



Comparing the psychometric characteristics of multiple choice tests and structural communication grids *

Esin BAĞCAN BÜYÜKTURAN **

Nükhet ÇIKRIKÇI DEMİRTAŞLI ***

ABSTRACT. Structural communication grid is a technic which is developed to minimize the limitations of multiple choice tests. The purpose of this study is to compare the psychometric characteristics of multiple choice test and structural communication grid which aim to test the same construct. The research was carried on 276 seventh grade students. Findings point out that structural communication grid items are easier than multiple choice test items. It has been found that structural communication grid is more reliable than multiple choice test, but no significant difference has been found between the criterion validity of two tests' scores.

Keywords: Structural communication grid, multiple choice items, psychometric characteristics.

* This study composed of the master thesis called "Comparing the psychometric characteristics of multiple choice tests and structural communication grids".

** Res. Assist. Dr., Abant İzzet Baysal University, Faculty of Education, Department of Measurement and Evaluation, Bolu, Turkey. E-mail: ybagcan@yahoo.com

*** Prof. Dr., Ankara University, Faculty of Educational Sciences, Department of Measurement and Evaluation, Ankara, Turkey. E-mail: mukhet@yahoo.com

SUMMARY

Purpose and significance: Educators have long been trying to get rid of the limitations of multiple choice tests. So they have been attempting to develop new item types, in addition to multiple choice item type. One of these types is structural communication grids. Educators discuss that structural communication grid method would get rid of limitations of multiple choice tests, like ignoring partial knowledge, effect of success by chance and lacking feedback about how students structured knowledge and they conducted various studies. However, there are no empirical studies on the comparison of multiple choice tests with structural communication grids. The purpose of this study is to compare the psychometric characteristics of multiple choice test and structural communication grid which aim to test the same construct.

Method: The research was carried on 276 seventh grade students at three primary schools. The multiple choice test has 16 items and the structural communication grid has 16 items. Those items which had been developed according to the content of the first unit of Science and Technology subject, was used in the research. The situation of students about taking multiple choice test or structural communication grid was determined randomly. The students' achievement scores related to the unit were obtained from the classroom teachers and they were taken as a validity criterion.

Results: Findings point out that structural communication grid items are easier than multiple choice test items. No significant difference has been found between the criterion validity of two tests' scores. It has been realized that structural communication grid is more reliable than multiple choice test. Consequently, it has been found out that the variance of multiple choice test scores is significantly higher than the variance of structural communication grid scores, and the mean of structural communication grid scores is significantly higher than the mean of multiple choice test scores.

Discussion and Conclusions: At the end of the study, it was suggested that more resources should be allocated to education in order to improve infrastructure, physical and technological facilities at schools opportunities should be provided to teacher in order to participate in in-service trainings about teaching and assessment methods and techniques, programs should be renewed and improved by considering the results of research and implications and besides the experimental studies aiming to evaluate programs, studies based on observations and interviews should be conducted.



Çoktan Seçmeli Testler ile Yapılandırılmış Gridlerin Psikometrik Özellikleri Bakımından Karşılaştırılması *

Esin BAĞCAN BÜYÜKTURAN **

Nükhet ÇIKRIKÇI DEMİRTAŞLI ***

ÖZ. Yapılandırılmış grid, çoktan seçmeli testlerin sınırlılıklarını azaltmak için geliştirilmiş bir tekniktir. Bu çalışmada aynı davranışları ölçmek üzere geliştirilmiş çoktan seçmeli test ile yapılandırılmış gridin psikometrik özellikler açısından karşılaştırılması amaçlanmıştır. Araştırma 7.sınıfta okumakta olan 276 ilköğretim öğrencisi üzerinde yürütülmüştür. Araştırmadan elde edilen bulgular, yapılandırılmış grid maddelerinin çoktan seçmeli maddelere göre daha kolay olduğunu göstermektedir. Araştırmada, yapılandırılmış grid testinden elde edilen ölçümlerin, çoktan seçmeli testten elde edilen ölçümlerden daha güvenilir olduğu, ancak iki testten elde edilen ölçümlerin ölçüt geçerlikleri arasında anlamlı fark olmadığı belirlenmiştir.

Anahtar Sözcükler: Yapılandırılmış grid, çoktan seçmeli maddeler, psikometrik özellikler.

* Bu çalışma I. Ulusal Eğitim Programları ve Öğretim Kongresi'nde sözlü bildiri olarak sunulmuştur.

** Arş. Gör. Dr., Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Ölçme ve Değerlendirme Anabilim Dalı Bolu, Türkiye. E-posta: ybagcan@yahoo.com

*** Prof. Dr., Ankara Üniversitesi Eğitim Bilimleri Fakültesi Ölçme ve Değerlendirme Anabilim Dalı Ankara, Türkiye. E-posta: mukhet@yahoo.com

GİRİŞ

Eğitim sürecini oluşturan temel boyutlar: eğitim-öğretim uygulamaları, öğretim programları ve ölçme ve değerlendirme etkinlikleridir. Bu boyutlarda yürütülen uygulamaların niteliği eğitim sürecinin de niteliğini etkiler. Bunlar içinde ölçme ve değerlendirme boyutu, öğrencinin öğrenmelerinin değerlendirilmesi ve niteliğine ilişkin geri bildirim sağlayan önemli bir boyuttur. McDonald (2002), eğitim ortamında doğru kararlar almanın geçerli ölçmeler ile mümkün olabileceğini, bu sebeple eğitimcilerin ölçmelerinde doğru aracı seçmeleri gerektiğini belirtmiştir.

Ölçme aracının geçerli ve güvenilir sonuçlar vermesinin yanında, uygulanması ve puanlanmasının kullanışlı olması tercih edilir. Uygun araç seçilirken, ölçme işleminin amacı, ölçme işlemine tabi tutulacak grup, aracın kapsamı gibi birçok durum göz önüne alınır.

Öğrenci başarısının ölçülmesinde kullanılan madde türleri açık uçlu maddeler, kısa cevap gerektiren maddeler, sınıflama gerektiren maddeler ve çoktan seçmeli maddelerdir. Açık uçlu maddelerde cevaplayıcı cevabı düşünür, bulur, tasarlar ve düzenler. Kısmi bilgiyi puanlamaya imkan verir ancak puanlaması güçtür ve puanlayıcı kanısından etkilenir. Kısa cevap gerektiren maddelerde cevaplayıcılar en fazla bir cümleden oluşan kısa cevaplar verir. Bu maddelerin puanlanması tamamen objektif değildir. Sınıflama gerektiren maddelerde yanıtlayıcıdan verilen önermeleri “doğru-yanlış” olarak iki ya da daha fazla sınıfa ayırması istenir. Şans başarısına açık olmasının yanında puanlaması kolay ve objektif bir yöntemdir. Çoktan seçmeli maddelerde doğru cevap seçeneklerin arasında verilmiştir ve yanıtlayıcı bunu bulmakla yükümlüdür (Turgut, 1992). Çoktan seçmeli maddeler objektif ve kolay puanlandıkları için sıkça tercih edilirler. Ancak çoktan seçmeli maddeler, 0-1 biçiminde puanlandıkları için kısmi bilgiyi yoklamada yetersizdir. Ayrıca öğrenciye anahtar yanıtın da içinde bulunduğu yapılandırılmış seçenekler hazır olarak sunulduğu için şans başarısına da açıktır (Turgut, 1992; McDonald, 2002; Abravomitz and Berezima, 2004; Çelen ve Çıkrıkçı-Demirtaşlı, 2006). Bu sebeple eğitimciler çoktan seçmeli testlerin sınırlılıklarını giderme çabası içindedirler.

