



Ölçme Değerlendirme Metinlerinin Seçimi ve Yazımına Yönelik Bir Araç: Okur Dostu Bilgilendirici Metin Dereceli Puanlama Anahtarı (ODBİMDEPA) *

Dilek Fidan ¹, Tuğba Çelik Korat ², Sercan Demirgüneş ³, Seçil Alaca ⁴, Nida Kırömeroğlu ⁵

Öz

Konusu, metin değerlendirmede kullanılan dereceli puanlama anahtarları (rubrik, rubric) olan bu çalışmada, hem eğitim ortamlarında hem de farklı sınavlar yapan kurumların (MEB, ÖSYM, Yunus Emre Enstitüsü/YEE gibi) sınavlarında kullanılan bilgilendirici metinlerin okur dostu olup olmadıklarının belirlenebilmesi için geliştirilen ve güvenilirlik çalışması yapılan *Okur Dostu Bilgilendirici Metin Dereceli Puanlama Anahtarı*'nın (ODBİMDEPA) alanyazına tanıtılması ve incelenen metinlerle ilgili gözlemlerin aktarılması amaçlanmıştır. Bu amaç doğrultusunda, çalışmada belge (doküman) inceleme yöntemi kullanılarak iki farklı veri tabanı oluşturulmuştur: Birinci aşamada kullanılan birinci belge seti, MEB tarafından 2018 yılında okutulan 5, 6, 7 ve 8.sınıf ders kitaplarındaki toplam 50 bilgilendirici metin; ikinci aşamada kullanılan ikinci belge seti, ÖSYM tarafından 2023 ve 2024 yıllarında yapılan YKS testlerindeki sorularda geçen toplam 37 bilgilendirici metinden oluşmaktadır. İlk belge setini birbirinden bağımsız 2 akademisyen; ikinci belge setini 4 alan için 11 uzman, 1 alan için ise 13 uzman değerlendirmiştir. Çözümlemelerde, çalışmanın birinci aşamasındaki puanların ağırlıklı kapa katsayısına bakılmış ve ağırlıklı ortalamasının kabul edilebilir olduğu ve ODBİMDEPA'nın güvenilirlik açısından anlamlı (*genellikle uyuşma*) olduğu sonucuna ulaşılmıştır. Çalışmanın ikinci aşamasında veri tabanı çok yüzeyli Rasch modeliyle çözümlenmiş ve değerlendirici güvenilirliği 0.95 bulunmuştur. Bu sonuç, geliştirilen puanlama anahtarının güvenilirliğinin oldukça yüksek olduğunu göstermektedir. Bu bulgular ışığında ODBİMDEPA kullanılarak MEB, ÖSYM gibi sınav uygulamaları yapan kurumların sınavlarında kullanılan ya da eğitim ortamlarında okutulan kitaplarda yer alan bilgilendirici metinlerin ne kadar okur dostu olduğu sayısal olarak ifade edilebilecek ve metinlerin sınavlarda ya da derslerde kullanılıp kullanılmamasına nicel olarak da karar verilebilecektir.

Anahtar Kelimeler

Ölçme ve değerlendirme
Bilgilendirici metin
Dereceli puanlama anahtarı
Okur dostu metin
ODBİMDEPA

Makale Hakkında

Gönderim Tarihi: 13.10.2024
Kabul Tarihi: 02.01.2025
Elektronik Yayın Tarihi: 03.03.2025

DOI: 10.15390/EB.2025.14166

* Bu çalışmanın bir bölümü 4-6 Ekim 2024 tarihleri arasında düzenlenen Uluslararası Ölçme, Seçme ve Yerleştirme Sempozyumu'nda sözlü bildiri olarak sunulmuştur.

¹ Kocaeli Üniversitesi, Eğitim Fakültesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Türkiye, dilek.fidan@kocaeli.edu.tr

² Bahçeşehir Üniversitesi, Eğitim Bilimleri Fakültesi, Temel Eğitim Bölümü, Türkiye, tugba.celikkorat@bau.edu.tr

³ Niğde Ömer Halisdemir Üniversitesi, Eğitim Fakültesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Türkiye, sdemirgunes@ohu.edu.tr

⁴ Hacettepe Üniversitesi, Türkçe ve Yabancı Dil Öğretimi Uygulama ve Araştırma Merkezi, Türkiye, secilalaca@hacettepe.edu.tr

⁵ İstanbul Üniversitesi, Yabancı Diller Yüksekokulu, Türkiye, nidakiromeroglu@gmail.com

Giriş

Bir metnin okur dostu (friendly-text/reader-friendly) olması, onun alıcı tarafından daha kolay işlemlendiği ve anlamlandırıldığı anlamına gelmektedir. “Okur dostu” kavramının alanyazındaki temellerinin, ilk defa Gray ve Leary (1935) ile birlikte atıldığı görülmektedir. Söz konusu çalışmada, bir kitabı okunabilir kılan dört faktörden söz edilmektedir. Bu faktörler; kitabın *içeriğini oluşturan etkenler*, kitabın *biçemi* (style), kitabın *biçimi/formatı* ve kitabın *genel düzenidir*. Metinlerin türleri ve türe özgü metinsel düzenlemelerin var olup olmaması da yukarıda bahsedilen alımlama ve işleme süreçlerinde etkilidir. Brewer (1980), Smith (1982) gibi çalışmalarda, metinlerin temelde, “anlatsal” ve “bilgilendirici” metinler olarak ikiye ayrıldığı görülmektedir. Çelik, Demirgüneş, Baştuğ, (2014), bilgilendirici ve anlatsal metinlerin biçimsel ve anlamsal niteliğini belirlemek üzere *Okur Dostu Metin Değerlendirme Rubriği* (ODMDR) geliştirmiştir. Dreher ve Singer (1989) tarafından “okur dostu” metinlerde bulunması gereken özellikler açıklanmıştır; ODMDR de bu çalışmadan yola çıkarak yapılandırılmış ve bu puanlama anahtarında yedi boyut tasarlamışlardır: *Metin düzenleme, zaman kullanımı, tutarlık-bağdaşıklık, ipucu verme, kavramsal yoğunluk, açıklama ve söylem*. Daha sonra, alanyazında ODMDR kullanılarak ders kitabı metinlerinin öğrenciler için uygunluğunu çeşitli açılardan değerlendiren çalışmaların (Aktaş ve Alıcı, 2018; Çelik, Demirgüneş ve Fidan, 2015; İşeri, Çelik ve Demirgüneş, 2015; Soyuçok ve Kartal, 2018; Yıldırım ve Ocak, 2016) yapıldığı görülmektedir. Demirgüneş (2015) her metin türünün farklı özellikler taşıdığını, ODMDR’nin boyutlarının her metin türüne bütünüyle uygun olamayacağını, bu nedenle puanlama anahtarının farklı metin türlerine göre geliştirilmesi gerektiğini önermiştir (s. 67). Müldür ve Şimşek’e (2020) göre alanyazın çalışmaları, eğitim ortamlarında kullanılan bilgilendirici metinlerin nicel ve nitel açıdan anlatsal metinler karşısında zayıf kaldığını ortaya koymaktadır. Bu verilerden yola çıkılarak ODMDR’nin, tüm metin türlerinin okur dostu olma/olmama düzeylerinin belirlenmesinde tek başına yeterli olmadığı; metin türüne göre özelleştirilerek, puanlama anahtarının geliştirilmesi gerektiği ortaya çıkmıştır.

Alanyazında metin sınıflamalarının çok sayıda olmasına bağlı olarak bilgi verici metinlerin alt tiplerinde de çeşitlenmeler görülmektedir. Duke ve Bennett Armistead (2003) bilgilendirici metni; salt bilgi verme amaçlı, karakterler ve olay örgüsü içermeyen; içinde indeks, fotoğraflar, dipnotlar, grafikler ve çizelgeler olabilen metinler olarak ve kurmaca olmayan metinlerin bir alt ulamı olarak tanımlamaktadır. Bilgi verme amacıyla üretilen metin tipleri *bilgilendirici* (informative), *açıklayıcı* (expository), *örneklendirici* (explanatory) ve *tartışmacı* (argumentative) olarak sınıflandırılmaktadır. Bilgilendirici metinler, herhangi bir yoruma ve değerlendirmeye yer vermeksizin ansiklopedik bilgilerin aktarıldığı metinlerdir. Açıklayıcı metinlerde ana bilgi, net bir biçimde düzenlenmiş metin yapısıyla “açıklama, ayrıntılandırma” ve “sebep-sonuç ilişkileri” ile sunulur. Metin, bölümlere ayrılır ve ana bilgiye özgü terimleri içerir. Çalışmalar (Beck, McKeown, Sinatra ve Loxterman, 1991; Britton ve Gulgoz, 1991; McNamara, 2001), metinlerde tutarlık ilişkilerinin iyi yapılandırılması; diğer bir ifadeyle metinlerde mantık hatalarının olmaması halinde bu durumun, metinlerdeki bilgilerin bellekte kalmasını ve anlamlandırılmasını olumlu yönde etkilediğini belirtmektedir. Örneklendirici metinlerde, temelde “*Nasıl?*” (Örn. “Musluk nasıl çalışır? Bilgisayar nasıl çalışır? Dağlar nasıl oluşmuştur?” gibi) ve “*Neden?*” (Örn. “Bazı cisimler neden batar ya da suyun üstünde kalır? Ozon tabakası neden incelmektedir?” gibi) sorularının yanıtları bulunur. Tartışmacı metinler, “tartışılacak konunun belirlenmesi” ile başlar ve “tartışma bölümü, ana bölüm, örnekler/savlar/görüşler” ile devam eder ve “kesin olmayan bir yanıt” ile sonlandırılır.

Millî Eğitim Bakanlığı ders kitaplarında metin türleri üçe (öyküleyici, bilgilendirici, şiir) ayrılmıştır ve Türkçe ders kitaplarında metinler bu ayrımla ele alınmıştır (MEB, 2024). Bu bağlamda, bu çalışmada tasarlanan dereceli puanlama anahtarı için, alanyazındaki dört farklı “bilgi verme amaçlı metin tipi” (bilgilendirici, açıklayıcı, örneklendirici ve tartışmacı), “bilgilendirici metin türü” adı altında birleştirilmiştir. Böylelikle bu çalışmada geliştirilmesi amaçlanan dereceli puanlama anahtarının adı “*Okur Dostu Bilgilendirici Metin Dereceli Puanlama Anahtarı* (ODBİMDEPA)” olarak belirlenmiştir. Bu alt türlerdeki zorunlu terimsel birleşim, sözü edilen dört tip metnin içerik yapısını, dereceli puanlama anahtarına yansıtmayı da beraberinde getirmiştir. Tasarlanan dereceli puanlama anahtarının öncelikli hedefi, eğitim ortamları ve eğitim öğretim sürecinin uygulama öğeleri (öğretmenler, merkezi sınav komisyonları ve öğrenciler vd.) için kullanılabilir/eğitsel amaçlı bir ölçme aracı oluşturmaktır.

Bilgilendirici metin türünün, ODBİMDEPA kapsamındaki boyutlarının ölçülebilir olması amacıyla bilgilendirici metinlerin tamamlayıcısı niteliğindeki “somut” ve “soyut” özellikler belirlenmiş ve Tablo 1’de özetlenerek kaynaklarıyla birlikte verilmiştir.

Tablo 1. Bilgilendirici metinlerin somut ve soyut özellikleri

Temel Nitelik	Tipik özellikler	Kaynak
Somut özellikler	İşaretleyiciler	Lorch, Lemarie ve Grant (2011)
	Başlıklar, özetler, önsöz, göstergeler, içerik tabloları, ajandalar (bilgi sunumunun tarihsel düzenleyicileri) ...	Fisher, Frey ve Lapp (2008)
	Evrensel nitelikler	Klink, Dengel ve Kieninger (2000)
	Sayılar, satır aralıkları, kenar boşlukları	Fisher ve diğerleri (2008)
Soyut özellikler	İlgi çekici içerikler	Fisher ve diğerleri (2008)
	Altyazılar, dipnotlar, yay ayaç açıklayıcıları, diyagramlar, grafikler ve diğer açıklayıcı görseller...	Fisher ve diğerleri (2008)
	Tipografik özellikler	McNamara, Kintsch, Songer ve Kintch (1996)
	Font bilgileri (yazı stili, punto...vs.)	Yeung (2007)
Soyut özellikler	Yapı ve Retorik	Aidinlou, Khodamard ve Azami (2012)
	Tutarlılık	Mann, Thompson (1988)
		McNamara, Louwerse, McCarthy ve Graesser (2010)
	Sözlüksel ve dilbilgisel içerik	Mann ve Thompson (1988)
Anlamsal içerik	Aidinlou ve diğerleri (2012)	
Sözdizimsel içerik	McNamara ve diğerleri (2010)	

Tablo 1 incelendiğinde, bir metnin bilgilendirici olarak kabul edilebilmesinde salt bilginin yeterli olmadığı, bu bilginin sunulduğunda ana bilginin anlamlandırılmasına destek olacak başlık, özet, grafik/ler, yazı stilleri, yazı büyüklükleri vb. gibi somut unsurların, soyut unsurlara (yapı ve retorik düzenlemeler, tutarlılık, sözlüksel ve dilbilgisel içerik) destek olması gerektiği görülmektedir. Bu çalışmada, Tablo 1’deki alanyazından hareketle ve ODMDR (2014) boyutlarını bilgilendirici metin özelinde güncelleyerek eğitim ortamlarında kullanılmak üzere yapılandırılmış, sadece bilgilendirici metinleri değerlendirmek üzere 8 boyut belirlenmiştir: *Metin düzenleme, Bilginin sunumu, Biçimsel örüntüler/Başlaşıklık, Anlamsal örüntüler/Tutarlılık, Açıklama ve gerçeklik, Tartışma, Biçim ve Evrensellik.*

Çalışmanın Önemi ve Amacı:

