

EĞİTİMDE PROGRAM DEĞERLENDİRME SÜRECİNİN TEMEL İŞLEMLERİ

Doç. Dr. Zeki KAYA*

Özet

Bu yazıda, daha işlevsel program değerlendirme çalışmalarının gerçekleştirilmesine katkı sağlamak amacıyla, değerlendirme sürecinde yer alan temel işlemler ve bu işlemlerin gerçekleştirilmesinde nelerin dikkate alınması gerektiği açıklanmıştır.

Anahtar sözcükler: Eğitim programı, program değerlendirme.

Abstract

It is explained in this present study what the basic procedures are and what should be considered while applying these procedures in the evaluation process. The aim in doing this is to make a contribution to carrying out a more functional program evaluation.

Key Words: Curriculum, curriculum evaluation

Giriş

Program değerlendirme, bir programdaki bütün boyutların ya da bir veya birkaç boyutun etkisinin, etkinliğinin ve sahip olabileceği tüm çıktılarının yargılanması için bilgilerin toplanması, çözümlenmesi ve yorumu olarak kabul edilebilir.

Program değerlendirme çalışmalarında sistematik bir sürece uymak gerekir. Sistematik bir sürece uyulması, değerlendirmenin, bir taraftan yapısallaştırılmasını sağlarken, diğer taraftan da daha kapsamlı hale getirilmesine yardımcı olmaktadır. Değerlendirmenin tasarlanmasından, tamamlanmasına kadar olan süreçte, birbirleriyle dinamik ilişkileri olan pekçok işlem vardır. Bu işlemlerden bazıları var ki, bunlar, değerlendirme sürecinde alınması gereken kararları yansıtır ve o nedenle de program değerlendirmenin temellerini oluşturan işlemler olarak nitelendirilirler. Program değerlendirmenin temelleri olabilecek işlemler şunlardır:

- Anlamın kesinleştirilmesi.
- Amacın belirlenmesi.
- Anahtar tarafların belirlenmesi.
- Olanakların ve engellerin belirlenmesi.
- Yanıt aranacak soruların belirlenmesi.
- Tasarımın kesinleştirilmesi.

* Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Eğitim Bilimleri Bölümü Öğretim Üyesi.

- Verilerin toplanması.
- Verilerin çözümlenmesi.
- Sonuçların yorumlanması ve ilgililere bildirilmesi.

Bu işlemler, değerlendirmenin yapacağı seçimleri ve seçimlere dayalı sonuçları doğrudan etkilemektedirler. Bu nedenle, sözü edilen işlemlerin daha açık hale getirilmesi uygun olacaktır.

Anlamın Kesinleştirilmesi

Birçok değerlendirme modeli ve bu modellerin dayandığı program değerlendirme yaklaşımları vardır. Gerek program değerlendirme modellerinin, gerekse program değerlendirme yaklaşımlarının her biri farklı görüşlere dayanmaktadır. Bu farklı görüşler, program değerlendirme hakkındaki felsefe ve inançlardan kaynaklanmaktadır. Farklı zamanlarda, farklı amaçlar için, farklı değerlendirme modellerine ve yaklaşımlarına uyulabilmektedir (Kaya, 1996, s. 9). Bu nedenle, program değerlendirmenin, değerlendirmeci ve diğer ilgililer için ne anlama geldiğinin açıkça ortaya konulması gerekir.

Değerlendirmenin, değerlendirme anlamı hakkında kendisine uyarladığı model ya da yaklaşım, değerlendirme çalışmasının tasarlanması ve düzenlenmesinde birçok temel kararlara ilişkin soruları yanıtlamada yardımcı olmaktadır. Başka bir deyişle, değerlendirmenin ne anlama geldiğinin belirlenmesi, temel kararlara ilişkin soruların yanıtlarını şekillendirmektedir. Örneğin, "yapılan değerlendirme, programın değerinin resmi bir yargısı mıdır veya programın çıktılarının, beklenen gibi olup olmadığının belirleyicisi midir?" ya da "programın değerini, değerlendirmeci tarafından varılan yargı mı ortaya koyar?", sorularının yanıtlanması ne tür verilerin toplanacağını ve bu verilerin kim ya da kimler tarafından toplanacağı hakkında verilecek kararları şekillendirmektedir.

