EPODDER

EĞİTİMDE VE PSİKOLOJİDE ÖLÇME VE DEĞERLENDİRME DERNEĞİ

BÜLTEN **HAZİRAN 2016** **SAYI: 2**

Başkan'dan 1

Editör'den 2

Bölüm I: Derneğimiz ve Çalışmalarımız 3

Bölüm II: Analiz 5

"Olmak Ya Da Olmamak":
Ölçek Geliştirme Sürecinin "Gariban Kalmışlığı" 5

Sivil Toplum Kuruluşlarında
Ölçme ve Değerlendirme Çalışmaları 12

Bölüm III: Lisansüstü Köşesi 17

Bölüm IV: Serbest Köşe... 23

Bölüm V: Tez Künyeleri 25

Bülten Hakkında

Bu bülten Eğitimde ve Psikolojide Ölçme ve Değerlendirme Derneği tarafından dernek çalışmalarını, alanımızdaki gelişme ve yenilikleri, fırsatlarını üyelerine/meslektaşlarına duyurmak amacıyla yılda iki defa yayınlanmak amacıyla hazırlanmıştır.

Bu amaç doğrultusunda bülten, üyelerimizden gelecek her türlü haber, bilgi, eleştiri ve önerilere açıktır. Sonraki bültende yayınlanmasını istediğiniz haber ya da bilgileri, varsa fotoğrafları ile birlikte bulten.epodder@gmail.com adresine göndermeniz yeterlidir.

Doç. Dr. İ. Alper KÖSE

Arş. Gör. Alperen YANDI

İçindekiler

Başkan'dan 1
Editör'den 2
Bölüm I: Derneğimiz ve Çalışmalarımız 3
Bölüm II: Analiz 5
"Olmak Ya Da Olmamak": Ölçek Geliştirme Sürecinin "Gariban Kalmışlığı" 5
Sivil Toplum Kuruluşlarında Ölçme ve Değerlendirme Çalışmaları 12
Bölüm III: Lisansüstü Köşesi 17
Bölüm IV: Serbest Köşe… 23
Bölüm V: Tez Künyeleri 25

Başkan'dan

Değerli üyelerimiz ve meslektaşlarımız;

Deneğimiz bünyesi altında çıkan bültenimizin ikinci sayısıyla yine karşınızdayız. Ülke olarak bugünlerde gerçek çok zor ve acı günler geçiriyoruz. Bu yazıyı kaleme aldığım dakikalarda bile yeni acılar yaşanıyor güzel ülkemde. Yaşanan bu acılar karşısında bizlere düşen sabırla ve azimle çalışmaya devam etmektir.

Dernek olarak kurulduğumuz günden bu yana önemli işlere imza attık, atmaya da devam edeceğiz. Derneğimiz 21-24 Nisan 2016 tarihinde İstanbul Kültür Üniversitesi ve PEGEM Akademi işbirliği ile düzenlenen 25. Ulusal Eğitim Bilimleri Kongresinde 20 eğitim derneği ile birlikte kongre düzenleme sürecine aktif şekilde katılmıştır. Bilimsel katkımız ve tanınırlığımız açısından derneğimize önemli katkılar sağladığını düşünmekteyiz.

Geçtiğimiz günlerde Ankara Üniversitesi ile birlikte "Anı Kitabı" projesini hayata geçirdik. Alanımıza ve bilime kakı sağlayan hocalarımızı ve anılarını topladığımız kitap projesi tamamlandı ve düzenlemiş olduğumuz bir kokteyl ile hocalarımıza ve akademik camiaya tanıtımı yapıldı.

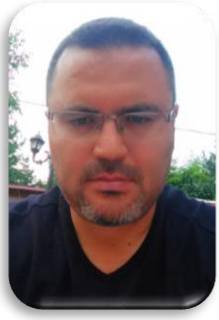
Bizim için ulusal düzeyde en önemli kongrelerden biri olan 5. Ulusal Ölçme ve Değerlendirme kongresi için hazırlıklarımız devam etmektedir. Akdeniz Üniversitesi ile yakın işbirliği içinde uluslararası katılımlı, nitelikli bir kongre için çalışmalara devam ediyoruz. Bu süreçte Anthony J. Nitko ve Jerome D'Agostino kongremize çağrılı konuşmacı olarak destek vereceklerdir. Kongremize katılımlarınızı bekliyoruz.

Dernek çalışmalarında emeği geçen yönetim kurulu üyelerimize, bültenin hazırlanma aşamasında çalışan arkadaşlarımıza ve önerileri ile bizlere destek olan tüm üyelerimize teşekkürlerimi sunuyor, alanımızda çalışan ve henüz üye olmamış meslektaşlarımızı da üye olmaya davet ediyorum.

Prof. Dr. Hülya KELECİOĞLU

EPODDER Başkanı

Editör'den



"İlk söz merhaba olsun…" seslenişiyle ilk sayımızı başlatmıştık. Ne mutlu bize ki ikinci sayımızla karşınızdayız. Önceki sayımızda da bahsettiğimiz gibi yılda iki defa yayınlayacağımızı duyurduğumuz bültenimiz Haziran ve Aralık aylarında çıkmaya devam edecektir.

İlk sayımızı çıkardıktan sonra sizlerden pozitif geri dönütler aldık. Bunlar bizi mutlu etti. Ama maalesef eleştiri ve önerilerinizden mahrum kaldık. Katkı sağlamak isteyen, yazlarıyla destek vermek isteyen, eleştirmek isteyen tüm üye ve meslektaşlarımızı bekliyoruz.

İkinci sayımıza yine dernek çalışmaları ile başlıyoruz. İlk sayımızdan bugüne kadar derneğimizin yapmış/yapmakta olduğu çalışmalardan bir özet sunacağız. Bu sayımızda yurt dışından önemli bir isimle yazılı röportaj planlamıştık. Ancak ilk görüşmeler olumlu olmasına rağmen, maillerimize geri dönüş alamadık. Bu nedenle kısa zaman zarfında bülteni aksatmamak için Adnan ERKUŞ hocamıza başvurarak bir yazı rica ettik. Kendileri de bizleri kırmayarak önemli bir konuyu ele aldığı yazısını gönderdi.

Lisansüstü köşemizde AIBU öğretim görevlilerimizden Ceren MUTLUER ilginç bir konuyu "uygulama güvenirliği"'ni tanıtan bir yazı ile katkı sağlamıştır. Unutulmamalıdır ki bu bölüm lisansüstü öğrenim gören siz değerli öğrencilerimiz ile gelişecektir. Katkılarınızı bekliyoruz…

Serbest köşe bölümünde farklı bir konumuz ve konuğumuz var. Öğrenci değişim programı ile ülkemizde bir özel üniversitede bir dönem eğitim alan bir arkadaşımızın deneyimlerini paylaştığı bir yazımız olacak. Çok önemli dersler çıkaracağımız noktalar var, dikkatinizi çekeceğine inanıyoruz.

Son bölümde ilk sayımızdan bu yana yayınlanmış, YÖK tez tarana merkezinden ulaşılabilen yüksek lisans ve doktora tezlerinin künyelerini paylaştık. Bu bölümde siz üyelerimiz/meslektaşlarımızı yakın geçmişte literatüre kazandırılan tezlerle ilgili haberdar etmeyi amaçlıyoruz.

İlgi ve destekleriniz bekliyoruz…

Doç. Dr. İbrahim Alper KÖSE

Bölüm I: Derneğimiz ve Çalışmalarımız

Bu bölümde üyelerimiz ve meslektaşlarımız için derneğimizin yapmış olduğu çalışmaları tanıtmak istiyoruz.

Derneğimizin bir süredir ulusal meslek standartların hazırlanması ile ilgili çalışmalar yaptığını önceki sayımızda duyurmuştuk. Geçtiğimiz süreçte Ölçme ve Değerlendirme uzmanı (seviye 7), taslak ulusal meslek standardı hazırlanmış olup ilgili taslak kişi, kurum ya kuruluşlara gönderilerek görüş talebinde bulunulmuştur. Gelen görüşlere göre taslak yeniden düzenlenip meslek yeterlik kurumuna eğitim sektör komitesinde değerlendirilmek üzere gönderildi. Şu an eğitim sektör komitesi ile derneğimizin yapacağı toplantı süreci beklenilmektedir. Bu süreç tamamlandığında MYK yönetim kurulunun onayı ile resmi gazetede yayımlanacaktır.