Alanyazında metin sınıflandırma çalışmalarının çeşitliliğine rağmen, eğitim ortamlarında kullanılan metinlerin niteliğini belirlemek üzere tasarlanmış dereceli puanlama anahtarları üzerine yapılan çalışma (Çelik, Demirgüneş ve Baştuğ, 2014; Çıralı Sarıca ve Koçak Usluel, 2016) sayısı sınırlıdır. 2014’te hem anlatsal hem bilgilendirici metinleri değerlendirmek üzere tasarlanmış olan ODMDR’nin (Çelik, Demirgüneş ve Baştuğ, 2014) tüm metin türleri için geliştirilmiş olması ve bu nedenle anlatsal ve bilgilendirici metin tiplerinin köklü farklılıklarını karşılayamaması (anlatsal metinlerde bilginin daha az olması ancak bu metinlerde söylemin daha güçlü olması...vb.) salt bilgilendirici metinler için yeni ölçme araçlarına ihtiyaç olduğunu ortaya çıkarmıştır. Bu nedenle bu çalışmada, eğitim ortamlarında kullanılmak üzere yazılan/yazılmış bilgilendirici metinlerin ya da özgün metinlerden uyarlanan bilgilendirici metinlerin biçimsel ve anlamsal yeterliklerini ölçmek üzere ODBİMDEPA geliştirilmiştir. Bu sayede; MEB, ÖSYM gibi kurumların sınavlarında kullanılan ya da eğitim

ortamlarında okutulan ders kitaplarında yer alan bilgilendirici metinlerin, ne kadar okur dostu olduğu sayısal olarak ifade edilebilecek ve bu metinlerin sınavda ya da derste kullanılmasının uygun olup olmadığına dair nicel verilere dayalı nesnel kararlar verilebilecektir. Bu bağlamda, çalışmanın temel araştırma sorusu, ODBİMDEPA'nın okur dostu bilgilendirici metinleri değerlendirmek için geçerli ve güvenilir bir araç olup olmadığının araştırılmasıdır. Bu amaç doğrultusunda, üç aşamalı bir çalışma gerçekleştirilmiştir: İlk aşamada, bir dereceli puanlama aracı geliştirilmiştir; ikinci aşamada ise iki alan uzmanı akademisyen tarafından MEB ders kitaplarında yer alan metinlerin puanlanması ile puanlama anahtarının güvenilirliği üzerinde çalışılmıştır. Çalışmanın üçüncü ve son aşamasında ise özellikle ölçme ve değerlendirme amacıyla kullanılan metinlerin okur dostu olması gerektiği düşünülerek ÖSYM tarafından yapılan YKS sınavlarında kullanılan toplam 37 bilgilendirici metnin, okur dostu olup olmadığının betimlenmesi için geliştirilen dereceli puanlama anahtarı kullanılmıştır. Okuduğunu kavrama üzerine yapılan çalışmalar (Kintsch, 1982; Kintsch, 1988; Dreher ve Singer, 1989; Beck vd., 1991), okuduğunu anlamının yalnızca önermelerin ve bilgilerin sunumu, sözdizimi gibi özelliklerle ilgili olmadığını, söylemin de süreçte etkisi olduğunu öne sürmektedir. Okuduğunu anlama yeterliklerini ölçen bir sınavda yer alan metinlerin öncelikle okunabilirliği yüksek, okur dostu olması gerekmektedir. Böylece, metne bağlı maddeler geliştirilebilmekte ve adayların sınav performansına ilişkin anlamlı sonuçlara ulaşılabilmektedir. Metnin okur dostu olmadığı bir ölçme aracında, maddelerin de anlamlı bir ölçme yapması beklenemez; çünkü adaylar okuduğunu kavrama ile ilgili bilişsel süreçlerde metinden kaynaklı sorunlar yaşayabilir. Bu nedenle çalışmanın son aşamasındaki temel amaç, bilgilendirici güvenilirliğinden daha çok metinlerin okur dostu olup olmadığının belirlenmesi olmuştur.

Yöntem

Araştırmanın Türü

Bilgilendirici metinlerin okur dostu olup olmadıklarını belirlemek amacıyla bir dereceli puanlama anahtarı geliştirilmesini ve ölçme aracı olarak kullanılan metinlerin okur dostu olup olmadığının belirlenmesini amaçlayan bu çalışma, yukarıda da belirtildiği gibi üç aşamalı olarak planlanmıştır: İlk aşamada dereceli puanlama anahtarının ölçütleri, puanları ve bileşenleri belirlenmiştir. İkinci aşamada, belirtilen amaca ulaşmak için geliştirilen dereceli puanlama anahtarının güvenilirliği test edilmiştir. Üçüncü aşamada ise geliştirilen dereceli puanlama anahtarı kullanılarak, veri tabanındaki toplam 37 bilgilendirici metnin okur dostu olup olmadıkları betimlenmiştir. Dereceli puanlama aracının geçerlik ve güvenilirliğinin test edilmesinde kullanılan MEB kitaplarındaki toplam 50 bilgilendirici metinden ve YKS'de çıkan sorulardaki toplam 37 bilgilendirici metinden oluşan veri tabanının oluşturulmasında, nitel araştırma yaklaşımlarından belge incelemesi yöntemine baş vurulmuştur. Yıldırım ve Şimşek'e (2008) göre bu yöntem, araştırılması hedeflenen olgu veya olgular hakkında bilgi içeren yazılı materyallerin analizini kapsar (s.187).

Bu yöntem kapsamında çalışmada, iki farklı belge seti hazırlanmıştır: İlk set, 2018'de MEB tarafından okutulan "Fen Bilimleri 4, 5, 6, 7, ve 8. sınıf ders kitapları", "Sosyal Bilgiler 4, 5, 6, ve 7. sınıf ders kitapları", "İnsan Hakları, Yurttaşlık ve Demokrasi 8. sınıf ders kitabı" ve "Türkçe 4, 5, 6, ve 7. sınıf ders kitapları"ndan oluşmaktadır. Bu kitaplar içinde yer alan metinler gözden geçirilmiş ve toplam 308 bilgilendirici metinden oluşan bir evrene ulaşılmıştır. Daha sonra bu evrenden, SPSS (2015) ile rastlantısal olarak 50 metin örneklem (evrenin %16.3'ü) olarak seçilmiştir.

İkinci belge seti, 2023 ve 2024 yıllarında ÖSYM tarafından yapılan ve ÖSYM'nin resmî web sitesinde yayınlanan YKS-Alan Yeterlilik Testi'ndeki ve YKS-Temel Yeterlilik Testi'ndeki toplam 528 soruluk evrenden, bilgilendirici metin türündeki paragrafların olduğu ve bu paragraflarla ilgili ayrıntıların "Sınırlılıklar" kısmında anlatıldığı kriterler göz önünde bulundurularak elde edilen toplam 37 sorudaki bilgilendirici metin seçilmiş ve örneklem olarak kaydedilmiştir. Bunun sonucunda çalışmanın kriterlerini karşılayan 10 Türk Dili ve Edebiyatı, 10 Felsefe, 7 Din kültürü ve Ahlak Bilgisi, 6 Sosyoloji ve 4 Tarih dersi sorusu metni örnekleme dahil edilmiştir.

Değerlendiricilerin Seçimi

ODBİMDEPA'nın geçerlik ve güvenilirlik çalışması iki farklı çalışma grubu ve iki farklı belge setiyle gerçekleştirilmiştir: İlk olarak, dereceli puanlama anahtarının işlerliğini görmek için ilk belge setini, iki akademisyen değerlendirici (her biri Türkçe eğitimi alanında doçentliğini almış ve en az 15 yıllık iş tecrübesi olan) değerlendirmiştir. Daha sonra, metinlerin okur dostu olup olmadığını belirlemek üzere 4 alan için 11 kişi, 1 alan için ise 13 kişi, ODBİMDEPA'ya göre ikinci belge setini değerlendirmiştir. Bu 13 kişinin tümü, dil öğretiminde lisansüstü dereceye sahip olup dereceli puanlama anahtarı değerlendirme deneyimine ve minimum 5 yıllık iş tecrübesine sahip kişilerden seçilmiştir.

Veri Toplama Araçları

Çalışmada, bilgilendirici metinlerin okur dostu olup olmadıklarının belirlenmesi amacıyla araştırmacılar tarafından geliştirilen ve temelini; Çelik, Demirgüneş ve Baştuğ (2014) tarafından geliştirilen ODMDR'den alan *Okur Dostu Bilgilendirici Metin Dereceli Puanlama Anahtarı/ODBİMDEPA*, veri toplama aracı olarak kullanılmıştır. Bu araç, toplam 8 boyuttan oluşmaktadır. Bu boyutlar sırasıyla "*Metin düzenleme, Bilginin sunumu, Biçimsel örüntüler (Bağlaşıklık), Anlamsal örüntüler (Tutarlık), Açıklama ve gerçeklik, Tartışma, Biçem, Evrensellik*" boyutlarıdır.

ODBİMDEPA'daki her bir boyutla ilgili açıklama, aşağıda verilmiş olup EK1'de, geliştirilen dereceli puanlama anahtarının güncel biçimi sunulmuştur. Dereceli puanlama anahtarı kullanılarak yapılacak puanlamalarda, aşağıdaki belirlemelere dikkat edilmesi gerekmektedir:

1.Boyut: Metin Düzenleme

Bilgilendirici metinlerde birden çok bilgi/düşünce yer alabilmektedir. Metindeki tüm bu bilgiler/düşünceler birbiriyle çelişmemeli, birbirini desteklemelidir. Çelişkili ifadeler ve bakış açıları, art arda sıralanan birbiriyle bağlantısı olmayan bilgi yığınları, okurun metinden uzaklaşmasına neden olur. Diğer bir ifadeyle, okurun metni anlamasına engel olur. Metnin genelinde ana bilginin/düşüncenin ağırlığı hissedilmeli, daha önce anılan diğer her bilgi ve düşünce onu destekler nitelikte olmalıdır. Ana bilgi/düşünce metne yayılamıyor, diğer bilgiler ya da düşünceler onun önüne geçiyorsa metin, asıl hedefinden ayrılmış olmaktadır. Birden çok bilgi, görüş, bakış açısı birleşerek bir ana düşünceyi oluşturabilmelidir.

2.Boyut: Bilginin Sunumu

Bilgilendirici metinlerde bilgi iki yolla sunulur. Birincisinde, sadece "**yalnız bir ana bilgi/düşünce**" iletilir ve ayrıntılandırma, karşılaştırma, açıklama...vb. çeşitli düşünceyi geliştirme yolları/tutarlık ilişkileri ile ana bilginin okurda kavramsallaştırılması sağlanır. İkinci yolda, "**bir bilgi/düşünce demeti**"nin sunumundan söz edilir. Bu yolda yazar, ana bilgiyi bir biçimde önceleyerek baskın kılar ve diğer bilgiler, ana bilginin destekleyicisi konumuna yerleşir. Bilginin sunumunda mutlaka ana bilginin/destekleyici bilgilerin varlığına ilişkin somutlaştırmalar, açıklamalar yapılmalıdır. Ayrıca okurun anlamasını kolaylaştıracak biçimde bilgilerin sınıflanması/gruplandırılması yoluna gidilmeli, bilgiler genelden-özele, basitten-karmaşığa gidecek biçimde aşamalı olarak düzenlenmelidir.

3.Boyut: Biçimsel Örüntüler (Bağlaşıklık)

Bilgilendirici metinlerin, biçimsel olarak diğer metin türlerinden farklılaşmalarını sağlayan ayrırcı tipik özellikleri vardır. Öncelikle metin; belirgin biçimde ara başlıklara ayrılmamış olsa bile, iletilmek istenen bilgiye ilişkin kuramsal/kavramsal bilgilerin sunulduğu bir bölüm, bu bilginin geliştirildiği/tartışıldığı bir diğer bölüm ve sonunda metin içinde anılan soru(n)ların çözümüne ilişkin ipuçları yer almalıdır.

Bu bölümler içinde ana bilgi ve ana bilgiyi destekleyecek yapılar; yazı karakteri, punto... vs. kullanımı ile okura ayırt edilebilir kılınmalıdır. Bunun yanı sıra, bilgiyi açıklayıcı her türlü görsele (tablolar, grafikler... vs.) yer verilebilir. Dilin tipolojik olanakları el verdiği biçimde, dile özgü ön ve art gönderimler, eksiltmeler, adil düşürmeler, yinelemeler kullanılabilir.

Burada asıl amaç, okurun ana bilgiyi kavramasına yönelik her türlü dilsel göstereni işe koşturmak. Bilgilendirici metinlerin özetlenmesi, anlamlandırılması ve kavranması diğer metin türlerine göre daha güç olduğundan, okura metni anımsatacak anahtar sözcüklerin çeşitli biçimlerde belirginleştirilmesi, metni daha kolay anlaşılır ve okur dostu kılacaktır.

4.Boyut: Anlamsal Örüntüler (Tutarlık)

Metinde biçimsel bir bütünlüğün yanı sıra, anlamsal bir bütünlük de olmalıdır. Bilginin sunumu eğer “**yalnız bir ana bilgi/düşünce**”nin iletilmesiyle gerçekleştiriliyorsa örneklendirme, karşılaştırma, tanık gösterme, tanımlama...vb. ilişkilerin ana bilgiyi destekleyecek savlarla örüntülenmiş olması gerekmektedir. Bilginin bu yolla sunulmasında, anlamsal olarak bütünlük sağlamak bir bilgi/düşünce demetinin sunumuna göre daha kolaydır. Ancak, ana bilgi, “**bir bilgi/düşünce demeti/öbeğinden**”nden destek alınarak ya da onlarla güçlendirilerek sunuluyorsa daha etkili bir yol kullanılmış demektir. Bu durumda, metinde ana bilgiyi destekleyecek diğer bilgilerin/düşüncelerin kendi arasında uyumlu olması ve ana bilgi ile çelişmemesi gerekmektedir.

5.Boyut: Açıklama ve Gerçeklik

Metinde okura sunulan her türlü bilginin, düşüncenin (okurun bilmediği varsayılan sözcükler başta olmak üzere) metin içinde ya da dışında (dipnot, metin sonu gönderimler, sözlük... vs.) bir açıklaması/tanımlaması olmalıdır. Bu açıklamalar/tanımlamalar, okurun belleğinde yeni belirsizlikler yaratmayacak biçimde somutlaştırma, örnekleme... vb. yollarla verilmelidir. Bilgilendirici metinlerde, bilgiyi/düşünceyi, anlamayı kolaylaştırmanın en etkili yolu, okurun dünyasına ulaşabilmek ve bilgiyi/düşünceyi okurun kendi gerçekliğine/dünyasına kazandırmaktır. Bu nedenle yazar, metindeki ana bilgiyi okurun dünya bilgisiyle bütünleştirmelidir.

6.Boyut: Tartışma

Her bilginin doğuşu bir merak dürtüsü ya da rastlantıya dayalıdır. Öğretim sürecinde rastlantıya yer vermeden bilgiye erişmek, okurda merak uyandırmakla mümkündür. Metin içinde sunulan bilgi ya da düşünce demeti, farklı düzeylerde okurun merakını uyandırabilmektedir. Yazarın görevi, okuru farklı tartışmalara, çelişkili düşüncelere terk etmek değil; metindeki ana bilgiye yönelik bir tartışma yaratmak ve tartışmayı net ya da tek bir çözüme ulaştırmadan okuru meraka sürüklemektir. Bu noktada tartışma için metinde birden çok bakış açısının ve görüşün bulunması gerekmektedir. Bu bağlamda, tartışmanın “**mutlaka tek bir çözüm**”e kavuşması gerekmez, merak duygusunun tetiklenmesi gerekir. Tartışma merak uyandırır; merak da ana bilgiyi kavrama yolundaki en önemli güdüleyici ögedir.

