



Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi
Ondokuz Mayıs University Journal of Faculty of Education

e-ISSN: 2548-0278 OMU EFD, June 2023, 42(1): 161-194

Özel Eğitim Okullarında Geliştirilen Bireyselleştirilmiş Eğitim Programlarının (BEP) Kalitesinin İncelenmesi

Quality of Individualized Education Programs (IEP)
Developed in Special Education Schools

Şükran ALAN¹, Veysel AKSOY²

¹ MEB

· sukranalan@gmail.com · ORCID > 0000-0002-8321-4543

²Anadolu Üniversitesi, Eskişehir

· vaksoy@anadolu.edu.tr · ORCID > 0000-0003-1020-8326

Makale Bilgisi/Article Information

Makale Türü/Article Types: Araştırma Makalesi/Research Article

Geliş Tarihi/Received: 09 Aralık/December 2022

Kabul Tarihi/Accepted: 28 Nisan/April 2023

Yıl/Year: 2023 | **Cilt-Volume:** 42 | **Sayı-Issue:** 1 | **Sayfa/Pages:** 161-194

Atıf/Cite as: Alan, Ş., Aksoy, V. "Özel Eğitim Okullarında Geliştirilen Bireyselleştirilmiş Eğitim Programlarının (BEP) Kalitesinin İncelenmesi-Quality of Individualized Education Programs (IEP) Developed in Special Education Schools"

Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, Ondokuz Mayıs University Journal of Faculty of Education, 42(1), June 2023: 161-194.

Sorumlu Yazar/Corresponding Author: Veysel AKSOY

Etik Kurul Beyanı/Ethics Committee Approv: "Araştırma için Anadolu Üniversitesi Sosyal ve Beşeri Bilimler Bilimsel Araştırma ve Yayın Etiği Kurulundan 26.02.2018 tarihli ve 19113 karar sayısı ile etik kurul izni alınmıştır."

<https://doi.org/10.7822/omuefd.1216755>

ÖZEL EĞİTİM OKULLARINDA GELİŞTİRİLEN BİREYSELLEŞTİRİLMİŞ EĞİTİM PROGRAMLARININ (BEP) KALİTESİNİN İNCELENMESİ

ÖZ

Bireyselleştirilmiş eğitim programı özel eğitimin temel ilkelerinden biridir. Bireyselleştirilmiş eğitim programında amaçların ve hizmetlerin belirlendiği biçimde uygulandığından ve izlendiğinden emin olmak için bireyselleştirilmiş eğitim programı kalitesinin belirlenmesi ve takip edilmesi gerektiği ifade edilmektedir. Bireyselleştirilmiş eğitim programlarının kalitesinin genel olarak düşük olduğu ifade edilmektedir. Ayrıca kalite incelemesi amacıyla kullanılan araçların psikometrik niteliklerinde sınırlılıklar söz konusudur. Bu araştırmanın amacı; bireyselleştirilmiş eğitim programı kalitesini değerlendirmeyi hedefleyen geçerli ve güvenilir bir ölçme aracı geliştirmek ve geliştirilen araçla özel eğitim okullarında geliştirilen bireyselleştirilmiş eğitim programının kalitelerini incelemektir. Araştırmanın birinci aşamasında Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeği geliştirilmiş ve 180 bireyselleştirilmiş eğitim programı kullanılarak geçerliliği ve güvenilirliği sınanmıştır. İkinci aşamasında Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeği kullanılarak, 211 bireyselleştirilmiş eğitim programının kalitesi incelenmiştir. Araştırma sonucunda Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeğinin geçerli ve güvenilir bir ölçek olduğu, bireyselleştirilmiş eğitim programlarının kalite puanlarının düşük olduğu, amaç sayılarının çok yüksek olduğu, okul kademesi yükseldikçe kalite puanlarının düştüğü, tanıya ve öğretmen mesleki deneyimine göre kalite puanının farklılaştığı görülmüştür. Elde edilen sonuçlar; öğretmenlerin bireyselleştirilmiş eğitim programı hazırlamakta güçlük çektiğini, eğitsel bir rehber olarak bireyselleştirilmiş eğitim programlarının kullanılmadığını ve bireyselleştirilmiş eğitim programlarının yalnızca yasal zorunluluk sebebiyle bir belge olarak hazırlandığını düşündürmektedir.

Anahtar Kelimeler: Özel Eğitim, Bireyselleştirilmiş Eğitim Programı, Bep Kalitesi, Bep Kalite Ölçeği.



QUALITY OF INDIVIDUALIZED EDUCATION PROGRAMS (IEP) DEVELOPED IN SPECIAL EDUCATION SCHOOLS

ABSTRACT

Individualised education programs one of the basic principles of special education. It is stated that individualised education programs quality should be monitored to ensure that the goals and services determined in individualised education programs are implemented and monitored as planned. Studies examining individualised education programs concluded that quality levels were generally low. However, it was also observed that there were limitations in the psychometric properties of the tools used in these studies. This study's purpose is to develop a valid and reliable tool to evaluate individualised education programs quality and examine the qualities of individualised education programs using the created tool. At the first stage of the study, the Individualised Education Program Quality Assessment Scale was developed, and its validity and reliability were tested using 180 individualised education programs. At the second stage, the quality of 211 individualised education programs was examined using the scale. Results indicated that the Individualised Education Program Quality Assessment Scale was a valid and reliable scale, the quality scores of individualised education programs were low, the number of goals was very high, the quality scores decreased as the school stage increased, and they differed according to the diagnosis and professional experience of teachers. The results suggest that teachers experience difficulty preparing individualised education programs, individualised education programs are not used as educational guides, and individualised education programs are prepared as documents only due to legal necessity.

Keywords: Special Education, Individualised Education Program, Iep Quality, Iep Quality Scale.



GİRİŞ

Bireyselleştirilmiş eğitim programları (BEP) öğrencinin var olan performans düzeyini, ölçülebilir yıllık amaçları, bu amaçlara ulaşmayı ve çocuğun eğitiminden yararlanmasını sağlayacak belirli özel eğitim ve ilişkili hizmetleri kapsar (Heward, vd., 2017). BEP eğitsel bir yol haritasıdır (Ruble vd., 2010; Sanches-Ferreira vd., 2013; Rakap, 2015; Gibb ve Dyches, 2016) ve öğrenciler, aileler, eğitimciler için kritik bir öneme sahiptir (Christle ve Yell, 2010). BEP bireylerin kendilerine özel olarak sunulan hizmetlere ulaşabilmeleri için temel bir araç niteliği taşır (Prettie-Frontzck ve Bricker, 2000) ve eğitim/öğretim, izleme ve değerlendirme, yönetme, iletişim ve sorumluluk gibi amaçlara hizmet eder (Bateman ve Cline,

2016). Ayrıca BEP planlamanın güvencesidir (Fiscus ve Mandell, 2002). BEP, öğrencilerin özel gereksinimlerine yanıt üreten bireyselleştirilmiş kısa ve uzun dönemli amaçların belirlenmesi ve izlenmesini mümkün kılan, kapsamlı bir araçtır (Ruble vd., 2010).

BEP'in ve hizmet ettiği amaçların önemi; araştırmaların BEP'lerin ne kadar işe yaradığını ya da amaca uygun yapıp yapılmadığını konu etmelerine neden olmuş (Örn., Boavida, vd., 2014; Hunt, vd., 1986; Prettie-Frontzcak ve Bricker, 2000; Rakap, 2015) ve BEP kalitesi kavramını ortaya çıkarmıştır. Araştırmalarda BEP'in kalitesinin, BEP'te yer alan hedeflerin ve hizmetlerin planlandığı şekilde uygulanması ve izlenmesinden emin olmak için incelenmesi gerektiği belirtilmektedir (Roach vd., 2009). BEP kalitesinin incelendiği çalışmalarda BEP kalitesiyle ilişkili olarak; var olan performans düzeyi (Le Salle vd., 2010; Ruble vd., 2010; Smith ve Simpson, 1989) işlevsellik (Boavida vd., 2010; Hunt vd., 1986; Rakap, 2015; Sanches-Ferreira vd., 2013) ölçülebilirlik-gözlenebilirlik (Boavida vd., 2010; Prettie-Frontzcak ve Bricker, 2000; Rakap, 2015; Ruble vd., 2010; Sanches-Ferreira vd., 2013), ölçüt (Boavida vd., 2010; Lynch ve Beare, 1990; Prettie-Frontzcak ve Bricker, 2000; Rakap, 2015; Ruble vd., 2010; Smith ve Simpson, 1989), genelleme, (Boavida vd., 2010; Hunt vd., 1986; Prettie-Frontzcak ve Bricker, 2000; Sanches-Ferreira vd., 2013), amaçların uygunluğu/uyumu (Catone ve Brady, 2005; Lynch ve Beare, 1990; Prettie-Frontzcak ve Bricker, 2000; Smith ve Simpson, 1989) gibi çeşitli BEP bileşenlerinin ele alındığı ancak açık (operational) bir tanımın yapılmadığı görülmektedir. Bu çalışmada çok bileşenli bir yapı olarak ele alınan BEP kalitesi; "Öğrenci gereksinimlerini temel alarak BEP'te bulunması gereken öğelerin bulunma durumu ve bu öğelere ilişkin ifadelerin yeterli, doğru ve açık olarak belirtilme düzeyi" olarak tanımlanmıştır. Çünkü, BEP kalitesine yönelik çalışmalarda BEP kalitesinin; BEP'in bulundurması gereken öğeleri içerme durumu (Hunt vd., 1994; Lucas vd., 2014; Lynch ve Beare, 1990; Rakap, 2015; Rowland, 2015; Ruble vd., 2010; Sanches-Ferreira vd., 2013; Smith ve Simpson, 1989) ve/veya bu öğelerin doğru olarak ifade edilme düzeyi (Boavida vd., 2014; Le Salle vd., 2010; Lowman 2016; Rakap, 2015) olarak iki temel boyutta ele alındığı görülmektedir.

Araştırmalar, incelenen BEP'lerin kalitesinin genel olarak düşük olduğunu (Boavida vd., 2010; Giangreco vd., 1994; Lych ve Beare, 1990; Prettie-Frontzcak ve Bricker, 2000; Rakap, 2015; Ruble vd., 2010; Sanches-Ferreira vd., 2013; Smith ve Simpson, 1989;), çok fazla amaç ifadesi içerdiğini (Boavida vd., 2010; Giangreco vd., 1994; Rakap, 2015; Sanches-Ferreira vd., 2013;), BEP kalitesinin tanı gruplarına göre farklılaştığını (Rakap, 2015), okul kademesi yükseldikçe BEP kalitesi puanlarının düştüğünü (Catone ve Brady 2005; Sanches-Ferriara vd., 2013) ortaya koymaktadır. Bununla birlikte BEP kalitesini incelemeyi amaçlayan araştırmalarda genel olarak BEP amaçları ele alınmış (Boavida vd., 2014; Boavida vd., 2010; Catone ve Brady, 2005; Giangreco vd., 1994; Hunt vd., 1986; Le Salle vd., 2010; Lowman, 2016; Lynch ve Beare, 1990; Prettie-Frontzcak ve Bricker, 2000; Rakap, 2015; Row-

land vd., 2015; Ruble vd. 2010; Sanches-Ferreira vd., 2013) yalnızca bir araştırmada (Smith ve Simpson, 1989) BEP holistik (bütünsel) olarak incelenmiştir. Ayrıca bu araştırmalarda kullanılan ölçme araçlarının psikometrik niteliklerinin açıklanmadığı ya da zayıf olduğu görülmüştür. Kullanılan ölçme araçlarının geçerliliğine ilişkin olarak; uzman görüş alındığı (Hunt vd., 1986), uyum geçerliliğine bakıldığı (Boavida vd., 2010; Lynch ve Beare, 1990; Rakap, 2015), yasaların/kurumsal göstergelerin temel alındığı (Ruble vd., 2010; Smith ve Simpson, 1989;), geçerli olduğunun ifade edildiği ancak başka bilginin bulunmadığı (Prettie-Frontzcak ve Bricker, 2000) ya da hiçbir bilginin bulunmadığı görülmüştür (Boavida vd., 2014; Le Salle vd., 2010; Lowman, 2016). Güvenirlik için genellikle değerlendiriciler/ puanlayıcılar arası güvenirlik analizleri ön plana çıkmaktadır (Boavida vd., 2014; Catone ve Brady, 2005; Hunt vd., 1986; Le Salle vd., 2010; Rakap, 2015; Ruble vd., Sanches-Ferriara vd., 2013). Sonuç olarak araştırmalarda BEP kalitesini inceleme amacıyla kullanılan araçların geçerlik ve güvenirlik bulgularının eksik olduğu ve psikometrik niteliklerinin zayıf olduğu görülmektedir.

Tüm bu bilgiler doğrultusunda BEP'in var olan performans düzeyi, kısa ve uzun dönemli amaçlar, öğretim yöntem ve teknikleri, materyaller, zaman gerekleri gibi birden çok öğeyi barındırmasına rağmen, BEP kalitesini bütüncül olarak ele alan çalışmaların sınırlı sayıda olması ve araştırmalarda kullanılan BEP kalitesi değerlendirme araçlarının psikometrik niteliklerinin zayıf olması göz önünde bulundurularak bu araştırma iki aşamalı olarak planlanmıştır. Araştırmanın ilk aşamasında BEP kalitesini bütüncül olarak ele alan geçerliği ve güvenirliliği sağlanmış bir ölçme aracı geliştirme, ikinci aşamasında geliştirilen araçla özel eğitim okullarında geliştirilen BEP'lerin kalitelerinin incelenmesi hedeflenmiştir. Bu kapsamda şu sorulara yanıt aranmıştır: (1) BEP Kalitesini Değerlendirme Ölçeğinin (BEPK-DÖ); (a) geçerliliği nasıldır?, (b) güvenirliliği nasıldır?, (2) Özel eğitim okullarında geliştirilmiş olan BEP'lerin kaliteleri nasıldır?, (3) Özel eğitim okullarında geliştirilmiş olan BEP'lerin kaliteleri; (a) amaç sayılarına (b) tanı gruplarına (c) okul kademelerine (d) öğretmenlerin özel eğitim deneyimlerine göre anlamlı düzeyde farklılaşmakta mıdır?