Amacın Belirlenmesi

Program değerlendirmede temel amacın, programın etkililiği hakkında yargıda bulunmak ve programdaki aksaklıkların programın hangi öge ya da öğelerinden kaynaklandığını belirleyerek gerekli düzeltmelerin yapılmasına olanak sağlamak olduğu söylenebilir (Erden, 1995, s.10). Bu temel amaç doğrultusunda gerçekleştirilen değerlendirme, bazan programın ilgililerinin gereksinimlerini ve koşullarını karşılamak, bazan da programın geliştirilmesine yardım etmek için yapılabilmektedir.

Değerlendirmenin başlangıcında, "niçin yapılmaktadır?" ve "kapsama alınacak olan konular, daha geniş bir konunun bir parçası mıdır?" gibi sorulara yanıt aramak gerekir. Ayrıca, çalışmanın kimler tarafından istenilebileceğini ve desteklenebileceğini belirlemek de önemlidir. Değerlendirmenin, değerlendirmeyi isteyenlerin ve destekleyenlerin ne çalışmadan ne beklediklerini bilmesi gerekir. Bu, değerlendirmenin;

- değerlendirmeyle ilgili tarafları belirlemesine,
- taraflar arasındaki ilişkiyi öğrenmesine,
- çalışma düzenini etkileyebilecek belirlenmemiş konuları belirlemesine ve
- elde edilen verilerin ve bulguların kullanımına, farklı grupların açık olup olmadığını belirlemesine, yardımcı olur.

Anahtar Tarafların Belirlenmesi

Program değerlendirme çalışmasının başlangıcında şu soruların yanıtlanması önemlidir:

- Değerlendirme etkinliğini kim başlatır?
- Değerlendirmenin giderlerini kim karşılar?
- Değerlendirmeyi kim izler?
- Değerlendirmeci kimin için çalışır?
- Değerlendirme raporu kime verilecektir?

Ne var ki, bu soruların yanıtları her zaman açık olmamaktadır. Sözü edilen sorulara ait yanıtların açık olmamasının, belli başlı iki nedeni vardır. Bunlardan birincisi, birçok değerlendirme çalışması, programı uygulayan kurum, bir bağış kurumu, bir grup ya da kişi tarafından maddi olarak desteklenebilir. Değerlendirilmeye aday program ise, tamamen farklı bir kurum, söz gelimi Milli Eğitim Bakanlığı, bir vakıf, bir şirket ya da grup tarafından uygulanıyor olabilir. Benzer şekilde, değerlendirmeyi maddi olarak destekleyenler ile değerlendirmeyi organize edenler ve değerlendirmeyi izleyenler aynı kurum ya da kişiler olmayabilir.

İkincisi, müşterinin, maddi destek sağlayanın ve değerlendirmeyle ilgili çıkar gruplarının güç ve yetkilerinin farklı olmasıdır. Değerlendirmeyle ilgili çıkar gruplarının istekleri müşterilerin isteklerinden farklı olabilir.

Değerlendirmecinin, çalışmanın başlangıcında, müşteri, eğer varsa maddi olarak destekleyen ve çıkar gruplarından oluşan anahtar tarafları belirlemesi önemlidir. Müşteri, çalışmayı başlatan ve değerlendirmeciye görevlendiren kişi ya da gruptur. Maddi olarak destekleyen, çalışmanın ödemesini yapan kişi ya da gruptur. Genelde program değerlendirme çalışmalarında maddi olarak destekleyen ile müşteri aynıdır. Buna, Milli Eğitim Bakanlığının uyguladığı bir programın belli bir ücret karşılığında değerlendirmesini yaptırması ya da o programın değerlendirilmesinde ölçme araçlarını çoğaltması, ölçme araçlarını uygulaması, istatistiksel çözümlenmesi yaptırması değerlendirmenin çoğaltımı ya da basımı gibi destekleri örnek olarak verilebilir. Bazı değerlendirmelerde müşteri ve maddi olarak destekleyen farklı olabilmektedir. Söz gelimi, maddi olarak destekleyen, bir bağış kurumu, müşteri ise Milli Eğitim Bakanlığı olabilmektedir. Buna da, eğitim uygulamalarının niteliğini artırmak amacıyla kurulmuş bir vakfın, Milli Eğitim Bakanlığımıza bağlı bir kurumda uygulanan bir programın değerlendirilme giderlerini karşılaması, örnek olarak verilebilir. Burada vakıf maddi olarak destekleyen, bakanlık ise müşteri konumundadır. Böyle durumlarda müşteri, çoğu zaman, bağışları alıp giderleri karşılayan ve projeyi düzenleyen yönetim birimi olmaktadır. Yani müşteri, değerlendirmeci ile birlikte çalışma kapsamını ve çalışmada yanıtları aranacak temel soruları belirlemek rolünü de üstlenebilmektedir. Değerlendirme raporunun kime verileceği açıkça belirlenmediğiçe, değerlendirme raporunun son şekli, genelde, müşteriye sunulur.

