

Ölçme ve Değerlendirme Dersi Kapsamı ve Gereklilikleri Hakkındaki Uzman ve Öğretmen Görüşleri: Pilot Çalışma

Derya ÇOBANOĞLU AKTAN*

Zafer ÇEPNİ†

Özet

Bu çalışmanın ilk amacı, öğretmenlerin ve ölçme değerlendirme uzmanlarının üniversitelerde verilen ölçme ve değerlendirme dersi konularının öğretmenler için gerekliliği konusundaki görüşlerini incelemek ve karşılaştırmaktır. İkinci amaç ise konuların ne ölçüde vurgulandığına dair öğretmen görüşlerini almaktır. 182 öğretmen yanında 47 ölçme ve değerlendirme uzmanından anketler yoluyla görüş alınmıştır. Öğretmenlerin en yüksek oranla gerekli bulduğu konular test hazırlama ve soru türleri ile standartlar olmuştur. Öğretmenler güvenilirlik hesaplama ve geçerlik türleri ile istatistik ve madde analizi konularını diğer konular kadar gerekli bulmamaktadır. Uzmanlar genel olarak bütün konuları gerekli görmektedirler. Ayrıca bazı konuları uzmanlar, öğretmenlerden daha yüksek oranda gerekli bulmuşlardır. Öğretmenlere göre, ölçme ve değerlendirme derslerinde en çok vurgulanan konular test hazırlama ve soru türleri ile ölçmenin temel kavramlarıdır. Buna karşın öğretmenlerin en az vurgulandığını düşündüğü konular ise performans değerlendirme, istatistik ve madde analizi, güvenilirlik hesaplama yöntemleri ve geçerlik türleridir. Uzmanlar tarafından çoğu ölçme ve değerlendirme konuları gerekli görülmesine rağmen, öğretmenlerin bu konuların yeterince vurgulanmadığını düşündüğü görülmüştür. Bazı iyileştirmeler için önerilerde bulunulmuştur.

Anahtar Kelimeler: Öğretmen yetiştirme, ölçme ve değerlendirme eğitimi, performans değerlendirme

Abstract

The first aim of this study is to investigate and compare the opinions of the teachers and measurement and evaluation specialists, regarding the necessity of the measurement and evaluation course topics for teachers. The second aim is to provide the opinions of teachers regarding the degree of the emphasis on these topics. 182 teachers and 47 measurement and evaluation specialists provided their opinions via surveys. The teachers find test construction, item formats and the standards topics the most necessary. Also the teachers consider statistics, item analysis, reliability estimation methods and validity types topics less necessary than other topics. Generally specialists consider almost all topics necessary. Also specialists find some topics more necessary than the teachers. According to teachers, the most emphasized topics are test construction, item formats and the fundamentals of measurement. The teachers think that the least emphasized topics are performance assessment, statistics, item analysis, reliability estimation methods and validity types. Specialists find almost all topics necessary. Although the specialists think that almost all topics are necessary, teachers believe that some of these topics are not emphasized enough in their measurement and evaluation courses. Some recommendations are provided for improvements.

Keywords: Teacher training, measurement and evaluation instruction, performance assessment

* Arş. Gör. Dr., Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, dcaktan@hacettepe.edu.tr

† Arş. Gör., Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, cepni@hacettepe.edu.tr

Giriş

Bu çalışmanın ilk amacı, öğretmenlerin ve ölçme değerlendirme uzmanlarının üniversitelerde verilen ölçme ve değerlendirme dersi konularının öğretmenler için gerekliliği konusundaki görüşlerini incelemek ve karşılaştırmaktır. İkinci amaç ise konuların ne ölçüde vurgulandığına dair öğretmen görüşlerini almaktır.

Eğitim fakültelerinde verilen ölçme ve değerlendirme dersinin neleri kapsamı gerektiği konusunda Amerika Birleşik Devletleri'nde 1960'lardan bu yana çeşitli çalışmalar yapılmış ve görüşler sunulmuştur (Gullickson ve Hopkins, 1987). Shepard'a (2006) göre yirminci yüzyılın büyük bir kısmında öğretmen adaylarına bilimsel ölçme prensiplerine göre kendi testlerini nasıl hazırlayacakları ve standart testleri nasıl kullanacakları öğretilmiştir. Ders kitapları ise günümüzde de hala çoğu ders kitabının temel kalıbını oluşturan geçerlik, güvenirlik, test geliştirme, değişik madde türleri, madde analizi ve test sonuçlarının istatistiksel analizini içermektedir. Yirminci yüzyılın sonlarında ise nicel sonuçlar almaya dayalı bu sınıf içi ölçme ve değerlendirme yaklaşımı, öğrencilerin neleri anladığını ön plana çıkaran ve de ölçme ve değerlendirmeyi öğrenme sürecinin biçimlendirici ve bütünleşik bir parçası olarak gören bir yaklaşıma dönüşmeye başlamıştır.

