yönergelerini verir. Eğer çocuk doğru tepkiyi verirse, uygulayıcı bir pekiştireç sağlar. Uygulayıcı ölçüt karşılanana kadar kaç deneme yapıldığını kaydetmektedir. Eğer çocuk Adım A’da (rastgele pozisyon I), ölçütü karşılar (ardı ardına 15 doğru), uygulayıcı bir Son Test uygulayarak çocuğun hem şekil hem de renk özelliklerini öğrenip öğrenmediğini belirleyebilmektedir. Öğrenciye “*Fişi tepsiye koy*” yönergesi verilerek ve daire şeklindeki beyaz fiş ve kare şeklindeki beyaz fişi sunarak, öğrencinin şekil özelliklerini öğrenip öğrenmediğini test edebilmektedir. Daha sonra “*Fişi tepsiye koy*” yönergesi verilerek ve daire şeklindeki siyah fişi ve daire şeklindeki beyaz fişi göstererek öğrencinin renk özelliklerini öğrenip öğrenmediğini test edebilmektedir.

ÖHÖKAF’nun özgün standardizasyon çalışmasında 81 uygulamacının gerçekleştirdiği uygulama çalışmalarında 124 katılımcıdan elde edilen verilerle standardizasyon işlemi yapılmıştır. Katılımcıların OSB ve ZY tanısı olan çocuklardan oluştuğu belirtilmiş ancak tanı kategorilerinde yer alan çocuk sayıları ve örneklemin yaş aralığı rapor edilmemiştir. Türkçe formun psikometrik niteliklerinin belirlenmesi çalışmasında ASIEP-3’te yer alan diğer informal araçların örneklemindeki yaş aralığı esas alınarak çalışma grubu oluşturulmuştur. Çalışmada eğitim öncesi ellere şekil verme ve A basamağında uygulamalar yapılarak aracın standardizasyonun gerçekleştirildiği belirtilmiştir. Özgün çalışmada aracın güvenilirliğine ilişkin herhangi bir istatistiksel analiz bilgisi rapor edilmemiş olup bu çalışmada geçerlik kapsamında literatür ve uzman görüşüne dayalı kapsam geçerliğinin yanı sıra yapı geçerliği bulguları rapor edilmiştir. Yapılan Kay-kare analizinin sonucunda ZYOSB olan öğrencilere göre çok daha az OSB olan çocuğun ölçütü karşıladıkları görülmüştür ( $\chi^2 = 9.05, p < .01$ ). ÖHÖKAF’nin A basamağıyla, EGİDEKAF toplam puanı arasında da güçlü bir ilişki bulunduğu rapor edilmiştir ( $r = -.78, p < .01$ ). Bu bulguya göre A basamağında EGİDEKAF da en düşük dil puana sahip katılımcılar öğrenme için en çok denemeye gereksinim duymaktadırlar (Krug ve diğ., 2008).

*Türkçe Erken Dil Gelişim Testi (TEDİL)*: TEDİL, Hresko, Reid ve Hammill (1999) tarafından geliştirilmiş olan TELD-3 isimli, 2-8 yaş arasında çocukların alıcı ve ifade edici dil becerilerini ölçme aracının, Türkçe uyarlamasıdır. TEDİL test seti şu bileşenlerden oluşmaktadır; TEDİL resim kitabı, Form A Uygulayıcı Kayıt Formu, Form B Uygulayıcı Kayıt Formu, TEDİL araç gereçleri (Bebek, bebek ayakkabısı, oyuncak araba, küçük top, kaşık, 5 tane bozuk para (kuruş/lira), 5 tane küp, uygulamacı kullanım kılavuzu).

TEDİL, A ve B olmak üzere iki paralel ölçme setinden oluşmaktadır. Her bir ölçme setinin resim kitapçıkları ve uygulama kayıt formları ayrıdır. Her bir set alıcı ve ifade edici olmak üzere iki alt test içermektedir. Her bir paralel formda dilin anlam bilgisi ve dilbilgisi alanlarını ölçen 76 madde bulunmaktadır. Bazı maddelerin alt şıkları bulunmaktadır. Her bir alıcı dil alt testi 37 madde içermektedir. Form A anlam bilgisini ölçen 24, madde bilgisini ölçen 13 maddeden oluşmaktadır. İfade edici alt dil testinin her iki formunda da 39’ar madde bulunmaktadır. Form A, 22 anlam bilgisi ölçen madde ve dilbilgisi ölçen 17 maddeden oluşmaktadır. Araştırmada ölçüt bağımlı geçerlik analizleri gerçekleştirmek amacıyla TEDİL-A formu kullanılmıştır.

Testin rapor edilen Cronbach Alfa katsayıları alıcı dil A ve B formları için .86 ile .98 arasında değişmektedir. İfade edici dil A formu .87 ile .98 arasında, B formu .86 ile .98 aralığındadır. Test tekrar test güvenilirliği Alıcı dil A formu için .96, B formu için .93; ifade edici dil A formu için .89, B formu için .83’dür (Güven, 2009; Güven ve Topbaş, 2009).

## **İşlem Süreci**

Araştırmaya başlanmadan önce psikometrik nitelikleri belirlenmesi amaçlanan araçların telif haklarına sahip olan ABD’deki kuruluştan yasal izinler alınmıştır. Çalışma kapsamında yer alacak araçların kullanımlarını içeren el kitabının çevirisi yapılarak her bir aracın nasıl kullanılacağına ilişkin detaylı bilgi elde edilmiştir. Araçların özgün formlarının çeviri ve yeniden çevirisi yoluyla Türkçeye çeviri işlemi gerçekleştirilmiştir. Türkçeye çeviri uygulamasından sonra her bir araç için gerekli olan uygulama materyalleri temin edilmiştir. Uygulamalar, bu kurumlarda çalışan ve araştırmaya katılmaya gönüllü olan öğretmenlere araçların tanıtımı, uygulanışı ve verilerin



toplanaşına ilişkin eğitimler verilmesi ve öğretmenlerin uygulamaları yaparak verileri toplaması, yoluyla gerçekleştirilmiştir. Her bir araç için ortalama bir saat süren eğitimler yapılmıştır. Bu eğitimlerde önce araçların amacı, kullanımı ve puanlanması anlatılmıştır. Daha sonra uygulamacı model olarak uygulamanın nasıl yapılacağını göstermiştir. Son evrede tüm öğretmenlerin aracın uygulanmasına ilişkin görüşlerini ve sorularını paylaştıkları etkileşimler gerçekleştirilmiştir. Eğitim oturumları öğretmenler aracın kullanımına ilişkin yeterli düzeyde bilgi sahibi olduklarını ifade edinceye kadar devam ettirilmiştir. Uygulamalar esnasında ise öğretmenler gözlemlenmiş ve çalışmada gözlemciler arası güvenilirlik verisi toplanmıştır.

Araştırmada veri toplama uygulamasını gerçekleştiren 33 öğretmenle her bir araç için 3 kez olmak üzere toplam 6 uygulamacı eğitimi toplantısı gerçekleştirilmiştir. EGİDEKAF için 126 (özgün çalışmada 132) ve ÖHÖKAF için de 131 (özgün çalışmada 124) olmak üzere toplam 257 oturumda performans verisi toplanmıştır. Ölçüt bağımlı geçerlik için 31 öğrenciden de TEDİL verisi alınmıştır. Ayrıca EGİDEKAF için 31 test-tekrar test verisi alınmıştır. Tüm testler birlikte alındığında bu araştırmada 319 ölçme işlemi gerçekleştirilmiştir.

