

Fizik Öğretiminde Kullanılan Yazılı Ölçme Türlerinin İtme-Momentum Konusu İçin Karşılaştırılması

The Comparing Of The Writing Measurement Formats Used At The Physics Education For The Impulse-Momentum Subject

Hasan Şahin KIZILCIK

G. Ü. Gazi Eğitim Fakültesi, O.F.M.A. Eğt. Böl., Fizik Öğretmenliği Anabilim Dalı, Ankara-TÜRKİYE

Mustafa TAN

G. Ü. Gazi Eğitim Fakültesi, O.F.M.A. Eğt. Böl., Fizik Öğretmenliği Anabilim Dalı, Ankara-TÜRKİYE

ÖZET

Bu çalışma, aynı hedef ve davranışlara yönelik olarak hazırlanmış, fizik eğitiminde kullanılan yazılı ölçme türlerinden objektif ölçme türü olarak bilinen, çoktan seçmeli test, kısa cevaplı test ve doğru-yanlış tipi testleri itme ve momentum konusu için karşılaştırmayı, aralarında anlamlı bir ilişkinin olup olmadığını tespit etmeyi amaçlamıştır. Aynı hedef ve davranışlara yönelik üç farklı türdeki ölçme aracı, Keçiören Kalaba Lisesi 10 Fen C sınıfında öğrenim gören 32 öğrenciye 2003-2004 Eğitim-Öğretim Yılı, II. döneminde uygulanmıştır. Araştırmanın sonucunda; İtme ve momentum konusu için uygulanan kısa cevaplı test ile çoktan seçmeli test arasında ve çoktan seçmeli test ile doğru-yanlış tipi test arasında anlamlı bir ilişki bulunmuştur. Uygulanan kısa cevaplı test ile doğru-yanlış tipi test arasında anlamlı bir ilişki bulunamamıştır. İtme ve momentum konusu için uygulanan kısa cevaplı test, çoktan seçmeli test ve doğru-yanlış tipi test ile öğrencilerin okullarında almış oldukları fizik dersi başarı notları arasında anlamlı bir ilişki bulunamamıştır.

Anahtar Kelimeler: Fizik Eğitimi, Ölçme ve Değerlendirme, İtme ve Momentum, Çoktan Seçmeli Test, Kısa Cevaplı Test, Doğru-Yanlış Testi.

ABSTRACT

This project, has been purposed of compare to the multiple-choice item test format, short-answer test format and true-false test format which are developed for measuring the same learning outcome. Those test formats are known objective test formats used for writing test formats at the physics education and purposed of establish correlation

among the tests for the impulse and momentum subject. The same learning outcomes and these three different sort of measurement formats which had been applied 32 senior students who attends on the science class 10-C in Keçiören Kalaba High School during the second semester of the 2003-2004 educational year. As a result; between the short-answer test format and multiple-choice item test format which are applied for impulse and momentum subject was found significant correlation. Between the short-answer test format and true-false test format which are applied was not found significant correlation. Between the true-false test format and multiple-choice item test format which are applied was not found significant correlation. Among the short-answer test format, multiple-choice item test format, true-false test format which are applied for impulse-momentum subject and the students' grade marks on physics lesson on their high school was not found significant correlation.

Key Words: Physics Education, Measurement and Assessment, Impulse and Momentum, Multiple-choice Item Test Format, Short-Answer Test Format, True-False Test Format.

1. Giriş

Eğitimde davranış değişikliği esastır. Eğitim sürecinin başarılı olup olmadığı, ne oranda başarılı olduğu, varsa aksaklıkları ve eksikliklerini görmek, bu konuda dönüt alabilmek oldukça önem taşımaktadır. Eğitim bir sistemdir ve her sistemde olduğu gibi, ölçme yapmak, bu ölçmeye bağlı olarak da değerlendirme yapmak gereklidir. Fizik eğitiminde de yapılan ölçme bu yüzden önem taşımaktadır.