Örneğin düzeltme formülü öğrencilerin çoktan seçmeli testlerde yanıtını bilmedikleri maddeleri tahminle yanıtlama girişimlerine engel olmak için geliştirilmiştir (Horst, 1966). Ancak geliştirilen bu formülün de eleştiri alan yönleri olmuştur. Düzeltme formülü uygulaması; özellikle, öğrencilerin tamamen emin olmadıkları maddeleri yanıtlamaktan kaçındıkları düşüncesiyle eleştiri almıştır (Ebel, 1965; Umay, 1998). Oysa öğrenme “ya hep ya hiç” meselesi değildir. Buna göre, bireylerden bildiklerini bilmediklerinden açıkça ayırmasını ve yalnızca kesinlikle emin oldukları

maddeleri yanıtlamasını istemek mantıklı değildir. Her problemin çözümü bir miktar akılcı tahmin gerektirir. Kararsızlığa düşmek ya da kararı ertelemek yerine doğruya yakın kararlar verilmesi gerekebilir. Ancak doğru yanıt konusunda hiç fikri olmayan ve maddeyi şansını kullanarak doğru yanıtlamak isteyen öğrencilerin maddelere yanıt verme eğilimlerinden söz ediliyorsa “tahmin” konusuna olumsuz bakılmalıdır (Ebel, 1965).

Yapılan araştırmalara göre çoktan seçmeli maddelerin sınırlılıklarına rağmen yaygın olarak kullanılmasının en önemli sebebi puanlanmasının kolay olmasıdır (Scouller, 1998, Bacon, 2003; Swartz, 2006). Eğitimciler çoktan seçmeli madde türünün üstünlükleri yanında bazı sınırlılıklarının var olması sebebiyle farklı madde türlerini kullanmayı da tercih etmişlerdir. Çoktan seçmeli maddelerin yanı sıra kullanılabilir bir madde türü de “yapılandırılmış grid” tir. Yapılandırılmış grid, bir konuda hazırlanmış birkaç sorunun cevaplarının numaralandırılmış dokuz ya da on iki gözenekli bir tabloya dağıtılmasıyla oluşturulur. Öğrencilerden soruların doğru yanıtlarını oluşturan seçenekleri bulmalarının yanında bu seçenekleri mantıksal olarak sıralamalarını da gerektiren bir tekniktir. Bu türün en önemli özelliği, anlamlı öğrenmeyi ölçmesi, kısmi bilgileri gözetmesi, öğrencinin bilişsel yapısındaki eksiklik ve yanlışlıkları ortaya koyan bir teşhis aracı niteliği taşımasıdır (Bahar, Öztürk ve Ateş, 2002). Yapılandırılmış gridin hazırlanmasında, öğretmen kendine söz konusu davranışı ölçecek bir soru sorar ve sorunun yanıtlarını tablodaki gözeneklere gelişigüzel yerleştirir. Daha sonra öğretmen kendine başka bir davranışı ölçecek ikinci bir soru sorar ve bu sorunun da yanıtlarını gözeneklere yerleştirir. Ancak bu sorunun yanıtı durumundaki ifadelerden halihazırda gözeneklerde yer alanlar varsa bunları tekrar başka gözeneklere yerleştirmez. Bu durumda birden fazla sorunun da ortak yanıtı durumundaki bazı gözenekler mevcut olabilir. Tablo dolana kadar yanıtları boş gözeneklere yerleştirme işlemine devam edilir. Daha sonra bu biçimde doldurulmuş olan grid öğrencilere verilir. Öğrenciden her sorunun cevabı olan gözeneklerin numaralarını seçmeleri, bazı sorularda ise bu numaraları mantıksal olarak sıralamaları istenir. Böylece öğretmen öğrencinin bilgiyi nasıl yapılandırıldığını da görür. Gridlerdeki gözenek sayısı sınıf seviyesine göre belirlenir. İlköğretim ikinci kademedan üniversite birinci sınıfa kadar olan öğrencilerde dokuz, lisans düzeyinden itibaren on iki gözenekli grid kullanılır (Johnstone, Bahar ve Hansell, 2000).

Şekil 1’de, bu çalışmanın ölçme aracında da kullanılan dokuz gözenekli bir yapılandırılmış grid örneği yer almaktadır.

1 su-kolonya	2 demir tozu-un	3 mürekkep-su
4 hidrojen-oksijen	5 cıva-zeytinyağı	6 karbon-oksijen
7 su-talaş	8 demir-oksijen	9 çakıl-kum

Şekil 1. Yapılandırılmış Grid Örneği

Bu yapılandırılmış gride ilişkin maddeler ve anahtar yanıtlar şöyledir;

Madde 1) Evde annenizin makarnayı sudan ayırmak için delikli süzgeç kullandığını görürsünüz. Tablodaki madde çiftlerinden hangilerinin oluşturduğu karışımı ayırmak için aynı yöntem kullanılır? Anahtar yanıt: 7, 9

Madde 2) Tablodaki madde çiftlerinden hangisinin/hangilerinin oluşturduğu bileşiği bileşenlerine ayırmak için elektroliz yönteminden yararlanır? Anahtar yanıt: 4

Bahar ve diğerleri (2002)'ne göre yapılandırılmış gridlerin puanlanma aşamaları şöyledir; Yapılandırılmış gridlerin puanlanmasında doğru gözeneklerin bulunması ve mantıksal olarak sıralanması adımları için ayrı sistemler kullanılır. Her sorunun cevabı için uygun gözeneklerin bulunması aşamasında aşağıdaki formül kullanılır.

$$\text{Maddeye ait puan} = \frac{C1}{C2} - \frac{C3}{C4}$$

C1: Doğru seçilmiş gözenek sayısı

C2: Toplam doğru gözenek sayısı

C3: Yanlış seçilmiş gözenek sayısı

C4: Toplam yanlış gözenek sayısı

Bu formül kullanılarak elde edilen madde puanları -1 ve +1 arasında değişir. Her maddeye ait puanı 10 üzerinden değerlendirmek için madde puanı önce 1 ile toplanır, sonra 5 ile çarpılır. Puanın 1 ile toplanmasındaki amaç negatif puanları ortadan kaldırmaktır. Gözenek numaralarının mantıksal veya işlevsel sıraya dizilmesi aşamasının değerlendirilmesinde aşağıdaki puanlama biçimi kullanılır;

Örneğin bir sorunun yanıtını oluşturan gözeneklerdeki olayların gerçekleşme sırası 7, 1 ve 5 şeklinde olsun. Öğretmen, öğrencinin vermiş olduğu yanıtları puanlarken kendine şu soruları sorar;

Soru 1: Öğrencinin yanıtında 7 numara 1'den önce mi geliyor? Cevap "evet" ise, art arda mı geliyor?

