

Bilimsel Yayınların Bulguları Şüpheden Uzak mı?

Prof. Dr. Atila Yüksel & Doç. Dr. Fisun Yüksel

Adnan Menderes Üniversitesi

“Modern bilimin ürettiği bilginin büyük bir bölümü yanlıştır” (Ioannidis, 2005).

“Likert ve semantik diferansiyel ölçekleriyle elde edilen tüm araştırma bulguları şüphelidir” (Rossiter, 2011).

“Tüm bulguların %90’ının yanlış olduğunu söylemek yanlış değildir” (Kollat, Engell ve Blackwell, 1972).

Rahatsız edici ifadeler olmakla birlikte günümüzde kendine hızla cevap bulması gereken yerinde bir uyarı. Yayın baskısı, daha fazla puan, teşvik ve fon bulma baskısı, kurumlar ve dergiler arası rekabet derken ve bilimsel üretim böylelikle nicelde artarken acaba üretilen “bilim” işe yaramaz bulgulardan başka bir şey değil mi yoksa? Bilimsel bilginin doğruluğu ve güvenilirliğini mercek altına alma zamanı çoktan gelmiş görünüyor. Belki de böylelikle yaygın ama bir o kadar yanlış bilim geleneklerinin son bulması mümkün. Doğruluğun ve güvenilirliğin, büyük ölçüde bilim insanlarının ahlâkî değerlerine bağlı olduğu iddia edilmekte. Yerinde, ama eksik bir iddia. Yanıtlayıcının bağışıklık siteminin yanlılık virüsüne karşı güçlü olduğunu mu düşünüyorsunuz? Yanıtlayıcıyı sadece soruya cevap veren insan (*homo sapiens* “bilen insan”) olarak görmek oldukça eksik bir bakış açısı. Çünkü karşımızdaki *homo sapiens*, belki de aslında bir *homo ludens*, *homo fabers*, *homo economicus*, *homo mentaxi*.

“Araştırmanın/çalışmanın sonuçlarına göre...” diye biten birçok bilimsel makaleye itimat iki nedenden dolayı şüphelidir. Birinci neden biraz felsefi: Makale “sonuç” üretir mi yoksa diğer bir araştırmanın yapılmasına “neden” mi üretir? TDK’ya göre sonuç “bir olayın doğurduğu başka bir olay veya durum, netice” olmakla birlikte sonuç kelimesi sosyal bilim araştırmalarında bu anlamını yitirmekte. Araştırmanın sonuçları ifadesinde yer alan “son-uç” kelimesine uygulamada sanki “kesin son, nihai nokta” anlamı yüklenmekte. Eğer her makalenin sonucu nihai son ise o konuda araştırma yapmaya gerek kalmamaktadır. Çünkü ilgili makale zaten konuya son noktayı koymuştur. Konu tamamen aydınlanmıştır. Daha fazla incelenmesine gerek kalmamıştır! Önceki makale incelenen konuya “kesin” nokta koyduysa sonraki araştırmacılar kesin sonlanan aynı konuları neden incelesin? Bu nedenle, “araştırma

sonuçlarının” uygun bir ifade olmadığını düşünmekteyiz. Araştırmanın son-ucu değil katkısı olabilir.

İkinci neden ise daha çok ölçüm hatası. Bulgu, neticede bir üründür. Bu ürünün salt ölçümden mi yoksa ölçüm dışı başka bir unsurdan mı meydana geldiğini ispatlamadan, araştırmanın genellenebilir sonuçlar ürettiği savı baştan sakattır. Neden mi? Bize öğretildiği üzere “bilim genellenebilir özellik taşıyan bilgilerden oluşmaktadır” (?). Ancak, tek defaya mahsus gerçekleşen davranış, olgu, durum olması durumunda bunun bilimsel bilgi içinde yer almaması kabul edilemez bir yanlış mı olacaktır (big bang gibi). TDK ya göre genellemek “ varlıklar veya olaylar arasındaki benzerlik bağlantılarını bir düşüncede toplamak, tamim etmek, belirlemektir”. “Makale sonuçlarının genellenebilirliği çoğunlukla ” su götürmez bir şekilde ayakları yere basmayan, sorgulamaya oldukça açık bir savdır. Genellenebilir olması için, diğer gerekliliklerin yanı sıra, araştırma bulgularının gerçeği yansıtması; gerçek olması için ise bulgunun hata kaynaklı olmadığını ispatlanması gereklidir. Kısaca bulgu **yanlılık eseri** olmamalıdır. Oysaki “yanlılık” bilimin bağımsızlık sistemine saldırmayı bekleyen bir virüs gibi doğal ve her zaman hazır olarak beklemektedir. Yanlılık, özellikle hakkında eğitim alınmamış yanlılık, bilinçli bir araştırmacının dahi gözünden kaçabilir.