Espinoza (2004), yapmış olduğu bir çalışmada, fiziğin bilimsel okur-yazarlığın gelişimi içerisindeki tartışılmaz merkezi rolünün, onun algılanan zorluğu yüzünden zayıfladığını ileri sürmüştür. Konuyla ilgili olarak Espinoza, lise öğrencilerinin momentum ve kuvvet kavramlarını kullanımı ile ilgili araştırmaları incelemiştir. Lise öğrencilerinin kuvvet ve momentumu kullanımları ve algılayışları ile ilgili analizlerin, fizik müfredatı içerisinde, koruma yasalarının; momentumun kuvvet üzerindeki bilişsel üstünlüğünden ötürü dinamik ve kinematikten daha önce gelmesi gerektiğini önemle vurguladığını belirtmiştir. Bu çıkarımı iki bulgu üzerine kurmuştur; (a) öğrenciler, bilgi verimi öncesindeki etkinliklerinde, momentum konusunda kuvvetten daha iyi bir performans sergilemiştir, (b) konuların sunum düzeni içerisindeki tersine çevrim, momentumun kuvvetten önce sunumunun, öğrencilerin mekaniği algılamalarındaki

gelişimi içeren standart yaklaşıma karşı üstün olduğunu gösterir. Bu yüzden, çalışma, günümüzün öğrenme teorileriyle tutarlı fizik öğretimi için pedagojik bir mantık sağlar.

Güneş, İnceç ve Taşar (2002), Gazi Eğitim Fakültesi'nde farklı anabilim dalı ve sınıflarda öğrenim gören 192 öğrencinin katılımı ile gerçekleştirdikleri bir çalışmada, itme ve momentum konusunda öğretmen adaylarında var olan kavramları incelemiştir. Verileri 2001 güz dönemi başlangıcında öğretmen adaylarına doğrudan tanım soruları sorularak toplamışlardır. Yazarlar, çalışmalarında, öğrencilerin momentum ve itme kavramlarını tanımlayabilme düzeylerini tespit etmeyi amaçlamışlar ve elde ettikleri bulgulara göre, özellikle liseden yeni gelen 1. sınıf öğrencilerinin bu kavramlara pek aşina olmadıklarını, fakat 4 yıllık öğrenimleri boyunca orta öğretim fizik öğretmenliği anabilim dalı öğrencilerinin tanımları anlamada önemli ölçüde gelişme gösterdiklerini tespit etmişlerdir. Elde ettikleri sonuçlar, üniversitede henüz fizik dersi almamış öğrencilerin itme kavramını tanımlayamadıklarını, fizik dersi alan öğrencilerin ise ancak yarıya yakınının bu tanımlamayı yapabildiklerini ortaya koymaktadır (İnceç ve diğerleri, 2002). Sorulara verilen cevaplarda gözlenen çeşitlilik ve farklılık bu kavramın gerçekten de anlaşılmasında büyük zorluk çekildiğinin bir göstergesi olarak görülmüştür. Bu kavramların öğretilmesi ve öğrenilmesini de gelişme sağlanması için hâlâ yapılması gerekenler olduğunu ifade etmişlerdir.

Eğitimde ölçme ve değerlendirme kabaca, “öğrenci hakkında bilginin toplanması, kaydedilmesi, yorumlanması ve kullanılması süreci” (Harlen ve diğerleri, 1991) olarak tanımlanabilir. Fizik eğitiminde de durum aynıdır. Ölçme ve değerlendirmenin temel amacı, öğrencinin öğrenmesine katkısını sağlamaktır. Bu temel amacı yerine getirebilmek için şekil verici bir değerlendirme kullanılması öngörülmektedir (Gipps 1994; Torrance ve Pryor, 1995; Black, 1998). Eğitim ile ilgili yapılan bütün etkinliklerin en son amacının öğrenci başarısı olduğu göz önüne alınırsa, yapılacak olan değerlendirmenin de bu amaca yönelik olması doğaldır. Bu nedenle yapılan araştırmalar genellikle öğrencilerin başarılarının nasıl arttırılacağı ve bununla ilgili alınması gereken önlemler üzerine yoğunlaşmıştır (Bakaç, 2003).

