



Farklı soru türlerinin güçlük ve ayırt edicilik düzeylerinin incelenmesi

Yrd. Doç. Dr. Murat AKYILDIZ^a
Yrd. Doç. Dr. Nejdett KARADAĞ^a

^a Anadolu Üniversitesi, Açıköğretim Fakültesi

Özet

Açık ve uzaktan öğrenme sistemlerinde çoktan seçmeli soruların çeşitli türleri sınavlarda kullanılmaktadır. Bunlar geleneksel çoktan seçmeli, öncüllü, boşluk doldurma türü ve eşleştirme türü sorulardır. Bu soruların herhangi birisinin daha zor ya da kolay olup olmadığını belirlemek önemlidir. Ayrıca bu soru tiplerinden birinin daha ayırt edici olduğunu belirlemek de sınav geçerliliği açısından önemlidir. Bu çalışmada güçlük ve ayırt ediciliklerin soru türüne göre anlamlı farklılık gösterip göstermediği belirlenmeye çalışılmıştır. Bu amaçla Anadolu Üniversitesi Açıköğretim sınavlarının 2017-2018 güz döneminde dönem sonu sınavında kullanılmış olan 9296 soruda geleneksel çoktan seçmeli, öncüllü, boşluk doldurma ve eşleştirme türü soruların güçlük ve ayırt edicilikleri birbirleriyle karşılaştırılmıştır. Boşluk doldurma sorularının diğer sorulara göre daha kolay ve daha ayırt edici olduğu bulunmuştur.

Anahtar Sözcükler: Madde güçlüğü, madde ayırt ediciliği, soru türleri, uzaktan öğrenme

Abstract

Different kinds of multiple choice questions are used in exams in open and distant education systems. These types are multiple choice, information dependent, completion and matching type questions. It is important to determine whether one of these types are harder than the other. Besides, it is important to determine whether one of these question types is more discriminating in terms of exam validation. In this study, we tried to determine whether difficulty and discrimination value differ based on question type. To this end, the difficulty and discrimination values of traditional multiple choice, information dependent, completion and matching type questions of 9296 questions used in the final exam in Anadolu University Open Education Faculty 2017-2018 fall semester. It is found out that completion type of questions are easier and more discriminating.

Keywords: Item difficulty, item discrimination, question types, distant education.

Giriş

Ölçme ve değerlendirme yoluyla bir yandan uygulanan eğitim programlarının etkililiği belirlenirken diğer yandan da programlara katılan öğrencilerin başarı durumları ölçülür. Öğrencilerin başarı durumlarının ölçülmesinde çeşitli ölçme araçları kullanılır. Alan yazında çeşitli sınıflamalara rastlansa da temel olarak bu araçlar geleneksel ve alternatif ölçme araçları olarak iki başlık altında toplanabilir. Geleneksel ölçme araçları daha çok öğrenme süreci sonunda düzey belirleme amaçlı değerlendirme için kullanılırken, alternatif ölçme araçları ise daha çok öğrenme sürecinde performansın ölçülmesi amacıyla kullanılmaktadır. Öğrenme çıktılarının niteliğini tam olarak belirleyebilmek için öğrencileri değerlendirmede çeşitli ölçme araç ve teknikleri işe koşulmalıdır (Popham, 2001).^[1] Ayrıca, öğrencilerin farklı öğrenme biçimlerine sahip olmaları ve değerlendirmenin farklı amaçlarla yapılması da öğrencilerin başarı durumlarının belirlenmesinde farklı ölçme araç ve tekniklerinin kullanılmasını gerektirmektedir (AL-Smadi, Wesiak, Guetl, Holzinger, 2012; Grover, 2009; Huba ve Freed, 2000). Amerikan Yüksek Öğretim Derneğinin (American Association for Higher Education-AAHE) öğrencilerin değerlendirilmesi süreciyle ilgili belirlemiş olduğu uygulama ilkelerinin birinde öğrenmenin karmaşık bir süreç olduğu, bilgi ve becerinin yanında değer, tutum ve alışkanlıkları da içerdiğinden hareketle, öğrencilerin eğitim deneyimlerini geliştirmek için farklı ölçme yöntemlerinin kullanılması gerektiğine vurgu yapılmaktadır (Astin ve diğ.,1996).

