

Okuduğunu Anlama Testinin Geliştirilmesi

Hakan ÜLPER¹, Gökhan ÇETİNKAYA², Nihat BAYAT³

Geliş Tarihi: 19.02.2016

Kabul Ediliş Tarihi: 31.10.2016

ÖZ

Bu çalışmanın amacı ortaokul öğrencilerinin okuduğunu anlama durumunu ölçmek için geçerli ve güvenilir bir ölçüm aracı geliştirmektir. Testi geliştirme sürecinde, kuramsal çerçeve bağlamında kavrama ve soru türü temelinde test maddeleri oluşturulmuştur. Test son biçimini almadan önce uzman görüşleri alınmış, yirmi öğrenciyle pilot çalışması yapılmış ve kapsam geçerliliği için test belirtke tablosu oluşturulmuştur. Okuma-Anlama Testi (OAT) son halini aldıktan sonra 2014-2015 Eğitim-Öğretim yılında Gaziantep, Burdur, Niğde, Ordu, Urfa ve İstanbul illerinde bulunan 7 devlet okulunun 5,6,7 ve 8. sınıfına devam eden 696 öğrenciye uygulanmıştır. İstatistiksel analizlerde Lawshe Tekniği ile kapsam geçerliliği 1.00 olarak hesaplanmıştır. Anlama testinde yer alan soruların iç tutarlılığını belirlemek için Cronbach's Alfa ve Sperman Brown iki yarı test test güvenilirliği hesaplanmıştır. Yapılan analizler sonucunda, test sonuçlarının güvenilirlik katsayısı (Cronbach's Alpha) 0,744 ve (Sperman Brown) .802 olarak bulunmuştur. Bu analizlere ek olarak okuma testi maddelerine %27'lik alt-üst grup karşılaştırmaları yapılmıştır. Bu açıdan, ölçeğin iç geçerliliğinin tespiti için bağımsız gruplar için t-testi uygulanmıştır. Ayrıca ölçekteki maddelerin madde güçlük ve madde ayırt edicilik değerleri hesaplanmıştır. Yapılan madde analizi (ITEMAN) ile testin ortalama madde güçlüğü 0,507 ve ortalama madde ayırt ediciliği 0,431 olarak hesaplanmıştır. Bu sonuçlar geliştirilen OAT ile geçerli ve güvenilir sonuçlara ulaşıldığını göstermektedir.

Anahtar kelimeler: Okuma-anlama, Türkçe eğitimi, test geliştirme, geçerlik, güvenilirlik.

Developing Reading Comprehension Test

ABSTRACT

The aim of this study is to develop a valid and reliable measurement tool to measure the status of middle school students' reading comprehension. In the test development process, test items were established according to types of comprehension and questions in the context of theoretical framework. Expert opinions were taken before the test was finalized. Pilot study was conducted with twenty students and table of test specifications were created for content validity. After the final version of Reading Comprehension Test was formed, it was fulfilled on 696 students who studied at 5th, 6th, 7th and 8th classes of governmental schools in Gaziantep, Burdur, Niğde, Ordu and Istanbul in 2014-2015 academic year. Item difficulty and item discrimination values were calculated with statistical analyze. Cronbach's Alpha was calculated to determine the internal consistency reliability of questions in comprehension test. According to the analyzes, the reliability coefficient (cronbach's alpha) of the test results was found to be 0,744. According to item analyzes (ITEMAN), the average item difficulty of the test is calculated to be 0,507 and the average item distinctiveness of the test is calculate to be 0,431. These results indicate that valid and reliable results were gained with the developed test.

¹ Doç. Dr. MAKÜ, Eğitim Fak., hakanulper@gmail.com

² Doç. Dr. Niğde Üniversitesi, Eğitim Fak., gokhancetinkaya76@hotmail.com

³ Doç. Dr. Akdeniz Üniversitesi, Eğitim Fak., nihatsbyat@gmail.com

Keywords: Reading-comprehension, Turkish language education, test development, validity, reliability.

GİRİŞ

İşlevsel bir okuduğunu anlama testi geliştirebilmek için nesnenin yani okuma-anlamanın doğru ve tüm yönleriyle tanımlanabilmesi gerekmektedir. Alan yazında okuduğunu anlamayı tanımlayıcı çok sayıda kuramın varlığı dikkat çekmektedir. Bu kuramlara göre artık günümüzde okuma tek boyutlu olarak değil farklı düzeylerde gerçekleşen zihinsel işlemlerden oluşan çok boyutlu bir süreç olarak kabul edilmektedir. Bu kuramlar içinde van Dijk ve Kintsch (1983) tarafından ileri sürülen kuramsal açıklamalar okuduğunu anlamayı çok yönlü bir biçimde tanımlaması bakımından dikkat çekmektedir. Bu bağlamda okurken dilsel, kavramsal ve durumsal olmak üzere üç ayrı düzeyde zihinsel gösterimler gerçekleşir. Dilsel düzeyde diğer bir deyişle yüzey yapıda sözcükler ve sözcük öbekleri arasındaki dilsel ilişkiler, kavramsal düzeyde sözcüklerin ve tümcelerin anlamı, durumsal düzeyde ise okurun genel bilgisiyle etkileşime giren metnin anlamına ilişkin zihinsel gösterimler vardır (aktaran Tracey, Morrow, 2012).

Metnin anlamsal ve retorik yapısıyla ilgili olan düzey, kavrama sürecinde oldukça önemli bir yer tutan bir ilişki ağı sunar. Bu düzeye metin tabanı (textbase) denir. Bu düzey daha derin bir anlamlandırma gerektiren durum modelinden ayrılır. Durum modelinde metindeki bilgi okurun ön bilgisi ile detaylandırılır ve onunla bütünleştirilir. Bu bakımdan bu üçüncü düzeyde okur, ön bilgileriyle metindeki bilgileri etkileşime sokarak üst düzeyde çıkarımlar yapmak zorundadır. Bu üç düzeye ilişkin bir yapılandırma gerçekleştirebilmek için bir yandan yerel (local) düzeyde diğer yandan ise bütüncül (global) düzeyde zihinsel işlem yapmak ve artalan bilgilerini metinle etkileşime sokmak gerekir. Yerel düzeydeki işlemlerde tümceler anlamsal olarak birbirleriyle ilişkilendirilirken bütüncül düzeydeki işlemlerde ise metnin bütüncül düzeydeki anlamını kurmak, anadüşünceye ulaşmak, metni özetlemek ve metne uygun bir başlık bulmak gerekir (van Dijk&Kintsch, 1983; Kintsch, 1994; 2004; Kintsch & Kintsch, 2005).

