



Uluslararası Sosyal Araştırmalar Dergisi

The Journal of International Social Research

Cilt: 10 Sayı: 49 Volume: 10 Issue: 49

Nisan 2017 April 2017

www.sosyalarastirmalar.com Issn: 1307-9581

MICHAEL SCRIVEN: DEĞERLENDİRME ÜZERİNE KAVRAMSAL DEVRİM ÖNERİLERİ

MICHAEL SCRIVEN: CONCEPTUAL REVOLUTIONARY SUGGESTIONS IN EVALUATION

Orhan İYİTOĞLU*

Mehmet GÜROL**

Öz

Bu çalışma Michael Scriven'in değerlendirmeyi kendine özgü bir disiplin ve diğer disiplin alanların özünü oluşturan bir alfa disiplin olarak gösterme ve var etme adına yaptığı katkıları incelemeyi amaçlamıştır. Bu bağlamda, doküman incelemesi yöntemi kullanılmış ve Scriven'in değerlendirme anlayışı *Alkin'in Değerlendirme Ağacı* metaforu baz alınarak temellendirilmiştir. Scriven'in değerlendirmeye ilişkin önerdiği ve devrim olarak nitelendiği kavramlar kapsamlı bir şekilde ele alınmıştır. Ayrıca değerden bağımsız olmaması gerektiğini belirttiği değerlendirme sürecinde kullanılması için geliştirdiği Anahtar Kontrol Listesi ve tamamlanmış değerlendirme çalışmalarının bazı ölçütlere göre değerlendirilmesi için geliştirdiği Değerlendirmenin Değerlendirmesi Kontrol Listesi de incelenmiştir. Sonuç olarak, Scriven değerlendirmeyi bir şeyin değer, kalite ve öneminin bazı standartlara dayanarak performans göstergelerinin ölçülmesi ve bunlarla ilgili yargıda bulunmayı gerektiren karmaşık profesyonel bir iş olarak tanımlamıştır. Ayrıca, uygulamalı bir alan olduğunu söylediği değerlendirmeyi, uygulama alanına taşımış ve sürecin karmaşıklığını da kontrol listeleri ile değerlendirmeye her boyutta ne yapması gerektiğini aşama aşama detaylı bir şekilde anlatarak azaltmaya çalışmıştır.

Anahtar Kelimeler: Değerlendirme; Program değerlendirme; Meta-değerlendirme, Değer; Önem.

Abstract

The main purpose of this article is to review Michael Scriven's contributions to "evaluation" to prove and survive it as an alpha discipline on its own. Within this perspective, adopting document analysis method, Scriven's view of evaluation was constructed based on Alkin's Evaluation Tree (Alkin, 2013). In the study, Scriven's revolutionary concepts of evaluation were reviewed in detail. Moreover, the study also covered Key Evaluation Checklist that he developed for use in the evaluation process which should not be value-free and Meta-Evaluation Checklist that he developed for use in the evaluation of completed evaluation process itself since he thinks this process must be evaluated based on some standards. As a result, Scriven (1981) defines evaluation as a complex process that determines merit, value and worth of things through identifying performance indicators based on some standards and asserts it as a profession where evaluation experts make judgements. In this respect, Scriven indicates evaluation as an applied discipline and aims to simplify evaluator's job by guiding him or her through every step of evaluation process thanks to checklists.

Keywords: Evaluation, Program evaluation, Meta-evaluation, Value, Significance.

1. Giriş

Değerlendirme konusunda alan yazında üzerinde henüz tam olarak uzlaşılmış bir tanım bulunmamaktadır. Buna rağmen, farklı değerlendirmeciler değerlendirme sürecinde değerlendirmenin nasıl bir rol üstlenmesi gerektiği konusunda benzer görüşlere sahiplerdir (Worthen, Sanders & Fitzpatrick, 1997). Değerlendirme sürecinin temel amacıyla ilgili yaygın ve ortak düşünce, değerlendirilen şeyin değeriyle ilgili varılan yargılama sonuçlarını ortaya koymaktır. Worthen ve arkadaşları (1997) değerlendirmeye yüklenen bu birincil vazifenin Scriven'in değerlendirme tanımı ve anlayışına dayandığını belirtmektedirler. Scriven (1981) değerlendirmeyi bir şeyin kalite, değer ve önemini ortaya koymak adına gerçekleştirilen süreç olarak tanımlamıştır. Scriven bu tanımla; değerlendirmeyi sosyal bilimlerin geleneksel araştırma kalıplarından kurtararak "maliyet, karşılaştırma, ihtiyaç, etik değerler, politika gibi boyutları olan bir hipotezi test etmekten ziyade güçlü bir değer yargısı ortaya koymayı amaçlayan çok boyutlu bir etkinlik" olarak resmetmiştir (Scriven, 1981, p. 53).

