



KAVRAM HARİTASI VE YAPILANDIRILMIŞ GRİDLE ELDE EDİLEN PUANLARIN GEÇERLİK VE GÜVENİRLİKLERİNİN İNCELENMESİ*

AN ANALYSIS ON THE VALIDITY AND RELIABILITY OF CONCEPT MAP AND STRUCTURAL COMMUNICATION GRID SCORES

Melek Gülşah EROĞLU**, Hülya KELECİOĞLU***

ÖZET: Bu çalışmada, “Newton’un Hareket Yasaları” ünitesinde öğrencilerin başarılarını ölçmek için kullanılan kavram haritası ve yapılandırılmış gridin geçerlik ve güvenilirlikleri araştırılmıştır. Çalışma grubunu, 2009-2010 öğretim yılında Fen Bilgisi Öğretmenliği ve Fizik Öğretmenliği programı 1. sınıfta okuyan 102 öğrenci oluşturmaktadır.

Kavram haritası ve yapılandırılmış gridin geçerliği için ölçüt olarak aynı konuda geliştirilen kısa cevaplı bir test kullanılmıştır. Kavram haritasından ve kısa cevaplı testten elde edilen puanlar arasındaki Pearson korelasyon katsayısı 0.57 bulunmuştur ($p < 0.05$). Kavram haritasının iç tutarlık anlamında güvenilirlik katsayısı olan Cronbach alfa katsayısı 0.69’dur. Yapılandırılmış gridden elde edilen puanlar ile kısa cevaplı testten elde edilen puanlar arasındaki ilişki 0.69’dur ($p < 0.05$). Yapılandırılmış gridin güvenilirliği için Cronbach alfa katsayısı hesaplanmış ve 0.77 bulunmuştur. Kavram haritası ve yapılandırılmış grid arasında pozitif yönde, orta düzeyde ve anlamlı bir ilişki olduğu görülmüştür (0.51; $p < 0.05$).

Anahtar sözcükler: kavram haritası, yapılandırılmış grid, tamamlayıcı ölçme araçları, geçerlik, güvenilirlik

ABSTRACT: This study aims to research on the validity and reliability of concept maps and structured communication grids used to assess students’ achievement regarding the unit called “Newton’s Laws of Motion”. The sampling consisted of 102 students studying in their Year One at Science Teaching and Physics teaching departments during 2009 – 2010 academic year. The responses to the short-answer test consisted of the main measure in checking the validity of the concept map and the structured communication grid. The Pearson Correlation Coefficient between the scores of the short-answer test and the concept map was calculated to be 0.57. In order to find the reliability state of the concept map, the calculation concluded as 0.69. The degree of the relationship between structured communication grid and short answer test scores was calculated (0.69; $p < 0.05$). For the reliability of the structured communication grid, Cronbach Alpha was calculated 0.77. A positive, medium level and significant relationship was observed between the structured communication grid and the concept map (0.51; $p < 0.05$).

Keywords: concept map, structural communication grid, alternative assessment tools, validity, reliability

1. GİRİŞ

Son yıllarda geliştirilen öğretim programlarının şekillenmesinde büyük bir rol oynayan yapılandırmacı yaklaşım, eğitim-öğretim faaliyetlerinin en önemli basamaklarından biri olan ölçme ve değerlendirme aşamasını da önemli ölçüde etkilemiştir. Yapılandırmacı yaklaşımı temel alan eğitim-öğretim uygulamalarında, değerlendirme süreci ağırlıklıdır. Ürüne değil daha çok süreç değerlendirmelerine ağırlık verilmektedir (Demirel 2007, s. 236). Böylece öğrencilere geleneksel ölçme araçları ve bu ölçme araçlarıyla elde edilen sonuçları, destekleyici ve tamamlayıcı nitelikte olan tamamlayıcı ölçme araçlarının birlikte kullanılmasıyla, çoklu ölçme ve değerlendirme fırsatları sunulmaktadır. Tamamlayıcı ölçme araçlarında; öğrencinin ne kadar bildiğinden çok, neyi bildiği, nerelerde eksikliği olduğu çok daha ön plandadır. Öğrencilerin verdiği yanıtların aslında daha geniş bir anlam taşıdığı ve farklı boyutlarda ele alınması gerektiği vurgulanmaktadır. Değerlendirme sadece öğrencilere not vermek ya da öğrencinin grup içindeki yerini belirlemek değil, öğrencilerin ilerleme aşamalarını ve tamamlamaları gereken eksiklikleri gösteren veya görmelerini sağlayan bir süreç olarak da düşünülmelidir. Dolayısıyla, hem süreç hem de ürün değerlendirmesinden söz edilmektedir. Bu çalışmada; tamamlayıcı ölçme araçlarından kavram haritası ve yapılandırılmış grid ile elde edilen puanların geçerlik ve güvenilirlikleri incelenmiştir.

* Bu çalışma, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalında hazırlanmış olan Yüksek Lisans tezini içermektedir.

** Gazi Üniversitesi, Gazi Eğitim Fakültesi, gulsah_eroglu@yahoo.com

*** Doç. Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, hulyaebb@hacettepe.edu.tr

Kavram haritaları 70’li yılların başında Cornell Üniversitesinde Novak ve arkadaşları tarafından, Ausubel’in “Anlamli Öğrenme” yaklaşımı temel alınarak geliştirilmiştir (Novak & Gowin 1984). Kavram haritaları, kavramların bağlantı cümleleri ile birbirlerine bağlandığı bir gösterim biçimidir. Kullanım amacına bağlı olarak çeşitli kavram haritası oluşturma biçimleri vardır (Edwards & Fraser 1983; Ruiz-Primo, Schultz, & Shavelson 2001). Haritalama bir veya birkaç anahtar kavram kullanılarak yapılabileceği gibi; belli bir sayıda kavram verip bunlar arasındaki ilişkilerin ortaya çıkarılması da sağlanabilir. Bu haritalama yöntemi, “sıfırdan kavram haritalama” olarak isimlendirilmektedir. “Hiyerarşik kavram haritaları” ve “balık kılıcı haritaları” da kaynaklarda diğeri adı geçen kavram haritalama çeşitleridir (Demirel 2002).