Yazılı metnin kavranmasıyla ilişkili her bir işlem farklı düzeyde bilişsel işleme ve farklı bilgi kaynaklarının kullanımını gerektirir (Basaraba, Yovanoff, Alonza ve Tindal, 2013). Bu bakımdan alan yazında metni kavrama genellikle yüzey (literal) ve çıkarım olmak üzere ikiye ayrılır. Yapı-bütünleme modelinde metni kavrama düzeyleri alışlagelen yüzeysel kavrama ve çıkarımsal kavrama ayrımı daha da genişletilmiştir. Özellikle çıkarımla ilgili olarak yapılan sınıflama dikkat çekmektedir. Kintsch (2004) çağırma-ekleme (retrieval) ve üretim (generation) olmak üzere çıkarımı iki ayrı ulamda tanımlamıştır. Çağırma-ekleme uzun süreli bellekteki metinle ilgili ön bilgiyi çağırıp metne eklemeye işaret eder. Üretim ise metinden yola çıkarak yeni bilgi üretmeye işaret eder.

Konuya ilişkin birçok açıklamada yer alan yaygın bileşen, metnin okuyucunun belleğinde bağdaşık bir biçimde yapılandırılmasını gerektiren bilişsel gösterimi içerir. Bu bilişsel gösterim okuyucunun metni anımsama, metinle ilgili soruları yanıtlama, metindeki bilgiye erişme ve başvurma gibi farklı değerlendirme görevlerinde işlettiği bir alt yapıdır (Kendeou, Papadopoulou ve Spanou, 2012).

Bu kuramsal bilgiler çerçevesinde kavramaya ilişkin hazırlanacak olan soruları metin tabanı ve çıkarım başlığı altında toplamak mümkündür. Metin tabanı ile ilgili sorular metindeki ana nokta ya da ayrıntı niteliğinde özgül (spesifik) bilgiyi hedefler. O nedenle bu nitelikteki sorular bellekteki ya da metindeki ilgili bilgiye ulaşarak yanıtlanır. Çoktan seçmeli, boşluk doldurma ya da eşleştirme türündeki sorular bu nitelikteki sorular için uygun soru türleridir. Bunun yanında özetleme de metin tabanlı bir sorudur. Fakat iyi yapılandırılmış yanıtların oluşturulmasını gerektirir. Bu bakımdan özetleme türü sorular birinci düzey sorulardır. İkinci düzeydeki sorular ise kısa yanıtli sorulardır. Üçüncü düzey sorular ise 5N 1K türü spesifik bilgiyi yoklayan ve yanıtları bir sözcük öbeği ya da bir tümceden ibaret olan sorulardır (Kintsch, 2005b).

Çıkarım gerektiren sorular ise metinde açıkça verilen bilginin ötesine geçmeyi, var olan bilgileri metindeki bilgiler aracılığıyla yorumlamayı, metindeki farklı tümceler ve düşünceler arasında yeni bağlantılar kurmayı, metindeki bilgiyi yeni bir probleme uygulamayı gerektirir (Allen, 1985; Cain & Oakhill, 1998; Kintsch, 2005b). Dört farklı türde gerçekleştirilebilir. Birincisi metindeki bilgiyle artalan bilgisini ilişkilendirerek ya da metnin farklı bölümlerindeki bilgiyi bağlayarak; ikincisi metinde açıkça sunulmamış kavramları, süreçleri, nedenleri ve sonuçları açıklayarak; üçüncüsü metinde daha sonra gelebilecek olan bilgileri tahmin ederek; dördüncüsü ise varsayımlar üzerine tahminlerde bulunarak. Bu nitelikteki sorular çoktan seçmeli soru türünde de hazırlanabilir, ancak genel olarak çıkarım gerektiren sorular açık uçlu soru türünde olur (Kintsch, 2005b).

Çıkarım gerektiren sorular diğer sorulara göre daha zordur ve daha çok bilişsel işleme yapmayı gerektirir. Bu bakımdan yetkin olmayan okuyucular açısından yanıtlanması daha zor olan sorular olarak değerlendirilebilir. Oakhill, (1984) yapmış olduğu deneysel çalışmada bu durumu ortaya koymuştur. Çocuklara kısa bir hikâyeye okutup ardından çıkarım gerektiren soruların yanıtlanmasını gerektiren bu çalışmada Oakhill, yetkin olmayan okurların çıkarım gerektiren soruları yanıtlamakta yetersiz kaldıklarını bulmuştur. Benzer bir bulguya Bowyer-Crane ve Snowling (2005) tarafından yapılan çalışmada da ulaşılmıştır. Bu sonuçlar yetkin ve yetkin olmayan okuyucuları belirlemede çıkarım gerektiren soruların da kullanılmasının ne denli gerekli olduğuna işaret etmektedir.

Okuduğunu anlama becerisini ölçmek için hazırlanacak olan testlerde farklı soru türlerinden ya da farklı soru tekniklerinden yararlanmak gerekmektedir. Çünkü her soru tekniğinin kendi açısından hem olumlu hem de olumsuz özellikleri bulunmaktadır. Bununla birlikte okuduğunu anlama sürecine ilişkin her soru

teknığının odaklandığı ve odaklanmadığı yerler bulunmaktadır. Yani farklı soru teknikleri okuduğunu anlama sürecinin farklı görünümünü ölçmektedir. Bununla birlikte her bir teknik neyi ölçmeye olanak sağlıyorsa ölçme işlemi onunla sınırlı olacaktır (Bkz. Ülper, 2010; Alderson, 2000). Diğer yandan Keenan ve Betjeman (2006) bir metne bağlı olan ve olmayan soruların karşılaştırmalı olarak bir incelemesini yaptıkları araştırmalarında metinden bağımsız olarak hazırlanan soruların okuma yetersizliklerini ölçmek amacıyla kullanılmasının uygun olmadığını bulgulamışlardır. Bu bağlamda bir kavrama testinde bulunacak olan soruların çeşitlilik göstermesi ve bir metne bağlı olarak hazırlanması gerektiği açıktır.