Öğrenme etkinliklerinin uzaktan eğitim yoluyla gerçekleştirildiği açık ve uzaktan eğitim sistemlerinde daha çok geleneksel ölçme araçlarından çoktan seçmeli testler kullanılmaktadır (Karadağ, 2014). Çoktan seçmeli testlerde en çok kullanılan soru türleri; soru kökü olumlu veya olumsuz soru kipinde olan sorular, soru kökü eksik cümle olan sorular, tek bir doğru cevap istenen sorular, en doğru cevap istenen sorular, cevabı işlem gerektiren sorular, alternatif seçmeli, doğru-yanlış, çoklu doğru-yanlış, eşleştirme, karmaşık çoktan seçmeli (K Tipi) sorular, ortak seçenekli sorular, ortak köklü sorular, olumlu, olumsuz, işlemleri ve birleşik sorular olarak sıralanabilir (Koçdar, Karadağ, Şahin ve Karadeniz, 2017; Haladyna, Downing ve Rodriguez, 2004; Tekin, 1994).

Dünyanın önde gelen mega üniversitelerinden biri olan Anadolu Üniversitesi'nde de Açıköğretim Sistemi'nde öğrencilerin başarı durumlarının belirlenmesinde ağırlıklı olarak çoktan seçmeli testlere yer verilmektedir. Açıköğretim Sistemi'nde çeşitli kurum sınavları da dahil edildiğinde bir yıl içinde 3.360 derste 128.000'den fazla soru kullanılmaktadır (Karadağ, Akyıldız, Kumtepe ve Akgün, 2017). Sınavlarda yer verilen sorular 4982 sayılı Bilgi Edinme Kanunu kapsamında sınavlardan sonra yayınlanmaktadır. Bu nedenle soru yazarlarından her

sınavda farklı/özgün soru hazırlamaları istenmektedir. Soru hazırlanacak ders kitaplarının içeriğinin sınırlı olması, farklı/özgün soru hazırlama konusunda soru yazarlarını kısıtlamaktadır. Bu nedenle birçok soru yazarından belli sayıda soru hazırlanan içeriklerden artık yeni soru üretmediklerine ilişkin geribildirim alınmaktadır. Bu soruna çözüm olabileceği düşüncesiyle 2017-2018 öğretim yılı güz döneminden itibaren soru yazarlarına yönelik olarak aynı yeterliği ölçen farklı türde soru örnekleri hazırlanmıştır. Bu sorular; senaryo soruları, eşleştirme soruları, öncüllü sorular ve boşluk doldurma sorularıdır.

Bu çalışmada, 2017-2018 Güz Döneminde soru yazarları tarafından hazırlanan ve Ara Sınavda kullanılan öncüllü, eşleştirme ve boşluk doldurma sorularının madde parametrelerinin incelenmesi ve geleneksel çoktan seçmeli sorular ile karşılaştırılması amaç edinilmiştir.

Araştırma Sorunsalı

Farklı soru türlerini kullanarak öğrenci başarısının gerçeğe daha yakın kestirilmesi açık ve uzaktan öğrenmede sıklıkla başvuru bir yoldur. Sınavların zorluk ve ayırt edicilik değerlerini sınavdan önce bilebilmek ve kontrol altında tutabilmek amacıyla madde güçlük ve ayırt edicilik değerleri önemli bir görev görmektedir. Bu çalışmada Anadolu Üniversitesi Açıköğretim sisteminde kullanılmakta olan çeşitli türlerden soruların farklı güçlük ve ayırt ediciliklere sahip olup olmadıkları test edilecektir.

Araştırmanın birinci sorusu, soru türlerine göre soruların güçlük düzeylerinin birbirinden anlamlı derecede farklılık gösterip göstermediğidir. Cevap aranan ikinci soru ise soruların ayırt edicilik değerlerinin soru türlerine göre anlamlı bir farklılık göstermediğidir.