### **Verilerin Analizi**

Verilerin analizine geçilmeden önce her bir veri formu incelenerek hatalı doldurulan veri formu olup olmadığı kontrol edilmiştir. Araştırmacı veri toplama sürecinde uygulamayı gerçekleştiren öğretmenlerle sürekli iletişim içinde olduğundan performans verilerinde herhangi bir eksikliğin olmadığı görülmüştür. Bazı veri formlarında çocukların doğum tarihleri ve cinsiyetleri bölümünün doldurulmadığı görülmüş bu formlar isimlerden kontrol edilerek kayıtlarda var olanlar araştırmacı tarafından olmayanlar da kurum yetkilileriyle ilişkiye geçilerek tamamlanmıştır. Bazı formlarda verilerin toplandığı ancak veri toplama formunun ön yüzünde toplam puanların işlenmediği görülmüş ve bu türden hesaplama işlemleri gerçekleştirilerek formlar bilgisayar ortamında istatistik programının veri setine işlenmeye hazır hale getirilmiştir.

İstatistik programında oluşturulan veri dosyaları üzerinde gerçekleştirilen analizlerde parametrik istatistikler yapılmadan önce ilgili istatistiklerin ön koşulu olan varsayımlar test edilmiştir. Araştırma kapsamında güvenilirlik ve geçerlikleri incelenen iki aracın ilgili analizlerinde ilk olarak katılımcıların betimleyici analizleri yapılmıştır. Araçların iki hafta arayla alınan test-tekrar test güvenilirlikleri için Pearson korelasyon katsayıları hesaplanmıştır. Değerlendiriciler arası güvenilirlik için gözlemciler arası güvenilirlik oranının yüzdelik olarak hesaplanması işlemleri gerçekleştirilmiştir. Geçerlik analizlerinde alt boyutlar arası korelasyon katsayıları Pearson korelasyon katsayısı ile hesaplanmıştır. Ölçeğin OSB olan çocuklarla OSB olan olmayan zihinsel yetersizliği olan çocukları ayırt etme düzeyleri t-testi kullanılarak gerçekleştirilmiştir. Ayrıca ölçeklerin ölçüt bağımlı geçerlikleri Pearson korelasyon katsayısı hesaplanması yoluyla analiz edilmiştir. İstatistiksel hesaplamaların yapılmasında SPSS paket programından yararlanılmıştır.

## **BULGULAR**

Bu bölümde bu araştırma kapsamında psikometrik nitelikleri belirleme çalışması gerçekleştirilen EGİDEKAF ve ÖHÖKAF informel değerlendirme araçlarından elde edilen puanların güvenilirlik ve geçerliklerine ilişkin yapılan analizlerin bulgularına yer verilmiştir.

### **EGİDEKAF'dan Elde Edilen Puanlara İlişkin Güvenirlik-Geçerlik Bulguları**

Özgün çalışmada EGİDEKAF güvenilirliğini incelemek üzere test-tekrar test ve gözlemciler arası kayıt güvenilirliği yöntemleri kullanılmıştır. Geçerlik boyutunda ise ölçüt bağımlı geçerlik ve yapı geçerliği yöntemleri kullanılmıştır. Bu çalışmada da test-tekrar test, ölçüt bağımlı geçerlik ve yapı geçerliği yöntemleri kullanılarak aracın Türkçeye uyarlanmış formunun güvenilirliği ve geçerliği incelenmiştir. EGİDEKAF'da eğitsel değerlendirme toplam puanına 12 puan katkı sağlayan çocuğun oturması alt boyutuna ilişkin bulgular rapor edilmemiştir. Çocuğun oturması değerlendirmenin önkoşulu olarak tanımlanmıştır ve tüm çocuklar bu alt boyuttan tam puan almışlardır. Bu nedenle bu alt boyutun istatistiksel olarak açıklayıcılığı yoktur.

*Kararlılık katsayısı.* Kararlılık katsayısı için test-tekrar test yöntemi ile 31 katılımcıdan 2 hafta arayla elde edilen veriler üzerinden Pearson korelasyon katsayısı hesaplanmıştır. Her bir kategori

için elde edilen test-tekrar test korelasyon katsayıları şu şekildedir: Alıcı Dil,  $r = .91$   $p < .01$ , İfade Edici Dil,  $r = .91$   $p < .01$ , Beden Kavramı  $r = .91$   $p < .01$ , Konuşma Taklidi  $r = .93$   $p < .01$  ve Eğitimsel Toplam Puan  $r = .98$   $p < .01$ 'dir. Özgün çalışmada 11 öğrenciyle gerçekleştirilen test-tekrar test analizine ilişkin bulgular ise şu şekilde rapor edilmiştir. Alıcı Dil,  $r = .93$   $p < .01$ , İfade Edici Dil,  $r = .95$   $p < .01$ , Beden Kavramı  $r = .99$   $p < .01$ , Konuşma Taklidi  $r = .97$   $p < .01$  ve Eğitimsel Toplam Puan  $r = .99$   $p < .01$ .

*Madde-toplam puan korelasyonu.* Özgün çalışmada yapılmamış olmakla birlikte bu çalışmada elde edilen veriler üzerinden EGİDEKAF'ın iç tutarlığı için hesaplanan Cronbach Alfa katsayısı  $\alpha = .87$ 'dir. Eğitsel değerlendirme toplam puanını oluşturan dört maddenin, madde toplam korelasyon değerleri Alıcı Dil .72, İfade Edici Dil .66, Beden Kavramı .79 ve Konuşma Taklidi .80 olarak hesaplanmıştır.

*Ölçüt bağımlı geçerlik.* EGİDEKAF'dan elde edilen puanların ölçüt bağımlı geçerliğini incelemek amacıyla TEDİL aracı uygulanan çocukların bu araçlardan aldıkları puanlarla EGİDEKAF'ın alt boyutlarından aldıkları puanlar arasındaki ilişkiye Pearson korelasyon katsayısı hesaplanarak bakılmıştır. TEDİL ve EGİDEKAF araçlarının birlikte uygulandığı 31 çocuktan elde edilen veriler üzerinden yapılan incelemede TEDİL alıcı ve ifade edici dil puanları ile EGİDEKAF'ın alt boyutları arasındaki korelasyon değerleri sırasıyla şu şekildedir: TEDİL Alıcı Dil puanı ile EGİDEKAF Alıcı Dil  $r = .72$   $p < .01$ , EGİDEKAF İfade Edici Dil  $r = .49$   $p < .01$ , EGİDEKAF Beden Kavramı  $r = .51$   $p < .01$ , EGİDEKAF Konuşma Taklidi  $r = .55$   $p < .01$ , ve EGİDEKAF Eğitsel Değerlendirme toplam puanı  $r = .62$   $p < .01$ . TEDİL İfade Edici Dil puanı ile EGİDEKAF Alıcı Dil  $r = .54$   $p < .01$ , EGİDEKAF İfade Edici Dil  $r = .74$   $p < .01$ , EGİDEKAF Beden Kavramı  $r = .53$   $p < .01$ , EGİDEKAF Konuşma Taklidi  $r = .61$   $p < .01$  ve EGİDEKAF Eğitsel Değerlendirme toplam puanı  $r = .66$   $p < .01$ .

Özgün çalışmada ölçüt bağımlı geçerlik için *Vineland Uyumsal Davranış Ölçeği: Görüşme Versiyonu* nun alıcı dil, ifade edici dil, iletişimsel alan ve günlük yaşam becerileri puanlarıyla EGİDEKAF puanlarının karşılaştırması yapılmıştır. Rapor edilen korelasyon katsayıları .47 ile .71 aralığındadır.

EGİDEKAF'ın özgün çalışmasında EGİDEKAF'ın eğitsel değerlendirme toplam puanı ile ÖHÖKAF'ın A basamağındaki toplam tepki sayısı arasındaki ilişkiye Pearson korelasyon katsayısı kullanılarak bakılmış ve bulguları yapı geçerliği başlığı altında verilmiştir. Bu değer  $r = -0.78$ ,  $p < .01$  olarak rapor edilmiştir. Aynı ilişkiye bu çalışmada da bakılmış ancak bu değer ölçüt bağımlı geçerlik başlığı altında rapor edilmesi uygun görülmüştür. Türkçeye uyarlama çalışmasında yapılan analizin bulgusu  $r = -0.53$ ,  $p < .01$ 'dir. EGİDEKAF'dan alınan toplam puanlar düştükçe ÖHÖKAF A basamağında verilen öğrenme etkinliğinde ölçütü karşılamak için ihtiyaç duyulan deneme sayısı artmaktadır. İki araç arasındaki negatif yönlü bu korelasyonun ölçüt bağımlı geçerlik bulgusu olarak değerlendirilmesinin uygun olacağı düşünülmüştür.