Bu ve buna benzer pek çok okuma kuramındaki bilgilerin benzerlerinin taksonomilerde de karşımıza çıktıkları görülmektedir. Birçok taksonomi okuduğunu anlamayı farklı bilişsel işleme basamakları içeren ulamlarla sınıflandırır (Barrett, 1976; Catala, Catala, Molina&Monelus, 2001). Bu taksonomilerde farklı sınıflandırmalar yer almakla birlikte, ortak noktalar da bulunmaktadır. Catala'ya (2001; akt. Cadime vd., 2013) göre bu taksonomiler dört noktada ortaklık taşımaktadır. Bunlar; yüzey kavrama (literal comprehension), çıkarımsal kavrama (inferential comprehension), yeniden düzenleme (reorganisation) ve eleştirel kavrama (critical comprehension) olarak belirtilmektedir. Öğrencinin başarısı konusunda geçerli ve güvenilir geribildirim sağlaması açısından testlerin farklı kavrama türlerini içermesi gerekir (Basabara vd., 2013).

Okuduğunu anlama becerisi ile ilgili Türkçe eğitimi alanında farklı sınıf düzeylerinde yapılmış birçok bilimsel çalışmanın var olduğu görülmektedir. Son dönemlerde yapılmış olan bazı çalışmalar örnek oluşturması amacıyla tarihi sıraya göre aşağıda verilmiştir.

Çelik, Demirgüneş ve Fidan (2015), belirli ölçütlere göre okuma uygun metinler olarak tanımladıkları metinlerin, okuduğunu anlama başarısındaki rolünü araştırdıkları çalışmalarında, kendileri tarafından geliştirdikleri okuduğunu anlama testini kullanmışlardır. Bu test, Tübitak Bilim Çocuk ve Doğan Kardeş dergilerinden seçilen ikişer metinden oluşmaktadır. Metinleri okuyan öğrencilerin okuduğunu anlama düzeylerinin belirlenmesi amacıyla Barrett taksonomisinin aşamalarına karşılık gelecek biçimde her metin için beşer soru hazırlanmıştır.

Ekrandan okumanın okuduğunu anlamaya etkisini inceleyen Yaman ve Dağtaş (2014) 8. sınıflar ile gerçekleştirdikleri bu araştırmalarında öğrencilerin okuduğunu anlama becerilerini ölçmek için kendileri bir okuma-anlama testi geliştirmişlerdir. Bu test 3'ü öyküleyici, 2'si öğretici ve 3'ü de kurmaca olmak üzere toplam 8 metinden oluşan bir testtir.

Baştuğ ve Kaman (2013) nörolojik etki yönteminin öğrencilerin akıcı okuma becerileri ve okuduğunu anlama başarıları üzerindeki etkisini incelemek

amacıyla yaptıkları çalışmalarında öğrencilerin anlama becerilerini bir diğer araştırmacı tarafından geliştirilen bir test ile ölçmüşlerdir. Bu test de yine açık uçlu sorulardan oluşan bir testtir.

İlkokul öğrencilerinin akıcı okuma becerilerinin okuduğunu anlamayı yordama düzeyi ve akıcı okuma becerileri ile okuduğunu anlama arasındaki ilişkiyi inceledikleri çalışmalarında Baştuğ ve Akyol (2012) anlamının değerlendirilmesi için sesli okutulan metinlere ait dördü yüzey metin dördü ise derin metne özgü sekiz açık uçlu soru kullanmışlardır.

Yıldız ve Akyol (2011) okuma motivasyonunun okuduğunu anlamayı doğrudan ve okuma alışkanlığı aracılığıyla dolaylı olarak etkileme durumunu inceledikleri bir araştırma yapmışlardır. Bu çalışmada öğrencilerin okuduğunu anlama düzeylerini belirlemek için 353 kelimedenden oluşan hikâye edici bir metin ile 295 kelimedenden oluşan bilgilendirici bir metin kullanılmışlardır. Bu metinlerden yola çıkarak hazırladıkları sorular Türkçe Öğretim Programı'ndaki okuduğunu anlama kazanımlarına dayandırılmıştır. Soru tipi olarak çoktan seçmeli sorular kullanılmıştır.

Hızlı okuma tekniklerinin ilköğretim sekizinci sınıf öğrencilerinin okuma hızlarına ve okuduğunu anlama düzeyine etkisini araştıran Dedebali ve Saracaloğlu (2010) öğrencilerin anlama becerilerini farklı metin türlerine göre hazırladıkları açık uçlu sorulardan oluşan bir test yoluyla ölçmüşlerdir.

Metinlerarası okumanın okuduğunu anlamaya etkisini belirlemek amacıyla yaptıkları çalışmalarında Köksal ve Ünal (2008), beşinci sınıf öğrencilerin anlama becerilerini ders kitaplarındaki metinlerden Bloom taksonomisine göre hazırladıkları yazılı yoklama sorularıyla ölçmüşlerdir.

Bu çalışmalara bakıldığı zaman okuduğunu anlama boyutunun ölçülmesinde kullanılan kavrama testlerinin genellikle çalışmayı yapan araştırmacılar tarafından o çalışma için geliştirildiği ya da daha önce geliştirilmiş olan bazı testlerin kullanıldığı görülmektedir. Burada kullanılan testlerin genel özellikleri, geliştirilme sürecinde çok dar bölgeden sınırlı sayıda bir örneklem üzerinden çalışılması ve bu yüzden ülke genelinde uygulanabilirliğinin olmamasıdır. Bununla birlikte kavrama türleri bakımından kavramın tek boyutuna odaklanılmış olması ve tek tip soru türü ile kavramanın ölçülmüş olması hatta bazı testlerin geçerlik ve güvenilirlik çalışmalarının bile yapılmamış olması önemli bir eksiklik olarak dikkat çekmektedir. Bu durum alan yazında okuduğunu anlama becerisinin ölçümünde tüm ülke çapında kullanılabilecek olan standart bir ölçme aracına duyulan gereksinime işaret etmektedir.

