



ISSN:1306-3111

e-Journal of New World Sciences Academy
2012, Volume: 7, Number: 1, Article Number: 1C0492

NWSA-EDUCATION SCIENCES

Received: May 2011

Accepted: January 2012

Series : 1C

ISSN : 1308-7274

© 2010 www.newwsa.com

Emrah Gül

Ankara University

egul@ankara.edu.tr

Ankara-Turkey

ONLINE DEĞERLENDİRME GÜVENİLİR MİDİR?

ÖZET

Online ölçme ve değerlendirmeler, öğretmen adaylarının gerçek sınıf ortamında ölçme-değerlendirme yapabilmelerini sağlamada ve ölçme-değerlendirme derslerinde, çağdaş yöntemlerini anlamalarında ve kullanmalarında etkili ve doğru bir yoldur. Bu durun şu soruyu akla getirmektedir: Online Değerlendirme Güvenilir Midir? Bu çalışmanın amacında, öğretmen adaylarının ölçme ve değerlendirme okuryazarlık seviyelerini belirlemede kağıt üzerinde yapılan uygulamalar ile online uygulamaların güvenilirliklerini karşılaştırmaktır. Araştırma sonuçları göstermiştir ki, online çözülen envanterin güvenilirlik katsayısı ($KR_{20}=0,803$), kağıt üzerinde çözülen envanterin güvenilirlik katsayısından ($KR_{20}=0,718$) daha yüksek olduğu görülmüştür.

Anahtar Kelimeler: Online Değerlendirme, Madde Analizi, Değerlendirme Okuryazarlığı, TAP, Güvenirlik

IS ONLINE ASSESSMENT RELIABLE?

ABSTRACT

Online based measurement and assessment are effective and accurate ways to provide the teacher candidates making measurement in actual classroom environment and to understand measurement and assessment methods at evaluation courses. This situation connotes this question: Is Online Assessment Reliable? The purpose of this study is to compare the determining of the levels of assessment literacy of teacher candidates on paper and online or web-based applications reliabilities. This research has shown that the reliability coefficient ($KR_{20}= 0.803$) of online or web-based inventory is higher than reliability coefficient ($KR_{20} = 0.718$) of on paper inventory.

Keywords: Online Assessment, Item Analysis, Assessment Literacy, TAP, Reliability

1. GİRİŞ (INTRODUCTION)

The World Wide Web, eğitim literatürünün gelişimi izlendiğinde birçok alanda olduğu gibi öğrenme ve öğretme alanında da büyük bir perspektif yaratmış ve birçok eğitimci-araştırmacı tarafından kullanılmaya başlanmıştır[14 ve 10]. Özellikle öğretimin ayrılmaz bir parçası olan ölçme ve değerlendirme uygulamaları ve yapılan araştırmalar online sistemlerin ne derece önemli olduğunu ortaya koymuştur. Öğrencilerin neleri öğrenmek istedikleri, ne kadar öğrenmek istedikleri, ne zaman öğrenmek istediklerini online değerlendirme sistemleri sağlamaktadır. Oysaki geleneksel değerlendirme yöntemleri öğrencileri belirli bir zaman ve mekân dilimi içerisine sokarak öğrencilerin gerçek anlamda neleri öğrenip öğrenmediklerini anlayamazlar. Birçok araştırmada ortaya konan bulgular online yapılan ölçme ve değerlendirmelerde öğrencilerin kendilerini daha rahat hissettikleri, öğrenmeye ve ölçme-değerlendirmeye karşı olumsuz tutumlarının ortadan kalktığını ve akademik başarılarının arttığını ortaya koymaktadır. Geleneksel yöntemlerde harcanan zaman, online yöntemlere nazaran çok daha fazladır. Bunun yansısı ortaya çıkan sonucun başarısı ise çok düşüktür. Kâğıt-kalem ile belirli bir sınıfın ve zamanın içerisinde yapılan değerlendirmeler öğrenciyi sınav psikolojisi içerisine sokarak bildiği bilgileri dahi unutturmaktadır. İnternet yolu ile yapılan ölçme ve değerlendirmeler ise öğrenciyi rahatlatmakta sınav stresinden uzak tutarak daha güvenilir bir ölçme ve değerlendirme yapmalarını sağlamaktadırlar. Kâğıt üzerinde yapılan ve internet yardımı ile yapılan değerlendirmeler karşılaştırıldığında sonuçlar internet destekli yapılan ölçmeler lehine çıkmaktadır[3ve10].

