



Anadolu-Sak Zekâ Ölçeği'nin (ASIS) Sosyal Geçerlik Çalışması *

Social Validity Study of the Anadolu-Sak Intelligence Scale (ASIS)

Ömer Faruk TAMUL¹, Bilge BAL SEZEREL², Uğur SAK³, Fatih KARABACAK⁴

• Geliş Tarihi: 11.06.2019 • Kabul Tarihi: 12.11.2019 • Çevrimiçi Yayın Tarihi: 12.11.2019

Öz

Bu çalışmada Anadolu-Sak Zekâ Ölçeği'nin (ASIS) sosyal geçerliği araştırılmıştır. Veri toplama aracı olarak 40 maddeden oluşan "ASIS Uygulayıcı Değerlendirme Anketi (ASISUDA)" kullanılmıştır. Araştırmanın katılımcılarını, ölçeğin norm çalışmasında uygulayıcı olarak yer alan ve ölçek geliştirildikten sonra formatörlük eğitimi almaya hak kazanan Mili Eğitim Bakanlığı'nın (MEB) görevlendirdiği 37 öğretmen oluşturmaktadır. Sosyal geçerlik düzeyi araştırmasında, uygulayıcıların tamamına ulaşıldığından evren ile çalışılmıştır. Evrenin sosyal geçerlik düzeyini belirlemek için ise elde edilen verilerin ortalaması ve standart sapmaları yorumlanmıştır. 6'lı likert anketin genel ortalaması 5,26 olarak belirlenmiştir. Bu değer ASIS'in sosyal geçerlik düzeyinin uygulayıcılar açısından oldukça yüksek olduğunu göstermektedir.

Anahtar sözcükler: Anadolu-Sak Zeka Ölçeği, geçerlik, sosyal geçerlik

Atıf:

Tamul, Ö. F., Sezerel, B. B., Sak, U. ve Karabacak, F. (2020). Anadolu-Sak zeka ölçeğinin (ASIS) sosyal geçerlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 49, 393-412. doi: 10.9779/pauefd.575479.

* Yapılan bu çalışma; "Anadolu - Sak Zeka Ölçeği'nin (ASIS) sosyal geçerlik ve güvenilirlik çalışması" adlı tezden oluşturulmuştur.

¹ Araş. Gör., Anadolu Üniversitesi, Eğitim Fakültesi, Özel Eğitim Bölümü, omerfaruktamul@gmail.com, <https://orcid.org/0000-0001-8884-6171>

² Araş. Gör. Dr., Anadolu Üniversitesi, Eğitim Fakültesi, Özel Eğitim Bölümü, bbal@anadolu.edu.tr, <https://orcid.org/0000-0001-7262-3563>

³ Prof.Dr., Anadolu Üniversitesi, Eğitim Fakültesi, Özel Eğitim Bölümü, usak@anadolu.edu.tr, <https://orcid.org/0000-0001-6312-5239>

⁴ Doç.Dr., ÜYEP Merkezi, fatihkarabacak@anadolu.edu.tr, <https://orcid.org/0000-0003-4925-512X>

Abstract

Social validity of intelligence tests has not become a subject of research though intelligence is a social construct. Its measurement by any instruments has not been socially investigated. In this study, the social validity of the Anadolu-Sak Intelligence Scale (ASIS) was investigated. The study involved testers' perceptions of the use of ASIS and its psychometric properties. The ASIS Practitioner Evaluation Form was used to collect data from the participants. The instrument, with a 6-point Likert scale, was composed of 40 items related to the efficiency of measuring intelligence and its components, practicality of use (administration, scoring, and interpretation), simplicity, and preference of testers. The participants included 37 testers who administered the ASIS to children during the norming study. A descriptive analysis was carried out to explore the level of participants' perceptions or satisfaction about each social validity item. The overall mean was found to be 5,26/6. This value shows a very high level of social validity. This study is the first of its kind in measuring the social validity of an intelligence test.

Keywords: Anadolu-Sak Intelligence Scale, validity, social validity

Cited:

Tamul, Ö. F., Sezerel, B. B., Sak, U. & Karabacak, F. (2020). Social validity study of the Anadolu-Sak intelligence scale (ASIS). *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 49, 393-412. doi: 10.9779/pauefd.575479.

Giriş

Günümüzde sosyal bilimlerin alt dallarından psikoloji ve eğitim bilimleri alanlarında kullanılan pek çok psikolojik ölçme aracı bulunmaktadır. Farklı alanlar için geliştirilen bu psikolojik ölçme araçlarının genel olarak iki amaca hizmet ettiği görülmektedir (Özgüven, 2014). Bunlardan ilki klinik ortamlarda tanılama, yerleştirme ve araştırma amaçlı kullanım, ikincisi ise eğitim ortamlarında ölçme, seçme ve yerleştirme amaçlı kullanım olarak düşünülebilir (Özgüven, 2015; Seçer, 2014). Güncel olarak kullanılan bu psikolojik ölçme araçlarının belirlenen amaçlara hizmet edebilmesi ve doğru sonuçlar ortaya koyabilmesi için taşınması gereken temel psikometrik (geçerlik, güvenilirlik gibi) özellikler bulunmaktadır (Seçer, 2015). Bu bağlamda araştırmanın amacı Anadolu – Sak Zekâ Ölçeği'nin (ASIS) geçerliği ile ilişkili olarak ölçeğin sosyal geçerlik düzeyini incelemektir.

Zekâ ölçekleri geliştirilirken ve standardizasyon çalışmaları yapılırken geçerlik ve güvenilirlik çalışmaları yapılması gerekmektedir. Ancak literatüre bakıldığında sosyal geçerliğin bir geçerlik türü olmasına rağmen zekâ ölçeklerinde incelenmediği görülmektedir. Oysa zekâ ölçeklerinin psikometrik özelliklerinin kapsamlı incelemesi ortaya çıkacak sonuçları iki yönden etkilemektedir. Bunlardan ilki tanılamadan elde edilen sonuçların doğruluk yüzdesi, ikincisi ise geliştirilen ürünün hedef kitle tarafından benimsenmesine bağlı olarak yaygın kullanımınıdır. Bu nedenle toplumun geliştirilen bir uygulama ya da yöntemi ne derece kabul ettiğini ortaya koymak için de sosyal geçerlik düzeyine bakılması gerekmektedir (Wolf, 1978). Bu sayede geliştirilen ürünün ne kadar sürdürülebilir olacağı ortaya konacaktır (Kazdin, 1977; Wolf, 1978).

Sosyal Geçerlik

Sosyal geçerlik kavramı literatürde kapsamlı ve sistematik olarak Wolf (1978) ve Kazdin (1982) tarafından açıklanmıştır. Wolf (1978) sosyal geçerlik kavramını bir ürüne yönelik toplumun verdiği değer olarak açıklamıştır. Diğer bir deyişle, sosyal geçerlik toplumun ürüne verdiği değeri belirlemek, hedef kitleye uygun amaçların seçilip seçilmediğini ortaya koymak, ürünlerin sosyal olarak kabul edilip edilmediğini araştırmak ve hedef kitlenin memnuniyet düzeyini belirlemek için geliştirilmiş bir kavramdır (Foster ve Mash, 1999; Wolf, 1978).

Geliştirilen bir uygulamanın sosyal geçerlik düzeyini belirlemek için 3 farklı bileşen incelenebilir. Bunlar; belirlenen amaçların anlamlılığı, uygulanacak yöntemlerin uygunluğu ve ortaya çıkacak sonuçların önemidir. Bu bileşenlerin düzeylerinin yüksek çıkması geliştirilen uygulamaların kullanıcılar tarafından benimseneceğinin ve uygulamaların etkin bir şekilde kullanılacağına göstergesidir (Wolf, 1978). Bu anlamda sosyal geçerlik kavramı kısaca toplumun ya da hedef kitlenin verdiği değer ve önem olarak açıklanabilir.

Sosyal geçerlik düzeyinin belirlenmesinde sadece sonuçlara değil süreçlere de odaklanmak gerekmektedir. Sosyal geçerlik düzeyi belirleme çalışmaları uygulama geliştirme sürecinin öncesinde, uygulama esnasında ve uygulama sonrasında yapılabilir. Ayrıca bütün bu değerlendirme süreçlerini toplumsal koşullar ile ilişkilendirmek de önemlidir (Finney, 1991; Schwartz ve Baer, 1991; Winnet, Moore ve Anderson, 1991). Üç farklı zamanda yapılan bu uygulamanın geliştirilen uygulamaya farklı yararları olacaktır. Örneğin, uygulama öncesinde yapılan bir sosyal geçerlik çalışması toplumun ihtiyacını belirlemek ve buna göre yatırım

yapmak açısından önemlidir. Uygulama ile eş zamanlı uygulanan bir çalışmada ise sürece yönelik aksilikler daha net bir şekilde belirlenmekte ve uygulama esnasında hızlı müdahaleler yapılabilmektedir. Uygulama tamamlandıktan sonra yapılan sosyal geçerlik çalışmalarında ise uygulamanın toplumun ihtiyaçlarını ne kadar karşıladığını belirlemek ve varsa eksik yanlarını tespit ederek ileriye yönelik düzeltme yapmak olanaklı hale gelmektedir.