Yukarıdaki araştırma sorularına verilecek yanıtlar belirli amaçlar için hazırlanan sınavlarda sınavların zorluk ve ayırt edicilik düzeylerini sınav öncesinde kontrol etmeye yardımcı bilgilerin elde edilmesine yardımcı olacaktır.

Yöntem

Araştırma Modeli

Araştırmada farklı soru türlerinin madde güçlüğü ve ayırt ediciliği değerleri karşılaştırılmıştır. Araştırmada kullanılan tüm değerler nicel değerlerdir ve bu değerlerin farklı gruplarda ortalamaları arasındaki ilişkiler incelenmiştir. Bu anlamıyla araştırma var olan bir durumu olduğu haliyle ortaya koymaya çalışan betimsel nicel bir araştırmadır.

Evren ve Örneklem

Araştırmaya dahil edilen sorular Açıköğretim sisteminde 2017-2018 güz dönemi öğretim yılında dönem sonu sınavında kullanılan 9286 sorunun tamamından oluşmaktadır. Bu sorular bir evren olarak değerlendirilmemiştir. Açıköğretim sisteminde ara ve dönem sonu sınavlarında kullanılan soruların tamamı bir evren olarak düşünülecek olursa bu çalışmaya dahil edilen sorular bir araştırma grubu olarak ele alınmıştır. Bu nedenle soruların güçlük ve ayırt edicilik değerleri istatistiksel analizlerle karşılaştırılmıştır. Aşağıda Tablo 1’de araştırmada dahil edilen soruların türlerine göre sayıları verilmiştir.

	N	Yüzde
Boşluk Doldurma	390	4.2
Eşleştirme	58	.6
Öncüllü	392	4.2
Geleneksel Çoktan Seçmeli	8456	91.0
Total	9296	100.0

Veri Toplama Araçları

Araştırmada kullanılan soruların türlerine, güçlüklerine ve ayırt ediciliklerine ilişkin verilere Açıköğretim sistemi İstatistik Bilgi Edinme ve Değerlendirme Birimi aracılığı ile erişilmiştir.

Veri Analizi

Elde edilen veriler SPSS 24 programına aktarılmıştır. Soruların türlerine bağlı olarak madde güçlüklerinin ve madde ayırt ediciliklerinin anlamlı farklılık gösterip göstermediği tek yönlü varyans analizi ile test edilmiştir. Varyans analizi, grup sayısının ikiden fazla ve gruplardaki obje sayısının 30’dan fazla olduğu durumlarda kullanılan güçlü bir parametrik istatistiktir.

Analiz sonuçlarının görselleştirilmesi için SPSS 24 ve R programlarından yararlanılmıştır.

Bulgular ve Yorumlar

Araştırmanın birinci sorusu, kullanılan soru türlerinin güçlük düzeylerinin birbirinden anlamlı bir farklılık gösterip göstermediğini ortaya koymaya yöneliktir. Bu amaçla farklı soru türlerinin güçlük ortalamalarının anlamlı farklılık gösterip göstermediğini test etmek amacıyla varyans analizi uygulanmış ve elde edilen sonuçlar aşağıda Tablo 2 ve Tablo 3'te gösterilmiştir.

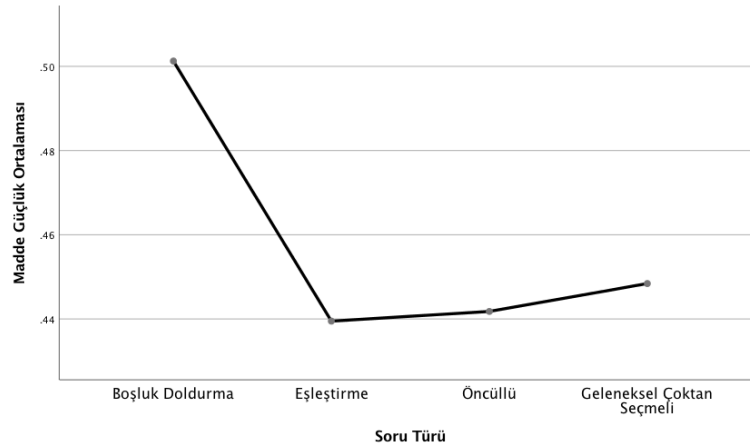
	N	Ortalama	Sdt. Sapma	Minimum	Maksimum
Boşluk Doldurma	390	.50	.21	.03	.97
Eşleştirme	58	.44	.19	.08	.88
Öncüllü	392	.44	.21	.02	.96
Geleneksel Çoktan Seçmeli	8456	.45	.21	.00	1.00
Toplam	9296	.45	.21	.00	1.00