*Yapı geçerliği. Alt gruplar arası farklılık.* EGİDEKAF'ın yapı geçerliği boyutunda özgün ölçekte de olduğu gibi EGİDEKAF'ın tanımlanmış davranış kategorilerine göre OSB olan ve ZY tanısı olan çocukları anlamlı düzeyde ayırt edebilme derecesi bağımsız örneklem t-testi analizi yoluyla incelenmiştir. T-testi sonuçları alıcı dil,  $t(121) = 6.24$ ,  $p < .01$ , ifade edici dil,  $t(121) = 5.86$ ,  $p < .01$ , beden kavramı,  $t(121) = 7.02$ ,  $p < .01$ , konuşma taklidi,  $t(121) = 7.28$ ,  $p < .01$  ve eğitsel değerlendirme toplam puanı için  $t(121) = 8.56$ ,  $p < .01$  şeklindedir. T-testi bulguları, OSB olan çocukların ZY tanısı olan çocuklara göre daha düşük alıcı dil becerisine ( $X_{(OSB)} = 6.44$ ,  $X_{(ZY)} = 9.26$ ), daha düşük ifade edici dil becerisine ( $X_{(OSB)} = 3.21$ ,  $X_{(ZY)} = 7.47$ ), daha düşük beden kavramına ( $X_{(OSB)} = 5.49$ ,  $X_{(ZY)} = 9.65$ ) ve daha düşük konuşma taklidi becerisine ( $X_{(OSB)} = 4.36$ ,  $X_{(ZY)} = 8.94$ ) ve daha düşük eğitsel değerlendirme toplam puanına ( $X_{(OSB)} = 30.71$ ,  $X_{(ZY)} = 47.22$ ) sahip olduklarını göstermektedir. Özgün çalışmada da tüm alt boyutlarda aynı manidarlık düzeyinde farklılık görüldüğü rapor edilmiştir. Bu çalışmadaki veriler özgün çalışmayla örtüşmektedir. OSB tanı ölçütleri göz önünde bulundurulduğunda OSB olan çocukların iletişim boyutunda düşük beceri göstermeleri bulguları destekler niteliktedir.

*Alt testler arası korelasyonlar.* Özgün çalışmada yapılmamış olmakla birlikte bu çalışmada, geçerlik incelemesi amacıyla, eğitsel değerlendirme toplam puanını oluşturan alt testler arasındaki korelasyonlar hesaplanmıştır. EGİDEKAF'ın alt testleri arasındaki korelasyon katsayıları şu şekildedir. Alıcı Dil alt testi; İfade Edici Dil ile  $r = .47$   $p < .01$ , Beden Kavramı ile  $r = .74$   $p < .01$ , Konuşma Taklidi ile  $r = .71$   $p < .01$ , Eğitsel Değerlendirme toplam puanı ile  $r = .85$   $p < .01$ 'dir. İfade Edici Dil alt

testi; Beden Kavramı ile  $r = .62$   $p < .01$ , Konuşma Taklidi ile  $r = .65$   $p < .01$ , Eğitsel Değerlendirme toplam puanı ile  $r = .60$   $p < .01$ 'dir. Beden Kavramı alt testi; Konuşma Taklidi ile  $r = .71$   $p < .01$ , Eğitsel Değerlendirme toplam puanı ile  $r = .86$   $p < .01$ 'dir. Konuşma Taklidi ile Eğitsel Değerlendirme toplam puanı  $r = .92$   $p < .01$ 'dir.

### **Eğitsel değerlendirme toplam ham puanlarının yüzdelerle dilimlerine çevrilmesi**

Özgün çalışmada EGİDEKAF aracından elde edilen eğitsel değerlendirme toplam ham puanlarının örneklem grubundaki dağılımına karşılık gelen yüzdelerle dilimler ham puanların kümülatif frekanslarının alınması ve oluşturulan puan aralıklarına göre frekans ortalamalarının hesaplanması yoluyla oluşturulmuştur. Bu yolla, değerlendirilen çocukların aldıkları ham puanın örneklem grubu yüzdelerle dilimdeki yeri belirlenebilmektedir. Ham puanların yüzdelerle dilimlere çevrilmesi yoluyla çocuğun eğitsel değerlendirme puanı örneklem grubuyla karşılaştırılabilmektedir. Bu çalışmada da elde edilen veriler üzerinden gerçekleştirilen yüzdelerle değerlere ilişkin sonuçlar Tablo 2'de verilmiştir.

**Tablo 2.** EGİDEKAF' da Eğitsel Değerlendirme Toplam Ham Puanlarının Yüzdelerle Dilimlerine Çevrilmesi.

Eğitsel Değerlendirme Toplam Ham Puanı	%	Eğitsel Değerlendirme Toplam Ham Puanı	%
60-59	< 9	23-19	58
58-54	10	18-14	67
53-49	11	13-9	79
48-44	25	8-4	93
43-39	36	3-0	98
28-24	49		

### **ÖHÖKAF' dan Elde Edilen Puanlara İlişkin Güvenirlik-Geçerlik Bulguları**

Özgün çalışmada ÖHÖKAF güvenirliliğine ilişkin herhangi bir analiz bulgusu rapor edilmemiştir. Geçerlik boyutunda ise yapı geçerliği boyutunda ÖHÖKAF uygulanan OSB' li ve ZY' li çocuklar öğrenme etkinliğinde ölçütü karşılayıp karşılayamama üzerinden Kay-kare analizi ile karşılaştırılmışlardır. Bu çalışmada güvenirlik boyutunda değerlendiriciler arası güvenirlik ve geçerlik boyutunda Kay-kare analizinin yanı sıra öğrencilerin A basamağındaki doğru tepki sayıları üzerinden ayırt ediciliği incelemek için t-testi gerçekleştirilmiştir. Ek olarak alt boyutlar arasında korelasyon analizi yapılmıştır.

*Güvenirlik. Değerlendiriciler arası güvenirlik.* Değerlendiriciler arası güvenirlik verisi için 15 uygulamada uygulamacılarla birlikte veri toplanmış ve ellere şekil verme ve A basamağı (Rastgele Pozisyon I) evrelerinde iki kaydın örtüşme yüzdelerine bakılmıştır. Her iki evrede de gözlemler arası güvenirlik % 100 bulunmuştur. Özgün çalışmada herhangi bir güvenirlik bulgusu rapor edilmemiştir.

*Yapı geçerliği. Alt gruplar arası farklılık.* Özgün çalışmada olduğu gibi bu çalışmada da OSB olan ve ZY' li çocuklar öğrenme etkinliğinde ölçütü karşılayıp karşılayamama üzerinden Kay-kare analizi ile karşılaştırılmışlardır. Yapılan Kay-kare analizinin bulgusu ( $\chi^2$  (sd=1, n=131) = 37.96,  $p < .01$ ) şeklindedir. Bulgu incelendiğinde OSB olan çocukların ÖHÖKAF A basamağında ölçütü karşılama düzeylerinin ZY tanısı olan çocuklara göre anlamlı düzeyde düşük olduğu görülmektedir.