Bu boyutta dikkat edilmesi gereken, okurun farklı bilgileri/düşünceleri kullanarak ana bilgiyi değerlendirmesini sağlamaktır. Öte yandan ana bilginin/düşüncenin dağılması, tartışmanın tek bir çözümle sonlandırılmaya çalışılması, tek çözümün/düşüncenin sergilendiği metinlerin oluşturulması, tartışmanın zayıf kaldığının göstergesidir.

7.Boyut: Biçem

Bütün metin türlerinin temel hedefi, yazar-okur etkileşimini güçlü kurabilmektir. Bilgilendirici metinlerde yazar, okurun güvenini sahip olduğu bilgilerle ve nesnel bakış açısıyla kazanmalıdır. Bilgilendirici metin yazarı, gerçek ve bilimsel bir tabana yaslandığı için nesnel olmak zorundadır. Bu bağlamda, yazar, okurun dünyasına öznel yapıları kullanmadan girmek zorundadır. “Bence”, “kim ne derse desin”, “tartışmaya kapalı”, “en önemli”, “tek koşul” gibi keskin ve nesnel olmayan dilsel yapılar, metne olan güveni zedeler. Bunun aksine yazarın “bir başka görüşe göre” “sayısal bakımdan” “gerçekçi bir tutumla söylersek” “kanıtlar gösteriyor ki” “1000 metreden fazla” “aralarındaki tek göçmen kuş olan” gibi detaylı anlatımlara başvurması; anlaşılır ve yetkin bir dile sahip olması gerekmektedir. Yazar “sonuç olarak” “başta söylediğimiz gibi” “raporlara bakılırsa” gibi özetleyici sözlerle okuru metne bağlayabilmelidir. Örneklerin ilgi çekici olması da okuru metne yaklaştırır. “Bir diğer nokta da (...)”,

“(...) kış mevsimi sıcaklıklarında ise (...)”, “Norveçliler için (...)”, “Osmanlı kültüründe de (...)” gibi dilsel ifadeler, yazarın farklı kültürlerle açık, kolay anlaşılacak isteyen, verilerle kanıt yoluna giden ve karşılaştırma yapabilen bir dile sahip olduğunu göstermektedir.

8.Boyut: Evrensellik

Bilgi evrensel bir olgudur, çeşitli değişkenler karşısında değişiklik göstermez. Uluslararası sözleşmelere uygun ifadeler (İnsan Hakları, Çocuk hakları vd.) kullanmak, bilgilendirici metinde evrensel bir bakış açısı oluşturmaya destek olur. Örneğin yazar, metninde dil-din-cinsiyet vs. ayrımı yapmamalıdır; böyle bir ayrım, okuyucunun metinde gördüğü herhangi bir olumsuz ifade nedeniyle ait olduğu kültür, inanç, cinsiyet durumlarından ötürü kendini ötekileştirilmiş hissetmesine neden olabilir. Ayrıca, yazar, okura farklı kültürlerle, farklı insan davranışlarına ilgi duymayı sağlayacak bir bakış açısı da sergileyebilmelidir. Örneğin metinde, ülke sınırları içinde farklı uzamlara ve burada yaşayan insanların özelliklerine yer verebilir; ömürleri çok uzun olan insanların ülkeleri hakkında bilgi verebilir, farklı hayvanların yaşam haklarını sezdirebilir, çokkültürlülüğün altını çizebilir, coğrafya, botanik gibi konuları ele alırken sürdürülebilirlik ya da çevreci olmayı vurgulayabilir. Evrensel olmayan bilgi/düşünce kurgusu içerisinde, evrensel bir bilgiyi/düşünceyi örüntüleyebilir. Öte yandan yazar metninde, insanlara karşı önyargı yaratacak ifadeler kullanmaktan, herhangi bir ayrımcılığı ya da olumsuz duygu/düşünce/bilgiyi normalleştirmekten, ayrımcılığa yol açacak ifadelerle yer vermekten kaçınmalıdır.

Veri Toplama Aracının Geçerlik ve Güvenirliği

Dereceli puanlama anahtarları için güvenirliliği Tuncel (2011), “değerlendirmeye tabi tutulan bir öğrencinin ödevinin/performansının her değerlendirilme ve her değerlendiren kişiden yine aynı puanı alması” olarak tanımlamaktadır (s. 222). Moskal ve Leydens’e (2000) göre dereceli puanlama anahtarının geçerliği “içerik, yapı ve ölçüt” yönüyle denetlenmelidir. Geçerliği denetlemeye ilişkin bu yönler, ilgili çalışmada şu şekilde açıklanmaktadır:

İçerik: Değerlendirme ölçütlerinin konu dışı herhangi bir içeriği tanımlayıp tanımlamadığı; ölçütlerin tasarlanan içeriğin tüm yönlerini tanımlayıp tanımlanmadığı; dereceli puanlama anahtarıyla değerlendirilecek ödevin ya da etkinliğin tanımlanmamış herhangi bir içerik alanının olup olmadığı ile ilgilidir.

Yapı: Amaçlanan yapının tüm önemli yönlerinin puanlama ölçütleriyle değerlendirilip değerlendirilmediği ve tüm değerlendirme ölçütlerinin ilgili yapıyla ilişkili olup olmadığıdır.

Ölçüt: Puanlama ölçütlerinin, ilgili performansı ya da gelecekteki başarının unsurlarını nasıl yansıttığı; puanlama ölçütlerine yansıtılmayan ilişkili performansın herhangi bir yönünün olup olmadığı gibi durumlardır.

ODBİMDEPA’nın geçerlik çalışmaları da Moskal ve Leydens (2000) temel alınarak “içerik, yapı ve ölçüt” yönlerine dayalı olarak üç aşamalı bir süreçle tamamlanmıştır:

Birinci aşamada ölçütler, okur dostu metinlerin özelliklerini belirleyen çalışmalar (Gray ve Leary, 1935; Dreher ve Singer, 1989; Nagabhand, Nation ve Franken, 1993) temelinde geliştirilmiştir. Böylelikle değerlendirme aracının içerik yönü denetlenmiştir.

İkinci aşama, daha önce de belirtildiği gibi dereceli puanlama anahtarının işlerliğini görmek için birinci belge setindeki (2018 yılı MEB’de kullanılan *Fen Bilimleri, Sosyal Bilimler, İnsan Hakları, Yurttaşlık ve Demokrasi* dersi ile *Türkçe* dersi kitaplarındaki bilgilendirici metinler) toplam 50 metin kullanılarak iki akademisyen değerlendirici ile yürütülmüş ve geliştirilen dereceli puanlama anahtarına ait güvenilirlik sonuçları, puanlama anahtarındaki puanlamanın sürerli (ordinal) olması nedeniyle Cohen kappa (Weighted kappa) analizleri yapılarak elde edilmiştir. Puanlama sonrası akademisyenlerden içerik ve yapı ile ilgili dönütler alınmış, daha sonra yapılan çalışma UTEOK 2018’de bildiri olarak sunulmuş dinleyicilerden de görüş alınmıştır. Dönütler ile puanlama anahtarının içerikleri yeniden düzenlenmiş böylece içerik, yapı ve ölçüt geçerliğinin geliştirilmesi sağlanmıştır.

Üçüncü aşama, ikinci belge setindeki (2023 ve 2024 yıllarındaki YKS-TYT'den ve YKS-AYT'den seçilen *Türk Dili ve Edebiyatı, Felsefe, Din kültürü ve Ahlak Bilgisi, Sosyoloji ve Tarih* dersi sorularındaki bilgilendirici metinleri) toplam 37 ölçme değerlendirme metni kullanılarak ikinci araştırma grubundaki Felsefe, Din kültürü ve Ahlak Bilgisi, Sosyoloji ve Tarih metinleri için 11, Türk Dili ve Edebiyatı metinleri için 13 değerlendirici ile yürütülmüş ve puanlama sonuçlarının analizinde, değerlendiriciler arası güvenilirliği değerlendirmek için iki farklı istatistiksel yaklaşım kullanılmıştır: R'da *irr* paketi ile yapılan "Fleiss' kappası" ve TAM paketi ile yapılan "çok yüzeyli Rasch modeli".

Eser ve Aksu'ya (2022, s.58-59) göre Fleiss'in kappası, üç veya daha fazla değerlendirici arasındaki uzlaşmanın ne düzeyde olduğunu belirlemenin bir yolu olup yaygınlık ve yanlılıktan etkilendiği için yüksek uzlaşma, ancak düşük kappası paradoksuna yol açabilmektedir. Ek olarak aynı çalışmada, Fleiss'in kappası kapsamında elde edilen sonuçlara ilişkin aralıkların, Cohen'in kappasına ilişkin sonuçların yorumlanmasında kullanıldığı belirtilmektedir.

Fleiss (1971) bu yöntemin, değerlendiricilerin kategorik verileri ne kadar tutarlı bir şekilde değerlendirdiğini gösterdiğini ve belirli ölçütler doğrultusunda bireylerin, puanlama anahtarlarını nasıl puanladığını ölçmek için uygun olduğunu bildirmektedir. Kappa istatistiği, şans eseri ortaya çıkabilecek uyumu hesaba katmakta ve sonuçlar -1 ile 1 arasında değişmektedir. 1, mükemmel uyumu, 0, şans eseri uyumu ve **negatif değerler**, şansa bağlı uyumdan daha kötü uyumu göstermektedir.

Diğer taraftan, çok yüzeyli Rasch modeli, bütün parametreleri kalibre ederek, sıralama ölçeğindeki gözlemleri eşit aralıklı bir logit ölçeğine dönüştüren doğrusal bir modeldir (Bond ve Fox, 2015).

Rasch modeli geleneksel ölçme yöntemlerine göre birtakım üstün özelliklere sahiptir (Rasch, 1960; Linacre, 1993). Bunları kısaca özetlemek gerekirse: Çok yüzeyli Rasch modeli ham puanlar yerine, kişilere ait (ölçme hatalarından arındırılmış) ölçme değerlerini kullanmaktadır. Çok yüzeyli Rasch modeli aynı zamanda, her bir yüzey için (örneğin bu çalışmada bilimsel araştırma ödevlerinin kalitesi, değerlendirme sorularının özellikleri, değerlendirici davranışı) yüzeyler arası olması gereken doğrusal bir bağlantının kurallarını belirlemekte ve bu bağlantıları oluşturmaktadır (Hambleton ve Swaminathan, 1985). Kısacası, çok yüzeyli Rasch modeli, yansız ve etkili bir ölçme elde etmek için yüzeyleri ortak bir düzlemde birleştirerek standart hale getirmekte ve bireylerin bilimsel araştırma becerilerini, soruların güçlüğünü ve değerlendiricilerin "katılık" ya da "cömertlik" davranışlarını aynı anda karşılaştırma olanağı sunmaktadır. Çalışmanın temel amaçlarından biri olan ölçme aracı olarak kullanılan metinlerin okur dostu olup olmadığını belirleyebilmek için değerlendiricilerin birbiri ile uyumundan çok değerlendiricilerin değerlendirme sürecinin genelinde "katı" veya "cömert" puanlama tutarlılığına odaklanılmıştır. Dereceli puanlama anahtarının kullanımı öncesinde herhangi bir eğitim verilmediği ve standardizasyon çalışması yapılmadığı için değerlendiriciler arası uyumun düşük olması göz ardı edilmiş ve tutarlılık temel alınarak değerlendirme yapılmıştır.

Verilerin Toplanması

Çalışmanın UTEOK 2018'de bir bildiri olarak sunulması sürecinde 5, 6, 7, ve 8. sınıf ders kitaplarındaki (Fen Bilgisi, İnsan Hakları ve Yurttaşlık Bilgisi, Sosyal Bilgiler, Türkçe) bilgilendirici metinler kullanılmış; dereceli puanlama anahtarının değerlendiricilerden gelen dönütler ve öneriler doğrultusunda ODBİMDEPA'nın boyutlarının ifadeleri ve açıklamaları geliştirilerek güncellenmiştir.

Bu çalışmanın ikinci aşamasında ODBİMDEPA'nın geçerlik çalışması için ilk olarak, yukarıda detayları anlatıldığı biçimde belge incelemesi yöntemi ile bilgilendirici okuma metinleri (ÖSYM'nin 2023 ve 2024 YKS sorularından) belirlenmiştir. Daha sonra uygun örnekleme yöntemiyle seçilen değerlendiricilerin her birine, daha önceden "ODBİMDEPA Kullanma Kılavuzu" verilerek dereceli puanlama anahtarını kullanarak metinleri nasıl puanlamaları gerektiği hakkında bilgilendirme yapılmıştır. Devamında, her değerlendiriciye, değerlendirecekleri metinlerin ve her metnin karşısında, dereceli puanlama anahtarının 8 boyutunun yer aldığı değerlendirme formları ODBİMDEPA ile verilmiştir.

Puanlamaların ardından, gerekli istatistik analizlerini yapabilmek için tüm değerlendiricilerin verileri bir araya getirilerek gerekli çözümlenmeler yapılmıştır.

Verilerin Çözümlemesi

İlk aşamada iki değerlendirici tarafından puanlama yapıldığı için iki değerlendiricinin güvenilirliğine ilişkin değerleri sunan Cohen Kappa (Weighted Kappa) yöntemi tercih edilmiştir. Cohen Kappa ile yapılan analizlere ilişkin bulguların değerlendirilmesinde Şencan'ın (2005) yorumu kullanılmıştır. Şencan (2005, s. 485), Cohen kappa katsayısından elde edilen verilerin şu şekilde yorumlandığını belirtmektedir:

<i>Zayıf uyuşma</i>	= < .20
<i>Kabul edilebilir uyuşma</i>	= .20-.40
<i>Orta Derecede uyuşma</i>	= .40-.60
<i>İyi uyuşma</i>	= .60-.80
<i>Çok iyi uyuşma</i>	= .80-1.00

Yukarıdaki katsayı aralıklarına göre katsayı .20 ve üstünde çıktığında değerlendiriciler arası güvenilirlikten söz edilebilmekte; katsayı .80- 1.00 arasında çıktığında ise güvenirlığın en üst düzeyde olduğu kabul edilmektedir.

Cohen Kappa yönteminde değerlendiriciler arasındaki uyum, her iki değerlendiricinin aynı puanı vermesi ile sağlanmaktadır. Halbuki bir metnin değerlendirildiği ODBİMDEPA'da değerlendiricilerden birinin 2, diğerinin 3 puan vermesi ile 1 ve 4 puan vermeleri arasındaki uyumsuzluk farklıdır. Değerlendiricilerin puanları incelendiğinde uyumsuzluk düzeyinin genellikle 1 puan farkla oluştuğu görülmüştür. Bu nedenle ilk aşamadaki değerlendirme sürecinin ardından değerlendiricilerin dönütleri ile yapı geçerliğini arttırmak için değişiklikler yapılmış ve daha fazla değerlendirici ile güvenilirlik çalışmasının yürütülmesine karar verilmiştir. Bunun yanı sıra metinlerin aldıkları puanların ortalamalarının da incelenmesi ve metinlerin okur dostu olup olmama durumlarındaki genel eğilimin belirlenmesi amaçlanmıştır.