Soru 2: Öğrencinin yanıtında 1 numara 5'ten önce mi geliyor? Cevap "evet" ise, art arda mı geliyor?

Bu iki sorudan toplam dört yanıt alınacaktır. Bir sorudan öğrenci en fazla 10 puan alabileceğine göre bu soruda her "evet" cevabı 10/4 yani 2,5 puan olacaktır. Her "hayır" cevabı için öğrenci 0 puan alır.

Yapılandırılmış gridlerde bir maddenin yanıtı birden fazla seçenekten oluşabilir. Ayrıca bir seçenek birkaç maddenin ortak yanıtı durumunda olabilir. Danili ve Reid (2005), çoktan seçmeli testlerde öğrenciler kimi zaman seçenekleri eleyerek ikiye indirip bu iki seçenek arasında karar verdiklerini, gridlerde ise cevabın kaç gözenekten oluştuğu bilinmediği için öğrenciler her gözeneğe "sorunun cevabı olabilir mi?" diye baktığını belirtmiştir. Böylece öğrenciler gridleri yanıtlarken çoktan seçmeli testlere göre daha çok muhakeme yaparlar. Remmers, Gage ve Rummel (1965) seçenek sayısı arttıkça, doğru yanıt tahminle bulunma olasılığının azalacağını belirtmiştir. Yapılandırılmış gridlerdeki seçenek sayıları çoktan seçmeli testlerden daha fazladır. Bu yüzden yapılandırılmış gridlerde doğru yanıt şans yoluyla bulma olasılığının daha düşük olması beklenir.

Ayrıca yapılandırılmış grid maddelerinin cevabının kaç gözenekten oluştuğunu bilmeyen öğrencinin eleme yoluyla doğru yanıtı ulaşması mümkün değildir. Eğitimciler yapılandırılmış grid yönteminin, kısmi bilgiyi ölçmesi, öğrencinin bilgiyi nasıl yapılandığı konusunda öğretmene geri bildirim verme konusunda çoktan seçmeli maddelerden daha güçlü olduğunu savunmuşlar ve bu konuda çeşitli araştırmalar yapmışlardır (Thomas, Ökten ve Buis, 2002; Danili ve Reid, 2005).

Çoktan seçmeli maddelerin objektif ve kolay puanlanıyor olmasının yanında kısmi bilgiyi ölçmede yetersiz kalma, şans başarısına açık olma gibi sınırlılıkları, eğitimcileri bu sınırlılıkları hafifletecek, çoktan seçmeli soruların yanı sıra kullanılabilir farklı madde türleri geliştirme ve deneme çalışmalarına yöneltmiştir. Bu kapsamda yapılandırılmış grid maddeleri ile çoktan seçmeli maddeleri görgül olarak karşılaştıran sınırlı sayıda araştırma bulunması sebebiyle yapılandırılmış grid ile çoktan seçmeli testlerin madde ve test istatistikleri bakımından farklı olup olmadığının araştırılmasına ihtiyaç duyulmuştur.

Araştırmanın Amacı

Bu araştırmada aynı davranışları ve konu kapsamını ölçmek üzere geliştirilen biri çoktan seçmeli maddelerden, diğeri yapılandırılmış grid maddelerinden oluşan iki testin psikometrik özellikleri açısından karşılaştırılması amaçlanmıştır. Bu kapsamda aşağıdaki sorulara yanıt aranmıştır:

İlköğretim 7. Sınıf Fen ve Teknoloji dersi için aynı davranışları ölçmek üzere geliştirilen

- 1) Çoktan seçmeli ve yapılandırılmış grid maddeleri arasında ortalama güçlük indeksleri bakımından manidar düzeyde bir fark var mıdır?
- 2) Çoktan seçmeli ve yapılandırılmış grid maddelerinin ayrııcılık indeksleri arasında manidar düzeyde bir ilişki var mıdır?
- 3) Çoktan seçmeli ve yapılandırılmış grid maddelerinden oluşan iki testten elde edilen puanların dersin aynı konusuna ilişkin öğretmen yapımı yazılı sınav puanlarıyla gösterdikleri korelasyonlar arasında manidar düzeyde bir fark var mıdır?
- 4) Çoktan seçmeli ve yapılandırılmış grid maddelerinden oluşan iki testin iç tutarlık anlamındaki güvenilirlik kestirimleri arasında manidar düzeyde bir fark var mıdır?

YÖNTEM

Araştırmanın Modeli

Bu araştırma farklı iki ölçme yönteminin karşılaştırılmasına dayalı tarama modelinde bir araştırmadır. Bu araştırmada var olan duruma herhangi bir müdahalede bulunulmadan, araştırmaya konu olan durum var olduğu biçimiyle tanımlanmaya çalışılmıştır. Karasar (2004)'a göre araştırmaya konu olan olay ya da nesneyi kendi koşulları içinde, değiştirmeden tanımlamayı amaçlayan araştırmalar tarama modelindedir.

Çalışma Grubu

Çalışma grubunu, araştırmacının ulaşım kolaylığı sebebi ile Ankara ili Etimesgut ilçesinin üç ilköğretim okulunun 7. sınıflarında okuyan 276 öğrenci oluşturmaktadır. Yapılandırılmış grid tekniğinin öğrenciler tarafından anlaşılması güç olduğundan ilköğretimin üst sınıflarından olan 7. Sınıf öğrencileri araştırma grubu olarak seçilmiştir. İlköğretim 8.sınıf öğrencileriyle çalışılmamasının sebebi bu öğrencilerin liselere giriş amaçlı sınavlara hazırlanıyor olmalarıdır. Araştırmaya dahil edilen öğrencilerin çoktan seçmeli ve yapılandırılmış grid olarak hazırlanmış testlerden hangisini alacakları seçkisiz olarak belirlenmiştir. Bu atama işleminde her iki testi alan kız ve erkek öğrenci sayılarının denk olmasına özen gösterilmiştir. Tablo 1'de araştırma grubundaki öğrencilerin aldıkları test türüne ve cinsiyete göre dağılımları verilmiştir.

Tablo 1. Araştırma grubunda yer alan öğrencilerin aldıkları test türüne ve cinsiyete göre dağılımı

Uygulanan Test Türü	Kız		Erkek		Toplam	
	(n)	%	(n)	%	(n)	%
Yapılandırılmış grid	65	24	71	26	136	49
Çoktan seçmeli	68	25	72	26	140	51
Toplam	133	48	143	52	276	100

Tablo 1'e göre yapılandırılmış grid testini alan gruptaki kız (%24) ve erkek (%26) öğrencilerin dağılımı, çoktan seçmeli testi alan gruptaki kız (%25) ve erkek (%26) öğrencilerin dağılımıyla benzerlik göstermektedir.