Bu çalışmada, ortaokul öğrencilerinin okuduğunu anlama durumunu ölçme amacıyla kullanılabilecek Okuduğunu Anlama Testi'nin (OAT) geliştirilmesi ve psikometrik niteliklerinin belirlenmesi amaçlanmıştır.

YÖNTEM

Araştırmanın Deseni ve Çalışma Grubu

Tarama modeline dayalı olarak yürütülen bu araştırma, Antalya, Burdur, Niğde, Ordu, İstanbul ve Urfa'da ortaokul 5, 6, 7 ve 8. sınıf düzeyinde öğrenimine devam eden 696 öğrenci üzerinde gerçekleştirilmiştir.

Tablo 1. Katılımcıların Sınıflara Göre Dağılımı

Sınıflar	Kız	Erkek	Öğrenci sayısı	Yüzde (%)
5	94	94	188	27,0
6	84	92	176	25,3
7	88	77	165	23,7
8	86	81	167	24,0
Toplam	352	344	696	100

Testin Geliştirilme Süreci

Araştırmacılar, okurların nasıl çıkarım yaptıklarını, tutarlılık ilişkilerini nasıl izlediklerini anlamak için okuma-anlama süreçlerini değerlendirirler. Bununla birlikte metnin zihinsel gösterimini değerlendirmek için okuma-anlama çıktılarını ölçerler. Okuma-anlama çıktıları okurun metinden ne öğrendiği ya da belleğine metinden ne aldığı ile ilgilidir. Okuma çıktıları anımsama, soru sorma ve çoktan seçmeli testler yoluyla değerlendirilir (Carlson, Seipel, McMaster, 2014).

Amaca bağlı olarak okuduğunu anlamayı değerlendirmek çok boyutlu bir görünüm sunar. Örneğin çıkartmalı test, çoktan seçmeli test ya da açık uçlu vb. soru türleri farklı olabilir; kağıt ile ya da bilgisayarla sunum biçimi farklı olabilir ve yüzeysel, çıkarımsal ve ana düşünce gibi okuma-anlama bileşenleri üzerine olabilir (Eason&Cutting, 2009). Bu bağlamda test hazırlama sürecinde tüm bu boyutların olumlu ve olumsuz yönleri üzerinde durmakta yarar vardır.

Kuramsal çerçeve bağlamında okuduğunu anlama testi için başlangıçta 50 soru hazırlanmıştır. Bu sorular alan uzmanlarından görüş alınarak 30'a indirilmiştir. Sayının 30'a indirilmesinde belirleyici olan etmenler soru türleri, kavrama türleri ve bir ders saatinde uygulanabilirlik olarak belirtilebilir. Hazırlanan taslağın son halinde bulunan otuz sorunun sekizi açık uçlu; on biri çoktan seçmeli; üçü boşluk doldurma ve sekizi de doğru yanlış türü sorulardan oluşmaktadır. Bu sorulardan dokuzu çıkarım, sekizi önermeler arası ilişkileri, üçü ana düşünce, sekizi yüzeysel metin ve ikisi de başlıkla ilgilidir. Taslak ölçme aracı bu son haliyle uygulamaya sokulmuştur.

Tablo 2. Okuma-Anlama Testinde Yer Alan Soruların Dağılımı

Soru alanı	Soru Türü	Soru Sayısı
Çıkarım	Açık uçlu	3
	Çoktan seçmeli	1
	Boşluk doldurma	
	Doğru-Yanlış	5

Önermeler arası ilişki	Açık uçlu	
	Çoktan seçmeli	5
	Boşluk doldurma	3
Ana düşünce	Doğru-Yanlış	
	Açık uçlu	3
	Çoktan seçmeli	
Yüzey metin	Boşluk doldurma	
	Doğru-Yanlış	
	Açık uçlu	2
Başlık	Çoktan seçmeli	3
	Boşluk doldurma	
	Doğru-Yanlış	3
Toplam	Açık uçlu	
	Çoktan seçmeli	2
	Boşluk doldurma	
	Doğru-Yanlış	
		30

Verilerin Çözümlemesi

Okuma-Anlama Testi'nde hem ikili hem de çoklu puanlar olduğu için güvenilirlik katsayısı olarak cronbach alfa katsayısı hesaplanmış ve Spearman Brown iki yarı testi uygulanmıştır. Ayrıca ölçeğin madde ayırt ediciliğini belirlemek için madde-toplam korelasyonu ve %27'lik alt-üst grup karşılaştırmaları yapılmıştır. Bu açıdan, ölçeğin iç geçerliliğinin tespiti için bağımsız gruplar için t-testi uygulanmıştır. Bu analizler için ITEMAN ve SPSS analiz programlarından yararlanılmıştır. Kapsam geçerliğini belirlemek amacıyla uzmanlar arası uyumlarını inceleyen Lawshe tekniği kullanılmıştır. Ölçek maddesinin anlaşılabilir olması, hedef kitleye uygunluğu vb. önsel çalışmalarda elde edilen uzman görüşleri arasındaki uyum/uyumsuzluk aynı zamanda kapsam ya da yapı geçerliği için birer kestirim niteliğinde kullanılmaktadır (Yurdugül,2005).