Çalışmanın ikinci aşamasında yukarıdaki değerlendirmelerden yola çıkılarak daha fazla değerlendiriciye ulaşılarak dereceli puanlama anahtarı ile YKS'den seçilen metinler puanlatılmıştır. İki'den fazla değerlendiricinin puanlama yaptığı durumlarda güvenilirlik için Cohen Kappa (Weighted Kappa) yöntemi kullanılmayacağı için Fleiss Kappa modeli tercih edilmiştir. Fleiss Kappa analizleri R programında *irr* paketi ile yapılmıştır. Değerlendiriciler arasındaki güvenilirliklerin istenen düzeyde olmadığı görülmüştür. Fleiss Kappa yönteminde Cohen Kappa'ya benzer biçimde değerlendiriciler arasındaki uyum her değerlendiricinin aynı puanı vermesi ile sağlanmaktadır. İlk aşamadaki puanlamaya benzer biçimde ikinci belge setinde yapılan puanlamalarda da değerlendiriciler arasındaki puan farklarının çoğunlukla 1 puan olduğu görülmüştür. Yukarıda da bahsedildiği üzere dereceli puanlama anahtarı için yürütülen güvenilirlik çalışmalarında puan aralığının düşük olması da önemli bir göstergedir. Bu nedenle değerlendiriciler arası tutarlılıktan ziyade değerlendiricilerin iç tutarlılığı ve "katılık" / "cömertlik" eğilimlerinin belirlenmesinin ve iç tutarlılığın yüksek çıkması durumunda metinlerin ortalama puanlar ile incelenmesinin uygun olacağına karar verilmiştir. Değerlendiricilerin iç tutarlılığına ilişkin değerleri elde edebilmek için de R programında *TAM* paketi ile çok yüzeyli Rasch analizi gerçekleştirilmiştir. Fleiss' kappanın aksine, çok yüzeyli Rasch modeli, değerlendiriciler arasındaki değişkenliği, modelin bir parçası olarak ele almaktadır. Rasch modeli, ölçülen performansı tahmin ederken aynı zamanda değerlendiricilerin "katılığı" veya "cömertliği" gibi faktörleri de hesaba katmaktadır (Rasch, 1960). Bu durum, değerlendiriciler arasında tutarsızlıklar olsa bile, modelin bu farklılıkları dengelemesi ve bu sayede daha yüksek güvenilirlik değerlerine ulaşılması anlamına gelmektedir.

Çok yüzeyli Rasch modeli ile ulaşılan modelin kestirim güvenirlığının yüksek olması puanlama anahtarının güvenilirliğine ilişkin başka bir bakış açısı sağlamıştır. Rasch modeli ile farklı konu alanlarındaki puanlamalardaki EAP güvenirlikleri ve değerlendiriciler arası tutarlılığı görebilmek için korelasyonlar incelenmiştir.

Değerlendiricilerin, metinlerin ve boyutların ilişkisini görebilmek için FACET programında 3 yüzeyli çok yüzeyli Rasch analizi yapılmıştır. Yüzey olarak puanlayıcılar, metinler ve boyutlar kullanılmıştır.

Değerlendiricilerin cömertlik-katılık eğilimlerini anlamak amacıyla her bir değerlendirici için bir değerlendirici katılık/cömertlik indeksi (Rater Severity Index-RSI) hesaplanmıştır.

$$RSI = \text{Ortalama puan} - \text{genel ortalama} / \text{genel standart sapma}$$

Her değerlendiricinin ortalama puanını z-skoruna dönüştürerek, puanlamalarının, genel ortalamadan kaç standart sapma gösterdiğinin bulunması amaçlanmıştır. Z-skoru standardizasyonu kullanmak, değerlendirici eğilimini ortak bir ölçek üzerinde karşılaştırmayı mümkün kılmıştır (Bond ve Fox, 2015; Field, 2013).

Değerlendiricilerin tutarlı davrandıkları sonucuna ulaşıldıktan sonra, metinlerin okur dostu olup olmadığının belirlenmesi amaçlanmıştır. Bu amaçla, değerlendiricilerin puanlarının ortalaması (M) alınarak metinlerin okur dostu olup olmadığı incelenmiştir. Her bir metin için her boyutta değerlendiricilerin verdiği puanların ortalaması alınmış ardından tüm boyutların ortalaması alınmıştır. Değerlendiricilerin her bir boyuttaki puan ortalamaları toplanarak toplam puan elde edilmiştir.

ODBİMDEPA kullanılarak puanlanan bir metin, her boyuttan en fazla 4 puan; toplamda da en fazla 32 puan alabilmektedir. Bu belirlemeler ışığında, bilgilendirici bir metnin “okur dostu” metin özellikleri taşıyıp taşımadığına aşağıdaki puan aralıklarına bakarak karar vermek mümkündür:

25 puan ve üzeri	: Okur dostu metin
16-24 puan	: Orta düzey okur dostu metin
15 puan ve altı	: Okur dostu olma özelliği taşımayan metin

Yukarıdaki puan aralıklarının belirlenmesinde, boyutlardan alınabilecek en yüksek ve en düşük puanlar göz önünde bulundurularak yüzdelik hesap yapılmıştır: Şöyle ki; bir metin, her boyuttan 1 puan alırsa toplam 8 puan; her boyuttan 2 puan alırsa toplam 16 puan; her boyuttan 3 puan alırsa toplam 24 puan alabilmektedir. Buna göre, ölçme değerlendirme sistemlerinde “teamülen” 4/4 mükemmel, 3/4 başarı elde eden, yeterli, 2/4 başarı elde eden orta düzey ve 1/4 yetersiz olduğundan ODBİMDEPA’dan 25 puan ve üstünü alan bir metin “okur dostu” olarak kabul edilirken 15 puan ve altında bir puanı alan metin ise okur dostu olarak kabul edilmemektedir.

Sınırlılıklar

Çalışmanın birinci sınırlılığı, sadece bilgilendirici metinlerin çalışma kapsamında değerlendirilmesi; anlatsal metinlerin kapsam dışı bırakılmasıdır. Çalışmanın ikinci sınırlılığında ise birinci belge setinde şiir metinleri, görsellerin bulunduğu ve/veya eksik kalmış/bitmemiş metinler; ikinci belge setinde paragraftaki boşlukların doldurulmasını isteyen ve/veya tek tümceden oluşan paragrafların olduğu metinler, görsel içeren metinler, şiirler, metnin akışını bozan ya da bir tümceden sonra araya girebilecek tümcenin sorulduğu metinler, paragrafın tamamının bir tanımdan oluştuğu metinler, bilginin maddeler halinde verildiği metinler ve Din kültürü ve ahlak bilgisi dersi sorularında, tamamı Kuran-ı Kerim’den alınmış bir ayetten oluşan metinler çalışma kapsamı dışında tutulmuştur.

Bulgular

Konusu, bilgilendirici metin dereceli puanlama anahtarı olan bu çalışmada, okur dostu bilgilendirici metinleri değerlendirmek üzere geliştirilen ODBİMDEPA'nın tanıtılması ve eğitim-öğretim, ölçme değerlendirme ortamlarında nasıl kullanılacağı hakkında bilgi vermesi amaçlanmıştır. ODBİMDEPA'nın bu amaçlar doğrultusunda geçerliliğini ve güvenilirliğini ölçmek için gerçekleştirilen süreçler sonucunda elde edilen bulgular, sırasıyla aşağıda açıklanmıştır:

1. Birinci çalışma grubundan elde edilen Cohen kappa (Weighted Kappa) analizi bulguları

Çalışmanın birinci çalışma grubu (bağımsız 2 akademisyen değerlendirici) tarafından 5, 6, 7 ve 8. sınıf ders kitaplarındaki bilgilendirici metinlerin, ODBİMDEPA'nın 8 boyutuna ait puanlarının ağırlıklı ortalama kappa değeri 0.53 [%95 GA (0.46, 0.60)] olarak hesaplanmıştır. Puanlamada 50 metin kullanılmıştır.

Birinci çalışma grubundan elde edilen Cohen kappa katsayıları, Cohen'e (1960) göre belirli oranda uyum içindedir. Bu bağlamda ODBİMDEPA'nın, dengeli ve uyumlu bir dağılıma sahip olduğu; diğer bir ifadeyle yüksek güvenilirliğe sahip olduğu, ancak daha yüksek bir güvenilirlik oranına ulaşmak için daha fazla değerlendiriciye gereksinim olduğu görülmüştür. Bu nedenle, çalışmanın ikinci bir çalışma grubuyla da yürütülmesine karar verilmiştir. İlerleyen bölümlerde, ikinci çalışma grubundan elde edilen bulgulara da değinilecektir.

Değerlendiriciler arasındaki genel uyumun yanı sıra her bir boyutta değerlendiricilerin uyumu da incelenmiştir. ODBİMDEPA'nın her bir boyutuna ait elde edilen kappa değerleriyle ilgili bulgular Tablo 2'de verilmiştir.

Tablo 2. Boyutlara göre Cohen (Weighted) Kappa Katsayıları

Boyut	Ölçüt	Cohen (Weighted) Kappa Katsayısı
1	Metin düzenleme	0.20
2	Bilginin Sunumu	0.17
3	Biçimsel Örüntüler/Bağlılık	0.34
4	Anlamsal Örüntüler/Tutarlılık	0.53 *
5	Açıklama ve Geçerlik	0.55 *
6	Tartışma	0.60 *
7	Biçem	0.57 *
8	Evensellik	0.14

* p<.05

Tablo 2'de her bir boyut için görülen değerlendirici güvenilirliğinin özellikle 2. ve 8. boyutta önemsiz uyuma düzeyinde olduğu, 1. ve 3. boyutta ise orta derecede uyuma olduğu görülmektedir. 2 boyutta değerlendirici güvenilirlikleri Cohen Kappa katsayıları düşük olsa da diğer boyutlar için değerlendiriciler arasındaki uyumun kabul edilebilir olduğu görülmektedir.

1.1. Birinci Belge Setindeki Bilgilendirici Metinlerden Alınan Puan Ortalamalarının Sınıflara Göre Dağılımı

Birinci belge setinden elde edilen veri tabanında yer alan metinlerin sınıflara göre metin sayıları ve ortalamaları Tablo 3'te gösterilmiştir.

Tablo 3. Bilgilendirici Metinlerin Puan Ortalamalarının Sınıflara Göre Dağılımı

Sınıf	Metin Sayısı	Puan Ortalaması
4. Sınıf	N=16	M= 2.3
5. Sınıf	N=8	M= 2.4
6. Sınıf	N=11	M= 2.3
7. Sınıf	N=11	M= 2.0
8. Sınıf	N=4	M=2.2
TOPLAM	50	2.24

Tablo 3 incelendiğinde, en çok bilgilendirici metnin (N=16) 4. sınıf düzeyinde; en az metnin (N=4) ise 8. sınıfta yer aldığı görülmektedir. Buna karşın en yüksek puan ortalaması 5. sınıfta 2.4 en düşük puan ortalaması 7. sınıfta 2.0'dır. Bu bulgular, metinlerin farklı yaş gruplarına göre uyarlanmış olduğu göz önünde bulundurularak farklı sınıflarda metinlerin anlaşılabilirlik ve dolayısıyla okur dostu olma düzeylerinin değişip değişmediğinin gözlenmesini sağlamaktadır: Küçük yaş gruplarında metinlerin daha iyi düzenlenmiş ve anlaşılabilirliklerinin daha yüksek olması beklenirken, elde edilen puan ortalamaları, yaş gruplarına göre metinlerin okur dostu olma düzeylerinin sınıflara göre değişmediğini göstermektedir.

1.2. Birinci Belge Setindeki Bilgilendirici Metinlerden Alınan Puan Ortalamalarının Derslere Göre Dağılımı

Birinci belge setindeki toplam 50 metnin derslere göre dağılımları ve her dersin puan ortalamaları Tablo 4'te gösterilmiştir.

Tablo 4. Bilgilendirici Metinlerden Alınan Puan Ortalamalarının Derslere Göre Dağılımı

Ders	Metin Sayısı	Puan Ortalaması
Fen Bilgisi	N=17	M=2.2
İnsan Hakları ve Yurttaşlık Bilgisi	N=2	M=1.9
Sosyal Bilgiler	N=15	M=2.2
Türkçe	N=16	M=2.4
TOPLAM	50	2.18

Tablo 4 incelendiğinde, bilgilendirici metin türünün en çok 17 metinle Fen Bilgisi ve 16 metinle Türkçe dersinde kullanılırken İnsan Hakları ve Yurttaşlık Bilgisi dersinde en az sayıda (N=2) kullanıldığı görülmektedir. Fen Bilgisi dersinde kullanılan bilgilendirici metin sayısının, Türkçe dersinde kullanılan bilgilendirici metin sayısından daha yüksek olmasına rağmen, Türkçe dersinin puan ortalaması (2.4) en yüksektir. En az bilgilendirici metin sayısına sahip İnsan Hakları ve Yurttaşlık Bilgisi dersinin ortalaması en düşük ortalama ile 1.9'dur.

1.3. Birinci Belge Setindeki Tüm Bilgilendirici Metinlerin ODBİMDEPA Boyutlarına Göre Aldıkları Puanların Ortalamaları

Birinci belge setindeki toplam 50 metnin, ODBİMDER'in 8 boyutundan aldığı puanlara karşılık gelen ortalamalar, Tablo 5'te gösterilmiştir.

Tablo 5. Tüm Bilgilendirici Metinlerin Boyutlara Göre Aldıkları Puanların Ortalamaları

Boyut	Puan Ortalaması
1- Metin düzenleme	2.6
2- Bilginin sunumu	2.24
3- Biçimsel örüntüler/Bağlaşıklık	2.17
4- Anlamsal örüntüler/Tutarlık	2.17
5- Açıklama ve gerçeklik	2.31
6- Tartışma	1.37
7- Biçem	1.94
8- Evrensellik	2.89
TOPLAM	2.22

Tablo 5 incelendiğinde, ODBİMDEPA'nın 2018'deki ilk versiyonunda en iyi işleyen boyutun 2.9 ortalama ile "Evrensellik" boyutu ve 2.7 puan ortalaması ile "Metin Düzenleme" boyutu olduğu; buna karşın 1.3 puan ortalaması ile en az işleyen boyutun "Tartışma" ve 1.9 ile "Biçem" olduğu görülmüştür.