1980'lerden sonra hesap verebilirlik amacıyla ABD'de yaygın olarak kullanılan standart testler, belirli madde türlerini kullanmaktaydı. Bu durum ise eğitimin kalitesini düşürmekteydi. Hesap verebilirlik amacıyla kullanılan testlerin böyle istenmedik sonuçlar doğurması, bahsi geçen yaklaşım değişimini tetiklemiştir (Black ve William, 1998; Gipps, 1999; Shepard, 2000; Shepard, 2006). Değişime yol açan diğer sebepler ise (1) ölçme ve alan uzmanlarının ders program hedeflerine daha uygun ölçme ve değerlendirme yöntemlerini geliştirmeleri ve (2) biliş ve motivasyon alanlarında yapılan

arařtırmaların deęiřimin gereklilięi konusunda sunduęu delil ve teorilerdir (U.S. Congress, Office of Technology Assessment, 1992; Shepard, 2006; Pellegrino, Chudowsky ve Glaser, 2001).

Yaklařımdaki bu donüřüm çerçevesinde, Airasian (1991) eęitimde ölçme kitaplarının ve derslerinin çoęunun öęretmenlerin sınıflarındaki uygulamalarına yeterince ışık tutmadıęını ve bu durumun deęiřmesi gerektięini belirtmiřtir. Gereki deęiřimin, ölçmenin temel prensipleri ile öęretmenlerin gereksinimlerinin bir araya getirilmesiyle mümkün olacaęını ifade etmiřtir. Airasian'a (1991) göre, ölçme deęerlendirme derslerinde ve kitaplarında var olan formal ölçme ve deęerlendirme yöntemlerinin yanı sıra öęretmenler için gerekli olan informal ölçme ve deęerlendirme yöntemleri üzerinde de durulmalıdır.

Bu konudaki öęretmenlerin ve uzmanların görüşlerini Gullickson (1987) incelemiřtir. Gullickson'un bu çalıřmasında öęretmenlere ölçme ve deęerlendirme derslerinde nelerin öęretilmesi gerektięi sorulmuř ve dersi veren öęretim elemanlarına da derslerde hangi konuları vurguladıklarını sorularak karřılařtırma yapılmıřtır. Bu iki grubun görüşleri, istatistik, test dıřı deęerlendirme teknikleri, biçimlendirici deęerlendirme ve düzey belirleme amaçlı deęerlendirme konularında istatistiksel olarak anlamlı farklılıklar göstermiřtir. Çalıřma sonucunda, ölçme ve deęerlendirme eęitiminin öęretmen ihtiyaçları doęrultusunda irdelenmesi gerektięi ortaya konmuřtur.

Bu deęiřimin yansımaları ölçme ve deęerlendirme ders kitaplarında da görülmektedir (Stiggins, 2008; Airasian ve Russel, 2008). Yurt dıřında yařanan bu geliřmelere paralel olarak, yeni öęretim programlarının da etkisi ile performans deęerlendirme konusu, "alternatif" veya "otantik" deęerlendirme yöntemleri

başlıklarıyla ülkemizde de ön plana çıkmaya başlamıştır (Kaptan, 1998; Korkmaz ve Kaptan, 2000; Öncü, 2009).

Ankara’da yapılan bir çalışmada, öğretmenlerin yeni ders programlarıyla daha çok vurgulanan ve uygulanması öngörülen değerlendirme yöntemleri konusunda kendilerini eksik hissettikleri, bu yöntemleri kullanırken zorluklar yaşadıkları ve bu yöntemleri geleneksel olarak adlandırılan yöntemlerden daha seyrek kullandıkları görülmüştür (Gelbal ve Kelecioğlu, 2007). Dolayısıyla, öğretmen adaylarına verilen ölçme ve değerlendirme derslerinin ne ölçüde öğretmen ihtiyaçlarını karşıladığının sorgulanması gerekmektedir. Bu gereklilik doğrultusunda, çalışmamız aşağıdaki araştırma sorularına yanıt aramaktadır.