Veri Toplama Araçları

Araştırmada, ilköğretim 7. sınıf Fen ve Teknoloji dersi “Maddenin İç Yapısına Yolculuk” teması kapsamında iki test hazırlanmıştır. Testlerden biri çoktan seçmeli maddelerden diğeri sadece yapılandırılmış grid maddelerinden oluşmaktadır. Testleri hazırlamak için ilgili dersin öğretim programı, ders ve öğrenci kitapları incelenmiştir. Daha sonra her iki testin geliştirilmesine dayanak olan belirtke tablosu hazırlanarak ilgili konudavranış ilişkisini temsil eden 35 çoktan seçmeli madde ve 35 yapılandırılmış grid maddesi yazılmıştır. Araştırma grubunun sınıf düzeyine uygun olarak çoktan seçmeli maddeler dört seçenekli, yapılandırılmış grid maddeleri dokuz gözenekli olarak hazırlanmıştır. Hazırlanan test maddeleri ve belirtke tablosu, fen ve teknoloji dersi öğretmenlerine gösterilerek madde-davranış uyumu konusunda görüş ve önerileri alınmıştır. Ayrıca öğretmenler, soruların dil ve anlatım özellikleri bakımından öğrenci seviyesine uygunluğunu da değerlendirmişlerdir. Bunun yanında test maddeleri, Fen Eğitimi ile Ölçme ve Değerlendirme alanlarında çalışan üç uzman tarafından, ölçme tekniğine, dil ve anlatım özelliklerine, ünitenin davranışlarına uygunlukları ve bilimsel doğrulukları bakımından denetlenmiştir. Alınan görüşler doğrultusunda bazı maddeler testten çıkarılmış, bazı maddeler önerilere göre düzeltilmiştir.

Bu düzeltmelerden sonra ortaya çıkan testin deneme formu, “Maddenin İç Yapısına Yolculuk” ünitesinin işlenmesinden sonra, araştırma grubunda yer alan ilköğretim okullarındaki öğrencilere benzer bir ilköğretim okulundaki 96 öğrenci üzerinde uygulanmıştır. Deneme uygulaması sonucunda madde ayırıcılık indeksi 0,30'un üstünde olan maddelerden ölçülmek istenen özelliği en iyi temsil eden, araştırmanın amacına en uygun olduğu düşünülen 16 çoktan seçmeli ve bunlara paralel 16 yapılandırılmış grid maddesi asıl form için seçilmiştir. Her iki testteki maddelerden 3'ü

bilgi, 9'u kavrama ve 4'ü uygulama düzeyindeki davranışları yoklar niteliktedir. İki test için de her bir maddeye verilen doğru yanıt 10 puan üzerinden puanlanmıştır. Buna göre her iki testten de alınabilecek en yüksek puan, 160'dır.

İşlem

Uygulamanın başında öğrencilere Fen ve Teknoloji dersi “Maddenin İç Yapısına Yolculuk” ünitesi ile ilgili konu değerlendirme testi alacakları söylenmiştir. Bu uygulama, her sınıfta öğretmenin “Maddenin İç Yapısına Yolculuk” ünitesi ile ilgili öğretim ve yazılı yoklama çalışmaları tamamlandıktan sonra yapılmıştır. Araştırma grubunda yer alan öğrencilerin seçkisiz olarak belirlenen yarısı yapılandırılmış grid testini, diğer yarısı da çoktan seçmeli testi almıştır.

Her sınıfta, uygulama öncesinde yapılandırılmış grid tekniği tanıtılmış ve öğrencilerle birlikte örnekler çözülmüştür. Uygulama esnasında sınıfın Fen ve Teknoloji dersini yürüten öğretmen de derslikte bulunmuştur. Öğretmenlerinin de onayı alınarak, öğrencilere; test sonuçlarının dönem sonu notuna etki edeceği söylenmiş, böylece testlerin ciddiyetle yanıtlanması sağlanmıştır. Bazı öğrencilerin alacakları testteki soruların, örneği çözülen yapılandırılmış grid türünde sorular olacağı belirtilmiştir. Bazılarının da çoktan seçmeli türde soruların yer aldığı bir test alacakları duyurulmuştur. Puanlamada yanlış yanıtlarının doğru yanıtlarını götüreceği belirtilerek, öğrenciler emin olmadıkları soruları yanıtlamamaları konusunda uyarılmışlardır. Öğrenciler testi yaklaşık bir ders saati içinde tamamlamışlardır. Bunun dışında araştırma amaçları bakımından bir geçerlik ölçütü olarak kullanılmak üzere, öğrencilerin söz konusu üniteye ilişkin yazılı sınav sonuçları dersin öğretmeninden alınmıştır.

Öğrencilerin yapılandırılmış grid maddelerine verdikleri yanıtlar, grid maddeleri puanlama formülüne göre puanlanmıştır. Analizler ham puanlar üzerinden gerçekleştirilmiştir. Yapılandırılmış grid test maddelerinin puanlanmasında düzeltme yapılmaktadır. Bu sebeple çoktan seçmeli test puanlanırken de aşağıdaki düzeltme formülü kullanılmıştır.

$$\text{Düzeltilmiş Puan} = D - \frac{Y}{n - 1}$$

D= doğru yanıt sayısı

Y=yanlış yanıt sayısı

n=her maddedeki seçenek sayısı

Verilerin Analizi

Birinci araştırma sorusuna yanıt bulmak üzere, iki testteki maddelerin madde güçlük indeksleri normal dağılmadığından, madde güçlük indeksleri ortalamaları arasında manidar fark olup olmadığı, “Mann Whitney U” testi kullanılarak analiz edilmiştir.

Bu çalışmada; maddenin ölçülmek istenen niteliğe sahip olanlarla olmayanları ayırma gücünü gösteren madde ayırıcılık indeksi çoktan seçmeli maddeler için, çift serili korelasyon formülü ile hesaplanmıştır. Yapılandırılmış grid maddelerinin puanları 0 ile 10 arasında değişmektedir ve yapılan ölçme işlemi kullanılan ölçek eşit aralıklı ölçektir. Bu yüzden, bu maddelere ait ayırıcılık indeksleri hesaplanırken, madde puanlarıyla test puanı arasındaki korelasyon, Pearson momentler çarpım korelasyon formülü kullanılarak bulunmuştur. İkinci araştırma sorusu kapsamında iki testteki maddelerin ayırıcılık indeksleri arasındaki ilişki Spearman Brown Sıra Farkları korelasyon formülü ile incelenmiştir. Puanlayıcıların verdikleri puanların kendileri değil de büyüklük sırasına sokulmuş değerleri temel alınmışsa ve sürekli veriler normal dağılmıyorsa korelasyon katsayısı olarak “Spearman – Brown Sıra Farkları Korelasyon Katsayısı” hesaplanmalıdır (Şencan, 2005). Bu çalışmada, hesaplanan Spearman Brown Sıra Farkları Korelasyon Katsayısı iki ölçme aracındaki maddelerin ayırıcılık indeksleri arasındaki ilişkinin bir ölçüsü olarak yorumlanmıştır.