Genel olarak; bir şeyin kalite, değer ve önemini belirlemek için sürdürülen sistematik çabayı ifade eden değerlendirme, program değerlendirme, personel değerlendirme şeklinde sıralanacak yirmiyi aşkın alt uygulama alanı olan çok kapsamlı ve derin bir etkinlik sahasını işaret etmektedir (Scriven, 1999). Bu çeşitliliğe uygun olarak da, değerlendirmenin teorisinin oluşturulması konusunda farklı kişiler tarafından ortaya atılan farklı görüşler bulunmaktadır. Bu alternatif değerlendirme yaklaşımları iki farklı şekilde gruplanmıştır. Worthen ve arkadaşları (1997) bu yaklaşımları bir matriste özellikleri, güçlü ve zayıf yanları

* Okutman, Gebze Teknik Üniversitesi, Yabancı Diller Bölümü, orhanmka@hotmail.com.tr

**Prof. Dr. Yıldız Teknik Üniversitesi, Eğitim Bilimleri Bölümü, mguro@yildiz.edu.tr

halinde sıraladıkları altı gruba ayırmışlardır. Hedef yönelimli, yönetim yönelimli, müşteri yönelimli, uzmanlık yönelimli, rakip yönelimli, katılımcı yönelimli olmak üzere farklı yaklaşımlara dayalı değerlendirme modelleri sunulmuştur. Worthen ve arkadaşları (1997) böyle bir sınıflamayla " her yaklaşımın güçlü, zayıf yanlarını ve temel kullanım alanlarını göstermeyi ve hiçbir yaklaşımın tam olarak kendi başına mükemmel olamayacağını ancak her yaklaşımın amaca göre değerli ve kullanışlı olacağını" gözler önüne sermeyi hedeflediklerini belirtmişlerdir (s. 182).

İkinci sınıflama ise Alkin (2013) tarafından değerlendirme teorisinin gelişimini ve amaçlarını ortaya koymak için ağaç metaforu kullanılarak yapılmıştır. Bu metaforda değerlendirme; sosyal sorumluluk, sosyal sorgulama ve epistemoloji olmak üzere üç farklı kökten filizlenen bir ağaç olarak gösterilmiştir. Her kökten de bir tane olmak üzere toplam üç dal türemiştir. Özellikle devlet tarafından organize edilen programların değerlendirmesi anlayışının alt yapısını oluşturan, sosyal sorumluluk kökünden çıkan kullanışlılık dalı bunlardan birincisidir. Kullanışlılık dalının yaprakları şeklinde gösterilen değerlendirmeciler karar verme sürecinde yöneticilere yardımcı olmayı ve organizasyon ve programla ilgili gerekli değişiklikleri teşvik etmeyi amaçlarlar. Pozitivist ve post-pozitivist anlayışın hakim olduğu sosyal sorgulama kökü daha çok bilimsel, ampirik ve doğrulanabilir yöntemler kullanmayı vurgular. Bu kökten ortaya çıkan yöntem dalındaki değerlendirme yaklaşımları deneysel, yarı deneysel gibi ölçme biçimlerini ortaya çıkarmıştır. Son olarak da bilginin doğası ve geçerliliğine ilişkin felsefi sorgulamaların yoğunlaştığı epistemoloji kökünden de değer vurgusu yapan dal türemiştir. Bu dalın yapraklarını işgal eden yaklaşımlara da Scriven ve Stake öncülük ederler. Bu değerlendirmeciler gerçeğin çoğulcu yapısını vurgulayarak değerlendirmecinin değerlendirilen şey ile ilgili bazı yargılarda bulunup sonuç çıkarması gerektiğini belirtirler (Alkin, 2013).

Bu bölümde, "değerlendirme"nin ayrı bir disiplin olarak algılanması için önemli çabalarda bulunan, ilgili alana "biçimlendirici (formative) ve özetleyici (summative) değerlendirme"; "bütüncül & analitik değerlendirme (holistic & analytic)", "değerlendirmenin değerlendirmesi (meta-evaluation)"; "hedeften bağımsız değerlendirme (goal-free)" gibi terimler kazandıran Scriven'in değerlendirme yaklaşımı incelenmektedir (Encyclopedia of Evaluation, 2005). Bu incelemede Worthen ve arkadaşlarının sınıflaması takip edilmeyecektir. Bu sınıflamada Scriven, müşterilerin ihtiyaçlarını temel alarak değerlendirmeyi tüm müşterilere yönelik daha önceden kararlaştırılan standartlara göre yapılması gerektiğini belirten Müşteri Yönelimli alana yerleştirilmiştir. Bu bölümde, daha güncel olması ve ağaç metaforuyla farklı yaklaşımlar arasındaki tarihsel ve fikrî ilişkileri daha net göstermesi bakımından Alkin'in (2013) Değerlendirme Teorisi Ağacı (An Evaluation Theory Tree) temel alınmıştır. Bu ağaç temel alınarak, Scriven'in değer vurgusu yapan değerlendirme anlayışı ve bu alana sunduğu terminolojik katkılar açıklanacaktır.

2. Michael Scriven

Scriven'in matematik ve felsefe alanında aldığı eğitim, değerlendirmeye ve değerlendirmecinin kendisine ilişkin görüşlerinde fazlaca hissedilir (Alkin, 2013). Melbourne Üniversitesi'nde matematik eğitimini tamamladıktan sonra, aynı üniversitede 1950 yılında matematiğin felsefesi bölümünde master yapmıştır. 1956 yılında Oxford Üniversitesi'nde felsefe alanında doktora eğitimini tamamlamıştır. On farklı alanda yayın yapan kırkı aşkın derginin hakem heyetinde görev yapmıştır. Dört yüzden fazla makalesi ile üretici bir akademisyen olan Scriven; Amerika, Avustralya ve Yeni Zelanda gibi farklı ülkelerde birçok üniversitede felsefe, mantık, disiplinler arası değerlendirme gibi alanlarda akademisyen olarak ders vermiştir (Encyclopedia of Evaluation, 2005). Değerlendirmeyle ilgili çeşitli kurumların ve dergilerin hem kuruculuğunu hem de yöneticiliğini yapmış ve bu alanda birçok ödül almıştır. Scriven, değerlendirmecinin rol ve sorumluluklarını açıkça belirterek değerlendirmenin, özünde bir değer yargısında bulunma olan disiplinler ötesi bir bilim dalı olarak kabul edilmesi için yoğun çaba harcamıştır (Shadish, 1998). Değerlendirmenin böylesi bir disiplin olduğu savını da mantık ve istatistik gibi disiplinlerin yaptıkları işlerin değerini belirlemek için bir değerlendirme sürecine ihtiyaç duyduğunu belirterek destekler (Alkin, 2013).