2. İkinci çalışma grubundan elde edilen Fleiss' Kappa ve çok yüzeyli Rasch modeli yaklaşımlarının bulguları

Çalışmanın ikinci çalışma grubu (4 branş için 11, 1 branş için 13 birbirinden bağımsız değerlendirici) tarafından, ÖSYM'nin 2023 ve 2024 YKS testlerindeki toplam 37 bilgilendirici metin, ODBİMDEPA'daki 8 boyuta göre puanlanmıştır. Puanlama sonuçlarının analizinde, **değerlendiriciler arası güvenilirliği** değerlendirmek için iki farklı istatistiksel yaklaşım kullanılmıştır: R'da *irr* paketi ile yapılan **Fleiss' kappa** ve TAM paketi ile yapılan **çok yüzeyli Rasch modeli**. Bu iki yöntem, verileri farklı açılardan ele almakta olup, sonuçları her modelin varsayımları ve güçlü yönleri bağlamında yorumlanmalıdır. Aşağıda, bulguların açıklamaları, değerlendiriciler arası güvenilirlik üzerindeki etkileri ve bu değerlerin nasıl raporlanması gerektiğiyle ilgili açıklamalar bulunmaktadır.

2.1. Fleiss' Kappa: Değerlendiriciler Arasındaki Uyumun Ölçülmesi

ÖSYM'nin YKS testlerinden oluşan veri tabanının bilgilendirici metinlerinin ODBİMDEPA'ya göre puanlanması sonucunda elde edilen Fleiss' kappa bulguları, Tablo 6'da gösterilmiştir.

Tablo 6. ODBİMDEPA'ya göre Puanlanan YKS Testlerindeki Bilgilendirici Metinlerin Fleiss' Kappa Değerleri

Metin İçeriği	Kappa Değeri
Din kültürü ve ahlak bilgisi testi metinleri	0.09 *
Felsefe testi metinleri	0.03 *
Sosyoloji testi metinleri	0.02
Türk dili ve edebiyatı testi metinleri	0.05 *
Tarih testi metinleri	0.03 *

* p<.05

Tablo 6 incelendiğinde, elde edilen Fleiss' kappa değerlerinin 0.02 ile 0.09 arasında değişmesi, değerlendiriciler arasında uyum olmadığını göstermektedir. Bazı sonuçlar istatistiksel olarak anlamlı olsa da (p<0.01), düşük kappa değerleri değerlendiricilerin, dereceli puanlama anahtarları aynı şekilde uygulamada tutarlı olmadığını göstermektedir. 0.09 kappa değeri (Din Kültürü ve Ahlak Bilgisi testi metinleri) oldukça düşük olup, şansa bağlı uyumdan anlamlı bir fark göstermemektedir. Bazı durumlarda (Sosyoloji, kappa = 0.02) uyum, istatistiksel olarak anlamlı değildir, bu da değerlendiriciler arasındaki uyumun rassal olduğunu göstermektedir. Birinci aşama ile ilgili uyumsuzluklarda yapılan açıklama burada da geçerlidir. 4 puanlı ve 8 ölçütlü bir puanlama anahtarında değerlendiricilerin uyumsuzluklarının derecesinin de önemli olduğu göz önünde bulundurulmalıdır. Örneğin Tablo 7'deki YKS Sosyoloji bölümü 1. metin 11 değerlendirici tarafından yapılan puanlamada değerlendiricilerin, 8 boyut için genel olarak birbirine yakın aralıklarda puan verdikleri görülmektedir.

Tablo 7. ODBİMDEPA'ya göre YKS Sosyoloji Testindeki Bilgilendirici 1. Metnin Değerlendirici (D) Puanları

Boyut	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11
1	4	4	4	4	4	4	2	4	4	4	4
2	3	4	3	3	4	4	2	3	4	4	3
3	4	3	2	4	2	4	3	4	4	4	2
4	4	3	3	4	3	4	2	4	4	4	3
5	3	3	3	3	4	4	3	3	4	3	4
6	4	4	2	4	4	2	1	3	4	4	3
7	2	4	3	4	3	1	1	4	4	4	3
8	4	3	4	3	4	3	4	3	4	4	4

Tablo 7'de değerlendiriciler 1. boyutta 7 numaralı değerlendirici (D7) dışında tüm değerlendiricilerin bu boyuta 4 puan verdiği görülmektedir. 2. boyutta değerlendiricilerin genellikle 3 ve 4 puan verdikleri, 3., 4., 5. ve 8. boyutta da benzer bir puanlama eğiliminin olduğu görülmektedir. 6. ve 7. boyutlarda puan aralığı daha açıktır.

2.2. Çok Yüzeyle Rasch Modeli: Değerlendirici Etkilerinin Düzeltilmesi

Yapılan çözümlenmelerde, tüm branşlarda (Türk dili ve edebiyatı, Felsefe, Sosyoloji, Din Kültürü ve Ahlak Bilgisi, Tarih) benzer sonuçlara ulaşılmıştır. Tablo 8’de, çok yüzeyle Rasch sonuçları gösterilmektedir:

Tablo 8. ODBİMDEPA’ya göre Puanlanan YKS Testlerindeki Bilgilendirici Metinlerin Çok Yüzeyle Rasch Bulguları

Çok Yüzeyle Rasch Öğeleri	Değer
EAP Güvenirlik	0.95
Varyans	0.07
Standart Sapma	0.27

Tablo 8 incelendiğinde, EAP güvenirliliği (A Posteriori Beklenen Güvenirlilik) olan **0.95**, modelin, metinlerin okur dostu olma ölçütlerine uygunluklarını **oldukça güvenilir** bir şekilde tahmin edebildiğini göstermektedir. Varyans (0.07) ve standart sapma (0.27) değerleri, puanlamalar arasında bir miktar varyasyon olduğunu, ancak modelin bu **farklılıkları telafi edebildiğini** ve böylece **güvenilir tahminler** elde ettiğini göstermektedir.

Bu sonuç, **ölçme açısından**, değerlendiriciler arasında tutarsızlıklar olsa dahi Rasch modelinin, metinlerin değerlendirilmesinde **geçerli çıkarımlar** yapabildiğini göstermektedir. Diğer bir ifadeyle Rasch modeli, dereceli puanlama anahtarı değerlendiricilerin tutarsız olduğunu varsaymakta, ancak her değerlendiricinin önyargılarını modelleyerek bu tutarsızlıkları dengelemektedir. Bu durum, değerlendirilen metinlerin okur dostu olup olmadığına ilişkin puanlarının, değerlendirici farklılıklarına rağmen güvenilir ölçümler üretildiğini göstermektedir.

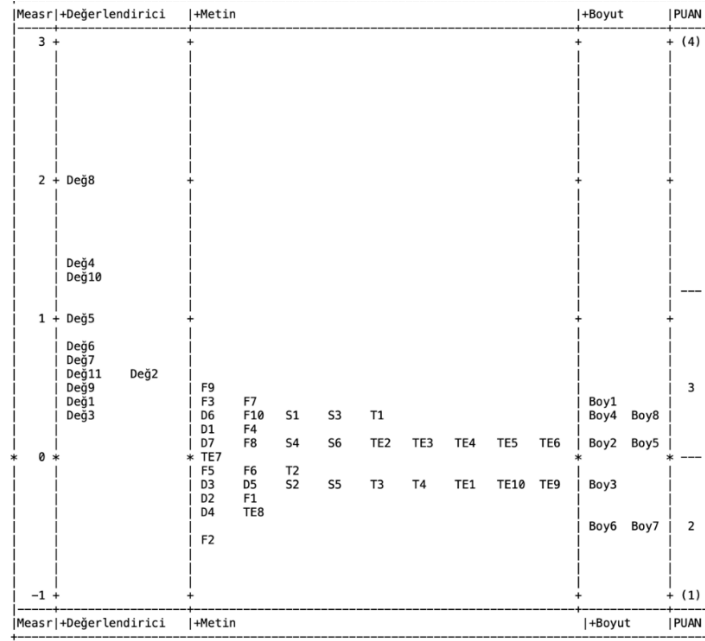
Çok yüzeyle Rasch modeli ile farklı YKS konu alanlarına ilişkin metinlerde, değerlendiricilerin her bir boyut için yaptıkları puanlamaya göre EAP güvenirlilikleri ve korelasyonlar da belirlenmiş ve Tablo 9’da sunulmuştur.

Tablo 9. YKS metinlerinin disiplinlere göre değerlendirici güvenirlilikleri

YKS Metinleri	EAP	Korelasyon
Sosyoloji	0.88	0.45
Tarih	0.95	0.27
Türk dili ve edebiyatı	0.88	0.40
Din kültürü ve ahlak bilgisi	0.92	0.62
Felsefe	0.92	0.59

Tüm branşlardaki metinlerde EAP güvenirliliğinin yüksek olduğu Tablo 9’da görülmektedir. EAP güvenirliliğinin yüksek olması modelin güvenilir kestirimler yapabildiğini göstermektedir. Tüm değerlendiriciler göz önünde bulundurularak yapılan değerlendiriciler arası güvenirlilikte ise özellikle tarih metinlerinde korelasyonun düşük olduğu görülmektedir. Veri seti incelediğinde iki değerlendiricinin “aykırı (outlier)” olduğu tespit edilmiştir. Korelasyon değerlerinin düşük olmasında bu değerlendiricilerin etkisi bulunmaktadır.

Dereceli puanlama anahtarı ile ilgili genel bir görünüm elde edebilmek için değerlendiricilerin her bir metin için puan ortalamaları ile uyumlarını gösteren analizin çıktısı Şekil 1’de gösterilmektedir. Yalnızca Türk Dili ve Edebiyatı metinlerini puanlayan iki değerlendirici bu analize dahil edilmemiş ve 11 değerlendiricinin tüm metin ve boyutlarla ilişkisi incelenmiştir. Tarih (T), Felsefe (F), Sosyoloji (S), Din Kültürü ve Ahlak Bilgisi (D), Türk Dili ve Edebiyatı (TE) olarak kodlanmış ve analizler FACET programı ile gerçekleştirilmiştir. Değerlendirici boyutu için güvenirlilik değeri 0.98, ayırma indeksi (separation index) 6.60 olarak bulunmuştur.



Şekil 1. Değerlendirici, Metin, Boyut Kalibrasyon Haritası

Şekil 1’de değerlendiriciler arası logit değerlerin 0 ve 2 aralığında değiştiği görülmektedir. Metinler bağlamında ise Felsefe (F1-F9) metinlerinden F2’nin diğer metinlere kıyasla düşük puan alarak ortalamasının altında kalmıştır. Öte yandan F3, F7 ve F9’un diğer metinlerden daha yüksek puan ortalamasına sahip olduğu dolayısıyla okur dostu olma düzeyinin daha yüksek olduğu söylenebilir. Dereceli puanlama anahtarında değerlendiricilerin 1. boyutta daha yüksek, 6. ve 7. boyutlarda ise daha düşük puanlama yaptıkları ve tüm boyutlarda genellikle 2 ve 3 puanı tercih ettikleri görülmüştür. Değerlendiricilerin en düşük ve en yüksek puanları daha az tercih etmeleri dereceli puanlama anahtarının metinlerin niteliklerine ilişkin soyut değerlendirmelerle ilgili olması ve değerlendiricilerin güvenli alanda kalma isteği olabilir. Değerlendiricilerin genellikle daha cömert bir puanlamaya eğilimli olduğu Şekil 1’de görülsede değerlendiricilerin katılık/cömertlik eğilimlerinin bireysel olarak incelenmesi, değerlendiricilerin tutarlılığı ile ilgili bilgi edinebilmek ve özellikle “aykırı (outlier)” olarak belirlenen değerlendiricilerin belirlenmesi için önemlidir.

Değerlendiricilerin bireysel eğilimlerini belirlemek için her bir değerlendiricinin ortalama puanı z-skoruna dönüştürülerek değerlendirici katılık/cömertlik değerleri belirlenmiştir. Tablo 10’da, değerlendiricilerin katılık/cömertlik eğilimleri ile ilgili bulgular verilmiştir.

Tablo 10. Değerlendiricilerin katılık/cömertlik değerleri (RSI)

Değerlendirici	Sosyoloji	Tarih	Din kültürü ve ahlak bilgisi	Felsefe	Türk dili ve edebiyatı
D1	-0.48	-0.68	-0.36	-0.01	-0.36
D2	-0.12	0.64	-0.25	-0.58	-0.05
D3	-0.57	-0.49	-0.32	-0.47	-0.37
D4	0.62	0.29	0.02	0.29	0.45
D5	0.10	0.14	0.28	0.20	-0.12
D6	-0.01	-0.04	-0.05	-0.10	-0.04
D7	-0.34	0.29	0.24	0.32	-0.62
D8	0.55	0.67	0.24	0.64	0.69
D9	-0.03	-0.52	0.36	-0.70	-0.44
D10	0.53	0.42	0.55	0.16	0.03
D11	-0.24	-0.74	-0.70	0.25	0.03
D12	-	-	-	-	0.37
D13	-	-	-	-	0.42

Tablo 10’da her bir değerlendiricinin farklı konu alanlarındaki metinler için yaptıkları puanlamadaki katılık ve cömertlik eğilimlerini gösteren değerler sunulmuştur. 0’dan uzaklaştıkça katılık/cömertlik eğilimi artmaktadır ve (-) değerler katılığı, (+) değerler cömertliği göstermektedir. Tablo 10 incelendiğinde D5’in genel eğilimi cömert iken Türk Dili ve Edebiyatı metinlerinde katı davrandığı görülmektedir. Ancak her iki eğilimde de değerlerin 0’a yakın olduğu görülmektedir. Dolayısıyla bu geçişkenliğin değerlendiricinin katılık/cömertlik eğilimindeki tutarlılığını belirgin bir biçimde olumsuz yönde etkilemediği söylenebilir. Ancak D7 ve D9 için aynı yorumu yapmak mümkün değildir. Her iki değerlendiricinin de katılık/cömertlik eğilimleri, metinlerin konu alanlarına göre rassal bir değişime sahip gibi görünmektedir. Söz konusu değerlendiriciler yukarıda bahsedildiği gibi “aykırı (outlier)” olarak kabul edilebilir. D11’in eğilimleri de farklı konu alanlarında değişse de D7 ve D9’dan farklı bir durum söz konusudur. Genel eğilimi katı olan bu değerlendirici, felsefe metinlerinde daha cömert davranmıştır. Türk Dili ve Edebiyatı metinlerinde de cömert bir eğilime sahip olsa da RSI değeri 0’a yakın olduğu için katılık/cömertlik eğilimindeki tutarlılığa olumsuz etkisinin yüksek oranda olmadığı söylenebilir. Felsefe metinlerindeki cömertlik eğilimi ise tutarlılığı etkileyecek orandadır. Ancak genel eğilimi göz önünde bulundurulduğunda D7 ve D9 gibi “aykırı” olarak kabul edilmemesi gerekmektedir. D12 ve D13, yalnızca bir konu alanında puanlama yaptığı için katılık/cömertlik eğilimleri konusunda bu değerlendiriciler ile ilgili yorum yapmak mümkün değildir.