YÖNTEM

Araştırma Modeli

Bu çalışmada nicel araştırma yöntemlerinden betimsel araştırma yöntemi kullanılmıştır. Araştırmanın özellikleri gereği model kesitsel tarama modeline karşılık gelmektedir. Belirli bir zaman aralığında ve tek bir defada örneklemeden veri alınması yoluyla yapılan araştırmalar bu modeli açıklamaktadır (Gay vd.,2012; Cresswell, 2012). Bu yolla yapılan çalışmaların amacı var olan durumu ortaya koymaktır (Gay vd.,2012).

Çalışma Grubu

Araştırmanın ilk aşaması için Eskişehir ilindeki devlete bağlı tüm özel eğitim okullarından 180 BEP temin edilmiştir. İlk aşamada ulaşılan 180 BEP yalnızca geliştirilen ölçme aracının standardizasyonunda kullanılmış, BEP kalitesiyle ilişkili olabilecek değişkenlere ilişkin veri toplanmamıştır. Araştırmanın ikinci aşamasında aynı ildeki devlete bağlı tüm özel eğitim okullarından 211 BEP temin edilmiştir. Toplamda 391 BEP incelenmiştir.

İkinci aşamada ilişkili olabilecek değişkenlere ilişkin veriler de toplanmıştır. BEP'leri incelenen öğrencilerin 71'i kız, 130'u erkektir, 10 öğrencinin ise cinsiyet verisine ulaşılamamıştır. BEP'lerin 93'ü (%44,1) hafif düzeyde zihin yetersizliği (HZY), 45'i (21,3) orta düzeyde zihin yetersizliği (OZY), 72'si (34,1) ağır düzey zihin yetersizliği (AZY) ya da otizm spektrum bozukluğu (OSB), biri (% 0,5) işitme yetersizliği tanısına sahip öğrencilere aittir. BEP'lerin; 11'i (% 5,2) okul öncesi, 67'si (%31,2) ilkokul (1-4), 58'i (%28) ortaokul (5-8) ve 75'i (%35,5) lise (9-12) düzeyinde eğitim gören öğrencilere aittir. BEP'leri geliştiren 75 özel eğitim öğretmenin 65'i kadın, 10'u erkektir. Kadın öğretmenlerin yaş ortalaması 34,54 (min=23, max=54, ss=7,04), erkek öğretmenlerin yaş ortalaması 39,10 (min=26, max=50, ss=8,86), tüm öğretmenlerin yaş ortalaması ise 34,15'tir (min=23, max=54, ss=7,05). Öğretmenlerin özel eğitim deneyimleri ortalaması 10,03 yıldır (min=1, max=29, ss=6,93).

Veri Toplama Araçları

Rubrik, performansın ya da ürününün değerlendirilmesinde kullanılan değerlendirme aracıdır (Mertler, 2001). Çok boyutlu ürün veya performansların bulunması halinde, boyutlar ayrıştırılabiliyor ve detaylı puanlama yapılması talep ediliyorsa analitik rubriklerin kullanılması, önerilmektedir (Haladyna, 1997). Bireyselleştirilmiş Eğitim Programı, BEP ekibi faaliyetlerinin bir ürünüdür ve birbirinden ayrıştırılabilen çok boyutlu bir yapıya sahiptir. Bu nedenle BEP kalitesinin, bir analitik rubrikle değerlendirilmesine karar verilmiştir.

BEP'e ilişkin alanyazın ve yasalar incelenerek BEP bileşenlerine ilişkin 19 rubrik maddesi hazırlanmıştır. Her bir maddeye ilişkin üçlü derecelendirmeler ve bu derecelendirmelere ilişkin de açıklamalar oluşturulmuştur. Birinci aşamada 180 BEP ile geçerlik ve güvenilirlik analizleri yapılmış, bu analizler sonucunda 16 maddeden oluşan *BEP Kalitesini Değerlendirme Ölçeği* (BEPKDÖ) geliştirilmiştir. İkinci aşamada BEPKDÖ ile 211 BEP'in kaliteleri değerlendirilmiştir.

Bilgi Formu

Geliştirilen bilgi formu ile araştırmanın ikinci aşamasında öğretmenlerden bilgi toplanmıştır. Geliştirilen bilgi formunda katılımcıların yaşları, cinsiyetleri ve özel eğitim alanındaki mesleki deneyimleri ile sınıflarında yer alan öğrencilerle olan deneyimlerine yönelik bilgiler yer almaktadır.

BEP Kalitesini Değerlendirme Ölçeği (BEPKDÖ)

Kalite ölçeği, BEP Öğretimsel Bileşenler Alt Ölçeği (BEPKDÖ-ÖB) ile BEP Tamamlayıcı Bileşenler Alt Ölçeği (BEPKDÖ-TB) şeklinde iki ayrı alt ölçekten oluşmaktadır. BEPKDÖ-ÖB'de; öğrenci performansı, kısa dönemli amaçlar (KDA), amaç ifadelerinde; ölçüt, koşul kullanımı, ölçülebilirlik - gözlenebilirlik, amaç-işlevsellik-ön koşulluk ilişkisi ve destek hizmetlerle ilgili maddeler yer almaktadır. BEPKDÖ-TB'de; uzun dönemli amaçlar (UDA), olumlu-olumsuz ifade kullanımı, genelleme, zaman, araç-gereçler, yöntem, değerlendirme ve davranış problemleri ile ilgili maddeler yer almaktadır. Her iki alt ölçekte 8'er madde olmak üzere toplamda 16 madde bulunmaktadır. Maddeler 0, 1 ve 2 şeklinde puanlanarak derecelendirilmiştir. BEPKDÖ'de bulunan her bir maddede sözü edilen puana karşılık gelen durum açıklanmış ve ölçekte yer verilmiştir. Değerlendirme sonucunda en az 0 puan; en fazla 32 puan alınabilir. Ölçekten alınan puan arttıkça BEP'in kalitesi artmaktadır.

Verilerin Analizi

Rubriklerin geçerliliğinin; kapsam geçerliliği, yapı geçerliliği ve ölçüt geçerliliği yönünden ve güvenilirliğinin ise değerlendiriciler/kodlayıcılar arası ve değerlendiriciler/kodlayıcılar içi güvenilirlik hesaplaması yoluyla incelenebileceği belirtilmektedir (Moskal ve Leydens, 2000). Araştırmanın birinci aşamasında kapsam geçerliliği için beş özel eğitim uzmanından uzman görüşü alınmıştır. Uzman görüşleri kapsam geçerliliği indeksiyle (Davis, 1992) analiz edilmiş, yapı geçerliliği için faktör analizi yapılmıştır. Güvenirlik için ise iç tutarlık katsayısı ve alt-üst 27 güvenilirliği analizleri yapılmıştır. Değerlendiriciler içi ve değerlendiriciler arası güvenilirlikte Kendall W uyum kat sayısı ve sınıf içi korelasyon katsayısı hesaplanmıştır. Birinci aşamada değerlendiriciler içi güvenilirlik için random olarak seçilmiş 60 BEP (%30) birinci yazar tarafından tekrar değerlendirilmiştir. Değerlendiriciler arası güvenilirlik için random seçilmiş 60 BEP (%30) doktora tez döneminde olan iki araştırma görevlisi tarafından puanlanmıştır.

İkinci aşamada 211 BEP birinci araştırmacı tarafından değerlendirilmiştir. Verilerin analizinde Tek yönlü varyans analizi (ANOVA), ilişkisiz ölçümler t testi, pearson korelasyon analizi kullanılmıştır. Kendall W uyum ve sınıf içi korelasyon katsayıları hesaplanmıştır. İkinci aşamada değerlendiriciler içi güvenilirlik için random olarak seçilmiş 70 BEP (%30) birinci yazar tarafından tekrar değeren-

dirilmiştir. Değerlendiriciler arası güvenilirlik için random seçilmiş 70 BEP (%30) doktora tez döneminde olan bir araştırma görevlisi tarafından puanlanmıştır.

Araştırmanın her iki aşamasında da ileri analizlere geçmeden önce normal dağılım verileri incelenerek verilerin normal dağıldığı belirlenmiştir.

Geçerlik, Güvenirlik ve Etik

Bu araştırma için bir devlet üniversitesinin sosyal ve beşeri bilimler yayın etiği kurulundan etik izin ve il milli eğitim müdürlüğünden uygulama izni alınmıştır.

Etik Kurul İzin Bilgileri

Yapılan bu çalışmada “Anadolu Üniversitesi Bilim Etiği Kılavuzu” kapsamında uyulması belirtilen tüm kurallara uyulmuştur.

Etik Değerlendirmeyi Yapan Kurul Adı: Anadolu Üniversitesi Sosyal ve Beşeri Bilimler Bilimsel Araştırma ve Yayın Etiği Kurulu

Etik Değerlendirme Kararının Tarihi: 20.02.2018

Etik Değerlendirme Belgesi Sayı Numarası: 19113

BULGULAR

Birinci Aşama

BEPKDÖ'nün Geçerliliği

Kapsam geçerliği analizi sonucunda 18 madde 1 puan, bir madde ise 0.80 puan almış ve 19 madde korunmuştur. Yapı geçerliliği için faktör analizine uygunluğun belirlenmesi amacıyla Kaiser-Meyer-Olkin (KMO) ve Bartlett testleri, maddeler arası korelasyon değerleri ve korelasyon matrisi determinanı ve anti-imağ korelasyonu incelenmiştir. KMO katsayısı .78 ve Bartlett küresellik testi anlamlılık değeri $p < .001$ bulundu ($\chi^2(120) = 1174.220$). Maddeler arası korelasyonda .90 üstü madde yoktur ve korelasyon matrisinin determinanı .001'dir. Anti-imağ korelasyonu incelemesinde .50 altında korelasyon değeri alan iki madde atılmış ve analiz tekrar gerçekleştirilmiştir. Analiz sonucunda elde edilen anti imaj korelasyonu değerleri .614 ve .846 arasındadır. Faktör analizi sonucuna göre; BEPKDÖ-ÖB alt ölçeği sekiz maddeden oluşmaktadır. Bu faktörde; özdeğer 4.09, açıklanan varyans % 25.56'dır. BEPKDÖ-TB alt ölçeği de sekiz maddelidir ve özdeğer 3.23, açıkladığı varyans % 20.19'dir. Ölçeğin toplam açıkladığı varyans % 45.76 olarak hesaplanmıştır. BEPKDÖ faktör yük değerleri ve açıklanan varyans oranları tablo 1'de verilmiştir.

Tablo 1. BEPKDÖ Faktör Analizinden Elde Edilen Değerler

Maddeler	Faktör Yük Değerleri			
	BEPKDÖ-ÖB	BEPKDÖ-TB	Özdeğer	Açıklanan Toplam Varyans %
M1	.67			
M3	.68			
M4	.81			
M5	.81			
M6	.54			25,569
M7	.82		4,091	
M16	.52			
M18	.47			
M2		.34		
M8		.61		
M11		.71		
M12		.35		
M13		.50		20,193
M14		.71	3,231	
M15		.83		
M19		.60		
Toplam			7,322	45,762

BEPKDÖ'nün Güvenirliği

BEPKDÖ-ÖB için iç tutarlılık katsayısı Cronbach Alfa değeri $\alpha=.84$, BEPKDÖ-TB için $\alpha=.64$, ölçeğin tamamı için $\alpha=.80$ olarak hesaplanmıştır. Alt üst %27 testi güvenilirliğiyle ilgili olarak yapılan t testinde alt grup ortalaması $\bar{X}=7.35$ ($sd=1,22$), üst grup ortalaması $\bar{X}=21.10$ ($sd=4,28$) olarak bulunmuştur ($p <.001$). Eta kare hesaplanarak bulunan etki büyüklüğü değeri ($\eta^2=.83$)'dir.

Değerlendiriciler arası güvenilirlik için iki yönlü karışık sınıf içi korelasyon katsayısı alfa değeri mutlak anlaşma ve tutarlık açısından hesaplanmış ve bu değer .99 olarak bulunmuştur ($p<.001$). Değerlendiriciler içi güvenilirlikte yapılan iki yönlü karışık analizde sınıf içi korelasyon katsayısı alfa değeri mutlak anlaşma için .99 ($p<.001$) olarak hesaplanmıştır. Kendall W için değerlendiricinin kendi içinde güvenilirliği ve diğer iki değerlendiriciyle birlikte üç değerlendiricinin uyum katsayısı Kendall W'de .99 olarak bulunmuştur ($p<.001$).

İkinci Aşama

BEP Kalite Düzeyi

BEP'lerin (n=211) kalite puanı ortalaması BEPKDÖ için $\bar{X}=9.66$ (ss=2,62), BEPKDÖ-ÖB için =4.98 (ss=2,10), BEPKDÖ-TB için $\bar{X}=4.67$ (ss=1,99) bulunmuştur. BEPKDÖ'den alınabilecek en yüksek puanın 32 en düşük puanın 0 olduğu göz önüne alındığında elde edilen kalite puanları oldukça düşüktür.