3. Scriven'in Değerlendirme Ve Değerlendirmeciye İlişkin Görüşleri

Scriven, değerlendirmeyi daha önceden belirlenen bazı standartlara uygun olarak ortaya konan performans göstergelerini temel alarak; sosyal bilimlerden alınan yöntemleri kullanarak değerlendirilen şeyin kalite, değer ve önemine ilişkin yargıda bulunma süreci olarak tanımlar (Scriven, 1983). Ancak bu noktada, Scriven (1999a) sosyal bilimlerden farklı olarak değerlendirmenin sadece gözlem, ölçüm ve hesaplamalarla sınırlı ampirik yöntemler sonucu elde edilen gerçek bilgiye göre yapılmasına karşıdır. Aksine, bu yöntemlerle elde edilen bilgilerin bazı standart ve değerlerle bütünleştirilerek daha kapsamlı bir değerlendirme yapılmasının gerekliliğine dikkat çeker. Bu bakımdan, Scriven'a göre, değerlendirme bazı sosyal problemleri çözmek için sosyal bilimlerden alınan yöntemlerin kullanılmasından ibaret görülemez.

Değerlendirmenin kendine özgü bir bilgi temeli olan ve diğer disiplinlere de hizmet eden bir alan olduğunu ileri süren Scriven (1986) değerlendirme aslında değerlendirilen şey ile ilgili iyi ya da kötü şeklinde bir karara varmak olduğunu şu şekilde belirtir: "Kötü kötüdür, iyi de iyidir ve neyin iyi neyin kötü olduğuna karar vermek de değerlendirmenin işidir" (s. 19). Buradan da anlaşılacağı üzere, Scriven değerlendirmeciliğin de tıpkı değerlendirme gibi özel ve uzmanlık gerektiren bir meslek olduğunu vurgular. Çok hassas bir toplum hizmeti yürüten değerlendirme rollerini de buna bağlı olarak belirtir (Scriven, 1981; 1996; 1999b). Sadece değerlendirilen şeyin paydaşları değil; olası tüm müşterilerin dikkate alındığı kapsamlı bir bakış açısıyla, uygun bütün değerlendirme sonuçlarını sıralaması gereken değerlendirme, bazı yöntem ve değerlendirme bilgisine, becerisine sahip olmalıdır. Scriven'in (1981; 1996; 1999b) değerlendirme için sahip olması gerektiğini düşündüğü özellikler şu şekilde özetlenebilir:

- ✓ Spesifik değerlendirme durumları için uygun teknik, yasal, etik değerleri belirlemek,
- ✓ Değerlendirmenin çeşitli öğelerini ve türlerini birbirinden ayırma mantığını kavramak,
- ✓ Değerlendirme sürecinde uygun verileri toplayacak ve sonuçları belirtecek yöntem bilgi ve becerisi,
- ✓ Standart olarak belirlenen değerler ile ampirik bilgileri bütünleştirip gerekli sonuçları çıkartacak analiz ve sentez becerilerine sahip olmak.

Tüm bunlardan anlaşılacağı üzere, Scriven değerlendirmeyi de değerlendirme uzmanı olarak konumlandırmıştır. Worthen ve arkadaşlarının (1997) yaptığı sınıflamada yönetim yönelimli değerlendirme yaklaşımlarına denk düşen modelleri eleştirilmesine, değerlendirme uzmanlarının, kendilerini hafife aldıklarından, genelde yaptıkları hatanın tüm değerlendirme süreci sonucunda toplanan bilgiyi yöneticilere, karar vericilere altın tepside sunmak olduğunu belirtir (Alkin, 2013). Hâlbuki ona göre iyi ve kötü konusunda son karar her zaman değerlendirme uzmanına aittir (Shadish, 1998).

4. Scriven'in Değerlendirmeye Kavramsal Katkıları

44 yılda değerlendirmeyle ilgili 100'den fazla çalışma yapan Scriven (2013), bu alanla ilgili üretilen yanlış bakış açılarıyla mücadele etmiş ve değerlendirme kendine özgü bir altyapısı ve mantığı olan diğer disiplin alanlarının muhtaç olduğu disiplinler ötesi bir alan olarak göstermeye çalışmıştır. Bu bağlamda; biçimlendirici ve özetleyici değerlendirme ile "The Methodology of Evaluation" adlı çalışmasında 1967 yılından başlayarak alana birçok kavram önermiş ve değerlendirme kavramsallaştırılmasına katkıda bulunmuştur. Bu özelliği de onun "ilk ve net değerlendirme teorisi olan değerlendirme" olarak gösterilmesini sağlamıştır (Shadish, 1998). Çalışmanın bu bölümünde Scriven'i böylesine önemli bir değerlendirme uzmanı olarak kavramsal katkıları kısaca anlatılacaktır:

4.1. Biçimlendirici & Özetleyici Değerlendirme: Scriven 1967 yılında biçimlendirici ve özetleyici değerlendirme arasında ayırım yapmış ve bu ayırım yaygın bir biçimde kabul görmüştür (Worthen ve diğerleri, 1997). Bu ayırımı, daha çok değerlendirme uzmanının amacı önemli rol oynamıştır. Biçimlendirici değerlendirme (formative evaluation) "bir ürün ya da programın geliştirilmesi aşamasında uygulanan... iç ya da bağımsız bir değerlendirme uzmanı tarafından yapılabilecek" bir değerlendirme (Scriven, 1981, s.63). Bu değerlendirme, bir programın planlandığı gibi uygulanıp uygulanmadığını ve uygulama sırasında ortaya çıkan problemleri keşfedip gerekli düzeltmeleri yapmak üzere gerçekleştirilir (Encyclopedia of Evaluation, 2005). Bir programın geliştirilmesi aşamasında yapılan pilot uygulamalar bu tarz bir değerlendirme örnek olarak gösterilebilir. Özetleyici değerlendirme (summative) ise daha çok "bir programın uygulanmasından sonra iç ya da bağımsız bir değerlendirme uzmanı tarafından farklı paydaşlar için ya da karar vericiler için gerçekleştirilir" (Scriven, 1981, s. 150). Bu değerlendirme karar vericilerin programın değeri, kalitesi ve akıbeti konusunda karar vermelerine yardımcı olur (Worthen ve diğerleri, 1997). Sene sonunda yapılan ülke çapındaki sınavların sonuçlarının, çeşitli kademelerdeki eğitim programlarının değerlendirilmesi amacıyla kullanılması böyle bir değerlendirme amacına hizmet etmektedir. Scriven (1981) Bob Stake'in bu ikili ayırımı şu şekilde netleştirdiğini belirtir: "Bir aşçı çorabayı tattığında biçimlendirici; misafirler tattığında özetleyici değerlendirme yapmış olur" (s.63) Her ne kadar ikili arasındaki ayırım net gibi görünse de Scriven uygulamada bu ikilinin iç içe geçebileceğinin altını çizer (Worthen ve diğerleri, 1997).

Bu iki değerlendirme biçimine ek olarak daha sonra, Scriven (2013) "Key Evaluation Checklist" adlı çalışmasında *atfedici* (ascriptive) değerlendirme söz eder. Bu değerlendirme belli konularda bilgi birikimini artırmak için ya da o konuya duyulan ilgiden ötürü yapılır. Bir savaşta atlı birliklerin kullanılmasının değerlendirilmesi bu tarz bir değerlendirme işaret eder.

4.2. Hedeften Bağımsız Değerlendirme: Scriven tarafından 1972 yılında geliştirilmiştir (Uşun, 2012). Bu değerlendirme yaklaşımında, değerlendirme programının hedeflerine dayalı değil; kendi belirlediği

çıktılarına ya da programın gerçek sonuçlarına dayalı olarak değerlendirme yapar. Scriven (1983) bu değerlendirme yaklaşımının "ihtiyaç temelli değerlendirme" ya da "müşteri yönelimli değerlendirme" şeklinde tanımlanabileceğini de belirtir. Bu görüşün temelinde, değerlendirmenin sadece hedeflenen kitleye değil toplumun tamamına karşı bir sorumluluğun olduğu varsayımı yatar. Bu bakımdan, "değerlendirmeci değerlendirmeye programın hedeflerinin ne olduğunu bilmeden başlar ve programın uygulamada ne yaptığını ortaya çıkarmaya çalışır" (s. 68). Hedeflerin, değerlendirmenin dayanak noktası olarak kabul edilmesinin değerlendirmenin tarafsızlığını ve geçerliliğini düşüreceğini belirtir. Scriven (1972) hedeften bağımsız değerlendirme anlayışının daha kapsamlı, güvenilir ve ön yargılardan uzak olduğunu belirtir. Çünkü hedefler genellikle değerlendirme sırasında görülen önemli sonuçlarla örtüşmez (Worthen ve diğerleri, 1997). Ayrıca sadece hedeflere yönelik bir değerlendirme, değerlendirmeyi programın paydaşları ve toplumun çeşitli kesimleri üzerinde göstereceği hedeflenmemiş etkilerini görmek ve tahlil etmekten alıkoyar.

Bu değerlendirme anlayışı ne yöneticiler ne de değerlendirmeciler tarafından çok sevilir (Scriven, 1983). Bunun nedeni; değerlendirmeyi yöneticinin ya da değerlendirmecinin hedefleriyle programın hedeflerinin yönlendirmiyor olmasıdır. Ayrıca, Scriven (1972) hedeften bağımsız değerlendirmenin, daha önceden belirlenmiş bir yapıya göre yapılmadığı için endişeyi ve hata oranını artırdığını ve beklenmedik sonuçlarla karar vericileri şaşırttığını belirtir.

Özetle, Worthen ve arkadaşlarının da (1997) belirttiği üzere, hedef yönelimli değerlendirmelerde olduğu gibi, bu değerlendirmede de, değerlendirmecinin bakış açısı hedeflerle sınırlandırılıp, hedeflerle doğrudan ilişkili olmayan önemli sonuçları es geçmesi engellenir. Aksine, daha geniş bir perspektifle, değerlendirme sürecinde kalite, değer ve önem ile ilgili yargılara daha önceden yapılan ihtiyaç analizi sonuçlarıyla program değerlendirme sonuçlarının karşılaştırılması ile varılır. Başka bir deyişle, program hedeflerini gerçekleştirdiği ölçüde değil, müşterilerinin ihtiyaçlarını karşıladığı ölçüde değerli kabul edilir.

