

EĐİTİM ve İNSANİ BİLİMLER DERĐİSİ

Teori ve Uygulama

Cilt: 10 / Sayı: 20 / Güz 2019

JOURNAL of EDUCATION and HUMANITIES

Theory and Practice

Vol: 10 / No: 20 / Fall 2019

**Üst Düzey OkuduĐunu Anlama Becerileri:
Başarı Testi Geliştirme Çalışması**

High-Level Reading Comprehension Skills:
A Research on Developing an Achievement Test

Makale Türü (Article Type): Araştırma (Research)

Derya YILDIZ
Ayşe Vildan ÜNAL
Muhammet Remzi BAYRAKCI
Mehtap POLAT

Üst Düzey Okuduğunu Anlama Becerileri: Başarı Testi Geliştirme Çalışması

Derya YILDIZ¹

Ayşe Vildan ÜNAL²

Muhammet Remzi BAYRAKCI³

Mehtap POLAT⁴

Öz: Uluslararası düzeyde geçerliği olan PISA sınavlarında OECD tarafından her üç yılda bir 15 yaş grubundaki zorunlu eğitimin sonuna yaklaşan öğrencilerin okuma, matematik ve fen becerileri ölçülmekte ve değerlendirilmektedir. Türkiye 2003 yılından beri katıldığı bu sınavlarda özellikle okuduğunu anlama alanında diğer ülkeler arasında son sıralarda yer almaktadır. Bu durumun nedenleri arasında öğrencilerimizin üst düzey okuduğunu anlama becerilerini ölçecek şekilde hazırlanmış sorularla karşılaşmamları yer almaktadır. Bu sebeple bu çalışmanın amacı üst düzey okuduğunu anlama konusunda öğrencilerin başarısını ölçmek için PISA formatında bir başarı testi geliştirmek olarak belirlenmiştir. İlk olarak üst düzey okuduğunu anlama alanında kazanım listesi oluşturulmuş ve bu kazanımlara uygun olarak 52 soru hazırlanmıştır. Sorular 2017-2018 eğitim-öğretim yılının bahar yarıyılında Konya ilindeki 8., 9., 10. ve 11. sınıfta öğrenim gören toplam 280 öğrenciye uygulanmıştır. Bu ön uygulamadan sonra geçerlik ve güvenilirlik analizleri yapılmış ve sorular yine aynı sınıf düzeylerindeki farklı 280 öğrenciye uygulanmıştır. Madde analizleri ile testin ortalama madde güçlük ve ayırt edicilik indeksleri 0.46 ve 0.45 olarak belirlenmiştir. Testin güvenilirlik analizi KR-20 ve KR-21 teknikleri ile hesaplanmıştır. Kuder Richardson-20 güvenilirlik katsayısı (r) 0.84; KR-21 ise 0.83 olarak bulunmuştur. Yapılan analizler sonucunda üst düzey okuduğunu anlama becerisini ölçen 24 soruluk geçerli ve güvenilir bir başarı testi geliştirildiği belirtilebilir.

Anahtar kelimeler: Üst düzey okuduğunu anlama, başarı testi, geçerlik, güvenilirlik.

Makale Hakkında:

Geliş Tarihi: 31.01.2019; Revizyon Tarihi: 11.05.2019; Kabul Tarihi: 02.07.2019

1) Doç. Dr., Necmettin Erbakan Üniversitesi, A.K. Eğitim Fakültesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Konya, deryacintasyildiz@gmail.com ORCID ID: 0000-0002-5419-8986

2) Arş. Gör., Necmettin Erbakan Üniversitesi, A.K. Eğitim Fakültesi, Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Konya, aysevildanunal@gmail.com ORCID ID: 0000-0001-64386545

3) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya, bayrakciremzi@gmail.com ORCID ID: 0000-0002-6434-3420

4) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya, mehtap8680@hotmail.com ORCID ID: 0000-0002-0212-7748

Giriş

Eğitim, bireyleri bilgilendirmenin yanında, onlara yaşam boyu gerekli olan becerileri kazandırma amacını taşır. Ekonomik İşbirliği ve Kalkınma Teşkilatının (OECD) düzenlediği Uluslararası Öğrenci Değerlendirme Programı – (PISA) eğitimin özellikle beceri edindirme işlevini ölçmek ve değerlendirmek amacıyla yapılan bir uygulamadır. PISA'ya, üç yılda bir 15 yaşındaki öğrenciler katılırlar. 2000 yılında başlayan uygulamaya Türkiye ilk kez 2003 yılında katılmıştır.

PISA'nın amacı, çocukları yakından tanıyarak ilgili yatırımlarda öngörüle bulunmak, gençlerin öğrenmeye karşı motivasyonlarını ve eğitimsel başarılarını belirlemektir. PISA'da öğrencilerin fen bilimleri, matematik okuryazarlığı ve okuma becerilerinin haricinde, motivasyonları, kendileriyle ilgili fikirleri, öğrenme şekilleri, okul ve aile ortamlarıyla ilgili bulgular da elde edilmektedir (OECD, 2016b).

PISA'daki “okuryazarlık” ifadesi, öğrencinin bilgi ve yetkinliklerini geliştirerek topluma etkili biçimde dahil olmasını kapsar. Okuma becerileri kavramı, okuma sırasında yapılan etkinlikleri, yazılı metinleri (ör. basılmış kitaplar, bilgi notları, çevrimiçi ortamlarda, haberlerde) ve bilgiyi bulmak gibi işlevsel yolları içermektedir (MEB, 2016). PISA okuma becerilerinin değerlendirilmesi üç biçimde ele alınır: Metin, okurun metne yaklaşımı ve metnin kullanım amacıdır (Coşkun, 2013). Sorular hem sürekli hem de süreksiz metinler şeklinde olabilmekte öğrencilerden çıkarımlar yapması, benzerlik ve farklılıkları bulması, grafik, şekil ve tablo gibi metinlerle günlük hayatta karşılaşılan durumların ilişkilendirilmesi istenebilmektedir. Kısacası okuma becerileri, öğrencilerin okuduklarını kendi hayatlarında kullanabilme becerisi üzerine odaklanmıştır. PISA okuma becerileri alanında belirlenen yeterli düzeyleri aşağıda verilmiştir:

1a Düzeyi: 1a düzeyindeki öğrenciler, metindeki belirli biçimde ifade edilen bir veya daha fazla bilgiyi tespit edebilir, konu ile ilgili bir metnin ana fikrini ve yazılış maksadını kavrayabilir ya da metindeki bilgilerle farklı bilgiler arasında ilişkilendirmeler yapabilir.

1b Düzeyi: Bu seviyedeki öğrenciler, bildiği bir bağlam hakkında yazılmış bir metindeki açıkça belirtilen bilgiyi tespit edebilir. Bu tarzdaki metinler okura bilginin tekrar edilmesi, görsel veya benzer sembollerin kullanılması türünde veriler sunar. Okuyucu birbiriyle çok ilişkili bilgiler arasında basit bağlamlar kurabilir.

2. Düzey: 2. düzeydeki öğrenciler metinden çıkarım yapabileceği bilgileri tespit edebilir. Ana fikri bulabilir, ilişkileri kavrayabilir veya fazla çıkarımda gerektirmeyen hallerde metindeki bir bölümden anlam çıkarabilir. Metindeki bir özelliğe göre farklılıkları veya benzerlikleri bulabilir.

3. Düzey: Bu seviyede farklı bilgiler arasındaki ilişkiyi belirler ve tanımlar. Ana fikri belirlemek, ilişkileri anlamak için metindeki bilgileri bir araya getirebilir. Öğrencilerin metinler arasında ilişkiler kurarak karşılaştırma yapabilmeleri ve açıklamalarla metinleri incelemeleri ve metnin özelliklerini değerlendirebilmeleri beklenmektedir.