Derneğimiz alanımıza katkı yapan değerli hocalarımızı unutmamış, Ankara Üniversitesi Ölçme ve Değerlendirme Uygulama ve Araştırma Merkezi (ANKÜDEM) ile Eğitimde ve Psikolojide Ölçme ve Değerlendirme Derneği (EPODDER) işbirliğinde, Yrd. Doç. Dr. Ömer KUTLU'nun editörlüğünde hazırlanan "Yaşantılar Evreninden Anılar Örneklemi Ölçmecilerden Geçerli ve Güvenilir Anılar" isimli kitabın tanıtımı ve dağıtımı 11 Nisan 2016, saat 15.00'da Ankara Üniversitesi Rektörlüğü 100. yıl salonunda gerçekleştirilmiştir. Web sitemizden de bilgi edinebilirsiniz. http://www.epodder.org/kitap/

Derneğimiz bünyesinde var olan etik komisyonumuzun hazırladığı etik ilkelerimiz yayımlandı. Amacı, eğitimde ölçme ve değerlendirme etkinliklerinin herhangi birinde faaliyet gösteren EPODDER üyelerinin mesleki çalışmalarına rehberlik etmek olan bu ilkelerin hazırlanmasında emeği geçen Prof. Dr. R. Nükhet DEMİRTAŞLI'ya Yrd. Doç. Dr. Emine ÖNEN'e ve Arş. Gör. Nermin KIBRISLIOĞLU'na teşekkür ederiz. Etik komisyon üyelerimiz ve ilkelerin tam metnini ilgili bağlantıdan edinebilirsiniz:

http://www.epodder.org/wp-content/uploads/2016/06/EPODDER_etik_ilkeler.pdf

Geçtiğimiz süreçte derneğimiz bünyesi altına yayın hayatına devam eden gergimiz EPOD, DOAJ'a kabul almıştır. Bunun yanında dergimizde yayımlanan makalelere DOI numarası verilmeye başlanmasıyla da önemli bir aşama geride bırakılmıştır.

Derneğimizin yapmış olduğu diğer çalışmalar ise kısaca aşağıda özetlenmiştir;

2017 yılında yapılacak olan Uluslararası Eğitim Bilimleri Kongresine, derneğimiz ulusal eğitim dernekleri platformu üyesi olarak katkı vermesine ilişkin protokol ilgili taraflarla birlikte imzalanmıştır.

Kaybetmiş olmamıza rağmen rahmetli M. Fuat Turgut hocamızın bizlere katkısı devam etmektedir. Hocamızın eşi Sayın Perihan TUGUT kitaplığında yer alan doküman ve kitapları derneğimize dolayısıyla alanımıza bağışlamıştır. Bunun yanında hocamızın "Şans Başarısının Test Puanlamasına Etkisi" isimli kitabı dernek tarafından basılmak ve alandaki meslektaşlarımıza dağıtılmak üzere işlemlere baskıya verilmiştir. Kendilerine teşekkür ediyoruz.

Derneğimiz köy okullarına kitap bağışına devam etmektedir. 2015 Kasım ayında İLKYAR vakfı aracılığıyla derneğimiz, Anadolu'da bir köy okuluna kitaplık (100 adet kitap) bağışında bulunmuştur.

Son olarak *1-3 Eylül* 2016 tarihlerinde yapılacak olan V. EPOD kongresi için Akdeniz Üniversitesi işbirliği ile birlikte derneğimiz kongre düzenleme hazırlıklarına devam etmektedir. Uluslararası katılımlı planladığımız kongremize siz değerli üyelerimizi/meslektaşlarımızı bekliyoruz.

Bölüm II: Analiz

"OLMAK YA DA OLMAMAK": ÖLÇEK GELİŞTİRME SÜRECİNİN "GARİBAN KALMIŞLIĞI"

Adnan ERKUŞ

(Az gelişmiş ülkenin az gelişmiş emekli akademisyeni)

Prof. Dr. Adnan ERKUŞ 1992 yılında Hacettepe Üniversitesi Psikoloji bölümünde lisans, 1996 yılında aynı üniversitenin psikometri alanında yüksek lisans öğrenimini tamamlamıştır. Yazar 1996 yılında Ankara Üniversitesi Ölçme ve Değerlendirme ana bilim dalından doktor unvanını almıştır. Mersin Üniversitesi'nden 2012 yılında Profesör unvanı ile emekli olana kadar birçok bildiri, makale, kitap yazmış; ilgili üniversitenin çeşitli birimlerinde yöneticilik yapmıştır. Yazar, evli ve 1 çocuk babasıdır.

Psikolojik ölçme tarihinde, matematik-istatistik ve psikoloji bilimindeki gelişmelere koşut olsa da, 'en önemli tarihsel aşamalar nedir' diye bakıldığında; Galton'un katkıları, fiziksel büyüklük ile algısal büyüklük arasındaki farkları ölçmeye çalışan psiko-fizik çalışmalar, Binet-Simon testinin psikometrik psikoloji anlayışıyla ilk kez geliştirilmesi ve en önemlisi de herhalde Spearman'ın faktör analizi ve diğer çalışmaları, Thurstone'un çok faktör çalışmaları ile Likert-Thurstone-Guttman tipi ölçek geliştirme çalışmalarının olduğu ileri sürülebilir. Guilford ve Gulliksen'in klasik test kuramının temellerini atması çok önemlidir, Lord'un modern test kuramını başlatması da çok önemlidir. Ancak, sanki bu kuramlarla birlikte, psikolojik ölçme asıl köklerinden ve amacından sapmaya başlamış gibidir: ***Ölçeği sağlam geliştirmek ya da örtük psikolojik özelliğin ne ve nasıl olduğunu ortaya çıkarmak***! Bu amaçtan sapmayı, özellikle bilgisayar teknolojisinin daha da hızlandırdığı ileri sürülebilir. Ama, köklerimize dönersek, bilgisayar teknolojisindeki gelişmeleri *asıl amacımız* için çok iyi bir şekilde kullanabiliriz. Sevindirici olan ise, McDonald ile birlikte, sanki tekrar "işin özüne" doğru dönmeye başlıyoruz *gibi*…

Gelin biraz aykırı sorular ve düşüncelerle işe başlayalım:

- Her türlü eşitliğin sağ tarafı, daima eşitliğin sol tarafını doğrulamak için bağımsız ve/veya yordayıcı değişkenler ("faktörler"; "faktör"ü iki anlamda kullanıyoruz; "etmen" ve "psikolojik yapı") ve onların bileşenlerinden oluşur; aksi halde zaten eşitlik olmaz. Öyleyse, eşitliğin sağ tarafını manipüle ederek yapılan her türlü çalışma, eşitliğin sol tarafını kanıtlamaya –zorunlu olarak- hizmet eder: Kendi kendini doğrulayan kehanet! Bu durum sizce de rahatsız edici ve bir kısır döngü değil mi? Örneğin,

  $$KR_{20} = \frac{n}{n-1}\left(\frac{S_X^2 - \sum pq}{S_X^2}\right) = \frac{n}{n-1}\left(1 - \frac{\sum pq}{S_X^2}\right)$$

  Bu formülde, "madde sayısı, madde güçlüğü veya toplam test varyansı değişirse, ne olur" diye merak edip, bu değişkenleri manipüle edersek, KR-20 *kesinlikle değişecektir*; KR-20'nin başka çaresi var mı?

- Tüm psikometri kitaplarında ve tüm ölçme kuramlarında, "*ölçme sonucuna göre,… ölçmeye karışan hatalar,…"* vb diye başlanır; madde istatistikleri, parametreleri hesaplanır, güvenirlik kestirimleri yapılır ve hatta bazı sayıltılarla ölçme kuramları oluşturulur vb. Kısacası, hangi ölçme kuramı olursa olsun, hangi güvenirlik yaklaşımı ve işlemi olursa olsun, her şey "***gözlenen puanlar***"dan başlar, sayıltılar, kestirimlerle vb devam eder. Peki bu "ölçme"den (gözlenen puanlardan) kastedilen nedir? "***Ölçek geliştirme aşamasındaki ölçme***" mi, "***ölçek geliştirildikten sonra, bu ölçme araçlarıyla ölçülen bireylerden elde edilen ölçmeler"*** mi? [Hani, son yıllarda "moda" bir soru gündemde ya… "*Ölçme aracı mı, ölçme sonuçları mı –aynı şey olmasa da puanlar mı- güvenilir*"? Sahi *biz* psikometristler hangisiyle ilgileniyoruz? Elbette ikisiyle de, ama öncelikle bizim için hangisi önemli, *ölçme aracı sağlam değilse ölçme sonuçları ne olur*?] Bu konuda açık bir kaynağa rastlamak hemen hemen olanaksız olmakla birlikte, kastedilenin çoğunlukla "geliştirilmiş bir ölçekle elde edilen ölçümler" olduğu veya hangisinin kastedildiğinin belli olmadığı yönündedir. Peki ölçek geliştirme aşamasındaki ölçümler ile geliştirilmiş bir ölçekten elde edilen ölçmelere aynı işlem ve mantık uygulanabilir mi? Bunlardan hangisinin temel alındığı, sayıltıları da, işlemleri de değiştirir mi veya sayıltı ve işlemlere gerek bırakır mı?

- Hata türlerini bilmek, tanımlamak güzel. Ama, asıl önemlisi hata kaynaklarını bilmek ve kontrol etmek ve elimizdeyse (ki o zaman, "hata", "*yanlış*" olur) onları hiç yapmamak değil midir? Ama hangi aşamadaki hata kaynakları? Daha "ölçme sonuçları"ndan, hangi aşamadaki elde edilen ölçme sonuçlarından söz ettiğimiz üzerine bile anlaşmaya varmış değiliz, tartışmış değiliz. Ölçek geliştirme süreci sağlıklı değilse, gerisi ne işe yarar?