BULGULAR

Alınan uzman görüşü sonrasında 30 maddeye indirilen testin ön deneme çalışması 20 öğrenci ile yürütülmüştür. Uzmanların beşi Türkçe eğitimi alanında, ikisi ölçme ve değerlendirme alanında çalışmalarını sürdürmektedir. Türkçe alanındaki uzmanların maddelere yönelik değerlendirmeleri sonucunda uygulanan Lawshe Tekniği ile hesaplanan kapsam geçerlik oranı 1.00'dir. Bu bulgu Veneziano ve Hooper (1997) tarafından tabloya dönüştürülen $\alpha=0,05$ anlamlılık düzeyinde kapsam geçerliği oranını yeterli kılmaktadır (Yurdugül,2005) Ayrıca geçerlik çalışmaları kapsamında madde analizi yapılmış her maddenin güçlük indeksi ve ayırt edicilik indeksleri hesaplanmıştır Başarı testi içerisindeki maddelerin her birinin madde güçlük ve ayırt edicilik değerleri Tablo 3'te gösterilmiştir.

Tablo 3. Okuma-Anlama Testi Madde Güçlük ve Ayırt Edicilik Değerleri

Madde Nu	Madde güçlük indeksi (P _{jx})	Madde Ayırt edicilik indeksi (r _{jx})
1	.67	.37
2	.69	.39
3	.89	.46
4	.50	.38
5	.59	.41
6	.46	.58
7	.80	.40
8	.18	.23
9	.41	.48
10	.27	.26
11	.32	.42
12	.50	.55
13	.51	.48
14	.41	.54
15	.39	.40
16	.63	.62
17	.44	.45
18	.54	.56
19	.59	.65
20	.68	.71
21	.52	.57
22	.14	-.21
23	.64	.52
24	.41	.39
25	.49	.42
26	.14	.58
27	.05	-.33
28	.79	.65
29	.82	.65
30	.74	.36

Her kazanıma ait maddeler seçilirken ayırıcılık gücü indeksi büyük olanlar tercih edilmeli ve ayırıcılığı 0,20 ve altında olanlar alınmamalıdır (Baykul, 2000). Bu doğrultuda, madde ayırt ediciliği 0,20 altında bulunan madde 22 ve 27 testten çıkarılmıştır.

Eğer bir soruyu çok az kişi yanıtlamışsa sorunun sayısal değeri 0,00'a yaklaşır ve soru zor, güç bir soru olarak yorumlanır (Özçelik 2010). Madde güçlük indeksi açısından sorular aldıkları değer doğrultusunda aşağıdaki şekilde değerlendirilir: - 0,00-0,19 arası çok güç, - 0,20-0,39 arası güç, - 0,40-0,59 arası orta güçlükte, - 0,60-0,79 arası kolay, - 0,80-1,00 arası çok kolay. Buna göre, 1,2,3,7,16,20,23,28,29 ve 30. maddeler kolay, 4,5,6,9,12,13,14,17,18,19,24 ile

25. maddeler orta güçlükte ve son olarak 8,10,11,15,26. maddeler zor olarak değerlendirilebilir. Böylelikle soruların kolaylık ve zorluk açısından dengeli bir biçimde dağılım göstermesi bakımından heterojen olduğu söylenebilir.

Testin ortalama madde güçlük ve ayırt edicilik indeksleri sırasıyla 0,507 ve 0,431 olarak belirlenmiştir.

Ölçeğin madde ayırt ediciliğini ve ölçeğin iç geçerliliğinin belirlemek amacıyla %27'lik alt-üst grup karşılaştırmaları yapılmıştır. Maddelerin başarılı ve başarısız olan gurubu ne ölçüde ayırt ettiğini hesaplamak amacıyla bu gruplar arasında bağımsız gruplar için t-testi uygulanmıştır. Öncelikle test puanları, küçükten büyüğe doğru sıralanmış ve grubun %27'lik alt ve üst kısmı hesaplanmıştır. %27'lik dilimde kalan 188 katılımcının aldığı toplam puanlar ve t-testi sonuçları tablo 4'te karşılaştırılmıştır.

Tablo 4. Okuma-Anlama Testi Alt Üst %27'lik Gruplar Karşılaştırma Sonuçları

Madde Nu	Gruplar	\bar{X}	S_x	T
1	Alt _{%27}	1.34	0.80	-6,703**
	Üst _{%27}	1.78	0.42	
2	Alt _{%27}	1.35	0.81	-5,480**
	Üst _{%27}	1.73	0.51	
3	Alt _{%27}	1.65	0.72	-5,817**
	Üst _{%27}	1.97	0.18	
4	Alt _{%27}	0.94	0.94	-7,006**
	Üst _{%27}	1.53	0.66	
5	Alt _{%27}	0.43	0.50	-6,109**
	Üst _{%27}	0.72	0.45	
6	Alt _{%27}	0.28	0.45	3,920**
	Üst _{%27}	0.12	0.33	
7	Alt _{%27}	0.71	0.52	-5,879**
	Üst _{%27}	0.97	0.29	
8	Alt _{%27}	0.43	0.57	-7,219**
	Üst _{%27}	0.96	0.84	
9	Alt _{%27}	0.24	0.43	-7,846**
	Üst _{%27}	0.62	0.49	
10	Alt _{%27}	0.21	0.41	-3,989**
	Üst _{%27}	0.40	0.49	

11	<i>Alt</i> _{%27}	0.54	0.76	-9,571**
	<i>Üst</i> _{%27}	1.30	0.79	
12	<i>Alt</i> _{%27}	0.26	0.44	-10,394**
	<i>Üst</i> _{%27}	0.73	0.44	
13	<i>Alt</i> _{%27}	0.32	0.47	-9,416**
	<i>Üst</i> _{%27}	0.76	0.43	
14	<i>Alt</i> _{%27}	0.19	0.40	-11,730**
	<i>Üst</i> _{%27}	0.71	0.46	
15	<i>Alt</i> _{%27}	0.25	0.43	-7,105**
	<i>Üst</i> _{%27}	0.59	0.49	
16	<i>Alt</i> _{%27}	0.35	0.48	-13,656**
	<i>Üst</i> _{%27}	0.90	0.30	
17	<i>Alt</i> _{%27}	0.24	0.43	-9,134**
	<i>Üst</i> _{%27}	0.67	0.80	
18	<i>Alt</i> _{%27}	0.47	0.42	-16,872**
	<i>Üst</i> _{%27}	1.77	0.81	
19	<i>Alt</i> _{%27}	0.57	0.51	-19,127**
	<i>Üst</i> _{%27}	1.86	0.72	
20	<i>Alt</i> _{%27}	0.32	0.18	-17,600**
	<i>Üst</i> _{%27}	0.97	0.94	
21	<i>Alt</i> _{%27}	0.27	0.66	-13,689**
	<i>Üst</i> _{%27}	0.84	0.50	
22	<i>Alt</i> _{%27}	0.38	0.45	-10,689**
	<i>Üst</i> _{%27}	0.98	0.45	
23	<i>Alt</i> _{%27}	0.45	0.33	-10,448**
	<i>Üst</i> _{%27}	0.96	0.52	
24	<i>Alt</i> _{%27}	0.26	0.29	-6,504**
	<i>Üst</i> _{%27}	0.57	0.57	
25	<i>Alt</i> _{%27}	0.30	0.84	-8,281**
	<i>Üst</i> _{%27}	0.70	0.43	
26	<i>Alt</i> _{%27}	0.06	0.49	-12,445**