4. Düzey: 4. düzeydeki öğrenciler metindeki bilgileri bulabilir ve gerekli olan verilere karar verip bağlamı oluşturabilir. Metindeki dilsel farklılıkları değerlendirebilir. Metni kavrayabilir ve tanıdık olmayan bağlamlara tasnifleri uyarlayabilir, hipotez kurabilir veya metni eleştirel olarak değerlendirebilir.

5. Düzey: Bu seviyedeki öğrenciler, metnin içindeki örtülü bilgileri bulabilir, metni düzenleyebilir. Metinle ilgili eleştirel değerlendirmelerde bulunabilir, hipotezler oluşturabilir. Farklı bağlamları detaylarıyla anlayabilirler.

6. Düzey: Bu seviyedeki öğrenciler, ayrıntılı biçimde benzerlik ve farklılıkları bulabilir, çıkarım yapabilirler. Metinleri bütüncül biçimde anladığını gösterir ve çok sayıdaki metinden ulaştığı bilgileri birleştirebilir. Önemli bilgileri içeren bir metindeki açıkça ifade edilmemiş kavramları anlayabilir ve soyut kavramları değerlendirebilir. Metindeki ayrıntıları fark eder ve analiz eder (MEB, 2016).

Bu yeterli düzeyleri PISA’da okuma becerisinin temel okuma düzeyinden üst düzey bilişsel becerileri kapsayacak biçimde değiştiğini göstermektedir. Aşağıdaki tabloda yıllara göre Türkiye’nin PISA’daki başarı sıralaması verilmiştir:

Tablo 1: PISA okuma alanında başarı sıralaması

	2003 (41 ülke)	2006 (57 ülke)	2009 (74 ülke)	2012 (67 ülke)	2015 (76 ülke)
	Finlandiya	Güney Kore	Şangay-Çin	Şangay-Çin	Singapur
	543 (1.)	556 (1.)	556 (1.)	570 (1.)	535 (1.)
OECD ortalaması	494	492	493	496	493
Türkiye	441 (33.)	447 (37.)	464 (41.)	475 (42.)	428 (50.)

(OECD-2004; OECD-2007; OECD-2010; OECD-2013; OECD-2016a)

Tablo 1’deki verilere göre 2003’teki PISA’da Finlandiya 1. olurken Türkiye 33., 2006’ daki PISA’da Kore 1. olurken Türkiye 37., 2009’daki PISA’da Şanghay Türkiye 41., 2012’deki PISA’da yine Şanghay 1. olurken Türkiye 42., 2015’teki PISA’da ise Singapur 1. olurken Türkiye 50. olmuştur. Bu veriler okuma alanında Türk öğrencilerin üst düzey becerilerinin diğer ülke öğrencilerinin gerisinde olduğunu göstermektedir.

Aşağıdaki tabloda Türkiye'nin PISA'da okuma becerileri alanında yeterlik sıralaması verilmiştir.

Tablo 2: Türkiye'nin okuma becerileri alanında yeterlik sıralaması

Yıl	1.Düzey altı	1. Düzey	2. Düzey	3. Düzey	4.Düzey	5. Düzey	6. Düzey
2003	12.5	24.5	30.9	20.8	7.7	3.8	
2006	10.8	21.4	31.0	24.5	10.3	2.1	
2009	6.4	18.1	32.2	29.1	12.4	1.8	0.0
2012	5.1	16.6	30.8	28.7	14.5	4.1	0.3
2015	13.2	26.8	32.6	21.1	5.7	0.6	0.0
OECD	6.5	13.6	23.2	27.9	20.5	7.2	1.1

PISA'da Türk öğrencilerin okuma becerileri alanında yeterlik düzeylerinin düşük olması öğrencilerin okuduğunu anlama ve yorumlama alanında sorun yaşadığını göstermektedir. 2009 ve 2015 PISA raporlarına göre altıncı düzeyde Türk öğrenci bulunmamaktadır. Yalnızca 2012 yılında öğrencilerin % 0,3'ü altıncı düzeyde yer almaktadır. Beşinci düzey için de benzer bir durum vardır. PISA 2012'de öğrencilerin % 4.1'i bu düzeydeyken 2015 PISA'da oran % 0.6'ya gerilemiştir (OECD, 2016b).

PISA 2015'te 1. düzey ve altındaki öğrenci sayıları PISA 2009 ile PISA 2012'ye kıyasla fazlaşmıştır. 2015'deki PISA'da alt düzeydeki öğrenci oranı OECD'de %20,1iken ülkemizde %30'dur. PISA 2015'de 5. düzey ile üzerindeki öğrenci sayıları ise PISA 2009 ve PISA 2012'ye nazaran azalmıştır. Üst düzeydeki öğrenci oranı OECD'de %8,3; Türkiye'de %0,06'dır (OECD, 2016a).

Üst düzey okuma becerileri sorgulama, analiz etme ve eleştirel düşünme gibi beceri alanlarını kapsar. Halpern (2003), farklı beceri ve tutumları kapsayan kompleks bir düşünme süreci olan eleştirel düşünme ve sorgulamanın, insanların ifade edilmemiş varsayımların analiz edilmesine, bilgi kaynağının güvenilirliğinin değerlendirilebilmesine ve bir problemin çözümünde en iyi yolun takip edilmesine imkân sağladığını belirtir. Üst düzey okuduğunu anlama becerilerine sahip bireyler bu beceri alanlarına sahip olacaklardır. Bu bağlamda özellikle Türkçe dersinde buna daha fazla önem verilmelidir.

Yıldız'ın (2015) araştırmasında Türkçe dersi sınav kâğıtlarındaki sorular yeniden yapılandırılan Bloom taksonomisine göre incelenmiş ve soruların genellikle bilgi boyutunda kavramsal bilgi basamağında; bilişsel süreç boyutunda da anlama basamağında olduğu ortaya çıkmıştır. Bu durum öğrencilerin karşılaştıkları soruların genellikle alt düzeyde olduğunu göstermektedir. Bu bağlamda ölçme-değerlendirmede kullanılabilecek üst düzey sorulara ihtiyaç olduğu açıktır.

Aşağıda 2018 Türkçe dersi öğretim programında yer alan 8. sınıf kazanımları arasındaki üst düzey olanları sıralanmıştır:

“T.8.1. Dinleme/İzleme: T.8.1.9. Dinlediklerinde/izlediklerinde tutarlılığı sorgular. T.8.1.11. Dinledikleri/izledikleri medya metinlerini değerlendirir. T.8.1.13. Konuşmacının sözlü olmayan mesajlarını kavrar. T.8.3. Okuma: T.8.3.21. Metnin içeriğini yorumlar. T.8.3.22. Metinde ele alınan sorunlara farklı çözümler üretir. T.8.3.23. Metinler arasında karşılaştırma yapar. T.8.3.25. Okudukları ile ilgili çıkarımlarda bulunur. T.8.3.29. Medya metinlerini analiz eder. T.8.3.32. Grafik, tablo ve çizelgeyle sunulan bilgileri yorumlar. T.8.3.35. Metindeki iş ve işlem basamaklarını kavrar. T.8.4. Yazma: T.8.4.5. Anlatımı desteklemek için grafik ve tablo kullanır. T.8.4.6. Bir işi işlem basamaklarına göre yazar. T.8.4.11. Formları yönergelerine uygun doldurur.” Türkçe dersi öğretim programında üst düzey kazanımlar olmasına rağmen öğrencileri PISA sınavında başarılı olamamaları uygulamada sorun olduğunu göstermektedir. Bu bağlamda öğrencilerin okuma alanında üst düzey becerilerini geliştirici uygulamaların yapılması gerekmektedir. Bu uygulamalar için de bu becerileri ölçen başarı testlerine ihtiyaç duyulmaktadır.

