

PAPER DETAILS

TITLE: Üretim, Dagitim ve Tüketim öğrenme alanina yönelik akademik basari testi gelistirme:
Geçerlik ve güvenirlik çalışması*

AUTHORS: Pinar TAGRIKULU,Ünsal BEKDEMİR,Elif Omca ÇOBANOGLU

PAGES: 1984-2004

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/1829819>


<http://kefad.ahievran.edu.tr>

Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi

ISSN: 2147 - 1037

Developing an Academic Achievement Test for Learning Area of Production, Distribution and Consumption: Validity and Reliability Study

Pınar Tağrikulu - Ünsal Bekdemir - Elif Omca Çobanoğlu

Article Information



CrossMark

DOI: 10.29299/kefad.953899

Received: 17.06.2021

Revised: 20.09.2021

Accepted: 20.10.2021

Keywords:

Achievement test,
Learning area of Production
Distribution and
Consumption,
Social Studies

Abstract

In this study, the aim was to develop an achievement test for 6th grade students in the learning area of Production, Distribution and Consumption, which is the 5th learning area in the 6th grade curriculum. In line with this goal, the process of developing an achievement test was carried out in accordance with the steps of preparation of the test plan, formation of the item pool, review of the items, application of the trial form and item analysis. When the curriculum is examined, it can be seen that this learning area consists of 6 acquisitions, and each acquisition is quite comprehensive in itself. Therefore, a detailed research and development process was carried out to develop an achievement test consisting of multiple choice questions covering each acquisition. In order to prepare questions suitable for the revised Bloom taxonomy, an item pool was created by examining sample questions, textbooks and question examples in exams. At this point, no questions were taken directly, and it was tried to ensure that all questions were prepared in an original way by the researcher. The validity and reliability study of the achievement test was carried out with 379 students studying at secondary school level in Samsun. Validity and reliability studies were started with 41 items, and 13 items were excluded from the test by looking at their difficulty and discrimination indices. The reliability level of the test is 0.897, the average difficulty level is 0.713 and the average discrimination level is 0.519. According to the data obtained from the study, a valid and reliable achievement test was obtained. It is thought that the developed achievement test can be used to determine students' prior knowledge and learning outcomes in the education process.

Üretim, Dağıtım ve Tüketim Öğrenme Alanına Yönelik Akademik Başarı Testi Geliştirme: Geçerlik ve Güvenirlik Çalışması

Makale Bilgileri



CrossMark

DOI: 10.29299/kefad.953899

Yükleme: 17.06.2021

Düzeltilme: 20.09.2021

Kabul: 20.10.2021

Anahtar Kelimeler:

Başarı testi,
Üretim Dağıtım ve Tüketim
Öğrenme Alanı,
Sosyal Bilgiler

Öz

Bu çalışmada 6. sınıf öğrencilerine yönelik, 6. sınıf öğretim programında 5. öğrenme alanı olarak yer alan Üretim, Dağıtım ve Tüketim öğrenme alanında bir başarı testi geliştirmek hedeflenmiştir. Bu hedef doğrultusunda test planının hazırlanması, madde havuzunun oluşumu, maddelerin gözden geçirilmesi, deneme formunun uygulanması ve madde basamaklarına uygun olacak şekilde başarı testi geliştirilmesi süreci yürütülmüştür. Öğretim programına bakıldığında bu öğrenme alanının 6 tane kazanımdan oluştuğu, her bir kazanımın ise kendi içinde oldukça kapsamlı olduğu saptanmıştır. Dolayısıyla her bir kazanımı kapsayan, çoktan seçmeli sorulardan oluşan bir başarı testinin geliştirilmesi için detaylı bir araştırma ve geliştirme süreci yürütülmüştür. Yenilenmiş Bloom taksonomisine uygun olacak sorular hazırlamak için sınavlarda çıkmış örnek sorular, ders kitapları ve soru örnekleri incelenerek bir madde havuzu oluşturulmaya çalışılmıştır. Bu noktada hiçbir soru doğrudan alınmamış, tüm soruların araştırmacı tarafından özgün şekilde hazırlanması sağlanmaya çalışılmıştır. Samsun ilinde ortaokul seviyesinde öğrenim görmekte olan 379 öğrenci ile başarı testinin geçerlik ve güvenilirlik çalışması yürütülmüştür. Geçerlik ve güvenilirlik çalışmalarına 41 madde ile başlanmış, 13 madde güçlük ve ayırt edicilik indekslerine bakılarak testten çıkarılmıştır. Güvenirlik düzeyi 0.897 olan testin, ortalama güçlük düzeyi 0.713 ortalama ayırt edicilik düzeyi ise 0.519'dur. Çalışmadan elde edilen verilere göre geçerli ve güvenilir bir başarı testi elde edilmiştir. Geliştirilen başarı testinin eğitim öğretim sürecinde öğrencilerin ön bilgilerinin ve öğrenme çıktılarının tespit edilmesinde kullanılabileceği düşünülmektedir.

Sorumlu Yazar : Pınar Tağrikulu, Arş. Gör., Ondokuz Mayıs Üniversitesi, Türkiye, pinar.tagrikulu@omu.edu.tr, ORCID ID: 0000-0002-5221-6888.

Ünsal Bekdemir, Prof. Dr., Giresun Üniversitesi, Türkiye, unsal.bekdemir@giresun.edu.tr, ORCID ID: 0000-0002-7767-9088

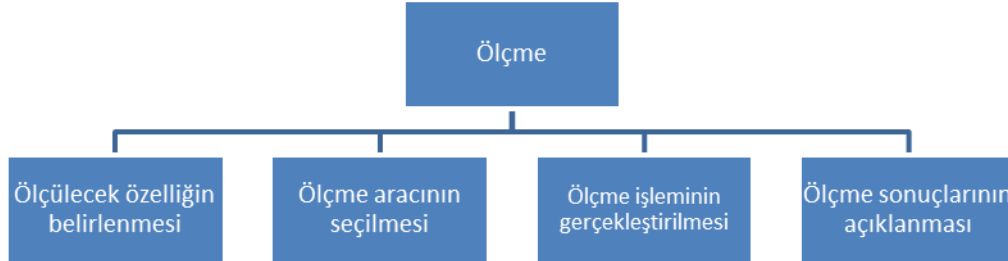
Elif Omca Çobanoğlu, Doç. Dr., Ondokuz Mayıs Üniversitesi, Türkiye, omcacobanoglu@gmail.com, ORCID ID: 0000-0002-3691-8273

Alt Bilgi: Bu çalışma Pınar Tağrikulu'nun Giresun Üniversitesi, Sosyal Bilimler Enstitüsü, Türkçe ve Sosyal Bilimler Eğitimi Anabilim Dalı'nda, Prof. Dr. Ünsal Bekdemir ve Doç. Dr. Elif Omca Çobanoğlu tarafından yürütülmekte olan doktora tezinden üretilmiştir.

Atıf için: Tağrikulu, P., Bekdemir, Ü., & Omca Çobanoğlu, E. (2021). Üretim, dağıtım ve tüketim öğrenme alanına yönelik akademik başarı testi geliştirme: geçerlik ve güvenilirlik çalışması. *Kırşehir Eğitim Fakültesi Dergisi*, 22(3), 1984-2004.