Kavram haritaları başlangıç itibariyle bir öğretim aracı olarak kullanılmış olsa da, son yıllarda öğrencilerin bilişsel yapılarını ortaya çıkarmada ölçme-değerlendirme aracı olarak da özellikle dikkat çekmektedir (Ruiz-Primo, 2004; Şen, Aykutlu 2008). Kavram haritasının puanlama biçimi, hazırlanma şekline göre değişmektedir. McClure, Sonak ve Suen (1999), literatürde en fazla kullanılan puanlama modellerini altı şekilde özetlemiştir: 1. Bütünsel Puanlama Modeli, 2. Uzman Harita ile Bütünsel Puanlama, 3. Yapısal Puanlama Modeli, 4. Uzman Harita ile Yapısal Puanlama Modeli, 5. İlişkisel Puanlama Modeli, 6. Uzman Harita ile İlişkisel Puanlama Modeli. Bütünsel puanlama modelinde, kavram haritasını oluşturan kişinin kavramları genel olarak anlayıp anlamadığı araştırılır (McClure, Sonak & Suen 1999). Oluşturulan her bir kavram bir puanlama cetveli üzerinde genellikle 1’den 10’a kadar olan bir aralıkta değerlendirilir. Hiyerarşik olarak hazırlanmış kavram haritalarının puanlanmasında, yapısal puanlama modeli olarak bilinen Novak’ın puanlama biçimi tercih edilir (Novak 1984). Bu yöntem, yüksek seviyeli bilişsel yapıları ve ifadelerin tanımlarını dikkate alır. Bu değerlendirmede, hiyerarşi seviyelerinin sayısı, kavramlar arası bağlar, çapraz bağlar ve varsa kullanılan örnekler puanlamaya katılır. Bunların her birisinin kendi içerisinde belirli bir puanlama değeri mevcuttur. Diğeri bir puanlama modeli, McClure ve Bell tarafından 1990 yılında ilişkisel puanlama olarak geliştirilmiştir (McClure, Bell 1990). Bu yaklaşımda; her bir kavram haritasındaki kavramlar ve bu kavramları birbirine bağlayan ifadeler, 0 ile 3 puan arasında puanlanır. McClure, Sonak ve Suen (1999) tarafından modellenen puanlama yöntemlerinde adı geçen uzman kavram haritası ile olan bölümlerde, konu alanında uzmanlar tarafından oluşturulan kavram haritaları ile öğrencilerin yapmış olduğu kavram haritalarının kıyaslanması söz konusudur.

Öğrencilerin öğrenme performanslarını belirlemede kullanılan tamamlayıcı ölçme araçlarından bir diğeri de yapılandırılmış griddir. Yapılandırılmış grid tekniğinin eğitimde ölçme-değerlendirme aracı olarak kullanılması ile ilgili çalışmalar, Egan (1972) tarafından başlatılmış ve sonrasında birçok araştırmacı tarafından başarı ile uygulanmıştır (MacGuire & Johnstone 1987; Bahar 2003; Durmuş & Karakırık 2005; Bahar, Aydın, Karakırık 2009). Egan (1972), yapılandırılmış grid tekniğini öğrencilerin kendilerine rastgele verilen bilgileri yeniden düzenlerken, bilgileri ve öğrendiklerini kullanarak öğretici ile iletişime geçmesi olarak tanımlamaktadır. Öğrencilerin verdikleri yanıtlarda, zihinlerinde oluşan fikir ve gerekçeleri gözlenmektedir (Johnstone & Ambusaidi 2000). Yapılandırılmış grid, zihindeki yapıların ve eksiklerin ortaya çıkmasına ışık tutan bir yaklaşımdır. Bu yöntemde, her bir kutucuğun içinde bir kavram, tanım, önerme, şekil v.b. yer alır. Böylece öğretmene zengin ve objektif değerlendirme imkanı sağlar (Johnstone, Bahar & Hansell 2000, Durmuş & Karakırık 2005). Bir grid ile ilgili birden fazla soru sorulabilir. Her bir soru için geçerli kutucuk sayısı birbirinden farklıdır. Her bir sorunun yanıtı genellikle birden fazla kutuda yer alır. Bir kutucuk birden fazla sorunun yanıtı olabilir. Kutucuklarda yer alan bilgiler birbiriyle ilişkilidir. Öğrencilerden her bir soru ile ilgili olan kutucukları, ilgili olmayan kutucuklardan ayırmaları istenir (Johnstone, Bahar & Hansell 2000). Yani; öğrenciden her bir soru için geçerli doğru yanıtı, diğeri soruların yanıtlarından ayırması beklenmektedir. Öğrencilerin verdiği yanıt o konudaki bilgi seviyesini, bilgi eksikliğini, kavramsal bağları veya yanlış kavramları gösterir (Bahar 2003). Böylece; yapılandırılmış grid sayesinde, öğrencinin zihninde oluşmuş bilişsel yapının zayıf ve güçlü yanlarını teşhis etme olanağı sağlanmaktadır (Hassan, Hill & Reid 2004).