- Örneğin, "*maddeler paralel, eşdeğer, eşbiçimli*" vb bir sayıltı mı olmalı, yoksa görgül verilere dayanan (hangisi ise o!) istatistiklere ve anlayışına mı dayanmalı? Peki maddeler, paralellikten-eşdeğerlilikten uzaklaşıyorsa, bu ölçeğin ölçmeyi amaçladığı yapısı için ne anlama gelir? Ölçek geliştirme sürecinde maddeler hangi amaçla neyi uyarmak için yazılıyor? Maddelerin paralel, eşdeğer, eşbiçimli vb olup olmamasına göre türetilen güvenirlik katsayıları arasındaki 0,01'lik farklılıklar "*ölçek kullanıcıları*" için pratikte ne anlam ifade eder, ama asıl önemlisi altta yatan ölçme sayıltıları açısından ne fark eder? Sahi sayıltıyı neye dayanarak türetmiştik, ölçeğin ölçtüğü psikolojik yapıya göre mi, *a priori* oluşturulan yapı ve maddelere göre mi, görgül verilere göre mi, yoksa sadece "kurama" göre mi? Peki, "*ölçme sonuçları*"nı *neyle hangi aşamada* elde ettik?
- Bir madde, ölçülen özelliğin uzaydaki yapısı gereği hem A boyutuna hem de B boyutuna (faktör analizi anlamında) yük veriyorsa veya aralarında pozitif ve dikkate değer korelasyon varsa (her ikisi de aynı kapıya çıkar ya, neyse), *ölçeğin ölçtüğü yapı gereği* mi öyle "çıkmakta", yoksa "*maddede bir sorun var*" sonucuna mı ulaşılacak (pratikte pek çok madde bu nedenle 'atılmaktadır'), yoksa, daha önemlisi ise "*dikkate aldığımız*" (*siyasi parti tutar gibi*) ölçme kuramına *bağlı* mı kalacağız? Yoksa kurduğumuz neden –sonuç zincirinde bir terslik mi var?

Hangi ölçme kuramına, yöntemine veya tekniğine dayanırsa dayansın, psikolojik ölçme araçları (her tekboyut için) birden çok maddeden oluşur ve sonuçta madde puanlarının birikici özelliğinden dolayı toplam puan birey hakkında karar vermede kullanılır. Bu bırakalım ölçmeyi vs.yi doğanın işleyiş yasası gereği böyledir: ***Aritmetikte, tek işlem vardır, o da toplanabilirliktir***! Toplanabilirlik ise, maddelerin aynı boyutta-vektörde yer almasıyla olanaklıdır. Çokboyutlu bir uzay, toplanabilirliği ihlal etmez; çünkü yine her altboyutun kendi içinde toplanabilirliği söz konusudur. Kaldı ki, psikolojik özelliklerin pek çoğu çokboyutludur. Sadece vektörler (boyutlar) arasındaki açı 90^0 olması durumunda (her bir vektörün (boyutun) değil), vektör*ler*in (boyutların) toplanabilirliğinden söz edilemez, çünkü ilişkisiz iki altboyut ile karşı karşıyayız demektir (tıpkı, elma+?armut gibi). Vektörler (boyutlar) arası açı birbirine yaklaştıkça, çokbileşenli (composite-faktörlü değil) ve toplanabilir yapıya doğru gidilir ki, bu maksimum açının ne olduğu (açının cutoff'u) ölçmecilerin, *kuramlardan önce hâlâ çözmeleri gereken en acil ve ciddi bir sorundur*. Ölçme kuramlarında gerek Spearman'ın genel yeteneği (G) ölçme girişimlerinden gerekse de çoğunlukla eğitimde başarı odaklı temellenmesinden sanki ölçülen hep tekboyutluymuş gibi bir algı oluşmuştur: Oysa, toplam puan elde etmek için tekboyutlu olma gerekliliği başka, ölçülenin çokboyutlu olması veya daha doğrusu boyutlarının neler olduğunun ortaya çıkarılması başka şeylerdir. *Maddeler iyi hazırlanmış, iyi veri toplanmış ve iyi analiz edilmişse, "o" yapı içindeki faktör yükleri ± 1,00'e (aynı yönde olmak koşuluyla) ve klasik madde ayırt edicilikleri +1,00'e, madde karakteristik eğrisi 90º'ye yaklaşır; bu da "o" boyutun pür (simple) tekboyutluluğa yaklaşması demektir. Böyle olursa ne olur, bu ne anlama gelir?* Peki, ama, madde veya ölçeğin ayrık veya bileşenli "çok boyutlu" olduğunu ne zaman ve neden irdelememiz gerekir; ölçek geliştirme aşamasında mı, yoksa bireyler hakkında değerleme (as-

sessment) yapma aşamasında mı? *Hangi aşamada*, bu "bilgiler" daha yaşamsaldır ve işlevseldir? Açıktır ki ölçeğin geliştirilme süreci bir önkoşuldur. Ölçeğin ölçmeyi amaçladığı yapıya uygun olmayan, uygun olsa da uygun üretilmeyen ölçek yapısı ve maddelerine KTK uygulansa ne olur, MTK uygulansa ne olur? Öyleyse "püf" noktası ne olabilir? En uygun psikolojik yapıyı ortaya çıkarmışsanız, madde analizi (ki, MTK'nın ek bilgilerinden başka) veya diğer işlemleri yapsanız ne olur? İlgili faktöre yüksek yük veren madde (o faktöre hizmet ediyor, o faktöre sahip olanlar ile olmayanları ayırt ediyor demektir-McDonald, 1999), KTK'na göre de ayırt edicidir, MTK'ya göre de (Fan, 1998); asıl önemlisi de maddenin o alt yapının önemli bir parçası olmasıdır. Bu bakımdan, ölçek geliştirme sürecinin ciddi bir şekilde 'masaya yatırılması', kuramlara gösterilen titizliğin bu sürece harcanması öncelikli sorunumuz olmalıdır.

Bu sürecin temelini, ***ölçeğin kavramsal yapısının iyi tanımlanması ve kurgulanması*** oluşturur. Aşağıdaki birkaç örneği inceleyerek bunun önemini görelim (*Alt çizgiler, vurgu yapmak için yazar tarafından çizilmiştir*):

"Madde 1: Duygusal sorunlarımı başkaları ile ne zaman paylaşacağımı bilirim

Madde 4: Bir sorunu çözmeye çalışırken ruh halimden etkilenmem"

1. Duygularımı tanıyorum. Üzüntümün ve korkumun farkındayım. Üzüldüğüm zaman sinirlendiğimin, korkunca saldırganlaştığımın farkındayım.

Evet Bazen Hayır

2. Duygularımı ifade ediyorum. Üzüntümü mizahla veya güçlü görünerek veya olduğu gibi dile getirebiliyorum. Gülüşüm sahte değildir.

Evet Bazen Hayır

Yukarıdaki örnek maddeler, ülkemiz kültürüne de sözde "uyarlanmış" iki ayrı "duygusal" zekâ "ölçeğine" ait maddelerdir. Aşağıdakiler de yine kültürümüze sözde "uyarlanmış" (aslında orijinalleri de öyle) "çoklu" zekâlar "ölçeği"nden bazı maddeleri içermektedir:

-*"İsimler, yerler, tarihler konusunda belleğim iyidir."*

- *"Kitap okumayı severim."*

- *"Kelimeleri doğru şekilde telaffuz ederim."*

- *"Aritmetik problemleri kafadan hesaplarım.*

- *"Matematik ve fen derslerinden hoşlanırım."*

-*"Doğaya karşı duyarsız olanlara kızarım."*

Zekâ bir sınır yeterlik(maksimum performans)tir. *Yukarıdaki gibi* bir kendini-değerlendirme (self-report) ölçeğiyle sınır yeterliğin ölçülmesini bir yana bıraksak bile, maddelerin uyardığı

psikolojik özellik, zekâ mıdır, ilgi, tutum vs midir? Elbette, zekâ olmadığı, olamayacağı çok açıktır. Şimdi, bu maddelerin a ve b parametrelerini MTK'na göre ortaya çıkarsak ne olur ve ölçeğin güvenirliğini çeşitli Cronbach katsayılarıyla veya test bilgi fonksiyonunuyla bulsak ne olur? Şunu hemen belirtelim, gayet iyi çıkardı ve çıkıyor da; çünkü, *"bunlar" aslında zekâyı falan değil bireyin kendisine ilişkin algısını (self-perception) "ölçmek"tedir(!)* ve bu nedenle tek faktörlüdürler ve iç tutarlık katsayıları da 0,90 ve üzerindedir! Ne olacak şimdi? Güvenilir mi güvenilir! Neye göre, neyin güvenirliği? "Ölçmeye" karışan hataları şu ya da bu kurama göre tanımlasak ve ona uygun analizler yapsak bir anlam ifade eder mi? Öyleyse, psikometristlerin önündeki asıl yakıcı sorun hangisidir; ölçek geliştirmenin temelleri mi, yoksa artık teknik olarak içinde "boğulunulan" kuramlar ve işlemler mi? *Böyle bir soru, ölçme kuram, model ve işlemlerinin işe yaramadığı sonucunu çıkarmaya yol açmamalıdır elbette, ama hangisi öncelikli sorunumuzdur*?