Uygulama sürecinde yapılan değerlendirmelerle birlikte değerlendirmeye katılan kişilere göre de sosyal geçerlik çalışmaları farklılık göstermektedir. Sosyal geçerlik çalışmaları programdan doğrudan etkilenen kişiler (dolaysız kullanıcılar), programdan doğrudan fayda sağlamayan ama dolaylı etkilenen kişiler (dolaylı kullanıcılar), bu kullanıcılarla yakın etkileşimde olan kişiler (yakın çevre) ve genel toplum üyeleri (geniş çevre) ile yapılabilmektedir (Hurley, Wehby ve Feurer, 2010; Schwartz ve Baer, 1991). Schwartz ve Baer'e göre (1991) uygulamalarda katılımcılar belirlenirken doğru olmayan kişilerle ya da sınırlı sayıda kişiyle çalışmak gibi hatalardan kaçınılması gerekir. Çünkü bu hatalar elde edilmek istenen görüşlerin doğru bir şekilde değerlendirilmemesine neden olabilmektedir. Sosyal geçerlik araştırmaları incelendiğinde genellikle araştırmaların sosyal bilimlerin çeşitli dallarında yapıldığı görülmektedir. Özellikle sosyal bilimlerin ekonomi ve iletişim bilimleri bunların başında gelmektedir. Yapılan bu çalışmalar öncelikle tüketicilerden (müşteri) bilgi toplamak amaçlıdır (Pollay, 1984). Örneğin reklam sektörü ya da yeni bir ürün geliştiren kurumların ürünü piyasaya sürmeden önce ya da sonra sosyal geçerlik çalışmalarından faydalandığı bilinmektedir. Çünkü bireylerin en başta neye ihtiyaç duydukları ya da üründen ne bekledikleri ortaya konmalı ve ihtiyaçlar ve beklentilere paralel olarak hareket edilmelidir. Diğer taraftan sosyal bilimlerin psikoloji ve eğitim bilimleri dalları için de sosyal geçerlik çalışmaları önemli bir araştırma konusudur. Örneğin bir zekâ ölçeği ya da bir eğitim uygulaması geliştirildiğinde de hedef kitlenin ürüne yönelik beklenti ve talepleri belirlenmelidir. Bu şekilde zekâ ölçeklerine yönelik yapılan eleştiriler net bir şekilde belirlenebilecek ve sorunlar ortadan kaldırılabilecektir. Ancak yukarıda da bahsedildiği üzere alan yazın incelendiğinde sosyal geçerliğin göz ardı edilen bir psikometrik özellik olduğu görülmektedir. Sosyal geçerlik araştırmasına geçmeden önce, araştırmanın konusu bir zeka ölçeğinin sosyal geçerliği olduğundan öncelikle zekâ ve zekânın ölçülmesine yönelik yapılan çalışmaları özetlemekte fayda vardır.

Genel Zekâ ve Ölçümü

Zekâ kavramlarına dönük olarak yapılan çalışmaları geçmişten günümüze tarihsel perspektiften incelemenin kavramların anlaşılabilirliğine yönelik daha kapsamlı bir bakış sunacağı düşünülmektedir. Bu bağlamda zekâ kavramının ne olduğu ile ilgili olarak yapılan alan yazın taramasında, literatürde kabul edilen ortak bir tanımın olmadığı rahatlıkla söylenebilir. Çünkü kavramın doğası gereği, zekânın evrensel ve sosyal bir etiket olarak değerlendirildiği görülmektedir (Carroll, 1997; Kontos, Carter, Ormrod ve Cooney, 1983). Diğer bir ifadeyle, zekâ toplumdaki topluma, kişiden kişiye, disiplinden disipline, zamandan zamana farklılık göstermekte ve dolayısıyla kavrama yönelik atfedilen tanım da değişiklik göstermektedir. Öyle ki 2000 yıl önce Aristo'nun yaptığı zekâ tanımından günümüzde yapılan çağdaş zeka tanımlarına kadar, kavramın ne olduğuna yönelik herhangi bir fikir birliği sağlanamamıştır (Goddard, 1946). Örneğin bilimsel bakış açısından önce, Yunanistan'da bireylerin üstün zekâlı olması "ilham perisi" tarafından edebiyat ve dans alanında verilen bir ayrıcalık olarak görülmekteyken (Phillipson, 2007), Çin'de cennetten gelen bir hediye olarak değerlendirilmekte

ve aileye bağlılık ile özdeşleştirilmektedir (Chan,2007). Bilimsel bilginin farklı disiplin alanlarına entegre olması ile birlikte zekâyâ yönelik fikirler ve tanımlar daha bilimsel bir hale gelmiştir.

Zekâ ile ilgili bilimsel araştırmalar incelendiğinde ilk bilimsel fikri 1884 yılında Galton'un ortaya attığı görülmektedir. Galton zekâyı basit algısal motor beceriler ve duyuşal süreçler olarak savunmuştur. Ayrıca zekânın yeni nesillere kalıtım yoluyla aktarıldığını belirtmiştir (Eysenck, 1967; Sternberg, 2012). Galton yaptığı çalışmalarda zekânın varlığını iddia etmiş ve zekâ ile beynin büyüklüğü arasında bir ilişki olup olmadığını sorgulamıştır (Raz, Willerman ve Yama, 1987). 1902 yılında Binet "Deneyşel Zekâ Çalışmaları" kitabında zekânın kavrama, karar verme ve akıl yürütme gibi üst düzey çalışmalarda ortaya çıkacağını vurgulamıştır (Goddard, 1946; Selçuk, Kayılı ve Okut, 2002). Spearman ise 1904 yılında, zekâyı özel ve genel yetenek olarak tanımlamıştır (Eysenck, 1967). Özel yetenek belirli bilişsel süreçlerin yürütülmesinde görev alırken, genel yetenek bütün bilişsel süreçlerde rol almaktadır (Jenden ve Weng, 1994). Zekâ kavramına yönelik bilimsel yaklaşımlarda 1921 yılında deęişim yaşanmıştır. Yaşanan deęişimin kaynağı The Journal of Education Psychology dergisinin 1921 yılında düzenlediğı sempozyumdur. The Journal of Education Psychology dergisi 1921 yılında düzenlediğı sempozyuma zeka ile ilgilenen dönemin önde gelen uzmanları katılmış ve zekayı tanımlamışlardır (Goddard, 1946; Sternberg, 2012). Sempozyumda Stern, zekâyı bireyin kendi fikirlerini karşılaştığı yeni durumlara uydurabilme becerisi; Thurndike, gerçek ve olgular arasında iyi seçimler yapabilme becerisi; Terman, soyut düşünme becerisi; Freeman, duyuşal kapasite, algılama kapasitesi, esneklik, hayal gücü ve hızlı cevaplama becerisi; Dearborn, öğrenme ve deneyimlerinden yarar sağlama becerisi olarak tanımlarken Colvin, öğrenme süreci ya da bulunduğu ortama uyum sağlayabilme becerisi; Pintner, göreceli olarak yeni durumlara adapte olabilme becerisi; Lester Ward ise akıl ve bilginin toplamı olarak tanımlamışlardır. 1946 yılına gelindiğinde ise Goddard (1946) öne çıkan zeka tanımlarını kronolojik olarak derlemiştir. Tanımlar incelendiğinde 1921 yılındaki sempozyumdan önce yapılan tanımlarda daha çok farklı çevresel koşullara adaptasyon vurgulanırken, sempozyumdan sonra yapılan tanımlarda ise algılama ya da yargılama gibi bilişsel beceriler vurgulanmıştır. Örneğin Gottfredson (1994) zekâyı; problem çözme, mantık, soyut düşünme, planlama, hızlı öğrenme, karmaşık fikirleri kavrama ve deneyimler yoluyla öğrenme gibi becerilerin birleşimi olan zihinsel bir süreç olarak tanımlamıştır. Broddy (1999) ise zekâyı geleneksel olarak IQ ile ifade edilen ve zekâ bölümlerinden oluşan bir yapı olarak ele almıştır. Zekâ bölümünü ise, zihinsel becerileri deęerlendirmek için farklı alt testlerin birleşimi ile hazırlanan testlerin karşılık geldiğı deęer olarak tanımlamıştır. Sak (2014) ise zekâyı beynin bir bütün olarak verimli ve etkili çalışmasının davranış üzerindeki etkileri olarak tanımlamıştır.