Kutucuklarda tek bir doğru yanıtın olmaması ve her bir kutucukta verilen yanıtın sorulardan en az biri için doğru olması, grid tekniğinin en önemli özelliğidir. Öğrencinin her bir soru için geçerli olan kutucuk sayısını bilmemesi, bu yaklaşımda tahminle doğru yanıtı ulaşma şansını düşürmektedir. Bundan dolayı öğrencinin seçtiği yanlış kutucuklardan öğrencinin zihninde oluşan kavram yanılgıları,

doğru kutucuklardan ise öğrendikleri tespit edilebilmektedir. Öğrencilerin kritik düşünme becerisini geliştirmesi, şans başarısını en aza indirmesi gibi birçok avantajından dolayı 2004 yılı MEB Fen ve Teknoloji dersi programında, tamamlayıcı değerlendirme yaklaşımlarından bahsedilirken, yapılandırılmış gride de özellikle yer verilmiştir (MEB 2008).

Grid tekniğinin analiz yönteminde 2 farklı puanlama türü kullanılır. Bu puanlama türlerinden en fazla kullanılan ve bu çalışmada da tercih edilen puanlama türü aşağıdaki şekildedir:

$$\frac{C1}{C2} - \frac{C3}{C4}$$

C1: Seçilen doğru kutucuk sayısı

C2: Toplam doğru kutucuk sayısı

C3: Seçilen yanlış kutucuk sayısı

C4: Toplam yanlış kutucuk sayısı

Bu formüle göre, öğrencilerin puanları (-1) ve (+1) arasında değişir. Formülden elde edilen puanlar 1 ile toplanarak negatif değerler pozitifte dönüştürülür. Daha sonra 5 ile çarpılarak, 0-10 aralığında puanlar elde edilir (Johnstone, Bahar, & Hansell 2000).

Son yıllarda geleneksel ölçme araçlarına tamamlayıcı nitelikte kullanılan kavram haritası ve yapılandırılmış gridin, geçerlik ve güvenirlikleri üzerine yapılan çalışmalar artarak devam etmektedir. Böylece; bu araçların bir ölçme aracında olması gereken nitelikleri ne ölçüde sağladığı görülecek ve bu sayede daha etkin bir şekilde ölçme-değerlendirme sürecinde pratik uygulamalarına katkı sağlanacaktır. Bu çalışmada, “Newton’un Hareket Yasaları” ünitesinde öğrencilerin başarılarını ölçmek için kullanılan kavram haritası ve yapılandırılmış gridin geçerlik ve güvenirlikleri araştırılarak, bu konudaki literatüre katkı sağlamak amaçlanmıştır.

2. YÖNTEM

Bu araştırma, tamamlayıcı ölçme araçlarından kavram haritası ve yapılandırılmış gridden elde edilen puanların geçerlik ve güvenirliklerini inceleyen betimsel bir çalışmadır. Çalışma grubunu, 2009-2010 öğretim yılı güz döneminde bir devlet üniversitesindeki Fen Bilgisi Öğretmenliği ve Fizik Öğretmenliği programı I. sınıfta okuyan 102 öğrenci oluşturmaktadır. Her iki grubun fen puanları ile üniversiteye yaklaşık eşit puanlarla girmiş olması ve derslerde benzer içerik ve kapsamı öğrenmiş olmaları nedeniyle, bu grupların eş değer olduğu kabul edilmiştir.

2.1. Veri Toplama Araçları

Bu çalışmada veri toplama araçları olarak; kısa cevaplı test, kavram haritası ve yapılandırılmış grid kullanılmıştır. Kısa cevaplı testin sonuçları, kavram haritası ve yapılandırılmış gridin geçerliğini test etmede ölçüt olarak alınmıştır.

Genel olarak Fen Bilgisi ve Fizik Öğretmenliği ders programlarının kur tanımlarında adı geçen “Newton’un Hareket Yasaları” ünitesi içerik olarak ele alınmıştır. Bilişsel düzeyin bilgi basamağında 7, kavrama basamağında ise 5 tane olmak üzere toplam 12 hedef-davranış belirlenerek belirtke tablosu oluşturulmuştur. Bu tabloda, klasik fiziğin temeli sayılan; kuvvet, ivme, eylemsizlik, sürtünme kuvveti ve Newton’un hareket yasaları kavramları ile bunlara yönelik örnekler bulunmaktadır. Aşağıda, belirtke tablosu esas alınarak hazırlanan kısa cevaplı test, kavram haritası ve yapılandırılmış gridin geliştirilme aşamaları ele alınmıştır.

2.1.1. Kısa Cevaplı Testin Geliştirilmesi

Belirtke tablosu doğrultusunda belirlenen her bir hedef-davranışa yönelik iki soru olmak üzere toplam 24 soru yazılmıştır. Kısa cevaplı testin geçerlik çalışması için sekiz uzmanın görüşlerine başvurulmuştur. Uzmanlardan dördü fizik eğitimi, biri fen eğitimi, biri de ölçme ve değerlendirme

alanında doktora eğitimine sahip kişilerdir. İkisi ise fizik eğitimi alanında yüksek lisans eğitimi almış ölçme ve değerlendirme uzmanlarıdır. Uzmanlardan gelen dönütler göz önüne alınarak, sorularda gerekli düzenlemeler yapılmış ve 24 soruluk deneme uygulaması son halini almıştır. Kısa cevaplı test için deneme uygulaması, bu araştırmadaki çalışma grubu ile benzer düzeydeki 138 öğrenci ile yapılmıştır. Testte yer alan her bir soru doğru yanıtlar için 1, yanlış yanıtlar için 0 olarak puanlanmış ve ITEMAN programında yapılan analizlerle test ve madde istatistikleri elde edilmiştir. Kısa cevaplı testin nihai uygulaması için, her bir hedef-davranışa yönelik yazılan 2 sorudan madde ayırt edicilik gücü ve güçlük indeksleri göz önüne alınarak bir tanesi seçilmiştir. Nihai uygulamaya alınan 12 maddenin ayırt edicilik gücü indeksleri 0.20-0.75 arasında değişmektedir. Madde ayırt edicilik indeksleri 0.20-0.30 arasında değişen 2 madde, uzman önerisi alınarak düzeltilmiştir. Maddelerin güçlük indeksleri ise 0.30-0.91 arasında yer almaktadır.