İkinci önemli aşama, örtük özelliğin somut davranışa dönüştüğü ***davranışsal göstergelerin iyi örneklenmesi***dir. Kavramsal yapı çok iyi yapılandırılmış olabilir; ancak, onun hangi koşul, ortam ve durumlarda gözlenebilir olduğu bulunamazsa, madde üretimi yine doğru olmayan temellere dayanmış olur. ***Örtük özelliğe uygun madde üretilmesi***, bir o kadar önemlidir. ***Deneme uygulaması koşullarında ölçmeye karışabilecek hataların kontrol edilmesi, bertaraf edilmesi*** yine büyük bir titizlik gerektirir. ***Ölçeğin yapısının ortaya çıkarılması*** ise üzerinde en çok durulması gereken aşamadır ve psikometristlerin bu aşamaya ilişkin derinlemesine çalışmalar yapmak gibi acil görevleri bulunmaktadır. Örneğin, klasik kitaplarımızda yazdığı gibi, madde-toplam test/ölçek korelasyonlarına veya alt-üst grup ortalamaları karşılaştırmalarına göre madde atarak ölçek geliştirmeye çalışmak doğru mudur? Hayır. Çünkü, bu madde analizi yöntemleri tekboyutlu yapılar için uygundur; daha yapısı ortaya çıkarılmamış hangi taslak ölçek madde yığınına bu yöntemleri uygularsanız uygulayın, bu şekilde madde ata ata her türlü psikolojik yapı tekboyuta indirgeniverir! Nerede kaldı, psikolojik yapının işevuruk tanımı? Hatta bir kişilik envanteri olan MMPI'ın tüm maddelerine bu şekilde analiz yapın, "kişilik tekboyutludur" sonucuna bile ulaşırsınız! Bu konular, Psikolojide Ölçme ve Ölçek Geliştirme kitabının ikinci cildinde incelenecektir.

Spearman hem KTT, KökTK (congeneric=kökdeşlik), hem de MTK'nın temelini oluşturur; neden? Çünkü, temel hep ölçeği oluşturan maddelerin kalitesi ve en önemlisi de "ölçeğin" (aslında *gerçek* psikolojik yapının) yapısı ile ilgilidir. Bunun temel yolu da Açıklayıcı Faktör Analizinden geçmektedir. KTT'ye göre ayırt edici olan bir madde, MTK'de de ayırt edicidir; KTT'ye göre güvenilir olan bir test/ölçek, MTK'ya göre de güvenilirdir. İster madde kovaryanslarına isterse faktör yüklerine (ilk faktörün özdeğerine vs) dayansın, *hepsi özünde maddelerin ilgili özelliği ölçme yeterliğine göre hazırlanıp hazırlanmadığına, deneme uygulamasına karışan hatalara, deneme uygulaması verilerine yapılan ölçek ve madde analizlerine, tepkide bulunanların büyüklüğüne (n) ve varyansına ve en önemlisi de maddeler arası ve bireyler arası varyansa bağlıdır*. O zaman neyin peşindeyiz? **Ölçmek istediğimiz ve doğrudan gözlenemeyen psikolojik (örtük) yapı ne, nasıl?** *Bunu bir sarraf titizliğiyle ortaya çıkarmadıktan sonra, ölçme kuramları ve dayandıkları sayıltılar ile onlara dayanarak türetilen güvenirlik işlemleri*

ne anlam ifade eder? ***Hangisi asıl sorunumuz; öncelikle*** *a priori* ***öngörülen psikolojik yapının ne ve nasıl olduğunu anlayıp tanımlamak mı, ölçme-güvenirlik kuramlarımızı kendi sayıltılarımızla ürettiğimiz formülleri doğrulamak mı?***

Güvenirlik ve geçerlik irdelemelerini de hatalar ve amaca hizmet açısından her iki aşama için ayrı irdelemeliyiz. Ölçek geliştirmenin denemelik uygulama aşamasındaki güvenirlik-geçerlik çalışmalarını, *gerçek* güvenirlik-geçerlik olmaktan çok, *ortaya çıkarılmaya çalışılan yapıya dolaylı destek amacıyla yapılan* ***öngüvenirlik-öngeçerlik*** çalışmaları; ayrı bir örneklemde ve seçilmiş-ortaya çıkarılmış ölçek ile toplanan kanıtları da ***gerçek güvenirlik-geçerlik*** kanıtları olarak ele almak durumundayız. *Deneme uygulamasına dayanarak* yapılan madde analizleri, güvenirlik ve geçerlik işlemleri, ortaya çıkarılmaya çalışılan yapıya ek kanıt sağlama işlemleri olarak görülmelidir. Zaten Açıklayıcı Faktör Analitik tekniklerle yapılan titiz çalışma sonucu ortaya çıkan yapının alt yapılarının diğer madde analizi göstergeleri de iyi görünecek ve içtutarlıkları da yüksek çıkacaktır; ölçek yapısı iyi bir şekilde ortaya çıkarılmamışsa bu göstergeler de zayıf çıkacaktır. Bu bakımdan, bu süreçteki hatalar ve *yanlışlar*, kuramların dayandığı hatalardan farklı bir şekilde ele alınmak ve (istatistiksel işlemlerdeki etkilerini dikkate alıp belirlemekten çok) hemen hiç yapılmamasının-ölçümlere karışmamasının önlemleri ele alınmak durumundadır. Bu aşama, olmazsa olmaz önkoşul bir aşamadır. Son zamanlarda moda olan, bir programın tuşlarına basarak bir DFA analiziyle ölçek geliştirdiğini sananları konu dışı bırakıyorum; çünkü, henüz bilinmeyen-ortaya çıkarılmamış olan bir yapının doğrulaması da geçerliği de olmaz! Ancak, ölçek geliştirme/uyarlama çalışmalarının bu denli basite indirgenmesinde, biz psikometristlerin bu aşamayı yeterince ayrıntılı incelememe, anlatmama (ve belki de bilmeme) gibi eksiklerinin payı olduğunu da göz ardı edemeyiz. Ölçek geliştirme/uyarlama çalışmaları günümüzde ne yazık ki pek çok akademisyen için kolay puan alma yolu olarak görülmekte ve ülkemiz giderek bir "ölçek çöplüğü"ne dönüşmektedir. Bu alanda yetişmenin ne kadar güç olduğunu, bu alandaki bilgi birikimini hiçe saymayı veya hafife almayı bir yana bıraksak bile, geliştirildiği sanılan ölçeklere dayanarak insanlar hakkında çok ciddi kararlar verildiğini, onlara dayanarak yapılan araştırmalarla bilgi üretildiğini, yanlış ölçmenin bireysel ve toplumsal olarak yüksek "maliyetini" kimse hafife alamaz.

Her türlü psikolojik değişkeni ölçmek için bir araç geliştirme, o psikolojik yapı ile ilgili bir denence ile başlar; ölçek geliştirme süreci özünde o denenceyi test etme sürecini ve ilgili psikolojik değişkenin işevuruk olarak tanımlanmasını içerir. Kuramlar, sayıltılar ve onlara dayalı işlemler (her ne kadar içinde görünüyorsa da) sonraki işlerdir; psikolojik yapı ne kadar sağlıklı ortaya çıkarılırsa, kuramlar da o kadar işlevini yitirir, daha doğrusu aynılaşır (zaten aralarında istatistiksel geçişler vardır). Önce örtük yapı (ki faktör analitik tekniklerden başkası uygun görünmüyor) ortaya çıkarılmalı, eğer çeşitli madde özellikleri ve ileri uygulamalar yapılacaksa da öncelikle bu yapıya katkı sunması için kullanılmalıdır. Ölçek geliştirme/uyarlama süreci sağlıklı yapılmazsa, ölçme kuramları birer "fantezi" olmaktan öteye geçemez.

"Ağaca bakarken ormanı, ormana bakarken ağacı" görebilmeliyiz; aksi halde, "*havanda su dövmeye*" devam ederiz.

Kaynaklar

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychologial Measurement, 58*(3), 357-381.

McDonald, R. P. (1999). *Test theory: Unified treatment.* Mahwah, New Jersey: Lawrence, Erlbaum Assoc. Pub.

SİVİL TOPLUM KURULUŞLARINDA ÖLÇME VE DEĞERLENDİRME ÇALIŞMALARI

Sonnur ÖZEL

1976 yılında Ardahan'da doğdu. Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme ABD Lisans programından 2001 yılında mezun oldu. Yüksek Lisans çalışmasını yine aynı üniversitenin Sosyal Bilimler Enstitüsü'nde tamamladı. Ahmet Yesevi Üniversitesi'nde Ölçme Değerlendirme Uzmanı olarak başladığı mesleki yaşamına, çeşitli özel okullarda Ölçme Değerlendirme Uzmanı ve Birim Yöneticisi olarak devam etti. Halen Öğretmen Akademisi Vakfı'nda Uzman Eğitimci ve Ölçme Değerlendirme Uzmanı olarak çalışmaktadır.

Öncelikle ölçme değerlendirme camiasına gönülden bir merhaba,

Değerli hocam Doç. Dr. İ. Alper Köse Ölçme Değerlendirme Bülteni için sivil toplum kuruluşlarında yürütülen ölçme değerlendirme çalışmalarına ilişkin deneyim ve gözlemlerimi anlatan bir yazı hazırlamamı istediğinden beri heyecan içindeyim.