	Üst _{%27}	.90	0.41	
27	Alt _{%27}	0.49	0.49	-11,492**
	Üst _{%27}	1.00	0.76	
28	Alt _{%27}	0.56	0.79	-10,116**
	Üst _{%27}	0.99	0.44	
29	Alt _{%27}	0.61	0.44	-10,053**
	Üst _{%27}	1.01	0.47	
30	Alt _{%27}	0.66	0.43	-6,856**
	Üst _{%27}	1.01	0.40	

**p=.000

Tablo 4 incelendiğinde maddelerin tümündeki ayırt edicilik değerlerinin anlamlı olduğu görülmektedir, (p=.000). Uzman görüşleri ve madde analizi sonunda 28 maddeye indirgenen başarı testinin geçerliliğinin yüksek olduğu saptanmıştır. Geçerlik çalışması kapsamında 2 madde çıkarıldıktan sonra nihai testin betimsel istatistikleri hesaplanmıştır. Bu değerler tablo 5’te verilmiştir.

Tablo 5. Nihai Teste İlişkin Betimsel İstatistikler

N	Ortalama	Medyan	Mod	Ss	Varyans	Min	Max	Basıklık katsayısı	Çarpıklık katsayısı
696	21,780	23,0	24,0	5,930	35,17	6.00	36.00	-.518	-.284

Testin ortalama güçlüğü 0,533 bulunmuştur. Testin orta güçlükte bir test olduğu söylenebilir. Nihai Testten alınabilecek en düşük puan 0, en yüksek puan 36 ‘dır. Teste katılan öğrencilerin testten aldıkları en düşük puan 6 en yüksek puan ise 36’dır.

Nihai teste ait çarpıklık ve basıklık değerleri -1 ile +1 arasında bulunmuştur. Dolayısıyla nihai teste ait puanların dağılımının normal dağılımdan önemli bir sapma göstermediği söylenebilir (Büyüköztürk, 2009).

OAT’ın iç tutarlılığını incelemeye yönelik olarak Cronbach alfa güvenilirliği hesaplanmış, güvenilirlik katsayısı 0,744 olarak bulunmuştur. Alpha değerinin en az 0.70 olması gerektiği (Anderson, 1988; Kline, 1994; Peers, 1996; Akt. Baş,2013) göz önünde bulundurulduğunda, OAT’da yer alan soruların kavrama türüne göre ölçülen özellikler bakımından iç tutarlılığa sahip olduğu, aynı zamanda testin uygulama koşullarından gelen tesadüfi hatalardan da arınık olduğu söylenebilir.

İkinci olarak test yarılama yöntemiyle ölçeğin güvenilirliği hesaplanmış, testin Spearman-Brown iki yarı test korelasyonu 0.802 olarak bulunmuştur. Alanyazında 0.80’in üzerindeki değerler “iyi” olarak nitelenmektedir

(Büyüköztürk, 2009). Test yarılama güvenilirliği, bir testi iki kere uygulamanın veya bir testin iki eşdeğer formunun hazırlanmasının güç olduğu ve testin tek bir değişkeni ölçtüğü durumlarda başvurulması gereken bir yöntemdir (Özen, Gülaçtı ve Kandemir, 2006). Testi yarılama yöntemi ile elde edilen güvenilirlik iç tutarlılık olarak da bilinmektedir.

TARTIŞMA ve SONUÇ

Ortaokul öğrencilerinin okuduğunu anlama durumlarını ölçmeye yönelik geçerli ve güvenilir bir test geliştirmeyi amaçlayan bu çalışmada kuramsal çerçeve bağlamında 50 soru hazırlanmış, daha sonra uzman görüşleri, ayrıca soru türü, kavrama türü ve bir ders saatinde uygulanabilirlik açısından bu sayı 30'a indirilmiştir. Hazırlanan taslağın son halinde bulunan sekizi açık uçlu, on biri çoktan seçmeli, üçü boşluk doldurma ve sekizi de doğru yanlış türü olmak üzere toplam 30 sorudan oluşan test, ortaokul 5-8. sınıfa devam eden 636 öğrenci üzerinde uygulanmıştır.

Yapılan çözümleme sonucunda, hesaplanan ortalama madde güçlük ve ayırt edicilik indeksleri çoktan seçmeli, boşluk doldurma ve doğru-yanlış türü sorulardan oluşan "Okuma-Anlama Testi"nin geçerli ve güvenilir olduğunu göstermektedir. Okuma-Anlama Testi'nin son biçimi 11 çoktan seçmeli, 3 boşluk doldurma ve 6 doğru-yanlış ve 8 açık uçlu olmak üzere toplam 28 sorudan oluşmuştur. Testten alınacak en yüksek puan 36; en düşük puan ise 0'dır. Testte yer alan çoktan seçmeli sorular 4 seçenekli olup, doğru yanıtla "1" yanlış ve boş yanıtla "0" puan; doğru-yanlış ve boşluk doldurma sorularında da doğru yanıtla "1", yanlış ve boş bırakılan yanıtla "0" puan"; açık uçlu sorularda ise bölümsel doğru yanıtla "1", bütünsel doğru yanıtla "2", yanlış ve boş bırakılan yanıtla "0" puan verilerek 36 puan üzerinden değerlendirme yapılmalıdır. Testten alınacak yüksek puanlar okuyucunun okuduğunu anlama durumu konusunda daha yüksek bir başarıya işaret eder.