BEP'te Yer Alan Amaç İfadesi Sayıları ve BEP Kalitesi

BEP'lerde (n=211) yer alan amaç ifadesi sayıları incelendiğinde uzun dönemli amaç (UDA) sayısının 1 ve 227 arasında değiştiği, ortalamanın $\bar{X}=31$ olduğu (ss=25.42), kısa dönemli amaç (KDA) sayısının 1 ve 429 arasında değiştiği, ortalamasının $\bar{X}=145$, (ss=92.68) ve toplam amaç sayısının 2 ve 595 arasında değiştiği ortalamasının $\bar{X}=177$, (ss=113.11) olduğu görülmüştür. BEP'lerde bulunan UDA, KDA ve toplam amaç sayıları ve BEPKDÖ puanları arasındaki ilişkiye Pearson korelasyon analiziyle il bakılmıştır. Analiz sonucunda UDA sayıları, KDA sayıları, toplam amaç sayıları ve BEP kalitesi arasında ölçeğin geneli ve BEPKDÖ-TB için anlamlı bir ilişki olmadığı görülürken BEPKDÖ-ÖB için aralarında anlamlı ve negatif yönde ilişki olduğu görülmüştür. BEPKDÖ, BEPKDÖ-ÖB ve BEPKDÖ-TB ve toplam amaç sayıları arasındaki korelasyon değerleri tablo 2'de verilmiştir.

Tablo 2. BEPKDÖ, BEPKDÖ-ÖB ve BEPKDÖ-TB ve Toplam Amaç Sayıları Arasındaki Korelasyon Değerleri

Değişkenler	UDA	KDA	Toplam
BEPKDÖ	-.116	-.042	-.060
BEPKDÖ-ÖB	-.210**	-.138*	-.160*
BEPKDÖ-TB	.068	.090	.089

**p<.01 *p<.05

Tanı Grupları ve BEP Kalitesi

Tek yönlü varyans analizi (ANOVA) sonucunda BEPKDÖ toplam puan ortalaması ile farklı tanı grupları arasında istatistiksel olarak manidar bir farkın olmadığı ($F_{(2, 207)}=.680$, $p > .05$) ancak BEPKDÖ-ÖB ($F_{(2, 207)}=33.140$, $p < .01$) ve BEPKDÖ-TB puan ortalamaları ($F_{(2, 207)}=52.184$, $p < .01$) için anlamlı fark olduğu görülmektedir. HZY tanıli öğrencilerin BEPKDÖ-ÖB puan ortalamaları ($\bar{X}=6.15$), OZY ($\bar{X}=4.02$) ve AZY/OSB ($\bar{X}=4.15$) gruplarından anlamlı derece yüksek bulunmuş ($\eta^2=0.24$). HZY tanıli öğrencilerin BEPKDÖ-TB puan ortalamaları ($\bar{X}=3.40$), OZY ($\bar{X}=5.51$) ve AZY/OSB ($\bar{X}=5.838$) gruplarından; anlamlı derece düşük bulunmuştur ($\eta^2=0.34$).

Okul Kademeleri ve BEP Kalitesi

Okulöncesi döneme ilişkin olarak geliştirilen sadece 11 tane BEP olması sebebiyle okul öncesi döneme analizde yer verilmemiştir. ANOVA sonuçlarında BEPKDÖ ($F_{(2, 197)} = 93.896, p < .01$), BEPKDÖ-ÖB ($F_{(2, 197)} = 25.064, p < .01$) ve BEPKDÖ-TB ($F_{(2, 197)} = 22.738, p < .01$) puan ortalamaları ve okul kademeleri arasında anlamlı bir fark olduğu görülmüştür.

BEPKDÖ toplam puanı için ilkökul öğrencilerinin BEP kalite puanı ortalamalarının ($\bar{X} = 11,74$), ortaokul ($\bar{X} = 9,44$) ve lise ($\bar{X} = 7,75$) öğrencilerinin BEP kalite puanı ortalamalarından anlamlı şekilde yüksek olduğu tespit edilmiştir. Ortaokul öğrencilerinin BEP kalitesi puan ortalamalarının da ($\bar{X} = 9,44$) lise grubundan ($X = 7,75$) anlamlı derecede yüksek olduğu görülmüştür ($\eta^2 = 0.49$).

BEPKDÖ-ÖB için ilkökul öğrencilerinin BEP kalite puanı ortalamalarının (\bar{X} ilkökul=6,33), ortaokul ve lise grubundan (\bar{X} ortaokul=4,59; \bar{X} lise=4,18) anlamlı derecede yüksek olduğu görülmüştür ($\eta^2 = 0.20$). BEPKDÖ-TB için ilkökul grubunun BEP kalite puanı ortalamalarının (\bar{X} ilkökul=5,41) ortaokuldan anlamlı derece yüksek olduğu (\bar{X} ortaokul=4,85), ortaokul grubunun kalite puanı ortalamalarının da (\bar{X} ortaokul=4,85) liseden (\bar{X} lise=3,56) anlamlı derecede yüksek olduğu görülmüştür ($\eta^2 = 0.19$). Okul kademesi arttıkça BEP kalitesinin düştüğü görülmektedir.

Öğretmenlerin Mesleki Deneyim Süreleri ve BEP kalitesi

Öğretmenlerin özel eğitim alanındaki mesleki deneyimi yıl bağlamında dört grupta incelenmiştir. Birinci grupta özel eğitim alanında 1-5 yıl, ikinci grupta 6-10 yıl, üçüncü grupta 11- 15 yıl ve dördüncü grupta 16 yıl ve üzerinde deneyimi olan öğretmenler yer almaktadır. Tek yönlü varyans analizi sonucunda BEPKDÖ ($F_{(2, 207)} = 8.945, p < .01$), BEPKDÖ-ÖB ($F_{(3, 207)} = 8.826$ ve $p < .01$) ve BEPKDÖ-TB ($F_{(3, 207)} = 15,224, p < .01$) puan ortalamaları ve öğretmenlerin özel eğitim deneyimleri arasında anlamlı fark olduğu görülmüştür. Birinci grupta yer alan öğretmenlerin hazırladığı BEP'lerin BEPKDÖ toplam puan ortalamaları (\bar{X} Birinci =11,04) ikinci (\bar{X} İkinci=9,56), üçüncü (\bar{X} Üçüncü=9,49) ve dördüncü gruptan (\bar{X} Dördüncü=8,25) anlamlı derecede yüksek bulunmuştur ($\eta^2 = 0.12$). BEPKDÖ-ÖB için birinci grubun hazırladığı BEP kalite puanlarının ortalaması (\bar{X} Birinci=8,26) dördüncü gruptan (\bar{X} Dördüncü=3,91) anlamlı derecede yüksek bulunurken ($\eta^2 = 0.11$) BEPKDÖ-TB için; birinci grubun hazırladığı BEP kalite puanlarının ortalaması (\bar{X} Birinci=5,86) üçüncü gruptan (\bar{X} Üçüncü= 3,73) ve dördüncü gruptan (\bar{X} Dördüncü= 4,34) anlamlı derecede yüksek bulunmuştur ($\eta^2 = 0.18$).

TARTIŞMA, SONUÇ VE ÖNERİLER

Bu araştırmanın birinci aşamasında BEPKDÖ'nün geçerlik güvenirlik analizleri yapılmıştır. Yapılan analizler sonucunda 16 maddeden oluşan iki boyutlu bir ölçek ortaya çıkmıştır. BEPKDÖ geçerlik bulguları yapı geçerliği ile kapsam geçerliği olmak üzere iki eksenle ele alınmıştır. BEP kalitesini ölçen geçerliliği ve güvenirliği bulunan başka bir Türkçe araca ulaşamadığı için ölçüt bağımlı geçerliliğe bakılamamıştır. Kapsam geçerliliğinde, tüm maddeler .80 üstü değer olarak kapsam geçerliliği sağlanmıştır (Davis, 1992). Faktör analizi için uygunluğun değerlendirilmesi amacıyla Bartlett küresellik testi anlamlılık değeri .000 ($p < .05$) testin sonuçlarının anlamlı olduğunu ve faktör analizi için uygun olduğunu göstermiştir (Field, 2009; Pallant, 2020; Can, 2019). KMO katsayısının .78 olması faktör analizi için uygunluğu sağlamıştır (Can, 2019; Tabachnick ve Fidell, 2014). Çalışmada .80 veya .90 üzerinde korelasyon değeri alan maddelerin olmaması faktör analizi yapabilmek için çoklu bağlantı sorunu olmadığını göstermiştir (Can, 2019; Field, 2009; Tabachnick ve Fidel, 2014). Anti-ımağ korelasyonu incelemesinde iki madde atıldıktan sonra yapılan analizle ulaşılan değerlerin .653 ve .845 arasında yer alması ($>.50$) maddelerin faktör analizi açısından uygun olduğunu göstermiştir (Can, 2019; Field, 2009). Faktör analizi sonucunda BEPKDÖ'nün iki faktörünün de öz değerinin 1'in üstünde olması (Büyüköztürk, 2014; Field, 2009; Huck, 2012; Kline, 1994; Tabachnick ve Fidell, 2014; Pallant, 2020) ve faktörlere ait öz değerlerin toplam için hesaplanan öz değerinin %5'inden büyük olması (Huck, 2012) faktör oluşumunun uygun olduğunu göstermiştir. Açıklanan varyans değerinin tek faktörlü yapılar için %30'un üstünde olması, çok faktörlü yapıdaysa daha fazla olması beklenmektedir (Büyüköztürk, 2014). Buna göre BEPKDÖ'nün iki faktörlü yapıda %45'lik varyansı yeterli düzeydedir. BEPKDÖ maddelerinin faktör yük değerleri .30'un üstünde olması (Büyüköztürk, 2014; Özdamar, 2017) ve faktör yükü olarak başka faktörlerde de anlamlı faktör yük değeri veren maddelerde faktör yüklerinde .10 seviyesinde bir fark olması (Büyüköztürk, 2014) maddelerin faktörlerle ilişkili olduğunu göstermiş ve faktör analizine uygunluğu desteklemiştir.

Güvenirlikle ilgili olarak BEPKDÖ'nün geneli için Cronbach alfa katsayısı $\alpha=.80$ ve BEPKDÖ-ÖB için $\alpha=.84$ değerleri iyi ya da yüksek derecede güvenilir olduğunu göstermiştir (Can, 2019; Cresswell, 2012; George ve Mallery, 2003; Ho, 2006; Özdamar, 2017). BEPKDÖ-TB için $\alpha=.64$ de kabul edilebilir düzeyde bulunmuştur (Can, 2019; George ve Mallery, 2003; Nunnally ve Bernstein, 1994; Özdamar, 2017). Alt üst %27 testi sonuçlarında her iki grup arasındaki farkın anlamlı olduğunu ($p < .01$), ölçülen niteliğe sahip olan ve olmayan BEP'leri iyi bir şekilde ayırt edebildiğini (Can, 2019) ve testin iç tutarlılığının bir göstergesi olduğunu göstermiş ve farkın etki büyüklüğü geniş etki büyüklüğünü göstermiştir (Büyüköztürk, 2014; Pallant, 2020). Sınıf içi korelasyon katsayısı için değerlendiriciler arası ve değerlendiriciler içi güvenirlik için .99 bulunan değerler çok iyi güvenirliğin sağlandığını (Koo ve Li, 2016) ve rubriğin güvenilir olduğunu göstermektedir (Kutlu, vd.,

2014). Kendall W uyumu katsayısı değerlendiriciler arası ve değerlendiriciler içi güvenilirlik için $p = .000$ olup, $p < .001$ düzeyinde anlamlıdır (Can, 2019) ve .99 olan Kendall's W değerleri güvenilirlik için olumlu yönde kanıt oluşturmaktadır (Can, 2019; Field, 2009; Legendre, 2005).

Bu ölçme aracında kapsam geçerliğinde gerekli niteliklerin sağlandığı, alan yazında bu alandaki ölçütlere göre yapı geçerliği için analize uygunluk ve uygun faktörleşmenin sağlandığı, bir derecelmeli puanlama anahtarı olarak BEPKDÖ'nün kapsam ve yapı geçerliliği psikometrik kanıtlarının ortaya konulduğu söylenebilir. BEPKDÖ bu alanda kalite göstergelerini ölçen araçlarda standardizasyonu bilimsel kanıtlarla sağlanmış bir araç olma özelliğini göstermektedir. BEPK kalitesini değerlendirmek için araştırmalarda kullanılan araçlar incelendiğinde: Çalışmalarda kullanılan ölçme araçlarından bir kısmının geçerlik bilgilerinin raporlanmadığı (Le Salle, vd., 2010; Lucas, 2014; Lowman, 2016; McWilliam, 2021; McWilliam, 2021; Rowland, 2015), geçerlik bilgileri raporlanan araştırmalarda da ölçme araçlarının yapı geçerliği analizlerinin sunulmadığı, bir kısmında sadece uyum geçerliğinin rapor edildiği (Boavida vd., 2010; Lynch ve Beare, 1990; Rakap, 2015) ya da yasalar ve alan yazından yola çıkılmış olması nedeniyle kapsam geçerliğinin sağladığı var sayıldığı (Ruble vd., 2010; Smith ve Simpson, 1989) göz önünde bulundurulduğunda BEPKDÖ'nün geçerliliği BEPK kalitesini incelemek adına geliştirilen hali hazırda var olan araçlardan daha güçlü bir şekilde ortaya konulmuştur. BEPKDÖ'nün güvenilirlikle ilgili analizlerde gerekli nitelikleri karşıladığı ve güvenilir bir araç olduğu görülmüştür. Güvenirlikle ilgili olarak iç tutarlılığın yanı sıra alt üst %27 testi güvenilirliği nitelikleri karşılayan ilk araştırma olması dolayısıyla aracın güvenilirliği güçlü olarak ortaya konulmuştur.