Yukarıda sözü edilen tanımlar dikkate alındığında bazı kavramların ortak olarak vurgulandığı göze çarpmaktadır. En sık vurgulanan kavramlar arasında; kolay öğrenbilme, problem çözme becerisi, çevreye uyum sağlama, soyut düşünme, deneyimlerden çıkarım yapma, seçim yapabilme ve ilişkileri algılama olduğu görülmektedir (Kızıltepe,2004).

Zekânın tanımlanmasından sonra zekânın nasıl ölçülebileceğı araştırmacıların ilgisini çekmiştir. Tarihsel süreç içinde zekâ testlerinin gelişimini en çok etkileyen faktör, pratik etmenlerdir. Örneğin, zihinsel olarak gelişim geriliğı gösteren bireylerin belirlenmesi ihtiyacının ortaya çıkması, sözlü ya da yazılı sınavların objektif olmaması ve objektif testlerin birden fazla

kişiyi hızlı bir şekilde tanılaması bu etmenlerdendir (Özgüven, 2015). Bu ihtiyaçların karşılanmasına yönelik araştırma ilgisi ise zekânın farklı boyutlarına değinen tanımların ve kuramların zekânın nasıl ölçülebileceğine dair çeşitli fikirlerin üretilmesine neden olmuştur.

Zekânın ölçülmesine yönelik ilk deneysel çalışmaları Galton yapmıştır. Darwin'in fikirlerinden etkilenen Galton yaptığı çalışmalarda zekânın kalıtım ile yeni nesillere aktarıldığını savunmuştur (Öner, 1996). Bundan dolayı zihinsel beceriyi, motor beceriler ve fiziksel yapının incelenmesi yoluyla ölçmeye çalışmıştır (Sternberg, 2012).

Galton'un zekâ ile ilgili fikirlerinden etkilenen Cattell, zekânın ölçülebilmesi için 50 farklı psiko-fiziksel test önermiştir (Sternberg, 2012). Bu testler hareket hızı, görme ve işitme durumları, farklı ağırlıkları olan cisimleri ayırt edebilme, ağrıya dayanma düzeyleri, kas gücü, bellek ve reaksiyon süresi gibi özellikleri içermektedir (Özgüven, 2014). Wissler (1901) ise Cattell'in testini kullanarak Kolombiya Üniversitesi'nde bir çalışma yapmıştır. Yaptığı çalışmada 21 öğrencinin okul puanları ile test puanlarını karşılaştırmıştır. Çalışma sonunda iki değişken arasındaki ilişki sifıra yakın çıkmıştır.

Wissler'in araştırmasının sonunda duyuşal özelliklerin zekâ ölçeklerinde ölçülmesinin gereksiz olduğu ortaya çıkmış ve yeni yaklaşımlara ihtiyaç olduğu görülmüştür (Özgüven, 2014; Sternberg, 2012; Wissler, 1901). Zekânın ölçülmesi için kullanılan duyuşal beceri testleri yeterli olmayınca ilerleyen zamanlarda daha doğru bilgi edinmek için algı kapasitesi, dil becerisi ve soyut düşünme gibi daha karmaşık beceriler dikkate alınmıştır (Gardner, 2011).

Spearman'ın yaptığı çalışmalar zekânın duyuşal aracılığıyla ölçümünden, bilişsel süreçlerin incelenmesi ile ölçümüne geçişteki önemli bir dönüm noktasıdır. Spearman 1904 yılındaki çalışmalarının sonucunda günümüzde kullanılan zekâ ölçeklerinin dahi benimsediği üç sonuca ulaşmıştır. Spearman'ın ulaştığı ilk sonuç yapılan farklı ölçümlerin birbirleriyle pozitif ilişki içinde olmasının gerekliliğidir. İkinci sonuç, genel zekâyı ölçebilmek için kullanılan ölçütlerin birbirinden farklılık göstermesidir. Son olarak ise yapılan farklı ölçümlerin birbiri ile ilişkilerinin artması ölçümlerin genel zekâ ile ilişkisinin arttığının göstergesidir. Kısaca Spearman (1904) birbiri ile ilişkili farklı ölçütlerin kullanılması ile genel zekâyı ölçebileceğimizi belirtmiştir. Ancak Spearman genel zekâ kavramının karmaşık olan ilişkileri açıklamada yetersiz kalmasından ötürü, genel zekâdan farklı zihinsel boyutlara da gereksinim olduğunu düşünmektedir (Brody, 1999; Wasserman, 2012).

Spearman'ın çalışmalarından sonra zekânın sayısal bir değer olarak ifade edilmesi süreci hızlanmıştır. Binet ve Simon 1905 yılında Paris'teki okullarda öğrencilerin zekâ düzeylerini belirlemek için zekâ ölçeği geliştirmiştir. Binet ve Simon'un geliştirdiği ilk zekâ ölçeğinde günümüzdeki zekâ ölçeklerine benzer olarak zekâ, yargı ve çıkarımda bulunma, kavrama gibi becerilerle tanılanmaya çalışılmıştır (Kaplan, 2008). Aynı dönemlerde Stern (1912) kronolojik yaş ve zekâ yaşı kavramlarını ortaya atmıştır. Zekâ bölümünü ise, Zekâ Bölümü (IQ)= Zekâ Yaşı (ZY) / Takvim Yaşı (TY) x 100 şeklinde bir sayısal formülle ortaya koymuştur (Aktaran: Suveren, 2006).

Geçmişten günümüze zekâ ölçekleri incelendiğinde kimi açılardan eleştirildikleri görülmektedir. 19. Yüzyılın ilk yarısında geliştirilen zekâ ölçeklerinin eleştirildiği noktalar, test süresinin gereğinden fazla uzun olması, dezavantajlı grupları tanılamadaki yetersizliği ve gerçek zekâyı ölçmedeki yetersizliğidir (Brim, Glass, Neulinger ve Firestone, 1969). Zekâ ölçeklerine

yönelik en büyük eleştirilerden bir diğeri ise zekâ ölçeklerinin farklı sosyo –ekonomik düzeydeki ve alt kültürdeki bireylerin zekâlarını ölçmede yeterli olmamasıdır (Snyderman ve Rothman, 1987). 19. Yüzyılın ikinci yarısı ve 20. Yüzyılda geliştirilen zekâ ölçeklerine yönelik eleştiriler incelendiğinde yapılan eleştirilerin azalmadığı görülmektedir. Dumont ve Willis (2010) WIAT-III (Wechsler İndividual Achievement Test-III- Wechsler Bireysel Başarı Testi-III) ve KTEA-II (Kaufman Test of Educational Achievement-II – Kaufman Eğitim Başarısı Testi – II) zekâ ölçeğini içerik olarak eleştirmişlerdir. Yapılan eleştiriler arasında kayıt formunun çok uzun olması (50 sayfa), maddelerin birden çok doğru cevabının olması, uygulayıcı kitabının karmaşık olması yer almaktadır. Eleştirilen noktaların uygulama sırasında ve öğrenci profillerinin yorumlanması sırasında sorunlara neden olduğu savunulmaktadır. Snyderman ve Rothman (1987) da yaptıkları çalışmada yapılan eleştirilere ek olarak, ilköğretim ve ortaöğretim öğrencilerinden elde edilen zekâ puanlarının yanlış yorumlanabileceğini belirtmiştir.

Dumont ve Willis'in (2013) WAIS-IV'e (Wechsler Adult Intelligence Scale- IV- Wechsler Yetişkinler İçin Zekâ Ölçeği – IV) yaptıkları eleştiriler; uygulayıcıya değerlendirme esnekliği sağlamaması, sadece alt testler temelinde yorumlar yapılması, testi yorumlayabilmek için gereken materyallerin yetersizliği ve teorik alt yapısının eksik olmasıdır. Roid ve Barram, (2004) ise Stanford Binet-5 Zekâ Ölçeği'nin (SB5) çeşitli yönlerden eleştirmektedir. SB5 zekâ ölçeği; alt testler arasında geçiş yapılırken materyallerin zorluk çıkarması, görsel alt testlerden bazılarının ifade edici ve alıcı dil becerisi gerektirmesi, tanıladığı zekâ aralığının dar olması (IQ 40-160) ve sonuçların yorumlanmasının ya da bilgisayarda değerlendirme yapılmasının fazladan maliyet getirmesi yönünden eleştirilmektedir.

Fletcher-Janzen ve Lichtenberger (2005) K-ABC-II Zekâ Ölçeği'ni eleştirmişlerdir. K-ABC-II ölçeğine yaptıkları eleştirilerde alt testlerdeki sonlandırma kurallarının farklılık göstermesinin uygulayıcılar açısından uygulama sorununa neden olacağını belirtmişlerdir. Sözel alt testlerin düşük zekâ seviyesindeki bireylerin düzeylerinin üzerinde olduğunu ve bu kişilerin sözel alt testleri anlamayacağını savunmuşlardır. Ayrıca zekâ puanı hesaplanırken ayrı bir sütunun olmamasının uygulama esnasında uygulayıcıları zorlayacağını belirtmişlerdir.