Katılımcının doğru ve gerçekten sapmasına neden olan bilişsel yanlılığın sadece özensiz çalışmaların bir ürünü olarak düşünülmesi, problemin büyüklüğünü küçümsemektir. Yanlılık, Batı da geliştirilen ölçüklerin Türkçeye doğru çevirileriyle aşılacak basit bir konu değildir. Yanlılık soru sormak-cevap almak uygulamasını gerektiren tüm araştırmaların doğasında bulunmaktadır. Uykuda olan virüs gibi uygun ortamı bulduğunda aktif duruma geçecektir. Peki, yanlılık araştırmacılar tarafından önemsenmekte midir? Bir soru daha ekleyelim. Sadece araştırmacılar değil, ilgili alanda bilimin birikimine ve yayın politikalarına yön veren dergilerin editörleri araştırmada yanlılığı önemsemekte midir? İkincisi de en az birincisi kadar önemlidir. Her iki taraf da yanlılığı önemsemelidir. Çünkü yanlılık olgular arasında aslında olmayan ilişkilerin varmış gibi görünmesine neden olmaktadır. Ya da aslında var olan ilişkilerin yokmuş gibi görünmesine. Yani korelasyon katsayıları, açıklanan varyanslar ya da güvenilirlik katsayıları aslında gördükleri gibi olmayabilir. Yanlılık eğer var ve araştırma analizleri bu yanlı bilgilere yapılıyorsa o zaman literatürde var olan hipotez testlerine dayalı üretilen “sonuçların genellenebilirliği” tartışmalıdır. Yanlılık üzerine inşa edilen hüküm, bataklık üzerine inşa edilmeye çalışılan binanın akıbetinden kaçamayacaktır.

Kötü bir haberimiz var! Üretilen bilginin doğru alana yerleştirilmesi, geçerliliğinin anlaşılması ve araştırmanın katkısının takdir edilmesi için gerekli olan “**yansızlığın ispatı**” maalesef sosyal bilim araştırmalarında umursanan yaygın bir gelenek olamamıştır. Neden mi? Çünkü mevcut yayınlanmış tüm makaleler içinde çok azı araştırmanın katkısının yorumlanmasında dikkatli olunması gerektiği, çünkü araştırmanın yanlılıktan mustarip olabileceğini tartışmaktadır. En prestijli bilim dergilerini kapsayan bir çalışmada Ioannidis (2007) tüm makaleler içinde araştırmanın sınırlılıklarından bahsedenlerin oranının %17 lerce olduğunu; benzer bir şekilde 14 bin makaleyi kapsayan çalışmada Mortel (2008) çalışmaların sadece binde 2 sinde sosyal beğenilirliğe ait yanlılığın izlerinin incelendiğini rapor etmektedir. Kabaca, “bulgularım yanlılık eseri midir?” kaygısını taşıyan bilim insanlarının oranı %1 bile değil! Ama şu ana kadar alanımızda 30.000 in üstünde makale bilimsel değerlendirmeden geçmiş ve yayımlanmış durumda. Bir şeyler yanlış mı gidiyor?