2.3. Metinlerin Puan Ortalamaları

Çalışmanın bu kısmında, “Sosyoloji, Tarih, Türk Dili ve Edebiyatı, Din kültürü ve Ahlak bilgisi ve Felsefe” derslerine ait testlerdeki bilgilendirici metinlerle ilgili puan ortalamaları ve toplam puanlara ait bulgulara yer verilmiştir.

YKS Sosyoloji testinden toplam altı bilgilendirici metin (N=6) belirlenmiş ve 11 değerlendirici (D=11) tarafından puanlanmıştır. Bu puanlamalara ait bulgular Tablo 11’de verilmiştir.

Tablo 11. Sosyoloji testi metinlerinin puan ortalamaları

YKS Sosyoloji Metinleri	Puan Ortalaması	Toplam Puan
1	3.38	27.09
2	3.09	24.72
3	3.42	27.36
4	3.28	26.27
5	3.11	24.89
6	3.26	26.09

Sosyoloji testindeki bilgilendirici metinlerin tümünün 25 puan ve üzerinde olduğu ve okur dostu metin olarak sınıflandırılabilmesi Tablo 11’de görülmektedir.

YKS Tarih dersi testinde toplam dört bilgilendirici metin (N=4) belirlenmiş ve 11 değerlendirici (D=11) tarafından puanlanmıştır. Bu puanlamalara ait bulgular Tablo 12’de verilmiştir.

Tablo 12. Tarih testi metinlerinin puan ortalamaları

YKS Tarih Metinleri	Puan Ortalaması	Toplam Puan
1	3.40	27.27
2	3.12	25.09
3	3.07	24.63
4	3.06	24.54

Tablo 12 incelendiğinde, Tarih testindeki bilgilendirici metinleri, Sosyoloji testindeki metinler ile benzer olarak okur dostu olarak sınıflandırmak mümkündür.

YKS Türk Dili ve Edebiyatı testinden toplam on bilgilendirici metin (N=10) belirlenmiş ve 13 değerlendirici (D=13) tarafından puanlanmıştır. Bu puanlamalara ait bulgular Tablo 13’te verilmiştir.

Tablo 13. Türk dili ve edebiyatı testi metinlerinin puan ortalamaları

YKS Türk dili ve edebiyatı Metinleri	Puan Ortalaması	Toplam Puan
1	3.13	25.08
2	3.32	26.54
3	3.40	27.23
4	3.39	27.15
5	3.35	26.77
6	3.38	27
7	3.24	25.92
8	2.93	23.46
9	3.16	25.31
10	3.22	25.77

Türk Dili ve Edebiyatı testinde Felsefe testi dışındaki diğer testlere göre daha fazla bilgilendirici metin yer almaktadır. Tablo 13'e göre, değerlendirilen metinlerin tümü, okur dostu sınıfında yer almaktadır.

YKS Din Kültürü ve Ahlak Bilgisi testinden toplam yedi bilgilendirici metin (N=7) belirlenmiş ve 11 değerlendirici (D=11) tarafından puanlanmıştır. Bu puanlamalara ait bulgular Tablo 14'te verilmiştir.

Tablo 14. Din Kültürü ve Ahlak Bilgisi testi metinlerinin puan ortalamaları

YKS Din kültür ve ahlak bilgisi metinleri	Puan Ortalaması	Toplam Puan
1	3.37	27
2	3.03	24.27
3	3.10	24.81
4	2.93	23.45
5	3.09	24.72
6	3.40	27.22
7	3.27	26.18

Tablo 14 incelendiğinde, Din Kültürü ve Ahlak Bilgisi testindeki 3. metnin sınırda okur dostu sınıfında; 4. metnin ise "Orta düzey okur dostu metin" sınıfında yer aldığı görülmektedir. Diğer metinlerin tamamı, okur dostu metin olarak sınıflandırılabilir.

YKS Felsefe testinden toplam on bilgilendirici metin (N=10) belirlenmiş ve 11 değerlendirici (D=11) tarafından puanlanmıştır. Bu puanlamalara ait bulgular Tablo 15'te verilmiştir.

Tablo 15. Felsefe testi metinlerinin puan ortalamaları

YKS Felsefe Metinleri	Puan Ortalaması	Toplam Puan
1	2.97	23.81
2	2.81	22.54
3	3.44	27.54
4	3.35	26.81
5	3.18	25.45
6	3.17	25.36
7	3.46	27.72
8	3.30	26.45
9	3.52	28.18
10	3.43	27.45

Felsefe testinde, Türk dili ve edebiyatı testi ile aynı sayıda ve diğer testlerden daha fazla bilgilendirici metin vardır. Tablo 15 incelendiğinde, 1. ve 2. metinlerin “Orta düzey okur dostu” metin sınıfında, diğer metinlerin ise “Okur dostu” metin sınıfında yer aldığı görülmektedir.

Sonuç ve Tartışma

Bu çalışmada hem eğitim ortamlarında hem de MEB, ÖSYM, YEE gibi ölçme değerlendirme yapan kurumların sınavlarında kullanılan bilgilendirici metinlerin okur dostu olup olmadıklarının belirlenebilmesi için geliştirilen ve güvenilirlik çalışması yapılan *Okur Dostu Bilgilendirici Metin Dereceli Puanlama Anahtarı*'nın-ODBİMDEPA alanyazına tanıtılması ve incelenen metinlerle ilgili gözlemlerin paylaşılması amaçlanmıştır.

İlk olarak Çelik, Demirgüneş ve Baştuğ (2014) tarafından geliştirilen ODMDR, bütün metin türlerinde kullanılmak üzere tasarlanmıştır. Bu puanlama anahtarı (ODMDR), alanyazında farklı araştırmacılar tarafından kullanıldıkça elde edilen dönütler, farklı türler için ayrı dereceli puanlama anahtarlarının hazırlanabileceği düşüncesini getirmiştir. ODMDR'nin yedi boyutundan biri olan “metin düzenleme” ODBİMDEPA'da korunmuş ancak içeriği bilgilendirici metinler için özelleştirilmiştir. ODMDR'deki “Zaman kullanımı” boyutu kaldırılmıştır; çünkü bilgilendirici metinlerde anlatsal metinlerde olduğu gibi zaman dilimleri açık ve takip edilebilir değildir/ yoktur. ODMDR'de (Çelik, Demirgüneş ve Baştuğ, 2014) aynı boyutta birleştirilerek sunulan “Tutarlık-Bağdaşıklık” boyutu ODBİMDEPA'da “Biçimsel Örüntüler (Bağdaşıklık)” ve “Anlamsal Örüntüler (Tutarlık)” boyutları olarak ikiye ayrılıp daha ayrıntılı hale getirilmiştir. Bunun nedeni anlatsal metinlerde bağdaşıklık ve tutarlığın farklı stratejilerle ve dilsel ifadelerle sağlanıyor olması; bilgilendirici metinlerde bu iki öğenin nasıl okur dostu hale getirilebileceğinin betimlenmesi gerekliliğidir. Yine ODMDR'nin boyutlarından olan “İpucu Verme” ODBİMDEPA'da kullanılmamış, bu boyutun bilgilendirici metin için gerekleri sözü edilen “Biçimsel Örüntüler/Bağdaşıklık” boyutuna aktarılmıştır. ODMDR'deki “Kavramsal Yoğunluk” boyutu kullanılmamış, onun yerine “Bilginin Sunumu” boyutu tasarlanarak bilgilendirici metinlerdeki bilginin sunulma biçiminin nasıl okur dostu hale gelebileceği ayrıntılandırılmıştır. ODMDR'deki “Açıklama” boyutu korunmuş ancak ona “Gerçeklik” boyutu da eklenmiştir; çünkü bilgilendirici metinde gerçeklik algısı okurda güçlü biçimde oluşturulmalıdır. ODMDR'nin son boyutu olan “Söylem” ODBİMDEPA'da “Biçem” boyutu olarak sunulmuştur. ODBİMDEPA için belirlenen “Evensellik” ve “Tartışma” boyutu ise ODMDR'de yer almamaktadır.

İlk versiyonu 2018'te sunulan ve daha sonra, değerlendiricilerden gelen dönütlerle birlikte, ODBİMDEPA'nın boyutları, biçimsel ve içerik açısından daha anlaşılır ifadelerle güncellenmiştir. Böylece hem dereceli puanlama anahtarının geçerlik çalışmalarının ilk aşaması yürütülmüş hem de gelecekteki kullanıcılarına, dereceli puanlama anahtarının boyutlarına göre herhangi bir metni değerlendirme sürecinde nelere dikkat etmeleri gerektiği de daha açık hale getirilmiştir.

Çalışmanın bundan sonraki kısmında, yukarıda anılan iki aşamanın her birinden elde edilen bulgularla ilgili sonuçlar paylaşılacak ve tartışılacaktır.

Bulguların Karşılaştırılması: Kappa ve Rasch Modeli

Fleiss' kappa ve Rasch modeli arasındaki temel fark, değerlendirici değişkenliğini nasıl ele aldıklarıdır. Fleiss' kappa, değerlendiricilerin doğrudan uyumunu ölçmekte ve değerlendiricilerin birbirleriyle tutarlı olması gerektiğini varsaymaktadır. Sonuçlar, değerlendiricilerin düşük uyum gösterdiğini ve dereceli puanlama anahtarlarını uygulamada tutarsız olduğunu göstermektedir.

Öte yandan, Rasch modeli, her bir değerlendiricinin önyargılarını hesaba katarak, yine de **yüksek güvenilirlik** elde etmektedir. Bu durum, değerlendiriciler değerlendirmede tutarlı olmasa bile, Rasch modelinin bu farklılıkları **telafi ettiğini** ve güvenilir kestirimler ürettiğini gösterir.

Çalışmada daha önce de belirtildiği gibi değerlendiriciler arasındaki puan farklarının genellikle 1 puan olması ve geliştirilen dereceli puanlama anahtarındaki puan tanımlayıcılarının soyut özelliklerle ilişkili olmasının değerlendirmeyi zorlaştırması gibi durumlar göz önünde bulundurularak “Çok yüzeyli Rasch Modeli” ile gerçekleştirilen analiz sonuçlarının **kabul edilebilir** olduğuna ve metinlere verilen puanların dikkate değer bir görünüm sunabileceğine karar verilmiştir.

Birinci Belge Setiyle İlgili Sonuç ve Tartışma

2018 yılında MEB tarafından okutulan derslerdeki 5, 6, 7 ve 8. sınıf ders kitaplarından oluşan veri tabanı incelendiğinde, bilgilendirici metinlerin sınıflara göre dağılımlarının dengeli olmadığı görülmektedir. Buna bağlı olarak veri tabanını oluşturan 50 bilgilendirici metinde, metin sınıflamasının yeterli olmadığı/gözden geçirilmesi gerektiği, alanyazındaki alt türlerden (en azından, örnekleyici, açıklayıcı, tartışmacı, savlayıcı...vs) bir ayırımın programda yer alması gerektiği belirlenmiştir.

ODBİMDEPA'nın 2018'deki ilk versiyonunda en iyi işleyen boyutun 2.9 ortalama ile "Evrensellik" ve 2.7 puan ortalaması ile "Metin Düzenleme" boyutu olması, inceleme nesnesi olarak seçilen metinlerin ilgili boyutların alt ulamlarını iyi derecede yansıtmaması ve boyutlara ilişkin içeriğin geçerli olması ile açıklanabilir. Benzer durum, tersinden düşünüldüğünde, "Tartışma" boyutunun en düşük puan ortalamasına (1.3) sahip olması, iki şekilde açıklanabilir. İlki, öğrenciler eğitim ortamlarında tartışma metinleriyle karşılaşırılmamış, bunun sonucu olarak da metnin dünyasından çıkıp gerçek dünyaya atılmaları sağlanamamıştır. İkincisi ise ders kitaplarında "tartışma" boyutunun alt öğeleri, içerik bakımından iyi kurgulanmamıştır; buna göre tartışma boyutunun alt öğeleri araştırmacılar tarafından yeniden belirlenerek güncellenmiş ve ikinci çalışmada "Tartışma" boyutu geçerli puana çıkabilmiştir. En az işleyen ikinci boyut ise 1.9 ile "Biçem" boyutudur. Biçem boyutu için de yine aynı yol izlenmiştir.

İkinci Belge Setiyle İlgili Sonuç ve Tartışma

Çalışmanın ikinci aşamasında, 2023 ve 2024 yıllarında ÖSYM tarafından yapılan YKS testlerindeki Türk Dili ve Edebiyatı, Felsefe, Din Kültürü ve Ahlak Bilgisi, Sosyoloji ve Tarih derslerine ait testlerden rassal olarak seçilen toplam 37 bilgilendirici metin Felsefe, Din Kültürü ve Ahlak Bilgisi, Sosyoloji ve Tarih metinleri için 11 bağımsız değerlendirici, Türk Dili ve Edebiyatı metinleri için 2 değerlendirici daha eklenmiş ve 13 bağımsız değerlendirici tarafından ODBİMDEPA'ya göre puanlanmıştır.