PISA'ya katılan ülkeler öğrencilerin bilişsel kapasitelerini ortaya çıkartacak ölçme-değerlendirme araçları geliştirmektedirler. Bu ölçme-değerlendirme araçlarının başında başarı testleri gelmektedir. Başarı testi geliştirirken öncelikle kazanımlar tespit edilir, bu kazanımlara uygun soru geliştirilir, uzman görüşleri alınarak maddeler yeniden yapılandırılır, geçerliği için madde analizi yapılır ve teste son şekli verilir (Akbulut & Çepni, 2013). Bu aşamalarla geliştirilen testin öğretim programı kazanımlarıyla örtüşmesi ve içeriği yansıtması önemlidir. Türkiye'nin üst düzey okuduğunu anlama becerisi test sonuçlarına göre bu alandaki ölçme - değerlendirme çalışmaları önem arz etmektedir (Yalçın, 2011). Alan yazını incelendiğinde başarı testi geliştirmeye yönelik çalışmaların (Akbulut & Çepni, 2013; Ayvaci & Durmuş, 2016; Demir, Kızılay & Bektaş, 2016; Saraç, 2018) fen ve matematik disiplinleri üzerinde yoğunlaştığı görülmektedir. Alan yazında üst düzey okuduğunu anlama becerilerini ölçen bir başarı testi yer almamaktadır. Bu sebeple geliştirilen başarı testi üst düzey okuduğunu anlama becerisinin gelişimine katkı sağlayacaktır.

Bu sebeple bu çalışmanın amacı PISA'daki örneklem grubuna uygun biçimde 8., 9., 10. ve 11. sınıfta okuyan öğrencilerin üst düzey okuduğunu anlama konusunda başarılarını belirlemeye yönelik PISA formatında güvenilirliği ve geçerliği yüksek bir başarı testi geliştirmek olarak belirlenmiştir.

Yöntem

Bu çalışma bir başarı testi geliştirmeyi amaçlamaktadır. Gömleksiz & Erkan (2010), başarı testi geliştirme çalışmalarının; hazırlık, testin uygulanması ve raporlaştırma aşamalarını içerdiğini belirtir.

Çalışma Grubu

2017-2018 eğitim-öğretim yılında Konya il merkezindeki amaçlı örnekleme yöntemiyle seçilmiş 8., 9., 10. ve 11. sınıfta öğrenim gören toplam 280+280 öğrenci çalışma grubunu

oluşturmaktadır. Tanrıoğen (2009), amaçlı örneklemede evrenden araştırmacının amacına en uygun olduğunu düşündüğü bir kümeyi örneklem olarak seçtiğini belirtir. 280 öğrenciye ön uygulama yapıldıktan sonra geçerlik ve güvenilirlik analizleri gerçekleştirilmiş ve sorular yine aynı sınıf düzeylerindeki farklı 280 öğrenciye tekrar uygulanmıştır. Toplam katılımcı sayısı olan 160 öğrencilerin 76'sı kız, 84'ü erkektir. Öğrenci sayıları PISA 2015'teki Türkiye'deki sınıf dağılımına (MEB, 2016) uygun biçimde belirlenmiştir. Buna göre öğrencilerin % 3.7'si (3 öğrenci) 8. sınıf, %21,1'i (17 öğrenci) 9. sınıf, %70'i (56 öğrenci) 10. sınıf, %5'i (4 öğrenci) 11. sınıftır.

Bulgular

Testin geçerlik çalışması

Testin kapsam geçerliğini sağlamak için üst düzey okuduğunu anlama alanında PISA'da üst düzey olan 4, 5 ve 6. düzeylerin amaçlarından hareketle kazanımlar oluşturulmuştur. Bu kazanım listesi için disiplinler arası bir yaklaşımla Türkçe ve Sosyal Bilgiler dersleri başta olmak üzere üst düzey okuduğunu anlama ile ilgili disiplinler incelenmiştir. Aşağıda kazanımlardan örnekler verilmiştir:

- Karmaşık metinlerdeki farklı sınıflandırmaları yapabilir.
- Metinleri eleştirel biçimde değerlendirebilir.
- Benzerlik ve farklılıkları bulabilir, çıkarımlarda bulunabilir.
- Hipotez kurabilir.
- En az iki metin aracılığıyla ulaşılan açıkça belirtilmeyen bilgileri bütünleştirebilir.

Daha sonra ise bu kazanımlara uygun olarak araştırmacılar tarafından PISA'daki okuma alanı yeterli düzeylerinin tümünü kapsayan 52 soru hazırlanmıştır. Hazırlanan 52 soru kapsam geçerliği için alan uzmanlarının incelemesine sunulmuştur. İncelemede doğruluk, belirlenen hedef kazanımlara uygunluk ve öğrenci seviyesi gibi ölçütler göz önüne alınmıştır. Türkçe Eğitimi, Sosyal Bilgiler Eğitimi ve Eğitim Bilimleri alan uzmanlarının görüşleri neticesinde bazı sorular uygun bulunmamış ve soru sayısı 41'e düşürülmüştür.

Testin uygulanması

Başarı testi, 2017-2018 öğretim yılı bahar yarıyılında Konya il merkezindeki amaçlı örnekleme yöntemiyle seçilmiş 8, 9, 10 ve 11. sınıfta öğrenim gören 280 öğrenciye uygulanmıştır. Daha sonra geçerlik ve güvenilirlik analizleri gerçekleştirilmiş ve sorular yine aynı sınıf düzeylerindeki farklı 280 öğrenciye tekrar uygulanmıştır. Testin puanlanması tam puan, kısmi puan ve sıfır puan biçimindedir.

Madde analizi

Madde güçlük ve ayırt edicilik analizi

Madde analizinde öncelikle öğrencilerin aldıkları puanlar azalan biçimde sıralanmış ve böylece % 27'lik alt - üst gruplar belirlenmiş ve madde güçlük indeksleri (P_{jx}) ile madde ayırt

edicilik indeksleri (R_{jx}) hesaplanmıştır. Madde ayırt edicilik indeksi, ölçülen davranışın cevaplayıcılar arasındaki ayırma gücünü belirtmektedir (Özçelik, 2013). Test geliştirmede, maddelerden, ölçülen niteliğe sahip olanları olmayanlardan ayırmaları beklenir. Madde güçlük indeksi 0 - 1 arasında değer alır. 1'e yaklaşma maddenin kolaylaşmasını, 0'a yaklaşma ise zorlaşmasını ifade etmektedir (Turgut & Baykul, 2010).

P_{jx} ve R_{jx} değerleri Tablo 3'te sunulmuştur.

Tablo 3: Testin ön uygulamadaki madde güçlük - ayırt edicilik indeksleri

Soru	P (Madde Güçlük İndeksi)	R (Madde Ayırt Edicilik İndeksi)	Sonuç
1	0.73	0.18	Testten çıkarıldı.
2	0.68	0.27	Testten çıkarıldı.
3	0.2	0.04	Testten çıkarıldı.
4	0.83	0.08	Testten çıkarıldı.
5	0.15	0.08	Testten çıkarıldı.
6	0.47	0.38	Düzeltilerek kullanıldı.
7	0.56	0.21	Testten çıkarıldı.
8	0.46	0.68	Testte kullanıldı.
9	0.31	0.44	Testte kullanıldı.
10	0.26	0.35	Düzeltilerek kullanıldı.
11	0.51	0.03	Testten çıkarıldı.
12	0.66	0.46	Testte kullanıldı.
13	0.25	0.56	Testte kullanıldı.
14	0.68	0.45	Testte kullanıldı.
15	0.47	0.51	Testte kullanıldı.
16	0.3	0.32	Düzeltilerek kullanıldı.
17	0.23	0.17	Testten çıkarıldı.
18	0.36	0.41	Testte kullanıldı.
19	0.8	0.22	Testten çıkarıldı.
20	0.36	0.41	Testte kullanıldı.
21	0.36	0.35	Düzeltilerek kullanıldı.
22	0.31	0.37	Düzeltilerek kullanıldı.
23	0.38	0.08	Testten çıkarıldı.
24	0.42	0.70	Testte kullanıldı.
25	0.36	0.58	Testte kullanıldı.
26	0.43	0.45	Testte kullanıldı.
27	0.50	0.51	Testte kullanıldı.
28	0.16	0.35	Düzeltilerek kullanıldı.
29	0.25	-0.01	Testten çıkarıldı.
30	0.31	0.11	Testten çıkarıldı.
31	0.21	0.18	Testten çıkarıldı.
32	0.25	-0.03	Testten çıkarıldı.
33	0.82	0.31	Düzeltilerek kullanıldı.
34	0.40	0.48	Testte kullanıldı.
35	0.27	-0.04	Testten çıkarıldı.
36	0.27	0.00	Testten çıkarıldı.
37	0.76	0.38	Düzeltilerek kullanıldı.
38	0.55	0.36	Düzeltilerek kullanıldı.
39	0.15	0.02	Testten çıkarıldı.
40	0.31	0.51	Testte kullanıldı.
41	0.19	0.54	Testte kullanıldı.