Online değerlendirme öğrenci ve öğretmen açısından çok uygun, kullanışlı ve rahattır. Bu nedenle öğrenme ve öğretmenin ayrılmaz bir parçası olduğu vurgulanmıştır. Online değerlendirmeler öğrencinin akademik başarısının değerlendirilmesinde kullanılabilir. Online değerlendirmeler öğrencilere öğrendikleri bilgi veya kavramla ilgili hızlı bir şekilde geri dönüt sağladığı için öğrencinin değerlendirilirken de öğrenmesine büyük katkı sağlayacaktır. Geleneksel sınıflardaki değerlendirme düşünüldüğünde, öğretmen bir soru sorduğunda 1 veya 2 öğrenci soruyu yanıtlayabilir ve öğretmen her bir öğrencinin anlayıp anlamadığı belirleyemeyecektir. Öğrenciye online olarak bir soru yöneltildiğinde, her bir öğrencinin cevabı online kurs devam ederken bile gelecektir ve öğretmen bu cevapları anında görüp hızlı bir şekilde geri dönüt vereceğinden öğretim kusursuz olacaktır [14]. Online değerlendirme yöntemlerinin bu kadar çok avantajı olmasına rağmen yükseköğretimde ve öğretmen yetiştiren kurumlarda hala etkili biçimde kullanılmamaktadır [18].

2. ÇALIŞMANIN ÖNEMİ (RESEARCH SIGNIFICANCE)

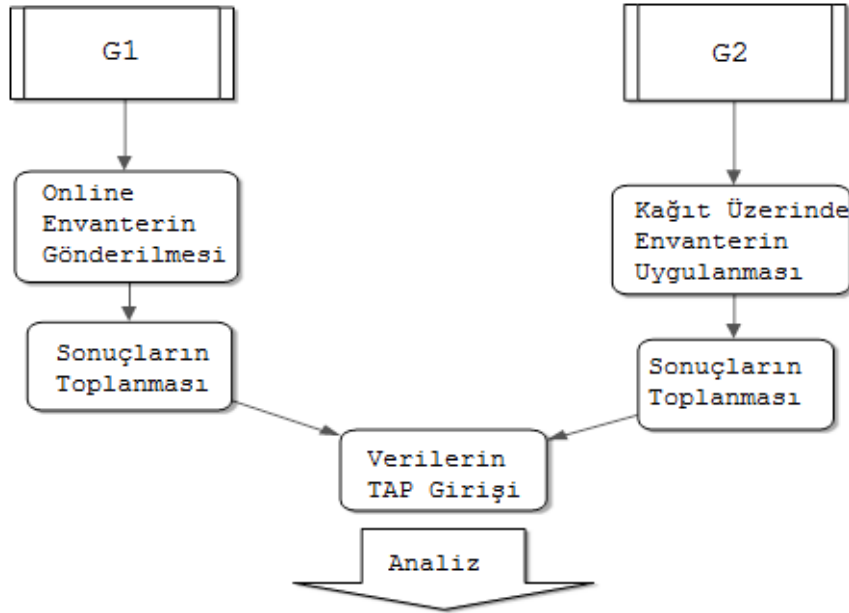
Son yıllarda eğitimde yapılan ölçme-değerlendirmeler, online veya web'e dayalı olarak internet ortamına aktarılmaya başlanmıştır. İnternet ölçme ve değerlendirme sisteminin ölçme ve değerlendirme işlemleri; en iyi, güvenilir, doğru, hatasız ve etkin şekilde yapılabileceği yapılan birçok araştırmada vurgulanmaktadır [17]. Ayrıca online olarak yapılan ölçme ve değerlendirme süreçlerinde öğrenciler çok hızlı geribildirim almakta, eksiklerini anında görmektedirler. Dünya literatürü incelendiğinde online sınav sistemlerinin çok fazla kullanıldığını fakat ülkemizde çok yaygın olmadığı görülmektedir. Akademisyenler, öğretmenler ve eğitim kurumları öğrencileri kâğıt-kalem testleri ile değerlendirmekte ve online ölçme ve değerlendirmenin avantajlarından faydalanmamaktadırlar. Bu nedenle bu araştırma özellikler öğretmen yetiştiren kurumlar açısından çok büyük önem taşımaktadır. Araştırmanın bir diğer önemi ise, "değerlendirme okuryazarlığı