Değerlendiriciler arası güvenilirliği test etmek için ilk önce Fleiss' kappası ve daha sonra çok yüzeyli Rash değerlerine bakılmıştır. Yapılan çözümlenmelerde, Fleiss' kappası değerleri düşük bulunmuştur. Bu bulgu, her ne kadar değerlendiriciler arasındaki uyumun azlığı nedeniyle "değerlendirici eğitimi" ya da "dereceli puanlama anahtarının iyileştirilmesi ihtiyacı"nı işaret ediyor olarak yorumlanabilse de şu durumların da dikkate alınması gerekmektedir: Kappa istatistiği, değerlendiriciler arasındaki şans dışı uyumu ölçmektedir ve uyumun yüksek olması için değerlendiricilerin her bir boyutta aynı puanı vermesi gerekmektedir. Halbuki bir metinle ilgili soyut niteliklere ilişkin, değerlendiricilerin karar vermesinin beklendiği ODBİMDEPA'da verilen puanlar arasındaki farkın da önemli olduğu göz ardı edilmemelidir. Her iki belge setindeki puanlamalarda değerlendiricilerin puanlarının arasında çoğunlukla 1 puanlık bir fark olduğu, puan farkının çoğu boyutta yüksek olmadığı gözlenmiştir. Bu durum Tablo 7'deki örnek puanlama kesitinden görülebilir. Tablo 7'deki puanlamalarda değerlendiricilerin özellikle 1. boyutta çoğunlukla 3 ve 4 puan arasında bir ayrılık yaşadığı görülebilir. Ek 1'de yer alan dereceli puanlama anahtarının 1. boyutundaki 3 ve 4 puan karşılığında açıklamalar incelendiğinde, değerlendiricilerin "Metinde en az 1 temel bilgi/ ana düşünce ve onu destekleyecek en az 1 yan düşünce yer almaktadır." ifadesini karşıladığında hemfikir oldukları ancak "Metnin tüm bölümleri ana ve yan düşünceyi destekler." ifadesini metnin karşılayıp karşılamadığı noktasında farklı düşündükleri çıkarımında bulunulabilir. Dereceli puanlama anahtarı ile nitel tanımlamaların puan karşılıklarını, özellikle bir metinde yer alan "destekleyici düşünceler" gibi soyut özellikleri belirlemenin kolay olduğu söylenemez. Nitekim bu nedenle dereceli puanlama anahtarı kullanımından önce değerlendirici eğitimleri, standardizasyon çalışmaları gibi uygulamalar takip edilir. Çalışmanın amacı puanlama anahtarının geliştirilmesindeki ilk adımları izlemek olduğu için değerlendiricilere sunulan eğitimin içeriği sınırlı tutulmuştur. Değerlendiricilerle puanlama öncesi yalnızca dereceli puanlama anahtarında yer alan boyutlarla ilgili açıklamaların yer aldığı içerikler paylaşılmıştır. Dolayısıyla değerlendiriciler arasındaki uyumsuzluk beklenen bir sonuçtur. Dolayısıyla, düşük bir Fleiss' kappası değeri, değerlendiricilerin dereceli puanlama anahtarı kriterlerini farklı yorumladığını gösterebilir. Ancak bu durum, dereceli puanlama anahtarının kendisinde bir sorun olduğu anlamına gelmez; aksine, yukarıda da belirttiği gibi değerlendiricilerin dereceli puanlama anahtarını nasıl yorumladığı konusunda daha fazla eğitim verilmesi veya netleştirilmesinin sağlanması gerektiğini gösterebilir. Bunların ötesinde, Fleiss' kappası değerleri arasında uyumun düşük çıkması; dilsel anlatımların, bireylerin artalan bilgilerine ve zihinsel şemalarına (schema) göre farklılık gösteriyor

olması, okurların metni farklı bakış açılarıyla değerlendirebilmeleri ve hatta bir okurun kendisinin yaptığı farklı okumalarda bile (aynı) metni farklı değerlendirebilmesi gerekçeleriyle açıklanabilir. Bu bağlamda dilsel bir ürün olarak metinlerin değerlendiricileri arasında uyumun ötesinde, ölçme aracının genel olarak metin üzerinde çalışıp çalışmadığına odaklanmak yerinde olacaktır. Bunun için de Fleiss' kappası değerlerine bakıldıktan sonra veri tabanı üzerinde çok yüzeyli Rasch modeli kullanılarak değerler kontrol edilmiştir.

Kappa değerinin tersine, Rasch modeli, ölçeğin ölçmek istediği yapıyı ne kadar iyi ölçtüğünü göstermektedir. Çok yüzeyli Rasch modeli sonuçları dereceli puanlama anahtarının iyi yapılandırıldığını ve puanlamaların belirli bir zorluk veya kalite ile ilgili olduğunu yansıttığını göstermektedir. Bu model, ölçme aracının tek boyutluluğunu ve maddelerin bu yapı içinde uygun işlev gördüğünü doğrulamaktadır. Çok yüzeyli Rasch modelinden elde edilen değerlendiricilerin tutarlılığını gösteren analiz sonuçları dereceli puanlama anahtarının genel anlamda amacına uygun çalıştığını ve ölçme yapısının sağlam olduğunu göstermektedir ki; bu çalışmanın bulgularında da EAP güvenilirliği 0.951 çıkmıştır. Diğer bir ifadeyle, yüksek EAP güvenilirliği, Rasch modelinin puanlayıcı değişkenliği olsa bile, dereceli puanlama anahtarı kullanımı sürecinde doğru ölçümler yapabildiğini göstermektedir. Dolayısıyla, bilgilendirici ya da herhangi bir türdeki metni okuyup anlama ve değerlendirme sürecinde değerlendiriciler arasında farklarının olabileceği, farklı ders içeriklerine, farklı eğitimlere tabi olmuş öğretmenlerin olduğu varsayıldığında, Rasch modelinin bu dereceli puanlama anahtarının güvenilirliğini kanıtlamak için doğru bir araç olduğu anlaşılmaktadır.

Değerlendiriciler, metinler ve boyutların genel ilişkisini görebilmek için FACET programı aracılığıyla yapılan 3 yüzeyli Rasch analizleri değerlendiriciler arası loigt değerlerinin 0 ile 2 arasında değiştiğini ve değerlendiricilerin 2 ve 3 puanı daha sık kullanma eğiliminde olduğunu göstermiştir. Değerlendiricilerin en düşük ve en yüksek puanları daha az kullanma eğiliminde olmasının nedeninin "güvenli aralıkta" kalma isteği olarak yorumlanmıştır. Ancak YKS'de yer alan metinlerin sözcük sayısının sınırlı olması, bir metnin yalnızca bir kısmının (bir/iki paragraf) metnin bütününden ayrılarak kullanılmasının da bu eğilimde etkisi olabilir. Nitekim dereceli puanlama anahtarında her boyut için en yüksek puan olan 4 puanı gösteren açıklamalar genellikle oldukça kapsayıcı tanımlardan oluşmaktadır. ODBİMDEPA metnin tümünü değerlendirmeye yönelik geliştirilmiştir ve metinden alıntılanan bir bölümün değerlendirilmesinde "**Metnin tüm bölümleri ana ve yan düşünceleri desteklemektedir.**" gibi nitelermeleri YKS'de kullanılan metin parçalarının karşılaması mümkün olmayabilir.

Değerlendiricilerin her birinin bireysel eğilimlerini görebilmek için R programında TAM programı ile 4 yüzeyli Rasch analizleri gerçekleştirilmiştir. Analiz sonuçları çeşitli durumlar ile ilgili yorumlar yapılabilmesini sağlamıştır. Öncelikle farklı konu alanlarındaki YKS metinleri için yapılan puanlamalar incelenmiş ve her birinde EAP güvenilirliğinin yüksek olduğu görülmüştür. Değerlendiriciler arası güvenilirliği gösteren korelasyon değerlerinin ise genel olarak orta düzeyde olduğu, Tarih metinlerinde ise düşük olduğu görülmüştür. Değerlendiricilerin katılık/cömertlik eğilimleri incelendiğinde de genel olarak değerlendiricilerin eğilimlerinin tutarlı olduğu görülmüştür. İki değerlendiricinin katılık/cömertlik eğilimlerinin rassal olduğu fark edilmiş ve bu nedenle söz konusu değerlendiriciler "aykırı (outlier)" olarak kabul edilmiştir.

İkinci belge setinde metinlere verilen toplam puanlar ve yöntem bölümünde belirlenen ODBİMPEYA göre yapılan puanlamalara göre metinlerin sınıflandırılması temelinde ikinci belge setinde yer alan metinlerin çoğunun "okur dostu metin", bir kısmının da "orta düzey okur dostu metin" olduğu görülmüştür. Metinlerin okur dostu olması, okuduğunu kavrama süreçlerinde olumsuz etkiyi azaltmaktadır. Metinlerin okur dostu olması bu metinlere bağlı geliştirilen maddelerin metnin anlaşılabilirliğinden bağımsız olmasını sağlayacaktır. Dolayısıyla "metin anlaşılabilirliğin düşük olması durumundaki ölçme sürecine yansıyan olumsuz etkinin" bulunmadığı değerlendirilmesi yapılabilir. Böylece ölçülmesi hedeflenen yeterliklere başka bir yeterliğin karışmaması sağlanmış olacaktır. YKS'de yer alan farklı disiplinlerin testlerindeki bilgilendirici metinlere ait puan ortalamalarının ve toplam puanların olduğu tablolar incelendiğinde, testlerin genel olarak okur dostu metinlerden oluştuğu saptanmıştır. Bu durumda, metinlerin ölçme aracı olarak kullanılması ve metinlere bağlı maddelerin geliştirilmesi, anlaşılabilirliğin ölçme sürecine etkisini azaltarak geçerliği yükseltebilir. Ancak bu önermenin doğrulanması için öncelikle bu metinlere bağlı maddelerin analizlerinin de incelenmesi gerekmektedir. Ayrıca okur dostu olarak değerlendirilen metinler ile orta düzey okur dostu kabul

edilen ya da okur dostu olmayan metinlere bağlı geliştirilmiş olan maddeler arasında güçlük, ayırt edicilik gibi verilerde fark olup olmadığı da incelenmelidir.

Sonuç olarak, yapılan çözümler sonucunda, ODBİMDEPA için dereceli puanlama anahtarı geliştirme yönteminin ilk adımının tamamlandığı söylenebilir. Puanlama anahtarı geliştirilmiş ve ilk güvenilirlik çalışmaları tamamlanarak değerlendiricilerin iç tutarlılığını sağlayan bir güvenilirlik değeri sağlamıştır. Değerlendiriciler arası uyumun daha yüksek olması için değerlendirici eğitimleri ile puanlama anahtarının kullanılmasında fayda olduğu düşünülmektedir. ODBİMDEPA, metinlerin okur dostu olup olmadığı ile ilgili genel bir sınıflandırma yapmayı sağlayan bir puanlama anahtarıdır. ODBİMDEPA'nın etkin bir puanlama anahtarı olarak öğretim materyallerinde yer alacak metinlerin seçiminde ve ölçme aracı olarak kullanılacak metinlerin seçiminde kullanılması için geliştirilen puanlama anahtarının değerlendirici eğitimleri ile tekrar farklı belge setlerinde uygulanarak güvenilirlik çalışmalarının sürdürülmesi hedeflenmektedir.

Öneriler

2018 yılında MEB tarafından okutulan derslerdeki 5, 6, 7 ve 8. sınıf ders kitaplarından oluşan veri tabanındaki bilgilendirici metinlerin sınıflara göre dağılımlarının dengeli olmamasına rağmen, 2024 yılından itibaren MEB tarafından uygulanmaya konan Türkiye Yüzyılı Maarif Modeli ile güncellenen öğretim programlarına göre güncellenen ders kitaplarındaki bilgilendirici metinlerin sınıflara ve derslere göre dağılımlarının ileri çalışmalarla incelenmesi ve bu kitaplardaki metinlerin ODBİMDEPA ile değerlendirilip seçilen metinlerin ne kadar okur dostu olduğunun belirlenmesi yararlı olacaktır.

Ayrıca 2024 yılından başlayarak 1, 5 ve 9. sınıflarda uygulanmaya başlayan yeni Model'de, ders kitaplarındaki bilgilendirici metin sınıflamasının ne kadar yeterli olduğunun ve alanyazındaki alt türlerden örnekleyici, açıklayıcı, tartışmacı, savlayıcı...vb türlerin bir ayrımının Model'de yer alıp almadığının ileri çalışmalarla araştırılması hem ders kitabı yazarlarına hem de eğitimcilere yol gösterecektir.

ODBİMDEPA'nın ilk versiyonunda incelenen ders kitaplarındaki bilgilendirici metinlerin, özellikle "Tartışma" ve "Biçem" boyutlarında en düşük ortalamaya sahip olmaları nedeniyle, Türkiye Yüzyılı Maarif Modeli ile uygulanmaya başlayan yeni öğretim programlarındaki ders kitaplarında da boyutların ne şekilde işlediğinin araştırılacağı ileri çalışmalara ihtiyaç vardır.

İkinci belge setindeki bilgilendirici metinlerin, güncellenmiş ODBİMDEPA'ya göre puanlanmasının ardından yapılan çözümlerinde her ne kadar Rasch modeli, dereceli puanlama anahtarının tutarlı ölçümler yaptığını göstermiş olsa da düşük Fleiss' kapa değerleri, değerlendiricilerin daha detaylı eğitilmeleri gerektiğini veya dereceli puanlama anahtarının iyileştirilmeye ihtiyacı olduğunu vurgulamaktadır. Dolayısıyla ODBİMDEPA'nın geniş kapsamlı kullanılması durumunda (YKS sorularında, ders kitabı metinlerinde...vb.) değerlendirici eğitime zaman ayrılması gerektiği anlaşılmaktadır. Bu eğitimlerin ve içerik güncellemelerinin takip edilerek alanyazına kazandırılması, bu konularda çalışan araştırmacılara yol gösterecektir.

Ölçme aracı olarak kullanılan metinlerin okur dostu olup olmadığı daha geniş bir veri seti ile incelenerek bu metinlere bağlı maddelerin madde analizleri ile karşılaştırmalı bir çalışma yapılmasına ihtiyaç vardır. Böylece metinlerin anlaşılabilirliğin ölçmeye etkisi ile ilgili sonuçlara ulaşılarak metinlerin seçiminde ne tür standartların belirlenmesi gerektiği ortaya konabilir.

Teşekkür

ODBİMDEPA'nın alanyazına kazandırılması sürecinde gece gündüz demeden metin değerlendiren değerlendiricilerimize gönülden teşekkür ederiz. Ayrıca çalışmada uzman görüşlerine başvuru alan Fahri Yılmaz'a ve Prof. Dr. Nuri Doğan'a emekleri için teşekkürü borç biliriz.

Kaynakça

- Aidinlou, N., Khodamard, N. ve Azami, J. (2012). The effect of textual cohesive reference instruction on the reading comprehension of Iranian EFL students. *International Journal English Linguistics*, 2(5), 18-24. doi:10.5539/ijel.v2n5p18
- Aktaş, M. ve Alıcı, D. (2018). Yazılan hikâyeyi değerlendirmeye yönelik analitik rubrik geliştirme: Geçerlik ve güvenilirlik çalışması. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14(2), 597-610. doi:10.17860/mersinefd.424198
- Beck, I. L., McKeown, M. G., Sinatra, G. M. ve Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 26, 251-276.
- Bond, T. G. ve Fox, C. M. (2015). *Applying the rasch model: Fundamental measurement in the human sciences* (3. bs.). Londra: Routledge.
- Brewer, W. F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. R. J. Shapiro, B. C. Bruce ve W. F. Brewer (Ed.), *Theoretical issues in reading comprehension* içinde (s. 221-239). Hillsdale, NJ: Erlbaum.
- Britton, B. K. ve Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329-404.
- Çelik, T., Demirgüneş, S. ve Baştuğ, M. (2014). Metin oluşturma ve değerlendirme çalışmalarına yönelik "okur dostu" metin değerlendirme rubriği. *International Journal of Eurasia Social Sciences*, 5(14), 65-83.
- Çelik, T., Demirgüneş, S. ve Fidan, D. (2015). Okur dostu metin olma özelliği ile okuduğunu anlama başarısı arasındaki ilişkinin incelenmesi. *Başkent Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 119-127.
- Çıralı Sarıca, H. ve Koçak Usluel, Y. (2016). Eğitsel bağlamda dijital hikâye anlatımı: Bir rubrik geliştirme çalışması. *Eğitim Teknolojisi: Kuram ve Uygulama*, 6(2), 65-84.
- Demirgüneş, F. (2015). *Türkçe ders kitaplarındaki metinlerin okur dostu metin bağlamında değerlendirilmesi* (Yayınlanmamış yüksek lisans tezi). Niğde Üniversitesi, Niğde.
- Dreher, M. J. ve Singer, H. (1989). Friendly texts and friendly teachers. Theory into practice. *Learning Through Text*, 28(2), 98-104.
- Duke, N. ve Bennett Armistead, V. (2003). *Reading and writing informational text in the primary grades: Research-based practices*. New York: Scholastic.
- Eser, M. T. ve Aksu, G. (2022). *R programlama dili ile puanlayıcılar arası uyum ve güvenilirlik* (N. Doğan, Ed.). Ankara: Pegem.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4. bs.). Thousand Oaks, CA: SAGE Publications.
- Fisher, D., Frey, N. ve Lapp, D. (2008). Shared reading: Modeling comprehension, vocabulary, text structures, and text features for older students. *The Reading Teacher*, 6(7), 584-556. doi:10.1598/RT.61.7.4
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382.
- Gray, W. S. ve Leary, B. E. (1935). *What makes a book readable*. Chicago: University of Chicago Press.
- Hambleton, R. K. ve Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- İşeri, K., Çelik, T. ve Demirgüneş, S. (2015). Türkçe ders kitaplarında okur dostu metinlerin incelenmesi. *E-Dil Dergisi*, 1(5), 227-242.
- Kintsch, W. (1982). Aspects of text comprehension. *Advances in Psychology*, 9, 301-312.