Zekâ ölçeklerine yapılan eleştiriler incelendiğinde ölçekler güvenilirlik ve geçerlik özelliklerini istatistiksel anlamda sağlamış olsalar da uygulama esnasında karşılaşılan zorlukların tanılamada sorunlara neden olacağı görülmektedir. Oysa zekâ ölçeğinin en önemli amacı hedef kitleye dâhil olan bireyi en doğru şekilde tanılamak ve ihtiyaçlarını belirlemektir. Bu nedenle sosyal geçerlik çalışması ile birlikte zekâ ölçeklerini uygulayacak kişilerin sürece dâhil edilmesi ve sorunların en doğru şekilde tespit edilmesi ve giderilmesi daha önemli bir hale gelmektedir. Böylece uygulama ve değerlendirme aşamasında yaşanan sorunlar azaldıkça, uygulayıcıların zekâ ölçeğini kabul düzeyi artacağı söylenebilir.

Yapılan bu çalışmada da Anadolu-Sak Zekâ Ölçeği'nin sosyal geçerlik düzeyi araştırılmıştır. Anadolu-Sak Zekâ Ölçeği incelendiğinde, ölçeğin Türkiye'de geliştirilmiş olan ilk bireysel zekâ ölçeği olma özelliği taşıdığı görülmektedir (Sak ve diğ., 2016). 4 yaş 00 ay 00 gün -12 yaş 11 ay 30 gün aralığındaki çocuklar için kullanılan ölçeğin uygulaması 20 – 45 dakika arasında sürmektedir. 7 alt testten oluşan ölçek 7 farklı endeks puanı vermektedir. Alt testler, Görsel Ardıl İşleyen Bellek (GAB), Sözel Analogik Muhakeme (SAM), Görsel Algısal Esneklik (GES), Görsel Analogik Muhakeme (GAM), Sözel Kısa Süreli Bellek (SKB), Görsel

Eş Zamanlı İşleyen Bellek, (GEB) ve Sözcükler Anlamlar (SAN) alt testlerinden oluşmaktadır. Ölçeğin endeks puanları ise Sözel Potansiyel Endeksi (SPE), Görsel Potansiyel Endeksi (GPE), Bellek Kapasitesi Endeksi (BKE), Sözel IQ (SZE), Görsel IQ (GZE), Tarama Endeksi (TIQ) ve Genel Zekâ Endeksi (GIQ) olmak üzere 7 türden oluşmaktadır. Ölçeğin geliştirilme sürecinde CHC kuramı temel alınırken, alt testlerde ise Luria'nın Nöropsikolojik Modeli ve Baddeley'in Bellek Modeli temel alınmıştır. ASIS'in bileşen ve alt test özellikleri ise şu şekildedir;

Genel Zekâ Endeksi; bütün zihinsel becerilerin bir bileşimidir. Bütün alt testlerin toplamı ile elde edilmektedir.

Sözel Potansiyel Endeks; SAM ve SAN alt testlerinden oluşan bir endekstir. SPE ile kristalize zekâ dâhil sözel muhakeme, sözel anlama, dil gelişimi, genel bilgi, semantik bilgi ve sözcük bilgisi ölçülmektedir. SAM alt testi problem çözme ve soyut muhakeme becerisini ölçmektedir. Alt test klasik analogi problemlerinden oluşmaktadır. SAN alt testi ise sözcük dağarcığı, sözcük bilgisi, dil gelişimi ve sözel anlama becerilerini ölçmektedir.

Görsel Potansiyel Endeks; GAM ve GES alt testlerinin bileşiminden oluşan bir endekstir. GAM alt testi soyut düşünme ve muhakeme becerisini görsel analogiler aracılığıyla ölçmektedir. Analogiler soyut şekillerden oluşmaktadır. Test ilerledikçe maddelerin soyutluğu artmaktadır. GES alt testinde, görsel algısal işleme, görselleştirme, uzamsal ilişkiler, zihinsel esneklik, görsel esneklik, algısal ayırt edicilik ve görsel manipülasyon becerileri ölçülmektedir. Şekiller iki boyutlu ve soyut şekillerdir.

Bellek Kapasitesi Endeksi; işleyen bellek ve kısa süreli bellekten oluşmaktadır. Endeks puanı ise GEB, GAB ve SKB alt testlerinden oluşmaktadır. GEB alt testi görsel uzamsal işleyen bellek ve görsel uzamsal eş zamanlı işleme becerisini ölçmektedir. GAB, görsel örüntülerden oluşan ve görsel kısa süreli bellek ile görsel ardıl işlemeyi ölçen bir alt testtir. SKB ise sözel kısa süreli belleği ölçmektedir.

Sözel IQ ve Görsel IQ; görsel IQ puanı, GAM, GAB, GEP ve GES alt testlerinden elde edilen puanla hesaplanırken sözel IQ puanı ise SAM, SAN ve SKB alt testlerinden elde edilen puanlarla hesaplanır. SPE ve GPE den farklı olarak bu puanlara bellek puanları da dâhil edilmektedir.

Tarama Endeksi; kısa formun uygulanması ile elde edilen puan türüdür. Kısa form SAN ve GAM alt testlerinden oluşmaktadır. İki şekilde tarama yapılabilir. İlk olarak çocuklara tarama alt testleri uygulanır ve belirlenen değer in altında kalan ya da üzerinde olan çocuklara tüm test uygulaması yapılır. Diğer bir yöntem ise ilk önce SAN ve GAM iki alt testi uygulanır. Uygun görülmesi halinde teste devam edilir.

ASIS'in standardizasyon çalışması yapılırken 2015 yılı Türkiye nüfus sayımı verileri dikkate alınarak yapılmıştır. 4-12 yaş arasındaki 11 farklı yaş grubunda yer alan 4641 öğrenci ile çalışılmıştır. Her bir yaş grubu 410-429 arasında çocuktan oluşmaktadır. Örneklem belirlenirken yaş, cinsiyet, okul türü, eğitim durumu ve yerleşim alanları dikkate alınarak tabakalı örneklem yöntemi kullanılmıştır.

ASIS'in güvenilirlik çalışmaları incelendiğinde iç tutarlılık katsayısı ve kodlayıcılar arası güvenilirliği hesaplanmıştır. İç tutarlılık güvenilirlik katsayıları Tablo 2.1'de verilmiştir.

Tablo 2.1. ASIS İç Tutarlılık Katsayıları

Alt testler ve endeksler	İç tutarlılık katsayıları
*GIQ	.99
*SPE	.99
*GPE	.95
*BKE	.95
*SZE	.99
*GZE	.97
*KF	.98
*GAB	.88
*SAM	.97
*GES	.84
*GAM	.95
*SKB	.81
*GEB	.94
*SAN	.98

Kaynak: Sak ve diğ., (2016)

*GIQ: Genel Zekâ Endeksi, SPE: Sözel Potansiyel Endeksi, GPE: Görsel Potansiyel Endeksi, BKE: Bellek Kapasitesi Endeksi, SZE: Sözel IQ, GZE: Görsel IQ, KF: Kısa Form, GAB: Görsel Ardıl İşleyen Bellek, SAM: Sözel Analogik Muhakeme, GES: Görsel Algısal Esneklik, GAM: Görsel Analogik Muhakeme, SKB: Sözel Kısa Süreli Bellek, GEB: Görsel Eşzamanlı İşleyen Bellek, SAN: Sözcükler Anlamlar

Tablo 2.1'de yer alan ASIS'in iç tutarlılık katsayıları incelendiğinde en düşük değer .81 ve en yüksek değer .99 olduğu görülmektedir. İç tutarlılık katsayısı en yüksek çıkan ve mükemmel sınırlar içinde yer alan değerler, SZE, GZE ve GIQ endeks değerleridir.

ASIS'in iç tutarlılık güvenilirlik katsayıları yaş grupları değişken olarak alındığında ise, 4-12 yaş arasında 11 farklı güvenilirlik katsayısı elde edilmiştir. Yaş grubuna göre iç tutarlılık katsayıları Tablo 2.2'de verilmiştir.