Madde güçlük indeksinin (P_{jx}); sıfıra yakın olması sorunun zor olduğu, bire yakın olması ise sorunun kolay olduğu anlamına gelmektedir. Bir testin konuyu bilen ve bilmeyenleri değişkenliği tanımlayabilmesi için orta güçlükteki (0.50 civarında) maddelerden oluşması gerekmektedir (Gömleksiz & Erkan, 2010; Tekin, 2016), Tablodaki değerlere göre sorular arasında kolay ve zor sorular da vardır ancak sorular genel olarak orta güçlüktedir. Ortalama güçlük indeksi incelendiğinde ($P_{jx} = 0.46$) testin orta güçlükte olduğu ifade edilebilir.

Madde ayırt ediciliği (R_{jx}); ölçülmek istenen nitelik ile ilgili bilenleri ve bilmeyenleri ayırt etme derecesidir. İndeks 0.40 ve üzerinde olursa, maddenin ayırt etme gücü yüksek, 0.30 - 0.39 arasındaysa maddenin ayırt etme gücü orta ve maddelerde geliştirme yapılması gerektiği, 0.30'un altında ise bu maddelerin testten çıkarılması ya da düzeltilmesi anlamını taşımaktadır (Büyüköztürk, 2012).

Tablo 4: Başarı testinin son uygulamadaki madde güçlük ve ayırt edicilik indeksleri

Soru	P (Ortalama Madde Güçlük İndeksi)	R Ortalama Madde Ayırt Edicilik İndeksi
1	0.68	0.37
2	0.47	0.48
3	0.46	0.76
4	0.31	0.45
5	0.26	0.19
6	0.66	0.61
7	0.25	0.50
8	0.68	0.59
9	0.47	0.54
10	0.30	0.32
11	0.36	0.52
12	0.80	0.29
13	0.36	0.33
14	0.36	0.36
15	0.31	0.31
16	0.42	0.70
17	0.36	0.64
18	0.43	0.57
19	0.50	0.61
20	0.16	0.31
21	0.82	0.43
22	0.40	0.61
23	0.76	0.08
24	0.55	0.30
P _{jx}		0.46
R _{jx}		0.45

Uygulanan testin madde ayırt edicilik indeksine ($R_{jx} = 0.45$) göre testin ayırt etme gücünün yüksektir. Madde analizleri neticesinde testin ortalama güçlüğü 0.46 ve ortalama ayırt etme gücü ise 0.45 olarak bulunmuştur. Bu sonuçlar, testin zorluk düzeyinin ve ayırt edicilik bakımından uygun olduğunu göstermektedir.

Güvenirlilik analizi

Güvenirlilik, testin ölçmeyi amaçladığı niteliği ne kadar doğru ölçtüğüyle ilişkilidir. Güvenirlilik değeri sıfır ile bir arasında farklılaşan bir sayıyla gösterilir. Bu sayının bire doğru yaklaşmasıyla öğrencilerin testten elde ettikleri puanların güvenirliliği de artmaktadır (Gömleksiz & Erkan, 2010).

Öğrencilerin testten aldığı puanların güvenirliliğini belirleyebilecek istatistik teknikleri mevcuttur. KR-20 ve KR-21, doğru cevaba bir, yanlış cevaba sıfır puan verildiği durumlarda kullanılır. KR-20 için testteki maddelerin güçlük indekslerinin bilinmesi gerekir. Atılğan (2013), doğru cevaba bir, yanlış cevaba sıfır puan verilen testlerde kullanılabilir ve KR-20 ile aynı mantıkla işlev gören olarak Cronbach Alfa yönteminin testten öğrencilerin puanlarının güvenirliliğinde kullanılacak bir başka istatistik tekniği olduğunu belirtir. Üst düzey okuduğunu anlama becerileri başarı testi çalışmasında güvenirlilik katsayısı (r), Kuder Richardson-20 yöntemi ve Kuder Richardson-21 teknikleri ile hesaplanmıştır. Kuder Richardson-20 güvenirlilik katsayısı (r) 0,84 olarak bulunmuştur. KR-21 ise 0.83 olarak hesaplanmıştır. Bu değerler öğrencilerin testten aldığı puanların güvenilir olduğunu göstermektedir.

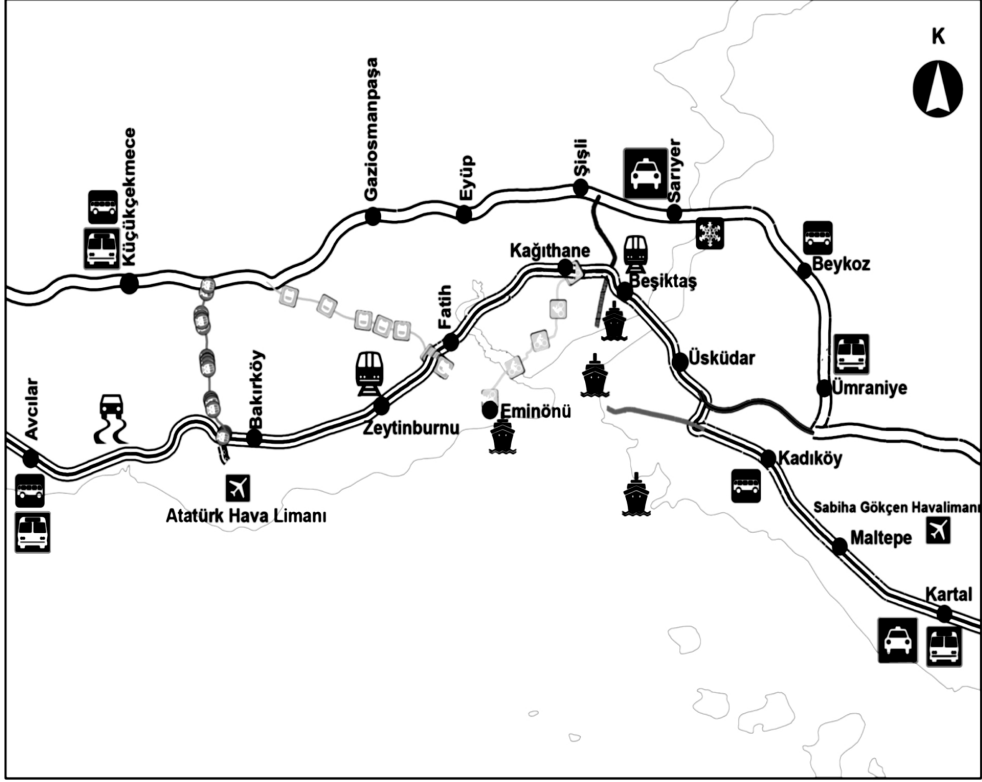
Cronbach Alfa veya güvenirlilik katsayısı 0.40'tan düşükse öğrencilerin testten aldığı puanlar güvenilir değildir. 0.40-0.60 aralığında ise öğrencilerin puanlarının güvenirliliği düşüktür, 0.60-0.90 arasında ise alınan puanlar oldukça güvenilirdir, 0.90'nın üstünde ise yüksek derecede güvenilirdir (Can, 2014). Üst düzey okuduğunu anlama becerileri başarı testi Cronbach alpha güvenirlilik katsayısı 0,837 olarak hesaplanmıştır.