Değerlendirmede, değerlendirmeye özel yarar sağlayan ve değerlendirmenin düzenlenmesinden ve/veya sonuçlarından etkilenebileceklerine inanan taraflar da vardır. Bunlar, programın çıkar grupları olarak kabul edilebilir. Çıkar grupları arasında genelde şunlar bulunur (Windham ve Chapman, 1990, s.42) :

- Programın planlayıcıları.
- Katılımcılar.
- Öğreticiler.
- Yöneticiler.
- Programın uygulanması için maddi destek sağlayan kişi ya da kurumun temsilcileri.
- Programla ilgili kurumun yetkilileri.
- Program yürütülmesine ayrılan kaynakların farklı şekilde kullanılabilmesine inananlar.
- Değerlendirmeden etkileneceklerine inananlar.
- Benzer değerlendirmeleri maddi olarak destekleme yeterliliğine sahip bağış kurumları.

Görülüyor ki, değerlendirmedeki çıkar grupları, müşteri ve maddi olarak destekleyenlere göre çok daha fazladır.

Çıkar gruplarının belirlenmesi işleminin, değerlendirmeci tarafından gerçekleştirilmesi gerekir. Bunun belli başlı üç nedeni vardır. Bunlardan birincisi, değerlendirmecinin, programın hizmet verdiği gruplardan, programın yararı konusundaki sorulara ne ölçüde ve nasıl yanıt aldığına dayanır. Bu grupların bilgi gereksinimi değişmektedir. Program yöneticilerinin, programı geliştirmek için bilmek isteyecekleri, anne-babaların, çocuklarının programda yer almasına izin vermeye karar verirken bilmek istediğinden (bireysel başarı, katılımın gizli maliyetleri) farklı olabilir. Çıkar gruplarının bilinmesi, değerlendirmeciye önemli yardımlar sağlar. Bu yardımlar, hangi tür verinin toplanacağına ve çalışma sonuçlarının nasıl bildirileceğinin, belirlenmesi konularında odaklanır (Windham ve Chapman, 1990, s.42).

İkincisi, çıkar grupları önemli olduğuna inandıkları ölçütlerde, uyguladıkları standartlarda ve güvendikleri ölçüm türleri konularında farklı düşünebilirler. Değerlendirmede, çıkar gruplarının gereksinimlerinin de karşılanması söz konusu olduğuna göre, onların düşünceleri arasındaki farkların anlaşılması gerekir.

Üçüncüsü, tarafların etki ve güç bakımından farklı olmalarıdır. Değerlendirmecinin, anahtar taraflar olan grupların arasında ve içinde etkin olan sosyal ve politik dinamikleri anlaması gerekir. Bu anlayış, en az politik etkiye sahip grupların gereksinimlerinin göz ardı edilmemesini sağlar (Windham ve Chapman, 1990, s.42). Değerlendirmecinin, taraflara ait sosyal ve politik dinamikleri hatasız olarak anlaması, taraflararası etkinliklerde değerlendirmeciye kılavuzluk eder.