Tablo 2.2. ASIS Yaş Gruplarına Göre İç Tutarlılık Güvenirlik Katsayıları

Yaş	*GIQ	*SPE	*GPE	*BKE	*SZE	*GZE
4	.85	.82	.65	.76	.82	.78
4,5	.88	.85	.72	.81	.86	.82
5	.92	.91	.79	.84	.90	.87
5,5	.92	.91	.77	.84	.90	.87
6	.94	.93	.84	.87	.93	.90
7	.96	.96	.90	.88	.95	.92
8	.97	.96	.92	.88	.96	.93
9	.97	.97	.92	.85	.96	.93
10	.97	.97	.92	.87	.97	.93
11	.98	.97	.93	.86	.97	.94
12	.97	.97	.92	.85	.97	.94

Kaynak: Sak ve Diğ., 2016

* *GIQ: Genel Zekâ Endeksi, SPE: Sözel Potansiyel Endeksi, GPE: Görsel Potansiyel Endeksi, BKE: Bellek Kapasitesi Endeksi, SZE: Sözel IQ, GZE: Görsel IQ*

Tablo 2.2’de yer alan iç tutarlılık güvenilirlik değerleri incelediğinde güvenilirlik değerlerinin oldukça yüksek olduğu görülmektedir. Endeks puanlarının yaş aralıkları dikkate alınarak iç tutarlılık kat sayıları hesaplandığında ise iç tutarlılık güvenilirlik katsayılarının oldukça yüksek olduğu bulgulanmıştır. Sak ve diğ. (2016) bu durumun nedenini kuramsal alt yapıyla uygun bir test olması ile ilişkilendirmiştir.

ASIS güvenilirlik çalışmaları kapsamında yapılan bir diğer çalışma ise kodlayıcılar arası güvenilirlik çalışmasıdır. Bu amaçla norm uygulamaları arasından rastgele seçilen 45 form, ASIS uygulayıcı eğitimi almış iki uygulayıcı tarafından bağımsız olarak puanlanmıştır. Analizler ise alt test puanları toplamı üzerinden yapılmıştır. Verilerin analizi sonucunda, SAN alt testi dışında diğer alt testlerin kodlayıcılar arası güvenilirliği 1.00 çıkmıştır. SAN alt testi ise .96 olarak hesaplanmıştır. Değerlerin bu kadar yüksek olması ise ASIS’in kolay uygulanabilir ve objektif kurallar içerdiğinin bir göstergesi olarak yorumlanabilir. Geçerlik ve güvenilirlik değerleri yüksek olan bir testin aynı zamanda sosyal geçerliğini de araştırmak literatüre uygulayıcıların memnuniyeti bağlamında önemli bir kazanım olarak düşünülmektedir. Bu bağlamda yapılan çalışmanın amacı ASIS’in sosyal geçerliğini araştırmaktır. Bu amaçla “Uygulayıcıların ASIS’e ilişkin memnuniyet algıları nasıldır?” sorusuna cevap aranmıştır.

Yöntem

Yapılan araştırmada tarama modeli kullanılmıştır. Sosyal geçerlik düzeyini belirlemek için ise betimsel tarama yaklaşımı uygulanmıştır.

Çalışma Grubu

Anadolu – Sak Zekâ Ölçeği’nin sosyal geçerlik düzeyini belirlemek amacıyla Milli Eğitim Bakanlığı tarafından ASIS’in norm çalışmasında uygulayıcı olarak yer alan ve formatör eğitimlerinde görevlendirilen öğretmenler ile çalışılmıştır. Araştırmada, uygulayıcıların tamamına ulaşıldığından evren ile çalışılmıştır. Evren, Bilim Sanat Merkezlerinde (BİLSEM) ve Rehberlik ve Araştırma Merkezlerinde (RAM) görev yapan 37 kişilik Psikolojik Danışmanlık ve Rehberlik öğretmeninden oluşmaktadır. Ancak ankette yer alan bazı maddelerde ASIS ile diğer zekâ ölçekleri arasında karşılaştırmaların yapılması istenmektedir. Bu nedenle ASIS dışında farklı bir zekâ testi uygulamayan 5 uygulayıcının verileri veri setinden çıkarılmış ve araştırma toplamda 32 katılımcı (17 erkek (%53,13), 15 kadın (%46,87)) ile gerçekleştirilmiştir.

Çalışmada Kullanılan Ölçme Araçları

Anadolu–Sak Zekâ Ölçeği uygulayıcı değerlendirme anketi

Çalışmada veriler “Anadolu–Sak Zekâ Ölçeği Uygulayıcı Değerlendirme Anketi (ASISUDA)” ile toplanmıştır. ASISUDA, ASIS uygulayıcılarının anket hakkındaki görüşlerini almak amacıyla geliştirilmiştir. Ankette ASIS’e ilişkin uygulanabilirlik algısı, beğeni algısı, yeterlik algısı ve anlaşılabilirlik algısı alt boyutlarından oluşan 40 madde bulunmaktadır. Uygulanabilirlik algısı alt boyutu 11 maddeden (2, 4, 9, 12, 13, 15, 22, 24, 30, 31, 34), beğeni algısı alt boyutu 6 maddeden (8, 11, 32, 35, 37, 38), yeterlik algısı alt boyutu 14 maddeden (1, 3, 5, 6, 14, 16, 19,

20, 25, 26, 27, 33, 36, 39) ve anlaşılabilirlik algısı alt boyutu 9 maddeden (7, 10, 17, 18, 21, 23, 28, 29, 40) oluşmaktadır. Uygulayıcıların, anketin tamamını değerlendirip değerlendirmediklerini ve değerlendirme esnasında objektif olup olmadıklarını kontrol etmek amacıyla ankette yer alan 40 maddeden 6 tanesi olumsuz (10, 12, 21, 33, 35, 39) madde olarak kodlanmıştır.

Uygulanabilirlik algısı alt boyutu, uygulayıcıların ASIS'i uygulamayı öğrenme düzeyi, alt testlerin ve alt testlerdeki maddelerin sıralamasının uygunluğu, uyarıcı kartın (ön ve arka yüzünde öğrencinin çeşitli tahminler yürütebilecekleri şekillerin olduğu ve öğrenciyi teste hazırlama amacıyla test öncesinde uygulanan kart) işlevselliği ve ASIS'in uygulanması, puanlanması ve yorumlanmasının kolaylık düzeyi değerlendirmektedir. Beğeni algısı alt boyutu ise maddelerin ve uyarıcı kartın çocukların dikkatini ne ölçüde çektiği ve çocukların testten sıkılıp sıkılmadıklarını değerlendirmektedir. Ayrıca ASIS'i diğer uygulayıcılara tavsiye edip etmeyecekleri ve bir çocuğa zekâ testi uygulayacağı zaman diğer zekâ testleri arasından ASIS'e öncelik verip vermeyecekleri değerlendirilmektedir. Yeterlik algısı alt boyutunda, ASIS'in zekâyı ve diğer bilişsel becerileri (hafıza, görsel yetenek, sözel yetenek) ölçmedeki yeterliliği, test ve alt test yönergelerinin hedef yaş grubu için uygunluğu, ASIS uygulayıcı kitabının testi uygulama ve yorumlama sürecinde uygulayıcılara ne ölçüde rehberlik ettiği, cevaplama sürelerinin çocuklara uygunluğu değerlendirilmektedir. Anlaşılabilirlik algısı alt boyutunda ise sözel madde kalıplarının uygunluğu, sözel ve görsel alt testlerdeki yönergelerin çocuklar tarafından ne kadar anlaşıldığı, aynı şekilde bu yönergelerin özel eğitim tanısı almış çocuklar için ne kadar anlaşılır olduğu değerlendirilmektedir. Ayrıca ASIS uygulayıcı kitabının uygulayıcılar açısından anlaşılabilirliği ve örnek maddelerin alt testin anlaşılabilirliğini ne ölçüde etkilediği değerlendirilmektedir.

ASISUDA'ya ek olarak uygulayıcılardan "demografik bilgi formu" aracılığıyla veriler toplanmıştır. Demografik bilgi formu aracılığıyla uygulayıcılardan ASIS dışında uyguladıkları diğer zekâ ölçeklerinin isim bilgisi istenmiştir. Buradan elde edilen bilgilere göre değerlendirmeye alınacak katılımcılar belirlenmiştir.