envanterinin" ülkemizde yok denecek kadar az kullanılmış olmasıdır. Envanterin online formuna rastlamamış olamaması da bu araştırmanın önemini arttırmaktadır [8].

3. DENEYSSEL YÖNTEM (EXPERIMENTAL METHOD)

Bu araştırmaya Fırat Üniversitesi 3. Sınıfının son döneminde ölçme ve değerlendirme dersi almış toplam 94 öğretmen adayı katılmıştır. Bu öğrencilerden online envanterleri çözmek üzere 48 ve kağıt üzerindeki envanterleri çözmek üzere 46 öğrenci katılmıştır. Bu öğrenciler sırası ile G_1 (online) ve G_2 (kağıt-kalem) olarak adlandırılmıştır. Envanter hem online hem de kağıt üzerinde toplam 34 sorudan oluşmaktadır. Online envanter bir web sitesi yardımı ile hazırlanmış ve öğrencilere mail yolu ile gönderilmiştir. Sonuçlar toplu halde araştırmacıya ulaşmıştır. Kağıt üzerinde yapılan envantere ise bir ders saati boyunca öğrenci motivasyonunu bozmayacak şekilde sınıfa sunulmuş ve sonuçlar işaretlendikten sonra toplanmıştır.

Öğretmen adaylarına online envanter mail yoluyla ve kağıt üzerinde yapılacak olan ise bir ders saati içerisinde verildi. Kağıt üzerinde yapılacak olan envanter için 45 dakika online envanter için ise mail gönderildikten en fazla 1 hafta sonra cevaplanması istendi. Araştırmada veriler, Brooks [2] (2003) tarafından geliştirilmiş TAP (Test Analiz Programı) programı kullanılarak analiz edilmiştir. Aşağıdaki şekil 1'de yapılan uygulamalar şematiğe edilmiştir.



Şekil 1. Prosedür
(Figure 1. Procedure)

4. DEĞERLENDİRME OKURYAZARLIĞI ENVANTERİ (ASSESSMENT LITERACY INVENTORY)

G_1 ve G_2 katılımcılarının değerlendirme okuryazarlıklarını karşılaştırmak için, her iki gruba da "Sınıfta Değerlendirme Okuryazarlığı Envanteri [Classroom Assessment Literacy Inventory (CALI)]" uygulanmıştır. Bu envanter 1987-1993 yılları arasında, 3 farklı organizasyonun[American Federation of Teachers (AFT), The National Council on Measurement in Education (NCME), The National Education Association (NEA)] yaklaşık 6 yıllık çalışması sonucunda

belirlenen ve bir öğretilende olması gereken ölçme ve değerlendirme yeterliğini içeren "Değerlendirme Okuryazarlığı" kavramı üzerine Plake [12] (1993) tarafından oluşturulmuştur. Bir öğretmenin sahip olması gereken 7 standart belirlenmiş ve her standart için 5'er soru oluşturulmuştur. Aşağıdaki Tablo 1 de standartlar ve soru numaraları verilmiştir.