	Kareler Toplamı	sd	Ortalama a Kare	F	p.	Fark
Gruplar Arası	1.078	3	.359	8.091	.000*	BD-Öncüllü
Gruplar İçi	412.764	9292	.044			BD-GÇS
Toplam	413.843	9295				

*p<0.05 BD=Boşluk Doldurma, GÇS=Geleneksel Çoktan Seçmeli

Tablo 2'den de görüleceği gibi farklı soru türlerinin güçlük düzeyleri ortalamaları arasında istatistiksel olarak anlamlı bir farklılık gözlenmiştir ($F_{3-9292;0.05}=8.091$). Buna göre, soru türleri arasında güçlük bakımından anlamlı farklılık bulunmaktadır. Bir başka ifadeyle bazı soru türleri diğerlerine oranla daha zordur. Gözlenen farklılığın hangi soru türleri arasından kaynaklandığını bulabilmek amacıyla Tukey post-hoc testi uygulanmıştır. Elde edilen sonuçlara göre Anova testinde gözlenen farklılık Boşluk Doldurma türü sorular ile Öncüllü ve Geleneksel Çoktan Seçmeli sorular arasındaki farktan kaynaklanmaktadır. Buna göre boşluk doldurma türü soruların güçlük ortalaması (.50) öncüllü soruların ortalamasından (.44) daha yüksektir. Ayrıca boşluk doldurma türü soruların güçlük ortalaması geleneksel çoktan seçmeli soruların ortalamasından (.45) daha büyüktür. Başka bir ifadeyle güçlük indeksi büyüdükçe sorunun kolaylaştığı göz önüne alındığında, boşluk doldurma türü sorular öncüllü sorulardan ve geleneksel çoktan seçmeli sorulardan daha kolaydır. Eşleştirme, öncüllü ve geleneksel çoktan seçmeli soruların güçlükleri birbirine denktir. Ortalamaları birbirinden büyük farklılık göstermekle birlikte eşleştirme türü sorularla boşluk doldurma soruları arasında

anlamli bir farklılık gözlenmemiştir. Soru türlerinin madde güçlük ortalamalarını gösteren grafik aşağıda Şekil 1’de verilmiştir.



Şekil 1. Farklı Soru Türlerinin Madde Güçlük Ortalamaları

Araştırmanın ikinci sorusu ise farklı soru türlerinin ayırt ediciliklerinin birbirinden anlamli bir farklılık gösterip göstermediğine odaklanmaktadır. Aşağıda Tablo 4 ve Tablo 5’te soru türlerine bağlı olarak ayırt ediciliklerin değişimlerini incelemek amacıyla uygulanmış varyans analizi sonuçları verilmiştir.

	N	Ortalama	Sdt. Sapma	Minimum	Maksimum
Boşluk Doldurma	390	.43	.20	-.26	.95
Eşleştirme	58	.37	.16	.00	.68
Öncüllü	392	.34	.17	-.20	.90
Geleneksel Çoktan Seçmeli	8456	.39	.18	-.20	1.00
Toplam	9296	.39	.18	-.26	1.00