Anketin geliştirilmesi ve kapsam geçerliği

Anketin geliştirilme aşamasında ilk olarak kapsam geçerliliği çalışmaları yapılmıştır. ASIS'in uygulama, puanlama ve yorumlama sürecindeki bütün adımlar dikkate alınarak maddeler yazılmıştır. Ayrıca zekâ ölçeklerine yapılan eleştiriler ve sosyal geçerlik ölçekleri incelenerek 32 maddelik bir havuz oluşturulmuştur. Maddeler sosyal geçerlik alanında çalışma yapmış üç uzman ile birlikte incelenmiştir. İncelemenin sonunda anket amaç, süreç ve ürün olarak üç bileşene ayrılmıştır. Bu aşamadan sonra literatür genişletilerek taranmış ve diğer sosyal bilim alanlarında bulunan sosyal geçerlik ölçekleri incelenmiştir. Yapılan incelemenin sonunda madde havuzu 53 maddeye çıkartılmıştır. 53 madde 8 bileşene ayrılarak incelenmiştir. Bu bileşenler ölçme, pratiklik, anlaşılabilirlik, yeterlilik, kolaylık-zorluk, ilgi çekme-dikkat, uygulanabilirlik ve kabul etme-onaylamadır. Madde sayısı 53 iken ilk değerlendirmeyi yapan iki uzman dil yapısı olarak anketi değerlendirmiştir. Uzman değerlendirmesi sonrasında ölçeğin dil yapısıyla ilgili düzeltmeler yapılmıştır. Bu süreçten sonra aynı uzmanlardan dil ve anketin genel yapısı ile ilgili farklı zaman dilimlerinde dört kez daha dönüt alınmıştır. Uzmanlardan alınan dönütlerle anket kapsamlı bir şekilde revize edilmiş ve dönütlere göre düzeltmeler yapılmıştır. Verilen düzeltmelerin sonunda anketteki madde sayısı 39 maddeye indirilmiştir. Anketin 8 bileşenli yapısı bazı bileşenlerin ortak olması nedeniyle 4 bileşene düşürülmüştür. Bu bileşenler yeterlik

algısı, uygulanabilirlik algısı, anlaşılabilirlik algısı ve beğeni algısıdır. Maddeler üretildikten sonra ASIS'i geliştiren ekipte yer alan doktora eğitimi yapmakta olan yedi araştırmacı ASISUDA'daki maddelerin ölçülmesi amaçlanan örtük yapılara hizmet edip etmemesine göre incelemiştir. Araştırmacıların verdikleri dönütler iki uzmanla tekrar değerlendirilmiştir. Yapılan değerlendirmenin sonunda anket 4 alt boyuttan ve 40 maddeden oluşturulmuştur. Ayrıca ankete verilen cevaplarda ortada yığılma ihtimalini ortadan kaldırmak için uzmanlarla birlikte anketin (1) kesinlikle yanlış ve (6) kesinlikle doğru olacak şekilde altılı derecelendirme ölçeği olması kararlaştırılmıştır.

1. Kesinlikle Katılmıyorum
2. Katılmıyorum
3. Kısmen Katılmıyorum
4. Kısmen Katılıyorum
5. Katılıyorum
6. Kesinlikle Katılıyorum

Şekil 1. Anadolu – Sak Zekâ Ölçeği Uygulayıcı Değerlendirme Anketi Derecelendirme Sistemi

Verilerin Toplanması

ASIS'in sosyal geçerlik çalışması kapsamında, MEB'in görevlendirmiş olduğu 35 uygulayıcıdan ve bir özel okulda görev yapmakta olan 2 Psikolojik Danışmanlık ve Rehberlik Öğretmeninden veri toplanmıştır. Katılımcıların belirlenmesinde ise belirli bir prosedür izlenmiştir. İlk olarak ASIS'in norm çalışmasında MEB 58 uygulayıcıyı görevlendirmiştir. Norm çalışması bittikten sonra MEB 58 uygulayıcıdan 35'ini formatör olması için görevlendirmiştir. Formatör olarak görevlendirilen kişilere ASIS'i geliştiren ekip tarafından Eskişehir'de beş günlük eğitim verilmiştir. Eğitimin sonunda formatörlere ASIS-UDA uygulanarak veriler toplanmıştır.

Verilerin Analizi

Sosyal geçerlik araştırması için toplanan verilerle istatistiksel analizler yapılmamıştır. Gorard (2013) evrenle yapılan çalışmalarda örnekleme değişme olmadığı için istatistiksel analiz yapmaya gerek olmadığını belirtmektedir. Bu nedenle anketteki maddelerin ortalama ve standart sapmaları rapor edilmiştir. Ayrıca yorumlamanın kolay yapılabilmesi için olumsuz olarak kodlanan 10, 12, 21, 33, 35 ve 39. maddelerde tersine kodlama yapılmıştır.

Bulgular

Uygulayıcıların ASIS'e İlişkin Memnuniyet Algıları Nasıldır?

Yapılan çalışmada ASIS uygulayıcılarının ASIS'e yönelik algıları incelenmiştir. Çalışmaya katılan uygulayıcıların verdikleri cevaplar doğrultusunda aritmetik ortalama ve standart sapmalar Tablo 1'de verilmiştir.

Tablo 1. Uygulayıcıların ASIS'e İlişkin Görüşleri

Maddeler	N	\bar{X}	SS
M1	32	5,25	0,88
M2	32	5,5	0,51
M3	32	5,59	0,50
M4	32	5,78	0,42
M5	32	5,5	0,57
M6	32	5,28	0,92
M7	32	5,22	0,75
M8	32	5,03	0,86
M9	32	5,59	0,56
M10	32	3,21	1,16
M11	32	5,72	0,77
M12	32	4,41	1,54
M13	32	5,5	0,72
M14	32	5,38	0,66
M15	32	5,47	0,92
M16	32	5,53	0,57
M17	32	5,63	0,49
M18	32	5,22	0,71
M19	32	5,34	0,70
M20	32	5,41	0,61
M21	32	3,53	1,19
M22	32	5,31	0,59
M23	32	5,41	0,50
M24	32	4,97	1,18
M25	32	5,41	0,76
M26	32	5,72	0,46
M27	32	5,19	0,90
M28	32	5,59	0,56
M29	32	5,5	0,51
M30	32	5,16	1,05
M31	32	5,66	0,55
M32	32	5,38	0,55
M33	32	5,03	1,20
M34	31	5,45	0,80
M35	32	4,81	1,03
M36	31	5,23	0,61
M37	32	5,53	0,57
M38	32	4,75	1,24
M39	32	5,72	0,58
M40	32	5,31	0,47

Maddeler incelendiğinde 10. maddenin 3.21 ortalama ile en düşük ortalamaya sahip olduğu (görsel alt testlerin yönergeleri özel eğitim tanısı almış -Otizm, Zihinsel gelişim geriliği vb.- çocuklar için yetersizdir), en yüksek ortalamanın ise 5.78 ile 4. madde olduğu (Alt test puanlarının standart puanlara dönüştürülmesi kolaydır.) görülmüştür. Ayrıca en düşük

ortalamaya sahip iki madde (10 ve 21) alt test yönergelerinin özel eğitim tanısı alan çocuklar için ne kadar yeterli olduğunu belirlemeye yönelik maddedir.

Maddelerin ortalamaları uygulayıcıların algılarına göre incelendiğinde 3. madde ve 5. maddenin kısmen katılmıyorum ile kısmen katılıyorum arasında olduğu, 12. madde, 24. madde, 35. madde ve 38. maddenin kısmen katılıyorum ile katılıyorum arasında olduğu ve geri kalan maddelerin katılıyorum ile kesinlikle katılıyorum arasında olduğu görülmüştür. Anketin ortalamasına göre bir değerlendirilme yapıldığında ise 2. madde, 3. madde, 4. madde, 5. madde, 6. madde, 9. madde, 11. madde, 13. madde, 14. madde, 15. madde, 16. madde, 17. madde, 19. madde, 20. madde, 21. madde, 22. madde, 23. madde, 25. madde, 26. madde, 28. madde, 29. madde, 31. madde, 32. madde, 34. madde, 37. madde, 39. madde ve 40. madde testin ortalamasından yüksek bir değere sahiptir.

Maddelerin standart sapmaları incelendiğinde 10, 12, 21, 23,30, 33 ve 35. maddelerin standart sapmalarının 1.00'den büyük olduğu görülmektedir. Bu maddelerde uygulayıcılar daha farklı görüşler bildirmiştir. Uygulayıcıların değerlendirme anketine verdikleri puanların ortalaması 5,26 olarak belirlenmiştir.

Tartışma ve Sonuç

Bu bölüm iki kısımdan oluşmaktadır. İlk kısımda elde edilen sonuçlar literatür ile karşılaştırılarak tartışılmıştır. İkinci kısımda ise araştırmacılara ve test geliştiricilere önerilerde bulunulmuştur.