Meslek yaşamımda çeşitli üniversite ve özel okullarda çalıştıktan sonra son bir yılımda çalışmalarını yürekten desteklediğim bir sivil toplum kuruluşlu olan Öğretmen Akademisi Vakfı'nda çalışmaya başladım.

Öğretmenlerin ve ülkede eğitim sorumluluğunu üstlenmiş tüm bireylerin mesleki ve kişisel gelişimlerine destek veren projeler oluşturup yürütmek hedefiyle kurulan Öğretmen Akademisi Vakfı (ÖRAV) ülkemizde bu alana odaklanmış ilk ve tek sivil toplum kuruluşu.

2008'de kurulan Vakıf, 7 yılda Türkiye'nin tüm illerinde 130,000'den fazla eğitimcinin kişisel ve mesleki gelişimlerini sürdürebilmeleri için destek vermiştir. Ülkemizde düşünen, sorgulayan ve araştıran bir eğitim modelinin uygulanmasına katkıda bulunmak amacıyla yürüten projeler devlet okullarına ücretsiz olarak ulaştırılmaktadır.

Sivil toplum kuruluşlarına yönelik literatürdeki tanımlamalara baktığımızda; belli amaçları gerçekleştirmek, belirli konularda kamuoyunu aydınlatmak veya yönlendirmek için çalışan ve gönüllülük esasıyla hareket eden kuruluşlara "Sivil Toplum Kuruluşları" denildiğini görüyoruz. Sivil toplum kuruluşları aynı zamanda üçüncü sektör olarak tanımlanmaktadır. Üçüncü sektör kuruluşları, devletin (kamu sektörünün) yetersiz kaldığı alanlarda topluma hizmet götürmek

amacıyla ortak hedefe yönelmiş insanlar tarafından kurulan ve kâr amacı gütmeyen yapılardır (aktaran Yurtseven, 2015, s. 114).

Sivil toplum kuruluşlarının faaliyet alanları; toplumun ortak sosyal problemlerinden, sağlık konularına; çevre sorunlarından, insan – kadın - çocuk haklarına; mesleki ya da kişisel gelişime yönelik çalışmalara kadar çeşitlilik göstermektedir.

Sivil toplum kuruluşları yürüttükleri faaliyetlerin hedef kitle üzerinde yarattığı değişimi belirlemeye yönelik ölçme değerlendirme çalışmalarıyla, yürüttükleri projelerin farklı insanlar / gruplar üzerinde oluşturduğu etkileri / sosyal faydayı belirlemeye çalışılırlar.

Ölçme değerlendirme çalışmaları bu ana hedefin yanı sıra;

- Yürütülen çalışmaların devamlılığını sağlamak,
- Yapılacak yeni çalışmalara karar vermek,
- Çalışanları ve proje destekçilerini motive etmek,
- Daha fazla insanın çalışmalara ilgi duymasını sağlamak,
- Şeffaflık ve hesap verebilirlik yoluyla paydaşlar, bağışçılar ve yatırımcılarla güven ilişkisi kurulmasını sağlamak,
- Kuruluşun güvenilir bir marka haline gelmesini sağlamak

gibi diğer hedeflere de önemli katkılar sunar.

Bu noktada yürütülecek faaliyetlerin ya da projenin ölçme değerlendirme planının oluşturulması ilk adımdır. Ölçme değerlendirme planı oluşturulurken yapılması gereken ilk çalışma değişim teorisinin belirlenmesidir. Değişim teorisi yürüttüğünüz tüm çalışmalar sonucunda ulaşacağınız nihai hedefinizdir. Bu nihai hedef olabildiğince net ve gerçekçi bir biçimde tanımlanmalı, kimleri kapsadığı belirlenmelidir. Daha sonra bu nihai hedefi gerçekleştirmek için ulaşılması planlanan alt hedefler oluşturulmalıdır. Bu hedeflere ulaşıldığını gösteren kanıtların / göstergelerin neler olduğunun belirlenmesi ve bu kanıtların hangi yol ve yöntemlerle toplanılacağı projenin en başında planlanmalıdır.

Yürütülen faaliyetlerin hedef kitle üzerinde yarattığı değişimi belirleme çalışmaları beraberinde pek çok zorluğu da barındırmaktadır. Bu zorlukların en temel olanlarından biri; oluşan etkinin yürütülen çalışma ile ilişkisinin tam olarak ortaya konulamaması durumudur. Yani projeden yararlanan kişilerde ortaya konan değişiminin ne kadarı proje faaliyetlerinin sonucudur? Bu sorun ölçme değerlendirmenin temel konularından biri olan geçerlikle ilgilidir.

Bu noktada kimi sivil toplum kuruluşlarının yaşadıkları zorluklar ya da başka nedenlerden dolayı sonuç odaklı yaklaşım (outcomes) yerine çıktı odaklı yaklaşımı (output) benimsendiği görülmektedir. Hazırlanan ölçme değerlendirme raporlarında proje faaliyetlerinde sağlanan hizmetlerin ya da uygulanan faaliyetlerin (eğitim sayısı, atölye sayısı, ulaşılan öğrenci sayısı vb.) çıktılarını ortaya konduğunu görülmektedir. Oysaki yürütülen faaliyetlerin yarattığı değişime odaklanmak (örneğin; verilen eğitimlerden faydalanan kişiler hangi becerilerini geliştirdiler ve

bu durum yaşamlarına nasıl katkı sağladı gibi) yani sonuç odaklı yaklaşımı benimsemek doğrudur.

KOÇ Üniversitesi Sosyal Etki Forumu tarafından 89 sivil toplum kuruluşu ile yürütülen "Türkiye'de Sivil Toplum Kuruluşlarının Sosyal Etki Ölçümleme Algısı ve Pratiği" (2015) adlı araştırmada sivil toplum kuruluşlarının %36'sı, yürüttükleri çalışmaların hedef grup/gruplar üzerindeki etkisini ölçmek istediklerini, bunun için gerekli becerilere sahip olduklarını ancak zaman ve / veya kaynakları olmadığını belirtmişlerdir. Katılımcı sivil toplum kuruluşlarının %26'sı da bu çalışmayı yapabilecek beceriye sahip olmadıklarını ifade etmişlerdir. Yanıtların dağılımı aşağıdaki grafikte sunulmuştur.

Yine aynı araştırmada katılımcılara "İzleme – değerlendirme ve / veya sosyal etkilerini ölçmede karşılaştıkları zorluklar" sorulmuştur. Aşağıdaki grafikte görüldüğü gibi katılımcı Sivil Toplum Kuruluşlarının %56'sı "Ölçümleme için yeterince uzmanlık ve becerilere sahip olunmaması" seçeneği işaretlemişlerdir. Bu oran ülkemizde sivil toplum kuruluşlarının faaliyetleri ile ilgili olarak ölçme değerlendirme konusundaki uzmanlık ihtiyacına dikkat çekmektedir.

Öğretmen Akademisi Vakfı'nda çalışmaya başladığımdan itibaren sivil toplum kuruluşlarındaki ölçme değerlendirme çalışmalarına ilişkin çeşitli konferans ve atölyelere katılma fırsatı buldum. Bu etkinlikler benim için çalıştığım bir alanın yeni bir terminolojiyle ifade edilişinin yanı sıra sivil toplum kuruluşlarının bu konuda ne kadar az deneyime sahip olduğunu da görmemi sağladı. Yukarıda bu süreçte edindiğim bilgi ve gözlemlerden söz etmeye çalıştım.

Bir Sivil Toplum kuruluşu olarak Öğretmen Akademisi Vakfı'nda da kurulduğu yıldan itibaren uygulanan projelerin katılımcıların (öğretmenlerin) mesleki ve kişisel yaşamında nasıl bir değişim yarattığını belirlemeye yönelik çalışmalar planlanmış ve uygulamaya konulmuştur. Eğitimlerin hedefleri doğrultusunda eğitimler öncesinde ve sonrasında çeşitli anketler / ölçekler aracılığıyla katılımcı öğretmenlerin ve onların öğrencilerin görüşleri alınmaya, eğitimin sınıf ortamına nasıl yansıdığı belirlenmeye çalışılmaktadır. Bu çalışmalardan elde edilen sonuçlar çeşitli akademik platformlarda eğitim dünyasıyla paylaşılmaktadır.

Bununla birlikte sivil toplum kuruluşlarının izleme değerlendirme çalışmaları sırasında yaşadıkları çeşitli zorlukları zaman zaman biz de yaşıyoruz.

Son söz olarak, yaşanan bu zorlukları en aza indirmek adına, sivil toplum kuruluşlarının ölçme değerlendirme çalışmalarında akademik çevrelere duyduğu desteğin altını çizmek isterim. Yapılacak işbirlikleri sivil toplum kuruluşlarına yürüttükleri faaliyetlerin sosyal etkisini belirlemede yol gösterirken, akademik çevreler için de eğitim bilimleri adına bir laboratuvar olanağı sunacaktır inancındayım.