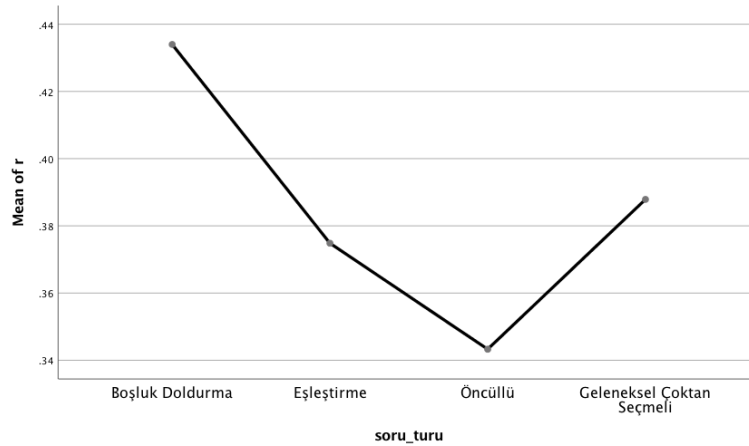
	Kareler Toplamı	sd	Ortalama Kare	F	p.	Fark
Gruplar Arası	1.618	3	.539	16.2*	.00*	BD-Öncüllü
Gruplar İçi	316.848	227.9*	.034			BD-GÇS
Toplam	318.466	9295				GÇS-Öncüllü

* Welch testi uygulanmıştır.

**p<0.05 BD=Boşluk Doldurma, GÇS=Geleneksel Çoktan Seçmeli

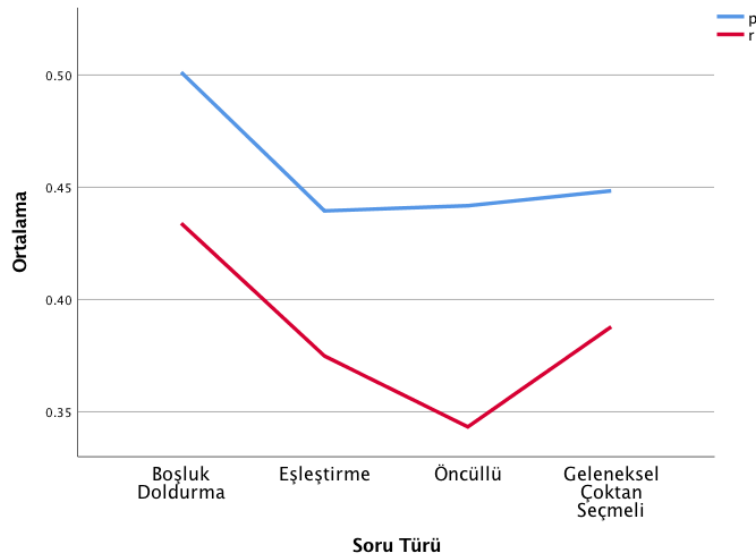
Tablo 5'ten de görüleceği gibi farklı soru türlerinin ayırt edicilik düzeyleri ortalamaları arasında istatistiksel olarak anlamlı bir farklılık gözlenmiştir ($F_{3-227.9;0.05}=16.2$). Buna göre, soru türleri arasında ayırt edicilik bakımından anlamlı farklılık bulunmaktadır. Bir başka ifadeyle bazı soru türleri diğerlerine oranla daha ayırt edicidir. Gözlenen farklılığın hangi soru türleri arasından kaynaklandığını bulabilmek amacıyla Tukey post-hoc testi uygulanmıştır. Elde edilen sonuçlara göre Anova testinde gözlenen farklılık boşluk doldurma türü sorular ile öncüllü ve geleneksel çoktan seçmeli sorular arasındaki farktan ve geleneksel çoktan seçmeli sorular ile öncüllü sorular arasındaki farktan kaynaklanmaktadır. Buna göre boşluk doldurma türü soruların ayırt edicilik ortalaması (.43) öncüllü soruların ortalamasından (.34) daha yüksektir. Ayrıca boşluk doldurma türü soruların ayırt edicilik ortalaması geleneksel çoktan seçmeli soruların ortalamasından (.39) daha büyüktür. Geleneksel çoktan seçmeli sorular öncüllü sorulardan, boşluk doldurma türü sorular öncüllü sorulardan ve geleneksel çoktan seçmeli sorulardan daha ayırt edicidir. Geleneksel çoktan seçmeli soruların ayırt ediciliği ise öncüllü sorulardan daha yüksektir. Eşleştirme sorularının ayırt edicilikleri öncüllü, boşluk doldurma türü ve geleneksel çoktan seçmeli soruların ayırt ediciliklerine denktir.