Okuduğunu anlamaya yönelik hazırlanan ölçme araçlarının farklı soru türlerini içermesi okuduğunu anlamaya ilişkin daha doğru çıkarımlara izin verebilir. Bunun nedeni, okuma ediminin karmaşık yapısıyla ilişkili olarak düşünülmelidir. Geliştirilen bu araçtaki farklı soru tipleri bu kaygıyla seçilmiştir. Buna karşın Çelik, Demirgüneş ve Fidan'ın (2015) yaptığı çalışmada yalnızca açık uçlu sorular kullanılmış ve farklı metin tiplerinde farklı sonuçlara ulaşılmıştır. Oysa bir ölçme aracının genellenebilir olması beklenir. Tek tip sorular, aracın genellenebilirlik düzeyini düşürebilir. Buna benzer biçimde Baştuğ ve Kaman'ın (2013) ve Baştuğ ve Akyol'un (2013) çalışmasında yalnızca açık uçlu sorular; Yıldız ve Akyol'un (2011) çalışmasında ise yalnızca çoktan seçmeli soru tipleri kullanılmıştır. Yaman ve Dağtaş'ın (2014) ekrandan okumanın anlamaya etkisine ilişkin yaptığı çalışmada geliştirilen okuduğunu anlama testinde ise farklı metin türleriyle ilgili çoktan seçmeli ve açık uçlu soruların birlikte kullanılmıştır. Ancak bu testte de her metin için 8 çoktan seçmeli, 1 açık uçlu soru kullanıldığından soru türleri arasında bir orantı sağlanmamıştır.

Ölçülecek beceri ile ölçme aracı arasındaki tutarlılık bilimsel çalışmaların daha doğru veriler üstünden yapılmasına olanak sağlar. Bu çalışmayla geliştirilen aracın farklı metin türlerine yönelik olmasıyla metinle ilgili unsurları, içerdiği farklı soru türleriyle de bireyle ilgili kimi boyutları kapsadığı söylenebilir. Okuma ediminin yeterli bir anlama durumuyla sonuçlanması, okuduğunu anlamının iki tarafını oluşturan metin-okur özelliklerini kapsamıyla olanaklıdır.

KAYNAKLAR

- Alderson, J., C. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Allen, J. (1985). Inferential comprehension: The effects of text source, decoding ability and mode. *Reading Research Quarterly*, 20(5), 603-615.
- Baş, G. (2013). Öğretmenlerin eğitim programı tasarımı yaklaşımı tercih ölçeği: geçerlik ve güvenilirlik çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 13(2), 965-992.
- Baştuğ, M & Akyol, H. (2012). The Level of Prediction of Reading Comprehension by Fluent Reading Skills. *Journal of Theoretical Educational Science*, 5(4), 394-411.
- Baştuğ, M. & Kaman, Ş. (2013). Nörolojik Etki Yönteminin Öğrencilerin Akıcı Okuma ve Anlama Becerilerine Etkisi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 13 (25), 291 – 309.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: ÖSYM Yayınları
- Bowyer-Crane, C., & Snowling, M. (2005). Assessing children's inference generation: What do tests of reading comprehension measure? *British Journal of Educational Psychology*, 75, 189–201.
- Büyüköztürk, Ş. (2009). *Sosyal Bilimler İçin veri Analizi*. (10. Baskı). Ankara: PegemA Yayıncılık
- Cadime, I., Ribeiro, I., Viana, F. L., Santos, S., Prieto, G. & Maira, J. (2013). Validity of a reading comprehension test for Portuguese students. *Psicothema*, 25(3), 384-389. doi: 10.7334/psicothema2012.288
- Cain, K., & Oakhill, J. (1998). Comprehension skill and inference-making ability: Issues of causality. In C. Hulme & R. M. Joshi (Eds.) *Reading and Spelling* (pp329-342). Mahwah, NJ: Erlbaum.
- Çelik, T.; Demirgüneş, S & Fidan D. (2015). Okur dostu metin olma özelliği ile okuduğunu anlama başarısı arasındaki ilişkinin incelenmesi. *Başkent University Journal Of Education*, 2(1), 115-122.
- Dedebali, N., C. & Saracaloğlu, A., S. (2010). Hızlı Okuma Tekniğinin Sekizinci Sınıf Öğrencilerinin Okuma Hızlarına ve Okuduğunu Anlama Düzeylerine Etkisi, *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 27, 171-183.
- Keenan, J. M. & Betjeman S. B. (2006). Comprehending the Gray oral reading test without reading it: Why comprehension tests should not include passage-independent items. *Scientific Studies od Reading*, 10(4), 363-380.
- Kintsch, E. (2005b). Comprehension Theory as a Guide for the Design of Thoughtful Questions. *Topics In Language Disorders*, 25 (1), 51-64.
- Kintsch, W. (2004). The construction-integration model of test comprehension and its implications. In R., B. Ruddel & N., J. Unrau (Eds.) *Theoretical Models and Process of Reading* (5th ed., pp1270-1328). Newark, DE: International Reading Association.