Sonuç ve tartışma

Yapılan çalışmada ASIS uygulayıcılarının bakış açısıyla ASIS'in memnuniyet algısı ya da diğer bir ifadeyle sosyal geçerliği araştırılmıştır. Araştırmanın katılımcılarını ASIS uygulayıcısı olup farklı zekâ testleri uygulaması da yapan uzmanlar oluşturmuştur. Farklı iki ya da daha fazla zekâ testi uygulayan uygulayıcılar zekâ testleri arasında kıyaslama yapma deneyimleri olması nedeniyle daha eleştirel ve doğru değerlendirme yapabilmektedir. Uygulayıcıların 1-6 arası puanlama içeren ankete verdikleri puanların ortalaması 5,26 olarak belirlenmiştir. Bu değer ASIS'in sosyal geçerlik düzeyinin oldukça yüksek olduğunu göstermektedir. Ayrıca anketin güvenilirlik katsayısı .87 olarak belirlenmiştir. Katılımcı sayısının az olmasına rağmen bu değer oldukça yüksek bir değerdir. Maddeler 1 (kesinlikle yanlış) ile 6 (kesinlikle doğru) arasında derecelendirilmiştir. Verilen cevaplar incelendiğinde ise 40 maddeden 2 tanesinin ortalama değeri 3 ile 4 arasında çıktığı görülmüştür. Bu maddeler 10.madde (*Görsel alt testlerin yönergeleri özel eğitim tanısı almış (Otizm, Zihinsel gelişim geriliği vb.) çocuklar için yetersizdir.*) ve 21.maddedir (*Sözel alt testlerin yönergeleri özel eğitim tanısı (Otizm, Zihinsel gelişim geriliği vb.) almış çocuklar için yetersizdir.*). Her iki madde de alt test yönergelerinin özel eğitim tanısı alan öğrenciler için yeterliliğini sorgulamaktadır. Bu bağlamda ASIS'in özel gruplar arası ayırt ediciliği incelendiğinde ayırt edicilik geçerliliğinin oldukça yüksek (Sak ve diğ., 2016, s.91) olmasına rağmen uygulayıcılar ASIS'in alt test yönergelerinin özel eğitim tanılı öğrenciler için yeterli görmemektedir. Bu yargı, uygulayıcıların henüz ASIS'i özel eğitim tanısı olan öğrencilere yeterince uygulamamış olmalarından kaynaklanıyor olabilir. Öte yandan alt test yönergelerinin özel eğitim grupları göz önünde bulundurularak yeniden düzenlenmesi ASIS'in geçerliğine ve güvenilirliğine ayrı bir katkı sağlayabilir. Bu şekilde özel eğitim

gruplarında hem ayırt ediciliği yüksek hem de uygulayıcıların kullanmak isteyeceği şekilde yönergeler düzenlenebilir. Bu sayede sosyal geçerlik düzeyi daha üst düzeylere çekilerek uygulayıcıların kullanırken tereddüt etmeyeceği bir zekâ ölçeği olması sağlanabilir.

Ortalamalar incelendiğinde ise 12. madde (*ASIS uygulaması çocuklara uzun gelmektedir.*), 24. madde (*Uyaran kart çocukların teste başlamasına yardımcı olmaktadır.*), 35. madde (*Çocuklar ASIS uygulanırken sıkılıyorlar.*), 38. madde (*Uyaran kart çocukların ilgisini çekmektedir.*) 4 ile 5 arasında değer almıştır. Maddeler incelendiğinde ise 12. madde ve 35. madde testin süresi ve çocukların sıkılmaları ile ilgilidir. Teste katılım gösteren çocuklar madde tiplerinden sıkılacağı gibi 12. maddede bahsedildiği gibi testin uzun sürmesinden kaynaklı da sıkılabilirler. Her iki maddenin de yakın değerlere sahip olması bu ilişkiyi güçlendirmektedir. 24. madde ve 38. madde incelendiğinde ise her iki maddede uyaran kart ile ilgili maddeler olduğu görülmektedir.

Uygulayıcılar uyaran kart hakkında olumlu değerlendirmede bulunsalar dahi elde edilen değerler uyaran kartların işlevini beklenen düzeyde yerine getirmediğini göstermektedir. Geri kalan maddeler ise 5 ve üzeri bir ortalama değere sahiptir. Bu maddeler doğru ile kesinlikle doğru arasında değerler almasından dolayı beklenen düzeyi sağlamışlardır.

En yüksek ortalamaya sahip madde ise 5.78 ortalama ile 4. maddedir (*Alt test puanlarının standart puanlara dönüştürülmesi kolaydır.*). 4. maddenin yüksek çıkmasının en başlıca nedeni ASIS'in uygulayıcı kitabında ham puanların endeks puanlarına dönüşümünün anlaşılır ve kolaylık sağlayacak şekilde sunulması olabilir. Buna ek olarak ASIS'in WEB tabanlı puan dönüşüm sistemine sahip olması ölçeğin endeks puanlarının hesaplanmasını kolaylaştırmaktadır. Ayrıca 11. madde (*ASIS'in ileride daha geniş yaş aralığı için geliştirilmesini isterim.*), 26. madde (*ASIS sorularının cevaplama süresi yeterlidir.*) ve 39. madde (*ASIS zekâyı ölçmez.*) 4.72 ortalama ile en yüksek ortalamaya sahip diğer maddelerdir. Maddeler incelendiğinde uygulayıcılar ASIS'in zekâyı doğru bir şekilde ölçtüğünü düşünmektedir. Buna bağlı olarak uygulayıcılar ASIS'in daha geniş bir yaş aralığını tanılayacak şekilde genişletilmesini talep etmektedir. Bu iki madde uygulayıcıların ASIS'e olan güvenlerinin yüksek olduğunu göstermektedir.

Anketteki maddeler incelendiğinde, uygulayıcılar ASIS'in zekâyı ölçme, zekâ profilleri hakkında bilgi verme, sözel, görsel ve bellek alanlarında doğru bir tanılama yaptığını düşünmektedir. Ayrıca ASIS'in geçerli ve güvenilir ölçümler yaptığını inanmaktadırlar. Uygulayıcılar ASIS'i uygulanabilir bir zekâ ölçeği olarak nitelendirmişlerdir. Bunun nedeni puan dönüşümlerinin kolay olması, kayıt formunun ve test materyallerinin uygulama sırasında kolaylık sağlaması olabilir. Ayrıca çocuklar açısından değerlendirildiğinde alt testlerin sıralanışını uygun ve uygulanabilir görmektedirler. Uygulayıcılar ASIS'i anlaşılır bulmuştur. Uygulamaya başlama yönergesi ve alt test yönergelerinin anlaşılır bir şekilde yazılması bu yargının oluşmasında etkilidir. Ayrıca örnek maddelerin ve açıklamaların da çocukların yaş grubuna uygun olarak yazılması uygulayıcıların yönergeleri anlaşılır bulmasında etkilidir.

Uygulayıcıların ASIS'e yönelik beğeni algıları yüksek bulunmuştur. Beğeni algısının yüksek olması ASIS'in uygulama sürecinin, puanlama ve yorumlama sürecinin uygulayıcılar ve çocuklar için uygun ve pratik olmasından kaynaklanmaktadır. Uygulayıcılar çocuklara zekâ testi uygulayacak olduklarında ASIS'i tercih edecekleri yönünde ağırlıklı görüş bildirmişlerdir. Ayrıca ASIS'i diğer zekâ testi uygulayıcılarına tavsiye edeceklerini belirtmişlerdir. Bu sonuçlar

ASIS'in uygulayıcılar tarafından nitelikli bir test olarak değerlendirildiğinin göstergesi olarak kabul edilebilir.

Öneriler

Araştırmadan elde edilen bulgular sonunda, gelecekte ASIS ile ilgili çalışma yapacak olan araştırmacılara ve uygulayıcılara aşağıdaki öneriler sunulabilir.

- Uygulayıcılar ASIS'in alt test yönergelerini özel eğitim tanısı almış öğrenciler için yetersiz görmüştür. Yapılacak nitel bir çalışma ile eksikliğin kaynağı tespit edilebilir ve yönergelerin düzenlenmesi sağlanabilir.
- ASIS'in sosyal geçerliği örneklem grubu çocuklardan oluşturularak araştırılabilir.

Anadolu-Sak Zekâ Ölçeği'nin (ASIS) Sosyal Geçerlik Çalışması başlıklı çalışmanın yazım sürecinde bilimsel, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamış, karşılaşılabilecek tüm etik ihlallerde "**Pamukkale Üniversitesi Eğitim Fakültesi Dergisi Yayın Kurulunun**" hiçbir sorumluluğunun olmadığı, tüm sorumluluğun Sorumlu Yazara ait olduğu ve bu çalışmanın herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiş olduğunu taahhüt ederim.