Özgün çalışmada rapor edilmemiş olmakla birlikte yapı geçerliği kanıtı olarak ÖHÖKAF' nin OSB' li ve ZY' li çocukları ayırt edebilme derecesi bağımsız örneklem t-testi kullanılarak da incelenmiştir. Bu amaçla çocukların tanı grupları ile A basamağında verdikleri doğru tepki sayısı karşılaştırılmıştır. Bu karşılaştırmanın yapılma nedeni bir çocuğun, ÖHÖKAF' nin A basamağında, ölçütü karşılayabilmesi için ardı ardına 15 kez doğru tepkide bulunma zorunluluğudur. OSB olan ve ZY tanısı olan çocukların doğru tepki sayıları arasında anlamlı fark bulunmaması durumunda aracın bu iki grubu sadece aynı tepkiyi ardı ardına 15 kez verip vermeme boyutunda ayırt ettiği düşünülebileceğinden doğru tepki sayıları üzerinden bir karşılaştırmayla ayırt ediciliğin daha etkili bir biçimde analiz edilmesi amaçlanmıştır. Yapılan analizde Levene testi sonuçları incelendiğinde

OSB olan çocukların ÖHÖKAF' nin A basamağında verdikleri doğru tepki sayısının varyansı ile ZY' li çocukların ÖHÖKAF' nin A basamağında verdikleri doğru tepki sayısının varyanslarının homojen olmadığı görülmüştür ( $F = 9.63, p < .01$ ). Bu nedenle varyansların eşit olmadığı varsayımında hesaplanan t değeri dikkate alınmıştır. Yapılan t-testi sonucu  $t(129) = 5.15, p < .01$ 'dir.

T-testi sonuçları incelendiğinde OSB olan çocukların ÖHÖKAF' nin A basamağında verdikleri doğru tepki sayıları ortalamaları ( $X_{(OSB)} = 15.5$ ) ile ZY' li çocukların ÖHÖKAF' nin A basamağında verdikleri doğru tepki sayıları ortalamaları ( $X_{(ZY)} = 23.07$ ) arasındaki farkın istatistiksel olarak manidar olduğu görülmüştür. Bu doğrultuda ÖHÖKAF' nin A basamağında verilen doğru tepkiler boyutunda OSB olan çocukların doğru tepki sayıları ortalamalarının ZY' li çocukların doğru tepki sayıları ortalamalarından anlamlı düzeyde düşük olduğu görülmüştür.

*Alt boyutlar arası korelasyon.* Özgün çalışmada yapılmamış olmakla birlikte aracın geçerlik kanıtlarını güçlendirmek amacıyla ellere şekil verme aşamasındaki doğru tepki yüzdesi, A basamağındaki doğru tepki yüzdesi, son test A doğru tepki yüzdesi ve son test B doğru tepki yüzdesi kategorileri arasındaki korelasyonlara bakılmıştır. Bu analizin yapılmasındaki amaç aracın kendi içinde benzer bir yapıyı ölçüp ölçmediğini kontrol edebilmektir. Dört farklı basamaktaki yüzdeler puanlar arasında anlamlı korelasyonun olması her basamağın benzer bir yapıya ya da aynı yapının ilişkili boyutlarını ölçüyor olduğunun kanıtı olarak düşünülmüştür. Bu puanlar arasındaki korelasyonlar Pearson korelasyon katsayısı kullanılarak hesaplanmıştır. Dört alt boyut arasındaki korelasyon katsayıları Tablo 7'de verilmiştir. Ellere Şekil Verme alt boyutu; A Basamağında Doğru Tepki ile  $r = .78, p < .01$ , Son Test A Doğru Tepki ile  $r = .62, p < .01$ , Son Test B Doğru Tepki ile  $r = .64, p < .01$ 'dir. A Basamağında Doğru Tepki alt boyutu; Son Test A Doğru Tepki ile  $r = .83, p < .01$ , Son Test B Doğru Tepki ile  $r = .74, p < .01$ 'dir. Son Test A Doğru Tepki ile Son Test B Doğru Tepki  $r = .85, p < .01$ 'dir.

### **Adım A için öğrenme hızı öngörüsü doğru tepki sayılarının yüzdeler dilimlerine çevrilmesi**

Özgün çalışmada ÖHÖKAF aracının A basamağında ölçütü karşılamak için gereksinim duyulan tepki sayısının örneklem grubundaki dağılımına karşılık gelen yüzdeler dilimleri oluşturulmuştur. Bu yolla değerlendirilen çocuğun ÖHÖKAF aracının A basamağında ölçütü karşılamak için gereksinim duyduğu tepki sayısının örneklem grubu yüzdeler dilimindeki yeri belirlenebilmektedir. Ham puanların yüzdeler dilimlere çevrilmesi yoluyla çocuğun ölçütü karşılamak için gereksinim duyduğu tepki sayısı örneklem grubuyla karşılaştırılabilmektedir. Bu çalışmada da elde edilen veriler üzerinden de bu işlem gerçekleştirilmiştir. Sonuçlar Tablo 3'de verilmiştir.

**Tablo 3. ÖHÖKAF Adım A İçin Öğrenme Hızı Öngörüsü Doğru Tepki Sayılarının Yüzdeler Dilimlerine Çevrilmesi.**

A Basamağı Toplam		A Basamağı Toplam	
Deneme Sayısı	%	Deneme Sayısı	%
15	4	32	27
16	5	33	28
17	6	34	29
18	8	35	30
19	9	36	33
20	11	37	34
21	12	38	35
22	13	39	35
23	13	40	39
24	14	41	44
25	18	42	49
26	20	43	53
27	22	44	59
28	23	45	69
29	24	46	77
30	25	47	86
31	26	48	100

## TARTIŞMA ve SONUÇ

Bu bölümde araştırma sonucunda elde edilen bulgular tartışılmış ve sonuçlar ortaya konulmuştur. Araştırmadan elde edilen bulgular tartışılmadan önce güvenilirlik ve geçerliğe ilişkin kısa bir kavramsal açıklama yapılmıştır.

Güvenirliğin, bireylerin test maddelerine verdikleri tepkiler olarak tanımlanabileceği belirtilmiştir (Büyüköztürk, 2010). Güvenirlik, testin ölçmeyi hedeflediği özelliği ne derece doğru olarak ölçtüğü ile ilgili bir kavram olup bir ölçme aracının güvenilirliği için iki temel ölçüt söz konusudur. Ölçme aracı farklı zamanlarda uygulandığında tutarlı sonuçları veriyorsa zaman boyutunda tutarlılıktan, tek bir uygulamadaki yanıtlayıcıların tepkileri birbiriyle tutarlıysa kendi içinde tutarlılıktan yani güvenilir bir ölçmeden söz edilebilir. Güvenirlikle geçerlik arasında yakın bir ilişki olduğu, güvenilirliğin yapısal geçerliğin bir yönünü oluşturduğu, güvenilirliğin geçerliğin ön koşulu olduğu ancak tek başına bir bilimsel araştırmada yeterli olmayacağı ve geçerlik çalışmalarının da yapılması gerektiği ifade edilmiştir (Şencan, 2005). Büyüköztürk (2010), bir testin güvenilirliğini incelemek için kullanılabilecek başlıca güvenilirlik türlerini şu şekilde sıralamıştır. Test-tekrar test güvenilirliği, paralel (eşdeğer) form güvenilirliği, iki yarı test güvenilirliği, Kuder Richardson ve Cronbach alfa güvenilirliği, madde-toplam puan korelasyonu ve ölçmenin standart hatası.