Tablo 5: Teste ait son verilerdeki istatistiksel analizler

Soru Sayısı	24
Uygulananların Sayısı	280
KR-20 (r)	0.84
KR-21 (r)	0.83
Madde güçlüğü (P _{jx})	0.46
Madde ayırt ediciliği (R _{jx})	0.45
Cronbach alpha güvenirlilik katsayısı	0.83

Tablo 5'e göre başarı testinin ortalama güçlük indeksi 0.46, ayırt edicilik indeksi ise 0.45'tir. KR-20 güvenirlilik katsayısı 0.84'tür. Bu değerler sonucunda üst düzey okuduğunu anlama becerileri için geçerli ve güvenilir bir başarı testi geliştirildiği ifade edilebilir.












Aşağıda geliştirilen başarı testinde örnek sorular verilmiştir:



AÇIKLAMALAR

Ana Yollar




ADI

-  Atatürk Havalimanı Caddesi
-  Barbaros Bulvarı
-  Hızlı Erişim Basın Express Yolu
-  Büyükdere Cad.
-  D-100
-  E-5 Otoyolu
-  Otobüs Yolu O-3
-  O-1
-  Vatan Caddesi
-  TEM Otoyolu
-  Bisiklet Yolu Şişli Büyükdere Caddesi

ÖZEL İŞARETLER

-  Metrobüs
-  İETT (İstanbul Elektrik Tren Tramvay)
-  Taksi
-  Uçak
-  Tramvay
-  İDO (İstanbul Deniz Otobüsü)
-  Kaygan Zemin
-  Gizli Buzlanma

Trafik Durumu

-  Çok Yoğun
-  Az Yoğun
-  Akıcı

• Hava Yolu Atatürk Hava Limanı (yurt dışı) ve Sabiha Gökçen Havalimanı (yurt içi) taşımacılığı aracılığıyla yapılmaktadır.

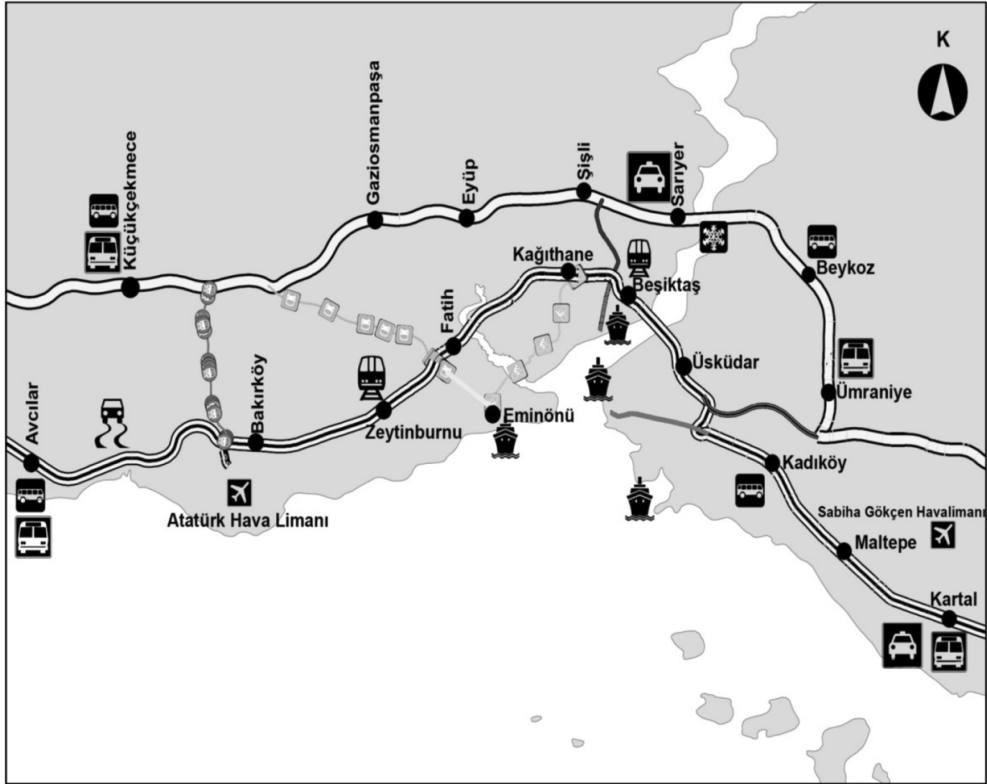
• Metrobüs ve Tramvay hatlarında çift yönlü gidiş-geliş yapılmaktadır.

• İstanbul Deniz Otobüsleri (İDO) Eminönü, Kadıköy, Beşiktaş, Üsküdar'dan çift yönlü gidiş-geliş seferleri düzenlenmektedir.

Soru 1: Küçükçekmece'den Kadıköy istikametine en kısa zamanda gitmek için hangi taşıtlar kullanılmalıdır?

Soru 2: Tem Otoyolu ile E-5 Otoyolu arasında Basın Express ve otobüs yolu açılma sebebi nedir?

Soru 3: Beykoz'dan Kâğıthane'ye giden en kısa rotayı bulmanız gerekmektedir. Gideceğiniz rotayı harita üzerinde çiziniz.



Desibel

Günlük şartlarda insanı rahatsız eden bazı sesler bulunmaktadır. Düşük sesler kulağa az şiddette, güçlü sesler ise yüksek şiddette gelir. Şiddet birimi desibel'dir (dB) . Desibel insan kulağının duyabildiği en küçük ses şiddetidir. Uzmanlar 85 desibel ve üstündeki sese uzun süre maruz kalmanın zararlı olabileceğini belirtmektedirler.

Ses Seviyesi Desibel	Bilinen Sesler
70 dB	Çamaşır makinesi, Bulaşık makinesi, normal tonda konuşma
80dB	Yoğun trafiğe maruz kalma, traş makinesi, elektrikli daktilo
90 dB	Çalar saat, saç kurutma makinesi, çim biçme makinesi
100 dB	Fabrika gürültüsü, kamyon trafiği, metro
110 dB	Kar aracı, çöp kamyonu, gece klübü

Gürültü Düzeyi	İşitme kaybı %		
	5 yıl sonra	10 yıl sonra	20 yıl sonra
dB			
80dB	0	0	0
90 dB	4	10	16
100 dB	12	29	42
110 dB	26	55	78

Soru 1: 2018 yılında kulağındaki şikâyet nedeniyle doktora giden Burak'ın işitme kaybı % 12 olarak belirlenmiştir. Doktor, hastasına ne kadar zamandır işini yaptığını sormuş ve 5 yıl cevabını almıştır. Buna göre söz konusu hasta aşağıdaki seslerden hangisine maruz kalmış olabilir?

A) Normal konuşma B) Kar aracı C) Fabrika gürültüsü D) Çamaşır makinesi

Soru 2: Yukarıdaki verilerden yararlanarak aşağıda verilen durumlardan hangisi doğrudur?

A) 2000 yılında kuaförlüğe başlamadan önce işitme yeteneğini ölçtüren birinin sürekli saç kurutma makinesi kullanması ve 2010 yılında kulağındaki şikâyet nedeniyle tekrar test yaptırdığında işitme kaybı durumu % 16'dır.

B) 2002 yılında işitme yeteneğini ölçtüren birinin işe girdikten sonra sürekli yoğun trafiğe maruz kalması ve kulağından duyduğu şikâyet nedeniyle 2009 yılında tekrar test yaptırdığındaki işitme kaybı durumu %5'tir.

C) 2010 yılında işitme yeteneğini ölçtürecek belediyede çöp kamyonu kullanmaya başlayan birinin 2020 yılında tekrar test yaptırdığındaki işitme kaybı durumu % 10'dur.

D) 2013 yılında bahçivanlığa başlayan birinin 2018'e kadar sürekli çim biçme makinesi kullanması ile birlikte yaşayacağı işitme kaybı %4'tür.