1. Öğretmenlerin üniversitedeki ölçme ve değerlendirme derslerindeki konuların vurgulanma dereceleri hakkındaki görüşleri nelerdir?
2. Uzman ve öğretmenlerin ölçme ve değerlendirme konularının öğretmenlerin meslek hayatlarındaki gereklilikleri hakkındaki görüşleri nelerdir?
3. Ölçme ve değerlendirme uzmanları ile öğretmenlerin, ölçme ve değerlendirme konularının öğretmenlerin meslek hayatları için gereklilikleri hakkındaki görüşleri istatistiksel olarak anlamlı bir farklılık göstermekte midir?
4. Ölçme değerlendirme konularının yeterince vurgulandığını ifade eden öğretmenlerin oranı nedir?

Yöntem

Araştırma, betimsel bir tarama çalışması olarak yürütülmüştür. Anketler ile veri toplanmış, frekanslar, yüzdeler ve ki-kare analizleri işe koşulmuştur.

Katılımcılar

Çalışmada ilk ve orta dereceli devlet okulları ve özel okullarda görev yapan öğretmenler yanında ölçme değerlendirme alanında en az lisans düzeyinde diploma sahibi uzmanlardan görüş alınmıştır. Katılımcılar yansız örnekleme ile belirlenmemiş olup, gönüllülük esasında çalışmada yer almışlardır. Çalışmaya katılan 182 öğretmenin yarısı internet üzerindeki anketi cevaplamış, diğer yarısı ise kâğıt formatındaki anketi cevaplayarak görüş bildirmiştir. Öğretmen katılımcıların %56'sı kadındır. Kâğıt formatında veri toplama sürecine katılan öğretmenlerin tümü Ankara'da görev yapmaktadır. İnternet üzerinden veri sağlayan gönüllü katılımcılar ise Türkiye'nin 36 farklı ilinde görev yapmaktadır. Katılımcılar 34 farklı branştan olup, 84'ü devlete bağlı bir ilköğretim okulunda, 81'i ise devlete bağlı bir lisede çalışmaktadır. Uzman katılımcılar, görüşlerini internet üzerindeki anketi cevaplayarak bildirmiştir. Ayrıca Hacettepe Üniversitesi'nde görev yapmakta olan katılımcılar kâğıt formatındaki anketi cevaplamışlardır. Uzman katılımcılar 20 erkek ve 27 kadından oluşmaktadır. Uzman katılımcıların 16'sı doktora, 19'u yüksek lisans derecesine sahip olup 22'si üniversitelerde 13'ü ise ilk ve orta dereceli özel okullarda görev yapmaktadır. Hem uzman hem de öğretmen katılımcıların geniş bir evreni temsil edecek biçimde seçilmemiş olması çalışmanın bir sınırlılığı olarak görülebilir.

Veri toplama araçları

Öğretmen ve uzman anketleri, ölçme ve değerlendirme dersinde işlendiğini düşündüğümüz ve ders kitaplarında yer alan 55 konu adını içeren bir listeden oluşmaktadır. Bu liste yedi gruptan oluşmaktadır. Bu gruplar: ölçmenin temel kavramları (14 madde), test hazırlama ve soru türleri (8 madde), güvenilirlik hesaplama