BEPKDÖ'den ve alt ölçeklerden elde edilen genel kalite puanları ortalamaları oldukça düşüktür. Bu sonuç geliştirilen BEP'lerin kalitelerinin düşük olduğunu göstermiştir ve alanyazınla tutarlıdır (Smith ve Simpson, 1989; Lynch ve Beare, 1990; Giangreco vd., 1994; Prettie-Frontzack ve Bricker, 2000; Boavida vd. 2010; Ruble vd., 2010; Sanches-Ferreira vd., 2013; Rakap, 2015). BEPKDÖ-ÖB'den elde edilen düşük puan ortalamaları alanyazında yer alan BEPK kalitesine ilişkin araştırma bulgularında; performans ifadelerinin yer almaması (Boavida vd., 2010; Le Salle vd., 2010; Smith ve Simpson, 1989), amaçlar ve performans arasında uyumsuzluğun bulunması (Catone ve Brady, 2005; Smith ve Simpson, 1989), işlevsellik bağlamında belirlenmiş olan amaçların yeterli olmaması (Boavida vd., 2010), BEPK'te ilişkili hizmetlerin tespit edilmesinde ve bahse konu hizmetlerin başlangıç ve bitiş tarihlerinin belirtilmesinde eksiklikler olması (Smith ve Simpson, 1989), amaç ifadelerinin anlamsız şekilde ve oldukça genel bir tarzda yazılması (Catone ve Brady, 2005; Giangreco vd., 1994; Lynch ve Beare, 1990), başarı ölçütünde eksikler oluşu (Boavida vd., 2010; Giangreco vd., 1994; Lynch ve Beare, 1990; Sanches-Ferreira vd., 2013), amaçların yazılmasında koşul ögesine yer vermede eksiklikler olması (Boavida vd., 2010; Giangreco vd., 1994; Lynch ve Beare, 1990; Rakap, 2015; Ruble

vd., 2010), hedef davranış belirtilirken ölçülebilir-gözlenebilir ifadeler kullanılmaması (Boavida vd., 2010; Ruble vd., 2010; Sanches-Ferreira, vd., 2013) sonuçlarıyla paralel ve tutarlı bulgular içerisinde olduğunu görülmektedir. BEPKDÖ-TB'den elde edilen düşük puan ortalamaları da alanyazında yer alan çalışmalarda belirlenen öğrenme stratejilerinin yeterli düzeyde ve belirgin olarak ortaya konulmaması (Lynch ve Beare, 1990; Ruble vd., 2010), zaman ifadelerinde eksiklikler olması (Boavida vd., 2010; Rakap, 2015; Ruble vd., 2010;), genellemeye dayalı amaçlarda eksiklikler (Boavida, vd., 2010; Rakap, 2015), değerlendirme sürecine ilişkin yeter-sizlikler (Smith ve Simpson, 1989) gibi sonuçlarla tutarlılık göstermektedir.

UDA, KDA ve toplam amaç sayıları ile BEPKDÖ ve BEPKDÖ-TB arasında anlamlı bir ilişki bulunamamıştır ancak BEPKDÖ-ÖB ile negatif yönde anlamlı ilişki görülmüştür. UDA, KDA ve toplam amaç sayıları arttıkça BEPKDÖ-ÖB puan ortalamasının düştüğü görülmektedir. Bu durum amaç yazım kalitesinin düştüğü anlamına gelmektedir. Ayrıca BEP'lerde belirlenen amaç sayılarının oldukça yüksek olması alanyazınla tutarlıdır (Boavida vd., Giangreco vd., 1994; 2010; Sanches-Ferreira vd., 2013; Rakap, 2015). Boavida vd. (2010) çalışmalarında UDA sayısının 4-95 arasında değiştiğini ($\bar{X}=38$); Rakap (2015) UDA/KDA sayısının 8-38 arasında değiştiğini ($\bar{X}=22$) Sanches-Ferreiravd., (2013) BEP'lerin %80'inde 20'den fazla UDA'nın yer aldığını belirtmiştir. Alanyazında amaç sayısının fazla olma sebebi; öğretmenlerin öncelikli ve önemli olan davranışları belirlemede güçlükler yaşadıklarını, dolayısıyla da öğrencinin kazanımını artırmak yerine sürece ilişkin gereklilikleri gerçekleştirmeye odaklanmaları olarak açıklanmaktadır (Shinn ve Shinn, 2000). Boavida vd. (2014) BEP amaçlarının kalitesinin artırılmasına yönelik olarak yaptıkları çalışmada, verilen eğitim sonrasında hazırlanan BEP'lerde belirlenen amaç sayılarında azalma olduğu ve geliştirilen BEP'lerin kalitelerinin arttığı sonucu bu durumu desteklemektedir. Ayrıca belirlenen amaç sayılarının gereğinden çok yüksek olmasının, öğretmenler açısından pek çok yeni gereklilik doğurduğu ve bu nedenle öğretmenlerin söz konusu amaçlara ilişkin uygulamaları izleme noktasında başarısız olma olasılığına neden olabileceği belirtilmektedir (Fuchs ve Shinn, 1989). Bununla birlikte belirlenen amaç sayılarının çok yüksek olması, BEP'in öğretmenler tarafından kullanılmadığı şüphesini akla getirdiği ifade edilmektedir (Giangreco vd., 1994). BEP için belirlenen amaçların çok olması, hem öğretmenlerin amaçları belirlemede zorluklar yaşadıklarını hem de BEP ekibinin BEP'i işlevsel bir araçtan ziyade yasal bir zaruret olarak gördüklerini düşündürmektedir.

BEPKDÖ'den elde edilen toplam puanlar incelendiğinde, tanı grupları bakımından BEP kalitesinin farklılaşmadığı görülmüştür. Bu bulgu Rakap'ın (2015) araştırmasında ulaştığı sonuçlarla tutarlılık göstermekte ancak Boavida vd.'nin (2010) çalışma bulguları ile örtüşmemektedir. Bununla birlikte bulgular alt ölçekler bazında incelendiğinde tanı grupları ve BEP kalitesi arasında anlamlı farklılıklar olduğu görülmüştür. Birinci alt ölçek için HZY grubunun BEP kalite puanlarının ortalaması OZY ve AZY-OSB gruplarından yüksekken, ikinci alt ölçekte

bu durumun tam tersi olarak HZY grubunun BEP kalite puanları ortalamasının, OZY ve AZY-OSB gruplarından daha düşüktür. HZY grubunun birinci alt ölçek kalite puanı ortalamalarının daha yüksek olma nedeni, öğretmenlerin bu çocuklar için nitelikli amaçlar oluşturmada kendilerini söz konusu özel gereksinim grubuna öğretim sunmada diğer yetersizlik gruplarına kıyasla daha yeterli hissetmeleri olabilir. Öğretmenler amaçlar hususunda daha az zorlanmaları ya da daha yeterli hissetmeleri sonucunda ikinci alt ölçekteki BEP bileşenlerine yeteri kadar ağırlık vermemiş olabilirler. Boavida vd., (2010), yetersizlikten yoğun düzeyde etkilenmiş olan öğrenciler için bir başka deyişle bağımsız yaşama ilişkin amaçlara daha yoğun düzeyde gereksinimi olan öğrenciler için öğretmenlerinin bağımsız yaşamı destekleyen işlevsel amaçlar belirleme konusunda daha çok gayret etmiş olabileceklerini belirtmiştir. Bu çalışmada öğretmenler işlevsel amaçlarla ilgili eksikliklerini BEPKDÖ-TB'de yer alan bileşenlerle tamamlayamaya çalışmış olabilir. BEPKDÖ-ÖB ve BEPKDÖ-TB etki büyüklüğü değerleri yüksek etki büyüklüğüdür (Pallant, 2020; Büyüköztürk, 2014). Alt ölçeklerde görülen bu zıt durum ölçeğin toplam puanında bir fark oluşmasını engellemiş olabilir.

Farklı okul kademelerinde BEP kalite puanlarının ortalamalarına bakıldığında, okul düzeyi arttığında BEP kalite puanlarının düştüğü ve etki büyüklüğünün büyük olduğu görülmektedir.. Bu sonucun alanyazındaki Le Salle vd.'nin (2010) bulgusuyla çeliştiği, bunun yanında başka araştırma bulgularıyla örtüştüğü görülmüştür (Catone ve Brady, 2005; Sanches-Ferriara, 2013). Ortaya çıkan bu durum öğrencinin okul düzeyinin artması ile birlikte akademik alanların karmaşıklığının ve çeşitliliğin de artmasıyla ilişkili olabilir (Sanches-Ferreira vd., 2013).

Öğretmenlerin özel eğitime ilişkin mesleki deneyimleri bağlamında BEP kalite puan ortalamaları incelendiğinde mesleki deneyimi 1-5 yıl aralığındaki öğretmenlerde, BEP kalite puan ortalamalarının diğer gruplara kıyasla yüksek ve manidar bir farklılık olduğu, bu farklılığa ilişkin etki büyüklüğü değerinin büyüğe yakın olduğu görülmüştür. Bu çalışmanın bulguları Ruble vd. (2010)'nin öğretmenlerin OSB'li çocuklarla mesleki deneyimleri ile BEP kalitesi arasında bir ilişki olmadığı bulgusuyla çelişmektedir fakat okuduğunuz bu çalışmada farklı tanı grupları ya da belirli bir tanı grubuna özgü deneyim değil, genel olarak özel eğitime ilişkin mesleki deneyim ele alınmıştır. Özel eğitim deneyimi 1-5 yıl olan öğretmenlerin, en güncel bilgilere sahip olması, bu gruptaki öğretmenlerin BEP kalitesi puanlarının ortalamalarında diğer gruplara kıyasla daha yüksek puanların alınmış olmasına yol açabileceği düşünülmektedir.

BEP; özel gereksinimi olan bütün öğrenciler için geliştirilmesi zorunlu ve yasalar ile sabit olan eğitsel bir yol haritası, bir rota niteliği taşımasının yanı sıra öğrenciye sağlanacak hizmetleri ve bu hizmetleri ile ilgili bilgileri kapsayan ve özel gereksinimli öğrenciyi izlemeye de olanak veren kapsamlı bir hizmet belgesidir. Bununla birlikte bu çalışma kapsamında ele alınan BEP'lerin sıralanan tüm bu

nitelikleri büyük ölçüde karşılamadığı da oldukça açık bir şekilde görülmektedir. BEP'in özel gereksinimli öğrenci için öğretimsel programı belirlemede oynaması beklenen aracı rol alanyazında sorgulanmaktadır (Berninger, 1997). Smith ve Simpson (1989) BEP'lerin gerekli bileşenleri içermeye yetersiz kalması durumunda, eğitsel anlamda rehberlik işlevine de gölge düştüğünü belirtmektedir. Benzer biçimde bu çalışmanın bulguları da etkili eğitimin tasarlanması sürecinde BEP'in rehberlik rolünü sağlamadaki yeterliliğini sorgulatabilir. BEP'in; özel gereksinimli öğrencinin ilişkili olduğu paydaşlar için (aileler, yardımcı öğretmenler; öğretmenler, okul personeli vb.) öğretim yöntem ve teknikleri ile bireysel gereksinimlere uygun biçimde planlanmış eğitim hizmetlerinin açık bir şekilde belirtildiği bir düzenek işlevine sahip olduğu (Ruble vd., 2010) ve gerçekleştirilen planlamalara ait güvence sağlayan bir nitelik taşıdığı belirtilmektedir (Fiscus ve Mandell, 2002). Ancak BEP'in kendisinde önemli eksikliklerin olması durumunda sağladığı güvence; çocuğun ailesi, gelecekteki öğretmenleri ve hizmet sağlayanlar için yetersiz olabilir ya da yanlış bilgilendirmelere yol açabilir. Alanyazında öğretmenlerin BEP hazırlamada güçlük yaşadıkları ve BEP'in içeriğine yönelik olarak yeterli düzeyde bilgi sahibi olmadıkları belirtilmektedir (Çıkılı vd., 2020; Öztürk ve Eratay, 2010). BEP kalite puanlarındaki genel düşüklüğün ve amaç sayılarında aşırı yüksekliğin sebebi öğretmenlerin konu hakkında yeterince bilgi sahibi olmamaları olabilir. BEP'lerin genel olarak kalitelerinin yükseltilmesi ve fazla sayıdaki amacın eğitim ihtiyaçları, zaman gerekleri ve öğrencinin var olan performans düzeyi ile ilişkili daha ulaşılabilir amaç sayılarına çekilebilmesi; BEP'lerin işlevsel olması bir başka deyişle eğitsel bir rehberlik rolü üstlenebilmesi özel gereksinimli öğrenci ile ilişkili tüm paydaşlar tarafından izlenebilmesi bakımından önem arz etmektedir. Bunu sağlamak için öğretmenlerin BEP'e dair bilgileri değerlendirilmeli ve gerekli görüldüğü durumlarda bilgileri yeterli düzeye getirme ve desteklemeye yönelik eğitim çalışmaları gerçekleştirilmelidir. Gelecekteki çalışmalarda bütünleştirici ortamlarda yer alan öğrencilerin BEP'lerinin kaliteleri incelenebilir. Ayrıca BEP'in kalitesi ve sınıflarda yürütülen uygulamalar arasındaki ilişkiler ve bunların öğrencilerin eğitsel düzeylerine etkisi incelenebilir.

Teşekkür ve Açıklamalar

Çalışmamıza katkı sağlayan Sayın Çetin TOPUZ ve Sayın Şerife ŞAHİN'e teşekkür ederiz.

Çıkar Çatışması

Makalenin yazarları arasında, çalışma kapsamında herhangi bir kişisel ve finansal çıkar çatışması bulunmamaktadır.