Kaynaklar

Yalçın, A.S. ve Güner, D. (2015). *Türkiye'de sivil toplum kuruluşlarının sosyal etki ölçümleme algısı ve pratiği.* İstanbul: KUSİF YAYINLARI

Yurtseven, M. (2015). *Bilgi toplumu olma yolunda kütüphane ve sivil toplum kuruluşu etkileşimi: Türk Kütüphaneciler Derneği örneği (TDK).* Yayımlanmamış yüksek lisans tezi, Marmara Üniversitesi, İstanbul.

Bölüm III: Lisansüstü Köşesi

UYGULAMA GÜVENİRLİĞİ

Ceren MUTLUER

1990 Bolu doğumlu olan yazar, Abant İzzet Baysal Üniversitesi, Fen Bilgisi Öğretmenliği lisans ve Eğitimde Ölçme ve Değerlendirme yüksek lisans tamamlamıştır. Yazar halen Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme doktora programına devam etmektedir. Yazar Abant İzzet Baysal Üniversitesi Eğitimde Ölçme ve Değerlendirme Ana Bilim Dalı'nda öğretim görevlisi olarak çalışmaktadır.

Giriş

Bilimsel araştırmalarda, araştırma planına bağlı kalarak çalışmanın amacına en uygun deseni seçmek önemlidir. Uygun ölçme araçları kullanarak veri toplayıp, merak ettiği sorulara cevap aramak istemektedir. Bu süreçte tasarlanan desene uygun olarak toplanmış verilerin, çeşitli ölçme araçlarında elde edilen ölçümlerin güvenirliğine ilişkin bilgiler raporlaştırılmasına rağmen, araştırma sürecindeki güvenirlik kavramına ilişkin nasıl bir uygulama yapıldığı genellikle atlanmakta ya da çok az çalışmada değinilmektedir.

Bilimsel araştırma sürecinde araştırmacının ve ölçme aracına yanıt veren katılımcıların psikolojik ve fizyolojik durumu davranış ölçümünde farklı sonuçların elde edilmesine sebep olmaktadır. Bu durum kullanılan ölçme aracının güvenirlik düzeyinden bağımsızdır Bear, Wolf ve Risley (1968)'e göre davranış ölçmelerinde gözlemlenen değişikliğin deneysel araştırmalardaki katılımcıların davranışlarındaki yansımalarındaki değişiklerin raporlandırılması gerektiğini belirtmiştir. (Akt. Billingsley, White, Munson, 1980, s. 229). Tutum araştırmacıları Bear, Wolf ve Risley (1968), gözlemci (puanlayıcı) güvenirliğinin önemine işaret ederek, yüksek düzeyde veri toplama yöntemi ile ortaya çıkacak ölçmedeki potansiyel hataya dikkat çemiştir. Bununla ilgili olarak, sadece şu anki tutum araştırmalarında kullanılan çoklu güvenirlik kontrolleri ve rapor edilen araştırmalardaki puanlayıcılar arası güvenirlik değil ayrıca puanlayıcı güvenirliklerinin de değerlendirilmesine vurguda bulunulmuştur (Akt. Billingsley, White, Munson, 1980).

Deneysel araştırmalarda araştırmacıların manipülasyonu sonucu bağımsız değişkenlerin bağımlı değişken üzerindeki etkileri incelenir. Araştırma sürecinde uygulayıcı, yöntemin etkililiği için katılımcılara yöntemi tanıtır. Araştırmacı, süreç boyunca katılımcılardan dönütler alarak bağımlı değişkenin bu bağımsız değişken(ler) lehinde anlamlı bir farklılığın olup olmadığını gözlemler. Bu süreç içerisinde katılımcılardan alınan puanlar bağımsız değişken(ler)in bağımlı değişkenler üzerinde anlamlı bir etkisi bulunmamışsa şu sorular ön plana çıkar: *"Bu çalışmada araştırmacı mı etkisiz oldu yoksa yöntem mi?"* ve *"Davranışta gözlenen değişimlerin nedeni gerçekten deneysel değişken mi? Bu değişime ne denli inanılabilir?"*. Bu sorular araştırmacıların deneysel araştırmalarda uygulama güvenirliğinin önemine dikkat çekmektedir.

Uygulama Güvenirliği

Uygulama güvenirliğine ilişkin araştırmacılar bir takım tanımlar yapılmıştır. Uygulama güvenirliği, öğretim planının ne ölçüde planlandığı ve uygulandığını değerlendiren güvenirlik hesaplamasıdır . Uygulama güvenirliği, uygulamacının ya da öğretmenin uygulama planına ne ölçüde bağlı kaldığının analiz edilmesidir (Billingsley, White ve Munson, 1980).

Yapılan araştırmalarda genellikle öğrencilerin çeşitli durumlardaki performanslarını çeşitli sınıf değerlendirme sistemleri ile belirtmeyi amaçlamıştır. Lindsley (1970) yetersiz operantının analizini şu şekilde göstermiştir:

E^A ----------M----------A----------E^c

(Öncüller/Girdiler) süreç düzenleme sonuç

E^A (girdiler) araştırma sürecinde planlanan hedef davranış konularını meydana getiren durumları; M (süreç) hedef davranış konularını; A(düzenleme) davranış oluşumlarını ve davranış sonuçları arasındaki ilişkiyi belirtmekte ve sonuç davranışları hedef davranışlarını takip eden olayları işaret etmektedir (Akt. Billingsley, White ve Munson, 1980, s. 232). Uygulama güvenirliği, araştırma planına uyularak yukarıda sunulan girdi, süreç, düzenleme ve sonuç aşamalarının gerçekleştirilme düzeyine ilişkin dönüt vermeyi esas alır.

Uygulama güvenirliği, sürecin başından itibaren araştırmacının yönteme ilişkin her bir davranışı ve bu davranışların yer aldığı planda durumlara ilişkin her bir doğru ya da hatalı sonuçları kapsar (Reed ve Codding, 2014). Açıklanmaya çalışılan kavram öğretimin ne ölçüde uygulandığı ile ilgili bilgi almak amacıyla gerçekleştirilmektedir (Tekin ve Kırcaali-İftar, 2001). Uygulama güvenirliğinde, araştırmada yer alan bağımsız değişkenin ne derece planlandığı gibi ne derece uygulandığını da belirlemek amaçlanmıştır. Diğer bir deyişle, uygulama güvenirliğinde planlanan uygulamanın ne kadar doğru olarak uygulandığına bakılmaktadır.

Uygulama Güvenirliğinin Belirlenmesi

Uygulama güvenirliğinde katılımcılar, araştırmacı, uygulamacı ve gözlemciler sürecin önemli elemanlarıdır. Uygulama güvenirliğinde araştırmacının yöntem hakkında yeterli bilgi donanımına sahip olması ve süreçte bu bilgileri uygulayabilmesi gerekir. Süreç içerisinde katılımcılara belirlediği yöntemin sonuçları etkilemesini bekliyorsa yöntemin nasıl uygulanacağı, uygulanma koşulları, yöntemin sınırlılıkları, katılımcıların gelişim dönemlerine uygunluğuna dikkat etmelidir. Örneğin; İlköğretim 3. Sınıf öğretmeni öğrencilere yaratıcı drama yöntemini kullanarak gelişim ve öğrenme ünitesinde hücre ve organeller konusunu işlemeyi ve drama yönteminin bu deneysel araştırma için etkililiğini araştırdığını düşünelim. Bu yöntemin etkililiğini test etmek için öncelikle öğretmenin yaratıcı drama yöntemine ilişkin bilgisinin olması gerekir. Yaratıcı drama için yaratıcılığın önemini kavramış olması beklenir. Süreç içerisinde öğrencilerinden hücrenin organellerinin görevlerini ezberletip yöntemi rol yapmaya çevirirse yani yaratıcılık düşüncesinden uzaklaşıp dersi işlerse bu yöntemin etkililiğinde doğacak sorun araştırmacının kendisinden kaynaklanmaktadır. Yani uygulama güvenirlik düzeyi düşük olacaktır. Uygulama güvenirlik düzeyi düşük olacaktır. Bu örnekte uygulama güvenirliği düzeyinin düşük olmasının nedeni yöntemin etkililiğini araştırma sırasında katılan hatalardan kaynaklanmaktadır.

Deneysel araştırmalardaki önemli noktalardan biri iç geçerliğin sağlanmasıdır ve uygulama güvenirliği iç geçerlik için önemlidir. İç geçerlik, araştırmada bağımlı değişkende gözlenen değişimlerin, bağımsız değişkenle açıklanabilirlik derecesidir (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel (2013). İç geçerliği etkileyen etmenler; dış etmenler, olgunlaşma, sınanma, ölçme, denek seçimi yanlılığı, denek yitimi, verilerin değişkenlik göstermesi, yapay ortam etkisi, uygulama güvenirliği ve çoklu uygulama etkisidir (Tekin-İftar ve Kırcaali-İftar, 2012).

Bu konuyla ilgili ülkemizdeki alanyazın incelendiğinde uygulama güvenirliği (procedural reliability), eğitim fakültesine ait bölümler arasında daha çok özel eğitim bölümlerinde deneysel bir araştırmada karşımıza çıkmaktadır. Sağlık bilimlerinde özellikle yurt dışındaki çalışmalarda uygulama güvenirliği "treatment fidelity, treatment integrity" gibi farklı adlarla arandığında deneysel desendeki önemine ilişkin pek çok dokümana ulaşılmaktadır.