Ölçme ve değerlendirmenin niteliği üzerine yapılmış olan bir çalışmada ise Banta (1997), ölçme-değerlendirme ile ilgili on prensip önermiştir. Bunlar; Öğrenci öğrenmelerinin ölçme ve değerlendirmesi, (1) eğitimsel değerlerle başlar ve en çok, çok boyutlu olarak bir öğrenme anlayışını yansıttığında, tamamlandığında ve zaman içerisindeki performansı açığa vurduğunda etkili olur, (2) programlar gelişmenin yollarını açıkça aradığı ve amaçları açıkça belirttiği zaman en iyi şekilde işler, (3) sonuçlara dikkat çekmeyi gerektirmesinin yanı sıra, bu sonuçlara ulaştıran deneyimlere de eşit oranda dikkat çekmeyi gerektirir, (4) kesintili değil, sürekli olduğunda en iyi sonucu verir, (5) eğitimci topluluklarından temsilcilere de yer verildiğinde daha geniş çapta gelişme sağlanır, (6) insanların gerçekten önem verdiği soruları aydınlatıldığı ve insanlara yarar sağlayacak konularla başladığı zaman fark yaratır, (7) değişimi sağlayan durumların daha büyük bir parçası olduğu zaman, gelişime yönelmesi en büyük olasılıktır, (8) eğitimciler öğrencileri ve halkı sorumluluklarıyla tanıştırır, (9) yeni fikirlere açık, bu fikirleri destekleyici ve gerçekleşmesini sağlayıcı bir ortam içerisinde gerçekleştiği zaman en etkili hale gelir.

Bakaç (2003), İzmir çapında ilköğretim II. kademe öğrencilerine ve öğretmenlerine, bugün fen derslerinde uygulanmakta olan ölçme-değerlendirme sistemini, değerlendirmelerini ve gelecekte nasıl bir ölçme-değerlendirme sistemi olması ile ilgili görüşlerini almak amacıyla açık uçlu bir anket uygulamıştır. Kompozisyon niteliğindeki görüşleri, bir uzman grubu ile inceleyerek, 20 maddeden oluşan bir üçlü likert anket hazırlamıştır. Bu anketi, çeşitli illerde 476 öğrenci ve 16 öğretmene ikişer kez uygulamıştır. Bakaç, sınav sorularının tüm konuları kapsamasının denekler tarafından önemsenme oranının çok yüksek olduğunu, sınavların zamana yayılarak yapılmasını, sonuçların mümkün olduğunca kısa sürede ve dönütleriyle birlikte öğrencilere ulaştırılma talebi olduğunu, öğrencilerin sınav kağıtlarının öğretmenlerle birlikte değerlendirilmesine yönelik istekleri olduğunu, bunun objektifliği arttıracığını düşündüklerini söylemiştir.

Şimşek'in (2000) fen bilimlerinde değerlendirme ile ilgili yapmış olduğu bir çalışmada; ülkemizde fen bilimleri ile ilgili yapılan ölçme ve değerlendirmenin,

genellikle öğrencilerin ezberleme yetenekleri ile elde ettikleri bilgi seviyelerini ortaya çıkarmak amacıyla yapıldığı; ölçme ve değerlendirmenin öğrenciye ve öğretmene sağladığı dönütlerin göz ardı edildiği belirtilmiştir.

Lawrenz, 2001 yılında yapmış olduğu bir çalışmada, farklı ölçme-değerlendirme formlarını kullanan farklı alt grup öğrenciler içinde fen başarısı sonuçlarını incelemeyi amaçlamıştır. Bunun için, ABD’de 13 liseden ulusal temsili örnek olarak fen bilimleri dersi okuyan yaklaşık 3500 adet 9. sınıf öğrencisi, ulusal fen eğitimi standartlarındaki seviyelerini ölçmek için tasarlanmış fen ölçme-değerlendirme serilerini tamamlamıştır. Ölçme-değerlendirmeler, çoktan seçmeli testler, yazılı şekilde fikirleri beyan eden testler, laboratuvar yetenekleri ile ilgili testler ve araştırma testlerini içermektedir. Lawrenz, farklı ölçme-değerlendirme formatlarındaki öğrenci sonuçlarını değerlendirdiğinde; daha yüksek başarılı öğrencilerin daha düşük seviyeli öğrencilere oranla daha fazla korelasyona sahip olduğu sonucuna varmıştır. Farklı kültürel gruplar için desenlerin, ölçme-değerlendirme formatlarında değişiklik gösterdiğini belirtmiştir. Cinsiyetler arası başarıda hiçbir fark bulunmamıştır. Sonuçlardan yola çıkarak, farklı değerlendirme formatlarının farklı yetenekleri ölçtüğünü ve farklı alt gruplardaki öğrencilerin başarılarının değerlendirme formlarına göre değişiklik gösterdiği fikri desteklenmiştir (Lawrenz, 2001).

