

Yapay Zeka Önyargı Ölçeğinin Geliştirilmesi

Güler ÇETİN, Merve UZUN ve Mehmet ŞATA

Özet

....

Anahtar Kelimeler: Yapay zeka, Ölçek geliştirme, Geçerlik, Güvenirlik, Önyargı.

Giriş

.....

Yöntem

Araştırmanın Modeli

Öğretmenlerin yapay zekaya yönelik önyargı düzeylerini ölçmeyi amaçlayan bir ölçek geliştirme çalışması olarak tasarlanan bu araştırma, nicel araştırma yaklaşımlarından uygulamalı ve betimsel araştırma modeli ile yürütülmüştür (Şata, 2020). Betimsel araştırmalar genel olarak büyük gruplar üzerinden ve herhangi bir müdahale yapılmadan yürütülen araştırma modelleri olarak tasarlanmaktadır.

Evren ve Örneklem

Bu araştırmanın evreni öğretmenlerden oluşmakta iken, örnekleme ise seçkisiz olmayan örnekleme yöntemlerinden uygun örnekleme yöntemi ile seçilen 320 öğretmenden oluşmaktadır. Araştırmaya katılım gönüllülük esasına bağlı olarak gerçekleştirilmiştir. Veri temizleme ve inceleme sürecinde kayıp değerin olmadığı tespit edilmiş ve herhangi bir değer ataması yapılmamıştır. Ayrıca uçdeğer analizi yapılmış ve herhangi bir uçdeğer olmadığı tespit edilmiştir. Araştırmaya katılan öğretmenlerin sosyo demografik değişkenlerine ait betimsel istatistikler tablo 1’de verilmiştir.

Tablo 1. Öğretmenlerin sosyo demografik değişkenlerine ait frekans ve yüzde değerleri

Değişkenler	Değişken düzeyi	f	%
Cinsiyet	Kadın	213	66.6
	Erkek	107	33.4
Yaş düzeyi	25 yaş ve altı	69	21.6
	26-30 yaş	108	33.8
	31-35 yaş	70	21.9
	36 yaş ve üstü	73	22.8
	Düşük	42	13.1
Sosyo ekonomik düzey	Orta	261	81.6
	Yüksek	17	5.3
	Okulöncesi	39	12.2
Çalıştığı okul kademesi	İlkokul	73	22.8
	Ortaokul	77	24.1
	Lise	131	40.9
	Lisans	256	80.0
Eğitim düzeyi	Lisansüstü	64	20.0
	Bekar	150	46.9
Medeni durum	Evli	163	50.9
	Boşanmış/dul	7	2.2
	Toplam	320	100.0

Tablo 1 incelendiğinde, kadın öğretmen sayısının erkek öğretmen sayısından iki kat daha fazla olduğu görülmektedir. Yaş dağılımları incelendiğinde en fazla 26-30 yaş aralığına sahip öğretmenin olduğu tespit edilmiştir. Öğretmenlerin büyük çoğunluğu kendisini orta düzeyde gelire sahip olarak görmektedir. Öğretmenlerin en fazla lisede en az ise okulöncesi düzeyinde çalıştıkları ve çok az bir kısmının lisansüstü eğitime sahip olduğu tespit edilmiştir. Medeni durumları incelendiğinde bekar ve evli öğretmenlerin birbirine yakın oranlara sahip olduğu tespit edilmiştir.

Veri Toplama Aracının Geliştirilmesi

Araştırmanın amacı doğrultusunda öğretmenlerin yapay zekaya yönelik önyargı düzeylerini belirlemek için geliştirilen ölçme aracının geliştirime aşamaları sistematik bir süreç içinde ele alınmıştır. İlk olarak alanyazın taraması yapılarak daha önce bir ölçme aracının geliştirilip geliştirilmediği belirlenmeye çalışılmıştır. Yapılan alanyazın taraması sonucunda, toplumun farklı kesimlerine yönelik yapay zekaya yönelik farklı ölçeklerinin olduğu fakat doğrudan öğretmenlere yönelik herhangi bir ölçme aracının olmadığı tespit edilmiştir. İkinci aşama olarak önyargı kavramının kuramsal yapısı araştırılmış ve kuramına göre alt boyutları belirlenmiştir. Üçüncü aşamada ise önyargının her bir alt boyutunu yeterli düzeyde temsil edecek şekilde maddeler yazılarak madde havuzu oluşturulmuştur. İlk olarak 32 madde yazılmış