Sonuç, tartışma ve öneriler

Türkiye'nin PISA sınavlarında diğer ülkelere kıyasla oldukça gerilerde yer alması öğrencilerin üst düzey okuduğunu anlama becerisi bakımından büyük eksiklikleri olduğunu göstermektedir. Petscher (2010), üst düzey okuma becerilerinin düşünme süreçleri ile yakından ilişkili olduğunu belirtir. Coşkun (2013) da ders kitaplarındaki metin ve değerlendirme çalışmalarında üst düzey okuma becerilerini geliştirecek etkinliklere yer verilmesi, öğretim programlarının bu verilere göre düzenlenmesi gerektiğini vurgulamıştır. Dolayısıyla eğitim süreçlerinde üst düzey okuma becerisini geliştirici sorulara yer verilmelidir.

Yıldız'ın (2017) araştırmasında üst düzey bilişsel becerileri ölçen açık uçlu maddelere Türk öğrencilerin aşına olmadığı, ders kitaplarındaki soruların genellikle alt okuma düzeyinde olduğu ve Türkçe dersindeki soruların üst düzey düşünme gerektiren okuma becerilerini geliştirmede metin türleri ve ölçme- değerlendirme bakımından yeterli olmadığı belirlenmiştir.

PISA sınavlarında sürekli ve süreksiz metin türleri bir arada kullanılmaktadır. Türkçe ders kitaplarında yer alan metinler ile ölçme ve değerlendirmede kullanılan metinler genellikle sürekli metin yapısındadır. Bu durum öğrencilerin sürekli ve süreksiz metinlerle karşılaşmaları gereğini ortaya koymaktadır.

Bu çalışmanın amacı, PISA sınavlarında öğrencilerin karşılaştıkları üst düzey okuduklarını anlama becerilerini ölçen sorularla paralellik gösteren, geçerliği ve güvenilirliği kanıtlanmış bir başarı testi hazırlamaktır. Bu amaçla test oluşturulurken alan yazındaki benzer çalışmalarda (Fidan, 2013; Güler, 2012; Tosun & Taşkesenligil, 2011) yapıldığı gibi testin amacı ve kapsamının belirlenerek maddelerin yazımı, ön uygulamanın gerçekleştirilmesi analizlerin yapılması ve testin oluşturulması, uygulanması ve puanlanması basamakları takip edilmiştir.

Bu amaçla hazırlanan testte geçerliği ve güvenilirliği sağlamak için bazı basamaklar izlenerek testin amacı ve kapsamı belirlenmiş, kazanımlar oluşturulmuş, kapsam geçerliği için uzman görüşüne başvurulmuş ve belirtke tablosu hazırlanmıştır. Alan yazını incelendiğinde bu basamaklara göre geliştirilen farklı alanlarda başarı testi çalışmalarının (Akbulut & Çepni, 2013; Demir, Kızılay & Bektaş, 2016; Gönen, Kocakaya & Kocakaya, 2011; Hauser & Kingsbury, 2009; Şen & Eryılmaz, 2011) mevcut olduğu görülmüştür. Ancak üst düzey becerileri ölçen bir başarı testine rastlanmamıştır.

Çalışmada, yukarıdaki test geliştirme basamakları sırasında geçerlik ve güvenilirlik çalışmaları da detaylı olarak yapılmıştır. Kapsam geçerliği için öncelikle, belirtke tablosu ve kazanımlar oluşturulmalı, uzman görüşü alınmalıdır (Güler, 2012). Benzer biçimde bu araştırmada da ilkin testin amacı belirlenmiş, üst düzey okuma becerisi ilgili kazanımlar oluşturulmuş ve bu kazanımlara uygun olarak araştırmacılar tarafından 52 soru sorular hazırlanmıştır. Soruların uygunluğunun belirlenmesi için uzman görüşü alınmıştır. Böylece testin kapsam geçerliliği elde edilmiştir. Aksoy (2010), Çalık & Ayas (2003) ve Kınır'ın (2011) çalışmalarında da başarı testinin kapsam geçerliğinin sağlanmasında aynı işlemler kullanılmıştır. Alan uzmanlarının görüşleri neticesinde bazı sorular uygun bulunmamış ve soru sayısı 41'e düşürülmüştür.

Uzman görüşleri neticesinde son haline ulaşan sorular, 2017-2018 eğitim-öğretim yılı bahar yarıyılında Konya ilindeki 8, 9, 10 ve 11. sınıfta öğrenim gören toplam 280 öğrenciye uygulanmıştır. Başarı testindeki bütün maddelerin ortalama madde güçlük ile madde ayırt edicilik indeksleri belirlenmiştir ve sorular yine aynı sınıf düzeylerindeki farklı 280 öğrenciye uygulanmıştır. Ortalama madde güçlük indeksi değeri 0.46 ve ortalama ayırt edicilik indeksi 0.45'tir. Testin ortalama madde güçlük indeksi değerinin 0.50'den küçük olması testin öğrencilere zor, 0.50'den büyük olması ise kolay geldiğinin göstergesidir. Bu sebeple madde ortalama güçlük indeksinin 0.50 seviyesinde olması istenir (Atılğan, 2013; Gömleksiz & Erkan, 2010; Tekin, 2016). Bu çalışmada testin ortalama madde güçlük indeksinin 0.50'ye

yakın olması, testin öğrencilerin düzeyi için ideal olduğunu belirtmektedir. Testin madde ayırt edicilik indeks değerinin 0.40 ve üzerinde olması maddenin ayırt ediciliğinin yüksek derecede olduğu anlamına gelmektedir (Tekin, 2016). Belirtilen değerler, başarı testinin ayırt ediciliğinin yüksek olduğu anlamını taşımaktadır.

Ölçme aracı geliştirmede yukarıdaki hususların dışındaki bir diğer önemli durum ise puanların güvenilir olmasıdır. Testten alınan puanların güvenilirliğini belirlemede en çok kullanılan yöntemler KR-20 ve Cronbach Alpha güvenilirlik katsayısıdır (Ayvaci & Durmuş, 2016; Demir, Kızılay & Bektaş, 2016; Saraç, 2018; Şen & Eryılmaz, 2011; Şener & Taş, 2017; Wilcox, 2003). Yapılan bu çalışmada da Kuder Richardson-20 güvenilirlik katsayısı (r); 0.841; Cronbach alpha güvenilirlik katsayısı 0.837'dir. Güvenirlik katsayısının 0.70'den büyük olması başarı testinden elde edilen puanların güvenilir olduğunu gösterir (Şencan, 2005; Can, 2014). Hesaplanan değer 0.60-0.90 arasında olması testteki puanların güvenilir olmasını ifade etmektedir.

Madde analizi bölümünde geliştirilen başarı testinde maddelerin ayırt edicilik ve güçlük indeksleri ile ortalama madde güçlük ve ayırt edicilik indeksleri belirlenmiştir. Analizden elde edilen veriler sonucunda testin ortalama madde güçlük indeksi 0.46 ve ortalama madde ayırt edicilik indeksi 0.45 olarak bulunmuştur. Bu noktadan hareketle testin ayırt ediciliğinin yüksek olduğu sonucuna ulaşılabılır. Testin ortalama madde güçlük indeksi değerinin 0.50'nin altında olması testin zor olduğu; 0.50'den yüksek olması ise testin kolay olduğu anlamına gelmektedir (Tekin, 2016). Bu değerlere göre öğrencilere uygulanan başarı testinin öğrenciye uygun olduğu söylenebilir. Madde ayırt edicilik indeksinin 0.40 ve yukarısında bir değer alması maddenin ayırt ediciliğinin yüksek olduğu sonucunu vermektedir (Özçelik, 2013). Alan yazın incelendiğinde başarı testi geliştirmeye yönelik çalışmalarda testin 0.40'ın üzerinde olan araştırmalar (Demir, Kızılay & Bektaş, 2016; Şener & Taş, 2017) olduğu görülmektedir.