Soru türlerinin madde ayırt edicilik ortalamalarını gösteren grafik aşağıda Şekil 2'de verilmiştir.



Şekil 2. Farklı Soru Türlerinin Madde ayırt edicilik Ortalamaları

Soru türlerine göre madde güçlük ve ayırt ediciliklerinin birlikte değişimlerini görmek amacıyla aşağıda Şekil 3 verilmiştir.



Şekil 3. Soru türlerine göre madde güçlük ve ayırt edicilik değerlerinin değişimi (p=güçlük; r= ayırt edicilik)

Yukarıda Şekil 3'te görüleceği gibi en kolay sorular boşluk doldurma türü sorulardır. Diğer soruların güçlükleri birbirlerine çok yakındır. Ayırt ediciliği diğer sorulardan daha fazla görünen soru türü de boşluk doldurma sorularıdır. Öncüllü soruların ayırt ediciliği ise diğer soru türlerine göre daha düşük görünmektedir.

Sonuçlar

Araştırma sonucunda elde edilen bulgular Açıköğretim Fakültesi sınavlarında kullanılmış olan boşluk doldurma türü soruların diğer soru türlerine göre daha kolay olduğunu göstermiştir. Öncüllü, eşleştirmeli ve geleneksel çoktan seçmeli sorular ise birbirlerine benzer zorlukta görünmektedir. Boşluk doldurma türü sorular bir bilginin önemli bir kısmın boş bırakıldığı türden sorulardır. Açık ve uzaktan öğrenme sistemlerine devam etmekte olan öğrencilerin bilgiyi olduğu haliyle ezberleme yoluna başvurmalarının yaygınlığı nedeniyle bu sorular ezber yoluna başvuran öğrencilere daha yakın dolayısıyla da daha kolay gelmiş olabilir.

Soruların türlerine göre ayırt edicilikleri incelendiğinde en ayırt edici soruların boşluk doldurma türü sorular olduğu görülmüştür. Ayırt ediciliği en düşük sorular ise öncüllü sorulardır. Geleneksel çoktan seçmeli soruların ayırt ediciliği ise öncüllü sorular ile boşluk doldurma türü sorular arasında kalmaktadır. Eşleştirme türü soruların ayırt edicilik değerleri ise geleneksel çoktan seçmeli sorulara denktir. Boşluk doldurma türü soruların zorluklarının orta düzeyde olduğu dikkate alındığında en ayırt edici sorular olmaları beklenen bir durumdur.

Öneriler

Çalışma sonucunda boşluk doldurma türü soruların zorluklarının orta düzeyde olduğu ve ayırt edicilik değerlerinin de diğer sorulardan daha yüksek olduğu görülmüştür. Bir başarı testindeki soruların orta zorlukta olması ve yüksek ayırt ediciliğe sahip olması istenen bir durumdur. Boşluk doldurma türü soruların testlerde daha yaygın kullanılmasının uygun olabileceği değerlendirilmektedir. Fakat boşluk doldurma sorularının bilgi düzeyinden öğrenmeleri ölçtüğü unutulmamalı testlereki yoğunluğu buna dikkat edilerek ayarlanmalıdır.

Ayırt ediciliği en düşük sorular öncüllü sorulardır. İleride yapılacak sınavlarda bu soruların ayırt eciliğini arttırıcı önelmeler alınarak çalışma tekrarlanmalıdır.