Altı üstü sorulan sorulara yanıt vermek, ne kadar zor olabilir? Tahmin ettiğimizden daha zor. Yanıt vermek bilişsel bir süreç. Dikkate, ilgiye, bilgiye ve kişinin bilişsel kapasitesine son derece bağlı. “Yanlılık” olmaması için sorulan sorunun yanıtlayan kişi tarafından önce anlaşılması, anlamlandırılması, sonra hafızada o bilgiye ulaşılması ve yanıtın oluşturulması gerekmektedir. Yanıtın düşünülerek verilmesi yanlılığa bir çözüm gibi gösterilmekte. Halbuki düşünülerek de veriliyor olsa, ortaya çıkan yanıt kişinin basitçe bir rakam tercihinin, sosyal beğenilme arzusunun, ön yargılarının, içgüdülerinin, yanlış algılamalarının, yazım hatasının, bilgisizliğinin sonucu olabilir. Şu soruyu düşünün. “Arkanız dönükken birinin size baktığını hissedersiniz? Cevabınız, bu konuyu bilimsel merakla inceleyen araştırma sonuçlarından sanırsanız farklı olmayacaktır. “Evet elbette hissedirim”. Önceki çalışmalar, hissedeceğini söyleyenlerin oranının %65-85 aralığında olduğunu söylüyor. Ancak, hissedeceğini söyleyen kişilerle yapılan deneyler aslında böyle bir hislerinin olmadığını gösteriyor. En ilginç örneklerden birinde kişi bu soruya yanıt verirken Goril kıyafeti giymiş bir kişinin uzun süredir arkasından dikkatlice baktığını deneklerin hissetmediğini gösteren goril deneyi. Amerika Birleşik Devletlerinde “Metalic Act” hakkında kişilere düşünceleri soruluyor. Olumlu ve olumsuz bulanların sayısı oldukça fazla. Ama gerçekte böyle bir Metalic Act diye bir yasa ABD de yok. Olmayan bir konu hakkında sırf soru sorulduğundan dolayı insanlar neden bu kadar cevap vermeye istekli? Avrupa’da değişik ülkeleri kapsayan sosyal içerikli bir araştırma. Sonuçlar açıklanıyor. Hollanda haricinde tüm ülkeler ortalama değerlerde. Hollanda ise resmen uç. Hollanda versiyonunda en alttaki seçenek çok yoğun seçiliyor. Araştırmacılar

acaba bu fark bir çeviri hatasından mı diye endişeleniyor. Ancak, çeviriler mükemmel, sorunsuz. Sonunda anlaşılıyor ki seçeneklere eşlik eden ve cevapların yazılacağı kutular diğer ülkelerdeki uygulamada hep aynı boydayken, Hollanda'da uygulanan anket formlarında yukarıdan aşağıya piramit şeklinde genişliyor. En alttaki cevap seçeneğinin kutusu en geniş. Hollandalı denekler daha çok en alttaki seçeneği işaretliyor. Sebep? Kutu daha uzun olduğu için daha çok yazmak isteyen insanın uzun kutuyu seçeceği inancı. Kısaca deneklerin ya da katılımcıların verdiği yanıtlara güvenecek miyiz? Ya toplanan veri farklı sebeplerden dolayı kirliyse?

Araştırmanızı uygularken aşağıdaki sorulardan biri ya da birkaçına evet diyorsanız yanlılık belirleme ve elimine etme süreçlerini devreye sokmadan doğruca analize gitmeniz hatalı olacaktır.

Yanıtın nedeni	Tipoloji	
Katılımcı araştırmacı tarafından eleştirilmemek, beğenilmek için yanıt verilebilir mi?	Sosyal beğenilirlik	Social desirability bias
Araştırmaya katılanların yanıtlarının araştırmaya katılmayanların yanıtlarından farklı olması bir yanlılık mıdır?	Katılmayan-katılan farkı	Non-response bias
Bir önceki soruya verilen yanıt daha sonraki sorulara verilen yanıtları etkiler mi?	Tanışık etkisi	Acquiescence bias
Soru içeriğine bakılmaksızın skalada uç noktaların daha fazla seçilmesi mümkün mü? 5 li skalada 1 ve 5 gibi	Ekstrem yanıt verme etkisi	Extreme-response rating
Soru içeriğine bakılmaksızın skalada orta değerler daha fazla seçilmesi mümkün mü?	Orta noktayı seçme eğilimi	Mid-point bias
Soruları okumadan tüm sorulara aynı cevap seçeneğinin işaretlenmesi mümkün mü?	Halo etkisi	Halo effect
Bilginin tek araçla toplanması verilen yanıtları etkiler mi?	Ortak-yöntem etkisi	Common-method bias
Yanıtlar aslında yazım hatalarının sonucu olabilir mi?	Baskı hataları	Graphical errors
Birbirine mesafe olarak yakın cevap seçenekleri daha uzak mesafede olan seçeneklerden daha çok tercih edilebilir mi?	Yakınlık etkisi	Nearness effect
Denek, içeriğinden daha ziyade konumundan dolayı üstte olan yanıt seçeneğinin altta olan seçeneğe göre daha fazla seçer mi?	Üst konum etkisi	Top means good
Denek sorulan konuyu bilmediği halde fiktif yanıt verebilir mi?	Bilgisizlik etkisi	Uninformed bias
Denek otomatikçe bağlayıp tüm sorulara devamlı evet ya da hayır cevabı verebilir mi?	Evet-hayır eğilimi	Yeah-nay saying