2.1.2. Kavram Haritasının Geliştirilmesi

Kavram haritasının oluşturulmasında, belirtke tablosu doğrultusunda belirlenen her bir hedef-davranışa yönelik bir kavram olmak üzere, toplam 12 kavram belirlenmiştir. Kavram haritasının istenilen hedef-davranışlar doğrultusunda yapılabilirliğini tespit etmek için, 40 öğrenci ile deneme uygulaması yapılmıştır. Deneme uygulamasının sonuçları iki uzman tarafından kavram haritalarının belirtke tablosu doğrultusunda hazırlanabilirliği ve öğrenciler tarafından anlaşılabilirliği açısından incelenmiştir. Uzmanlardan birisi fizik eğitimi alanında doktora eğitimini yapmış diğeri ise fizik eğitimi alanında yüksek lisans eğitimi almış ölçme ve değerlendirme uzmanıdır.

2.1.3. Yapılandırılmış Gridin Geliştirilmesi

Yapılandırılmış grid için, belirtke tablosu doğrultusunda belirlenen her bir hedef-davranışa yönelik bir soru olmak üzere toplam 12 soru yazılmıştır. Sorular iki bölüme ayrılmış olup, her bir bölümde yer alan soruların yanıtları 16'lı hazırlanmış gridin kutucuklarına rastgele yerleştirilmiştir. Yapılandırılmış gride yer alan soruların anlaşılabilirliğini ve cevaplanabilirliğini görmek için 40 öğrenci ile deneme uygulaması yapılmıştır. Uygulama sırasında öğrencilerden gelen dönütler göz önüne alınarak, anlaşılmadığına kanaat getirilen kutucuklar, doktorasını fizik ve fen eğitiminde yapmış olan iki uzmandan gelen görüşler doğrultusunda tekrar gözden geçirilerek düzeltilmiştir. Ek 1'de, uygulanan yapılandırılmış gridlerden biri örnek olarak verilmiştir.

2.2. Ölçme Araçlarının Uygulanması ve Puanlanması

Testlerin nihai uygulamaları 2009-2010 öğretim yılı güz döneminde bir devlet üniversitesinin Fen Bilgisi Öğretmenliği ve Fizik Öğretmenliği programındaki 102 öğrenci üzerinde gerçekleştirilmiştir. Uygulamalar öncesinde, çalışmaya katılan bütün öğretmen adaylarına gerek kavram haritası gerekse yapılandırılmış grid hakkında yaklaşık yarım saat süren bir bilgilendirme yapılmış ve katılımcıların her iki uygulamayı da yapabileceklerine kanaat getirildiğinde araştırmaya başlanmıştır.

Nihai uygulama için; 12 sorudan oluşan kısa cevaplı testte yer alan her bir soru doğru yanıtlar için 1, yanlış yanıtlar için 0 olarak puanlanmış, daha sonra test ve madde istatistikleri hesap edilmiştir. Kısa cevaplı testin nihai uygulamasından öğrencilerin aldığı puanların ortalaması 9.83 ve standart sapması 2.23'tür. Öğrencilerin çoğunluğu, ortalamanın üstünde başarı göstermişlerdir. Kısa cevaplı testin iç tutarlılık anlamındaki güvenilirliği için KR-20 güvenilirlik katsayısı hesaplanmış ve 0.74 bulunmuştur. Kehoe (1995)'e göre, 10-15 maddeden oluşan kısa testlerde KR-20 güvenilirlik katsayısının 0.50 ve üzerinde değer alması yeterlidir. Buna göre kısa cevaplı testin güvenilirliğinin yeterli olduğu söylenebilir. Kısa cevaplı testin puanlayıcı güvenilirliğini incelemek amacıyla, araştırmacılarından biri ve bir fizik eğitimi uzmanı tarafından yapılan puanlamalar arasındaki korelasyon katsayısı 0.92 ($p < 0.05$) olarak bulunmuştur. Birinci puanlayıcı ve ikinci puanlayıcının verdiği puanlar arasında pozitif yönde, yüksek düzeyde ve anlamlı bir ilişki olduğu görülmektedir. Birinci puanlayıcının verdiği puanla, ikinci puanlayıcının verdiği puanların birlikte değişim gösterdiği ve büyük ölçüde birbirine yakın olduğundan kısa cevaplı testin puanlayıcı güvenilirliğinin de yüksek olduğu belirtilebilir.

Ek 2’de bir örneği verilmiş olan, öğrencilerin hazırlamış oldukları kavram haritalarının değerlendirilmesinde; her bir hedef-davranışı içeren kavramların kavram haritalarında olup olmamasına ve bu kavramlar arasındaki ilişkilerin fiziksel olarak doğruluk derecesine bağlı olarak Tablo 1’deki puanlama modeli kullanılmıştır. Bu puanlama yöntemi, McClure ve Bell (1990) tarafından geliştirilen ilişkisel puanlama modeli temel alınarak geliştirilmiştir.

Tablo 1: Kavram Haritası Puanlama Cetveli

	0	1	2
Kavram yok-ilişki yok	X		
Kavram var-ilişki yok	X		
Kavram var-ilişki yanlış	X		
Kavram var-ilişki kısmen doğru		X	
Kavram var-ilişki doğru			X

Öğrencilerin yapılandırılmış gridden aldıkları puanlar, giriş bölümünde yapılandırılmış grid için verilen formül kullanılarak hesaplanmıştır.