2. Amaç

Eğitim sistemimiz içinde değişik düzeylerde ve değişik alanlarda karşılaşılan problemler mevcuttur. Bu problemlerden biri de fen eğitiminde görülen başarısızlık yüzdesidir (Bakaç, 2003). Bu başarısızlığın nedenlerinden biri, değerlendirme tekniklerinin yapısı ve bu sürecin uygulamalarındaki farklılıklar olabilir mi? Bu araştırılması gereken önemli bir sorudur.

Öğrenci hakkında bilgi toplamanın çok değişik yolları olmasına rağmen kolaylığı, nesnelliği ve ispatlanabilir olma özelliğinden ötürü en çok kullanılan metot, yazılı sınav yöntemidir. Bu yöntemin çok kullanılması ve zaman, yazılı sınav yönteminde de kendi içinde çeşitliliğe yol açmıştır. Bu çeşitlilik de kendi içinde gitgide farklılaşma eğilimi göstermektedir. Bu farklı yazılı sınav türleri farklı amaçlarla kullanılabilirler gibi,

aynı amaç için veya amaca göre en uygun olanı seçilmek sûretiyle de kullanılmaktadır (Tekin, 1996).

Öğrencilerin başarı düzeylerinin tespiti amacıyla kullanılan bazı sınav türleri şöyledir: Yazılı yoklama, Sözlü sınavlar, Kısa cevaplı testler, Doğru-Yanlış testleri, Çoktan seçmeli testler, Eşleştirmeli testler.

Yukarda belirtilen sınav türlerinden Kısa cevaplı testler, Doğru-Yanlış testleri, Çoktan seçmeli testler ve Eşleştirmeli testler puanlamada objektiflikleri çok yüksek olduğu için bu sınav türleri literatürde objektif sınavlar olarak da adlandırılmaktadırlar. Ölçme yapılırken ideal durum, öğrencilerin ilgili hedef-davranışını tam olarak, objektif, doğru ve güvenilir bir biçimde ölçebilmektir. Bu yüzden kullanılan ölçme türünün hedef-davranışı ölçmede etkisi olup olmadığı sorusuna cevap bulunmalıdır. Ölçme türü değiştiğinde aynı öğrencinin, aynı hedef-davranışlara göre hazırlanmış bir başka testte aynı başarıyı gösterip gösteremeyeceği önem taşımaktadır.

Bu araştırma, “İtme-momentum konusu için hazırlanan, fizik öğretiminde kullanılan farklı ölçme türlerinin aralarında anlamlı bir ilişki var mıdır?” sorusuna cevap aramaktadır.

3. Yöntem

Bu çalışmada, lise 2. sınıf öğrencilerini temsilen seçilmiş 32 kişilik (7 kız, 25 erkek) bir grup örneklem olarak belirlenmiştir. Örneklem olarak, 2003-2004 Eğitim-Öğretim Yılı'nda Ankara Keçiören Kalaba Lisesi 10 Fen C sınıfında öğrenim görmekte olan 32 öğrenci seçilmiştir.

Verilerin elde edilmesinde, fizik eğitiminde sık kullanılan ölçme türlerinden objektif ölçme türleri olarak bilinen, “Kısa Cevaplı Test”, “Çoktan Seçmeli Test” ve “Doğru-Yanlış Tipi Test” türleri kullanılmıştır.

Araştırmada kullanılmak üzere, “Kısa Cevaplı Test”, “Çoktan Seçmeli Test” ve “Doğru-Yanlış Tipi Test” türündeki üç ölçme aracı geliştirmek amacıyla, ilk aşama olarak konuyla ilgili hedef-davranışlar tespit edilmiştir. Söz konusu hedef ve davranışlar hazırlanırken Milli Eğitim Bakanlığı, Eğitimi Araştırma ve Geliştirme Dairesi

Başkanlığı'nın yayınlamış olduğu "Ortaöğretim Kurumları Fizik Dersi Taslak Öğretim Programı"ndan yararlanılmıştır (Milli Eğitim Bakanlığı Yayınları, 1998). İkinci aşama olarak, bu hedef-davranışlara ait belirtke tablosu çizilmiştir. Farklı türlerdeki her üç test de aynı hedef davranışlara yönelik, aynı sayıda soru maddesinden oluşmaktadır. Her bir test 21'er adet test maddesinden oluşmaktadır.