Denek geçmiş bir olgu, durum, davranışı hatırlarken bilişsel zorluklar yaşayabilir ve sonunda çarpıtılmış bilgi verebilir mi?	Hatırlayabilme yanlılığı	Recall bias
Cevap seçeneklerinde, oluş sırasında ya da soru sıralamasında ilk olanlar daha çok hatırlanabilir/ tercih edilebilir mi?	İlkler etkisi	Primacy effect
Cevap seçeneklerinde, oluş sıralamasında ya da soru sıralamasında son olanlar daha çok hatırlanabilir/tercih edilebilir mi	Sonlar etkisi	Recency effect
Bazı çalışanların üstün performansından memnun olma sonucu tüm çalışanların performansı kötü olanlarında dahil olumlu görülebilir mi?	Genelleme etkisi	Leniency bias
Yanıtlar cevap seçeneklerinin dikey-yatay skalada sunumundan; etiketlenmiş-etiketlenmiş skalanın ürünü olabilir mi?	Format/ölçek yanlılığı	Scale bias
Verilen yanıt gerçeği mi yansıtır yoksa basit bir rakam tercihi midir?	Rakam tercihi etkisi	Digit preference

Süregelen hakim bilim geleneği son derece sıkıcı, kasvetli, kurallara boğulmuş durumda. Aslında bilimin kendisi yerine kurallar bilim olmuş. Özellikle çok yanlış anlaşılan ve kullanılan olasılık değeri (p). (%95 güvenilirlik aralığı) çoğu bilim insanı için kible ve farz olmuş. Merak ve bulma arzusu kuralın dışına çıkamıyor. Belki de bu yüzden araştırmalar genel olarak daha çok Batı'da yapılanların taklitleri. Ne kadar düşündürücü. Kompleks analizler, hayatta olması muhtemel olmayan varsayımlar, ilişkiler, acabalar vs. Bilim ve ürünü basit, anlaşılır ve en önemlisi eğlendirici olmalı (entelektüel eğlenceyi kastediyoruz). Lisans, Yüksek lisans ve doktora öğrencileri içinde geleceğe yön verecek bilim insanların yer aldığını gözden kaçırmamalıyız. Daha ilk anda korkunç istatistikler ve anlaşılmaz yöntemlerle bilimden soğutursak, gelecek çok verimsiz olacaktır. Soruları eğlenceli hale getirmek ve böylelikle ilgiyi artırmak yanlılık gidermede bir yaklaşıma dönebilir.

Bu amaçla uzun zamandır daha çok öğrencilerin yer aldığı basit testler yapıyoruz. Testlerin amacı bilimsel yaklaşımı basitleştirmek, mevcut korkutucu karmaşıklığından uzaklaştırmak. Kısaca bilimsel üretimi öğrencilere sevdirmek. Literatürden esinlenerek bir basit testte öğrencilerimize bir bisikletin nasıl çalıştığını biliyor musunuz diye sorduk. Sizde kendinize 5 TL nin ön ve arka yüzünde neler olduğunu şu an sorabilirsiniz. Önce çok iyi bilirim-hiç bilmem seçeneklerinin olduğu bir skalada yanıtları topladık. Sonra aynı öğrencilerimize pedalı, zinciri ve tekerleği görünür olmak şartıyla bir bisiklet çizin dedik. Sonuç inanılmaz! Her gün karşılaşılan ve kullanılan somut bir nesneyi tarif ederken dahi denekler sıkıntı yaşıyor. Tutum, tatmin, bağlılık, duygusal imaj ve daha çok sayıda olan karmaşık, soyut kavramlar