Giriş

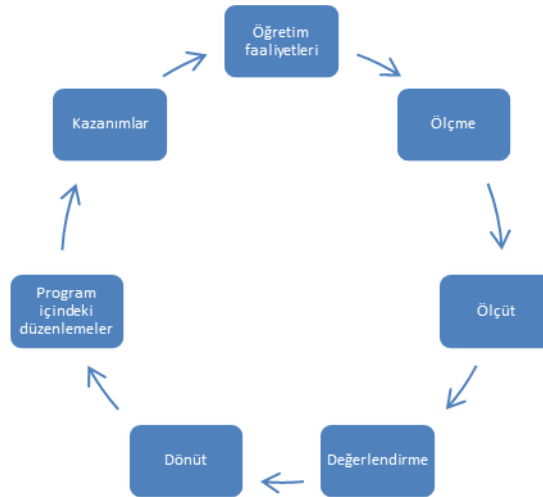
Ölçme, ölçülecek olan özelliğe göre farklı yapıda olmakta, ölçme yapılması için ölçülecek olan özelliğin tanımlanması, ölçülecek olan özelliğe uygun nitelikte sayı veya semboller kümesi belirlenmesi ve ölçülecek olan özelliğe verilecek olan sayı veya sembollerin hangisi olduğunu gösteren nitelikte kural ve kuralların gösterilmesi gerekmektedir (Yaşar, 2017). Ölçme sürecinin aşamaları ise şu şekilde gösterilebilir:



Şekil 1. Ölçme sürecinin aşamaları (Yaşar, 2017).

Tüm ölçmeler, bir değerlendirmeye hizmet etmekte; bir ölçütten yola çıkarak yargıya ve karara ulaşma süreci değerlendirme olarak ifade edilmektedir (Semerci, 2014). Değerlendirmenin sonunda ölçülen özellik ya da nitelik hakkında, az ya da çok, yeterli ya da yetersiz, kabul edilebilir ya da edilemez gibi bir kanaate ulaşılır (Semerci, 2014). Değerlendirme sayesinde başarı seviyesinin yükselmesi, farklı yöntemlerin ne derece etkili olduğuna ulaşılması, eğitimde kalite kontrolü sağlanması mümkün olurken etkili bir öğretim de gerçekleşir (Başol, 2018).

Öğretim sürecinde ölçme ve değerlendirmenin yeri şu şekilde gösterilebilir:



Şekil 2. Öğretim sürecinde ölçme ve değerlendirmenin yeri (Başol, 2018).

Bu döngüden de anlaşılacağı üzere ölçme ve değerlendirme bir bütündür. Bir bütün olmasının da ötesinde her bir basamağı bir diğer basamağı ile yakından ilgili olan ölçme ve değerlendirme, eğitim öğretim sürecinin olmazsa olmazıdır. Bir diğer ifade ile ölçme ve değerlendirme kavramları birbiri ile anılan, birbirini tamamlayan süreçler olarak değerlendirilmekte; değerlendirme işleminin yapılması bir ölçme yapılmasını gerektirmektedir (Özer Özkan, 2019). Eğitim

etkinliklerinin ne ölçüde başarılı olduğu, yıl veya dönem içinde kaydedilen ilerlemenin öğretmenin öğretim hedeflerini kazandırmada ve öğretim sürecinde öğrencilere rehberlik etmede ne ölçüde başarılı olduğu ölçme ve değerlendirme yapılarak saptanır (Kilmen, 2019).

Öğrencilerin duyuşsal, bilişsel ve psikomotor özellikleri bakımından tanınarak birbirleri ve çevreleri ile dengeli bir iletişim kurabilmeleri adına belli başlı becerilere sahip olmaları eğitim açısından çok önemli iken gözlenemeyecek nitelikteki özellikleri gözlenebilecek hale getirmek ve eğitimde kazandırılması istenen özelliklerin kazanılıp kazanılmadığını saptayabilmek amacı ile ölçme araçlarından faydalanılır (Kan, 2017). Öğrencilerin başarıları, sınavlarda kullanılan ölçme araçları ile ölçülebilen bu ölçme araçları, uzun cevap gerektiren sınavlar, çeşitli performans testleri, kısa cevaplı testler, çoktan seçmeli testler olabildiği gibi sözlü yoklamalar şeklinde de olabilir (Semerci, 2014). Ölçme araçları, ölçme işlemini kolaylaştırması, ölçme sonucunda elde edilen sonuçların niteliklerini belirleyici özelliklere sahip olmakla birlikte bunlar, ölçme sonuçlarının matematiksel özellikleri olup bir ölçek tanımlar (Kan, 2019). Objektif sınavlar, optik okuyucu veya delikli anahtarlar kullanılarak okunduğu için, puanlama güvenilirliği yüksek olan sınavlardır ve bu tür sınavlarda sübjektiflik bir tehdit değildir (Başol, 2018). Objektif puanlanabilen maddeler, açık veya kapalı uçlu soru şeklinde olabilirken kapalı uçlu soruların yoruma kapalı olduğu belirtilebilir ve bu tür sorularda soruyu cevaplayan kişilerin cevabı verilenler arasından seçmek durumunda olduğu ifade edilebilir ve çoktan seçmeli testler bu testler içinde en çok kullanılan olarak başı çekmektedir (Başol, 2018). Çoktan seçmeli testler, çok sayıda soru sorulması, seçme ve yerleştirme gerekmesi, sınav hazırlamak için yeterli vaktin olduğu ve sonuçların çabuk bildirilmesi gereken, bu testleri hazırlamak için gereken donanımına sahip olunan, testin geçerlik ve güvenirliğinin yüksek olması gereken durumlarda tercih edilmektedir (Güler, 2019).

Çoktan seçmeli maddeler ile farklı türde beceri ve bilgiyi, üst düzey zihinsel becerileri ölçmek, çok sayıda kişiye uygulama yapabilmek, bilgi düzeyinden değerlendirme düzeyine kadar farklı türde sorular hazırlayabilmek, üst düzey bilgi ve beceri alanlarını ölçmek, fazla sayıda soru sormak mümkün iken kapsam geçerliğinin ve testten alınacak puanların güvenirliğinin artması da mümkündür (Çakan, 2017). Bunların yanı sıra çoktan seçmeli bir sınavın hazırlanmasını zorlaştıran etmenler de bulunmakta bunlar, soruların ve seçeneklerin oldukça açık ve kısa yazılması gerekliliği, olumsuz köklerin vurgulanması, madde kökleri ve çeldiricilerin ipuçları içermemesi, çeldiricilerin öğrenciler tarafından kolayca elenememesi, aynı seçeneğin üst üste doğru cevap olarak kullanılmasından kaçınılması ve anahtarlanmış yanıtların bir örüntüye sahip olmaması gerektiği olarak ifade edilebilirler (Aydın, 2019). Doğru cevabın çoğu zaman madde içinde yer alması, şans başarısı, ileri düzey davranışların ölçülmesinin zor olması ve hatırlama süresinin uzun, zahmetli ve uzmanlık gerektirmesi de çoktan seçmeli testlerin dezavantajları arasında yer almaktadır (Güler, 2019). Bu dezavantajların yanı sıra soru sayısı arttıkça, okuma hızı değişkeni de ölçmeye katılmakta, istatistiksel yollarla düzeltme imkanlarının olmasına karşın şans başarısının tamamen ortadan