4.3. Biçim İşlemcisi (Modus Operandi): "Bir etkiyi yaratan sebebi ortaya çıkarmak için olaylar zincirini incelemeyi gerektiren prosedürler" bütünü olarak tanımlanmıştır (Scriven, 1981, s.98). Deneysel ve yarı-deneysel desenlemelere alternatif olarak Scriven tarafından sunulan bu yöntem, bir etkiyi yaratan durumun derin bir şekilde analiz edilmesini sağlar. Bu yöntemde değerlendirmeci, bir etkiyi ortaya koyan tüm olası sebepleri sıralar. İkinci aşamada ise, bu sebeplerden hangilerinin etkinin oluşmasından önce var olduğunu ayırt edip en sonunda gerçek sebebi bulmaya çalışır. Bu yöntemde; bir suçlunun suçu nasıl işlediğini bulmak amacıyla yapılan polis soruşturması örnek olarak gösterilebilir. Böylece, değerlendirmeci değerlendirilen faaliyetin, etkinliğin, alanın, sahanın, programın ve durumun sebep-sonuç zincirleriyle birlikte kapsamlı bir fotoğrafını çekmiş olur.

5. Değerlendirmede Kavramsal Devrim Önerisi

Scriven (2013) değerlendirmenin kavram haritasını netleştirmek ve ona bilimsel bir özerklik kazandırmak adına yoğun bir çaba sarf etmiş ve yukarıda sıralanan kavramları önermiştir. Bu kavramlar bu alanda bazı gelişmeleri teşvik etmiştir (Worthen ve diğerleri, 1997). Değerlendirmenin bir disiplin ya da disiplin ötesi bir alan olduğu, olması gerektiğiyle ilgili farklı görüşler vardır. Worthen ve arkadaşları (1997) "değerlendirmeyi bir disiplinden çok çeşitli materyal, program ve çabanın değerinin belirlenmesini amaçlayan sosyal bir süreç ya da aktivite" olarak görmüşler ve onu farklı disiplinlerden yararlanan bir uygulama alanı olarak nitelendirmişlerdir. Ancak; Scriven (2013) bir disiplinin sahip olması gereken özelliklerin ışığında, değerlendirmeyi çeşitli disiplinlere katkıda bulunan disiplinler ötesi bir alan ve disiplin olarak kabul etmektedir. Disiplinler ötesi bir alan olarak kabul görmesi için, değerlendirmeye ilişkin kavram haritamızda yapmamız gereken 3 aşamalı devrim önermektedir (Scriven, 2013):

- ✓ *Birinci devrim*, değerlendirmeyi bilimle ilgisi olmayan bir etkinlik olarak görmekten ziyade, bilimsel olarak saygın bir etkinlik olarak görmeyi gerektirmektedir. Değerlendirmenin bilimsel bir değerinin olmadığını savunan görüş, daha çok bilimin değerden bağımsız olması gerektiği anlayışına dayanmaktadır. Ancak, Scriven (2013) bu görüşü bilim adamlarının çalışacakları alanı seçerken, yapacakları uygulamalara karar verirken bir değer yargısı taşıdıklarını ve bu bakımdan bilimin değerden bağımsız olamayacağını ileri sürerek çürütür.
- ✓ *İkinci devrim*, değerlendirmeyi bilimsel olarak saygın bir konumdan alfa disiplin olarak görmeyi gerektirir. Alfa disiplin olarak değerlendirme, diğer disiplinlere iyiyi kötüden ayırt etme kalitelerini kanıtlamak, verilerinin, hipotezlerinin meşruiyetini ortaya koymak için kaynaklık eder. Scriven (2013) değerlendirmenin disiplinler arası rolünü vurgulamak için, bir güneş sistemi metaforunu kullanır. Diğer disiplinleri yıldızlar; değerlendirmeyi de onların muhtaç

olduğu güneş olarak metaforlaştırır. Son zamanlarda anestezi ve çeşitli ilaçlarla ilgili araştırmaların sonuçlarının yanlış olduğunun ortaya çıkması, her gün altı bardak su içmenin sağlığa yararı ve ıspanakla ilgili bildiğimiz şeylerin gerçek dışı olduğunun anlaşılması, disiplinlerin profesyonel anlamda değerlendirmeye vermeleri gereken önemi gözler önüne sermektedir (Scriven, 2013). Panellerin ya da akran değerlendirmesinin bilimsel anlamda geçerlilik sıkıntısı vardır. Bu bakımdan, disiplinlerin, değerlendirmenin öğretilmesi konusunda hassas olmalarını ve ona gereken önemi vermeleri belirtilir.

- ✓ *Üçüncü devrim* ise değerlendirmeyi bir alfa disiplinden öte uygulamalı sosyal bilimler için bir paradigma, disiplinler ötesi bir alan olarak görmeyi gerektirir. Böylesi bir alan, diğer disiplinlerin kendi modellerini oluşturmaları için geçerli bir model sunar. Bu da sosyal ilimlerdeki her disiplinin öncelikle değerlendirmeye dayalı olan ve olmayan boyutları olduğunu kabul etmelerini gerekli kılar. Scriven (2013) bu durumu açıklarken, sosyal bilimler için nasıl ki istatistik kullanmamak mümkün değilse değerlendirme yapmamanın da mümkün olmadığını belirtir. Çünkü uygulamaya dayalı, ekonomi, eğitim, sosyoloji gibi disiplinlerin değerlendirme gerektiren sorular sorarlar ve bu soruların cevapları da öncelikle değerlendirmeyle ilgili böylesi paradigmatik bir kabullenışı gerektirir.