Üçüncü araştırma sorusunu yanıtlamak üzere, her iki testten alınan puanlar ile geçerlik ölçütü olarak kabul edilen “Maddenin İç Yapısına Yolculuk” ünitesine ait öğretmenlerin yaptığı yazılı sınav puanları arasındaki ilişkiyi belirlemek üzere, Pearson Momentler Çarpım korelasyon tekniği kullanılmıştır. İki tür testten elde edilen puanların, ünite başarı notları ile gösterdiği korelasyonlar arasında manidar bir fark olup olmadığını bulmak için bu korelasyon katsayılarının z değerlerine dönüşümü yapılmış ve iki korelasyon katsayısı arasında manidar fark olup olmadığı Fisher’in Z’si ile test edilmiştir. Howell (1987)’a göre ilişkisiz iki korelasyon arasındaki farkın manidarlığı Fisher’in Z’si ile test edilir.

Testlerden elde edilen ölçümlerin iç tutarlılık anlamındaki güvenilirlik düzeylerine ilişkin kanıtlar Cronbach (1984)’ın alfa formülü kullanılarak elde edilmeye çalışılmıştır. Cronbach alfa hem 1-0 biçiminde puanlanan çoktan seçmeli maddelerden; hem de geniş puan aralıklarında ve farklı puan ağırlıklarında puanlanan açık uçlu maddelerden oluşmuş testlerin iç tutarlılık anlamındaki güvenilirlik kestirimini veren bir hesaplama formülüdür. Benzer bir formül olmasına rağmen KR 20, sadece 1-0 biçiminde puanlanan maddelerden oluşmuş ve düzeltme formülü kullanılmayan testlerin iç tutarlılık anlamındaki güvenilirlik kestirimlerinin hesaplanmasında kullanılır (Crocker

ve Algina, 1986). Bu sebeple çalışmada Cronbach α formülü kullanılmıştır. Her ne kadar Cronbach α bir korelasyon katsayısı olmasa da, güvenilirlik katsayılarının karşılaştırılmasına ilişkin bir formül bulunmadığından bu bir korelasyon katsayısı gibi düşünülerek dördüncü araştırma sorusuna yanıt bulmak için iki güvenilirlik katsayısı arasındaki farkın manidarlığı Fisher Z testi ile test edilmiştir. Verilerin analizinde 0,05 manidarlık düzeyi esas alınmıştır.

BULGULAR

Tablo 2’de 7.sınıf Fen ve Teknoloji dersi Maddenin İç Yapısına Yolculuk ünitesinin davranışlarını yoklamak üzere geliştirilmiş olan 16 maddelik yapılandırılmış grid testi ile 16 maddelik çoktan seçmeli testten alınan puanlara ilişkin betimleyici istatistikler yer almaktadır. Her iki testte de bir maddeden alınabilecek en yüksek puan 10, testten alınabilecek en yüksek puan 160’tır. Ayrıca yapılandırılmış grid ve çoktan seçmeli test puanlarının normale yakın dağılım sergilediği görülmüştür. Tablo 2 yapılandırılmış grid ve çoktan seçmeli test puanlarına ilişkin betimsel istatistikleri göstermektedir.

Tablo 2. *Yapılandırılmış grid testi ve çoktan seçmeli test puanlarının betimsel istatistikleri*

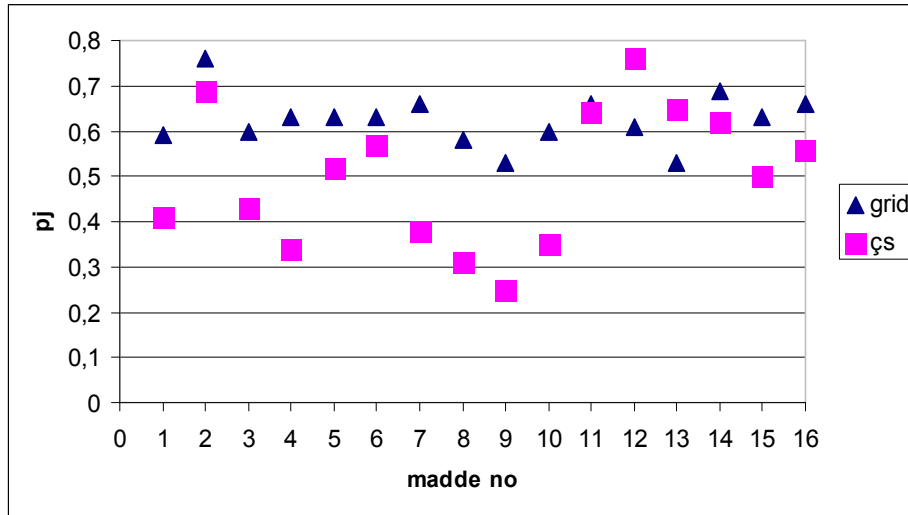
İstatistikler	Yapılandırılmış grid testi	Çoktan seçmeli test
n	136,00	140,00
\bar{X}	100,12	54,05
mod	98,27	26,67
ortanca	96,89	53,33
standart sapma	18,51	43,58
basıklık	-0,43	-0,76
çarpıklık	0,56	0,09
en büyük puan	148,29	46,67
en küçük puan	64,42	-40,00
dizi genişliği	83,87	186,67

Tablo 2’ye göre yapılandırılmış grid testinde ortalama ve ortanca değerleri birbirine oldukça yakındır. Çoktan seçmeli testte de benzer bir durum söz konusudur. Tablo 3’te testlerde yer alan maddelerin güçlük indeksleri yer almaktadır. Ayrıca iki testten alınan puanların dizi genişliğinin birbirinden oldukça farklı olduğu da görülmektedir.

Tablo 3. Yapılandırılmış grid testi ve çoktan seçmeli testteki maddelerin güçlük (P_j) indeksleri

	Yapılandırılmış grid testi	çoktan seçmeli test
en küçük p_j	0,53	0,25
en büyük p_j	0,76	0,76
ortalama p_j	0,62	0,50
medyan p_j	0,63	0,54

Tablo 3'te görüldüğü gibi yapılandırılmış grid testindeki maddelere ait madde güçlük indeksleri 0,53 ve 0,76, çoktan seçmeli test maddelerine ait madde güçlük indeksleri ise 0,25 ve 0,76 arasında değişmektedir. Yapılandırılmış grid ve çoktan seçmeli test maddelerine ait madde güçlük indeksleri dağılımı Şekil 2'de verilmiştir.



Şekil 2. Fen ve Teknoloji Dersi "Maddenin İç Yapısına Yolculuk" Ünitesine ait yapılandırılmış grid maddeleri ile çoktan seçmeli maddelerin madde güçlük indeksleri dağılımı

Yapılandırılmış grid ve çoktan seçmeli testteki maddelerin güçlük indeksleri ortalamaları arasındaki farkı incelemek üzere yapılan "Mann Whitney U" testi sonuçları Tablo 4'te verilmiştir.