PISA okuma becerileri soruları metin üzerinden varılacak basit yargılara ulaşmaktan çok bir dilin söz varlığını, gramerini ve metin yapısını bilmeyi gerektirmektedir. Bununla beraber PISA okuma becerilerinde öğrencilerin okuduğunu anlama becerilerine yönelik soruları çözerken doğru stratejiyi kullanılabilirliklerini ve üst düzey becerilerini de değerlendirmektedir. Altı farklı düzeyden oluşan sorularda en zor sorular olarak kabul edilen 5 ve 6. düzey sorular, öğrencilerin birden fazla metinleri bir araya getirmesi, soyut kavramları yorumlaması, hipotez kurabilmesi ve eleştirel değerlendirme yapabilmelerini gerektirir (OECD, 2016a). Çalışmada hazırlanan başarı testi sorularının öğrencinin üst düzey okuduğunu anlama becerisini ölçmeye dayalı olması araştırmanın PISA ile benzer hedeflere sahip olduğunu göstermektedir. Geliştirilen başarı testinin PISA üst düzey okuduğunu anlama becerilerini ölçmeye

yönelik olması çalışmanın gerekliliğini ortaya koymaktadır.

Üst düzey düşünme becerilerinin değerlendirildiği ve dünya genelinde 72 ülkeden 15 yaş grubu öğrencilerin katıldığı bu sınavda ülkemizin almış olduğu sonuçlar Singapur, Finlandiya ve Kore gibi başarılı sonuçlar elde eden ülkelerin çok gerisindedir. MEB tarafından yayımlanan 2016 PISA ulusal raporunda öğrencilerden okuma becerileri sorularında üst düzey sorulara cevap verme yüzdesi 0,05 olarak ifade edilmektedir. Alan yazını incelendiğinde ülkemizin aldığı sonuçların diğer ülkelerle karşılaştırmalı olarak değerlendirilmesi, eğitim sistemimizdeki aksaklıkların giderilmesi için bazı çalışmalar (Batur & Ulutaş 2013; Bozkurt, 2014; Coşkun 2013; Durak & Seferoğlu, 2015; Eraslan, 2009; Güçlü & Yılmaz, 2014; Öznusul & Kaya, 2014; Yalçın 2011) yapılmış ve birtakım sonuçlara ulaşılmıştır. Coşkun (2013), PISA sınavlarındaki okuma ölçütleriyle Türkçe ders kitaplarını karşılaştırarak ders kitaplarının PISA üst düzey okuma becerileri sorularında yetersiz kaldığını ifade etmiştir. Batur & Ulutaş (2013) ise Türkçe dersi kazanımlarının yeniden gözden geçirilerek PISA üst düzey okuma becerileri sorularına uygun hale getirilmesi gerektiğini belirtmişlerdir. Epstein & Kernberger (2006), üst düzey düşünme becerilerinin günümüz dünyasında bütün bireylerin sahip olmaları gereken beceriler arasında olduğunu ifade eder. Erdağı & Toksun (2019) da okunan bir metni tam ve doğru olarak anlama da üst düzey düşünme becerilerinin kullanımına dikkat çeker. Bu bağlamda üst düzey düşünme becerilerinin gelişiminde başarı testlerinin kullanımı önem taşımaktadır.

Alan yazını incelendiğinde PISA üst düzey okuduğunu anlama becerileri üzerine geliştirilen bir başarı testinin mevcut olmadığı görülmektedir. Bu durum yapılan çalışmayı önemli ve gerekli kılmaktadır. Sonuç olarak sürekli ve süreksiz metinlerle üst düzey okuma becerilerini ölçecek PISA'daki okuma alanı yeterlik düzeylerinin tümünü kapsayan alan yazına katkı sağlayacak 24 soruluk güvenilir ve geçerli bir başarı testi geliştirilmiştir.

Kaynakça

- Akbulut, H. İ. & Çepni, S. (2013). Bir üniteye yönelik başarı testi nasıl geliştirilir? : ilköğretim 7. sınıf kuvvet ve hareket ünitesine yönelik bir çalışma. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 18-44.
- Aksoy, M. (2010). *Ortaöğretim kimya dersindeki çözümlülük konusunun kavram haritaları ile öğretilmesinin öğrencilerin başarı ve tutumlarına etkisi*. Selçuk Üniversitesi Eğitim Bilimleri Enstitüsü, Yayınlanmamış Yüksek lisans tezi.
- Ayvacı, H. Ş. & Durmuş, A. (2016). Bir Başarı Testi Geliştirme Çalışması: Isı ve Sıcaklık Başarı Testi Geçerlik ve Güvenirlik Araştırması. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 35(1), 87-102.
- Atılğan, H. (2013). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Batur, Z. & Ulutaş, M. (2013). PISA ile Türkçe öğretim programındaki okuduğunu anlama kazanımlarının örtüşme düzeylerinin incelenmesi, *International Journal of Social Science (JASSS)*, 6(2),1549-1563.
- Bozkurt, Ü. (2014). Development of reading literacy in South Korea from PISA 2000 to PISA 2009. *Education and Science*. 39(173),140-154.
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara. Pegem A Yayıncılık.
- Can, A. (2014). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Ankara: PegemA Yayıncılık.
- Coşkun, Y. D. (2013). Türkçe ders kitaplarının PISA okuma ölçütleri açısından incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 13(26), 22-43.
- Çalık, M. & Ayas, A. (2003). Çözeltilerde kavram başarı testi hazırlama ve uygulama. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 14(2), 1-17.
- Demir, N., Kızılay, E. & Bektaş, O. (2016). 7. sınıf çözeltiler konusunda başarı testi geliştirme: Geçerlik ve güvenirlik çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 10(1).
- Durak, H. Y. & Seferoğlu, S. S. (2015). PISA sonuçlarının sayısal uçurumun göstergeleri açısından karşılaştırılması: Türkiye, Finlandiya ve Güney Kore örnekleri. *Kastamonu Eğitim Dergisi*, 24(1), 1-16.
- Epstein, R. L. & Kernberger, C. (2006). *Critical thinking*. Canada: Thomson Wadsworth.
- Eraslan, A. (2009). Finlandiya'nın PISA'daki başarısının nedenleri: Türkiye için alınacak dersler. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 2(3), 238-248.

- Erdağı Toksun, S. & Toprak, F. (2019). Türkçe öğretmenlerinin okuma stratejileri bilişsel farkındalık becerilerini kullanma düzeylerine ilişkin bir araştırma (Kars ili örneği). *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi*, 8(2), 1137-1157.
- Fidan, E. (2013). *İlkokul öğrencileri için matematik dersi sayılar öğrenme alanında başarı testi geliştirilmesi*. (Yayımlanmamış Yüksek Lisans Tezi), Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Gömleksiz, M. & Erkan, S. (2010). *Eğitimde ölçme ve değerlendirme (2. Baskı)*. Ankara: Nobel Yayın Dağıtım.
- Gönen, S., Kocakaya, S. & Kocakaya, F. (2011). Dinamik konusunda geçerliliği ve güvenilirliği sağlanmış bir başarı testi geliştirme çalışması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 8 (1), 40-57.
- Güçlü, N. & Yılmaz, G. (2014). Uluslararası öğrenci değerlendirme programlarına göre bazı ülkelerin eğitim sistemlerinin incelenmesi. *Eğitim Bilimleri Araştırmaları Dergisi*, 1(4).
- Güler, N. (2012). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi Yayıncılık, Ankara.
- Halpern, D. F. (2003). *Thought & knowledge: An introduction to critical thinking (Fourth Edition)*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Hauser, C. & Kingsbury, G. G. (2009). *Individual score validity in a modest-stakes adaptive educational testing setting. Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego, CA.*
- İdin, Ş. & Aydoğu, C. (2016). Kuvvet ve hareket ünitesi başarı testi geçerlik ve güvenilirlik araştırması. *Eğitim, Bilim ve Teknoloji Araştırmaları Dergisi*, 1(1), 14-33.
- Karabulut, A. (2017). MEB'in yayınlamış olduğu örnek PISA sorularının 2015 Türkçe öğretim programı okuma anlama kazanımları çerçevesinde analizi. *Uluslararası Eğitim Bilim ve Teknoloji Dergisi*. 3(3) . 166 -174.
- Kıngır, S. (2011). *Using the science writing heuristic approach to promote student understanding in chemical changes and mixtures*. ODTÜ, Doktora Tezi.
- MEB (2011). *PISA Türkiye*. Ankara: EğiTek Yayınları.
- MEB (2016). *PISA 2015 ulusal raporu*. Ankara: EARGED Yayınları.
- OECD (2004). A Profile of student performance in reading and science in learning for tomorrow's World: first results from PISA 2003, OECD publishing. 12 Eylül 2018 tarihinde <http://dx.doi.org/10.1787/97892640064016-7en> adresinden erişildi.
- OECD (2007). PISA 2006, science competencies for tomorrow's world, 08 Aralık 2018 tarihinde <http://www.oecd.org/dataoecd/30/18/39703566.pdf> adresinden erişildi.