"Kısa Cevaplı Test" için hesaplanan Cronbach-Alpha güvenilirlik katsayısı, 0,65 olarak bulunmuştur. "Çoktan Seçmeli Test" için hesaplanan KR-20 güvenilirlik katsayısı, 0,65 ve "Doğru-Yanlış Tipi Test" için hesaplanan KR-20 güvenilirlik katsayısı 0,53 olarak bulunmuştur. Kehoe'ye göre (1995), 10-15 civarı maddeden oluşan testler için 0,50 civarında olan KR-20 güvenilirlik katsayısının yeterli olacağını, test 50 maddeden fazla ise, bu katsayının 0,80 olması gerektiğini belirtmiştir. Bu durumda hazırlanan testlerin güvenilirliği sağlanmıştır denilebilir.

Geliştirilen üç farklı türdeki ölçme aracı için hedef-davranışlar aynıdır. Yani farklı tipte olmasına rağmen test maddeleri, aynı davranışı ölçmeye yönelik hazırlanmıştır. Örnek test maddeleri Tablo 1'de verilmiştir.

Geliştirilen farklı türdeki üç ölçme aracı, aynı gruba birer gün ara ile uygulanmak suretiyle veriler elde edilmiştir. Testlerin birer gün ara ile uygulanması ile, testlerin uygulanması arasında geçen sürede, konu ile ilgili ek öğrenmelerin en aza indirgenmesi amaçlanmıştır. Her öğrenciye bir "Öğrenci No." verilmek suretiyle, öğrenciler numaralandırılmış, böylelikle işlemler esnasında öğrenciler ayırt edilmiştir. Ayrıca öğrencilerin fizik dersi başarı notları (FDBN), uygun katsayı ile genişletilerek, uygulanan testlerle karşılaştırılmıştır. Okullarında öğrencilerin FDBN verilirken "Yazılı Yoklama Tipi Test" kullanılmıştır.

Puanlama, ÇST ve DYT maddelerinde, doğru cevaplar için "2", yanlış cevap ve boş bırakılan cevaplar için ise "0" (Kod 0) verilerek; KCT maddelerinde ise, fiziksel tanımlarla uyum gösteren cevaplar için "2" (Kod 2), fiziksel tanımlarla tam örtüşme görülmemekle birlikte uyumlu unsurlar içeren cevaplar için "1" (Kod 1) ve boş

birakılan cevaplar da dahil olmak üzere, fiziksel tanımlarla uyumsuz cevaplar için “0” (Kod 0) verilerek yapılmıştır.

Tablo 1. Uygulanan testlerde kullanılan sorulara örnekler

No	Test Tipi	Test Maddesi
1	Çoktan Seçmeli Test (ÇST)	Soru: İtme kavramı için; I. Bir cismin üzerinde, büyük bir kuvvet, küçük bir kuvvetten daha büyük bir itme oluşturur. II. İki cisme aynı sürede etkiyen kuvvetler eşitse, itmeler de eşittir. III. İtme bir çeşit kuvvettir. yargularından hangisi yada hangileri kesinlikle doğrudur? A) Yalnız I B) Yalnız II C) I-II D) I-III E) I-II-III
	Kısa Cevaplı Test (KCT)	Soru: m ve $2m$ kütleli iki cisim çarpıştıklarında birbirlerine uyguladıkları itmelerin şiddetlerinin oranı I_1/I_2 ne olur?
	Doğru-Yanlış Tipi Test (DYT)	Soru: () Bir cismin üzerinde, büyük bir kuvvet, küçük bir kuvvetten her zaman daha büyük bir itme oluşturur.
2	Çoktan Seçmeli Test (ÇST)	Soru: Durgun bir bomba patladığında, üç parçaya ayrılıyor. Bu olayda: I. Parçalardan biri kuzeye, ikisi güneye fırlar. II. Parçalardan biri durur, diğer ikisi fırlar. III. Parçalardan ikisi durur, biri fırlar. IV. Parçaların ikisi kuzeye, biri doğuya fırlar. V. Parçaların biri güneydoğuya, biri kuzeye, biri batıya fırlar. durumlarından hangisi yada hangileri gerçekleşemez? A) Yalnız III B) III-IV C) I-II-V D) I-II E) II-III-IV
	Kısa Cevaplı Test (KCT)	Soru: Durgun bir bomba patlayarak üç eşit parçaya ayrılıyor. Parçalardan biri kuzeye diğeri doğuya fırlarsa, üçüncü parçanın yönü ne olur?
	Doğru-Yanlış Tipi Test (DYT)	Soru: () Doğrultuları arasında 60° açı bulunana eşit şiddette momentuma sahip eşit kütleli iki cisim çarpışıp kenetlenirlerse, ikisi de önceki doğrultularından 30° saparlar.