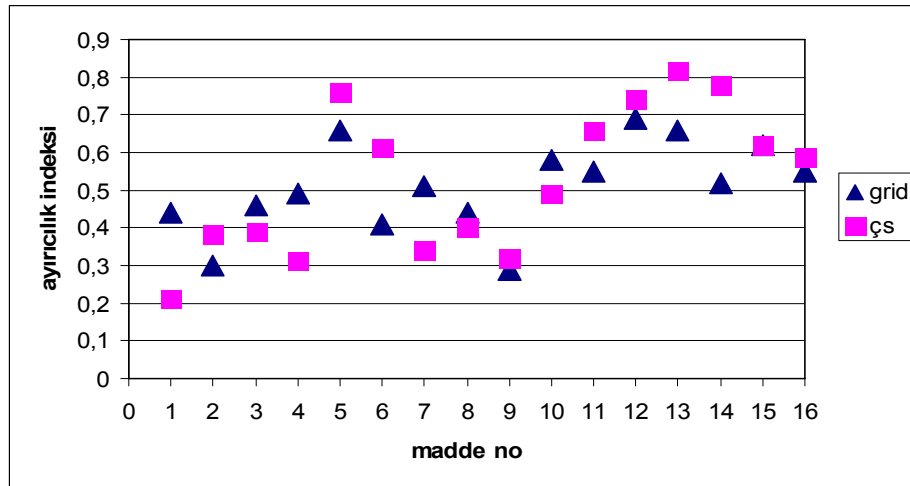
Tablo 4. Yapılandırılmış grid maddeleri ile çoktan seçmeli maddelerin güçlük indeksleri ortalamalarının karşılaştırılmasına ilişkin Mann Whitney U Testi sonuçları

Grup	k	Sıra Ortalaması	Sıra Toplama	U	p
grid	16	20,56	329,00	63,00	0,014
çs	16	12,44	199,00		

k : madde sayısı

Tablo 4'e göre yapılandırılmış grid ve çoktan seçmeli test maddelerine ait madde güçlük indeksleri ortalamaları arasında anlamlı bir fark olduğu bulunmuştur ($U= 63,00$, $p<.05$). Sıra ortalamaları dikkate alındığında yapılandırılmış grid maddelerinin madde güçlük indeksleri ortalamasının, çoktan seçmeli maddelerin madde güçlük indeksleri ortalamasından daha yüksek olduğu anlaşılmaktadır. Bu durum yapılandırılmış grid maddelerinin daha kolay olduğunun göstergesidir.

Yapılandırılmış grid maddelerinin ayırıcılık indeksleri 0,29 ile 0,69 arasında, çoktan seçmeli maddelerin ayırıcılık indeksleri ise 0,21 ile 0,74 arasında değişmektedir. Şekil 3'te çoktan seçmeli maddelerin çift serili korelasyon formülü ile, yapılandırılmış grid maddelerinin pearson momentler çarpım korelasyon formülü ile hesaplanan ayırıcılık indekslerinin dağılımı görülmektedir.



Şekil 3. Fen ve Teknoloji Dersi "Maddenin İç Yapısına Yolculuk" Ünitesine ait yapılandırılmış grid maddeleri ile çoktan seçmeli maddelerin madde ayırıcılık indeksleri dağılımı

Şekil 3'te görüldüğü gibi iki test türüne ait maddeler ayrıricılık değerleri bakımından benzer dağılım göstermektedir. Hesaplanan Spearman Brown sıra farkları korelasyon katsayısı da ($r=0,715$) ayrıricılık indekslerinin büyüklük sıralarının büyük ölçüde tutarlı olduğunu göstermektedir. Ayrıricılık indeksinin 0,40 ve daha büyük değerler alması, maddenin ayırt etme gücünün çok iyi olduğu anlamına gelir (Ebel, 1965). Buna göre çoktan seçmeli maddelerin 10 tanesi, yapılandırılmış grid maddelerinin ise 14 tanesi ayırt etme gücü yüksek maddelerdir.

Yapılandırılmış grid ve çoktan seçmeli maddelerden oluşan testlerden elde edilen puanların ölçüt ölçümleri ile ilişkisini belirlemeye yönelik hesaplanan korelasyon katsayılarının karşılaştırılmasına ilişkin Z testi sonuçları Tablo 5'te verilmiştir.

Tablo 5. *Yapılandırılmış grid ve çoktan seçmeli test puanları ile Fen ve Teknoloji dersinin ilk ünitesine ait başarı puanları arasındaki korelasyonların karşılaştırılmasına ilişkin Z testi sonuçları*

Test Türü	n	r	Zr	Z
Yapılandırılmış grid	136	0,691	0,848	0,33
Çoktan seçmeli	140	0,670	0,811	

Tablo 5'te görüldüğü gibi yapılandırılmış grid testinden alınan puanlarla öğrencilerin Fen ve Teknoloji dersinin ilk ünitesine ait başarı puanları arasında 0,70 e yakın ($r=0,691$, $p<.05$) yüksek sayılabilecek düzeyde, pozitif ve manidar düzeyde bir ilişki olduğu bulunmuştur. Benzer şekilde, çoktan seçmeli test puanlarıyla öğrencilerin Fen ve Teknoloji dersinin ilk ünitesine ait başarı puanları arasında yüksek sayılabilecek düzeyde, pozitif, manidar düzeyde bir ilişki olduğu bulunmuştur ($r=0,670$, $p<.05$). Bulunan korelasyon katsayılarına Z dönüşümü yapılarak, iki korelasyon değeri Fisher'in Z'si ile karşılaştırılmıştır. Z testi sonuçlarına göre iki farklı test türünden alınan puanların Fen ve Teknoloji dersinin ilk ünitesine ilişkin başarı puanları ile korelasyonları arasında anlamlı bir fark bulunmamıştır. Bu durum, öğrencilerin Fen ve Teknoloji dersinin ilk ünitesine ait başarı puanları geçerlik ölçütü olarak kabul edildiğinde, iki testin benzer düzeyde geçerli ölçümler sağlayabildiğine işaret etmektedir.

Yapılandırılmış grid ve çoktan seçmeli testten elde edilen ölçümlerin güvenilirliklerine ilişkin karşılaştırma Tablo 6'da yer almaktadır.

Tablo 6. Fen ve Teknoloji Dersi “Maddenin İç Yapısına Yolculuk” Ünitesine ait yapılandırılmış grid testi ile çoktan seçmeli testin iç tutarlılık anlamındaki güvenilirliklerinin karşılaştırılmasına ilişkin Z testi sonuçları

Test Türü	N	α	Zr	Z
Yapılandırılmış grid	136	0,82	1,16	2,25
Çoktan seçmeli	140	0,71	0,89	

Tablo 6’ya göre yapılandırılmış grid testi ile elde edilen ölçümlerin iç tutarlılık anlamındaki güvenilirlik katsayısı 0,82; çoktan seçmeli test ile elde edilen ölçümlerin iç tutarlılık anlamındaki güvenilirlik katsayısı ise 0,71 bulunmuştur. Bu iki güvenilirlik katsayısı arasında manidar düzeyde bir fark olduğu bulunmuştur. Sonuç olarak aynı davranışları yoklamak üzere geliştirilmiş yapılandırılmış grid testinin iç tutarlılık anlamında, çoktan seçmeli testten daha güvenilir ölçümler sağladığı görülmüştür.