2.3. Verilerin Analizi

Kavram haritasının geçerlik katsayısının hesaplanmasında, ölçüt geçerliği belirlenmiş ve araştırmacılar tarafından geliştirilen kısa cevaplı ölçüt olarak test kabul edilmiştir. Kavram haritası ve kısa cevaplı test arasındaki ilişkinin derecesi için Pearson Momentler Çarpımı korelasyon katsayısı hesaplanmıştır (Baykul 2000). Güvenirlik katsayısının hesaplanması için, iç tutarlılık ve puanlayıcı güvenilirliği belirlenmiştir. Kavram haritalarının iç tutarlılık güvenirlilik hesaplanmasında Cronbach alfa katsayısı kullanılmıştır. Puanlayıcı güvenirliklerinin araştırılması için, araştırmacıardan biri ve fizik eğitiminde doktora eğitimine sahip bir öğretim üyesi tarafından kavram haritaları puanlanmıştır. Puanlayıcıların verdikleri puanlar arasındaki ilişkinin derecesi Pearson Momentler Çarpımı korelasyon katsayısı ile hesaplanmıştır.

Yapılandırılmış gridin geçerlik katsayısının hesaplanmasında da, ölçüt geçerliliği belirlenmiş ve ölçüt olarak, araştırmacılar tarafından geliştirilen kısa cevaplı test kabul edilmiştir. Yapılandırılmış grid ve kısa cevaplı test arasındaki ilişkinin derecesi için Pearson Momentler Çarpımı korelasyon katsayısı hesaplanmıştır. Güvenirlik belirlenmesinde, iç tutarlılık güvenirlilik katsayısı olan Cronbach alfa hesaplanmıştır.

3. BULGULAR

Yapılan çalışmada, Fen Bilgisi Öğretmenliği ve Fizik Öğretmenliği öğrencilerinin kavram haritası ve yapılandırılmış gridden elde edilen başarı puanları, kısa cevaplı testten elde edilen başarı puanları ile karşılaştırılmıştır. 102 öğrenciye uygulanan kavram haritası ve yapılandırılmış gridden elde edilen puanlara ilişkin betimsel istatistikler Tablo 2’de toplu halde verilmiştir.

Tablo 2: Ölçme Araçlarına İlişkin Betimsel İstatistikler

Ölçme Aracı	N	\bar{X}	X_{tepe}	S_x	X_{min}	X_{mak}
Kavram Haritası	102	16.9	16.5	4.38	4.5	24
Yapılandırılmış Grid	102	102.8	91.7	8.86	72.6	114.7

Uygulanan ölçme araçlarına ilişkin betimsel istatistikler incelendiğinde; öğrencilerin kavram haritasından aldığı puanların ortalaması 24 tam puan üzerinden 16.9 dur (Tablo 2). X_{tepe} değeri 16.5 olduğundan, öğrencilerin çoğunluğu, ortalamaya yakın ve birbirinden farklı başarı ($S=4.38$) sergilemişlerdir. Tepe değer en büyük frekansın hangi puanda olduğunu göstermektedir ve aritmetik ortalama gibi istatistiklerle karşılaştırılarak, frekans dağılımlarının karşılaştırılmasında kullanılır (Turgut 1992). Kavram haritasına ilişkin, madde-test korelasyonları ranjı 0.41 ile 0.60 arasında değişmektedir.

Yapılandırılmış gridden öğrencilerin aldığı puanların ortalaması 120 tam puan üzerinden 102.8'dir (Tablo 2). X_{tepe} değeri 91.7 olduğundan, öğrencilerin çoğunluğu, ortalamanın altında ve birbirinden farklı başarı ($S=8.86$) göstermişlerdir. Yapılandırılmış gride ilişkin, madde-test korelasyonları ranjı da 0.32 ile 0.71 arasında değişmiştir.

3.1. Kavram haritasına ilişkin bulgular

Kavram haritasının geçerliği için kavram haritasından ve kısa cevaplı testten elde edilen puanlar arasındaki Pearson korelasyon katsayısı hesaplanmıştır. Elde edilen korelasyon katsayısı 0.57 ($p<0.05$) bulunmuştur. Kavram haritası ve kısa cevaplı test arasında pozitif yönde, orta düzeyde ve anlamlı bir ilişki olduğu görülmektedir. Buna göre, öğrencilerin kavram haritası ve kısa cevaplı testten elde ettikleri başarı puanlarının birlikte değişimi orta düzeydedir. Kavram haritasının ölçmek istediği özellikleri ölçmede amacına hizmet ettiği söylenebilir.

Kavram haritasının güvenilirlik düzeyine yanıt bulabilmek için; iç tutarlılık anlamındaki güvenilirlik katsayısı olan Cronbach alfa hesaplanmış ve 0.69 bulunmuştur. Nunnally (1978)'a göre; başarı testleri için Cronbach alfa güvenilirlik katsayısının 0.70 ve üzerinde değer aldığı durumlarda ölçme aracı güvenilir kabul edilebilir. Bu çalışmada kavram haritasından elde edilen alfa değeri 0.70 değerine çok yakın olmasından dolayı, kullanılan ölçme aracının güvenilir olduğu belirtilebilir. Buna göre kavram haritasının iç tutarlılığının yeterli olduğu ve kavram haritasında yer alan kavramların birbiri ile ve ölçme aracının bütünüyle uyumlu olduğu görülmektedir.

Kavram haritasının puanlayıcı güvenilirliğini incelemek amacıyla, iki puanlayıcının verdiği puanlar arasındaki korelasyon katsayısı hesaplanmıştır. Birinci ve ikinci puanlayıcının verdiği puanlar arasında pozitif yönde, yüksek düzeyde ve anlamlı bir ilişki olduğu görülmektedir ($r=0.76$; $p<0.05$). Buna göre, birinci puanlayıcının verdiği puanla, ikinci puanlayıcının verdiği puanların birlikte değişiminin yüksek olduğu ve birbirleriyle yüksek oranda tutarlı olduğu söylenebilir.