Kaynakça

- AL-Smadi, M.; Guetl, C.; Holzinger, A.; Wesiak, G. (2012). Assessment for/as learning: Integrated automatic assessment in complex learning for self-directed learning, in Proceedings of the 2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems, (CISIS 2012), Jul 4-6 2012, pp. 929-934. Palermo, Italy: IEEE Computer Society.
- Astin A.W.; Banta, T.W.; Cross, K.P.; El-Khawas, E.; Ewell, P.T.; Hutchings, P.; Marchese, T.J.; McClenney, K.M.; Mentkowski, M.; Miller, M.A.; Moran, E.T. ve Wright, B.D. (1996). 9 principles of good practice for assessing student learning. AAHE Assessment Forum, July 25, 1996. <http://www.aahe.org/princip1.htm>
- Grover, K.T. (2009). Student assessment: An exploratory mixed methods study of teachers' perceptions and resulting practices. Yayınlanmamış Doktora Tezi. St Charles, MO: Lindenwood University.
- Haladyna, T.M. (2004). Developing and validating multiple-choice test items. (3 ed.) Mahwah, NJ: Lawrence Erlbaum Associates.
- Huba, M. E. ve Freed, J.E. (2000). Learner-centered assessment on college campuses: Shifting the focus from teaching to learning. Needham Heights, MA: Allyn & Bacon.
- Karadağ, N. (2014). *Açık ve Uzaktan Eğitimde Ölçme ve Değerlendirme: Mega Üniversitelerdeki Uygulamalar*. Yayınlanmamış Doktora Tezi. Anadolu Üniversitesi Sosyal Bilimler Enstitüsü. Eskişehir.
- Karadağ, N., Akyıldız, M., Kumtepe, A. T. ve Akgün, H. R. (2017). Anadolu Üniversitesi Açıköğretim Sistemi soru yazarlarının ölçme ve değerlendirme seminerlerine ilişkin görüşleri. *AUAd*, 3(3), 9-46.
- Koçdar, S.; Karadağ, N.; Şahin, M.D.; Karadeniz, A. (2017). Uzaktan eğitimde çoktan seçmeli soruların güçlük ve ayırt edicilik değerlerinin soru türlerine göre incelenmesi. Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 32(1), 168-184.
- Popham, W. J. (2001). The truth about testing: An educator's call to action. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tekin, H. (1994). Eğitimde ölçme ve değerlendirme. Ankara: Yargı Yayınları.

Yazarlar Hakkında

Yrd. Doç. Dr. Murat AKYILDIZ



Yrd. Doç. Dr. Murat Akyıldız, psikoloji lisans eğitimi aldıktan sonra yüksek lisansını ve doktorasını ölçme ve değerlendirme alanında tamamladı. Çok boyutlu ve çok parametrelili lojistik modellerle bireylerin bilişsel ve duyuşsal özelliklerinin kestirilmesi, Dr. Akyıldız'ın temel ilgi alanıdır. Yazarın diğer ilgi alanları arasında psikometri, istatistiksel analiz, bilim yöntemi ve felsefesi yer almaktadır. Dr. Akyıldız, halen Anadolu Üniversitesi Açıköğretim Fakültesi öğretim üyesi olarak görev yapmaktadır.

Posta adresi : Anadolu Üniversitesi Açıköğretim Fakültesi

Tel (İş) : +90 222 3350580/2410

Eposta : muratakyildiz@anadolu.edu.tr

Yrd. Doç. Dr. Nejdet KARADAĞ



Yrd. Doç. Dr. Nejdet Karadağ, Anadolu Üniversitesi Eğitim Fakültesi Fransızca Öğretmenliği Bölümü mezunudur. Dr. Karadağ, yüksek lisans ve doktora eğitimini Anadolu Üniversitesi Sosyal Bilimler Enstitüsü Uzaktan Eğitim Anabilim Dalı'nda tamamlamıştır. Yazar, halen Anadolu Üniversitesi Açıköğretim Fakültesi Yaygın Öğretim Bölümü'nde Yardımcı Doçent olarak çalışmaktadır. Ayrıca Açıköğretim Fakültesi'nde Test Araştırma Birimi'nin yöneticisi olarak görev yapmaktadır. Dr. Nejdet Karadağ'ın ilgi alanları açık ve uzaktan öğrenmede ölçme ve değerlendirme,

öğretim tasarımı ve mega üniversitelerdir.

Posta adresi : Anadolu Üniversitesi Açıköğretim Fakültesi, Yunusemre Kampüsü, Eskişehir, 26470

Tel (İş) : +90 222 3350580/2716

Eposta : nkaradag@anadolu.edu.tr