daha sonra benzer anlama sahip maddelerin olduđu belirlenerek 23 maddeye düşürülmüştür. Elde edilen taslak form 4 PDR uzmanı, 2 Türk Dili uzmanı ve 2 Ölçme ve değerlendirme uzmanı olmak üzere sekiz uzmana gönderilerek maddelerin kapsam geçerliğine ilişkin kanıtların toplanması sağlanmıştır. Uzman görüşü Lawshe tekniğı kullanılarak yapılmıştır (Lawshe, 1975). Lawshe tekniğinde sekiz uzman için her bir maddenin eşik değeri .693 olduđu belirlenmiş (Wilson vd., 2012) ve bu eşik değerin altındaki maddeler taslak formdan çıkartılmıştır. Her bir madde için kapsam geçerlik oranı hesaplanmış ve tüm maddenin kapsam geçerlik oranlarının .693 eşik değerinden yüksek olduđu belirlenmiş ve ölçme aracından herhangi bir madde çıkartılmamıştır. Ayrıca uzman görüşleri doğrultusunda maddelerin gerekli ama düzeltilmeli dediğı dört madde düzeltilmiş ve bir madde ise iki maddeye dönüştürülmüştür. Sonuç olarak uzmanlara 23 madde olarak gönderilen taslak form 24 maddeye çıkartılmış ve uygulama öncesi son halini almıştır. Ölçme aracının geliştirme süreci tamamlandıktan sonra veri toplama sürecine geçilmiştir.

Verinin Toplanması

Araştırmacılar tarafından geliştirilen ölçme aracı, çevrimiçi ortamda uygulanmış ve veri toplama süreci bu platformun sağladığı birçok avantajdan faydalanmıştır. Çevrimiçi veri toplamanın tercih edilmesinin temel nedenlerinden biri, hedef kitlenin çevrimiçi anketlere uyum sağlama yeteneğidir. Bu yöntem, katılımcılara zaman ve mekan açısından büyük bir esneklik sunarak, anketin istedikleri zaman ve mekanda tamamlanabilmesine olanak tanır. Ayrıca, internet üzerinden yapılan anketler, coğrafi sınırlamalardan bağımsız olarak geniş bir kitleye ulaşma imkanı sağlar ve bu, veri toplama sürecini daha kapsayıcı hale getirir. Maliyet açısından, çevrimiçi anketler geleneksel yöntemlere göre daha tasarruflu olup, yazılı materyal ve postalama gibi ek giderleri ortadan kaldırır. Veri toplama ve analiz süreci de hızlanır; dijital formatta toplanan veriler anlık olarak analiz edilebilir. Katılımcılar, anonim olarak yanıt verebildikleri için daha dürüst ve objektif cevaplar sağlayabilirler. Ayrıca, çevrimiçi platformlar üzerinden anket içeriğinde yapılacak güncellemeler ve değişiklikler hızlı bir şekilde uygulanabilir, bu da veri toplama sürecinin esnekliğini artırır. Bu nedenlerle, araştırmanın veri toplama sürecinde çevrimiçi yöntem tercih edilmiştir. Veri toplama süreci yaklaşık olarak iki hafta sürmüştür.

Veri Analizi

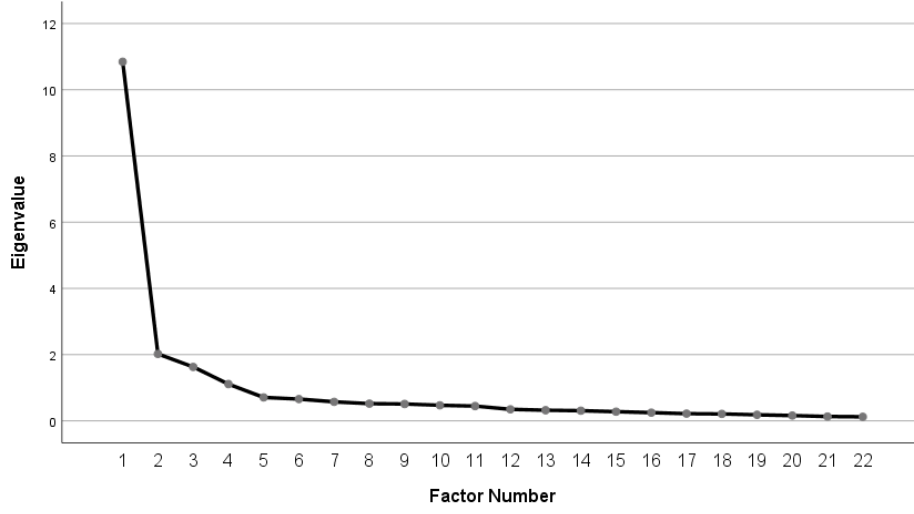
Veri analizinde ilk olarak betimsel istatistik ölçüleri kullanılmıştır. Daha sonra ölçme aracından elde edilen ölçümlerin kapsam geçerliği için kapsam geçerlik oranı kullanılmıştır. Ölçme aracından elde edilen ölçümlerin yapı geçerliğine kanıt sağlamak amacıyla açımlayıcı faktör analizi yapılmıştır. Ölçme aracından elde edilen ölçümlerin güvenirliğine kanıt sağlamak amacıyla da Cronbach alfa ve McDonald omega katsayıları kullanılmıştır. Son olarak madde-toplam puan korelasyon katsayıları ile maddelerin ayırt edicilikleri incelenmiştir.