TARTIŞMA VE SONUÇ

Bu çalışmada aynı konuda aynı davranışları ölçmek üzere geliştirilmiş, biri yapılandırılmış grid diğeri çoktan seçmeli maddeler içeren iki testten elde edilen sonuçlar karşılaştırılmıştır. Araştırmadan elde edilen bulgular, kullanılan testlerdeki yapılandırılmış grid maddelerinin çoktan seçmeli maddelere göre daha kolay olduğunu göstermektedir. Araştırmada, yapılandırılmış grid testinden elde edilen ölçümlerin, çoktan seçmeli testten elde edilen ölçümlerden daha güvenilir olduğu, ancak iki testten elde edilen ölçümlerin ölçüt geçerlikleri arasında anlamlı fark olmadığı belirlenmiştir. Literatürde yapılandırılmış grid tekniğini konu alan görgül araştırmaların sınırlı olduğu görülmüştür.

Thomas, Ökten ve Buis (2002), yapılandırılmış grid yaklaşımını inceledikleri araştırmada Matematik öğretmenliğinde okuyan 30 öğrenciye grid maddeleri ile açık uçlu maddelerden oluşan bir kalem kağıt testi vermişler ve öğrencilerin bu iki madde türüne ilişkin performanslarını karşılaştırmışlardır. Sonuç olarak öğrencilerin grid maddelerinde gösterdikleri performans ile açık uçlu maddelerdeki performanslarının benzer olduğu ($r = 0.82$) görülmüştür.

Danili ve Reid (2005), kısa cevaplı, yapılandırılmış grid ve çoktan seçmeli madde formatlarını karşılaştırdıkları araştırmalarında aynı içerikte ancak farklı madde formatlarını içeren 3 test hazırlamışlardır. Bu araştırmanın sonuçlarına göre bireyler en yüksek performansı çoktan seçmeli testte göstermişlerdir. Test puanları arasındaki korelasyonlara bakıldığında

ise en yüksek korelasyonun objektif testler olan çoktan seçmeli test ve yapılandırılmış grid testi arasında olduğu görülmüştür.

Yukarıdaki çalışmalarda yapılandırılmış grid ve çoktan seçmeli testi alan grup aynı olduğundan korelasyon hesabı yapılabilmektedir. Ancak bu çalışmada iki testi iki ayrı grup almış olduğundan korelasyon hesaplanamamıştır.

Yapılandırılmış grid ve çoktan seçmeli testleri psikometrik özellikler açısından karşılaştıran kapsamlı çalışmaların sınırlı olması sebebiyle yapılandırılmış grid maddelerinin çok seçenekli ve birden fazla anahtar yanıtı olma özellikleri gözetilerek bu türe benzer çoktan seçmeli maddeleri konu alan araştırmalar incelenmiştir. Çoktan seçmeli maddelerde seçenek ve doğru yanıt sayısının madde ve test istatistikleriyle ilişkisi üzerine yapılan çalışmalar; çoktan seçmeli maddelerde seçenek sayısını arttırmanın madde gücünü, testin geçerliğini ve güvenilirliğini arttırdığını, birden fazla doğru yanıtı olan maddelerden oluşan bir testin ise tek doğru yanıtı maddelerden oluşan bir teste göre daha güvenilir sonuçlar verdiğini göstermektedir. Çalışmalarda bu durum, seçenek sayısının ve doğru yanıt sayısının artmasının şans başarısını düşürdüğü biçiminde yorumlanmıştır (Turgut, 1971; Bayır, 1998; Araz, 2001). Bu çalışmada da yapılandırılmış grid testinden elde edilen ölçümlerin, çoktan seçmeli testten elde edilen ölçümlerden daha güvenilir olduğu bulunmuştur. Bu durum da yapılandırılmış grid maddelerinin çok seçenekli ve birden fazla anahtar yanıtı olma özelliği sebebiyle şans başarısında düşüş sağladığı bu sebeple daha güvenilir ölçümler sağladığı biçiminde yorumlanabilir.

Bu çalışmada okul öğrenmelerinin ölçülmesi konu alınmıştır. Okul öğrenmelerinin ölçülmesinde test puanları üzerinde düzeltme formülü kullanılmadığı bilinmekle birlikte bu çalışmada, yapılandırılmış grid maddelerinin madde puanını veren formülü kendi içinde bir düzeltme formülü içerdiğinden çoktan seçmeli maddeler için de puanlar üzerindeki işlemler bakımından denkliği sağlamak üzere, düzeltme formülü kullanılmıştır. Yapılandırılmış grid maddelerini puanlamak için kullanılan formüle göre, negatif madde puanını ortadan kaldırmak amacıyla, sıralama gerektiren maddelerde madde puanına 2,5; sıralama gerektirmeyen maddelerde ise madde puanına 5 puan eklenir. Madde puanlarına yapılan bu eklemeler, her öğrencinin sıralama gerektiren maddelerden fazladan 2,5, sıralama gerektirmeyen maddelerden ise fazladan 5 puan almasına sebep olmaktadır. Böylece maddeyi cevaplamayan öğrenci de o maddeden puan alır. Bu durum madde istatistiklerine yansımış, yapılandırılmış grid maddelerinin güçlük değerlerini arttırarak, bu maddelerin çoktan seçmeli maddelerden daha kolay görünmesine yol açmıştır. Benzer biçimde yapılandırılmış grid madde puanı formülü gereği negatif olamaz ve öğrenci,

maddeyi yanıtsız bıraksa dahi bu maddeden puan alır. Çoktan seçmeli testlerde ise düzeltme formülü kullanıldığı takdirde, $1/a$ (seçenek sayısı) ölçüsünde yanlış yanıtlar, doğru yanıtlardan çıkarılmaktadır. Bu çalışmada dört seçenekli sorulardan oluşan çoktan seçmeli testte öğrencinin test puanından 3,33 puan eksiltmekte yanıtsız maddeler puan almamaktadır. Bu durum aynı grubun yapılandırılmış grid testinden alınan puan ortalamalarının artmasına sebep olmuştur. Oysa yapılandırılmış grid maddelerinde seçenek sayısının çoktan seçmeli maddelerdekinden fazla olması sebebiyle, yapılandırılmış gridlerdeki madde puanlarının ve dolayısıyla test puanının azalması beklenmektedir. Yapılan çalışmalar çoktan seçmeli maddelerde seçenek sayısının artması durumunda madde puanlarının ve madde güçlük indekslerinin azaldığını göstermektedir (Yaralıoğlu, 1987 ; Bayır 1998; Araz, 2001).

Araştırmada kullanılan yapılandırılmış grid testinde sıralama gerektiren iki madde yer almaktadır, bunlar 12. ve 13. maddelerdir. Şekil 2 incelendiğinde, yapılandırılmış grid testindeki maddelerden yalnızca 12. ve 13. maddelerin güçlük değerlerinin çoktan seçmeli testteki paralel maddelerinkinden düşük olduğu, yapılandırılmış grid testinin diğer maddelerine ilişkin güçlük değerlerinin ise çoktan seçmeli test maddelerinin güçlük değerlerinden yüksek olduğu görülmektedir. Aynı şekilde bu durumun da, yapılandırılmış grid maddelerini puanlama formülüne göre, sıralama gerektiren maddelerde madde puanına fazladan 2,5, sıralama gerektirmeyen maddelere ise fazladan 5 puan eklenmesinden kaynaklandığı sonucuna ulaşılmıştır.