- Kintsch, W. (2005a). An overview of top-down and bottom-up effects in comprehension: The CI perspective, *Discourse Process*, 39:2-3, 125-128.
- Kintsch, W., & Kintsch, E. (2005). Comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Current Issues onreading comprehension and assessment* (pp. 71-92).
- Köksal, K. ve Ünal, E. (2008). Metinler arası okumanın okuduğunu anlamaya etkisi. *Elektronik Sosyal Bilimler Dergisi*, 7(26), 154-169.
- Mahwah, NJ: Erlbaum. Köksal, K. & Ünal, E. (2008). Metinler arası okumanın okuduğunu anlamaya etkisi. *Elektronik Sosyal Bilimler Dergisi*, 7(26), 154-169.
- Oakhill, J. V. (1984). Inferential and memory skills in children's comprehension of stories. *British Journal of Educational Psychology*, 54, 31-39.
- Özçelik, D. A. (2013). *Test Hazırlama Kılavuzu (5. Baskı)*. Ankara: Pegem Akademi
- Özen, Y., Gülaçtı, F. ve Kandemir, M. (2006). Eğitim bilimleri araştırmalarında geçerlik ve güvenilirlik sorunsalı. *Erzincan Eğitim Fakültesi Dergisi*, 8 (1), 69-89.
- Tavşacı, E. (2006). Tutumların Ölçülmesi ve SPSS ile Veri Analizi. Ankara: Atlas Yayınları.
- Tracey, D, H., Morrow, L.,M. (2012). Lenses on Reading An Introduction to Theories and Models, New York: The Guilford Press.
- Ülper, H. (2010). Okuma ve Anlamlandırma Becerilerinin Kazandırılması. Ankara: Nobel Yayıncılık.
- van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. San Diego, CA: Academic Press.
- Yaman H. & Dağtaş, A. (2014). Ekrandan okumanın okuduğunu anlamaya etkisi. *Okuma Yazma Eğitimi Araştırmaları*, 1(2), 64-79.
- Yıldız, M. & Akyol, H. (2011). İlköğretim 5. sınıf öğrencilerinin okuduğunu anlama, okuma motivasyonu ve okuma alışkanlıkları arasındaki ilişki. *Gazi Eğitim Fakültesi Dergisi*, 31, (3) 793-815.
- Yurdugül, H. (2005). Ölçek geliştirme çalışmalarında kapsam geçerliği için kapsam geçerlik indekslerinin kullanılması. *XIV. Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi*, Denizli.

SUMMARY

At the present time, reading is accepted as a multidimensional process that consists of mental operations at different levels. Accordingly, mental operations come true in linguistic, conceptual, and situational levels while reading. For a configuration related to these three different levels, it is necessary to make a mental operation at local and global degrees and interacting background knowledge by the text. While sentences are being associated each other in the processes at the local degree; structuring the meaning of the text at the global degree, grabbing the main idea, summarizing the text, and defining a proper title for the text are made at the global degree (van Dijk&Kintsch, 1983; Kintsch, 1994; 2004; Kintsch & Kintsch, 2005).

In regard to the theoretical knowledge, it is possible that the questions for comprehension can be placed under the title of text-based and inference. The questions about text-base are oriented towards the main idea or details in the text. For this reason, that kind of questions are answered by gaining the related information from the mind or text. Multiple-choice, gap filling, matching questions are appropriate for this aim. Summarization can be taken text-based quality question on condition that it requires well-structured answers. In this regard, summarization questions are first-level questions. Second-level questions are those which has short answers. Third-level questions are the questions examining specific information such *wh* questions and have answers consisting of a lexical bundle or one sentence (Kintsch, 2005b). Inference questions need passing beyond of the external information in the text, interpreting the knowledge by the given content, establishing new connections among different sentences and thoughts, applying the information in the text to a new problem (Allen, 1985; Cain & Oakhill, 1998; Kintsch, 2005b).

It is necessary to use different kinds of questions or techniques in reading comprehension scales. Each kind of questions has advantages and disadvantages. Further, each technique focuses on different points in reading comprehension process. In other words, different question techniques measure different aspects of reading comprehension process. In addition to this, measurement is restricted by what the technique enables to measure of reading (See Ülper, 2010; Alderson, 2000).

This study was conducted in survey model. The participants were 696 students at 5, 6, 7, and 8. grades at schools in Antalya, Burdur, Niğde, Ordu, İstanbul, and Urfa. At the beginning, 50 questions were prepared for the reading comprehension scale. 20 questions were discarded after expert opinions by different reasons such as question types, comprehension types, applicability in a reasonable time. Rest of the questions in the last draft were 30 consisting of 8 open-ended, eleven multiple-choice, 3 gap filling, and 8 true-false questions. The draft scale had 9 questions on inferences, 8 questions on predicting the relations between propositions, 3 questions about the main idea, 8 questions on the surface

of the text, and 2 questions about finding title for the text. The last draft was applied in this form.

Cronbach's Alpha and Spearman-Brown split half correlation were calculated in the study for reliability co-efficient. For item differentiation of the scale, item total correlation was calculated and the differences between mean scores of upper 27% and lower 27% groups were examined. Independent sample t-test was applied for internal validity and Lawshe technique was applied for the content validity of the scale.

The item difficulty index of the test is calculated as 0,505 and item differentiation index as 0,431. The calculation indicates that all items' differentiation values are meaningful, ($p=.000$). The validity of the achievement test degraded to 28 items after expert opinion and item analysis was determined high. The mean of difficulty of the test is 0,533. This score indicates that the test is in the medium difficulty. The highest score of the test is 36 and the lowest 0. The students participated to the research got 6 scores as lowest, 36 as highest. Skewness and kurtosis values are found between -1 to +1 of the final test. Thereby, it can be said that the distribution of the final test's scores did not deviate from the normal distribution significantly.

Cronbach alpha coefficient was found as 0,744. This value indicated that the questions in the test had internal consistency in terms of the features measured for comprehension types, at the same time; it was purged from the accidental errors because of the applying circumstances. Afterwards, reliability of the scale was calculated by split half method, Spearman-Brown two sub-test correlations are found as 0.802.

It is understood that Reading Comprehension Test is a reliable and valid test in respect to the mean of item difficulty and item differentiation calculation. The last version of the test has 11 multiple-choice, 3 gap filling, 6 true-false and 8 open ended; in total 28 questions. The highest score from the test is 36, the lowest is 0. The multiple choice questions have four choices. The right answers should be marked as "1" and the wrong answers as "0" in all question types. However, in open-ended questions, the local right answers should be marked as "1", and the global right answers as "2", wrong and non-response questions as "0". The higher scores from the test indicate higher achievement on reading comprehension.