kalkması da mümkün olamamaktadır (Turgut ve Baykul, 2019). Tüm bu dezavantajlara karşılık uygulama sürecindeki pratikliği göz önünde bulundurulduğunda, çoktan seçmeli testler, eğitim öğretim sürecinde olmazsa olmaz, kullanışlılığını her daim koruyacak testlerdir. Bu çalışmada da eğitim öğretim sürecinde işe koşulabilecek, pratik, farklı tiplerde yazılmış sorulardan oluşan, 6. sınıf düzeyinde geçerli ve güvenilir bir başarı testi geliştirmek hedeflenmiştir. Bu öğrenme alanına yönelik geliştirilmiş başarı testleri bulunmakta (Akgün ve Arıkan, 2015) ancak Sosyal Bilgiler dersinin çok yönlülüğü ve geniş bir konu alanı kapsamı olması, pek çok farklı soru yazmaya ve ölçme aracı geliştirmeye imkan sağlamaktadır. Bir konu alanının birbirinden farklı sorularla ölçme ve değerlendirilmesinin yapılması da yine farklı ölçme araçlarının geliştirilmesi ile mümkün olabilir. Böylece Sosyal Bilgiler dersinin 6. sınıf seviyesi, üretim, dağıtım ve tüketim öğrenme alanında ölçme aracı bakımından zenginlik sağlanabileceği düşünülmektedir.

Araştırmanın Amacı

Öğrencilerin çok yönlü düşünebilmesini gerektiren ve disiplinlerarası yapıda bir ders olan Sosyal Bilgiler dersi, farklı öğretim model ve yaklaşımları ile işlenmeye oldukça uygun yapıda dolayısıyla birbirinden farklı pek çok çeşitlilikte ölçme ve değerlendirme yaklaşımlarının da kullanılabileceği oldukça kapsamlı bir derstir. Bu yaklaşımlar doğrultusunda kullanılabilecek testlerden bir tanesi de çoktan seçmeli testlerdir. Çoktan seçmeli testlerin daha önce de ifade edilen avantajları göz önünde bulundurulduğunda, bu ders için 6. sınıf Üretim, Dağıtım ve Tüketim öğrenme alanına yönelik, çoktan seçmeli, bu öğrenme alanında yer alan tüm kazanımlara hitap eden, geçerli ve güvenilir bir başarı testi geliştirmek amaçlanmıştır.

Yöntem

Bu çalışmada bir başarı testi geliştirilmeye çalışılmış bu amaçla tarama yöntemine uygun hareket edilmiştir. Tarama araştırması, büyük evrenlerde, çok miktarda veri ile gerçekleştirildiğinden nicel verilere ve verilerin istatistiksel olarak analiz edilmesine dayalı bir yöntemdir (Özdemir, 2015).

Çalışma Grubu

Bu çalışma, 2018-2019 eğitim öğretim yılının bahar döneminde yapılmış, çalışmada Samsun ilinde öğrenim görmekte olan 379 öğrenciden, geçerli verilere ulaşılmıştır. Çalışmaya katılımda gönüllülük esasına uygun hareket edilmiş, çalışmaya katılmak isteyen öğrenciler ile süreç yürütülmüştür. Çalışmanın 379 öğrenci ile yapılan uygulama çalışmasından önce iki öğrenci ile uygulaması yapılarak, soruların açıklığı ve anlaşılabilirliği yordanmaya çalışılmış, öğrencilerden gelen dönütler doğrultusunda sınav süresinin de 65 dakika olarak belirlenmesi sağlanmıştır. Ardından 41 soruluk başarı testi 379 öğrenciye dağıtılmış, cevaplamaya geçilmeden önce test hakkında gerekli açıklamalar yapılarak öğrencilerin motive edilmesi sağlanmıştır.

Veri Toplama Aracı

Bir ölçek geliştirme çalışması olarak bu çalışmada veri toplama aracı olarak araştırmacı tarafından geliştirilen 6. sınıf Üretim, Dağıtım ve Tüketim öğrenme alanına yönelik başarı testi kullanılmıştır. Bu formun geliştirilmesindeki amaç öğrencilerin ilgili öğrenme alanına yönelik akademik başarılarını tespit edebilecek nitelikte bir araç geliştirilmesinin istenmesidir. Başarı testinin hazırlanmasında dikkat edilen basamaklar şu şekildedir:

Test planının hazırlanması

Madde havuzunun oluşturulması

Maddelerin gözden geçirilmesi

Deneme formunun uygulanması

Madde analizi (Başol, 2018: 236-245).

Bu çalışmada öncelikle test planı hazırlanmıştır. Geliştirilen testin 6. sınıf öğrencilerinin Üretim, Dağıtım ve Tüketim öğrenme alanı kazanımlarına ulaşip ulaşmadıklarını ortaya koyması hedeflenmiştir. Madde havuzu oluşturulurken Milli Eğitim Bakanlığı'nın öğretim programında (URL1) yer verdiği kazanımlar dikkate alınarak yola çıkılmıştır. Bloom'un yenilenmiş taksonomisine uygun olarak çalışılan başarı testine soru hazırlanırken soru örnekleri, ders kitapları, sınavlarda çıkmış örnek sorular incelenmiş ve bir madde havuzu oluşturulmaya çalışılmıştır. Ancak hiçbir soru doğrudan alınmamış, tüm sorular araştırmacı tarafından hazırlanmıştır. Burada bulunan altı tane kazanımın yenilenmiş Bloom taksonomisine göre bilişsel süreç ve bilgi boyutu bakımından hangi basamakta yer aldığı tespit edilerek bir belirtke tablosu oluşturulmuştur. Ardından bu altı kazanıma yönelik olarak her bir kazanım boyutunda 9 soru hazırlanmış, ilk etapta 54 soruluk bir madde havuzu oluşturulmuştur. Geliştirilen başarı testi, bilişsel basamaklara uygun olup olmama, dil ve anlatım kurallarına uygunluk, öğrenci seviyesine uygunluk bakımlarından uzman incelemesine sunulmuştur. Uzman incelemeleri sonrasında bazı maddelerde sorunun Bloom taksonomisinde bulunduğu basamağı değiştirmeyecek şekilde ifade düzeltmeleri yapılmıştır. 4., 5., 9., 16., 24., 27., 33., 35., 40., 45., 48., 53. ve 54. maddeler, öğretmenlerden gelen görüşler doğrultusunda sorulmaktan vazgeçilerek çıkarılmıştır. Nihayetinde 54 sorudan oluşan test, 41 maddeye düşürülmüş ve maddelerin geçerlik ve güvenirlikleri bu 41 madde ile tespit edilmeye çalışılmıştır.