3.2. Yapılandırılmış gride ilişkin bulgular

Yapılandırılmış gridin geçerlik düzeyine yanıt bulabilmek için, yapılandırılmış grid ve kısa cevaplı testten elde edilen puanlar arasındaki ilişkinin derecesi hesaplanmıştır. Elde edilen korelasyon katsayısı 0.69 ($p<0.05$) bulunmuştur. Yapılandırılmış grid ve kısa cevaplı test arasında pozitif yönde, yüksek düzeyde ve anlamlı bir ilişki olduğu görülmektedir. Buna göre, öğrencilerin yapılandırılmış grid ve kısa cevaplı testten elde ettikleri başarı puanlarının birlikte değişiminin yüksek olduğu söylenebilir. Yapılandırılmış gridin ölçmek istediği özellikleri ölçmede amacına hizmet ettiği görülmektedir.

Yapılandırılmış gridin güvenilirlik düzeyine yanıt bulabilmek için, iç tutarlılık anlamındaki güvenilirlik katsayısı olan Cronbach alfa hesaplanmış ve 0.77 bulunmuştur. Buna göre yapılandırılmış gridin iç tutarlılığının yüksek olduğu ve gride yer alan her bir maddenin birbiri ile ve ölçme aracının bütünüyle uyumlu olduğu görülmektedir.

3.3. Kavram Haritası ve Yapılandırılmış Grid Arasındaki İlişki

Kavram haritası ve yapılandırılmış griden elde edilen puanlar arasındaki korelasyon katsayısı 0.51 ($p<0.05$) olarak bulunmuştur. Kavram haritası ve yapılandırılmış grid arasında pozitif yönde, orta düzeyde ve anlamlı bir ilişki olduğu görülmektedir. Buna göre, öğrencilerin kavram haritası ve yapılandırılmış gridden elde ettikleri başarı puanlarının birlikte değişiminin orta düzeyde olduğu söylenebilir.

Yapılandırılmış gridin geçerlik ve güvenilirlik katsayılarının (0.69-0.77) kavram haritasına (0.57-0.69) kıyasla yüksek olduğu görülmektedir. Buradan, yapılandırılmış gridin kullanılış amacına hizmet etme, iç tutarlılık ve tesadüfi hatalardan arınıklık derecelerinin kavram haritasına göre daha yüksek olduğu sonucu çıkarılabilir.

4. TARTIŞMA ve SONUÇ

Araştırmanın sonucunda elde edilen bulgular; kavram haritasının fen bilgisi ve fizik derslerinde geçerli ve güvenilir bir araç olarak kullanılabilceğini göstermiştir. Kavram haritasının diğer ölçme araçları ile ilişkilerinin belirlenmesine, literatürde pek çok araştırmada rastlanmaktadır. Bu çalışmalarda, kavram haritasının diğer yöntemlerle arasında değişik düzeylerde ilişki bulunmuştur. Örneğin; Novak, Gowin ve Johansen (1983) kavram haritalarından kendi yaklaşımlarına göre elde ettikleri puanları, okuldaki başarı notları ile karşılaştırmışlar ve aralarında yaklaşık sıfır düzeyinde ilişki tespit etmişlerdir. Rye ve Rubba (2002) ise, kavram haritaları ile California başarı testini karşılaştırmış ve aralarında 0.72 düzeyinde ilişki bulmuşlardır. Şahin (2003) çalışmasında, üç farklı şekilde puanlanan kavram haritaları ile kısa cevaplı test puanlarını karşılaştırmış ve aralarında 0.70 düzeyinde ilişki belirlemiştir. Kandil-İngeç (2008), üç farklı türde puanlanan kavram haritalarının her birisini çoktan seçmeli başarı testi puanları ile karşılaştırmış ve aralarında 0.35, 0.36, 0.22 düzeyinde ilişki bulmuştur. Literatürde kavram haritalarının güvenilirlik belirleme çalışmalarına bakıldığında yine değişik değerler aldığı görülmektedir. McClure, Sonak ve Suen (1999), öğrencilere uygulanan kavram haritalarını 6 farklı puanlama yöntemini kullanarak değerlendirmişler ve elde edilen güvenilirlik katsayılarının $r=0.23$ den $r=0.76$ 'ya kadar farklı değerler aldığını tespit etmişlerdir. Rye ve Rubba (2002) çalışmalarında, kavram haritalarının puanlayıcı güvenilirliğini hesaplamış ve 0.92 olarak bulmuşlardır. İyilik (2007), doldurma ve oluşturma kavram haritalarının iç tutarlılık güvenilirliğini sırasıyla 0.90 ve 0.96; puanlayıcı güvenilirliklerini ise 1.00 ve 0.79 olarak bulmuştur. Şu ana kadar yapılan araştırmalar incelendiğinde, kavram haritalarının geçerlik ve güvenilirlik çalışmalarında, farklı puanlama türlerinin diğer geleneksel ölçme araçları ile olan ilişkilerinin daha fazla irdelenmesi gerekliliğini ortaya koymaktadır.

Ayrıca, bu araştırmadan elde edilen bulgular; yapılandırılmış gridin, fen bilgisi ve fizik derslerinde geçerli ve güvenilir bir araç olarak kullanılabilceğini göstermiştir. Yapılandırılmış gridin diğer ölçme araçları ile ilişkilerinin belirlenmesine ilişkin literatürde çok az örneğe rastlanmaktadır. Hassan, Hill ve Reid (2004) çalışmalarında, 12'li grid ve dönem sonu sınavını karşılaştırmışlar ve bazı sorular arasındaki korelasyonu 0.19 ile 0.34 arasında değişen bir aralıkta bulmuşlardır. Danili ve Reid (2005), çoktan seçmeli, kısa cevaplı test ve yapılandırılmış grid arasında 0.30-0.71 aralığında değişen korelasyon tespit etmişlerdir. Yapılan kaynak taramasında, yapılandırılmış gridin güvenilirliğine ilişkin bir çalışmaya rastlanılmamıştır. Bu çalışma kapsamında elde edilen sonuçlar, literatürde bu alandaki boşluğu doldurmada önemli bir katkı sağlayacaktır.