Geçerlik ise bir testin bireyin ölçülmek istenen özelliğini ne derece doğru ölçtüğüyle ilgili bir kavramdır. Bu bağlamda kullanılan ölçüm aracının ölçülmek istenen özelliğe uygun olması, verilerin ölçülmek istenen özelliğin niteliğini tam olarak yansıtması ve aynı zamanda verilerin amaca yönelik olarak yararlı olmasıdır. Bu nedenle geçerlik kısaca "test puanlarının sonuç çıkartmak için uygun, anlamlı ve yararlı olması" biçiminde tanımlanmıştır (Şencan, 2005). Bununla birlikte bir testten farklı sonuçlar çıkartılabilir ve belirli bir çıkarsamayı destekleyen verilerin toplanması için pek çok yol vardır. Verilerin çeşitli yollarla elde edilebilmesine karşılık, geçerliğin, daima verilerin test puanlarına dayalı çıkarsamaları ne ölçüde desteklediğini gösterdiği ve testin kendisinin değil, testin özel kullanımına ait çıkarsamaların geçerliğini ortaya koyduğu ifade edilmiştir (TPD, 1997). Bir testin geçerliğini belirlemede değişik sınıflandırmalar olduğu bu sınıflandırmalar içerisinde en çok Kapsam Geçerliği, Ölçüt-bağımlı Geçerlik ve Yapı Geçerliği sınıflandırmasının tercih edildiği belirtilmektedir (Büyüköztürk, 2010).

*Güvenirlik;* Bu çalışmada güvenilirliği incelemek amacıyla yapılan ilk istatistiksel analiz çalışması test-tekrar test güvenilirliği çalışmasıdır. Test-tekrar test güvenilirliği bir testin aynı gruba belirli aralıklarla iki kez uygulanmasıyla elde edilen puanlar arasındaki korelasyon ile açıklanmaktadır ve testin zamana bağlı olarak ne derece kararlı ölçümler verdiğini yorumlamak amacıyla kullanılmaktadır (Büyüköztürk, 2010).

EGİDEKAF aracı için 31 katılımcıdan 2 hafta arayla elde edilen veriler üzerinden Pearson korelasyon katsayılarının hesaplanması yoluyla test-tekrar test güvenilirliği analiz edilmiştir. Eğitsel değerlendirme toplam puanını oluşturan dört alt boyut ve toplam puan için hesaplanan korelasyon katsayıları .91 ile .98 arasında değişmektedir. Test-tekrar test güvenilirlik katsayısının .80 ve üzerinde olması gerektiği bildirilmiştir (Innes ve Straker, 2003). Bununla birlikte, 70'lik test-tekrar test güvenilirliğinin ya da güvenilirlik katsayısının da kabul edilebilir olduğunu gösteren kaynaklar mevcuttur (Büyüköztürk, 2010; Şencan, 2005). Alanda psikometrik özellikleri incelenmiş Türkçeye uyarlama çalışmaları için rapor edilen değerler incelendiğinde .72 ile .99 arası test-tekrar test korelasyonlarının güvenilirlik için kabul edilebilir değerler olarak rapor edildiği görülmektedir (Alev, 2011; Atamaz, Yağız On ve Durmaz, 2007; Colay, 2006; Colay, Topbaş ve Tanrıdağ, 2007; Diken, Ardıç, Diken ve Gilliam, 2012; Düver, 2006; İncekaş, 2009; Kapçı, Küçüker ve Uslu, 2010; Korkmaz, 2009; Köse, Bora, Erermiş ve Aydın, 2010; March-Göçer, 1996; Maviş, Savaşır, Sezgin ve Erol, 2006; Yalaz, Anlar ve Bayaoğlu, 2010;). Bu bilgiler doğrultusunda EGİDEKAF için hesaplanan test-tekrar test güvenilirliği katsayısının aracın güvenilir kabul edilmesini sağlayacak düzeyde olduğu söylenebilir.

Değerlendiriciler arası ya da gözlemciler arası güvenilirlik katsayısı güvenilirlik analizi türlerinden eş değerlik analizlerinin bir alt türüdür. Paralel formlar güvenilirliğine benzer olarak aynı özelliği ölçmek üzere iki değerlendirmeci/gözlemcinin aynı performansı puanlamasına dayanmaktadır (Büyüköztürk, 2010; Şencan, 2005). İki değerlendirme ya da gözlem arasındaki ilişkinin derecesi örtüşme yüzdeleri ya da Pearson korelasyon katsayıları ile rapor edilebilmektedir. Özel eğitim uygulamalarında gözlemciler arası güvenilirliğin, özellikle tek denekli araştırmalarda,

genellikle yüzdellik türünde rapor edildiği görülmekte ve önerilen güvenilirliğin yüzde yüze yakın olmasının istendik olduğu ifade edilmektedir (Erbaş, 2012). ÖHÖKAF'ın psikometrik niteliklerinin belirlenmesi çalışmasında değerlendiriciler arası güvenilirlik araştırmacının 15 uygulamada uygulamacıyla birlikte aynı performansı gözleyerek puanlaması yoluyla elde edilmiş ve sonuç %100 bulunmuştur. ÖHÖKAF'ta ki performans görevi tek basamaklı bir öğrenme görevi olduğundan tepkinin doğru ya da yanlış olduğuna ilişkin karar oldukça açıktır. Bu nedenle özel eğitim meslek uzmanlarının bu tepkinin doğru olup olmadığına ilişkin karar vermede fikir ayrılığı yaşamaları oldukça nadir rastlanabilecek bir durum olarak değerlendirilebilir.

Madde-toplam puan korelasyonu test maddelerinden alınan puan ile testin toplam puanı arasındaki ilişkiyi açıklamaktadır. Madde-toplam korelasyonun pozitif ve yüksek olması maddelerin benzer davranışları örneklediğini ve testin iç tutarlılığının yüksek olduğunu gösterir. Ölçekler için hesaplanacak Cronbach alpha katsayısının .70 den yüksek olması gerektiği, maddelerin madde toplam korelasyon katsayısının da .30 dan düşük olmaması gerektiği belirtilmiştir. Bununla birlikte Cronbach alpha katsayısının ölçekteki madde sayısından etkilendiği, az sayıda (on maddeden az) madde içeren ölçeklerde bu değer düşük çıkabileceği bu nedenle az sayıda madde içeren ölçeklerde maddeler arası korelasyon katsayısı değerlerinin de rapor edilmesinin uygun olacağı ifade edilmiştir (Büyüköztürk, 2010; Pallant, 2001).

Ölçeğin özgün çalışmasında yapılmamış olmakla birlikte, ölçeğin iç tutarlılığına ilişkin görgül kanıt sağlaması, benzer olarak gerçekleştirilen birçok çalışmada yaygın olarak rapor edilmesi ve bir anlamda evrensel bir güvenilirlik katsayısı sağlaması nedenleriyle bu çalışmada EGİDEKAF için madde-toplam korelasyonu analizi yapılmıştır. EGİDEKAF aracının iç tutarlılığı için hesaplanan Cronbach alpha katsayısı  $\alpha = .87$ ' dir. Maddelerin her birinin madde-toplam korelasyonu da .66 ile .80 arasında değişmektedir. Ölçek geliştirme Türkçeye uyarlama çalışmalarına bakıldığında .32 ile .99 arası Cronbach alpha ve madde-toplam korelasyon katsayılarının rapor edildiği görülmektedir (Alev, 2011; Colay, 2006; Ege, Acarlar ve Turan, 2005; Kabil, 2005; Kapıcı, Küçükler ve Uslu, 2010; Köse, Bora, Erermiş ve Aydın, 2010; Maviş, Colay, Topbaş ve Tanrıdağ, 2007; Savaşır, Sezgin ve Erol, 2006).

EGİDEKAF aracında, eğitsel değerlendirme toplam puanını oluşturan dört maddeden, .66 ile en düşük madde-toplam puan korelasyonuna sahip olan ifade edici dil boyutunun toplam puana olan katkısının diğer maddelere göre düşük olması beden kavramı, konuşma taklidi gibi boyutlarda çocuğun sahip olduğu alıcı dil düzeyinin yönergeleri anlama ve yerine getirmede ifade edici dilden daha önemli olmasıyla açıklanabilir. Bulgular incelendiğinde de görülebileceği gibi alıcı dil, beden kavramı ve konuşma taklidi hem eğitsel değerlendirme toplam puanına sağladıkları katkı hem de birbirleriyle ve toplam puanla olan korelasyonları yönünden ifade edici dilden daha yüksek değerlere sahiptirler. Büyüköztürk (2010), madde-toplam korelasyonunu yorumlamada istatistiksel anlamlılığın da ölçüt olarak alınabileceğini ifade etmektedir. İfade edici dil boyutunun hem madde-toplam korelasyonu hem de diğer maddeler ve toplam puanla korelasyonun görece olarak daha az olmasına rağmen istatistiksel olarak .001 düzeyinde anlamlı bir değere sahiptir. Bu nedenle eğitsel değerlendirme toplam puanını oluşturan maddelerin tamamının beklenen düzeyde güvenilirlik değerine sahip olduğu söylenebilir.