Verilerin Analizi

Madde analizinden elde edilen sonuçlar, kusurlu maddelerin belirlenmesine yardımcı olurken, öğrencilerin yanlış kavramlar ile ilgili edinmiş oldukları bilgileri ve ek çalışma yapmaları gereken konuları gösterebilir (Bayrakçı, 2014). Deneme uygulamasının ardından testte yer alan her bir madde, madde seçiminde yol gösterebilecek olan testin ölçtüğü varsayılan özelliği her bir maddenin ne kadar ölçtüğünü gösteren madde ayırıcılık gücü indeksi ve her bir maddenin zorluk

derecesi ile uygun güçlük düzeyine sahip olup olmama durumunu gösteren madde güçlük indeksi kriterleri bağlamında incelenmelidir (Kan, 2017). Bu amaç ile deneme uygulamasına katılan her bir öğrencinin her maddeye vermiş oldukları cevaplar, doğru ise 1; yanlış, boş ya da birden fazla seçenek işaretlenmiş ise 0 şeklinde puanlanmaktadır (Atılgan, 2019). Bu çalışmadan elde edilen 379 tane testin madde analizinin yapılmasına karar verilmiş; elde edilen veriler, bir test ve madde analizi programı olan ITEMAN ile analiz edilmiştir. Madde analizi yapılırken 379 öğrencinin cevap kâğıtları incelenerek doğru cevaplar 1 yanlış cevaplar ise 0 olacak şekilde kodlanmış ve analizler gerçekleştirilmiştir. Bu aşamada madde güçlük ve ayırt edicilik indeksleri incelenmeye çalışılmıştır.

Araştırmanın Etik İzinleri

Yapılan bu çalışmada “Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi” kapsamında uyulması belirtilen tüm kurallara uyulmuştur. Yönergenin ikinci bölümü olan “Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemler” başlığı altında belirtilen eylemlerden hiçbirisi gerçekleştirilmemiştir.

Bulgular

Samsun ilinde öğrenim görmekte olan, 379 öğrenci tarafından doldurulan 41 soruluk deneme formu ile yapılan çalışmadan elde edilen veriler ITEMAN programı ile analiz edildiğinde madde güçlük ve ayırt edicilik indekslerine ulaşılmıştır. Başarı testinin deneme formu ile yapılan çalışmadan madde güçlük ve ayırt edicilik indeksleri Tablo 1’de sunulmuştur:

Tablo 1. Başarı testinin madde güçlük ve ayırt edicilik indeksleri

Pilot uygulama madde no	Madde güçlük indeksi	Madde ayırtıcılık indeksi	Pilot uygulama madde no	Madde güçlük indeksi	Madde ayırtıcılık indeksi
1	.52	.38	22	.71	.47
2	.67	.56	23	.44	.25
3	.83	.41	24	.77	.44
4	.62	.67	25	.89	.32
5	.89	.22	26	.79	.51
6	.64	.62	27	.30	.16
7	.78	.47	28	.67	.69
8	.71	.35	29	.87	.37
9	.53	.47	30	.79	.49
10	.81	.41	31	.61	.62
11	.69	.65	32	.81	.51
12	.82	.49	33	.64	.75
13	.67	.62	34	.59	.59
14	.54	.43	35	.85	.40
15	.91	.32	36	.77	.57
16	.41	.13	37	.85	.48
17	.85	.39	38	.51	.36
18	.82	.44	39	.80	.45
19	.62	.63	40	.86	.34
20	.30	.33	41	.76	.52
21	.75	.51			

Tablo 1 madde güçlük ve ayırtıcılık indekslerini göstermektedir. Buradan elde edilen verilere dayanarak maddeler arasından bir eleme işlemine gidilmiştir. En yüksek 1.00 en düşük 0 değerini alan madde güçlük derecesi, madde ne kadar az kişi tarafından doğru cevaplanırsa o kadar zor olarak adlandırılıp güçlük indeksi 0’a yakinken; madde ne kadar çok kişi tarafından doğru cevaplanırsa maddenin o kadar kolay olduğu ifade edilebilir ve maddenin güçlük derecesi 1’e yakın olur ve maddenin 0.40 ile 0.60 arasında bir güçlüğüünün olması ise o maddenin orta güçlükte olduğunu gösterir (Başol, 2018). 0.0 ile 1.0 arasında bir değer alabilen madde güçlük indeksinde değer sıfıra yaklaştıkça maddenin zor, bire yaklaştıkça kolay olduğu ifade edilebilir ve başarı testlerinde madde güçlük indeksinin .50 civarında olması ise istenen bir durumdur (Bayrakçeken, 2014). Bir maddenin madde güçlük indeksinin bir gruptan diğerine değişkenlik gösterebileceği ise unutulmamalıdır (İlhan, 2019).

Madde ayırt edicilik gücünün 0.40 ve üzerinde olması maddenin çok iyi bir madde olduğunu gösterir (Başol, 2018). +1 ile -1 arasında bir değer alan madde ayırt edicilik indeksi, +1’e yaklaştıkça

maddenin istenen özelliği ölçtüğü ifade edilebilir ve madde ayırt edicilik indeksi 0'a yaklaştıkça maddenin istenen özelliği ölçemediği, negatif değer aldığı anda ise maddenin başka bir özelliği ölçtüğü ifade edilebilir (Kan, 2017). Madde ayırt ediciliğinin negatif değer alması durumu daha açık bir ifade ile şu şekilde belirtilebilir: ölçülen özelliğe düşük düzeyde sahip olan bireyler maddeden daha yüksek; yüksek düzeyde sahip olan bireyler ise maddeden daha düşük puan alırlar (İlhan, 2019).

Bu çalışmada başarı testinin madde güçlük ve ayırıcılık indekslerine bakılarak, nihai teste ne çok zor ne de çok kolay olan maddelerin alınmamasına karar verilmiştir. Çok kolay olan maddelerin teste alınması durumunda ise bu maddenin ayırıcılık indeksinin yüksek olmasına dikkat edilmiştir. Nihai teste madde seçimi yapılırken madde ayırıcılık gücü .40 ve üzeri olan maddeler seçilmiş, madde ayırıcılık gücü .40'ın altında olan maddeler nihai teste alınmamıştır. Bu anlamda nihai teste madde alımı yapılırken maddelerin bilen ile bilmeyeni iyi bir şekilde ayırt etmesine dikkat edilmiştir.

Elde edilen verilere göre, 1., 5., 8., 15., 16., 17., 20., 23., 25., 27., 29., 38. ve 40. maddeler testten çıkarılmıştır. 16. ve 27. maddeler madde ayırıcılık gücünün düşük olmasının yanında, öğrenciler tarafından doğru cevabın dışındaki şıklardan bir tanesinin daha fazla doğru cevap olarak işaretlendiği maddelerdir. 13 maddenin madde güçlük ve ayırt edicilik indekslerine bakılarak testten çıkarılmasına karar verilmiştir. Böylece 28 maddelik bir başarı testi ortaya çıkmıştır.