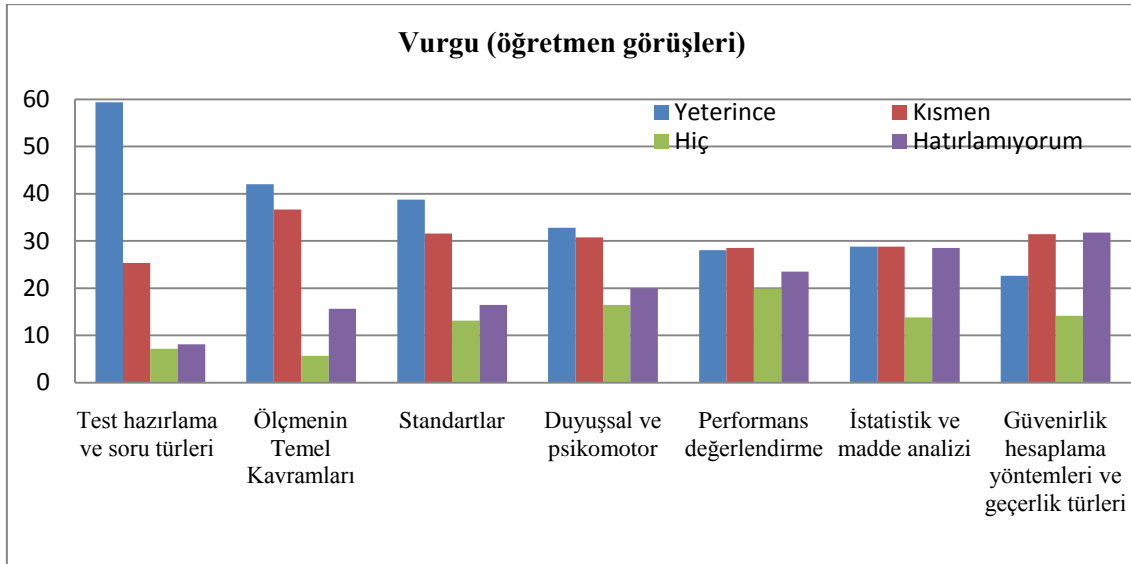
yöntemleri ve geçerlik türleri (10 madde), istatistik ve madde analizi (9 madde), standartlar (7 madde), performans değerlendirme (5 madde), duyuşsal ve psikomotor özelliklerin ölçülmesidir (2 madde). Standartlar grubundaki maddeler, Amerikan Ulusal Eğitimde Ölçme Kurulu (NCME), Amerikan Öğretmenler Federasyonu ve Amerikan Ulusal Eğitim Birlięi'nin öğretmenler için belirledięi ölçme ve değerlendirme alanındaki öğretmen yeterlik standartlarıdır (AFT, NCME ve NEA, 1990). Anketin içerięi Hacettepe Üniversitesi'nde görev yapan üç uzman tarafından gözden geçirilmiştir. Öğretmen anketinde, öğretmenlere üniversitede aldıkları ölçme ve değerlendirme dersinde konuların ne derece vurgulandığı ve bu konuların meslek hayatları için gereklilięi hakkındaki görüşleri sorulmuştur. Uzman anketinde ise, uzmanlara konuların öğretmenlerin meslek hayatları için gereklilięi hakkındaki görüşleri sorulmuştur. Öğretmenler ve uzmanlar konuların gereklilięi hakkındaki görüşlerini “evet”, “hayır” veya “fikrim yok” seçeneklerinden birini seçerek bildirmişlerdir. Öğretmenler konuların vurgulanmasına dair görüşlerini “yeterince”, “kısmen”, “hiç” veya “hatırlamıyorum” seçeneklerinden birini seçerek bildirmişlerdir. Veri toplama araçlarımız[‡] bir toplam puan verecek şekilde tasarlanmış ve bir yapıyı ölçmek üzere geliştirilmiş bir ölçek olmadığından veri toplama araçlarının geçerlięi uzman kanılarına dayanarak sağlanmışır (Gelbal ve Kelecioęlu, 2007).

Bulgular ve yorumlar

Birinci araştırma sorusuyla ilgili olarak anketteki konuların üniversitedeki ölçme ve değerlendirme derslerinde ne derecede vurgulandığı hakkındaki öğretmen görüşleri, analiz ve yorumlama kolaylığı sebebiyle, konu grubu bazında ortalama yüzdeler

[‡] Çalışmada kullanılan anketler <http://yunus.hacettepe.edu.tr/~dcaktan/oanket.pdf> ve <http://yunus.hacettepe.edu.tr/~dcaktan/uzanket.pdf> adreslerinden incelenebilir.

hesaplanarak incelenmiştir. Şekil 1’de konu gruplarının ortalama yüzdeleri görülmektedir. Öğretmenlere göre, ölçme ve değerlendirme derslerinde en çok vurgulanan konular test hazırlama ve soru türleri ile ölçmenin temel kavramlarıdır. Öğretmenlerin yaklaşık yarısı bu iki konunun derslerde yeterince vurgulandığını ifade etmiştir. Buna karşın öğretmenlerin en az vurgulandığını düşündüğü konular ise performans değerlendirme, istatistik ve madde analizi, güvenilirlik hesaplama yöntemleri ve geçerlik türleridir. Ankete katılan öğretmenlerin yaklaşık üçte biri bu üç konunun derste bahsi geçtiğini hatırlamadıklarını ifade etmiştir. Performans değerlendirme konusu ile ilgili olarak katılımcı öğretmenlerin %20’si konunun hiç vurgulanmadığını, % 24’ü ise bu konunun derste işlenip işlenmediğini hatırlamadığını ifade etmiştir. Diğer bir deyişle katılımcıların %44’ü ölçme değerlendirme dersinde performans değerlendirme konusunda bilgi edinmemiştir.