- OECD (2010). PISA 2009 Results: What Students Know and Can Do - Student Performance in Reading, Mathematics and Science. 12 Ekim 2018 tarihinde <http://www.oecd.org/pisa/pisaproducts/48852548.pdf> adresinden erişildi.
- OECD (2013). PISA 2012 Results in Focus: What 15-Year-Olds Know and What They Can Do with What They Know. 21 Ekim 2018 tarihinde <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf> adresinden erişildi.
- OECD (2016a). *PISA 2015 assessment and analytical framework: science, reading, mathematical and financial literacy*. PISA, Paris: OECD Publishing.
- OECD, (2016b). *PISA 2015 Results (Volume II): Policies and practices for successful schools*. PISA, Paris: OECD Publishing.
- Özçelik, D. A. (2013). *Test hazırlama kılavuzu*. Ankara: Pegem Akademi Yayıncılık.
- Öznusul, M. & Kaya, A. (2014). Türkiye'nin PISA 2009 ve 2012 sonuçlarına ilişkin karşılaştırmalı bir analiz. *Journal of European Education*, 4(1), 23-40.
- Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, 33(4), 335-355.
- Saraç, H. (2018). Fen bilimleri dersi 'maddenin değişimi' ünitesi ile ilgili başarı testi geliştirme: Geçerlik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 416-445.
- Şen, H. C. & Eryılmaz, A. (2011). Bir başarı testi geliştirme çalışması: Basit elektrik devreleri başarı testi geçerlik ve güvenilirlik araştırması. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi*, 8(1), 1-39.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlik*. Ankara: Seçkin Yayıncılık.
- Şener, N. & Taş, E. (2017). Developing achievement test: A research for assessment of 5th grade biology subject. *Journal of Education and Learning*, 6(2), 254.
- Tanrıoğen, A. (Ed.). (2009). *Bilimsel araştırma yöntemleri*. Ankara: Anı Yayıncılık.
- Tekin, H. (2016). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınevi.
- Tosun, C. & Taşkesenligil, Y. (2011). Revize edilmiş Bloom'un taksonomisine göre çözeltiler ve fiziksel özellikleri konusunda başarı testinin geliştirilmesi: Geçerlik ve güvenilirlik çalışması. *Kastamonu Eğitim Dergisi*, 19(2), 499-522.
- Turgut, M. F. & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Pegem Akademi. Ankara.

- Wilcox, R. (2003). *Applying contemporary statistical methods*. San Diego, CA: Elsevier Science.
- Yalçın, K. (2011). Almanya ve Türkiye’de 2003-2006-2009 PISA Araştırma Sonuçlarına Göre Alınan Önlemler ve Uygulamalar. *Türk Kütüphaneciliği*. 25(4) . 494-508.
- Yıldız, D. (2015). Türkçe dersi sınav sorularının yeniden yapılandırılan Bloom taksonomisine göre analizi. *Gaziantep University Journal of Social Sciences*, 14(2):479-497.
- Yıldız, D. (2017). Türk Öğrencilerin PISA Sınavlarındaki Okuma Becerileri. *Uluslararası Türk Dünyası Sempozyumu Bildiri Kitabı*, Ankara: Halk Kültürü Araştırmaları Kurumu Yayınları. 197-205.
- Yıldız-Durak, H. & Seferoğlu, S. S. (2015). PISA sonuçlarının sayısal uçurumun göstergeleri açısından karşılaştırılması: Türkiye, Finlandiya ve Güney Kore örnekleri. *Kastamonu Eğitim Dergisi*, 24(1),1-16.

High-Level Reading Comprehension Skills: A Research on Developing an Achievement Test

Extended abstract

Introduction

PISA (Programme for International Student Assessment) tests having international validity are organized by the OECD every three years and applied to 15-year-old students who come to the end of compulsory education in order to measure and assess their skills in reading, mathematics and science. Our country has participated in these tests since 2003, and our country is one of the lowest-performing OECD countries especially in reading comprehension. It has been reported that not encountering questions that would measure high-level reading comprehension skills of students is one of the reasons for this situation. It is observed that in PISA tests, both continuous and discontinuous texts are included, intertextuality is emphasized, situations which might be encountered in our daily lives are presented in question format through texts given as graphics, shapes and tables which depend on making deductions and include similarities and differences. Therefore, the purpose of this research is to develop an achievement test in PISA format in order to measure the achievement of students in terms of high-level reading comprehension.

Method

First of all, an acquisition list was prepared in terms of high-level reading comprehension. In order to create this acquisition list, disciplines related to high-level reading comprehension, notably Turkish and Social Studies, were examined with an interdisciplinary approach. Next, 52 questions were prepared in accordance with these acquisitions. After asking opinions of experts from Turkish Instruction, Social Studies Instruction and Educational Sciences, some of the questions were not considered appropriate, so the number of questions were decreased to 41. The prepared questions were conducted on 280 students who were studying in grades 8,9,10 and 11 in the spring semester of the 2017-2018 academic year in Konya Province. After this pre-application, validity and reliability analyses of the questions were done and the questions were applied to 280 students at the same grade level again. A table of specifications was prepared for the content validity of the achievement test.

Results

To provide the content validity of the test, the acquisitions were created considering the objectives of 4th, 5th and 6th levels which are the highest levels in reading comprehension in PISA. The prepared 52 questions were presented to the experts for the content validity. While investigating, criteria such as accuracy, suitability to the target acquisitions and student levels

were taken into consideration. Taking the opinions of the experts from Turkish Language Instruction, Social Studies Instruction and Educational Sciences into consideration, some of the questions were not found appropriate, and the number of the questions was decreased from 52 to 41.

For the pilot scheme, the achievement test was conducted on high school students in order to determine comprehensibility of the items and their suitability to the levels and necessary corrections were made in accordance with the obtained feedbacks. The power of distinctiveness of the conducted test is high according to the item distinctiveness index ($R_{jx} = 0.45$). As a result of item analyses, the average difficulty of the achievement test was found as .46 and the average power of distinctiveness was found as 0.45. These results show that the achievement test is appropriate in terms of its difficulty level and distinctiveness. In the conducted item analysis, the average item difficulty and distinctiveness indices of the test were calculated as 0.48 and 0.46 respectively. Kuder Richardson-20 technique was used in the reliability analysis of the achievement test. Kuder Richardson-20 reliability coefficient was detected as (r) 0.84. As a result of the conducted analyses, a valid and reliable achievement test measuring high-level reading comprehension skills and including 24 questions has been developed.

Discussion and conclusion

In consequence of the conducted analyses, it is likely to say that a valid and reliable achievement test has been developed for high-level reading comprehension skills in the research. Achievement tests are significant for students in terms of providing quick feedback to the system. In this concern, increasing the number of studies related to reading comprehension skills oriented to PISA test will show current faults in the educational system of our country and make it possible to correct them. Different achievement tests might be developed in other domains of PISA as it is developed in the domain of high-level reading comprehension skills.