Araştırmanın sonunda elde edilen bulgular; kavram haritası ve yapılandırılmış grid arasında pozitif yönde, orta düzeyde ve anlamlı bir ilişki olduğunu göstermiştir. Ancak kavram haritasının güvenilirlik ve geçerliğinin, yapılandırılmış gride kıyasla daha düşük olduğu görülmüştür. Bu nedenle; yapılandırılmış gridden elde edilen puanların iç tutarlılık açısından daha güvenilir ve daha geçerli sonuçlar verebileceği, bu çalışma sınırları kapsamında, söylenebilir. Yapılandırılmış gridin cevaplandırılması öğrencilerin aşına olduğu çoktan seçmeli teste benzemekte iken, kavram haritalarının oluşturulması daha fazla beceri gerektirmektedir. Yapılandırılmış gridin puanlanması da daha önceden verilen puanlama formülleri kullanıldığı için kavram haritasına göre daha objektiftir. Bu nedenlerden dolayı, yapılandırılmış gridden elde edilen puanların geçerlik ve güvenilirlik katsayıları kavram haritasından daha yüksek çıkmış olabilir. Ancak, kavram haritasının özellikle görsel olarak bilişsel düzeydeki yapının ortaya çıkarılmasında daha ön plana çıktığı söylenebilir.

Bu araştırmadan elde edilen sonuçlar ışığında, aşağıdaki öneriler sıralanabilir: Bu çalışmada, konu alanı olarak fen bilgisi ve fizik dersleri seçilmiştir. Diğer konu alanları için de benzer çalışmalar yapılabilir. Bu çalışmada, yapılandırılmış grid için iki farklı tip puanlamadan sadece bir tanesi kullanılmıştır. Kavram haritaları için de birden fazla puanlama türü geçerlidir. Burada kullanılan puanlamalardan farklı puanlama türleri kullanılarak sonuçlara bakılabilir. Ayrıca, yapılacak başka

çalışmalarda, ölçüt olarak kısa cevaplı test dışında, diğer geleneksel veya tamamlayıcı ölçme araçları da kullanılıp, sonuçlar karşılaştırılabilir. İlköğretim ve ortaöğretim basamaklarında da benzer çalışmalar yapılarak elde edilen sonuçların yaşa ve seviyeye göre değişimi incelenebilir.

KAYNAKLAR

- Bahar, M. (2003). A study of pupils ideas about the concept of life. *Kastamonu Eğitim Dergisi*, 11, 93-104.
- Bahar, M., Aydın, F., & Karakırık, D. (2009). A diagnostic study of computer application of structural communication grid. *The Turkish Online Journal of Educational Technology*, 8 (2), 5-20.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulamaları*. Ankara: ÖSYM Yayınları.
- Danili, E., & Reid, N. (2005) Assessment formats: do they make a difference? *Chemistry Education Research and Practice*, 2005, 6 (4), 204-212.
- Demirel, Ö. (2002). *Planlamadan değerlendirmeye öğretme sanatı*. Ankara: PegemA Yayıncılık.
- Demirel, Ö. (2007). *Eğitimde program geliştirme*. Ankara: Pegem A Yayıncılık.
- Durmuş, S., & Karakırık, E. A (2005). Computer assessment tool for structural communication grid. *The Turkish Online Journal of Educational Technology*, 4, 1303-6521.
- Edwards, J., & Fraser, K. (1983). Concept maps as reflectors conceptual understanding. *Research in Science Education*, 13(1), 19-26.
- Egan, K. (1972). Structural communication- a new contribution to pedagogy. *Program Learning and Educational Technology*, 1, 63-78.
- Hassan, A.K., Hill, R.A., & Reid, N. (2004). Ideas underpinning success in an introductory course in organic chemistry. *The Higher Education Chemistry Journal of The Royal Society of Chemistry*, 8 (2), 40-52.
- İyilik, H. (2007). *İki farklı kavram haritası tekniğinden elde edilen puanların güvenilirlik ve geçerliğinin incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.
- Johnstone, A. H., & Ambusaidi, A. (2000). Fixed response: what are we testing? *Chemistry Education: Research and Practice in Europe*, 1(3), 323-328.
- 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.
- Kandil-İnceç, Ş. (2008). Kavram haritalarının değerlendirme aracı olarak fizik eğitiminde kullanılması. *Hacettepe University Journal of Education*, 35, 95-206.
- Kehoe, J. (1995). Basic item analysis for multiple-choice tests. *Practical Assessment, Research & Evaluation*, 4(10). 12.04.2010 tarihinde <http://PAREonline.net/getvn.asp?v=4&n=10> adresinden alınmıştır.
- MacGuire, P.R.P. & Johnstone, A.H. (1987). Techniques for investigating the understanding of concepts in science. *International Journal of Science Education*, 9, 565-577.
- McClure, J.R., & Bell, P.E. (1990). *Effects of an environmental education related STS approach instruction on cognitive structures of pre-service science teachers*. University Park, PA: Pennsylvania State University. (ERIC Document Reproduction Services No. ED 341 582).
- McClure, J.R., Sonak, B., & Suen, H.K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36, 4, 475-492.
- Milli Eğitim Bakanlığı, (2008). *Program Dosyaları*. 14.04.2010 tarihinde http://ttkb.meb.gov.tr/ogretmen/modules.php?name=downloads&d_op=viewdownload&cid=1 adresinden alınmıştır.
- Novak, J. D. (1984). Application of advances in learning theory and philosophy of science to the improvement of chemistry teaching. *Journal of Chemical Education*, 61, 607-612.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- Novak, J. D., Gowin, D. B., & Johansen, G. T. (1983). The use of concept mapping and knowledge: Vee mapping with junior high school science students. *Science Education*, 67(5), 625-645.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Ruiz-Primo, M.A. (2004). Examining concept maps as an assessment tool. *Computer and Information Science*, 36-43.
- Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Teaching*, 38(2), 260-278.
- Rye, J. A., & Rubba, P. A. (2002). Scoring concept maps: An expert map-based scheme weighted for relationships. *School Science and Mathematics*. 102(1), 33-42.
- Şahin, B. (2003). *Matematik dersinde kavram haritası yöntemini kullanarak öğrenci başarısının değerlendirilmesi*. Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Şen, A. İ., & Aykutlu, I. (2008). Using concept maps as an alternative evaluation tool for students' conceptions of electric current. *The Eurasian Journal of Educational Research*, 31 (8), 75-92.
- Turgut, M.F. (1992). *Eğitimde ölçme ve değerlendirme*. Ankara: Saydam Matbaacılık.