Testlere ilişkin betimsel istatistiklere bakıldığında yapılandırılmış grid testinden alınan puanların dizi genişliğinin 83,87, çoktan seçmeli testten alınan puanların dizi genişliğinin ise 186,67 olduğu görülmektedir. Bu çalışmada, çoktan seçmeli bir maddenin doğru ve/veya yanlış yanıtlanması sonucunda test puanına katkısı -3,33 ile 10 puan arasında değişirken yapılandırılmış grid maddesinin test puanına katkısı 0 ile 10 puan arasında değişmiştir. Böylece yapılandırılmış grid maddelerinden alınan puanların dolayısıyla da test puanlarının ranjı daralmış, yapılandırılmış grid testinin standart sapması düşük çıkmıştır. Öğrencilerin hangi testi alacakları seçkisiz olarak belirlendiğinden çoktan seçmeli testi alan öğrenciler ile yapılandırılmış grid testini alan öğrencilerin ölçülen davranışlar açısından başarı dağılımlarının denk olması beklenmekteyken, iki test türünden elde edilen puanların varyanslarının homojen olmadığı görülmüştür. Bu durumun, iki test türüne ilişkin puanlama yöntemlerinin birbirinden farklı olmasından kaynaklandığı yorumu yapılabilir. Yapılandırılmış grid testini alan grubun testten aldıkları puanlar açısından nispeten homojen olması maddelerin ayıricılık değerlerinin daha düşük çıkmasına da yol açmıştır.

Genel olarak değerlendirildiğinde, yapılandırılmış grid testi ile çoktan seçmeli testin benzer psikometrik özelliklere sahip olduğu görülmektedir. Bununla birlikte, yapılandırılmış grid tekniğinin kısmi bilgiyi yoklama ve doğru yanıtı şansa bulma olasılığını azaltma özellikleri dikkate alındığında, bu türün çoktan seçmeli testlerle birlikte kullanılması önerilebilir. Bunun dışında grid türü maddeler kullanılarak, kavrama; yorumlama, ilişki kurma, çözümlenme gibi üst düzey zihinsel özelliklerin ölçülmesi mümkün olabilir. Sonuç olarak grid türü maddeler çoktan seçmeli maddelerin yanı sıra geçerli ve güvenilir ölçme yapacak bir madde türü olarak okul öğrenmelerinin ölçülmesinde kullanılabilir. Yapılandırılmış grid maddelerinin hazırlanması ve puanlanması zahmetli olduğundan bu tür maddelerin hazırlanmasında ve puanlanmasında yardımcı olacak bilgisayar temelli programların geliştirilmesi önerilebilir. Aydın(2004)'ın çalışmasında da yapılandırılmış grid tekniğinin değerlendirme sürecinin öğretmenler tarafından zahmetli, maliyetli ve zaman alıcı olarak değerlendirildiği sonucuna ulaşılmıştır.

Bu araştırmada sadece 9 gözenekli yapılandırılmış grid maddeleri kullanılmıştır, gözenek sayısı değiştirilerek araştırma tekrarlanabilir. Ayrıca çoktan seçmeli ve grid türü maddelerin bir arada kullanıldığı testler ile sadece bir türün kullanıldığı testlerin psikometrik özellikleri karşılaştırılabilir. Yapılandırılmış grid maddelerine ilişkin madde puanını hesaplama formülünün öğrencilere ek puan vererek sabit hataya neden olduğu sonucuna ulaşılmıştır. Bu formülün yeniden düzenlenerek tekniğin kullanılması önerilir.

KAYNAKLAR

- Araz, G. (2001). *Aynı Davranışı Ölçmeye Yönelik Kısa Cevaplı, Üç ve Beş Seçenekli Çoktan Seçmeli Testlerin Madde ve Test Özelliklerinin Şans Başarısı ile Birlikte İncelenmesi*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Aydın, F. (2004). *Ölçme Değerlendirme Tekniği Olarak Yapılandırılmış İletişim Gridi ve Bilgisayar Ortamında Uygulanabilirliğine İlişkin Görüşler*. Yayımlanmamış yüksek lisans tezi, Abant İzzet Baysal Üniversitesi Sosyal Bilimler Enstitüsü, Bolu.
- Bahar, M., Öztürk, E. ve Ateş, S. (16-18 Eylül 2002). Yapılandırılmış Grid Metodu İle Lise Öğrencilerinin Newton'un Hareket Yasası, İş, Güç ve Enerji Konusundaki Anlama Düzeyleri ve Hatalı Kavramların Tespiti. V. Ulusal Fen Bilimleri ve Matematik Kongresinde sunulan bildiri, Ankara.
- Bayır, D. (1998). *Sayısal Yetenek Testlerinde Seçenek Sayısının Madde ve Test İstatistikleri Üzerindeki Etkisinin Şans Başarısı ile Birlikte İncelenmesi*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Cronbach, L. (1984). *Essentials of Psychological Testing*. New York: Harper & Row.
- Çelen, Ü. ve Çıkrıkçı-Demirtaşlı, N. (2006). Düzeltme Yönergesinin Testin Psikometrik Özelliklerine Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. Sayı: 30, 82-91.
- Danili, E. & Reid, N. (2005). Assessment Formats: Do They Make a Difference?. *Chemistry Education Research and Practice*, 6(4), 204-212.
- Ebel, R. (1965). *Measuring Educational Achievement*. New Jersey: Prentice Hall.
- Howell, D. C. (1987). *Statistical Methods for Psychology* (second edition). Bostun: Duxbury Press.
- Johnstone, A. H , Bahar, M. & Hansell, M. H. (2000). Structural Communication Grids: A Valuable Assessment and Diagnostic Tool for Science Teachers, *Journal of Biological Education*, 34(2), 87-89.
- Karasar, N. (2004). *Bilimsel Araştırma Yöntemleri*. (13. Baskı). Ankara: Nobel Yayın Dağıtım
- Remmers, H.H., Gage, N.L, Rummel, J.F. (1965). *A Practical Introduction to Measurement and Evaluation* (second edition). New York: Harper & Row Publishers.
- Şencan, H. (2005). *Sosyal ve davranışsal Ölçümlerde Güvenirlik ve Geçerlik* (birinci baskı). Ankara: Seçkin Yayınevi.

- Thomas, D., Ökten, G., and Buis, P. (7-12 July,2002). On-line Assessment of Higher-order Thinking Skills: A Java Based Extension to Closed Form Testing. Refereed Proceedings of the Sixth International Conference on Teaching Statistics, Durban, South Africa.
- Turgut, F. (1992). *Eğitimde Ölçme ve Değerlendirme Metotları* (9. Baskı). Ankara: Saydam Matbaacılık.
- Umay, A. (1998). Seçmeli Testlerde Yanıtlayıcı Davranışları ve Şans Başarısının Elimine Edilmesi İşlemlerine İlişkin Bazı Öneriler. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. 14, 54-61.
- Yaralıoğlu, O. (1987). *Üç ve Beş Seçenekli Maddelerden Oluşan Testlerin Psikometrik Özelliklerinin Karşılaştırmalı Olarak İncelenmesi*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