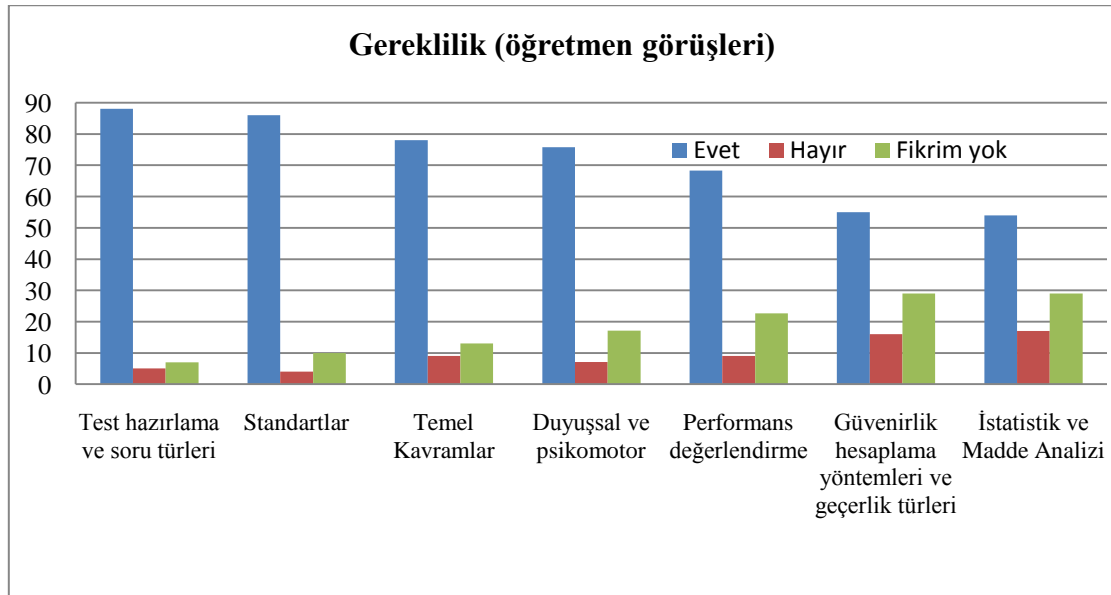


Şekil 1. Öğretmenlerin ölçme ve değerlendirme konularının derslerde vurgulanma dereceleri hakkındaki görüşlerinin ortalama yüzdeleri.

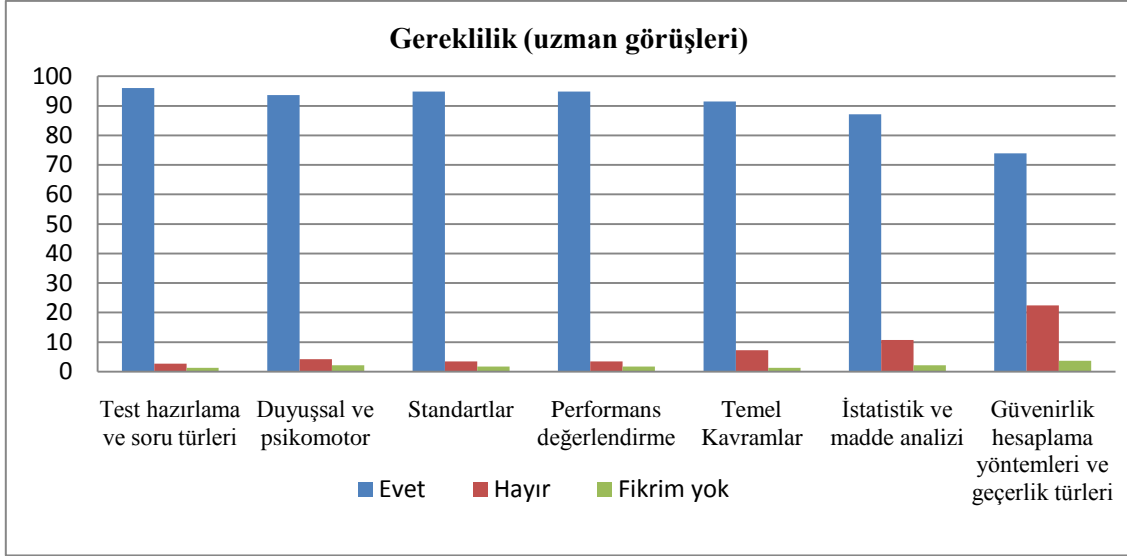
İkinci araştırma sorusu için uzman ve öğretmenlerin ölçme ve değerlendirme konularının öğretmenlerin meslek hayatlarındaki gereklilikleri hakkındaki görüşleri