Değerlendirmede, her çıkar grubu eşit derecede öneme sahip değildir. Bu nedenle, çıkar gruplarının hepsinin aynı ölçüde değerlendirme kapsamına alınması gerekmez. Kapsamda hangi çıkar grubunun daha çok yer alacağına ilişkin kararlar, değerlendirme hazırlığının başlamasıyla daha kolay verilebilir. Bu kararlarda değerlendirilmecinin etkin rol alabileceği söylenebilir. Değerlendirmeci karar vermede bazı etmenlerden etkilenebilir. Bu etmenlerin başlıcaları; grupların programa yakınlığı, program çıktılarının davranışları ve sosyal adalettir.

Olanakların ve Engellerin Belirlenmesi

Müşteriler genelde, değerlendirme etkinliklerinde yer alan işlemler, bu işlemlerin ne kadar süreceği ve ne kadar maliyet getireceği konularında fazla fikre sahip değildirler. Değerlendirme çalışmasının olası maliyetini ve bitirme zamanını belirlemek, değerlendirmenin görevleri arasında yer alır. Değerlendirmenin kendi ülkesi dışında gerçekleştirileceği program değerlendirme çalışmaları için, bunun zor olduğu söylenebilir. Çünkü, değerlendirmeci verileri toplamak için gerekli kaynakları ve zamanı, verileri çözümlene olanaklarını, yükleme ve postalama zamanını, yerel kurumlardan ve bağış kurumlarından beklenen ortaklık derecesi ve benzerlerini yeterince bilmeyebilir. Lojistik desteklerle ilgili olan bu belirsizliklerin yanı sıra, özellikle az gelişmiş ülkelerdeki değerlendirme çalışmaları için, gelişmiş ülkelerdeki değerlendirme çalışmalarında yer almayan bazı işlemler ve adımlar gerekli olabilir. Örneğin, çalışmaya ilişkin politika ve program belgelerinin o ülkenin diline çevrilmesi ve yapılan çevirilerin doğrulanması gerekebilir (Windham ve Chapman, 1990, s.43). Ne var ki, gelişmekte olan ülkelerde bile, bu konuya gerekli özenin gösterilmediği izlenimi edinilmektedir. Buna, Türkiye ile Dünya Bankası arasında gerçekleştirilen ikraz antlaşmasına dayalı olarak 1985 yılında uygulamasına başlanmış olan Yüksek Öğretim Kurulu / Dünya Bankası Endüstriyel Eğitim Projesi kapsamında çevirisi yapılan bazı ölçme araçları örnek gösterilebilir. Bu tip projelerin değerlendirme boyutuyla ilgili ölçme araçlarının çevirilerinin doğrulanması için yeterli zaman ayrılmalıdır. Yapılan çevirilerde de, gerekli özen gösterilmelidir. Çünkü, bu çeviriler çalışmayı doğrudan etkileyebilmektedir.

Programların ve program değerlendirme çalışmalarına ilişkin belgelerinin çevrilmesi, çevirmenden kaynaklanan bazı risklere neden olabilir. Doğru olan, değerlendirmenin değerlendireceği programın ait olduğu ülkenin dilini iyi bilmesidir. Değerlendirmeci, değerlendireceği programın ait olduğu ülkenin dilini iyi bilmezse, çeviri süreci bir filtre işlevi görebilir.

Değerlendirilecek programla ilişkili üst düzey görevliler, değerlendirmenin dilini konuştuğunda, değerlendirmeci çevirinin gereksiz olduğuna inanabilir ve bu görevlilerle doğrudan konuşmayı yeğleyebilir. Bu durumda, değerlendirmeci, programda önemli yerleri olduğundan değil, kendisiyle en açık iletişimi kurabildikleri için belli görevlilere dayanmayı deneyebilir.

Yanıt Aranacak Soruların Belirlenmesi

Bazı değerlendirme çalışmaları, önemli soruları yanıtlamadığı için eleştirilebilir, hatta işe yaramaz kabul edilebilir ve kesinlikle karar almada kullanılamaz. Değerlendirmelerin, temel konuları ele almada, diğer bir deyişle temel sorulara yanıt bulmada başarısız olmasının dört nedeni vardır:

1. Hiçbir değerlendirme sorusu düşünceyi anlatacak şekilde belirtilmemiştir. Değerlendirmenin ve/veya müşterinin çabalarını yansıtan değerlendirme sorularını açıkça belirtmede başarısız olunmuştur.
2. Değerlendirmeci, müşterinin karşılaştığı temel sorunları ve kararları açıklamada başarısız olmuştur ve uygun olmayan değerlendirme soruları belirlemiştir.