Kaynakça

- A. S. Kaufman., E. O. Lichtenberger, E. Fletcher – Janzen, & N. L. Kaufman. (Eds.) (2010). *Essentials of KABC - II assessment*. New Jersey: John Wiley & Sons, Inc.
- B. Devlin, S. E. Fienberg, D.P. Resnick, & K. Roeder. (Eds.). (1997). *Intelligence, genes, and success*. New York: Springer.
- Brim, O. G., Glass, D. C., Neulinger, J., & Firestone, I. J. (1969). *American beliefs and attitudes about intelligence*. New York: Russell Sage Foundation.
- Brody, N., (1999). What is intelligence?. *International Review of Psychiatry*, 1(1), 19- 25.
- Büyüköztürk, Ş. (2016). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi.
- Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2015). *Bilimsel araştırma yöntemleri*. Ankara: Pegem Akademi.
- Çüm, S. ve Gelbal, S. (2015). Kayıp veriler yerine yaklaşık değer atamada kullanılan farklı yöntemlerin model veri uyumu üzerindeki etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 1(35), 87-111.
- Eysenck, H. J. (1967). Intelligence assessment: A theoretical and experimental approach. *British Journal of Educational Psychology*, 37(1), 81-98.
- Finney, J. W. (1991). On further development of the concept of social validity. *Journal of Applied Behavior Analysis*, 24(2), 245-249.
- Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of consulting and clinical psychology*, 67(3), 308.
- Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. Newyork: Basic books.
- Goddard, H. H. (1946). What is intelligence?. *The Journal of social psychology*, 24(1), 51-69.
- Gorard, S. (2013). *Research design: Robust approaches for social science*. London: SAGE.
- Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence*, 24(1), 13-23.
- Hurley, J. J., Wehby, J. H., & Feurer, I. D. (2010). The social validity assessment of social competence intervention behavior goals. *Topics in early childhood special education*, 30(2), 112-124.
- Jensen, A. R., & Weng, L. J. (1994). What is a good g?. *Intelligence*, 18(3), 231-258.
- I. B. Weiner. (Eds.). (2012). *Handbook of Psychology*. New Jersey: John Wiley & Sons.
- Kaplan, A. (2008). *Raven'in ilerleyen matrisler plus testinin 12-13 yaş çocukları üzerinde geçerlik, güvenilirlik ve ön norm çalışmalarına göre üstün zekâlı olan ve olmayan öğrencilerin mantıksal düşünme yeteneklerinin karşılaştırılması*. Yayınlanmamış Yüksek Lisans Tezi. İstanbul: İstanbul Üniversitesi Sosyal Bilimler Enstitüsü.
- Karasar, N. (2014). *Bilimsel araştırma yöntemi*. Ankara: Nobel.
- Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behavior Modification*, 1(4), 427-452.
- Kızıltepe, Z. (2004). *Öğretim: Eğitim psikolojisine çağdaş bir yaklaşım*. İstanbul: Ofset Yayınevi.
- Kontos, S., Carter, K. R., Ormrod, J. E., & Cooney, J. B. (1983). Reversing the revolving door: A strict interpretation of Renzulli's definition of giftedness. *Roeper Review*, 6(1), 35-39.
- O. Lichtenberger, & A. S. Kaufman, (Eds.). (2013). *Essentials of WAIS – IV assessment*. New Jersey: John Wiley & Sons, Inc.
- Öner, N. (1996). *Türkiye'de kullanılan psikolojik testler*. İstanbul: Boğaziçi Üniversitesi Yayınları.
- Önöz, B. ve Bayazit, M. (2003). The power of statistical tests for trend detection. *Turkish journal of engineering and environmental sciences*, 27(4), 247-251.
- Özgüven, İ. E. (2014). *Psikolojik testler*. Ankara: Nobel yayınları.
- Pollay, R. W. (1984). Twentieth-century magazine advertising: Determinants of informativeness. *Written Communication*, 1(1), 56-77.

- Raz, N., Willerman, L., & Yama, M. (1987). On sense and senses: intelligence and auditory information processing. *Personality and Individual Differences*, 8(2), 201-210.
- Roid, G. H., & Barram, R. A. (2004). *Essentials of Stanford-Binet intelligence scales (SB5) Assessment*. New Jersey: John Wiley & Sons.
- Sak, U. (2014). *Üstün zekâlılar: Özellikleri, tanılanmaları, eğitimleri*. Ankara: Vize.
- Sak, U., Bal Sezerel, B., Ayas, B., Tokmak, F., Özdemir, N., Demirel Gürbüz, Ş. ve Öpengin, E. (2016). *Anadolu Sak Zekâ Ölçeği (ASIS) uygulayıcı kitabı*. Eskişehir: Anadolu Üniversitesi ÜYEP Merkezi.
- Schwartz, I. S., & Baer, D. M. (1991). Social validity assessments: Is current practice state of the art?. *Journal of applied behavior analysis*, 24(2), 189-204.
- Seçer, İ. (2015). *Psikolojik test geliştirme ve uyarlama süreci: SPSS ve LISREL uygulamaları*. Anı yayıncılık.
- Selçuk, Z., Kayılı, H. ve Okut, L., (2002), *Çoklu Zekâ Uygulamaları (2.baskı)*, Ankara: Nobel Yayınları.
- Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. *American Psychologist*, 42(2), 137-144.
- Suveren, S. (2006). *Anasınıfına devam eden çocuklar arasından üstün yetenekli olanların belirlenmesi*. Yayınlanmamış Yüksek Lisans Tezi. Bolu: Abant İzzet Baysal Üniversitesi.
- Weiner, B.I. (Eds). (2012). *Handbook of Psychology Second Edition*. New Jersey: John Wiley & Sons.
- Winett, R. A., Moore, J. F., & Anderson, E. S. (1991). Extending the concept of social validity: Behavior analysis for disease prevention and health promotion. *Journal of Applied Behavior Analysis*, 24(2), 215-230.
- Wissler, C. (1901). The correlation of mental and physical tests (Abstract). *The Psychological Review: Monograph Supplements*, 3(6), i-62.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of applied behavior analysis*, 11(2), 203-214.

Extended Abstract

Introduction

Social validity of intelligence tests has not become a subject of research though intelligence is a social construct. Its measurement by any instruments has not been socially investigated. The aim of this research was to examine the social validity of the Anadolu Sak Intelligence Scale through an investigation of its users' perceptions. Psychological and educational measurement tools serve two purposes (Özgül, 2014). The first one can be considered as diagnostic, placement and research use in clinical settings. The second one as measurement, selection and placement in special educational environments (Özgül, 2015; Sezer, 2014). The purpose of a measurement tool and its capacity to achieve this purpose partially determine its social acceptance.

Methodology

In order to investigate the social validity of the Anadolu Sak Intelligence Scale, its field testers were surveyed. These testers were working for the Ministry of Education. The participants included 37 testers who administered the ASIS during its norming study. They had sufficient experience with the use of ASIS because they had a five-day training for using the ASIS and administered a certain number of tests to children, including the ASIS. Because five of the testers did not have experience with another test, except the ASIS, their data was excluded from data analysis. The final sample included 32 participants (17 males (53.13%) and 15 females (46.87%). ASIS Practitioner Evaluation Form was used to collect data. The instrument included 40 items scored on a 6-point Likert type scale and related to the efficiency of measuring intelligence and its components, practicality of use (administration, scoring, and interpretation), simplicity, and preference of test users. In addition, some items were related to a comparison of ASIS to other tests of intelligence that the testers had experience before. Data was collected after the testers completed the norming testing. A descriptive analysis was carried out to explore the participants' perceptions about the ASIS and to find out the social acceptance of the ASIS by the testers. Data analysis was carried out item by item and by total score. A classification of rating means was made as very high between 5 and 6, high between 4 and 5, medium between 3 and 4, low between 2 and 3, and very low between 1 and 2.

Findings

Analysis showed that the overall mean of the questionnaire items was 5.26/6.00, a very high level between agree and strongly agree. The item 10 had the lowest mean, a medium level of acceptance (3.21; instructions visual subtests of the ASIS are insufficient for children with a disability (e.g., intellectual or autism). The item 4 had the highest mean, a very high level of acceptance (5.78; It is easy to convert raw scores into standard scores). Two items with the lowest means were related to the adequacy of the subtest instructions for children diagnosed with intellectual disability. Furthermore, the items 3 and item 5 were rated to be between partially disagree and partially agree, respectively. The items 12, 24, 35 and 38 were rated to be in between partially agree and agree. The rest of the items were rated to be between agree and strongly agree. The following items had a mean score higher than the overall mean: items 2, 3,

4, 5, 6, 9, 11, 13, 14, 15, 16, 17, 19, 20, 21, 22, 23, 25, 26, 28, 29, 31, 32, 34, 37, 39, and 40. The findings showed a high level of social acceptance in general.

Discussion

According to the findings, test users in general believe that ASIS's measures of children's intelligence is valid and reliable. It measures children's capacity in verbal, visual and memory domains. It provides sufficient profile analyses. It is easy to learn, administer, score and interpret. Test users consider the order of the subtests and items within these subtests in the test booklet to be appropriate for administration easiness and for children not to be bored or disinterested. According to their perceptions, test items are appropriate for the age range of the test and children are highly engaged when solving ASIS problems. They showed a high level of preference to administer the ASIS among other intelligence tests when they need to test a child. They like to see the ASIS to include a wider age in future revisions. Furthermore, the findings imply adaptations in instructions and beginning items for children with low functioning.