Yazar Katkı Oranları

Çalışmanın Tasarlanması (Design of Study): ŞA(%50), VA(%50)

Veri Toplanması (Data Acquisition): ŞA(%50), VA(%50)

Veri Analizi (Data Analysis): ŞA(%50), VA(%50)

Makalenin Yazımı (Writing Up): ŞA(%50), VA(%50)

Makalenin Gönderimi ve Revizyonu (Submission and Revision): ŞA(%50), VA(%50)

KAYNAKLAR

- Berninger, V. W. (1997). Introduction to intervention for Students with learning and behavior problems: Myths and realities. *School Psychology Review*, 26(3), 326-332 <https://doi.org/10.1080/02796015.1997.12085870>
- Boavida, T., Aguiar, C., McWilliam, R. A., & Pimentel, J. S. (2010). Quality of individualized education program goals of preschoolers with disabilities. *Infants and Young Children*, 23(3), 233-243. <https://doi.org/10.1097/IYC.0b013e3181e45925>
- Boavida, T., Aguiar, C., & McWilliam, R. A. (2014). A training program to improve IFSP/BEP goals and objectives through the routines-based interview. *Topic in Early Childhood Special Education*, 33(4), 200-211. <https://doi.org/10.1177/0271121413494416>
- Büyükköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). Bilimsel araştırma yöntemleri. (18. Baskı). Pegem Akademi.
- Can, A. (2019). SPSS ile bilimsel araştırma sürecinde nicel veri analizi. Pegem Akademi.
- Catone, W. V., & Brady, S. A. (2005). The inadequacy of individual educational program (IEP) goals for high school students with word-level reading difficulties. *Annals of dyslexia*, 55(1), 53-78. <https://doi.org/10.1007/s11881-005-0004-9>
- Christle, C. A., & Yell, M. L. (2010). Individualized education programs: Legal requirements and research findings. *Exceptionality*, 18(3), 109-123. <https://doi.org/10.1080/09362835.2010.491740>
- Creswell, J. W. (2012). Educational research planning, conducting and evaluating quantitative and qualitative research (4th ed.). Pearson Education, Inc
- Çikılı, Y., Gönen, A., Bağcı, Ö. A., & Kaynar, H. Özel eğitim alanında görev yapan öğretmenlerin Bireyselleştirilmiş Eğitim Programı (BEP) hazırlama konusunda yaşadıkları güçlükler. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 15(1), 5121-5148.
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, 5(4), 194-197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- DeVellis, R. F. (2017). Scale development: Theory and applications. (3th ed.). SAGE Publications Inc.
- Field, A. (2009). Discovering statistic using SPSS (3rd Edition). Sage Publications.
- Fiscus, E. D., & Mandell, C. J. (2002). Bireyselleştirilmiş eğitim programlarının geliştirilmesi BEP. (2. Baskı). (G. Akçamete, H. G. Şenel ve E. Tekin, çev.). Anı Yayıncılık.
- Fuchs, L. S. & Shinn, M. R. (1989). Writing CBM IEP objectives. In M. R. Shinn (Ed.), (pp.130- 152). Guilford Press.
- Gay, L. R., Mills, G. E. & Airasian, P. W. (2012) Educational research: Competencies for analysis and applications (Tenth Edition). USA: Pearson Education, Inc.
- George, D., & Mallery, P. (2003). SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.). Allyn and Bacon.
- Giangreco, M. F., Dennis, R. E., Edelman, S. W., & Cloninger, C. J. (1994). Dressing your IEPs for the general education climate analysis of IEP goals and objectives for students with multiple disabilities. *Journal for Special Educators*, 15(5), 288-296. <https://doi.org/10.1177/074193259401500504>
- Gibb, G. S. & Dyches T. T. (2016). IEPs: writing quality individualized education programs (3rd. ed.). Pearson Publications.
- Goodrich-Andrade, H. (2000). Using rubric to promote thinking and learning. *Educational leadership*, 57(5), 13-19.
- Haladyna, T. M. (1997). Writing Test Item to Evaluate Higher Order Thinking. Allyn and Bacon.
- Heward, W. L., Alper-Morgan, S.R., & Konrad M. (2017). Exceptional children: An introduction to special education (11th ed.). Pearson Publications.

- Ho, R. (2006). Handbook of univariate and multivariate data analysis and interpretation with SPSS. Chapman and Hall/CRC.
- Huck, S. W. (2012). Reading statistic and research. (6th ed.). Pearson Educations, Inc.
- Hunt, P., Goetz, L., & Anderson, J. (1986). The quality of IEP objectives associated with placement on integrated versus segregated school sites. *Research and Practice for Persons with Severe Disabilities*, 11(2), 125-130. <https://doi.org/10.1177/154079698601100206>
- Kline, P. (1994). An easy guide to factor analysis. Routledge Taylor and Francis Group
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kutlu, Ö., Doğan C. D. & Karakaya, İ. (2014). Ölçme ve değerlendirme. Pegem Akademi.
- Le Salle, T. P., Roach, A. T., & McGrath, D. (2010). The relationship of IEP quality to curricular access and academic achievement for students with disabilities. *International Journal of Special Education*, 28(1), 135-144.
- Legendre, P. (2005). Species associations: the Kendall coefficient of concordance revisited. *Journal of agricultural, biological, and environmental statistic*, 10(2), 226-245 <https://doi.org/10.1198/108571105X46642>
- Lowman, J. J. (2016). A comparison of three professional development mechanisms for improving the quality of standards-based IEP objectives. *Communication Disorders Quarterly*, 37(4), 211-224. <https://doi.org/10.1177/1525740115595338>
- Lucas, A., Gillaspay, K., Peters, M. L., & Hurth, J. (2014). Enhancing recognition of high quality, functional IEP goals: A training activity for early childhood special education staff. <http://www.ectacenter.org/~pdfs/pubs/rating-IEP.pdf>
- Lynch, E. C., & Beare, P. L. (1990). The quality of IEP objectives and their relevance to instruction for students with mental retardation and behavioral disorders. *Remedial and Special Education*, 11(2), 48-55. <https://doi.org/10.1177/074193259001100207>
- McWilliam, R. A. (01.04.2021). Goal Functionality Scale. https://www.utoledo.edu/education/grants/direct/pdf/Goal_Functionality_Scale.pdf
- McWilliam, R. A. (01.04.2021). Goal Functionality Scale III. https://ectacenter.org/~pdfs/topic/families/GoalFunctionalityScaleIII_2_.pdf
- Mertler, C. A. (2001). Designing scoring rubric for your classroom. *Practical Assessment, Research and Evaluation*, 7(25), 1-10
- Moskal, B. M. & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research, and Evaluation*, 7(10), 71-81.
- Nunnally, J.C. & Bernstein, I.R. (1994). Psychometric theory (3rd. Ed.). McGraw-Hill
- Özdamar, K. (2017). Ölçek ve test geliştirme yapısal eşitlik modellemesi. Nisan Kitapevi.
- Öztürk, C. Ç. & Eratay, E. (2010). Eğitim uygulama okuluna devam eden zihnen engelli öğrencilerin öğretmenlerinin bireyselleştirilmiş eğitim programı hakkında görüşlerinin belirlenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 10(2), 145-159.
- Pallant, J. (2020). SPSS Survival Manual. A step by step guide to data analysis using SPSS (7th ed.). Routledge Taylor and Francis Group.
- Pretti-Frontczak, K., & Bricker, D. (2000). Enhancing the quality of individualized education plan (IEP) goals and objectives. *Journal of Early Intervention*, 23(2), 92-105. <https://doi.org/10.1177/105381510002300204>
- Rakap, S. (2015). Quality of individualised education programme goals and objectives for preschool children with disabilities. *European Journal of Special Needs Education*, 30(2), 173-186. <https://doi.org/10.1080/088562572014.986909>
- Roach, A. T., Chilungu, N., La Salle, T. P., Talapatra, D., & Vignieri, M., (2009). Opportunities and options for facilitating and evaluating access to the general curriculum for students with disabilities. *Peabody Journal of Education*, 84(4), 511-528. <https://doi.org/10.1080/01619560903240954>
- Rowland, C. M., Quinn, E. D., & Steiner, S. A. (2015). Beyond legal: crafting high-quality IEPs for children with complex communication needs. *Communication Disorders Quarterly*, 37(1), 53-62. <https://doi.org/10.1177/1525740114551632>
- Ruble, L. A., McGrew, J., Dalrymple, N., & Jung, L. A. (2010). Examining the quality of IEPs for young children with autism. *Journal Of Autism and Developmental Disorders*, 40(12), 1459-1470. <https://doi.org/10.1007/s10803-010-1003-1>
- Sanches-Ferreira, M., Lopes-dos-Santos, P., Alves, S., Santos, M., & Silveira-Maia, M. (2013). How individualised are the Individualised Education Programmes (IEPs): an analysis of the contents and quality of the IEPs goals. *European Journal of Special Needs Education*, 28(4), 507-520. <https://doi.org/10.1080/088562572013.83043>
- Shinn, M. R., & Shinn, M. M. (2000). Writing and evaluating IEP goals and making appropriate revisions to ensure progress and participation in general curriculum. In C. Telzrow ve M. Tankersley (Eds.), IDEA Amendments of 1997: Practice guidelines for school-based teams (pp. 351-382). National Association of School Psychologists.
- Smith, S. W., & Simpson, R. L. (1989). An Analysis of Individualized Education Programs (IEPs) for Students with Behavioral Disorders. *Behavioral Disorders*, 14(2), 107-116. <https://doi.org/10.1177/019874298901400206>

QUALITY OF INDIVIDUALIZED EDUCATION PROGRAMS (IEP) DEVELOPED IN SPECIAL EDUCATION SCHOOLS

ABSTRACT

Individualised education programs one of the basic principles of special education. It is stated that individualised education programs quality should be monitored to ensure that the goals and services determined in individualised education programs are implemented and monitored as planned. Studies examining individualised education programs concluded that quality levels were generally low. However, it was also observed that there were limitations in the psychometric properties of the tools used in these studies. This study's purpose is to develop a valid and reliable tool to evaluate individualised education programs quality and examine the qualities of individualised education programs using the created tool. At the first stage of the study, the Individualised Education Program Quality Assessment Scale was developed, and its validity and reliability were tested using 180 individualised education programs. At the second stage, the quality of 211 individualised education programs was examined using the scale. Results indicated that the Individualised Education Program Quality Assessment Scale was a valid and reliable scale, the quality scores of individualised education programs were low, the number of goals was very high, the quality scores decreased as the school stage increased, and they differed according to the diagnosis and professional experience of teachers. The results suggest that teachers experience difficulty preparing individualised education programs, individualised education programs are not used as educational guides, and individualised education programs are prepared as documents only due to legal necessity.

Keywords: Special Education, Individualised Education Program, Iep Quality, Iep Quality Scale.



ÖZEL EĞİTİM OKULLARINDA GELİŞTİRİLEN BİREYSELLEŞTİRİLMİŞ EĞİTİM PROGRAMLARININ (BEP) KALİTESİNİN İNCELENMESİ

ÖZ

Bireyselleştirilmiş eğitim programı özel eğitimin temel ilkelerinden biridir. Bireyselleştirilmiş eğitim programında amaçların ve hizmetlerin belirlendiği biçimde uygulandığından ve izlendiğinden emin olmak için bireyselleştirilmiş eğitim programı kalitesinin belirlenmesi ve takip edilmesi gerektiği ifade edilmektedir. Bireyselleştirilmiş eğitim programlarının kalitesinin genel olarak düşük olduğu ifade edilmektedir. Ayrıca kalite incelemesi amacıyla kullanılan araçların psikometrik niteliklerinde sınırlılıklar söz konusudur. Bu araştırmanın amacı; bireyselleştirilmiş eğitim programı kalitesini değerlendirmeyi hedefleyen geçerli ve güvenilir bir ölçme aracı geliştirmek ve geliştirilen araçla özel eğitim okullarında geliştirilen bireyselleştirilmiş eğitim programının kalitelerini incelemektir. Araştırmanın birinci aşamasında Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeği geliştirilmiş ve 180 bireyselleştirilmiş eğitim programı kullanılarak geçerliliği ve güvenilirliği sınanmıştır. İkinci aşamasında Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeği kullanılarak, 211 bireyselleştirilmiş eğitim programının kalitesi incelenmiştir. Araştırma sonucunda Bireyselleştirilmiş Eğitim Programı Kalitesini Değerlendirme Ölçeğinin geçerli ve güvenilir bir ölçek olduğu, bireyselleştirilmiş eğitim programlarının kalite puanlarının düşük olduğu, amaç sayılarının çok yüksek olduğu, okul kademesi yükseldikçe kalite puanlarının düştüğü, tanıya ve öğretmen mesleki deneyimine göre kalite puanının farklılaştığı görülmüştür. Elde edilen sonuçlar; öğretmenlerin bireyselleştirilmiş eğitim programı hazırlamakta güçlük çektiğini, eğitsel bir rehber olarak bireyselleştirilmiş eğitim programlarının kullanılmadığını ve bireyselleştirilmiş eğitim programlarının yalnızca yasal zorunluluk sebebiyle bir belge olarak hazırlandığını düşündürmektedir.

Anahtar Kelimeler: Özel Eğitim, Bireyselleştirilmiş Eğitim Programı, Bep Kalitesi, Bep Kalite Ölçeği.



INTRODUCTION

Individualised education programs cover the current performance of students, measurable annual goals, certain special education and related services that will enable them to achieve these goals and benefit from education (Heward, Alper-Morgan, and Konrad 2017). IEP is an educational roadmap (Ruble et al. 2010; Sanches-Ferreira et al. 2013; Rakap 2015; Gibb and Dyches 2016) and is very im-

portant for educators, families, and students (Christle and Yell 2010). IEP is the main tool for accessing individualised services (Prettie-Frontzcak and Bricker 2000) and serves education/training, communication, management, responsibility, monitoring, and evaluation purposes (Bateman and Cline 2016). Moreover, IEP is the assurance of planning (Fiscus and Mandell 2002). As a comprehensive tool, IEP enables schools to record and monitor individualised annual goals and short-term objectives that respond to students' special needs (Ruble et al. 2010).