* İsteyen araştırmacılara test maddelerinin tamamı verilecektir.

4. Bulgular ve Yorumlar

Araştırma çerçevesinde, öğrencilere uygulanan, KCT, ÇST ve DYT'lerine verilen cevaplar tek tek incelenmiştir. Testleri oluşturan 21 soru maddesinin geneli ile ilgili istatistiksel veriler Tablo 2'de verilmiştir.

Tablo 2. Uygulanan testler ile ilgili istatistiksel sonuçlar

İstatistiksel Nicelikler	Kısa Cevaplı Test (KCT)			Çoktan Seçmeli Test (ÇST)		Doğru-Yanlış Testi (DYT)	
	Kod 2	Kod 1	Kod 0	Kod 2	Kod 0	Kod 2	Kod 0
Topl. Kod Sayısı	139	89	444	196	476	383	289
Ort. Kod Sayısı.	4,34	2,78	13,88	6,13	14,88	11,97	9,03
Standart Sapma	2,50	1,68	3,05	3,00	3,00	2,92	2,92
	20,69	13,24	66,07	29,17	70,83	56,99	43,01
Kodların Yüzdeleri (%)							

Testler için kodların dağılımına bakıldığında görülmektedir ki, “Kod 0”ın temsil ettiği, “boş bırakılan veya fiziksel tanımlarla uyumsuz cevaplar” ilk iki test türünde yoğunluk göstermektedir. Üçüncü test türünde ise “Kod 2”, “Kod 0”a göre daha yoğundur.

Öğrencilerin her biri için kodlar puana dönüştürülmüş ve öğrencilerin her bir testten aldıkları puanlar Tablo 3’de verilmiştir.

Tablo 3. Öğrencilerin ölçme araçlarından ve fizik dersinden aldıkları başarı puanları.

Ö. No	KCT	ÇST	DYT	GFBN	Ö. No	KCT	ÇST	DYT	GFBN
1	6	4	22	16,8	17	9	4	24	16,8
2	7	10	20	8,4	18	6	12	22	25,2
3	8	2	22	25,2	19	7	12	20	16,8
4	6	12	20	16,8	20	10	8	28	8,4
5	6	4	22	16,8	21	19	20	24	8,4
6	8	6	14	16,8	22	14	22	40	25,2
7	8	12	22	25,2	23	3	4	28	8,4
8	13	12	26	16,8	24	13	14	16	33,6
9	16	16	28	25,2	25	4	16	18	8,4
10	16	14	24	8,4	26	13	6	16	16,8
11	8	6	24	8,4	27	11	14	14	8,4
12	11	18	28	8,4	28	14	20	22	8,4
13	17	18	20	8,4	29	18	18	28	16,8
14	21	20	30	25,2	30	13	14	28	16,8
15	25	18	26	33,6	31	18	2	24	8,4
16	12	16	28	8,4	32	7	18	38	8,4
					Ort.	11,47	12,25	23,94	15,75
					St. Sp.	5,31	3,00	2,92	7,91

Öğrencilerin testlerden almış oldukları puanlar ve fizik dersinden aldıkları başarı notları ile uygulanan testler arasında bir ilişki olup olmadığı incelenmiştir. Öğrencilerin uygulanan ölçme araçlarından elde edilen veriler ve fizik dersi başarı notları ikiye ikiye karşılaştırılmış, aralarında ilişki olup olmadığı, Pearson Momentler Çarpımı Korelasyon Katsayısı kullanılarak hesaplanmıştır. Ölçme araçlarının ikili korelasyonları ile ilgili istatistiksel veriler Tablo 4’te verilmiştir.