13 madde testten çıkarıldıktan sonra analizler tekrarlanmış ve başarı testinin nihai halinin madde güçlük ve ayırt edicilik indeksleri Tablo 2'de sunulmuştur:

Tablo 2. Başarı testinin nihai halinin madde güçlük ve ayırt edicilik indeksleri

Madde	Madde güçlük indeksi	Madde ayırıcılık indeksi	Madde	Madde güçlük indeksi	Madde ayırıcılık indeksi
1	.66	.60	15	.70	.44
2	.82	.40	16	.75	.47
3	.61	.59	17	.78	.50
4	.63	.57	18	.66	.66
5	.78	.48	19	.78	.50
6	.53	.45	20	.60	.61
7	.81	.38	21	.80	.49
8	.69	.65	22	.63	.75
9	.81	.49	23	.58	.62
10	.67	.62	24	.84	.41
11	.54	.44	25	.76	.54
12	.81	.43	26	.83	.47
13	.61	.62	27	.79	.42
14	.74	.55	28	.74	.51

Tablo 2 incelendiğinde, 7. madde dışındaki maddelerin tamamının madde ayırıcılık gücünün .40'ın üzerinde olduğu dolayısıyla ayırt ediciliğinin yüksek olduğu görülebilir. 7. madde de madde ayırıcılık gücü bakımından iyi bir maddedir. Testin nihai halinin bilen ile bilmeyeni ayırt etme gücü bu anlamda yüksektir. Madde güçlük indekslerinin ise birbirinden farklı olması sağlanmaya çalışılarak testte güçlük derecesi farklı maddelerin yer alması sağlanmaya çalışılmıştır. Böylece farklı

zorluk derecelerine sahip, aynı zamanda bilen ile bilmeyeni ayırt edebilen bir test geliştirilmiştir.

Testin istatistiki değerleri Tablo 3'te sunulmaktadır:

Tablo 3. Nihai ölçeğin istatistikleri

Değerler	Sonuçlar
Madde sayısı	28
Ölçeğin uygulandığı kişi sayısı	379
Ortalama	19.966
Varyans	40.788
Standart sapma	6.387
Çarpıklık	-0.758
Basıklık	-0.395
Minimum	3.000
Maksimum	28.000
Ortanca	21.000
Alpha	0.897
SEM	2.046
Ortalama güçlük	0.713
Ortalama ayırt edicilik	0.519
Çift serili ortalaması	0.705
Max Score (Low)	17
N (Low Group)	115
Min Score (High)	25
N (High Group)	122

Tablo 3 incelendiğinde, nihai testin alpha değerinin 0.897 olduğu dolayısıyla testin oldukça güvenilir olduğu görülmektedir. 0.713 olan ortalama güçlük derecesi bu testin testin kolay bir test olduğunu, 0.519 olan ortalama ayırt edicilik seviyesi ise testin bilen ile bilmeyen öğrencileri ayırt ettiğini göstermektedir. Bu anlamda 6. sınıf Üretim, Dağıtım ve Tüketim öğrenme alanına yönelik geçerli ve güvenilir bir test elde edildiği ifade edilebilir.

Tartışma, Sonuç ve Öneriler

Üretim, Dağıtım ve Tüketim öğrenme alanı Sosyal Bilgiler ders programı içinde 6 haftalık bir süreyi kapsamakta, bu süre içinde 6 kazanımın öğrencilere verilmesi amaçlanmaktadır. Bu anlamda bu öğrenme alanına ilişkin kavramların öğrencilere öğretilmesi için çalışmalar yapılmış (Dere ve Aktaşlı, 2019; Ada, 2020; Kasım, 2019) ve başarı testleri geliştirilmiştir (Uzun, 2019; Akgün ve Arıkan, 2015). Ancak yapılan her bir çalışma birbirinden farklı yapılar arz etmekte, geliştirilen başarı testleri de kendi içinde özgün yapılara sahip olmaktadır. Bu anlamda her bir araştırmacının geliştirdiği başarı testleri de farklı yapılarda olmaktadır. Bu çalışmada 6. sınıf düzeyindeki öğrencilerin, Üretim, Dağıtım ve Tüketim öğrenme alanına yönelik başarı durumlarını tespit edebilecek bir başarı testi geliştirilmiştir. Bu test nihai hali ile 28 çoktan seçmeli sorudan oluşan, öğrenme alanında bulunan her bir kazanıma yönelik soru içeren yapısı ile geçerli ve güvenilir bir başarı testidir. Bu öğrenme alanında hazırlanan başarı testleri olsa da (Akgün ve Arıkan, 2015) çalışmamızda geliştirilen başarı testi, öğrencilerin ön ve son bilgilerini yoklayabilecek nitelikte, öğrenciler ile uygulanacak olan uygulamalı bir eğitimin öğrenme çıktılarını ölçmeye yarayabilecek özgün olarak geliştirilmiş bir başarı testidir. Bir

başarı testinde yer alan soruların çoktan seçmeli olmasının öğrenciler ve değerlendirme süreci açısından, objektif puanlama yapılabilmesi, çok sayıda bireye kısa sürede sınava katılma şansı tanınması, kolay ve çok kısa bir sürede puanlama yapılabilmesi gibi avantajları bulunmaktayken, bu maddelerin her yaş seviyesinden öğrenciye uygulanması mümkündür (Çakan, 2017). Bunların yanı sıra çoktan seçmeli testlerin uygulandığı sınavlarda öğrenciler vakitlerinin çoğunu maddeyi okuma ve cevabı bulmaya harcamakta dolayısıyla sınav süresini cevap yazmakla harcamak durumunda kalmamaktadırlar (Güler, 2019). Her ne kadar uygulama sürecinden kaynaklı zorlukları bulunmakta ise de çoktan seçmeli testler ölçme ve değerlendirme süreci için oldukça pratik testlerdir.

Bu çalışmada 6. sınıf Üretim, Dağıtım ve Tüketim öğrenme alanına yönelik 28 maddeden oluşan geçerli ve güvenilir bir başarı testi elde edilmiştir. Bu başarı testi, öğretmenler tarafından öğrencilerin ön ve son bilgilerinin yoklanmasıyla kullanılabilir. Testte verilen yanıtlardan yola çıkarak öğrencilerin hangi konularda eksik öğrenmeleri olduğunu tespit eden öğretmenin öğretim sürecini buna göre planlayabilmesi veya öğrencilere, öğrenme çıktılarına uygun dönütler vermesi mümkün olabilir. 6. sınıf Üretim, Dağıtım ve Tüketim öğrenme alanı kazanımlarına yönelik 6 haftalık ders işlenip, bu derslerin öncesinde ön test, sonrasında son test olarak kullanılacak nitelikteki bu test ile öğrencilerin başarılarının süreç içindeki seyri de ortaya konulabilir. Geliştirilmiş olan test, farklı şubelerde öğrenim görmekte olan 6. sınıf öğrencilerinin akademik başarılarını ortaya koyarak bu şubelerin kendi aralarındaki başarı seviyelerindeki farklılıklarının ortaya konulmasında kullanılabilir.



<http://kefad.ahievran.edu.tr>

Ahi Evran University Journal of Kırşehir Education Faculty

ISSN: 2147 - 1037

ENGLISH VERSION

Introduction

Measurement has a different structure depending on the feature to be measured. For measurement, the feature to be measured should be described. A cluster of numbers or symbols suitable for the feature to be measured should be determined. Rules should be given showing which number or symbols are given to the feature to be measured (Yaşar, 2017).