Extended Abstract

In line with the effects of the constructivist approach on teaching programs, modern perspectives about the assessment and evaluation techniques have emerged. Evaluation in teaching and learning practices within the constructivist approach emphasize the process. Students are provided with multiple assessment and evaluation opportunities through the utilization of traditional assessment tools and their results together with supplementary and complementary assessment tools. This study involves the analysis on the reliability and validity of the scores obtained from concept maps and structural communication grid as two of the complementary assessment tools.

Concept maps were developed by Novak et al. at Cornell University at beginning of 70s based on the “Meaningful Learning” approach introduced by Ausubel. Concept maps are methods of display where concepts are linked to each other through linking sentences. Concept maps, which were initially used as teaching tools, have recently been utilized as assessment tools for revealing the cognitive structures of students. The scoring patterns of the concept maps differ according to their methods of preparation. Another complementary assessment tool used for the determination of students’ learning performances is the structural communication grid. The studies on utilizing structural communication grids as assessment tools in education were initiated by Egan (1972) and practiced by various researchers afterwards. Egan defines the structural communication grid technique as students’ communicating with the teacher while reorganizing the randomly given information using what s/he has learnt and the knowledge.

Recently, studies on the validity and reliability of concept maps and structural communication grids as complementary assessment tools have increased in quantity. Therefore, it is made clear to what extent these tools address the requirements of an assessment tool, which allows for more effective contributions to practical applications during the assessment and evaluation process. This descriptive study aims to contribute to the literature through a research on the validity and reliability of concept maps and structured communication grids used to assess students’ achievement regarding the unit called “Newton’s Laws of Motion”. The sampling consisted of 102 students studying in their Year One at Science Teaching and Physics teaching departments of a state university during the fall semester of the 2009 – 2010 academic year.

The data collection tools used in the study consisted of the short-answer test, concept map and structured communication grid. The responses to the short-answer test were made us of as the main measure in checking the validity of the concept map and the structured communication grid. “Newton’s Laws of Motion” was chosen as the target content referring to its presence in both Science and Physics Teaching syllabus. 12 target behaviors were identified to form the indication table consisting of 7 statements belonging to the knowledge stage of the cognitive level and 12 statements belonging to the comprehension level.

The Pearson Correlation Coefficient between the scores of the short-answer test and the concept map was calculated in order to obtain the validity of the concept map. The correlation coefficient calculated was found to be 0.57 ($p < 0.05$). A positive, medium level and significant relationship was observed between the concept map and the short answer test. In order to find the reliability state of the concept map, Cronbach Alpha, as the internal consistency coefficient, was calculated to be 0.69. It was observed that the internal consistency of the concept map was rather high. In order to examine the scoring reliability of the concept map, the correlation coefficient of both score providers was calculated and a positive, high and significant relationship was observed amongst ($r = 0.76$; $p < 0.05$). Therefore it is possible to say that the changing of the first score provider’s scores and the second score provider’s scores are rather high and strongly consistent with each other. Determination of the relationship between concept maps and other assessment tools is a very common topic in the literature. These studies concluded with relationships between concept maps and other methods at different levels. Looking at the research so far, it could be said that in reliability and validity studies of concept maps, the relationship between different scoring types and other traditional assessment tools should be examined.

In order to find out the validity level of the structured communication grid, the degree of the relationship between structured communication grid and short answer test scores was calculated. The obtained correlation coefficient was 0.69 ($p < 0.05$). A positive, high level and significant relationship between the structured communication grid and the short answer test. For the reliability of the structured communication grid, Cronbach Alpha was calculated and found to be 0.77, which indicated that the structured communication grid was reliable and had a high level of internal consistency. Despite the few studies in the literature about the relationship between the structured communication grid and the other assessment tools, no studies were found on its reliability. The conclusions of this study would contribute to the lack in the literature.

The degree of the relationship between the structured communication grid and the concept map was calculated and the correlation coefficient was found to be 0.51 ($p < 0.05$). A positive, medium level and significant relationship was observed between the structured communication grid and the concept map. Therefore it is possible to say that the changing of the scores of students obtained from the structured communication grid and the concept maps were at medium level. Responding to the structured communication grid is similar to the multiple-choice tests that students are familiar with, whereas constructing concept maps require better skills. The scoring system of the structured communication grid is more objective than that of the concept map because the previously set scoring formulas are used. That could be the reason why the reliability and validity coefficients of the structured communication grid were higher than that of the concept map.

Science and physics courses were selected as the field of study in this research. Similar studies could be applied to other subject areas. Other studies may compare and contrast with scoring systems different than the systems used in this research. Moreover, further studies may make use of other traditional or complementary assessment tools than short answer tests as measures and compare the results.