sorulduğunda yanıtlayıcıların beyinlerinde neler olduğunu düşünmek bile istemiyoruz. Ya gerçekten bu kavramları tanımlayamıyorlarsa? Kişiler pedal, zincir ve teker ilişkisini gösteremiyor. Örneğin zincir arka tekerden başlayıp ön tekerleğe ulaşıyor. Pedal var zincir yok, ya da pedal ve zincir var ama zincir pedaldan ön tekerleğe bağlı. Bu, basitçe çizime karşı yeteneksizlik değil. İlginç. Bisikletin nasıl çalıştığını bilmem diyenlerin çoğunda çizim hatası yok. Ama durun. Çok iyi anlarını diyenlerin çoğunda ise çizim hatası daha çok. Denning ve Kruger etkisi. Kifayetsizliğin azmi bu olsa gerek. Bu basit testten önemli bir katkı olabilir. Sadece 1. soruyu yeterli gören araştırmancının katkısı belki de “23-25 yaş aralığındaki gençler bir bisikletin nasıl çalıştığını çok iyi biliyorlar” olacaktı. Halbuki, 2. soru bunun böyle olmadığını gösterdi. Testte ortaya çıkan bir durum şu. Kişiler ne kadar bildiğini abartma ya da azaltma eğilimi gösterebilir. Bu tehlikeli. Özellikle araştırmalarda olduğundan daha bilgili görünmek ya da olduğundan daha az bilgili olmayı tercih eden yanıtlayıcılarında olabileceğini düşündüğümüzde. Diğer bir testimiz ise ortalama insan olma arzusunun ne kadar yüksek olduğuyla ilgili. Bazı araştırmalar çalışanlara “sizce müşterileriniz aşağıdakilerden hangisine ne kadar önemser” gibi sorular sormakta. Yani kendini bir başkasının yerine koyarak yanıtlamak. Öğrencilerimizden birbirini tanımayan takımlar oluşturduk. 2 kişiden oluşan bu takımlar takımdaki diğer arkadaşının seçimini yazılı olarak tahmin ediyor. Örneğin 1 den 7 ye kadar bir sayı seçmesi gerekseydi takım arkadaşın kaç seçerdi? Siz kaç seçerdiniz. Evet çok da şaşırtıcı olmayacak şekilde en fazla tercih edilen rakam 3-4-5. Yatay olarak sıralanmış yedi kareden ortağınız hangisini seçerdi sorusuna en çok yanıt alan seçenekler 4.,5. ve 6. Ortağınız buluşmak isteseydi sizinle günün hangi saatinde buluşurdu sorusuna yanıt 13-14-15.00. Yani 24lük bir gün diliminin ortasının biraz üstü. Bu bulgular gözün herhangi bir doğruya baktığında doğruyu iki ucundan katlayarak ortasını bulup, orta noktaya daha fazla odaklanmasından ötürü olabilir. Ya da 7’li skalada işaretlenen 3-4 gibi yanıtlar gerçek bir fikri yansıtmaktan daha çok sadece insanın tipik insanlar gibi görünme arzusunun kaynaklanmış olabilir.

Kısaca yayına sunmaya, değerlendirmeye göndermeden önce bir araştırmada yapılması gerekenler listesi yaygın uygulamadakinin daha uzun. Sosyal bilim şu an bir kavşakta ve doğru yöne döneceği ise şüpheli. İşe soru sormaktan başlamalı.

Kaynakça

Ioannidis JPA (2005) Why Most Published Research Findings Are False. PLoS Med 2(8): e124.

Ioannidis, J. P. A. (2007). Limitations are not properly acknowledged in the scientific literature. *Journal of Clinical Epidemiology*, 60, 324–329.

Kollat, D. T., R. D. Blackwell, and J. E. Engel. (1972). "The Current Status of Consumer Behaviour Research: Development during the 1968-1972 Period." In *Proceedings of the 3rd Annual Conference of the Association of Consumer Research*, edited by M. Venkatesan, pp. 576-84

Mortel, F. (2008). Faking it: social desirability response bias in selfreport research. *Australian Journal of Advanced Nursing*, 25(4): 40-48

Rossiter, J. R. (2011). *Measurement for the Social Sciences: The C-OAR-SE Method and Why It Must Replace Psychometrics*. New York: Springer