The stages of the measurement process can be shown as follows:

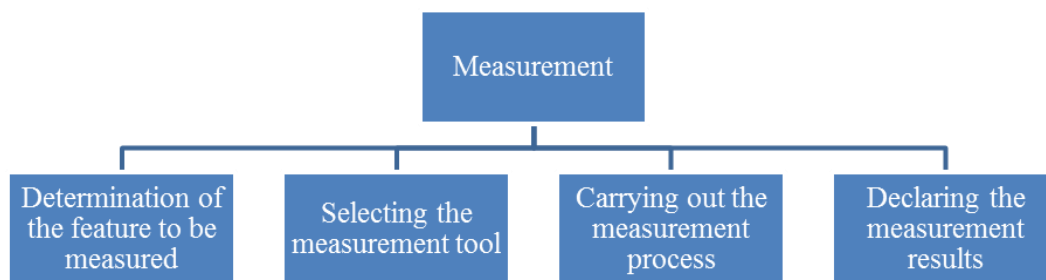


Figure 1. Stages of the measurement process (Yaşar, 2017).

All measurements serve as an evaluation; the process of reaching judgment and decision based on a criterion is expressed as evaluation (Semerci, 2014). At the end of the evaluation, a judgment is reached about the measured feature or quality like more or less, sufficient or insufficient, acceptable or unacceptable (Semerci, 2014). Through evaluation, it is possible to increase the level of success, to find out how different effective methods are and to provide quality control in education, while effective teaching also takes place (Başol, 2018).

The place of measurement and evaluation in the teaching process can be shown as follows:

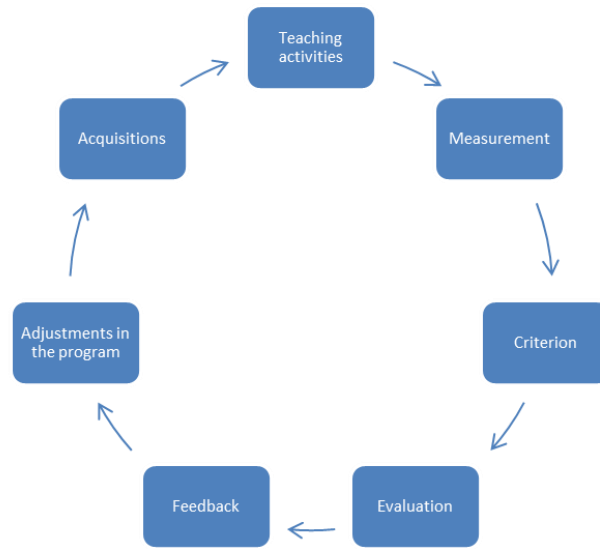


Figure 2. The place of measurement and evaluation in the teaching process (Başol, 2018).

As can be understood from this cycle, measurement and evaluation are a whole. Beyond being a whole, measurement and evaluation, each step of which is closely related to the other, is an indispensable part of the education process. In other words, measurement and evaluation concepts are evaluated as processes that are mentioned together and that are mentioned together and complement each other; evaluation requires a measurement (Özer Özkan, 2019). The extent to which educational activities are successful, the extent to which the progress made during the year or during the term is successful in helping the teacher gain teaching goals and guiding students during the teaching process is determined by measurement and evaluation (Kilmen, 2019).

While it is very important in terms of education for students to have certain skills so that they are recognized for their affective, cognitive and psychomotor characteristics and they establish a balanced communication with each other and their environment, measurement tools are used to make observable the features that cannot be observed and to find out whether the features that are desired to be gained in education are gained or not (Kan, 2017). Students' success can be measured with measurement tools used in exams and while these measurement tools can be exams that require long answers, various performance tests, short answer tests and multiple choice tests, they can also be in the form of oral exams (Semerci, 2014). Measurement tools facilitate the measurement process and they have characteristics that determine the qualities of the results obtained as a result of measurement; these are the mathematical characteristics of measurement results and they define a scale (Kan, 2019). Since optical reader or keys score objective exams, they are exams that have high scoring reliability and subjectivity is not a threat in such exams (Başol, 2018). Items that can be scored objectively can be in the form of open or close-ended questions, it can be stated that close-ended questions are close to interpretation and it can be stated that in such questions those who answer the question have to choose the answer from the ones given and multiple choice tests are the most used among these tests (Başol, 2018). Multiple choice tests can be preferred in cases when a large number

of questions are asked, when selection and placement is required, when there is enough time to prepare the test, when the results should be reported quickly and when the validity and reliability of the test should be high (Güler, 2019).

While it is possible to measure different kinds of skills and knowledge and high level mental skills, to apply the tests to a large number of people, to prepare different kinds of questions ranging from the level of knowledge to the level of assessment, to measure high level knowledge and skill areas and to ask too many questions with multiple choice items, it is also possible to increase content validity and the scores to be taken from the test (Çakan, 2017). In addition to these, some factors make it challenging to prepare a multiple choice exam; these can be stated as the necessity of writing questions and choices very clearly and concisely, emphasizing negative roots, item roots and distractors' not including clues, the inability of students to eliminate distractors easily, avoiding using the same choice as the correct answer successively and the necessity of the answer key not having a pattern (Aydın, 2019). The fact that the correct answer is within the item most of the time, the factor of luck, the fact that it is difficult to measure advanced level behaviours and the recall period is long, complicated and requires expertise are also among the disadvantages of multiple choice tests (Güler, 2019). In addition to these disadvantages, the variable of reading speed is also included in measurement as the number of questions increases and although it is possible to correct through statistical means, it is not possible to eliminate the factor of luck totally (Turgut and Baykul, 2019). Despite all these disadvantages, considering the practicality in the process of application, multiple choice tests are indispensable tests that will always be used. The aim of this study is to develop a valid and reliable achievement test for 6th grade, consisting of practical, different types of questions that can be used in the education process. There are achievement tests developed for this learning area (Akgün and Arıkan, 2015); however, the fact that social studies lesson is multifaceted and has a wide scope of subject area allows for writing many different questions and developing measurement tools. Measuring and evaluating a subject area with different questions can also be possible with the development of different measurement tools. Thus, it is thought that the 6th grade social sciences lesson production, distribution and consumption learning area can be enriched in terms of measurement tool.

Purpose of the study

Social studies lesson, which is an interdisciplinary lesson that requires students to think multi-dimensionally, is a very comprehensive lesson suitable for teaching with different teaching models and approaches. Therefore, various measurement and evaluation approaches can also be used. One of the tests that can be used in line with these approaches is the multiple choice test. Considering the previously expressed advantages of multiple choice tests, the aim was to develop a valid and reliable

multiple choice achievement test for the 6th grade production, distribution and consumption learning area, which addressed all acquisitions in this learning area.

Method

This study aimed to develop an achievement test and therefore survey method was used. Since a survey is conducted with a large amount of data on large populations, it is a method based on quantitative data and statistical analysis of data (Özdemir, 2015).

Study Group

This study was conducted in the spring term of the 2018-2019 academic year and valid data were reached from 379 students who were studying in Samsun province. Participation was voluntarily in the study and the process was carried out with students who volunteered to participate in the study. Before the study was conducted with 379 students, an application was made with two students. The clarity and comprehensibility of the questions were tested. The duration of the exam was determined as 65 minutes in line with the feedback from the students. Following this, the 41-question achievement test was given to 379 students, and before the students started to answer the test, the students were motivated by making the necessary explanations.

Data Collection Tool

In this scale development study, achievement test for 6th grade production, distribution and consumption learning area developed by the researcher was used as data collection tool. The purpose in developing this test was to develop a tool that could find out the academic achievement of students in the related learning area. The steps to be considered in the preparation of the achievement test are as follows:

Preparing the test plan

Forming the item pool

Reviewing the items

Applying the trial form

Item analysis (Başol, 2018: 236-245).