Tablo 1. Değerlendirme standartları ve soru numaraları
[Table 2. The classroom assessment literacy inventory (CALI)]

1. Uygun değerlendirme yöntemlerini seçebilme	1, 2, 3, 4, 5
2. Uygun değerlendirme yöntemlerini geliştirebilme	6, 7, 8, 9, 10
3. Değerlendirme sonuçlarını yorumlama, puanlama ve yönetme	11, 12, 13, 14, 15
4. Eğitimsel kararlar alırken değerlendirme sonuçlarını kullanabilme	16, 17, 18, 19, 20
5. Geçerli not verme statüsü geliştirme	21, 22, 23, 24, 25
6. Değerlendirme sonuçlarını paylaşabilme	26, 27, 28, 29, 30
7. Değerlendirmede etik olmayan durumları bilme	31, 32, 33, 34, 35

5. SONUÇLAR (CONCLUSIONS)

Tablo 2. Online ve kâğıt üzerinde yapılan değerlendirme okuryazarlığı envanterine ilişkin madde analizi sonuçları

(Table 2. Item analysis results on paper and online assessment literacy inventory)

	Testin Tüm Verilerine İlişkin		Üst ve Alt % 27'lik Gruplar İçin			
	Doğru Cevaplayanların Sayısı (I_i)		Ayırt Edicilik İndeksi (r_j)		Zorluk İndeksi (p_j)	
	G_2	G_1	G_2	G_1	G_2	G_1
1.	10	19	0,34	0,47	0,49	0,61
2.	23	26	0,37	0,31	0,61	0,55
3.	31	39	0,39	0,40	0,68	0,68
4.	22	27	0,44	0,41	0,56	0,65
5.	23	26	0,32	0,36	0,50	0,52
6.	19	22	0,35	0,40	0,67	0,60
7.	35	41	0,26	0,30	0,48	0,49
8.	26	29	0,47	0,45	0,57	0,59
9.	16	18	0,44	0,49	0,79	0,81
10.	29	33	0,29	0,36	0,23	0,38
11.	18	23	0,39	0,34	0,79	0,85
12.	24	26	0,56	0,54	0,59	0,56
13.	20	27	0,36	0,41	0,49	0,39
14.	19	24	0,44	0,41	0,68	0,59
15.	31	42	0,31	0,35	0,20	0,15
16.	11	13	0,48	0,46	0,79	0,88
17.	19	23	0,39	0,41	0,64	0,56
18.	26	29	0,33	0,38	0,44	0,43
19.	10	15	0,41	0,45	0,71	0,59
20.	13	19	0,49	0,48	0,58	0,61
21.	9	13	0,29	0,33	0,77	0,62
22.	16	22	0,33	0,31	0,42	0,51
23.	18	23	0,35	0,37	0,59	0,56
24.	19	24	0,40	0,41	0,44	0,45

25.	26	30	0,38	0,34	0,44	0,50
26.	17	20	0,63	0,59	0,57	0,58
27.	21	29	0,62	0,55	0,44	0,49
28.	29	36	0,32	0,41	*0,17	0,35
29.	25	27	0,46	0,43	0,39	0,42
30.	29	34	0,45	0,47	*0,19	0,38
31.	23	25	0,39	0,51	0,50	0,49
32.	19	23	0,43	0,46	0,55	0,51
33.	17	20	0,57	0,51	0,55	0,59
34.	22	21	0,33	0,36	0,49	0,45

Kağıt üzerinde yapılan envanterin analiz sonuçları incelendiğinde KR_{-20} güvenilirlik katsayısının (0,718), online envanter (0,803) ile karşılaştırıldığında daha düşük olduğu görülmektedir. Ayrıca yapılan madde analizi sonuçlarına bakıldığında online envanterin genel olarak ta kağıt üzerinde yapılan envanterden üstün olduğu görülmektedir.