Scriven (2013), birikimli olarak ilerleyen bu kavramsal devrimlerin gerçekleşmesinin zor bir süreci beraberinde getireceğini belirtir. Ancak, bu devrimler daha fazla kaynak israfı ya da yanlış değerlendirme sonuçlarının getirdiği bedelleri ödememek ve topluma ödetmemek için gerçekleşmelidir. Bunun ön şartı da, üç aşamalı devrimi kabul ederek, sosyal bilimlerin er ya da geç değerlendirmenin, disiplinlerinin en temel yapı taşlarından birisi olduğunu anlaşılmasıdır.

6. Anahtar Değerlendirme Kontrol Listesi (Key Evaluation Checklist-Kec)

Scriven (1981), değerlendirmeyi daha önceden belirlenen standartlara dayanarak değerlendirilen şeyin kalitesi, değeri ve önemi ile ilgili bir yargıya varma süreci olarak tanımlamıştı. Bu görüşüne uygun olarak çok detaylı hazırladığı kontrol listesinde de, değerlendirme sürecinde neden bir kontrol listesi kullandığını şu şekilde açıklamıştır: "Değerlendirme çok karmaşık bir iştir ve değerlendirilen alanların gerektirdiği deneyim de geçerli kontrol listelerinin kullanımını zorunlu kılmaktadır" (Scriven, 2005, s.1). Scriven (2005) bu kontrol listesinin değerlendirmenin ne kadar karmaşık ve çok boyutlu bir aktivite olduğuna işaret ettiğini ifade eder. Aynı zamanda, değerlendirmenin, sosyal bilimlerin geleneksel yöntemlerinin tekdüze bir şekilde uygulandığı bir süreçten çok daha fazlası olduğunu da göstermektedir (Scriven, 1981).

Ürün, süreç, kurum, otel, restoran, eğitim, hizmet, politika gibi birbirinden farklı birçok şeyin değerlendirmesinde kullanılacak bu liste, değerlendirmeyi değerlendirme sürecinde ne yapması ve nelere dikkat etmesi gerektiği konusunda bilgilendirir (Scriven, 2013). Bu liste, değerlendirme sürecini; hazırlık, temellendirme, değerlendirme ve sonuç olmak üzere dört ana boyuta ayırır. Her bir boyutta değerlendirmeye nelere dikkat etmesi ve ne yapması gerektiğiyle ilgili bilgiler veren toplam 18 adet alt aşama bulunmaktadır. Bu aşamaları içeren dört ana boyut aşağıdaki gibi özetlenebilir (Scriven, 2013, KEC):

Hazırlık: Bu boyutta değerlendirmeci, değerlendirilen şeyin bağlamını, paydaşlarını ve müşterilerini belirler ve bunlarla ilgili özet yapar. Tüm bağlam, paydaş ve müşteriler netleştirildikten sonra, uygun değerlendirme yaklaşımına ve yöntemine karar verilir.

Temellendirme: Bu aşamada değerlendirilen programın tarihi ve güncel bağlamında başarı ve başarısızlığına neden olabilecek faktörler belirlenir. Programa ilişkin tüm bileşenler ve terimler tanımlanır. Programın ana müşterileriyle birlikte programdan dolayı olarak etkilenecekler de belirlendikten sonra, kaynakların ne denli verimli kullanıldığı ortaya koyulur. Son olarak da, programın paydaşlarının, müşterilerinin ve dolaylı olarak programdan etkileneceklerin ve toplumun tamamının değer yargıları, tepkileri belirlenir.

Değerlendirme: Değerlendirme sürecinde toplanan verilere dayalı olarak, sürecin kalite, değer ve önemine ilişkin değerlendirme yapılır. Daha sonra, programın iyi ya da kötü; hedeflenen ya da hedeflenmeyen; kısa ya da uzun vadeli etkileri tespit edilir. Bunu nitel ve nicel yöntemlerle yapılan maliyet analizi izlemektedir. Program benzer sonuçları olan benzer kaynaklarla ya da daha aزیyla gerçekleştirebilecek başka programlarla karşılaştırılır. Son olarak da programın, kendi zamanının ötesinde transferinin mümkün olup olmadığına bakılıp bir genelleme yapmaya çalışılır.

Sonuç: Bu aşamada ilk olarak değerlendirmeci, daha önceki basamaklarda elde edilen sonuçların paydaşlara nasıl aktarılacağına, sunulacağına karar verir. SWOT bu sunumda takip edeceği yöntemlerden biri olabilir. Bu sunum paydaşlara, programın aksayan ve güçlü yönlerini aktarmalı, onları bekleyen sorunlar konusunda bilgilendirmeli ve sorunların nasıl çözüleceği ya da programın nasıl revize edileceğiyle

ilgili tavsiye niteliğinde görüşleri de içermelidir. Farklı paydaşlara uygun zaman ve yerlerde sonuçların aktarılmasıyla birlikte, değerlendirmeci değerlendirmenin sorumluluğunu almalıdır. Değerlendirme de son olarak değerlendirmenin değerlendirmesi yapılır. Çünkü insan hataya meyillidir ve akran değerlendirmesi değerlendirmenin güvenilirliğini artırabilir. Etik kaygılarla yapılması gereken değerlendirmenin değerlendirmesi geçerlilik, güvenilirlik, maliyet, kullanılabilirlik, uygunluk, dayanıklılık gibi kıstaslar takip edilerek gerçekleştirilir.