Geliştirilen ölçme aracının faktör yapısını faktör analitik yöntemler ile incelemeye önce karşılanması gereken bazı varsayımlar test edilmiştir. İlk olarak geliştirilen ölçeğin maddelerinin bir faktör altında toplanabilir olup olmadığını belirlemek için Kaiser-Meyer-Olkin (KMO) ve Bartlett küresellik testi yapılmıştır. Yapılan analizlere göre ölçek maddelerinin faktörleşebilir yapıda olduğu tespit edilmiştir (KMO = .942; Bartlett testi χ^2 (sd) = 5753.604 (276); $p < .001$). Bu bulgulara göre ölçek maddelerinin faktörleşebilir yapıda olduğu görülmektedir. Çünkü KMO değeri .50 ve üzeri ayrıca Bartlett testi de istatistiksel olarak anlamlı olduğu tespit edilmiştir (Field, 2009; Büyüköztürk, 2012). Faktör sayısına karar vermek için faktör özdeğerleri, saçılma grafiği (Scree Plot), paralel analiz, uzman görüşü ve maddelerin içerikleri dikkate alınmıştır. Her bir maddenin ilgili faktörle ilişkisinin bir ölçüsü olarak ele alınan faktör yükü sınır değeri olarak 0.40'ın alınması uygun görülmüştür (Tabachnick ve Fidell, 2013). Madde-toplam korelasyon katsayısı değerinin ise .20 ve üzerinde değer alması maddenin ölçeğin geneliyle uyumlu şekilde çalıştığına işaret etmektedir (Crocker ve Algina, 2006). Son olarak yeterli örneklem büyüklüğü için madde başı en az beş katılımcının olması gerektiği belirtildiğinden (24 madde x 5 = 120 katılımcı) bu varsayımın da sağlandığı tespit edilmiştir. Ayrıca karmaşık olmayan modellerin test edilmesi için örneklem büyüklüğünün 200 olması yeterli olarak görülmektedir (Brown, 2015; Bentler & Chou, 1987). Veri analizinde SPSS (versiyon 25) ve JASP (versiyon 0.19.1.0) paket programları kullanılmıştır.