*Geçerlik;* Psikometrik nitelikleri belirleme çalışması yapılan araçların geçerliğini incelemek üzere kullanılan ilk yöntem ölçüt bağımlı geçerlik yöntemidir. Ölçüt bağımlı geçerlik, test puanlarının belirlenen bir veya birkaç dış ölçütle ilişkisini inceleyen geçerlik tekniğidir. Bu geçerlik tekniği zamandaş (concurrent) geçerliği ve yordama (predictive) geçerliği olmak üzere ikiye ayrılmaktadır. Bu çalışma kapsamında da kullanılmış olan eşzaman geçerliğinde katılımcıların geliştirilmek istenen bir testten aldıkları puanlarla aynı davranışı ölçen eski bir test ya da ilişkili bir başka özelliği ölçen bir testten aldıkları puanlar arasındaki korelasyonlara bakılmaktadır (Büyüköztürk, 2010).

Şencan (2005), geçerlik katsayılarının güvenilirlik katsayıları kadar yüksek olmadığını ifade etmektedir. Uygulamada geçerlik ve tahmine ilişkin r değerinin .30 ile .50 arasında değiştiğinin görüldüğünü nadiren .60'ın üzerine çıktığını belirtmektedir. Bu nedenle geçerlik ve belirliliğe ilişkin .30 ile .50 arasında bir değer elde edilmişse testin geçerli olduğu sonucuna varılmaktadır. Ölçüt bağımlı karşılaştırmalarda "aynı" ve "ilgili" yapılar arasındaki korelasyona bakılmaktadır. Aynı kavramsal yapıyı ölçen testlerle elde edilen korelasyon sayılarının .70- .80 gibi güçlü değerler olması beklenirken ilgili kavramsal yapıları ölçen testlerle yapılan korelasyonlarda orta derecede ilişkiyi

gösteren .50 ile .70 arası değerlerin geçerlik kanıtı olarak değerlendirilebileceğini belirtmiştir. Bununla birlikte Büyüköztürk (2010) .30'un üzerindeki korelasyon katsayılarının ilişki aranan özelliğe bağlı olarak değişmekle birlikte testin geçerliliğinin kanıtı olarak kullanılabileceğini ifade etmektedir.

Ölçüt bağımlı geçerlik yöntemiyle EGİDEKAF aracının geçerliği incelenmiştir. TEDİL ve EGİDEKAF'ın alt boyutları arasındaki korelasyonlar ise .49 ile .72 arasında hesaplanmıştır. Özgün çalışmada EGİDEKAF'ın alt boyutları için ölçüt bağımlı geçerlik katsayısı olarak .47 ile .71 aralığında korelasyon katsayıları rapor edilmiştir. Benzer ölçek geliştirme çalışmalarında .40 ile .91 arasında ölçüt bağımlı geçerlik bulgularının rapor edildiği görülmektedir (Alev, 2011; Korkmaz, 2009; İncekaş, 2009; Savaşır, Sezgin ve Erol, 2006).

EGİDEKAF ile TEDİL aracı arasındaki korelasyon bulguları incelendiğinde TEDİL aracıyla arasında elde edilen değerlerin beklenen değer aralığında olduğu ve aracın geçerliği için kanıt sağladığı görülmektedir. EGİDEKAF'ın ifade edici dil alt boyutuyla TEDİL alıcı dil puanları arasında da 0.49'luk bir korelasyon katsayısı söz konusudur. Bu ilişkili iki yapı arasında da .01'lik fark ölçme hatasından kaynaklanmış olabileceği gibi TEDİL aracının doğrudan dil gelişimini ölçmeye duyarlı bir araç olmasına karşılık EGİDEKAF'ın sadece iki alt boyutta dil gelişim düzeyini ölçen informel bir eğitsel değerlendirme aracı olması dolayısıyla daha düşük duyarlılıkta ölçüm yapıyor olmasıyla açıklanabilir. Ayrıca .30 ve üzerindeki korelasyon katsayılarının testin geçerliğine kanıt sağladığını ifade eden kaynaklar göz önünde bulundurulduğunda elde edilen korelasyon katsayılarının geçerlik için kanıt sağladığı görülmektedir. Bu bilgiler doğrultusunda EGİDEKAF ile TEDİL arasında ölçüt bağımlı geçerlik için elde edilen korelasyon katsayılarının EGİDEKAF'ın geçerli kabul edilmesini sağlayacak düzeyde olduğu söylenebilir.

Psikometrik nitelikleri belirleme çalışması yapılan araçların geçerliğini incelemek üzere kullanılan ikinci yöntem yapı geçerliği yöntemidir. Yapı geçerliği testin ölçülmek istenen davranış bağlamında soyut bir kavramı (faktörü) doğru bir şekilde ölçebilme derecesini göstermektedir. Yapı geçerliğini incelemek amacıyla faktör analizi, küme analizi, iç tutarlık analizi ve hipotez testi tekniklerinden yararlanılabileceği ifade edilmiştir. Hipotez testi tekniği eşzaman geçerliği anlamına da gelmektedir yani eşzaman geçerliği bağlamında yapılan ölçüt bağımlı geçerlik çalışmaları aynı zamanda yapı geçerliği için de kanıt sağlamaktadır. Yine belirli özelliklere sahip grupları ayırt edebilme düzeyi de bu başlık altında ele alınmaktadır (Büyüköztürk, 2010). Şencan (2005), gruplar arası farklılık yoluyla yapısal geçerliğin analiz edilebileceğini ifade etmiştir. Örneğin, bu çalışmada OSB ve ZY grubunda yer alan çocukların araçların toplam puanlarından ya da alt boyut puanlarından aldıkları puanların grup ortalamasının anlamlı düzeyde farklı olmasının yapı geçerliği için kanıt oluşturduğu söylenebilir. Yine aynı yapıyı oluşturan alt boyutların birbirleriyle olan ilişkilerinin düzeyinin de yapı geçerliği için kanıt oluşturduğu ve alt boyutların kendi aralarında .60 ve üzerinde korelasyon katsayısı değerine sahip olmaları durumunda boyutların birbirleriyle bağımlı olduğu ve hepsinin birlikte aynı kavramsal yapıyı ölçtüğünden söz edilebileceğini belirtmektedir. Yine güvenilirlik analizi için kullanılmakta olan Cronbach alpha değerlerinin de yapı geçerliği için kanıt sağladığı ve geçerlik analizi amacıyla kullanılabilmesi ifade edilmektedir.

Psikometrik nitelikleri belirleme çalışmasında özgün çalışmada gerçekleştirilen analizlere paralel olarak gruplar arası farklılıklar t-test ve Kay-kare analizleri kullanılarak ve alt boyutlar arası ilişkiler Pearson korelasyon katsayısı kullanılarak incelenmiştir. EGİDEKAF ve ÖHÖKAF araçlarında t-test ve ÖHÖKAF aracının bir boyutunda da Kay-kare analizleri kullanılarak yapılan gruplar arası farklılık incelemelerinde grup ortalamaları arasında anlamlı farklılıklar olduğu görülmüştür. Bu nedenle, yukarıda yapısal geçerliğe ilişkin yapılan kuramsal açıklamalar da göz önünde bulundurulduğunda, araçların OSB olan ve ZY tanısı olan çocukları ayırt edebildiği ve bu bulguların araçların yapı geçerliğine kanıt sağladığı görülmektedir.