Tüm bu boyutlara ek olarak 1971 yılında iki sayfa halinde ortaya çıkan son hali 51 sayfa olan kontrol listesinde, Scriven (2013) ayrıca 11 madde halinde yapılması gerekenleri özetlemiştir:

- ✓ Değerlendirilen şeyin bağlamını tanımlamak
- ✓ Değerlendirilen şey ile ilgili kalite, değer standartlarını belirlemek
- ✓ Bu standartları önemine göre sınıflamak
- ✓ Her boyutta performans göstergelerini ölçecek uygun ölçekleri belirlemek
- ✓ Her boyutu uygun tanımlamalar, ihtiyaçlar kullanarak ve karşılaştırmalar yaparak çeşitli müşterileri tekrarlamak
- ✓ Her boyutta başarı ve başarısızlığı gösterecek ampirik göstergeleri belirlemek
- ✓ Her göstergeye ilişkin başarıyı ölçmek
- ✓ Gösterge puanlarını bütünleştirerek her boyutla ilgili bir başarı düzeyi belirlemek
- ✓ Gerek duyulduğunda, her boyutun önem derecesini hesaplayarak bu boyutlardan alınan puanları bütünleştirip genel başarı düzeyini belirlemek
- ✓ Genel kıstasları uygulamak
- ✓ Sonuçları raporlaştırmak

7. Değerlendirmenin Değerlendirmesi

Scriven (1981), değerlendirmenin değerlendirmesini tamamlanan bir değerlendirmenin bazı kıstaslara bağlı olarak başka bir değerlendirici tarafından yeniden değerlendirilmesi olarak tanımlar. Değerlendirmeyi yapan kişinin yapmış olabileceği hatanın başka biri tarafından görülerek yanlış değerlendirme sonuçlarının bertaraf edilmesi ve böylece güvenirliliğin artırılması gibi sebeplerden değerlendirmenin değerlendirmesini yapmak etik bir zorunluluktur (Scriven, 2011). Scriven (2011) tarafından değerlendirilen kalite ve değeriyle ilgili yargıya varmak için yapılan bir etkinlik olarak tanımlanan değerlendirilen değerlendirmesi daha önce bahsedilen Anahtar Kontrol Listesi (KEC) kullanılarak ilk değerlendirme ile ikinci değerlendirilen sonuçları arasındaki tutarlılığa bakılarak ya da 6 kıstastan oluşan Değerlendirmenin Değerlendirmesi Kontrol Listesi kullanılarak gerçekleştirilebilir. Bu altı kıstas ve ifade ettikleri aşağıdaki gibi özetlenebilir:

Geçerlilik: Değerlendirmenin bazı bilimsel standartlara, kurallara, amacına ve bağlamsal şartlarına uygun olarak yapılıp yapılmadığı değerlendirilir.

Güvenirlilik: Değerlendirmecinin yeterli düzeyde tecrübe ve bilgisiyle paydaşlara gerekli güveni verip vermediği değerlendirilir. Ayrıca, değerlendirilmecinin paydaşlarla ya da yöneticilerle değerlendirilen tarafsızlığını etkileyecek herhangi bir bağının olup olmadığı değerlendirilir.

Açıklık: Değerlendirme ve uygulama aşamasında hatalara neden olabilecek paydaşlara, müşterilere ve programdan dolaylı olarak etkilenenlere ilişkin bilgilerin açıkça belirtilip belirtilmediğine bakılır.

Uygunluk: Değerlendirmenin yasal, etik ve kültürel hassasiyet ve kurallara uygun bir şekilde yapılıp yapılmadığı değerlendirilir.

Maliyet: Bu aşamada, fayda-maliyet analizi ile bireysel ve toplumsal kazançlar ve maliyetler karşılaştırılarak bir değerlendirme yapılır.

Genellik: Bu değerlendirilen kalitesi ve değeri için bir zorunluluk değil artı bir puandır. Genellik, kıstasların değerlendirilmesi yapılan değerlendirilen aynı gruba başka bir zamanda uygulandığında aynı sonuçları verip vermediğini ifade eden tekrar kullanılabilirlik; başka programların değerlendirilmesinde kullanılabilirliğini ifade eden uyarlanabilirlik; programda yapılacak küçük değişikliklere rağmen benzer sonuçlar verme anlamına gelen dayanıklılık gibi alt gereklilikleri vardır.

Son kıstas zorunlu olmayan ancak değerlendirilen değerinin ekstra önemini ve kalitesini gösteren genellik kıstası ile birlikte 6 kıstasa ya da standarda göre değerlendirilmesi yapılmış bir değerlendirilen değer, kalite ve önemini belirlemeyi amaçlamaktadır. Bu süreç değerlendirilen sonuçlarının hizmet ettiği amaçlar düşünüldüğünde, bu konuda yapılacak olası hataların maliyeti dikkate alındığında sadece değerlendirilmeci ve yönetici için değil, tüm paydaşlar, müşteriler, programdan dolaylı olarak etkilenebilecek tüm kişiler ve hatta toplumun büyük bir kısmı için hayati önem taşımaktadır.