The significance of IEP and the goals it serves has caused research to discuss the extent to which IEPs are useful and whether they are suitable for their purpose (e.g., Boavida et al. 2014; Hunt, Goetz, and Anderson 1986; Prettie-Frontzcak and Bricker 2000; Rakap 2015), which has introduced the concept of IEP quality and its discussion. It is stated that IEP quality should be monitored to ensure that the goals and services determined in IEP are implemented and monitored as planned (Roach et al. 2009). In studies examining IEP quality, it is observed that various IEP components have been addressed in relation to IEP quality, such as current performance (Le Salle, Roach, and McGrath 2010; Ruble et al. 2010; Smith and Simpson 1989), functionality (Boavida et al. 2010; Hunt, Goetz, and Anderson 1986; Rakap 2015; Sanches-Ferreira et al. 2013), measurability-observability (Boavida et al. 2010; Prettie-Frontzcak and Bricker 2000; Rakap 2015; Ruble et al. 2010; Sanches-Ferreira et al. 2013), criterion (Boavida et al. 2010; Lynch and Beare, 1990; Prettie-Frontzcak and Bricker, 2000; Rakap, 2015; Ruble et al. 2010; Smith and Simpson 1989), generalisation, (Boavida et al. 2010; Hunt, Goetz, and Anderson 1986; Prettie-Frontzcak and Bricker 2000; Sanches-Ferreira et al. 2013) and goals/objectives appropriateness (Catone and Brady 2005; Lynch and Beare 1990; Prettie-Frontzcak and Bricker 2000; Smith and Simpson 1989), but an operational definition has not been established. IEP quality, addressed as a multi-component structure in this study, was defined as "The availability of the components that should be included in IEP based on student needs and the level of expressing statements about these components sufficiently, accurately, and clearly." This definition is formed on the basis that studies on IEP quality show that it is addressed in two main dimensions; the status of including the components that IEP should have (Hunt, Goetz, and Anderson 1986; Lucas et al. 2014; Lynch and Beare 1990; Rakap 2015; Rowland 2015; Ruble et al. 2010; Sanches-Ferreira et al. 2013; Smith and Simpson 1989) and/or level of expressing these components correctly (Boavida, Aguiar, and McWilliam 2014; Le Salle, Roach, and McGrath 2010; Lowman 2016; Rakap 2015).

Studies have revealed that the quality of the IEPs examined is generally low (Boavida et al., 2010; Giangreco et al. 1994; Lych and Beare 1990; Prettie-Frontzcak and Bricker 2000; Rakap 2015; Ruble et al. 2010; Sanches-Ferreira et al. 2013; Smith and Simpson 1989), they contain too many goal statements (Giangreco et al. 1994; Boavida et al. 2010; Sanches-Ferreira et al. 2013; Rakap 2015), IEP quality differs by diagnostic groups (Rakap 2015), and the IEP quality scores decrease as

the school stage increases (Catone and Brady 2005; Sanches-Ferriara et al. 2013). Furthermore, studies aiming to examine IEP quality have generally discussed goals in IEP (Boavida, Aguiar, and McWilliam 2014; Boavida et al., 2010; Catone and Brady 2005; Giangreco et al. 1994; Hunt, Goetz, and Anderson 1986; Le Salle, Roach, and McGrath. 2010; Lowman 2016; Lynch and Beare 1990; Prettie-Frontzcak and Bricker 2000; Rakap 2015; Rowland, Quinn, and Steiner 2015; Ruble et al. 2010; Sanches-Ferreira et al. 2013), and only one study (Smith and Simpson 1989) has examined IEP holistically. Additionally, it was observed that the psychometric properties of the measurement tools used in these studies were not explained or were poor. Concerning the validity of the measurement tools used, it was observed that expert opinion was obtained (Hunt, Goetz, and Anderson 1986), convergent validity was examined (Boavida et al. 2010; Lynch and Beare 1990; Rakap 2015), laws/institutional indicators were taken as a basis (Smith and Simpson 1989; Ruble et al. 2010), and they were expressed to be valid, but no information was found (Prettie-Frontzcak and Bricker 2000). Regarding reliability, inter-rater reliability analysis usually comes to the forefront (Boavida, Aguiar, and McWilliam 2014; Catone and Brady 2005; Hunt, Goetz, and Anderson 1986; Le Salle, Roach, and McGrath 2010; Rakap 2015; Ruble et al. 2010; Sanches-Ferriara et al. 2013). As a result, it is observed that the validity and reliability findings of the tools used to examine IEP quality in research are lacking and their psychometric properties are weak.

In line with this information, this study was planned in two stages, considering the fact that there are a limited number of studies addressing IEP quality holistically, although IEP contains more than one component, and the psychometric properties of the IEP quality assessment tools used in the studies are weak. The first stage of the study aimed to develop a measurement tool that provided validity and reliability and addressed IEP quality holistically. The second stage aimed to examine the quality of IEPs developed in special education schools using the created tool. In this context, answers to the following questions were sought: (1) What is (a) the validity of the IEP Quality Assessment Scale (IEPQAS)? (b) What is its reliability? (2) What is the quality of the IEPs developed in special education schools? (3) Does the quality of the IEPs developed in special education schools differ significantly by: (a) the number of goals, (b) diagnostic groups, (c) school stages, (d) and special education experiences of teachers?

METHOD

This study was conducted in two stages. At the first stage, a valid and reliable tool was developed to assess the quality of IEPs. At the second stage, the created tool was used to examine the quality of the IEPs prepared in special education schools.

Sample

For the first stage of the study, 180 IEPs were obtained from all public special education schools in the province where the researchers resided. The 180 IEPs reached at the first stage were used only in the standardisation of the measurement tool developed, and data on variables that might be related to IEP quality were not collected. At the second stage of the study, 211 IEPs were obtained from all public special education schools in the province where the researchers resided. A total of 391 IEPs were used.

At the second stage, data on variables that might be related were collected. Of the students whose IEPs were examined, 71 were female, 130 were male, and data on the gender of 10 students was not obtained. Of the IEPs, 93 (44.1%) belonged to students with mild intellectual disability (MID), 45 (21.3%) to students with moderate intellectual disability (MoID), 72 (34.1%) to students with severe intellectual disability (SID) or autism spectrum disorder (ASD), and one (0.5%) belonged to students with a hearing disability diagnosis. Of the IEPs, 11 (5.2%) belonged to preschool students, 67 (31.2%) to elementary school (1-4) students, 58 (28%) to middle school (5-8) students, and 75 (35.5%) to high school (9-12) students. Of the 75 special education teachers who developed IEPs, 65 were female, and 10 were male. The mean age of female teachers was 34.54 (min=23, max=54, sd=7.04), the mean age of male teachers was 39.10 (min=26, max=50, sd=8.86), and the mean age of all students was 34.15 (min=23, max=54, sd=7.05). The mean special education experience of the teachers was 10.03 years (min=1, max=29, sd=6.93).

Process and Data Collection Tools

A rubric is an assessment tool used to evaluate performance or product (Mertler 2001). In case there are multidimensional performances or products, and if these dimensions can be separated and detailed scoring is required, it is recommended to use analytic rubrics (Haladyna 1997). IEP is a product of the activities of the IEP team and has a multidimensional structure that can be separated from each other. Therefore, it was decided that in this research, IEP quality would be evaluated using an analytic rubric.

By reviewing the literature and laws on IEP, 19 rubric items related to IEP components were prepared. Tripartite ratings for each item and explanations for these ratings were created. At the first stage, validity and reliability analyses were conducted with 180 IEPs. As a result of these analyses, the IEPQAS consisting of 16 items, was developed. At the second stage, the quality of 211 IEPs was evaluated using the IEPQAS.

Information Questionnaire

Information was collected from teachers at the second stage of the study, using the developed information questionnaire. The information questionnaire contains information on the participant's gender, age, professional experience in special education, and his/her experience with students in his/her class.

IEPQAS

The IEPQAS has two subscales, the IEP instructional components subscale (ICS) and the IEP complementary components subscale (CCS). The ICS contains items related to student's performance, short-term objectives (STOs), goal statements, criterion, condition, usage, measurability-observability, goal-functionality-prerequisite relationship, and support services. The CCS contains items related to annual goals (AGs), the use of positive-negative statements, generalizability, time, tools, method, assessment, and behavioural problems. There are a total of 16 items, 8 items in each subscale, and the scores of the items are graded as 0, 1, and 2. For each item, the condition corresponding to this score is included in the scale as explained. As a result of the assessment, a minimum of 0 and a maximum of 32 points can be obtained. IEP quality increases as the score obtained from the scale increases.

Data Analysis

It is stated that the validity of rubrics can be examined as content validity, construct validity, and criterion-related validity, while its reliability can be examined as interrater and intrarater reliability (Moskal and Leydens, 2000). At the first stage of the study, expert opinion was obtained from five special education experts for content. Expert opinions were analysed using the content validity index (Davis 1992). Factor analysis was conducted for construct validity. For reliability, the internal consistency coefficient and upper-lower 27% reliability analyses were performed. To examine interrater and intrarater reliability, Kendall's W coefficient of concordance and the intraclass correlation coefficient was calculated. At the first stage, 60 IEPs (30%) randomly selected for intrarater reliability were re-evaluated by the first author. For interrater reliability, 60 randomly selected IEPs (30%) were scored by two research assistants in their doctoral dissertation period.

At the second stage, 211 IEPs were evaluated by the first researcher. One-way ANOVA, the independent samples t-test, and Pearson's correlation analysis were used in the data analysis. To examine intrarater and interrater reliability, Kendall's W coefficient of concordance and the intraclass correlation coefficient was analysed. At the second stage, 70 IEPs (30%) randomly selected for intrarater reliability were re-evaluated by the first author. Seventy IEPs (30%) randomly selected for interrater reliability were scored by a research assistant in his doctoral dissertation period.

At both stages of the study, it was determined that the data were normally distributed by examining the normal distribution data before proceeding to advanced analyses.

Ethics Committee Permission Information

In this study, all the rules specified in the “Anadolu University Science Ethics Guide” were followed.

Name of the Ethical Evaluation Committee: Anadolu University Social and Human Sciences Scientific Research and Publication Ethics Committee

Date of Ethics Evaluation Decision: 20.02.2018

Ethics Evaluation Document Issue Number: 19113

RESULTS

First Stage

Validity of the IEPQAS

As a result of the content validity analysis, 18 items received 1 point, one item received 0.80 points, and 19 items were preserved. To determine the suitability for factor analysis for construct validity, the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests, inter-item correlation values and correlation matrix determinant, and anti-image correlation were examined to determine suitability for factor analysis for construct validity. The KMO coefficient was .78, and the significance value of Bartlett's test of sphericity was found to be $p < .001$ ($\chi^2(120) = 1174.220$). There is no item above .90 in the inter-item correlation, and the determinant of the correlation matrix is .001. In the anti-image correlation examination, two items with a correlation value below .50 were discarded, and the analysis was repeated. The anti-image correlation values obtained from the analysis vary between .614 and .846. As a result of the factor analysis, a sixteen-item two-factor scale with a total variance of 45.76% was formed, together with the first eight-item factor with an eigenvalue of 4.09 and explained variance ratio of 25.56% and the second eight-item factor with an eigenvalue of 3.23 and explained variance ratio of 20.19%. The first factor was named the 184IEP instructional components subscale-ICS-, and the second factor was named the 184IEP complementary components subscale-CCS-. The factor loading values and explained variance ratios of the IEPQAS are presented in table 1.

Table 1. The IEPQAS Factor Loading Values and Explained Variance Ratios

Items	Rotated Factor Loading Values			Total Variance Explained %
	IEPQAS-ICS	IEPQAS-CCS	Eigenvalue	
M1	.67			25,569
M3	.68			
M4	.81			
M5	.81			
M6	.54			
M7	.82		4,091	
M16	.52			
M18	.47			
M2		.34		
M8		.61		
M11		.71		
M12		.35		
M13		.50		
M14		.71	3,231	
M15		.83		
M19		.60		
Toplam			7,322	45,762

Reliability of the IEPQAS

Cronbach's alpha of the internal consistency coefficient was calculated to be $\alpha=.84$ for the ICS, $\alpha=.64$ for the CCS, and $\alpha=.80$ for the overall scale. In the t-test performed on upper-lower 27% reliability, the mean value of the lower group was found to be $\bar{X}=7.35$ ($sd=1.22$), and the mean value of the upper group to be $\bar{X}=21.10$ ($sd=4.28$) ($p<.001$). The effect size of the difference was calculated with eta squared ($\eta^2=.83$).

For interrater reliability, the two-way mixed intraclass correlation coefficient alpha was found to be .99 when calculated for both absolute agreement and consistency ($p<.001$). As a result of the two-way mixed analysis for intrarater reliability, the intraclass correlation coefficient alpha of absolute agreement was found to be .99 ($p<.001$). For Kendall's W, intrarater reliability and Kendall's W coefficient of concordance among three raters were found to be .99 ($p<.001$).

Second Stage

IEP Quality Level

The mean quality score of IEPs (n=211) was found to be $\bar{X}=9.66$ (sd=2.62) for the IEPQAS, $\bar{X}=4.98$ (sd=2.10) for the ICS, and $\bar{X}=4.67$ (sd=1.99) for the CCS. Considering that the highest score that can be obtained from the IEPQAS is 32 and the lowest score is 0, the obtained quality scores are relatively low.

Number of goals of IEPs and IEP quality

When the numbers of IEP goals (n=211) were examined, it was observed that the number of AGs varied between 1 and 227, the mean was $\bar{X}=31$ (sd=25.42), the number of STOs varied between 1 and 429, the mean was $\bar{X}=145$ (sd=92.68), the total number of goals and objectives varied between 2 and 595, and the mean was $\bar{X}=177$ (sd=113.11). The correlation between the numbers of AGs, STOs and total goals and objectives in IEPs and the IEPQAS scores was checked with Pearson's correlation analysis. No significant correlations were observed between the number of AGs, number of STOs, total goals and objectives and IEP quality for the overall scale and the CCS. A significant negative correlation was observed between them for the ICS. Correlation values between the IEPQAS, ICS, and CCS and total goals and objectives are given in table 2.