benzer biçimde Şekil 2’ve Şekil3’te özetlenmiştir. Öğretmenlerin en yüksek oranla gerekli bulduğu konular test hazırlama ve soru türleri (%88) ile standartlar (%86) olmuştur. Öte yandan öğretmenler güvenilirlik hesaplama ve geçerlik türleri(%55) ile istatistik ve madde analizi (%54) konularını diğer konular kadar gerekli bulmamaktadır.

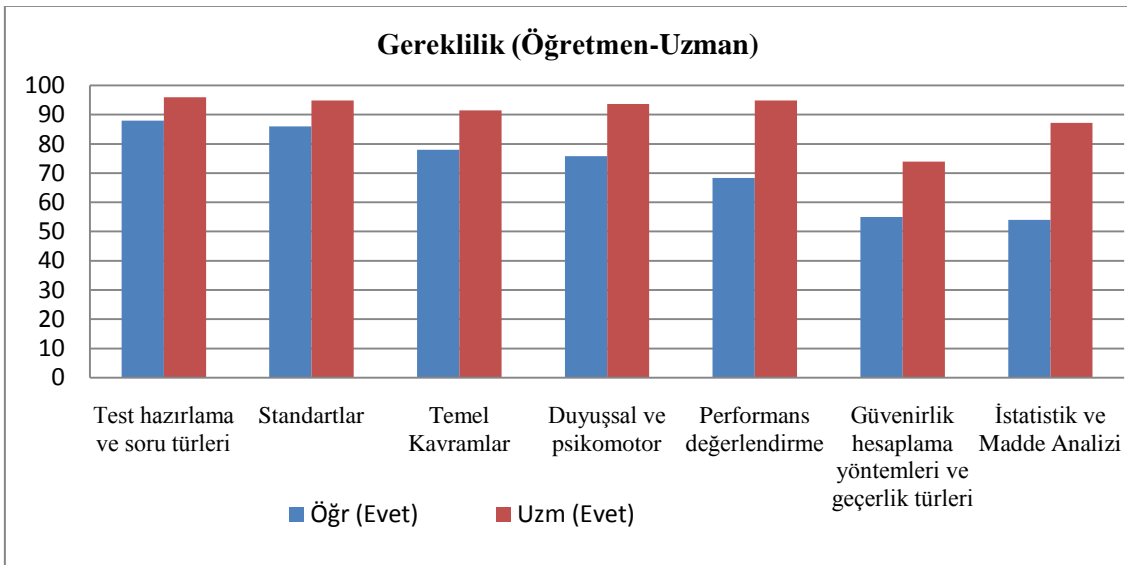
Uzmanlar genel olarak bütün konuları gerekli görmektedirler. Öğretmenler gibi uzmanlar da sıralamalarında güvenilirlik hesaplama ve geçerlik türlerini sona yerleştirmişlerdir. Üçüncü araştırma sorusu gereği bu iki meslek grubunun gereklilikler hakkındaki görüşleri görsel ve istatistiksel olarak da kıyaslanmıştır. Görsel kıyaslama Şekil 4’te sunulmaktadır. Gerekli görme oranlarındaki farkın en büyük olduğu konular; performans değerlendirme, duyuşsal ve psikomotor becerilerin ölçülmesi, güvenilirlik hesaplama ve geçerlik türleri ile istatistik ve madde analizidir. Bu dört konuyu uzmanlar, öğretmenlerden daha yüksek oranda gerekli bulmuşlardır. Bu konulardaki farklar ki-kare analizlerinde istatistiksel olarak da anlamlı bulunmuştur. Ki-kare analizi sonuçları Tablo 1’de verilmiştir.



Şekil 2. Öğretmenlerin ölçme ve değerlendirme konularının meslek hayatlarındaki gereklilikleri hakkındaki görüşlerinin ortalama yüzdeleri.



Şekil 3. Uzmanların ölçme ve değerlendirme konularının öğretmenlerin meslek hayatlarındaki gereklilikleri hakkındaki görüşlerinin ortalama yüzdeleri.