EGİDEKAF ve ÖHÖKAF araçları için alt ölçekler arası korelasyonların incelenmesi yoluyla yapılan yapısal geçerlik analizleri gerçekleştirilmiştir. EGİDEKAF aracında eğitsel toplam puanını oluşturan dört alt boyutun birbirleriyle olan korelasyon değerleri .47 ile .92 arasındadır. ÖHÖKAF aracında dört alt boyutun birbirleriyle olan korelasyon değerleri .62 ile .85 arasındadır. Her iki araca ilişkin olarak özgün çalışmada bu türden bir bulgu rapor edilmemiştir.

İki aracın alt boyutları için hesaplanan korelasyon katsayısı değerleri incelendiğinde eğitsel değerlendirme ifade edici dil puanı ile alıcı dil puanı arasında .47 değeri .60'lık değer altında kalan

tek ilişki katsayısını oluşturmaktadır. Bununla birlikte ifade edici dil alt boyutunun eğitsel değerlendirme toplam puanıyla olan korelasyon katsayısı .60'dır. Benzer Türkçeye uyarlama, psikometrik nitelikleri belirleme ve ölçek geliştirme çalışmaları incelendiğinde .60 değerinin altında korelasyon katsayısı rapor edilen çalışmalar olduğu görülmektedir (Köse, Bora, Erermiş ve Aydın, 2010; Korkmaz, 2009; Düver, 2006; Savaşır, Sezgin ve Erol, 2006; Temel, Ersoy, Avcı ve Turla 2005). İfade edici dil puanı ile alıcı dil puanı arasındaki diğer alt boyutlar arası korelasyonlara göre, düşük korelasyon katsayısı, güvenilirlik tartışmalarının yapıldığı kısımda da ifade edildiği gibi çocukların eğitsel performanslarının yönergeleri anlama ve yerine getirmeye dayalı olduğu bir informel eğitsel değerlendirme performans testinde diğer alt boyutlara göre daha düşük düzeyde katkı sağlıyor olması olasılığından kaynaklanabileceği düşünülebilir. Ayrıca gelişim döneminde çocukların alıcı dil düzeyleri ifade edici dil düzeylerinden daha yüksek olmakta ve ifade edici dil gelişimi OSB tanısı ve ZY tanısı olan çocuklarda gecikebilmekte bazen de ortaya çıkmayabilmektedir. Bu nedenlerle iki boyut arasındaki korelasyon katsayısı diğer alt boyutlar arası korelasyonlara göre daha düşük çıkmış olabileceği düşünülebilir. Bu olasılıklarla birlikte ifade edici dil puanlarının ölçek toplam puanıyla arasındaki korelasyon katsayısının .60 olması ayrıca alt grup farklılıklarında anlamlı düzeyde ayırt ediciliğe sahip olması ve alan yazında .47'den daha düşük korelasyon katsayılarının da rapor edilmiş olması göz önünde bulundurulduğunda bu alt boyutun diğer alt boyutlarla birlikte aynı yapıyı ölçtüğünü kabul etmemize kanıt sağladığı düşünülmüştür.

Yapılan gruplar arası farklılık ve alt boyutlar arası korelasyon analizlerinin sonuçları ve açıklamalar göz önünde bulundurulduğunda, psikometrik nitelikleri belirleme çalışması yapılan iki aracın yapı geçerliği için elde edilen bulguları, araçların geçerli kabul edilmesini sağlayacak düzeyde olduğunu göstermektedir.

Sonuç olarak psikometrik nitelikleri belirleme çalışması yapılan araçların güvenilirlik ve geçerlik bulguları incelendiğinde, araçların uzman görüşüne dayalı olarak gerçekleştirilen Türkçeye çevri çalışmaları ve veri toplama (performans alma) uygulamaları sonucunda elde edilen bulguların ortaya koyduğu güvenilirlik ve geçerlik kanıtları doğrultusunda her bir araç için elde edilen sonuçlar şu şekildedir.

EGİDEKAF aracının güvenilirliğini incelemek amacıyla test-tekrar test güvenilirliği ve madde-toplam puan korelasyonu yöntemleri kullanılarak analizler gerçekleştirilmiştir. Geçerlik boyutunda yapılan analizlerde ise ölçüt bağımlı geçerlik ve yapı geçerliği kapsamında alt boyutlar arası korelasyon ve alt grup farklılıklarının incelenmesi yöntemleri kullanılmıştır. Elde edilen bulgular yorumlandığında aracın Türkçe versiyonunun psikometrik özellikleri aracın güvenilir ve geçerli olduğunu ortaya koymaktadır.

ÖHÖKAF aracının güvenilirliğini incelemek amacıyla değerlendiriciler arası güvenilirlik yöntemi kullanılarak analiz gerçekleştirilmiştir. Geçerlik boyutunda yapılan analizlerde ise yapı geçerliği kapsamında alt boyutlar arası korelasyon ve alt grup farklılıklarının incelenmesi yöntemleri kullanılmıştır. Elde edilen bulgular yorumlandığında aracın Türkçe versiyonunun psikometrik özellikleri aracın güvenilir ve geçerli olduğunu ortaya koymaktadır.

Bu çalışmada psikometrik nitelikleri belirleme çalışması gerçekleştirilen araçların özellikleri dikkate alındığında araçlar OSB tanısında eğitsel değerlendirmeyi destekleme amacı ile kullanılabilirler. EGİDEKAF ve ÖHÖKAF araçları çocukların eğitsel yerleştirme kararlarının alınmasında ve çocuklara uygulanacak eğitim programlarında başlama düzeyi, programda ilerleme ve program tamamlandığında yapılacak son değerlendirmede program temelinde oluşturulmuş ölçüt bağımlı değerlendirmelere ek olarak nesnel değerlendirme araçları olarak kullanılabilirler. İleri araştırmalarda OSB'ye müdahalede kullanılan programların etkililiklerinin test edilmesi araştırmalarında ve OSB'ye yönelik benzer hedefleri olan etkililikleri kanıtlanmış programların etkililiklerinin karşılaştırıldığı araştırmalarda performans değişimlerinin nesnel ölçümleri amacıyla araçların kullanılabilmesi düşünülmektedir.

## KAYNAKÇA

Alev, G. (2011). *Pragmatik Dil Becerileri Envanteri'nin Türkçe standardizasyon çalışması*. Yayınlanmış yüksek lisans tezi, Anadolu Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir, Türkiye. <http://tez.yok.gov.tr/UlusalTezMerkezi> adresinden elde edilmiştir. (Tez no: 288468)