Bulgular

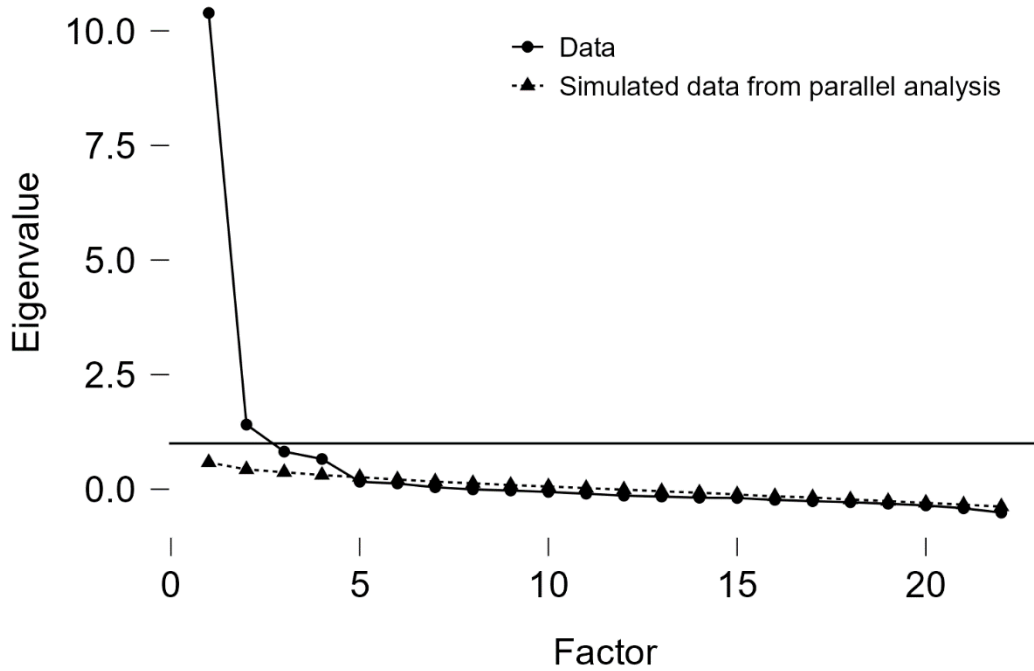
Araştırmanın amacı doğrultusunda geliştirilen ölçme aracından elde edilen ölçümlerin yapı geçerliğine kanıt sağlamak amacıyla ilk olarak açımlayıcı faktör analizi gerçekleştirilmiştir. Yapılan açımlayıcı faktör analizine göre faktör özdeğerleri 1.00'den büyük olan dört faktör olduğu ve toplam varyansın %62.39'unu açıkladığı tespit edilmiştir. İlk faktör analizinde iki maddenin (m2 ve m20) hiçbir faktör altında yük vermediği ve .40 yük değerinden düşük olduğu

belirlendiğinden sırayla her iki madde çıkartılarak analizler tekrarlanmıştır. İlk analize ilişkin saçılma grafiği şekil 1’de verilmiştir.



Şekil 1. Geliştirilen ölçme aracının faktör yapısına ait saçılma grafiği

Şekil 1 incelendiğinde dördüncü faktörden sonra çizginin düzleştiği görülmektedir. Buna göre hem saçılma grafiği hem de kuramsal yapı göz önüne alınarak paralel analiz yapılmış ve faktör sayına karar verilmiştir. Paralel analize göre yapının dört faktörlü olduğuna karar verilmiştir. Paralel analiz sonucu elde edilen saçılma grafiği şekil 2’de verilmiştir.



Şekil 2. Paralel analiz sonucu elde edilen saçılma grafiği.

Şekil 2 incelendiğinde gerçek veri doğrusu ile paralel analiz sonucu oluşturulan veri doğrusunun kesiştiği noktanın üzerinde dört faktör olduğu görülmektedir. Bu sonuç ölçme aracının dört faktörlü bir yapıya sahip olduğunu göstermektedir. Faktör sayısına karar verildikten sonra maddelerin faktör yük değerlerine bakılmış ve bir maddenin (m12) düşük faktör yüküne sahip oldukları ve üç maddenin (m4, m13 ve m18) ise binişik faktör yüküne sahip oldukları tespit edilmiş ve ölçme aracından çıkartılmıştır. Her madde sırasıyla çıkartılmış ve analizler tekrarlanmıştır. Sonuç olarak 18 madde ve dört faktörden oluşan bir ölçme aracı elde edilmiştir. Ölçme aracının son haline ait AFA sonuçları tablo 3’de verilmiştir.

Tablo 3. AFA analizi

	Faktör 1		Faktör 2		Faktör 3		Faktör 4	
	No	λ	No	λ	No	λ	No	λ
	M19	.662	M06	.979	M11	.503	M01	.698
	M21	.680	M07	.894	M14	.654	M03	.710
	M22	.766	M08	.765	M15	.896	M05	.567
	M23	.705	M09	.558	M16	.959		
	M24	.681	M10	.602	M17	.871		
Özdeğer	8.668		2.011		1.558		1.056	
Açıklanan varyans (%)	46.545		9.130		6.196		4.508	
Toplam açıklanan varyans (%)					66.378			