Table 2. Correlation Values Between the IEPQAS, ICS, and CCS and Total Goals and Objectives

Variables	AG	STO	Total
IEPQAS	-.116	-.042	-.060
IEPQAS-ISC	-.210**	-.138*	-.160*
IEPQAS-CCS	.068	.090	.089

**p<.01 *p<.05

Diagnostic Groups and IEP Quality

As a result of the one-way ANOVA analysis, no significant difference was noted between the mean IEPQAS total score and diagnostic groups ($F(2, 207)=.680$, $p>.05$). However, a significant difference was observed for the mean ICS score ($F(2, 207)= 33.140$, $p<.01$) and the mean CCS score ($F(2, 207)=52.184$, $p<.01$). The mean ICS score of students diagnosed with MID ($\bar{X}_{MID}=6.15$) was found to be significantly higher than the MoID ($\bar{X}_{MoID}=4.02$) and SID/ASD ($\bar{X}_{SID/ASD}=4.15$) groups ($\eta^2 =0.24$). The mean CCS score of students diagnosed with MID ($\bar{X}_{MID}=3.40$) was found to be significantly lower than the MoID and SID/ASD groups ($\bar{X}_{MoID}=5.51$; $\bar{X}_{SID/ASD}=5.83$) ($\eta^2 =0.34$).

School Stages and IEP Quality

Since there were only 11 IEPs for the preschool period, this school stage was excluded from the analysis. As a result of the one-way ANOVA analysis, a significant difference was observed between the mean IEPQAS ($F(2, 197)=93.896$, $p<.01$), ICS ($F(2, 197)=25.064$, $p<.01$) and CCS ($F(2, 197)= 22.738$, $p<.01$) scores and school stages. For the IEPQAS total score, it was revealed that the mean quality score of IEPs of elementary school students ($\bar{X}_{\text{elementary school}}=11.74$) was significantly higher than the middle school ($\bar{X}_{\text{middle school}}=9.44$) and high school ($\bar{X}_{\text{high school}}=7.75$) groups and the mean quality score of IEPs of middle school students ($\bar{X}_{\text{middle school}}=9.44$) was significantly higher than the high school group ($\bar{X}_{\text{high school}}=7.75$), ($\eta^2 =0.49$). For the ICS, the mean quality score of IEPs of elementary school students ($\bar{X}_{\text{elementary school}}=6.33$) was observed to be significantly higher than the middle school and high school groups ($\bar{X}_{\text{middle school}}=4.59$; $\bar{X}_{\text{high school}}=4.18$), ($\eta^2 =0.20$). For the CCS, the mean quality score of the elementary school group ($\bar{X}_{\text{elementary school}}=5.41$) was significantly higher than the middle school group ($\bar{X}_{\text{middle school}}=4.85$), and the mean quality score of the middle school group ($\bar{X}_{\text{middle school}}=4.85$) was significantly higher than the high school group ($\bar{X}_{\text{high school}}=3.56$), ($\eta^2=0.19$). As the school stage increased, IEP quality was observed to decrease.

Teachers' Special Education Experience and IEP Quality

The special education experiences of the teachers were examined by creating four groups by year. IEPs prepared by teachers with 1-5 years, 6-10 years, 11-15 years, and 16 years+ of special education experience were included in the first, second, third, and fourth groups, respectively. As a result of the one-way ANOVA analysis, a significant difference was found between the mean IEPQAS ($F(3, 207)=8.945$, $p<.01$), ICS ($F(3, 207)=8.826$ and $p<.01$) and CCS ($F(3, 207)=15.224$, $p<.01$) scores and special education experiences of teachers. The mean IEPQAS total score of the IEPs prepared by teachers in the first group ($\bar{X}_{\text{First}}=11.04$) was detected to be significantly higher than the second ($\bar{X}_{\text{Second}}=9.56$), third ($\bar{X}_{\text{Third}}=9.49$) and fourth ($\bar{X}_{\text{Fourth}}=8.25$) groups ($\eta^2 =0.12$). For the ICS, the mean quality score of the IEPs prepared by the first group ($\bar{X}_{\text{First}}=8.26$) was revealed to be significantly higher than the fourth group ($\bar{X}_{\text{Fourth}}=3.91$) ($\eta^2 =0.11$). For the CCS, the mean quality score of the IEPs prepared by the first group ($\bar{X}_{\text{First}}=5.86$) was observed to be significantly higher than the third group ($\bar{X}_{\text{Third}}= 3.73$) and the fourth group ($\bar{X}_{\text{Fourth}}=4.34$) ($\eta^2 =0.18$).

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

At the first stage of this study, the validity and reliability analyses of the IEPQAS were conducted. The analyses put forward a two-dimensional scale consisting of 16 items. Validity findings of the IEPQAS were discussed in terms of content validity and construct validity. Criterion-related validity could not be checked as no other Turkish tools with confirmed validity and reliability that measured IEP quality, could be accessed. In content validity, validity was ensured as all items received a value above .80 (Davis, 1992). To evaluate the suitability for factor analysis, the significance value of .000 ($p < .05$) of Bartlett's test of sphericity showed that the results of the test were significant and suitable for factor analysis (Field 2009; Pallant 2020; Can 2019). The KMO coefficient of .78 ensured suitability for factor analysis (Can 2019; Tabachnick and Fidell 2014). The absence of items with a value above .80 or .90 in the study indicated that inter-item correlation values were acceptable for factor analysis, and there were no multicollinearity problems (Can 2019; Field 2009; Tabachnick and Fidel 2014). In the anti-image correlation examination, values obtained from the analysis performed after excluding two items varied between .653 and .845 ($> .50$), which showed that items were suitable for factor analysis (Can, 2019; Field, 2009). As a result of the factor analysis, the eigenvalues of both factors of the IEPQAS were above 1 (Buyukozturk 2014; Field 2009; Huck 2012; Kline 1994; Tabachnick and Fidell 2014; Pallant 2020), and the eigenvalue of each of the factors was higher than 5% of the total eigenvalue (Huck 2012). This indicated that factor formation was appropriate. The value of explained variance is expected to be higher than 30% for single-factor constructs, and a higher value is expected for constructs with more than one factor (Buyukozturk 2014). Accordingly, 45% variance of the IEPQAS in the two-factor construct is at an adequate level. Factor loading values of the IEPQAS items above .30 (Buyukozturk 2014; Ozdamar 2017) and a difference of .10 between factor loadings for the items creating factor loading in more than one factor (Buyukozturk 2014), showed that the items were associated with factors and supported the suitability for factor analysis.

Concerning reliability, Cronbach's alpha coefficient of $\alpha = .80$ for the overall IEPQAS and $\alpha = .84$ for the ICS indicated a good or high level of reliability (Can 2019; Cresswell 2012; George and Mallery 2003; Ho 2006; Ozdamar 2017). For the CCS, $\alpha = .64$ was also found at an acceptable level (Can 2019; George and Mallery 2003; Nunnally and Bernstein 1994; Ozdamar 2017). In upper-lower 27% test results, it was shown that the difference between the two groups was significant ($p < .01$), those with and without the characteristics planned to be measured were differentiated well (Can, 2019), the internal consistency of the test was shown to be an indicator, and the effect size of the difference pointed to a large effect size (Buyukozturk 2014; Pallant 2020). Considering the intraclass correlation coefficient, values of .99 obtained for intrarater and interrater reliability showed that very good reliability was provided (Koo and Li, 2016) and the rubric was reliable

(Kutlu et al., 2014). Kendall's *W* coefficient of concordance is $p=.000$ for interrater and intrarater reliability and significant at the level of $p<.001$ (Can 2019). Kendall's *W* values of .99 provide positive evidence for interrater and intrarater reliability (Can 2019; Field 2009; Legendre 2005).

In this study, it can be stated that the requirements for content validity were provided and the criteria emphasised in the literature for providing construct validity were met in the formation of factors for construct validity, in general. Therefore, the content and construct validity of the IEPQAS, developed as a rubric, can be said to have been provided. The IEPQAS is the only tool with ensured construct validity among the tools that examine IEP quality in the literature. When the tools developed to examine IEP quality were reviewed, the validity of the IEPQAS was put forward more strongly than all tools developed to investigate IEP quality, because the validity information of some of the tools could never be accessed in studies (Le Salle, Roach and McGrath 2010; Lucas 2014; Lowman 2016; McWilliam 2021; McWilliam, 2021; Rowland 2015). The construct validity of the tools was not examined in any study with provided validity information, only convergent validity was examined in some studies (Boavida et al. 2010; Lynch and Beare 1990; Rakap 2015), or their content validity was accepted to be provided in view of laws and the literature (Ruble et al., 2010; Smith and Simpson, 1989). The IEPQAS was observed to be a reliable tool fulfilling the requirements in the reliability analyses. The reliability of the tool was shown to be strong because it was the first study meeting the reliability qualities of the upper-lower 27% test as well as internal consistency.

The mean general quality scores obtained from the IEPQAS, and subscales are quite low. This result showed that the quality of the IEPs developed is low and is consistent with the literature (Smith and Simpson, 1989; Lynch and Beare, 1990; Giangreco et al., 1994; Prettie-Frontzack and Bricker, 2000; Boavida et al. 2010; Ruble et al., 2010; Sanches-Ferreira et al., 2013; Rakap, 2015). The low mean scores obtained from the ICS were consistent with the lack of performance expressions (Boavida et al., 2010; Le Salle et al., 2010; Smith and Simpson, 1989), incompliance between goals and performance (Catone and Brady 2005; Smith and Simpson 1989), inadequacy of the goals in terms of functionality (Boavida et al. 2010), deficiencies in determining the associated services in IEP and the start and ending dates of services (Smith and Simpson 1989), goals written in unclear and general terms (Catone and Brady 2005; Giangreco et al. 1994; Lynch and Beare 1990), deficiencies in success criteria (Boavida et al. 2010; Giangreco et al. 1994; Lynch and Beare 1990; Sanches-Ferreira et al. 2013), deficiencies in the conditional component of goals (Boavida et al., 2010; Giangreco et al. 1994; Lynch and Beare 1990; Rakap 2015; Ruble et al. 2010) and the target behaviour not being presented with a measurable or observable expression (Boavida et al. 2010; Ruble et al. 2010; Sanches-Ferreira et al. 2013). These were among the results of the studies

on IEP quality in the literature. The low mean scores obtained from the CCS were consistent with results such as inadequate emphasis laid on learning strategies in studies in the literature (Lynch and Beare 1990; Ruble et al. 2010), deficiencies in time expressions (Boavida et al. 2010; Rakap 2015; Ruble et al. 2010), deficiencies in generalisation (Boavida et al. 2010; Rakap 2015), and inadequacies in evaluation (Smith and Simpson 1989).

No significant correlations could be detected between the number of AGs, STOs and total goals and objectives and the IEPQAS and CCS, but a significant negative correlation with the ICS was observed. It was observed that, as the number of AGs, STOs and total goals and objectives increased, the mean ICS score decreased. This suggests a decrease in goal writing quality. Furthermore, it was observed that the number of goals of IEPs was quite high, and this result is consistent with the literature (Giangreco et al. 1994; Boavida et al. 2010; Sanches-Ferreira et al. 2013; Rakap 2015). In their study, Boavida et al. (2010) reported that the number of AGs varied between 4 and 95 ($\bar{X}=38$). Rakap (2015) stated that the number of AGs/STOs varied between 8 and 38 ($\bar{X}=22$). On the other hand, Sanches-Ferreira et al. (2013) indicated that 80% of IEPs had more than 20 AGs. Therefore, the number of goals in this study is extremely high. The reason for the high number of goals in the literature is explained as teachers' concentration on the completion of students' procedural requirements rather than increasing their acquisitions due to the difficulty they experience in determining which behaviours are important (Shinn and Shinn 2000). This explanation is supported by the conclusion of Boavida et al. (2014) that the number of goals of IEPs prepared after the education given to increase the quality of IEP goals decreased, and their quality increased. Moreover, it is stated that a high number of goals brings along many requirements for teachers and may cause teachers to fail in monitoring practices (Fuchs and Shinn 1989). Additionally, it is asserted that such a high number of goals raises the suspicion that teachers do not use IEP (Giangreco et al. 1994). Such a high number of goals both brings into doubt that IEPs are not used and suggests that the whole team, especially teachers, due to the difficulties they face in determining goals, prepare the IEP not as a functional tool but as a document that is not possible to achieve, far from being an educational guide, and prepared only due to legal necessity.

Considering the total score obtained from the IEPQAS, IEP quality did not differ according to diagnostic groups. This result is consistent with the result obtained by Rakap (2015) in his study, but it contradicts the research results of Boavida et al. (2010). However, when the subscales were examined, significant differences were observed between IEP quality and diagnostic groups. While the mean quality score of the MID group is higher than the MoID and SID-ASD groups in the first subscale, the mean quality score of the MID group is lower than the MoID and SID-ASD groups in the second subscale. The reason for the higher mean quality score of the MID group in the first subscale may be that teachers feel more com-

petent and have less difficulty in creating quality goals for these children. They may not have given enough weight to IEP components in the second subscale as a result of experiencing less difficulty or feeling more competent in goals. Boavida et al. (2010) stated that teachers might have given more effort to writing functional goals for students who had greater needs or severer disabilities and whose needs dominantly related to the goals of independent life. In this study, teachers may have tried to complement their deficiencies concerning functional goals with the components in the CCS. The ICS and CCS effect size values are large effect sizes (Palant, 2020; Buyukozturk, 2014). This opposite situation observed in the subscales may have prevented the emergence of a difference in the total score of the scale.