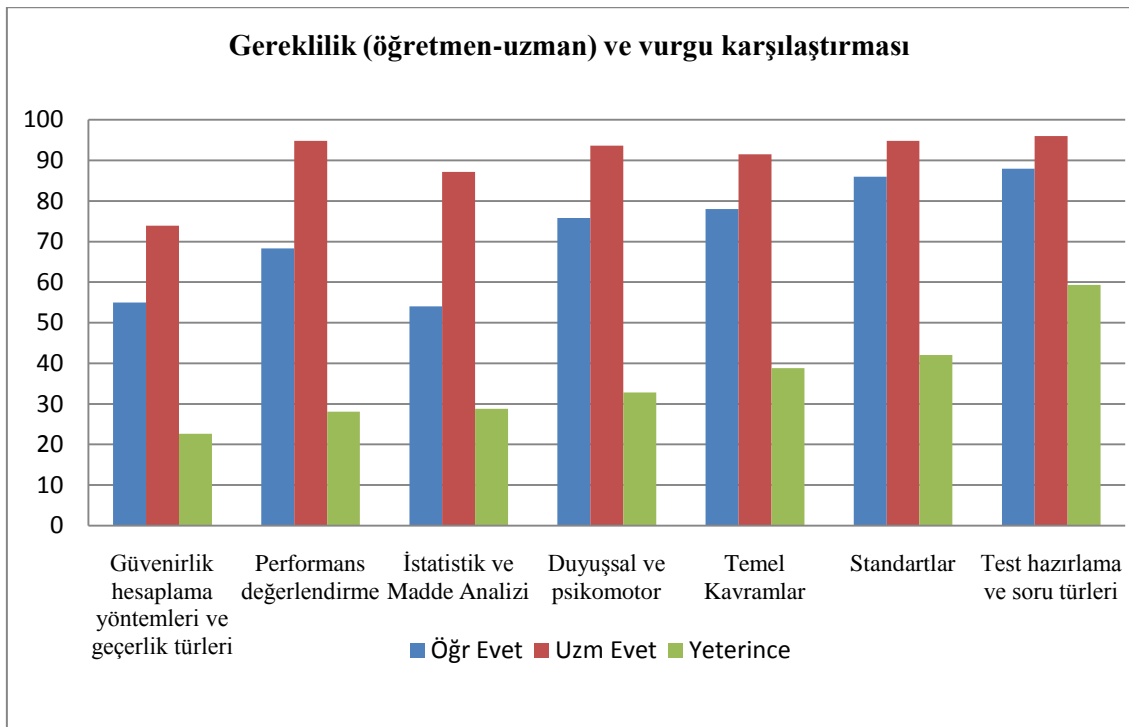


Şekil 4. Öğretmenlerin ve uzmanların ölçme ve değerlendirme konularının öğretmenlerin meslek hayatlarındaki gereklilikleri hakkındaki görüşlerinin kıyaslanması.

	Ki-kare (sd=2)	p
Temel Kavramlar	5,294	0,071
Güvenirlik Hesaplama ve Geçerlik Türleri	12,700	0,002
Test Hazırlama ve Soru Türleri	2,274	0,321
İstatistik ve Madde Analizi	19,061	0,000
Standartlar	4,990	0,082
Performans Değerlendirme	12,723	0,002
Psikomotor ve Duyuşsal Becerilerin Ölçülmesi	8,109	0,017

Tablo 1. Katılımcıların meslekleri ile görüşleri arasındaki ilişkinin sınındığı ki-kare analizi istatistikleri

Dördüncü araştırma sorusuna istinaden ankette yer alan konuları, öğretmenlerin ve uzmanların gerekli görme oranlarının yanı sıra bu konuların **yeterince** vurgulandığını ifade eden öğretmenlerin oranları da Şekil 5'te sunulmuştur. Bu konuların derslerde yeterince vurgulandığını ifade eden öğretmenlerin oranı %23 ile %59 arasında değişmektedir. Bu değerler gerekli görülme oranlarıyla kıyaslandığında düşük kalmaktadır.