Tablo 4. Öğrencilerin farklı türdeki testlerden aldıkları puanlar arasındaki korelasyonlara ait istatistikler.

Ölçme Araçları		KCT	ÇST	DYT	GFBN
Kısa Cevaplı Test	r	1	0,508	0,221	0,244
	P	0	0,003*	0,224	0,178
Çoktan Seçmeli Test	r	0,508	1	0,379	0,074
	P	0,003*	0	0,032*	0,687
Doğru-Yanlış Testi	r	0,221	0,379	1	0,022
	P	0,224	0,032*	0	0,905
G. Fizik Dersi Başarı Notları	r	0,244	0,074	0,022	1
	P	0,178	0,687	0,905	0

* P < 0,05

KCT ve ÇST için hesaplanan korelasyon katsayısı değeri, $r = 0,51$ olarak bulunmuştur. Anlamlılık değeri olarak $\alpha = 0,05$ alınmıştır. Korelasyon katsayısı ise, $P = 0,003$ olasılık değerini vermektedir. Olasılık değeri, anlamlılık değerinden küçük olduğu için ($P < \alpha$), bu iki test arasında anlamlı bir ilişki var denilmektedir.

KCT ve DYT için hesaplanan korelasyon katsayısı değeri, $r = 0,22$ olarak bulunmuştur. Korelasyon katsayısı ise, $P = 0,224$ olasılık değerini vermektedir. Olasılık değeri, anlamlılık değerinden büyük olduğu için ($P > \alpha$), bu iki test arasında anlamlı bir ilişki yoktur.

ÇST ve DYT için hesaplanan korelasyon katsayısı değeri, $r = 0,38$ olarak bulunmuştur. Korelasyon katsayısı ise, $P = 0,032$ olasılık değerini vermektedir. Olasılık değeri, anlamlılık değerinden küçük olduğu için ($P < \alpha$), bu iki test arasında anlamlı bir ilişki var denilmektedir.

KCT, ÇST ve DYT türlerinden oluşan ölçme araçları ile öğrencilerin fizik dersi başarı notları (GFBN) arasındaki ilişkiyi incelemek için hesaplanan korelasyon katsayısı değerlerinin verildiği Tablo 4'e bakıldığında, her üç testin de olasılık değerlerinin anlamlılık değeri olan, $\alpha = 0,05$ 'den daha büyük olduğu görülmektedir. Bu da

uygulanen testlerin hiçbirisi ile öğrencilerin fizik dersinden almış oldukları başarı puanları arasında anlamlı bir ilişki olmadığını gösterir.

5. Sonuç ve Öneriler

Fizik eğitiminde kullanılan yazılı ölçme türlerinden bazılarını, itme ve momentum konusu için karşılaştırmayı amaçlayan bu araştırmanın verilerinden aşağıdaki sonuçlar elde edilmiştir.

- İtme ve momentum konusu için uygulanan KCT ile ÇST arasında anlamlı bir ilişki bulunmaktadır.
- İtme ve momentum konusu için uygulanan KCT ile DYT arasında anlamlı bir ilişki bulunmamaktadır.
- İtme ve momentum konusu için uygulanan ÇST ile DYT arasında anlamlı bir ilişki bulunmaktadır.
- İtme ve momentum konusu için uygulanan KCT, ÇST ve DYT ile öğrencilerin okullarında almış oldukları fizik dersi başarı notları arasında anlamlı bir ilişki bulunmamaktadır.