Uygulama güvenirliğinde sınıfın mevcudu diğer önemli unsurdur. Araştırma kapsamında uygulayıcının, katılımcıların süreç içerisindeki davranışlarını tek tek gözlemlemesi gerekmektedir. Eksik ya da yanlış öğrenmelerde anında düzeltmeler yapılmalıdır. Eğer sınıf mevcudu fazla ise araştırmacı tek tek katılımcılarla ilgilenemeyecektir.

Uygulama güvenirliğinde gözlemci, sürecin işleyişine katılan bir diğer elemandır. Süreçte katılımcılar araştırmayı yapan(lar) ve gözlemciler bulunmalıdır. Çünkü puanlayıcılar arası güvenirlik, uygulama güvenirliğinde de güvenirlik hesaplanırken önemli bir yer tutar. Muhtemelen, uygulama güvenirliğin sonuçlarının değerlendirilmesi çok çeşitli yollarla rapor ve analiz edilmiş olacaktır. Çoğu yöntemler son zamanlarda veri setine ilişkin yapılan güvenirlikler uygulama güvenirliği sonuçları olarak kabul edilmektedir. Buna rağmen çoğu yöntemlerin sonucu olan çıktılar iki ya da daha çok puanlayıcı tarafından kontrol edilmektedir.

Araştırmacı sürecinin video kaydını alarak da farklı gözlemcilerden de puanlayıcılar arası güvenirliği hesaplamaya dâhil edebilir.

Uygulama güvenirliği, bağımsız değişken güvenirliği olarak da bilinmektedir. Uygulama güvenirliğinde, araştırmada yer alan bağımsız değişkenin ne derece planlandığı gibi ne derece uygulandığı da belirlenmektedir (Wolery ve Ledford, 2013). Diğer bir deyişle, uygulama güvenirliğinde planlanan uygulamanın ne kadar doğru olarak uygulandığına bakılmaktadır. Uy

gulama güvenirliğinde, uygulama ile ilgili tüm değişkenlerin uygulama planında hedeflendiği/yazıldığı gibi uygulanıp uygulanmadığı değerlendirilmektedir (Stormont ve Reinke, 2013; Tawney ve Gast, 1984; Tekin İftar, Kurt ve Çetin, 2011). Bu amaçla uygulama planında gerçekleşmesi beklenen uygulama davranışları belirlenir ve uygulama güvenirliği formu hazırlanır. Gözlemci, uygulamacının bu davranışları yerine getirip getirmediğini forma işaretler.

Aşağıda örnek bir uygulama güvenirliğini formu yer almaktadır.

X Kavramının Öğretimine İlişkin Uygulama Güvenirliği Formu

Katılımcının Adı - Soyadı: Tarih:
Uygulayıcının Adı- Soyadı: Oturum No:
Gözlemcinin Adı- Soyadı:

Etkinlikler	Deneme Sayısı	(+) (-)	Dikkat çekme	Öğretimle ilgili açıklama yapma	Araç-gereç ile kavramı ilişkilendirme	Öğrenciden kendi cümlesiyle durumu açıklamasına izin verme	Doğru tepkiler sonunda öğrenciyi pekiştirme
	1						
1. Etkinlik	2						
	3						
	4						
	5						
	1						
2. Etkinlik	2						
	3						
	4						
	5						
	1						
3. Etkinlik	2						
	3						
	4						
	5						

Örnek olarak sunulan uygulama güvenirliğini incelersek bu formda, yöntemi uygulayan kişinin süreci işlerken geçerliğin sınırları aşılmadan tüm davranışları yer almalıdır. Puanlayıcılar veya

sınıf içerisindeki gözlemciler, uygulayıcının davranışlarını içeren uygulama güvenirliği listesindeki davranışlar doğrultusunda uygulayıcının tam ve doğru yapmış olduğu davranışları işaretleyerek ve aşağıda yer alan uygulama güvenirliği katsayısını hesaplayarak uygulama güvenirliğine ilişkin sonuç elde eder.

Ölçmeden kaynaklanabilecek tehditleri kontrol altına alabilmek için tüm oturumların (yoklama, öğretim, izleme ve genelleme) en az %30'unda gözlemciler arası güvenirlik ve uygulama güvenirliği verisi toplanmalıdır (Tekin ve Kırcaali-İftar, 2001).

Uygulama güvenirliğinde uygulama ile ilgili tüm değişkenlerin uygulama planından hedeflendiği/yazıldığı gibi uygulanıp uygulanmadığı değerlendirilmektedir (Leblebici, 2012). Bu amaçla uygulama planında gerçekleşmesi beklenen uygulama davranışları belirlenir ve uygulama güvenirliği formu hazırlanır. Gözlemci, uygulamacının bu davranışları yerine getirip getirmediğini forma işaretler. Daha sonra aşağıda verilen eşitlikten yararlanarak uygulama güvenirliği katsayısı hesaplanır.

$$Uygulama\ Güvenirliği\ Katsayısı = (\frac{Gözlenen\ Uygulayıcı\ Davranışı}{Planlanan\ Uygulayıcı\ Davranışı}) \times 100$$

Formülü kullanılarak hesaplanır (Kırcaali- İftar ve Tekin, 1997; Billingsley, White ve Munson, 1980). Elde edilen değeri bildiğimiz güvenirlik yorumlar gibi yorumlayıp deneysel desendeki bu dokümanda giriş bölümünde belirtilen soruya ilişkin cevabı elde etmiş oluruz. Bizim deneysel araştırmalardan beklediğimiz, bağımsız değişken ya da değişkenlerin bağımlı değişkenler için anlamlı bir farklılık oluştururken araştırmacının süreç içerisinde beklenen davranışları doğru ve tam bir şekilde yaparak araştırma sonuçlarına uygulama güvenirlik düzeyini olumsuz şekilde etkilememesi ve iç geçerliği yükseltmektir.

Kaynaklar

Billingsley, F., White, O. R., Munson, R. (1980) Procedural Reliability: A Rationale and an Example. *Behavioral Assessment, 2*, 229-241.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, , Ö. E., Karadeniz, Ş., ve Demirel, F. (2013) *Bilimsel Araştırma Yöntemleri* (15. Baskı). Ankara: Pegem Akademi.

Kırcaali- İftar, G. ve Tekin, E. (1997). *Tek Denekli Araştırma Yöntemleri*. Ankara: Türk Psikologlar Derneği Yayınları.

Kırcaali-İftar, G., ve Tekin-İftar, E. (2012). Otizm spektrum bozukluklarına yönelik program örnekleri. E. Tekin (Ed.). *Otizm spektrum bozukluğu olan çocuklar ve eğitimleri içinde*, (s. 239-265) Ankara: Vize Yayıncılık.

Leblebici, T. (2012). *Zihinsel Engelli Öğrencilere Galoş Yapma Becerisinin Öğretiminde Eş Zamanlı İpucuyla Öğretimin Etkililiği*. (Yayımlanmış Yüksek Lisans Tezi). Dokuz Eylül Üniversitesi/ Eğitim Bilimleri Enstitüsü: İzmir.

Stromont, M. Ve Reinke, W. M. (2013). Providing Performance Feedback for Teachers to Increase Treatment Fidelity.*Intervention in School and Clinic. 49(4),* 219-224.

Tawney, J. W., & Gast, D. L. (1984). *Single subject research design in special education*. Columbus, OH: Merril.

Tekin. E., & Kırcaali-İftar. G. (2001). *Özel eğitimde yanlışsız öğretim yöntemleri* (2. Baskı). Ankara: Nobel Yayıncılık.

Tekin-İftar, E., Kırcaali-İftar, G., Kurt, O.,ve Çetin, Ö. (2011). Yüksek ve Düşük Uygulama Güvenirliğiyle Sunulan Sabit Bekleme Süreli Öğretim Uygulamalarının Karşılaştırılması. *Kuram ve Uygulamada Eğitim Bilimleri, 11(1)*, 363-381.

Reed, D. D., ve Codding, R. S. (2014). Advancements in Procedural Fidelity Assessment and Intervention: Introduction to the Special Issue. *Journal of behavioral education, 23*, 1-18.

Wolery, M., Ledford, J. R. (2013). Procedural Fidelity: An Analysis of Measurement and Reporting Practices. *Journal of Early Intervention, 35 (2)*, 173-193.

Bölüm IV: Serbest Köşe…

Kübra ESEN

Kübra Esen 1995 yılında Almanya‘nın Gelsenkirchen şehrinde doğmuştur. Şu an Dortmund şehrinde 2015 den beri Sosyal Hizmet okuyor ve bu senenin yaz döneminde İstanbul’da özel bir üniversitede ERASMUS programı ile bir dönem eğitim almıştır. Bu yazı ile Alman ve Türk üniversitelerini yaşadığı kısa deneyimle karşılaştırmayı amaçlamaktadır.