8. Sonuç

Tüm bu anlatılanlar özelinde, değerlendirmenin bilimsel olarak hak ettiği ilgiyi görmesi için Scriven düşüncelerimizde bazı devrimsel nitelikte değişiklikler yapmamız gerektiğini mantıksal bir çerçevede açıklamaktadır. Bu bağlamda, Trochim (1998), Scriven'in değerlendirme alanına yaptığı katkıları üç başlıkta toplaması temel alınabilir. Bu üç nokta, çalışma boyunca anlatılmaya çalışılan Scriven'in değerlendirme alanına getirdiği değişiklikleri ve yaptığı katkıları somut bir şekilde özetlemek için faydalı olacaktır.

Bu üç temel katkının ilki, "Minimalist Theory" başlığı altında özetlenmiştir. Teorileri "içsel" ve "dışsal" şeklinde sınıflandıran Scriven, program değerlendirme alanında programın nasıl çalıştığı ile ilgilenen içsel teoriye değil etki-tepki ilişkisi dahilinde programın istenen ve istenmeyen etkilerinin ortaya konduğu dışsal teoriye önem vermektedir. Bu da neticede değerlendirme alanında minimum düzeyde teori kullanılmasını önerdiği ve alanı uygulamaya daha çok yaklaştırdığı bir anlayışın hakim olmasını teşvik etmiştir.

İkinci olarak, değerlendirmeyi yeniden tanımlayarak onu disiplinler ötesi bir alanda konumlandırmasıdır. Bu bağlamda, alana çeşitli kavramsal katkılarda bulunmuştur. Bilimin her alanında, her safhada kullanılmasına ihtiyaç duyulan böylesi bir değerlendirme anlayışı değerlendirmeyi değerlendirilenin önem ve kalitesinin ampirik veri ve kanıtlara dayalı olarak belirlendiği bir uzmanlık mesleği olarak tanımlamıştır.

Scriven'in değerlendirmeye getirdiği üçüncü temel katkı kendisinin değerlendirme tanımının doğal bir sonucudur. Değerlendirme sürecinde, değerlendirmenin çok boyutlu yapısını göz önünde bulundurarak madde ve kriter odaklı kontrol listeleri arasında ayırım yapmıştır. Bu doğrultuda, değerlendirme sürecinde tüm süreçleri kapsayacak şekilde detaylı kriter odaklı kontrol listelerinin kullanılmasına hassasiyet göstermiştir. Tüm bu anlayışına ve yaptığı tanıma tutarlı olarak, hem değerlendirme hem de değerlendirmenin değerlendirilmesi için bazı standartlar içeren detaylı kontrol listeleri hazırlamıştır. Böylece, uygulamalı bir alan olduğunu söylediği değerlendirmeyi, uygulama alanına taşımış ve sürecin karmaşıklığını da değerlendirmeciye her boyutta ne yapması gerektiğini aşama aşama detaylı bir şekilde anlatarak azaltmaya çalışmıştır.

Sonuç olarak, değerlendirmenin kendine has bir yapısı olan ve diğer disiplinleri de besleyen uygulamalı disiplinler arası bir alan olduğunu belirten Scriven değerlendirmeyi bir şeyin değer, kalite ve öneminin bazı standartlara dayalı olarak performans göstergelerinin ölçülmesi ve bunlarla ilgili yargıda bulunmayı gerektiren karmaşık ve bir uzmanlık işi olarak tanımlamıştır (Scriven, 2007).

KAYNAKÇA

- ALKIN, M. C. (2013). *Roots of evaluation: A wider perspective of theorists' views and influences* (2nd ed.), Newbury Park, CA: Sage
- MATHISON, S. (2005). *Encyclopedia of evaluation*, Newbury Park, CA: Sage
- SCRIVEN, M. (1981). *Evaluation thesaurus* (3rd ed.), California: Edge press
- SCRIVEN, M. (1986). "New frontiers of evaluation". *American Journal of Evaluation*, 7 (1), pp. 7-44.
- SCRIVEN, M. (1996). "Types of evaluation and types of evaluator". *American Journal of Evaluation*, 17 (2), pp. 151-161.
- SCRIVEN, M. (1999a). "The nature of evaluation part i: Relation to psychology". *Practical Assessment, Research & Evaluation*, 6 (11).
- SCRIVEN, M. (1999b). "The nature of evaluation part ii: Training". *Practical Assessment, Research & Evaluation*, 6 (12).
- SCRIVEN, M. (2005). "The logic and methodology of checklists". Интернет-ресурсы http://www.wmich.edu/evalctr/checklists/papers/logic&methodology_dec07.pdf
- SCRIVEN, M. (2007). "The logic of evaluation", In H.V. Hansen, et. al. (Eds), *Dissensus and the Search for Common Ground*, CD-ROM (pp. 1-16). Windsor, ON: OSSA.
- SCRIVEN, M. (2012). *Evaluating evaluations: A meta-evaluation checklist*.
- SCRIVEN, M. (2013). *Conceptual revolutions in evaluation: Past, present and future. Evaluation roots: A wider perspective of theorists' views and influences*. Newbury Park, CA: Sage
- SCRIVEN, M. (2013). Key evaluation checklist (KEC).
- SHADISH, W. R. (1998). "Evaluation theory is who we are", *American Journal of Evaluation*, 19 (1), pp. 1-19.
- TROCHIM, W. M. (1998). "An evaluation of Michael Scriven's "Minimalist theory: The least theory that practice requires". *The American Journal of Evaluation*, 19 (2), pp. 243-249.
- WORTHERN, B. R., SANDERS J. R. & FITZPATRICK J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*, New York: Longman, Inc.
- UŞUN, S. (2012). *Eğitimde program değerlendirme: Süreçler, yaklaşımlar ve modeller*, Ankara: Anı Yayıncılık