- Kintsch, W. (1988). The role of knowledge in discourse processing: A construction-integration model. *Psychological Review*, 95, 163-182.
- Klink, S., Dengel, A. ve Kieninger, T. (2000). *Document structure analysis based on layout and textual features*. Proceedings of International Workshop on Document Analysis Systems, DAS2000.
- Linacre, M. (1993). *Generalizability theory and many-facet Rasch measurement*. Annual Meeting of the American Educational Research Association' da sunulan bildiri, Atlanta, GA.
- Lorch, R., Lemarié, J. ve Grant, R. (2011). Signaling hierarchical and sequential organization in expository text. *Scientific Studies of Reading*, 15(3), 267-284. doi:10.1080/10888431003747535
- Mann, W. ve Thompson, S. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse*, 8(3), 243-281. doi:10.1515/text.1.1988.8.3.243
- McNamara, T. (2001). Language assessment as social practice: challenges for research. *Language Testing*, 18, 333-349. doi:10.1177/026553220101800402
- McNamara, D. S., Kintsch, E., Songer, N. B. ve Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14(1), 1-43. doi:10.1207/s1532690xci1401_1
- McNamara, D. S., Louwrese, M. M., McCarthy, P. M. ve Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes*, 47(4), 292-330. doi:10.1080/01638530902959943
- Millî Eğitim Bakanlığı. (2024). Ortaokul Türkçe dersi öğretim programı (5, 6, 7 ve 8. sınıflar). <https://mufredat.meb.gov.tr> adresinden erişildi.
- Moskal, B. M. ve Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10). https://www.researchgate.net/publication/242637290_Scoring_Rubric_Development_Validity_and_Reliability adresinden erişildi.
- Müldür, M. ve Şimşek, N. D. (2020). Türkçe ders kitaplarındaki bilgilendirici metinlerin ve yapılarının incelenmesi üzerine bir araştırma. *Ana Dili Eğitimi Dergisi*, 8(3), 843-867.
- Nagabhand, S., Nation, I. ve Franken, M. (1993). Can text be too friendly?. *Reading in a Foreign Language*, 9(2), 895-907.
- Rasch, G. (1960). *Probabilistic model for some intelligence and achievement tests*. Copenhagen: Danish Institute for Educational Research.
- Smith, F. (1982). *Writing and the writer*. London: Heinemann.
- Soyuçok, M. ve Kartal, H. (2018). İlkokuma-yazma ders kitaplarının metin oluşturma basamağında oluşturulan metinlerin değerlendirilmesi. *Mehmet Akif Ersoy Üniversite Eğitim Fakültesi Dergisi*, 47, 435-460.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçmelerde güvenirlik ve geçerlilik*. Ankara: Seçkin Yayınları.
- Tuncel, G. (2011). Sosyal bilgiler dersinde rubriklerin etkili kullanımı. *Marmara Coğrafya Dergisi*, 23, 213-233.
- Yıldırım, H. ve Ocak, G. (2016). İlkokul sosyal bilgiler dersi "okur dostu" metinlere yönelik rubrik hazırlama çalışması. *XV. Uluslararası Sınıf Öğretmenliği Sempozyumu bildiri özet kitabı* içinde (s. 422-423). Ankara: Pegem Akademi.
- Yeung, L. (2007). In search of commonalities: Some linguistic and rhetorical features of business reports as a genre. *English for Specific Purposes*, 26(2), 156-179.
- Yıldırım, A. ve Şimşek, H. (2008). *Sosyal bilimlerde nitel araştırma yöntemleri* (6. bs.). Ankara: Seçkin Yayıncılık.

Ek 1

OKUR DOSTU BİLGİLENDİRİCİ METİN DEĞERLENDİRME RUBRİĞİ (ODBİMDER)

Boyutlar	1	2	3	4
Metin Düzenleme	Metinde yalnızca bir bilgi/düşünce okura açıkça anlatılmaktadır. Metinde bu bilgi/düşünce ile çelişecek birden fazla içerik bulunmaktadır.	Metinde birden çok bilgi/düşünce okura açıkça sunulmaktadır. Metinde bu bilgiler/düşünceler ile çelişecek birden fazla içerik bulunmaktadır.	Metinde en az 1 temel bilgi/ ana düşünce ve onu destekleyecek en az 1 yan düşünce yer almaktadır. Metinde, temel bilgi/düşünce ve yan bilgi/düşünceler ile çelişecek içerikler bulunmaktadır.	Metinde en az 1 ana düşünce ve onu destekleyecek en az 1 yan düşünce yer almaktadır. Metnin tüm bölümleri ana ve yan düşünceleri desteklemektedir.
Bilginin Sunumu	Metindeki bilgiler gelişigüzel sıralanmış, öncelik-sonralık ilişkisine dikkat edilmemiştir. Somutlaştırılmamış, açıklanmamış ve örneklendirilmemiş birden fazla bilgi/düşünce bulunmaktadır.	Metindeki bilgilerin/düşüncelerin bazıları sınıflandırılabilir. Somutlaştırılmamış, açıklanmamış ve örneklenmemiş 1'den fazla bilgi/düşünce bulunmaktadır.	Metindeki bilgiler genelden özele-özelden genele, basitten karmaşığa, eskiden yeniye vs. aşamalı bir düzenle sunulmaktadır. Somutlaştırılmamış, açıklanmamış en fazla 1 bilgi bulunmaktadır.	Metinde iletilmek istenen ana bilgi/düşünceyi destekleyen diğer bilgiler/düşünceler aşamalı olarak sunulmuştur ve somutlaştırılmamış, örneklenmemiş bilgi bulunmamaktadır.
Biçimsel Örüntüler/ Bağlılık	Metinde ana biçimsel düzenlemenin (punto, yazı karakteri, satır aralığı... vs.) dışında başka biçimsel düzenlemelere yer verilmemiştir. Metinde, temel biçimsel düzenlemelerde (punto, yazı karakteri, satır aralığı, yazım, noktalama... vb.) eksikler/hatalar bulunmaktadır.	Metinde ana biçimsel düzenlemelerle birlikte, iletilmek istenen temel bilgiyi/düşünceyi işaretleyecek (italik, kalın yazı karakteri... vb.) temel düzeyde işaretlemelere yer verilmiştir. Metinde, temel biçimsel düzenlemeler (punto, yazı karakteri, satır aralığı, yazım, noktalama... vb.) doğru uygulanmıştır. Öte yandan ayrıntılı ve metnin içeriğine göre özgülendirilmiş alt başlıklarımlar, işaretleyiciler vd. uygulanmamıştır.	Metinde ana biçimsel düzenlemelerle birlikte, metne özgü alt başlıklara, bölümlere (kuramsal bilgi, problem durumu, çalışma alanı, problemin çözümü, örnekler, öneriler... vb.) yer verilir. İletilmek istenen ana bilgi ve onu destekleyecek bilgiler/düşünceler farklı işaretleyiciler (kalın, italik... vs. sözcük, dipnot ... vs.) ile düzenlenmiştir.	Metinde ana biçimsel düzenlemelerin etkin kullanımıyla birlikte, metne özgü başlıklarımlara, bölümlendirmelere yer verilmiştir. Dilin tipolojik bağlaşıklık olanakları (adıl eksiltme, art/ön gönderimler... vs.) hatasız kullanılmıştır.

Anlamsal Örüntüler/ Tutarlık	Metnin tümceleri arasında bir bütünlük bulunmamaktadır. Farklı bilgiler birbirini bütünlüyici biçimde sunulmamıştır.	Metnin tümceleri arasında kısmen bir bütünlük bulunmaktadır. Tümceler arasında anlamsal bağlantı kuracak bağlayıcılara (edat, bağlaç...vb.) kısmen yer verilmiştir. Bilgi demetleri arasında kısmen bir bütünlük bulunmaktadır.	Metnin tümceleri arasında bir bütünlük bulunmaktadır. Bağlayıcılara yer verilmiştir, bağlayıcıların kullanılmadığı durumlarda, tümceler arasında anlamsal bütünlük sağlanabilmiştir. Metin, bütüncül olarak ana bilginin çıkarımına kısmen izin vermektedir.	Metnin tümceleri arasında, anlamsal bir bütünlük bulunmaktadır. İletilmek istenen ana bilgiyi destekleyici bilgiler, metnin bütünlüğüne katkı sağlar niteliktedir. Metin, bütüncül olarak ana bilginin çıkarımına izin verir.
Açıklama ve Gerçeklik	Metinde iletilmek istenen ana bilginin/düşüncenin ve diğer bilgilerin (kişiler, yer, zaman, terim, genel bilgi demeti...vb.) açıklamalarına yer verilmemiştir.	Metinde iletilmek istenen bilgiye yönelik sunulan diğer bilgilerin (kişiler, yer, zaman, terim, genel bilgi demeti...vb.) açıklamalarına kısmen yer verilmiştir. Ana bilgiyi anlamayı kolaylaştıracak örnekleme, karşılaştırma, somutlaştırma...vb. yapılmamıştır.	Metinde iletilmek istenen bilgiye yönelik sunulan diğer bilgilerin (kişiler, yer, zaman, terim, genel bilgi demeti...vb.) açıklamalarına örnekleme, karşılaştırma, somutlaştırma...vb. destekleyiciler çoğunlukla kullanılmıştır. Verilen temel bilgi/düşüncenin okurun gerçek yaşam dünyasıyla örtüşen karşılıklarına rastlanmamaktadır.	Metinde iletilmek istenen temel bilgi/düşünce ve destekleyici bilgi/yan düşünce, örnekleme, karşılaştırma, somutlaştırma, farklılıkları / benzerlikleri ortaya koyma, problem ve çözümü, özel sözcüklerin açıklamaları ile birlikte, okurun gerçek yaşam dünyasıyla örtüşecek biçimde sunulmuştur.
Tartışma	Metinde sunulan bilgi/düşünce, aşamalı olarak düzenlenmemiştir. İçerik, okurun belleğinde merak uyandıracak bir tartışma ortamı oluşturulamamıştır.	Metinde sunulan bilgi/düşünce demeti içerik olarak okurda bir merak uyandırmaktadır ve okuru kısmen bir tartışmaya götürmektedir. Tartışmanın çözümüne yönelik öneriler/ıpuçları yer almamaktadır.	Metinde iletilmek istenen temel bilgi ve destekleyici bilgi/ana düşünce/yan düşünce demeti, okurda merak uyandırmaktadır ve okuru, belleğinde bir tartışma yapmaya götürmektedir. Tartışmanın çözümüne yönelik öneriler/ıpuçları yer almamaktadır.	Metinde iletilmek istenen ana bilgi, destekleyici bilgi demeti etkisiyle okurda zihinsel bir tartışma başlatmaktadır. Tartışmaya ilişkin öneriler/ıpuçları okurun çözüme ulaşmayı sağlayacak niteliktedir.

Biçem	Yazar, okurla iletişim kuramamıştır. Bilgiyi ya da düşünceyi yalnızca aktarmış, herhangi bir görüş, bakış açısını özgün ifadelerle ortaya koyamamıştır.	Yazar okurla kısmen iletişim kurmuştur. Yazar kendi deneyimlerini sıkça paylaşmış, “bence”, “aslında”, “bana göre” gibi birden çok öznel yargı kullanarak nesnellikten uzaklaşmıştır. Yazar bilgi/düşüncesini belirtmiş ancak tartışmaya ve başka düşünce ve bilgilere kapalı biçimde sunmuştur.	Yazar okurla iletişim kurmuştur. “Görüldüğü gibi” “öte yandan” gibi ifadelerle bilgi ve düşünceye olan tutumunu görünür kılmıştır. Kendi deneyimlerini değil, gerçek, kanıtlanmış bilgiyi açık bir dille metne aktarmıştır. “Bence”, “aslında”, “görülüyor ki”, “bana göre” gibi öznel yargı bildiren ifadelere kısmen rastlanmıştır.	Yazar okurla güçlü bir iletişim kurmuştur. Bu iletişim, gerçekliği kanıtlanmış örnek/karşılaştırmaların net, yazarın deneyimlerinin “bence”, “aslında”, “görülüyor ki”, “bana göre” gibi öznel yargı bildiren ifadelere gerek duyulmadan kesin bir bilgi/düşünce ve kanaat bildiren göstergelerle gerçekleşmiştir.
Evrensellik	Metinde insan haklarına, hayvan haklarına, sürdürülebilirlik, çevrecilik gibi evrensel duyarlıklara aykırı birden çok durum, olay, dil kullanımı vs. yer almaktadır (dil- din- cinsiyet ayrımı; savaş ve her tür şiddet övgüsü vs.)	Metinde insan haklarına, hayvan haklarına, sürdürülebilirlik, çevrecilik gibi evrensel duyarlıkları sezdirenen bir dil yoktur. Dil- din- cinsiyet ayrımı; savaş ve her tür şiddet övgüsünü içeren az 1 bilgi/düşünce beyanı söz konusudur.	Metinde insan, hayvan ve çevre haklarına vs. aykırı bir durum bulunmamaktadır. Metin farklı kültürlerle, canlı türlerine ilişkin ilgi, merak, sevgi uyandıracak nitelikte değildir. Okurun farklı kültürlerle, ilişkilere ilgi duyacağı içerik açımları yapılmamıştır.	Metinde farklı kültürlerle, canlı türlerine ilişkin ilgi, merak, sevgi uyandıracak öğeler yer almaktadır. İçerik bakımından metin, her kültürde kabul edilir bir içerikle evrensellik değeri taşımaktadır.

ODBİMDER'in Puan Aralıkları

- 25 puan ve üzeri** : Okur dostu metin
- 16-24 puan** : Orta düzey okur dostu metin
- 15 puan ve altı** : Okur dostu olma özelliği taşımayan metin