In this study, first of all, a test plan was prepared. It was aimed for the developed test to show whether 6th graders reached the acquisitions of the production, distribution and consumption learning area. While forming the item pool, the objectives included in the curriculum of the Ministry of National Education (URL1) were taken into account. While preparing questions that were in line with the revised Bloom taxonomy, example questions, text books, example questions in exams were examined, and an item pool was created. However, no questions were taken directly as they were, all of the questions were prepared by the researcher. It was determined in which cognitive process and

knowledge dimension step the six acquisitions found here were according to the revised Bloom taxonomy and a table of specifications was formed. 9 questions were prepared for each acquisition, and an item pool of 54 questions was created. The developed achievement test was submitted to expert examination regarding whether they were suitable for cognitive steps, rules of language and expression and students' levels. After expert examination, corrections were made in wording in some of the items in a way that did not change the step of the question in Bloom taxonomy. Items 4, 5, 9, 16, 24, 27, 33, 35, 40, 45, 48, 53 and 54 were not used in line with the views taken from teachers. Finally, the 54 question test was reduced to 41 items and the validity and reliability of the items were found with these 41 items.

Data Analysis

The results obtained from item analysis can help to identify defective items and also show the knowledge that students have acquired about misconceptions and the subjects they need to study more (Bayrakçeken, 2014). After the trial application, each item in the test should be examined in terms of item discrimination power index which shows to what extent each item measures the feature that the test is supposed to measure and item difficulty index criteria, which show the difficulty level of each item and whether each item has the suitable difficulty level (Kan, 2017). With this purpose, the answers of each student who participated in the trial application given to each item are scored as 1 if they are correct and as 0 if the answer is incorrect, not given or if more than one choice is selected (Atılğan, 2019). It was decided to perform item analysis of 379 tests obtained from this study and the data obtained were analyzed with ITEMAN, which is a test and item analysis program. While performing item analysis, answer sheets of 379 students were examined, correct answers were coded as 1 and incorrect answers were coded as 0 and the analyses were carried out. Item difficulty and discrimination indices were examined at this stage.

Ethical Considerations

All of the rules specified to be followed within the scope of "Higher Education Institutions Scientific Research and Publication Ethics Directive" were complied within this study. None of the actions specified under the heading "Actions Contrary to Scientific Research and Publication Ethics" in the second part of the directive was taken in the study.

Results

When the data obtained from the study conducted with 41question trial form filled in by 379 students studying in Samsun province were analyzed with the ITEMAN program, item difficulty and discrimination indices were reached. Item difficulty and discrimination indices obtained from the study conducted with the trial form of the achievement test are shown in Table 1.

Table 1. *Item difficulty and discrimination indices of the achievement test*

Pilot application number of item	Item difficulty index	Item discrimination index	Pilot application number of item	Item difficulty index	Item discrimination index
1	.52	.38	22	.71	.47
2	.67	.56	23	.44	.25
3	.83	.41	24	.77	.44
4	.62	.67	25	.89	.32
5	.89	.22	26	.79	.51
6	.64	.62	27	.30	.16
7	.78	.47	28	.67	.69
8	.71	.35	29	.87	.37
9	.53	.47	30	.79	.49
10	.81	.41	31	.61	.62
11	.69	.65	32	.81	.51
12	.82	.49	33	.64	.75
13	.67	.62	34	.59	.59
14	.54	.43	35	.85	.40
15	.91	.32	36	.77	.57
16	.41	.13	37	.85	.48
17	.85	.39	38	.51	.36
18	.82	.44	39	.80	.45
19	.62	.63	40	.86	.34
20	.30	.33	41	.76	.52
21	.75	.51			

Table 1 shows the item difficulty and discrimination indices. Based on the data obtained here, some of the items were excluded. Item difficulty level has the highest value of 1.00 and the lowest value of 0. It can be stated that the less an item is answered correctly, the more difficult that item is and the difficulty index is close to 0, while the more an item is answered correctly, the easier that item is and the difficulty index is close to 1 and a difficulty index between 0.40 and 0.60 shows that the item has moderate difficulty (Başol, 2018). In item difficulty index, which is between 0.0 and 1.0, the item can be stated to be difficult as it gets closer to 0 and easy as it gets close to 1 and an item difficulty index of about .50 is a desirable level in achievement tests (Bayrakçıken, 2014). It should also be remembered that the item difficulty index of an item can be different from one group to another (İlhan, 2019).

An item discrimination power of 0.40 and higher shows that the item is very good (Başol, 2018). Item discrimination index can have a value between +1 and -1; as this index gets closer to +1, it can be stated that the item measures the desired feature, and as this index gets closer to 0, it can be stated that the item does not measure the desired feature. It can be stated that the item measures another feature when it has a negative value (Kan, 2017). Negative item discrimination value can be expressed more clearly as follows: individuals who have a lower level of the measured feature get a higher score from the item. In comparison, individuals who have higher level of the measured feature get lower scores from the item (İlhan, 2019).

In this study, item difficulty and discrimination indices of the achievement test were examined and it was decided not to include items that were either too easy or too difficult. In case of including too easy items, care was taken for these items to have a high discrimination index. While choosing items for the final test, the items with a discrimination power of $\geq .40$ were chosen and the items with an item discrimination power of $\leq .40$ were not included in the final test. In this sense, while including items in the final test, care was taken for the items to discriminate well between the participants who knew the answers and those who did not.

According to the data obtained, items 1, 5, 8, 15, 16, 17, 20, 23, 25, 27, 29, 38 and 40 were excluded from the test. In addition to having low item discrimination power, items 16 and 27 were items in which one of the incorrect answers was chosen more than the correct answer by students. These 13 items were decided to be excluded from the test by checking their item difficulty and discrimination indices. Thus, a 28 item achievement test was obtained.

The analyses were repeated after 13 items were excluded from the test and item difficulty and discrimination indices of the achievement test's final version are shown in Table 2:

Table 2. *Item difficulty and discrimination indices of the achievement test's final version*

Item	Item difficulty index	Item discrimination index	Item	Item difficulty index	Item discrimination index
1	.66	.60	15	.70	.44
2	.82	.40	16	.75	.47
3	.61	.59	17	.78	.50
4	.63	.57	18	.66	.66
5	.78	.48	19	.78	.50
6	.53	.45	20	.60	.61
7	.81	.38	21	.80	.49
8	.69	.65	22	.63	.75
9	.81	.49	23	.58	.62
10	.67	.62	24	.84	.41
11	.54	.44	25	.76	.54
12	.81	.43	26	.83	.47
13	.61	.62	27	.79	.42
14	.74	.55	28	.74	.51

When Table 2 is examined, it can be seen that all of the items except item 7 have a discrimination power higher than .40; therefore, they have high discrimination. Item 7 is also good in terms of item discrimination power. In this sense, the test has a high power to discriminate between those who know the answers and those who do not. The study aimed to include items that had different difficulty degrees by trying to have different item difficulty indices. Thus, a test with different difficulty degrees was developed to discriminate between those who knew the answer and those who did not. Statistical values of the test are shown in Table 3:

Table 3. Statistics of the final scale

Values	Results
Number of items	28
Number of individuals to whom the scale was applied	379
Mean	19.966
Variance	40.788
Standard deviation	6.387
Skewness	-0.758
Kurtosis	-0.395
Minimum	3.000
Maximum	28.000
Median	21.000
Alpha	0.897
SEM	2.046
Mean difficulty	0.713
Mean discrimination	0.519
Mean biserial	0.705
Max Score (Low)	17
N (Low Group)	115
Min Score (High)	25
N (High Group)	122

When Table 3 is examined, it can be seen that alpha value of the final test is 0.897, which shows that the test is quite reliable. Mean difficulty level, which is 0.713, shows that the test is easy, and mean discrimination level, which is 0.519, shows that the test discriminates students who know from those who do not. In this sense, it can be stated that a valid and reliable test for the 6th grade Production, Distribution and Consumption learning area was obtained.