When the mean IEP quality scores were examined according to school stages, IEP quality scores decreased as the school stage increased, and the effect size was large. This result contradicts the finding of Le Salle et al. (2010) in the literature and is supported by the findings of some studies (Catone and Brady 2005; Sanches-Ferriara 2013). This can be thought to result from the increase in the diversity and complexity of academic fields as the school stage increases (Sanches-Ferreira et al. 2013).

When the mean IEP quality scores were examined according to teachers' special education experiences, it was observed that the mean IEP quality score of teachers with 1-5 years of special education experience was significantly higher than the other groups, and the effect size of the difference was close to large. The findings of this study contradict the finding of Ruble et al. (2010), indicating that there is no relationship between teachers' experience of working with children with ASD and the quality of IEPs. However, in this study, different diagnostic groups and the general professional experience in special education, not the experience specific to a diagnosis, were addressed. It is thought that teachers with 1-5 years of special education experience have the most up-to-date knowledge and approaches, and therefore, can ensure a significantly higher mean IEP quality score in this group compared to the other groups.

IEP is an educational roadmap that needs to be developed for all students with special needs, and this requirement is expressed legally. It is also a tool that contains information about the services to be received by students and enables observing the student. However, it can be clearly stated that the examined IEPs do not have these characteristics. It is reported that the role of IEP in determining the instructional program as a tool is investigated in the literature (Berninger 1997). Smith and Simpson (1989) stated more clearly that the inability of the IEP to provide the necessary components compromised its function of educational guidance. Likewise, the results obtained in this study may question the ability of IEP to provide its guidance role in designing effective education. IEP is reported to be a mechanism through which, teaching methods and education planned specially

are clearly described for teachers, teachers' assistants, school personnel and families (Ruble et al. 2010) and the assurance of the plans made (Fiscus and Mandell 2002). However, the assurance provided by the IEP may be inadequate or lead to misinformation for the child's future teachers, service providers and family, due to its significant deficiencies. In the literature, it is stated that teachers have difficulties in preparing IEP and do not have adequate information about the content of IEP (Çıkılı et al. 2020; Ozturk and Eratay 2010). The reason for the overall low IEP quality scores and the extremely high number of goals may result from teachers' insufficient knowledge about the subject. Increasing the quality of IEPs in general and reducing the high number of goals to an achievable number in line with educational needs, time and student performance is important for IEP to assume the role of educational guidance. In other words, these changed would allow IEP to be functional, and for teachers, families and relevant service providers to follow it. To achieve this, teachers' knowledge of IEP should be evaluated, and training should be provided to bring their knowledge to a sufficient level and support it if required. The quality of IEPs of students who are in integrative environments can be investigated in future studies. Moreover, the relationships between IEP quality and classroom practices and their effect on students' educational levels can be examined.

Thanks and Explanations

We would like to thank Çetin TOPUZ and Şerife ŞAHİN who contributed to our study.

Conflict of Interest

There is no personal or financial conflict of interest between the authors of the article within the scope of the study.

Author Contribution

Design of Study: ŞA(%50), VA(%50)

Data Acquisition: ŞA(%50), VA(%50)

Data Analysis: ŞA(%50), VA(%50)

Writing Up: ŞA(%50), VA(%50)

Submission and Revision: ŞA(%50), VA(%50)

REFERENCES

- Berninger, V. W. (1997). "Introduction to intervention for students with learning and behavior problems: Myths and realities." *School Psychology Review*, 26(3): 326-332 <https://doi.org/10.1080/02796015.1997.12085870>
- Boavida, T., C. Aguiar, R. A. McWilliam, and J. S. Pimentel. (2010). "Quality of Individualized Education Program Goals of Preschoolers with Disabilities." *Infants and Young Children*, 23(3): 233-243. <https://doi.org/10.1097/IVC.0b013e3181e45925>
- Boavida, T., C. Aguiar, and R. A. McWilliam. (2014). "A Training Program to Improve IFSP/IEP Goals and Objectives through the Routines-Based Interview." *Topics in Early Childhood Special Education*, 33(4): 200-211. <https://doi.org/10.1177/0271121413494416>
- Büyükoztürk, S., E. Kılıç Çakmak, O. E. Akgün, S. Karadeniz, and F. Demirel. (2014). *Bilimsel Araştırma Yöntemleri* [Scientific Research Methods]. (18th ed.). Ankara: Pegem Akademi.
- Can, A. (2019). *SPSS ile Bilimsel Araştırma Sürecinde Nicel Veri Analizi* [Quantitative Data Analysis in Scientific Research Process with SPSS]. Ankara: Pegem Akademi.
- Catone, W. V., and S. A. Brady. (2005). "The Inadequacy of Individual Educational Program (IEP) Goals for High School Students with Word-Level Reading Difficulties." *Annals of Dyslexia*, 55(1): 53-78. <https://doi.org/10.1007/s11881-005-0004-9>
- Christle, C. A., and M. L. Yell. (2010). "Individualized Education Programs: Legal Requirements and Research Findings." *Exceptionality*, 18(3): 109-123. <https://doi.org/10.1080/09362835.2010.491740>
- Creswell, J. W. (2012). *Educational Research Planning, Conducting And Evaluating Quantitative And Qualitative Research* (4th ed.). Boston: Pearson Education.
- Çıkkılı, Y., A. Gonen, O. A. Bağcı, and H. Kaynar. "Özel Eğitim Alanında Görev Yapan Öğretmenlerin Bireyselleştirilmiş Eğitim Programı (BEP) Hazırlama Konusunda Yaşadıkları Güçlükler." *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 15(1): 5121-5148, <https://doi.org/10.26466/opus.659506>
- Davis, L. L. (1992). Instrument Review: Getting the Most from A Panel of Experts. *Applied Nursing Research*, 5(4): 194-197. [https://doi.org/10.1016/S0897-1897\(05\)80008-4](https://doi.org/10.1016/S0897-1897(05)80008-4)
- DeVellis, R. F. (2017). *Scale development: Theory and applications*. (3th ed.). California: Sage Publications Inc.
- Field, A. (2009). *Discovering Statistics Using SPSS (and Sex and Drugs and Rock 'n' Roll)* (3rd Ed.). USA: Sage Publications.
- Fiscus, E. D., and C. J. Mandell. (2002). *Bireyselleştirilmiş Eğitim Programlarının Geliştirilmesi* [Developing Individualized Education Programs]. (2nd ed.). Translated and edited in Turkish by Gonul Akcamete, Hatice Günayer Şenel and Elif Tekin. Ankara: Anı Yayıncılık.
- Fuchs, L. S. and Shinn, M. R. (1989). "Writing CBM IEP Objectives." In *Curriculum-Based Measurement: Assessing Special Children*, Edited by Mark R. Shinn, 130-152. New York: The Guilford Press.
- George, D., and P. Mallery (2003). *SPSS for Windows Step by Step: A simple guide and reference. 11.0 update* (4th ed.). USA: Allyn and Bacon.
- Giangreco, M. F., R. E. Dennis, S. W. Edelman, and C. J. Clonger. (1994). "Dressing Your IEPs for the General Education Climate Analysis of IEP Goals and Objectives for Students with Multiple Disabilities." *Journal for Special Educators*, 15(5): 288-296. <https://doi.org/10.1177/074193259401500504>
- Gibb, G. S., and T. T. Dyches (2016). *IEPS: Writing Quality Individualized Education Programs* (3rd. ed.). USA: Pearson Publications.
- Haladyna, T. M. (1997). *Writing Test Item to Evaluate Higher Order Thinking*. USA: Allyn and Bacon.
- Heward, W. L., S. R. Alper-Morgan, and M. Konrad M. (2017). *Exceptional Children: An Introduction to Special Education* (11th ed.). USA: Pearson Publications.
- Ho, R. (2006). *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*. New York: Chapman and Hall/CRC.
- Huck, S. W. (2012). *Reading Statistics and Research*. (6th ed.). Boston: Pearson Educations, Inc.
- Hunt, P., L. Goetz, and J. Anderson. (1986). The Quality of IEP Objectives Associated with Placement on Integrated Versus Segregated School Sites. *Research and Practice for Persons with Severe Disabilities*, 11(2): 125-130. <https://doi.org/10.1177/154079698601100206>
- Kline, P. (1994). *An Easy Guide to Factor Analysis*. USA and Canada: Routledge Taylor and Francis Group
- Koo, T. K., and Li, M. Y. (2016). "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research." *Journal of Chiropractic Medicine*, 15(2): 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kutlu, O., Dogan, C. D. ve Karakaya, İ. (2014). *Ölçme ve Değerlendirme* [Measurement and Assessment]. Ankara: Pegem Akademi.
- Le Salle, T. P., A. T. Roach, and D. McGrath (2010). "The Relationship of IEP Quality to Curricular Access and Academic Achievement for Students with Disabilities." *International Journal of Special Education*, 28(1): 135-144.

- Legendre, P. (2005). "Species Associations: the Kendall Coefficient of Concordance Revisited." *Journal of Agricultural, Biological, and Environmental Statistics*, 10(2): 226-245 <https://doi.org/10.1198/108571105X46642>
- Lowman, J. J. (2016). "A Comparison of Three Professional Development Mechanisms for Improving the Quality of Standards-Based IEP Objectives." *Communication Disorders Quarterly*, 37(4): 211-224. <https://doi.org/10.1177/1525740115595338>
- Lucas, A., K. Gillaspay, M. L. Peters, and J. Hurth. (2014). "Enhancing recognition of high quality, functional iep goals: A training activity for early childhood special education staff." Accessed 5 May 2021 <http://www.ectacenter.org/~pdfs/pubs/rating-iep.pdf>
- Lynch, E. C., and P. L. Beare. (1990). "The Quality of IEP Objectives and Their Relevance to Instruction for Students with Mental Retardation and Behavioral Disorders." *Remedial and Special Education*, 11(2): 48-55. <https://doi.org/10.1177/074193259001100207>
- McWilliam, R. A. (2005). "Goal Functionality Scale." Accessed 5 May 2021 https://www.utoledo.edu/education/grants/direct/pdf/Goal_Functionality_Scale.pdf
- McWilliam, R. A. (2014). Accessed 5 May 2021. "Goal Functionality Scale III." https://ectacenter.org/~pdfs/topics/families/GoalFunctionalityScaleIII_2_.pdf
- Mertler, C. A. (2001). "Designing Scoring Rubrics for Your Classroom." *Practical Assessment, Research and Evaluation*, 7(25): 1-10.
- Moskal, B. M. and J. A. Leydens. (2000). "Scoring Rubric Development: Validity and Reliability." *Practical Assessment, Research and Evaluation*, 7(10): 71-81.
- Nunnally, J.C. and I. R. Bernstein. (1994). *Psychometric Theory* (3rd. Ed.). New York: McGraw-Hill
- Ozdamar, K. (2017). *Öçek ve Test Geliştirme Yapısal Eşitlik Modellemesi*. [Scale and Test Development Structural Equation Modeling] Eskisehir: Nisan Kitapevi.
- Ozturk, C. C. and E. Eratay. (2010). "Eğitim uygulama okuluna devam eden zihin engelli öğrencilerin öğretmenlerinin bireyselleştirilmiş eğitim programı hakkında görüşlerinin belirlenmesi." *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 10(2): 145-159.
- Pallant, J. (2020). *SPSS Survival Manual. A Step by Step Guide to Data Analysis Using SPSS* (7th ed.). London: Open University Press
- Pretti-Frontczak, K., and D. Bricker. (2000). "Enhancing the Quality of Individualized Education Plan (IEP) Goals and Objectives." *Journal of Early Intervention*, 23(2): 92-105. <https://doi.org/10.1177/105381510002300204>
- Rakap, S. (2015). Quality of Individualised Education Programme Goals and Objectives for Preschool Children with Disabilities. *European Journal of Special Needs Education*, 30(2): 173-186. <https://doi.org/10.1080/088562572014.986909>
- Roach, A. T., N. Chilungu, T. P. La Salle, D. Talapatra, and M. Vignieri, (2009). "Opportunities and Options for Facilitating and Evaluating Access to the General Curriculum for Students with Disabilities." *Peabody Journal of Education*, 84(4): 511-528. <https://doi.org/10.1080/01619560903240954>
- Rowland, C. M., E. D. Quinn, and S. A. Steiner. (2015). "Beyond Legal: Crafting High-Quality IEPs for Children with Complex Communication Needs." *Communication Disorders Quarterly*, 37(1): 53-62. <https://doi.org/10.1177/1525740114551632>
- Ruble, L. A., J. McGrew, N., Dalrymple, and L. A. Jung. (2010). "Examining the Quality of IEPs for Young Children with Autism." *Journal of Autism and Developmental Disorders*, 40(12): 1459-1470. <https://doi.org/10.1007/s10803-010-1003-1>
- Sanches-Ferreira, M., P. Lopes-dos-Santos, S., Alves, M. Santos, and M. Silveira-Maia. (2013). "How Individualised are the Individualised Education Programmes (IEPs): an Analysis of the Contents and Quality of the IEPs Goals." *European Journal of Special Needs Education*, 28(4): 507-520. <https://doi.org/10.1080/088562572013.83043>
- Shinn, M. R., and M. M. Shinn. (2000). "Writing and Evaluating IEP Goals and Making Appropriate Revisions to Ensure Progress and Participation in General Curriculum. In *IDEA Amendments of 1997: Practice Guidelines For School-Based Teams*, Edited by Cathy F. Tetzlow and Melody Tankersley, 351-382. USA: National Association of School Psychologists.
- Smith, S. W., and R. L. Simpson, R. L. (1989). "An Analysis of Individualized Education Programs (IEPs) for Students with Behavioral Disorders." *Behavioral Disorders*, 14(2): 107-116. <https://doi.org/10.1177/019874298901400206>
- Tabachnick, B. G., and L. S. Fidell. (2014). *Using Multivariate Statistics*. (6th ed.) USA: Pearson Education