- Allen, R. A., Robins, D. L., & Decker, S. L. (2008). Autism spectrum disorders: Neurobiology and current assessment practices. *Psychology in the Schools*, 45(10), 905-917.
- Amerikan Psikiyatri Birliđi (2001). *Psikiyatride hastalıkların tanımlanması ve sınıflandırılması elkitabı (DSM-IV-TR)*(4. Baskı). (Çev: E. Körođlu) Ankara: Hekimler Yayın Birliđi
- Amerikan Psikiyatri Birliđi. (2013) *Ruhsal bozuklukların tanınal ve sayımsal elkitabı (DSM-5)* (5. Baskı).(Çev: E. Körođlu) Ankara: Hekimler Yayın Birliđi.
- Atamaz, F., Yađız On, A. & Durmaz, B. (2007). EGE Aphasia Test: Description of the test and performance in normal subjects. *Turkish Journal of Physical Medicine Rehabilitation*, 53(1), 5-10.
- Avcıođlu, H. (2011). *Özel eğitime gereksinim duyan öğrencilerin eğitsel ve davranışsal değerlendirilmesi*. Ankara: Vize Yayıncılık.
- Büyüköztürk, Ş. (2010). *Sosyal bilimler için veri analizi el kitabı* (12. Baskı). Ankara: Pegem Akademi.
- CDC(2014). Prevalence of autism spectrum disorder among children aged 8-years Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *MMWR Surveillance Summaries*, 63(2)
- Colay, İ. K. (2006). *Gülhane Afazi Testi (GAT-2)'nin standardizasyon ve geçerlik ve güvenilirlik çalışması*. Yayınlanmış yüksek lisans tezi, Anadolu Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir, Türkiye. <http://tez.yok.gov.tr/UlusalTezMerkezi> adresinden elde edilmiştir. (Tez no: 193095)
- Diken, İ. H., Ardıç, A., Diken, Ö. & Gilliam, J. E. (2012). Exploring the validity and reliability of Turkish Version of Gilliam Autism Rating Scale-2: Turkish standardization study. *Eđitim ve Bilim*, 166(37), 318-328.
- Dover, C. J., & Le Couteur, A. (2007). How to diagnose autism. *Archives of Disease in Childhood*, 92, 540-545.
- Düver, E. (2006). *5-7 yaş grubu normal gelişim gösteren ve özel gereksinimi olan çocukların dil kullanım (pragmatik) becerilerinin karşılaştırılması*. Yayınlanmamış yüksek lisans tezi, Hacettepe Üniversitesi,..... Ankara, Türkiye.
- Ege, P., Acarlar, F. & Turan, F. (2005). *Ankara Artikülasyon Testi El Kitabı*. Ankara: Ankara Üniversitesi Yayınları.
- Erbaş, D. (2012). Güvenirlik. Elif Tekin-İftar (Ed.) *Eđitim ve Davranış Bilimlerinde Tek Denekli Araştırmalar içinde* Ankara: TPD yayınları.
- Gay, L. R., Mills, G. E., & Airasian, P. (2006). *Educational research: Competencies for analysis and application* (8th ed.). New Jersey: Prentice Hall.
- Glosser, R. (2011). *Examination of the relationship between the child behavior checklist/6-18 and the social responsiveness scale parent forms using individuals with high functioning autism*, Doctoral Dissertation. Retrieved from <https://dspace.iup.edu/handle/2069/352?show=full>
- Gürsel, O. (2012). *Özel eğitimde değerlendirme*. İbrahim H. Diken (Ed.) *Özel eğitime gereksinimi olan öğrenciler ve özel eğitim içinde*. (5. Basım). Ankara: PEGEM Akademi.
- Güven, O. S. (2009). *Erken dil gelişimi testi-Üçüncü edisyon'un Türkçeye uyarlama güvenilirlik ve geçerlik ön çalışması*. Yayınlanmamış yüksek lisans tezi, Anadolu Üniversitesi, Sağlık Bilimleri Enstitüsü, Eskişehir Türkiye. <http://tez.yok.gov.tr/UlusalTezMerkezi> adresinden elde edilmiştir. (Tez no: 236587)
- Güven, O. S. & Topbaş, S.(2009). *Validity and reliability measures of TELD-3 to Turkish*. ASHA Convention, 18-21 Nov., New Orleans, US.
- Hedrick, D., Prather, E. & Tobin, A. (1975). *Sequenced inventory of communication development*. Seattle: University of Washington Press.
- Hresko, W. P., Reid, D. K., & Hammill, D. D. (1999). *Test of Early Language Development: Teld 3: Examiner's Manual*. USA: Proed.
- Innes, E., & Straker, L. (1999). Reliability of work-related assessments. *Work: A Journal of Prevention, Assessment and Rehabilitation*, 13(2), 107-124.
- İncekaş, S. (2009). *Çocukluk Otizmini Derecelendirme Ölçeđi geçerlik ve güvenilirlik çalışması*. Yayınlanmış tıpta uzmanlık tezi, Dokuz Eylül Üniversitesi, Tıp Fakültesi, İzmir, Türkiye. <http://tez.yok.gov.tr/UlusalTezMerkezi> adresinden elde edilmiştir. (Tez no: 230129)
- Kabil, A. T. (2005). *A validity study of the Checklist for Autism in Toddlers (CHAT) on a Turkish sample*. (Yayınlanmamış yüksek lisans tezi). Bođaziçi Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul.
- Kapçı, E. G., Küçüker, S. & Uslu, S. (2010). How applicable is the Ages and Stages Questionnaires to Turkish Children? *Topics in Early Childhood Special Education*, 30(3), 176-188.
- Karasar, N. (2008). *Bilimsel araştırma yöntemi* (18. Baskı). Ankara: Nobel Yayın Dađıtım Tic. Ltd. Şti.
- Korkmaz, E. (2009). *6-24 aylardaki normal gelişim gösteren ve gelişimsel gecikmesi olan bebeklerin iletişim ve sembolik davranışlarının karşılaştırılması*, Yayınlanmamış yüksek lisans tezi, Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü, Ankara, Türkiye.
- Köse, S., Bora, E., Erermiş, S. & Aydın, C. (2010). Otizm Spektrum Anketi Türkçe Formu'nun piskometrik özellikleri. *Anadolu Psikiyatri Dergisi*, 11, 253-260.
- Krug, D. A., Arick, J. R. & Almond, P. J. (2008). *ASIEP-3 Autism Screening Instrument for Educational Planning*. Third Edition. Austin, Texas: Pro-ed.

- Maviş, İ., Colay, K., Topbaş, S. & Tanrıdağ, O. (2007). Gülhane Afazi Testi (GAT-2)'nin standardizasyon ve geçerlik ve güvenilirlik çalışması. *Türk Nöroloji Dergisi*, 13(2), 89-98.
- March-Göçer, E. (1996). Frenchay Afazi Tarama Testi: Türk nöroloji hastaları için standardizasyon çalışması. *Türk Psikoloji Dergisi*, 11(38), 56-63.
- Pallant, J. (2001). *SPSS Survival Manual. A step by step guide to data analysis using SPSS*. Philadelphia: Open University Press.
- Peterson, N. L. & Meier, J. H. (1988). Assessment and evaluation processes. In Peterson, N. L. (Ed.). *Early intervention for handicapped and at risk children an introduction to early childhood-special education* (pp. 275-326). USA: Love Publishing Company.
- Savaşır, I., Sezgin, N. & Erol, N. (2006). *AGTE, Ankara Gelişim Tarama Envanteri*. Ankara: Türk Psikologlar Derneği.
- Sucuoğlu, B. (2012). Otizm spektrum bozukluğu olan çocukların değerlendirilmesi. Elif Tekin İftar (Ed.). *Otizm spektrum bozukluğu olan çocuklar ve eğitimleri* içinde. Ankara: Vize Basın Yayın.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlik*. Ankara: Seçkin Yayıncılık.
- Temel, F., Ersoy, Ö., Avcı, N. & Turla, A. (2004). *Gazi Erken Gelişimi Değerlendirme Aracı "GECDA"*. Ankara: Rekmay Ltd. Şti.
- Tincani, M. (2010). Autism spectrum disorder. In Smith, D. D. & Tyler, N. C. (Eds.). *Introduction to special education* (pp.402-431). NJ: Pearson&Merilll.
- TPD (Türk Psikologlar Derneği), (1997). *Eğitimde ve psikolojide ölçme standartları*. Amerikan Eğitim Araştırmaları Birliği, Amerikan Psikoloji Birliği, Eğitim Ölçümleri Uluslararası Konseyi Yayını. (Çev: Hovardaoğlu, S. ve Sezgin, N.) Ankara: Türk Psikologlar Derneği ve Öğrenci Seçme ve Yerleştirme Merkezi Yayını.
- Yalaz, K., Anlar, B. & Bayaoğlu, B. U.(2010). *Denver II Gelişimsel Tarama Testi Türkiye Standardizasyonu*.Ankara:Gelişimsel Çocuk Nörolojisi Derneği Yayınları.