Sonuç olarak, aynı hedef-davranışları ölçmeyi amaçlayan farklı üç tipteki ölçme aracı kısmen farklı sonuçlar vermiştir. Bu da, itme ve momentum konusu için ölçmenin, aynı hedef-davranışları ölçmeyi amaçlarsa bile, ölçme türü değiştiğinde farklı sonuçlar verebileceğini göstermektedir. Burada, test türlerinin puanlama şekilleri, şans faktörü ve öğrencilerin test türleri üzerindeki bireysel yeteneklerinin ölçüme etkisinin olduğu söylenebilir. Örneğin, yazılı ifade yeteneği yeterli olmayan bir öğrenci, düşüncelerini kısa cevaplı testte ifade etmekte zorlanabilmekte, aynı şekilde şans faktörü daha yüksek olan çoktan seçmeli ve doğru-yanlış tipi testlerde öğrenci şansına ve sezgilerine güvenerek doğru cevabı bulabilmektedir.

Ayrıca, öğrencilerin fizik dersinde almış oldukları başarı notları ile itme ve momentum konulu testlerin sonuçları ile uyumsuz oluşunun bir diğer nedeni, öğrencilerin ÖSS'ye hazırlanıyor olmaları ve bu sebeple sıklıkla çoktan seçmeli tipte sorularla ve lisede ise

yazılı yoklama tipi testlerle değerlendiriliyor olmaları olabilir. Çünkü öğrenciler, ÖSS ve lise yaşamlarında Doğru-Yanlış Testleri ve Kısa Cevaplı Testleri ile genelde karşılaşmamaktadırlar. Bunun sonucu olarak farklı tiplerde sorulan soruları algılamakta zorlanabilmekte, hatta yanılıya düşebilmektedirler.

Ölçmenin daha sağlıklı yapılabilmesi için sıklığı artırılmalı ve her konu sonunda gerçekleştirilmelidir. Ölçme yaparken farklı ölçme türlerinin farklı sonuçlar verebileceği de dikkate alınmalıdır. Şans faktörünün azaltılması ve öğrencilerin test türleri üzerindeki bireysel yeteneklerinin ölçmeye etkisinin asgari olması için, ölçme araçları tek bir ölçme türü yerine, her ölçme türünü içerecek şekilde hazırlanan testlerle yapılmalıdır.

Kaynaklar

- Bakaç, M. (2003) Fen Bilgisi Öğretiminde Ölçme-Değerlendirme Üzerine Bir Çalışma. Milli Eğitim Dergisi, Kış 2003, 157.
- Banta, T. (1997). Moving assessment forward: Enabling conditions and stumbling blocks. *New Directions In Higher Education*. 100: p.79-92.
- Black, P. Testing: friends or foe? Theory and practice of assessment and testing. London: Falmer Press. 1998.
- Espinoza, F. (2004) Enhancing mechanics learning through cognitively appropriate instruction. *Physics Education*, 39 (2)
- Gipps, C. V. Beyond testing. London: The Farmer Press. 1994.
- Güneş, P., İnceç, Ş., Taşar, M. F. (2002) Momentum ve İmpuls Kavramlarını Anlama – I: Öğretmen Adaylarının Açık Uçlu Sorularla Momentum ve İmpulsu Nasıl Tanımladıklarının Belirlenmesi. G.Ü. Gazi Eğitim Fakültesi Dergisi, Cilt 22, Sayı 3, s.121-138
- Harlen, W. & et al (1991). Assessment and the improvement of education. *The Curriculum Journal*, 3(3), p.215-230.
- İnceç Ş., Güneş P., Taşar M. (2002) Öğrencilerin impulsu tanımlamaları ve bir probleme uygulamaları. V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi <<http://www.fedu.metu.edu.tr/ufbmek-5/>>
- Kehoe, J. (1995). ERIC/AE Digest Series EDO-TM-95-11

- Lawrenz, F., Huffman, D., Welch, W. (2001) The Science Achievement of Various Subgroups on Alternative Assessment Formats. *Science Education* 85:279–290, 2001.
- Milli Eğitim Bakanlığı, (1998). Ortaöğretim Kurumları Fizik Dersi Taslak Öğretim Programı, Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı, Ankara.
- Şimşek, S. (2000). Fen Bilimlerinde Değerlendirmenin Önemi. *Milli Eğitim Dergisi*, 148, s.30-32.
- Tekin, H., Eğitimde Ölçme ve Değerlendirme Yargı Yay. 9. Baskı, Ankara-Mayıs 1996.
- Torrance, H. & Pryor, J. (1995). Investigating teacher assessment at key stage 1: methodological problems and emerging issues. *Assessment in Education*, 2 (3), p.305-320.