Tablo 3. TAP analiz sonuçları
(Table 3. TAP analysis results)

G2	$KR_{-20} = 0,718$	$KR_{-20} = 0,803$	G1
	$p_j = 0,610$	$p_j = 0,525$	
	$r_j = 0,374$	$r_j = 0,401$	

Online değerlendirme güvenilir midir? Bu soru birçok eğitimci ve öğretmen tarafından yıllarca tartışılmıştır. Kâğıt-kalem ile yapılan değerlendirmeler öğrenci başarısını değerlendirmede ve eğitimsel kararlar almada çok yavaş ve yeterince etkin olmayan bir yoldur. Elektronik değerlendirmelerin ne derecede yararlı olduğu açıktır.

Öğretmenler e-değerlendirmeyi kullanarak daha hızlı sonuçlar aldıkları gibi bilgisayarları kullanmak eğitim ve öğretimi daha iyi bir yapıda şekillendirmeyi de sağlayacaktır. Bu sayede geleneksel bir eğitim ortamından, öğrenciyi değerli kılan ve öğrenmesine katkı sağlayan bir ortam oluşturacaktır. Son yıllarda eğitim alanında geleneksel ölçme ve değerlendirme yöntemlerinin yerini alternatif yöntemler olarak ölçme ve değerlendirmeye yeni bir boyut kazandırmıştır. Eğitim teknolojileri ile ilgilenenler alternatif ve otantik olarak bilgisayarların ve internetin ölçme ve değerlendirmede kullanılmasını yaygın hale getirmeye çalışmaktadır [5, 7, 11, 13 ve 15].

Chase [4] (1999) zorluk değerleri 0,20 ile 0,85 arasında olan maddelerin psikometrik açıdan iyi olarak değerlendirilebileceğini belirtmişlerdir. Kağıt üzerinde yapılan envanterdeki maddelerin ortalama zorluk değeri 0,610 olarak bulunmuştur. Bayrakçeken [1] (2008) ve Tekin [16] (1996) ortalama güçlük değerinin 0,50 olması gerektiğini vurgulamışlardır. Online envanterin zorluk indeksinin 0,525 olması 0,5 e daha yakın olması gereği psikometrik açıdan daha iyi olduğu söylenebilir.

Nunnally ve Berstein [9] (1994) ve Chase [4] (1999), ayırt edicilik değeri 0,30 ve daha fazla olan maddeleri psikometrik açıdan iyi olarak ifade etmişlerdir. Ayırt edicilik değeri 0,40 ve üstü olan maddelerin çok iyi, 0,20 ile 0,40 arasındaki maddelerin iyi, 0,0 ile 0,20 arasındaki maddelerin zayıf, 0,00'dan az olan maddelerin atılması veya gözden geçirilmesi gerektiğini belirtilmiştir. Online envanterdeki 22 maddenin ayırt ediciliği 0,40 üstü yani çok iyi

derecede olduğu buna karşın, kâğıt-kalem ile çözülen envanterin ise 15 maddesinin çok iyi olduğu görülmektedir.

Bu araştırma, kağıt üzerinde yapılan değerlendirme ile online olarak yapılan değerlendirmeler arasında önemli ve anlamlı bir farklılık olduğunu gösteriyor. Yapılan araştırmalarda öğrencilerle online yapılan sınavlarla ilgili fikirleri sorulduğunda kendilerini daha rahat hissettikleri ve değerlendirilirken kendilerini daha iyi ifade ettiklerini ifade etmişlerdir [6]. Elde edilen sonuçlarda literatürü destekler niteliktedir.