Merhaba, şu an 21 yasındayım ve İstanbul’ da özel bir üniversitede ERASMUS programı sayesinde bir dönem okumaya fırsatım oldu. Ondan önceki eğitim hayatimin tümünü Almanya’da okudum. Türkiye ve Almanya’daki yükseköğretim sistemlerinin farklı olduğunu biliyordum, hele ki üniversite özel olduğu için, daha da farklı olacağını biliyordum. Ama çok şaşırtıcı farklılıkların olması bu yazının doğmasına neden olmuştur. Daha ülkeme gelmeden zorluklar başlamıştı. Almanya’da alıştığım düzen birden yok olmaya başladı ve çok zorlandım. Ders programını bile ilk haftalarda oluşturamadım, böylelikle ilk haftalar derslere katılamadım. Almanya’daki sistem söyle: Internet üzerinden üniversite bilgileriyle kayıt oluyorsun ve tüm dersleri istediğin saate göre seçebiliyorsun. Yani bir ders haftada 2-3 kere veriliyor. Ders programını kendi özel hayatına göre ayarlayabiliyorsun. Geldiğim üniversitede, bunun tamamen farklı olduğunu anladım. Orada bize hocalarla birlikte sınav programımızı hazırlamamızı söylediler. Ama hangi hocaya neden gideceğimiz söylenmemişti. İlk haftalar böylelikle çok koşturmalı geçti çünkü Almanya’ya göndermem gereken evrakları yetiştirmekle meşguldüm. Hocalar beni ERASMUS öğrencisi olduğum için, diğer „Türk“ öğrenciler gibi sıkmadılar. Benim için faydalı olan şey Türkçe bilmemdi, bu nedenle bazı konularda çok yardim gerekmiyordu. Yine de çoğu şeyleri anlamakta zorluk çekiyordum, hocalar ve çalışanlar Türkçe bilmem nedeniyle anlamamı bekliyorlardı. Bazı öğretim görevleriyle İngilizce anlaşmaya çalıştım ama o da maalesef olmadı. İngilizce konuşma konusuna gelince: Almanya’da ilkokuldan beri öğrenciler İngilizce dersi görmek zorunda. Bu nedenle öğrencilerin çoğu süper olmasa bile İngilizce konuşup anlayabiliyorlar. Türkiye’de gördüğüm ve tanıdığım öğrencilerde bunu maalesef göremedim. İngilizce biliyorum dendiğinde bile şaşıranlar vardı. Bazı hocalar bile sadece kırık İngilizce konuşabiliyorlardı.

İlk aylarda gerçekten çoğu şeye aklım almıyordu ve hep Almanya'daki sitemle karşılaştırıyordum. Bunu dönemin son aylarında bıraktım çünkü böyle daha çok zorluk çekiyordum. Üniversite özel olduğu için öğrencilerin çoğunun maddi durumu iyiydi. Almanya'da özel üniversite çok az olduğu için böyle okulları tanımıyordum. Almanya'daki üniversitelerde 6 aylık bir harç ödeme zorunluluğu var. Bu harç da tramvay, otobüs vs. kullanmak için bileti karşılıyor. Çok az bir miktar üniversiteye veriliyor. Bunu Türkiye'deki özel üniversiteyle karşılaştırdığım zaman çok büyük bir fark çıkıyordu ortaya. Üzüldüğüm şeyde bu paraya aldığın eğitim. Derslere girdiğim zaman basta gerçekten çok zorlanıyordum çünkü hocanın sadece önde durup slaytlardan okumasını anlayamıyordum yani hiç görmemiştim. Öğrenciler de hocaya karşı saygıyı unutmuş durumdalardı. El kaldırıp hoca ismini söyledikten sonra konuşmak vardır sadece Almanya'da. Benim gördüğüm öğrenciler ise hocanın lafını kesip istedikleri gibi konuşabiliyorlardı.

Ülkemde ders saatlerinin varlığının anlamını hala çözebilmiş değilim. Derslerin başlama ve bitiş saatlerinin de anlamı yok. 8:30 dersi 9:00 da başlayıp 10: 30 da bitiyor genellikle. Ara verdiğimiz sandığımızda aslında dersin bitmiş olduğunu görüyoruz. Ders saatlerinin hocalar kadar, öğrenciler için de bir anlamı yok. Öğrenciler ders başladıktan sonra istedikleri zaman kapıyı vurup "pardon" diyerek derse gelebiliyorlar ve hocalar da buna göz yumuyor. O zaman ders saatlerine ne gerek var?

İstediğin bölümü büyük miktarda para verip okuyabiliyorsun ama paran olmazsa sınavlara çok çalışıp denemen gerekiyordu. Tabi ki bir öğrenci olarak ders çalışmalısın, ama bu her öğrenciye geçerli olmalıdır. Zengini fakiri ayırt edilmemelidir. Bu ayrımı maalesef İstanbul'da olduğum süre çok fark ettim. Almanya'da hocalarla belirli bir mesafe vardır ve o mesafe daima korunur. Böylelikle o saygı hep devam eder. Türkiye'de bazı hocalar öğrencileriyle arkadaş gibiler ve bunun sonucu olarak hoca öğrenci ilişkisi pek yaşanmıyor. Buda bence dersi pek ciddiye almamalarını sağlıyor. Sonuç olarak her iki ülke eğitim sistemlerine baktığımda Alman ve Türk eğitim sistemlerinin ne kadar farklı olduğunu ve Türk eğitim sisteminde bir şeylerin değişme zamanı geldiğini düşünüyorum.

Bölüm V: Tez Künyeleri

Bu bölümde önceki bültenimizde yer verilmemiş olan 2015 ve 2016 yılı içerisinde Eğitimde Ölçme ve Değerlendirme lisansüstü eğitim programlarında öğrenim gören öğrencilerin tarafından literatüre kazandırılan tezlerin künyeleri yer almaktadır.

Tez Adı	Türü	Yazar	Danışman	Okul
Uluslararası Öğrenci Değerlendirme Programı (PISA - 2012) Okuryazarlık Testlerinin Ölçme Değişmezliğinin İncelenmesi	DR	Betül KARAKOÇ ALATLI	Doç Dr. Ömay ÇOKLUK	Ankara
Tek Boyutlu ve Çok Boyutlu Madde Tepki Kuramı'na göre Çok Boyutlu Yapıların İncelenmesi	DR	Emrah GÜL	Prof. Dr. Nizamettin KOÇ	Ankara
Çok Kategorili Puanlanan Maddelerin Psikometrik Özelliklerinin Farklı Test Koşullarında Parametrik Olmayan Madde Tepki Kuramı Modellerine Göre İncelenmesi	DR	Asiye ŞENGÜL AVŞAR	Prof. Dr. Ezel TAVŞANCIL	Ankara
Farklı Yıllarda Uygulanan Yabancı Dil Seviye Tespit Sınavı Puanlarının Eşitlenmesi	YL	Yaşar Mehmet ZOR	Doç. Dr. Ömay ÇOKLUK	Ankara
TEOG Kapsamında Uygulanan Matematik Alt Testi İle Matematik Mazeret Alt Testinin İstatistiksel Eşitliğinin Sınanması	YL	Mustafa ÖZKAN	Doç. Dr. Ömay ÇOKLUK	Ankara
Akademik ortamlarda olumsuz değerlendirilme korkusu ölçeğinin geliştirilmesi	YL	Volkan ALKAN	Doç. Dr. Ömay ÇOKLUK	Ankara
Test Yönergesi ve Risk Alma Eğiliminin Seçmeli Testlerin Psikometrik Özelliklerine ve Tahminle Yanıtlama Davranışına Etkisi	YL	Ercan ÇOBAN	Prof. Dr. R. Nükhet DEMİRTAŞLI	Ankara
Çoktan seçmeli ve karma test uygulamalarına ilişkin öğrenci başarıları ile öğrenci ve öğretmen görüşlerinin karşılaştırılması	YL	Buket EREN	Prof. Dr. R. Nükhet DEMİRTAŞLI	Ankara
Ekonomik bakımdan dezavantajlı öğrencilerin akademik yılmazlık düzeylerinin bazı koruyucu faktörler açısından incelenmesi	YL	Hatice Çiğdem YAVUZ	Yrd. Doç. Dr. Ömer KUTLU	Ankara

Çok Değişkenlik Kaynaklı Rasch Ölçme Modeli ve Hiyerarşik Puanlayıcı Modeli ile Kestirilen Parametrelerin Karşılaştırılması	DR	Müge ULUMAN	Prof. Dr. Ezel TAVŞANCIL	Ankara
Akademik başarıyı etkileyen faktörlerin etkili okul kavramı bağlamında incelenmesi	YL	Selin ALAMDAR	Yrd. Doç. Dr. Ömer KUTLU	Ankara
PISA 2003 ve PISA 2012 matematik okuryazarlığı testlerinin madde yanlılığı bakımından Türkiye ve Kore uygulamalarında karşılaştırılması	DR	Esra EMİNOĞLU ÖZMERCAN	Prof. Dr. R. Nükhet DEMİRTAŞLI	Ankara
Bilişsel tanı modellerinin değişen koşullar altında karşılaştırılması: DINA, RDINA, HODINA ve HORDINA modelleri	DR	Ömür Kaya KALKAN	Prof. Dr. Hülya KELECİOĞLU	Hacettepe