Şekil 5. Konuları öğretmenlerin gerekli görme oranları, uzmanların gerekli görme oranları ve bu konuların yeterince vurgulandığını ifade eden öğretmenlerin oranları

Sonuç ve öneriler

Bulgulara göre uzmanlar ve öğretmenler ölçme ve değerlendirme konularının bazılarının gereklilikleri hususunda hemfikir olurken bazı konular hakkında da farklılık sergilemektedirler. Öğretmenler test hazırlama, soru türleri ve standartlar konularını gereksinim bakımından ön plana koyarken güvenirlilik hesaplama ve madde analizi

konularını aynı derecede gerekli bulmamaktadır. Bu bulgu, Gullickson'ın (1987) bulgularıyla örtüşmektedir. Uzmanlar hemen hemen bütün konuları gayet yüksek oranda gerekli bulurken, güvenilirlik hesaplama ve madde analizi konularını daha az oranda gerekli bulmaktadırlar. Öğretmenler ve uzmanlar, performans değerlendirme, duyuşsal ve psikomotor becerilerin ölçülmesi, güvenilirlik hesaplama ve geçerlik türleri, istatistik ve madde analizi konularının gerekliliği hakkında istatistiksel ve pratik olarak anlamlı farklar göstermektedir. Uzmanlar tarafından çoğu ölçme ve değerlendirme konuları gerekli görülmesine rağmen öğretmenlere göre bu konular yeterince vurgulanmamaktadır. Bu farklılıkların bir nedeni konuların teorik ağırlıklı olarak işlenmesi olabilir. Eğer ölçme ve değerlendirme dersleri sınıf içi uygulamalara daha uygun hale getirilip uygulama ağırlıklı işlenirse, öğretmenlerin bu konuları öğrenmeleri ve mesleklerinde etkin şekilde kullanma olasılıkları artacaktır.

Öğretmenlerin performans değerlendirme konusunu gerekli görme oranının (%68) test hazırlama ve madde türlerini gerekli görme oranından (%88) daha düşük olması Shepard'ın (2006) açıkladığı değişimin ülkemiz öğretmenlerinde de başladığı fakat ön plana çıkmadığını göstermektedir. Öğrenci merkezli öğretim programlarının ülkemizde benimsenmesi ve bununla birlikte ölçme ve değerlendirmenin öğrenmeye destek olacak biçimde derse entegre edilmesi performans değerlendirme ve benzeri yöntemlerle mümkün olduğu için, öğretmenlerin bu konuları gerekli görme oranlarının yükselmesi öngörülmektedir. Bunun sağlanması için öğretmenlerin bu konudaki görüşleri ve ihtiyaçları belirlenmelidir (Gelbal ve Kelecioğlu, 2007). Bu bilgiler ışığında, ölçme ve değerlendirme dersleri öğretmenlerin gereksinimlerine cevap verecek biçimde düzenlenmelidir.

Kaynaklar

- Airasian, P. W. ve Russell, M. (2008). *Classroom assessment : concepts and applications* (6th ed.). Boston: McGraw-Hill Higher Education.
- Airasian, P. W. (1991). Perspectives on Measurement Instruction. *Educational Measurement: Issues and Practice*, 10(1), 13-16,26.
- Black, P. ve Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Gelbal, S. ve Keleciođlu, H. (2007). Öğretmenlerin Ölçme ve değerlendirme Yöntemleri Hakkındaki Yeterlik Algıları ve Karşılaştıkları Sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Gipps, C.V. (1999). Socio-cultural aspects of assessment. *Review of Research in Education*, 24, 355-392.
- Gullickson, A. R. (1986). Teacher Education and Teacher-Perceived Needs in Educational Measurement and Evaluation. *Journal of Educational Measurement*, 23(4), 347-354.
- Kaptan, F. (1998). Fen Öğretiminde Kavram Haritası Yönteminin Kullanılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 14, 95-99.
- Kaptan, F. ve Korkmaz, H. (2000). Fen Öğretiminde Tümel (Portfolio) Değerlendirme. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 19, 212-220.
- Öncü, H. (2009). Ölçme ve Değerlendirmede Yeni Bir Yaklaşım: Portfolyo Değerlendirme. *Türkiye Sosyal Araştırmalar Dergisi*, 1, 103-130.
- Pellegrino, J. W., Chudowsky, N. ve Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Shepard, L.A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29 (7), 4-14.
- Shepard, L.A. (2006). *Classroom assessment*. In Brennan, RL (Ed.), *Educational measurement*. (4th ed.). (pp. 624-646). Westport, CT: Praeger.
- Stiggins, R. J. (2008). *Student-involved assessment for learning* (5th ed.). Upper Saddle River, NJ: Pearson / Merrill Prentice Hall.