Yapılacak olan araştırmalar, yeni ölçme ve değerlendirme yöntemlerini kapsayarak ve mutlaka internet ve bilişim destekli yapılmalıdır. Ayrıca araştırmalar yapıldıktan sonra öğrenci algısı ölçülmeli ve öğrenci ihtiyaçları göz önünde bulundurularak yeniden düzenlenmelidir. Bu anlamda eğitim fakültelerine düşen görev daha nitelikli öğretmenler yetiştirmek ve onları bilişim çağına uygun hale getirmektir. Online ölçme ve değerlendirme yönteminin en güzel taraflarından biri de yeni öğretim yaklaşımlarına uygun olmasıdır. Öğrencilerin bilgiyi yapılandırmaları sürecinde en büyük destekçisi olacaktır.

NOT (NOTICE)

Bu çalışma, 22-24 Eylül 2011 tarihleri arasında Elazığ'da düzenlenen "(ICITS-2011) 5. Uluslararası Bilgisayar ve Öğretim Teknolojileri Sempozyumu"nda sözlü bildiri olarak sunulmuştur.

KAYNAKLAR (REFERENCES)

1. Bayrakçeken, S., (2008). Test geliştirme. İçinde, E. Karip (Ed.), Ölçme ve değerlendirme, Ankara: Pegem Akademi Yayıncılık ss:244-274
2. Brooks, G.P. and Johanson, G.A., (2003). Test analysis program. Applied Psychologica Measurement, 27, pp:300-306.
3. Calvert, E.S. and Waterfall, R.C., (1982). A Comparison of Conventional and Automated Administration of Raven's Standard Progressive Matrices. International Journal of Man-Machine Studies, 17, pp:300-310.
4. Chase, C.I., (1999). Contemporary assessment for educators, New York: Longman
5. Davies, P., (2000). Computerized peer assessment. Innovations in Education and Teaching International 37, pp: 346-355.
6. Dermo, J., (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. British Journal of Educational Technology 40 (2), pp: 203-214
7. Lawrenz, F., Huffman, D., and Welch, W., (2001). The science achievement of various subgroups on alternative assessment formats. Science Education 85, pp: 279-290.
8. Mertler, A.C. and Campbell, C., (2005). Measuring teachers knowledge & application of classroom assessment concepts: development of the assessment literacy inventory. Annual Meeting of American Educational Research Association, Quebec, Canada, 11-15, April.
9. Nunnally, J.C. and Bernstein, I.H., (1994). Psychometric theory (3th edition). New York: Mc Graw-Hill.
10. Olsen, J.B., Maynes, D.M., Slawson, D.A., and Ho, K., (1986). Comparison and Equating of Paper-administered, Computer-administered and Computerized Adaptive Tests of Achievement, paper presented at the meeting of the American Educational Research Association, San Francisco.
11. Özmen, H., (2006). E-Sınav Sistemi, Akademik Bilişim, Pamukkale University, Denizli

12. Plake, B.S. and Impara, J.C., (1993). Teacher assessment literacy questionnaire. University of Nebraska- Lincoln. In cooperation with the National Council on Measurement in Education and the W.K. Kellogg Foundation.
13. Reeves, T.C., (2000). Alternative assessment approaches for online learning environments in higher education. Journal of Educational Computing Research 23, pp:101-111.
14. Robles B. and Braathen S., (2002). Online assessment techniques. Delta Pi Epsilon Journal 44 (1)
15. Ross, S.J., (2005). The impact of assessment method on foreign language proficiency growth. Applied Linguistics 26, pp:317-342.
16. Tekin, H., (1996). Eğitimde Ölçme ve Değerlendirme. Ankara: Yargı Yayınları
17. Tsai, C.C., (2001). The interpretation construction design model for teaching science and its applications to internet-based instruction in Taiwan. International Journal of Educational Development 21,pp:401-415.
18. Wen, M.L. and Tsai, C.C., (2006). University students' perceptions of and attitudes toward (online) peer assessment. Higher Education 51, pp:27-44.