Discussion, Conclusion and Recommendations

Production, Distribution and Consumption learning area covers a period of 6 weeks within Social Studies curriculum and during this period, the objective is to teach 6 acquisitions to students. In this sense, studies have been conducted to teach students concepts about this learning area (Dere and Aktaşlı, 2019; Ada, 2020; Kasım, 2019) and achievement tests have been developed (Uzun, 2019; Akgün and Arıkan, 2015). However, each study conducted has different structures, and the developed achievement tests have unique structures within themselves. In this sense, the achievement test developed by each researcher has different structures. An achievement test was developed to determine the achievement of 6th grade students in Production, Distribution and Consumption learning area. In its finalized form, this test is a valid and reliable achievement test that includes 28 multiple choice questions, with its structure including questions for each acquisition in the learning area. Although there are achievement tests in this learning area (Akgün and Arıkan, 2015), the achievement test developed in our study is a uniquely developed achievement test that can examine the pre and post-knowledge of students and that can measure the learning outcomes of an applied education with students. While having multiple choice questions in an achievement test has advantages for students and the evaluation process, such as objective scoring, providing the large

number of individuals to take the exam in a short time and easy scoring in a concise time, it is also possible to apply these items to students of all ages (Çakan, 2017). In addition to these, in exams with multiple choice questions, students spend most of their time reading the item and finding the answer, and therefore they do not have to spend the exam time writing answers (Güler, 2019). Although they have difficulties resulting from the application process, multiple choice tests are very practical tests for the process of measurement and evaluation.

In this study, a 28 item valid and reliable achievement test was obtained for 6th grade Production, Distribution and Consumption learning area. This achievement test can be used in examining students' pre and post-knowledge. Based on the answers given to the test, it can be possible for teachers who find out in which areas students have difficulties to plan the teaching process according to this or to give students feedback suitable for learning outcomes. With this test which can be used by teachers as pre-test and post-test before and after 6 weeks of teaching on 6th grade Production, Distribution and Consumption learning area acquisitions, the course of students' achievement in the process can also be shown. The developed test can be used to find out the academic achievement of 6th grades in different classes and show the differences between the students' achievements in these classes.

References

- Ada, S. (2020). 4. sınıf Sosyal Bilgiler dersi üretim, dağıtım ve tüketim öğrenme alanında kavram karikatürü kullanımının ders başarısına etkisi. Yayınlanmamış Yüksek Lisans Tezi, Düzce Üniversitesi, Sosyal Bilimler Enstitüsü, Düzce.
- Akgün, İ. H. & Arıkan, İ. (2015). Altıncı sınıf öğrencilerinin sosyal bilgiler dersi üretim, dağıtım ve tüketim öğrenme alanına yönelik akademik başarılarının çeşitli değişkenler açısından incelenmesi. *The Journal of International Social Research*, 8(40), 511-518.
- Atılğan, H. (2019). Test geliştirme. İçinde Atılğan, H. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss. 281-314). Ankara: Anı Yayıncılık.
- Aydın, B. (2019). Çoktan seçmeli testler. İçinde Atılğan, H. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss.201-232). Ankara: Anı Yayıncılık.
- Başol, G. (2018). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Bayrakçıken, S. (2014). Test geliştirme. İçinde Karip, E.(Ed.), *Ölçme ve değerlendirme* (ss.291-322). Ankara: Pegem Akademi.
- Çakan, M. (2017). Eğitim sistemimizde yaygın olarak kullanılan sınav türleri. İçinde Tekindal, S. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss. 87-122). Ankara: Pegem Akademi.
- Dere, İ. & Aktaşlı, İ. (2019). Üretim, dağıtım ve tüketim kavramlarının öğretimine ilişkin bir eylem araştırması. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 23(3), 1389-1406.
- Güler, N. (2019). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- İlhan, M. (2019). Madde analizi ve madde ile test istatistikleri arasındaki ilişki. İçinde Çetin, B. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss.269-298). Ankara: Anı Yayıncılık.
- Kan, A. (2017). Ölçme aracı geliştirme. İçinde Tekindal, S. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss.241-277). Ankara: Pegem Akademi.
- Kan, A. (2019). Ölçmenin temel kavramları. İçinde Atılğan, H. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss. 19-42). Ankara: Anı Yayıncılık.
- Kasım, Ş. (2019). Üretim, dağıtım, tüketim öğrenme alanına yönelik react modeline göre hazırlanan rehber materyallerin etkililiğinin değerlendirilmesi. Yayınlanmamış Yüksek Lisans Tezi, Trabzon Üniversitesi, Lisansüstü Eğitim Enstitüsü, Trabzon.
- Kilmen, S. (2019). Ölçme ve değerlendirmede temel kavramlar. İçinde Çıkrıkçı, R. N. (Ed.), *Eğitimde ölçme ve değerlendirme* (ss.25-56). Ankara: Anı Yayıncılık.
- Özdemir, E. (2015). Tarama yöntemi. İçinde Metin, M. (Ed.), *Kuramdan uygulamaya eğitimde bilimsel araştırma yöntemleri* (77-97). Ankara: Pegem Akademi.
- Özer Özkan, Y. (2019). Ölçme ve değerlendirme ile ilgili temel kavramlar. İçinde Çetin, B. (Ed.), *Eğitimde ölçme ve değerlendirme* (s.1-22). Ankara: Anı Yayıncılık.

- Semerci, Ç. (2014). Eğitimde ölçme ve değerlendirme. İçinde Karip, E.(Ed.), *Ölçme ve değerlendirme* (ss.1-15). Ankara: Pegem Akademi.
- Turgut, M. F. & Baykul, Y. (2019). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Uzun, E. (2019). *Ters yüz sınıf modelinin 7. sınıf Sosyal Bilgiler dersi üretim dağıtım ve tüketim ünitesinde uygulanmasının akademik başarıya etkisinin incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Aksaray Üniversitesi, Sosyal Bilimler Enstitüsü, Aksaray.
- URL1: *Sosyal Bilgiler Dersi Öğretim Programı* (İlkokul ve Ortaokul 4, 5, 6 ve 7. Sınıflar). (2018). Retrieved from <http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=354>
- Yaşar, M. (2017). Ölçme ve değerlendirme ile ilgili temel kavramlar. İçinde Tekindal, S. (Ed.), *Eğitimde ölçme ve değerlendirme* (s.9-40). Ankara: Pegem Akademi.