Tablo 3 incelendiğinde, dört faktörün toplam açıkladıkları varyansın yaklaşık olarak %66 olduğu görülmektedir. Buna göre öğretmenlerin yapay zekaya yönelik önyargı düzeylerindeki değişimin %66’sının geliştirilen ölçme aracı ile açıklandığı tespit edilmiştir. Alt faktörler incelendiğinde birinci faktörün en fazla dördüncü faktörün ise en az varyans değişimine sahip olduğu tespit edilmiştir. Maddelerin faktör yük değerleri incelendiğinde, .503 ile 0.979 aralığında değiştiği ve minimum olan .40 değerinden yüksek oldukları tespit edilmiştir. Faktörleri oluşturan maddeler incelendiğinde, faktörler sırasıyla değişen insan-makine ilişkisi, toplumsal ve bireysel riskler, etik kaygılar ve mesleki otomasyon riski şeklinde isimlendirilmiştir. Sonuç olarak ölçme aracından elde edilen ölçümlerin yapı geçerliğine ilişkin kanıtlar sağlanmış ve ölçümlerin güvenilirliği için Cronbach alfa ile McDonald omega katsayıları hesaplanmış ve tablo 4’de verilmiştir.

Tablo 4. Ölçme aracından elde edilen ölçümlere ait güvenilirlik değerleri

Faktör	Madde sayısı	McDonald ω (%95 GA)	Cronbach α (%95 GA)
Değişen insan-makine ilişkisi	5	.916 (.901-.931)	.912 (.896-.927)
Toplumsal ve bireysel riskler	5	.917 (.902-.931)	.916 (.900-.930)
Etik kaygılar	5	.901 (.884-.918)	.900 (.881-.917)
Mesleki otomasyon riski	3	.691 (.632-.750)	.687 (.623-.742)
Ölçeğin bütünü	18	.931 (.920-.942)	.928 (.917-.939)

Tablo 4 incelendiğinde, mesleki otomasyon riski dışındaki alt faktörlerin ve ölçeğin bütünü için hesaplanan güvenirlik değerlerinin .70 ve üstü olduğu görülmektedir. Buna göre ölçeğe aracından elde edilen ölçümlerin yüksek güvenirliğe işaret ettiği belirlenmiştir (Salvucci vd., 1997). Ölçeğe aracından elde edilen ölçümlerin geçerlik ve güvenirliklerine kanıt sağlandıktan sonra, madde-toplam puan korelasyon katsayıları incelenmiş ve elde edilen bulgular tablo 5’de verilmiştir.

Tablo 5. Madde-toplam puan korelasyonları

Madde No	Madde Ortalaması ($\bar{X}$)	Düzeltilmiş Madde-Toplam Korelasyonu (r)	Madde No	Madde Ortalaması ($\bar{X}$)	Düzeltilmiş Madde-Toplam Korelasyonu (r)
M01	2.241	.213	M14	4.009	.705
M03	2.869	.391	M15	3.438	.661
M05	2.300	.265	M16	3.034	.597
M06	4.163	.671	M17	3.450	.682
M07	4.434	.762	M19	4.872	.528
M08	4.291	.777	M21	4.488	.708
M09	3.650	.696	M22	4.766	.699
M10	4.672	.768	M23	4.522	.746
M11	3.741	.712	M24	4.691	.717
Ölçeğin Geneli İçin Madde Ortalaması				:	3.868
Ölçeğin Geneli Madde-Toplam Korelasyon değerleri				:	.628

Tablo 4 incelendiğinde, ölçeğe aracındaki maddelerin ortalamaları 2.241 ile 4.872 aralığında değiştiği görülmektedir. Ölçeğin geneli için ortalama değer ise 3.868 düzeyinde olduğu ve önyargı düzeylerinin ortalama üstü olduğu tespit edilmiştir. Madde-toplam puan korelasyon değerleri incelendiğinde ise .213 ile .777 aralığında değiştiği ve maddelerin toplam puan ile orta ve yüksek düzeyde ilişkiye sahip oldukları başka bir değişle ayırt edicilik düzeylerinin yüksek olduğu tespit edilmiştir.

Tartışma ve Sonuç

....

Kaynakça

- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi.
- Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Belmont: Wadsworth Pub Co.
- Field, A. (2009) *Discovering Statistics Using SPSS*. London: Sage Publications Ltd.

- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Salvucci, S., Walter, E., Conley, V., Fink, S., & Saba, M. (1997). *Measurement Error Studies at the National Center for Education Statistics*.
- Şata, M. (2020). Nicel araştırma yaklaşımları. E. Oğuz. (Ed.), *Eğitimde araştırma yöntemleri* (1.baskı) içinde (s. 77-90). Eğiten Kitap Yayınları.
- Tabachnick, B. G. & Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson.
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197-210. <https://doi.org/10.1177/0748175612440286>