3. Değerlendirme sürecinin başlangıcında, yanıtı aranacak sorular uygun olarak belirlenmiş, fakat zamanla, müşterinin ilgileri ve kaygıları değişmiştir.
4. Değerlendirme sonuçları, çalışmanın alanını aşan sorunlar ortaya çıkarmıştır.

Bu nedenler, değerlendirmeciye, müşterinin karşılaştığı önemli sorunları açık hale getirmede, değerlendirme sorularını temel taraflara üstlendirmede ve bu sorulara ortaklaşa karar vermede ve değerlendirmede yanıtlanacak soruların sınırlarını çizmede başarısızlığa sevk eder.

Tasarımın Kesinleştirilmesi

Tasarım kavramı, genellikle, bir yapıtı oluşturmak için fikir, işlev ve ürün öğelerinin birleştirilip bir şekle konulması anlamında kullanılmaktadır (Saylan, 1995, s.13). Bu kavram, program yapma sürecinde olduğu gibi, program değerlendirme sürecinin de odağında yer almaktadır. Değerlendirme süreci tasarısı, yanıtlayıcıların seçimi ve veri toplama işlemlerine ilişkin kesin stratejiyi gösterir. Tasarımın kesinleştirilmesi, değerlendirmenin düzenlenmesinde, en önemli konular arasında yer alır. Çünkü kullanılan bir değerlendirme tasarısı, büyük oranda, çalışma sonuçlarından çıkarılabilecek iddiaların yapısını ve güçlü yönlerini belirler.

Değerlendirme tasarısı, çalışmanın amaçlarını ve kurallarını ortaya koyar. Tasarımın önemi, değerlendirilmekte olan programın değeri konusunda ilgili ve yeterli veriyi vermeyi ve daha bilgili karar alınması için ilgili taraflara olanaklar sağlamayı gerçekçi bir şekilde hedeflemesi ile ölçülebilir.

Bir tasarımın hazırlanmasında, araştırma ile değerlendirme arasındaki farkları akılda tutmak gerekir. Araştırma temeldir; bir arama, öğrenme, bilinmeyenini biliniyor yapma, karanlığa ışık tutma, kısaca bir aydınlanma sürecidir (Gürkan, 1994, s.56). Bu süreç, değerlendirmeye çok benzemekle birlikte, değerlendirmeden farklıdır. Araştırma ile değerlendirme arasındaki iki önemli farkı Windham ve Chapman şu şekilde açıklamaktadırlar (Windham ve Chapman, 1990, s.45):

1. *Önemli olan programın değeri ve yararı hakkındaki fikirlendir. Değerlendirmenin amacı, önemli görülen ölçütlere dayalı olarak, program değerine ilişkin yargıları desteklemektedir. Değerlendirme, işlem ve etki arasındaki rastlantısal ilişkilerin ötesinde önemli ve anlamlı konuları hedefler.*
2. *Değerlendirme sonuçlarının, genelleştirilmesi değerlendirmenin gerekli bir hedefidir. Değerlendirmeler, söz konusu programa ilişkin kesin kararları ele almak için planlanır. Sonuçları, diğer ortamlara genelleştirme çabaları, değerlendirmenin, programın uygulandığı belli bir ortamla ilişkisini içerir. Genellenen bulguların önemi, çalışma kapsamındaki belli değerlendirme sorularıyla değişecektir.*

Bu açıklamalardan, araştırmanın köşe taşı olan genellenmenin, değerlendirmede her zaman temel hedef olmadığı anlaşılmaktadır. Yine aynı açıklamalardan, program değerlendirmenin anahtar çıktısının, program hakkında verilmesi gereken kararları desteklemek olduğu, araştırmanın anahtar çıktısının ise, her zaman kararları desteklemek olmadığı anlaşılmaktadır.

Programın müşterileri ya da maddi destek sağlayanları, programın etkileri hakkında genellenebilen iddialar aradığında, bu konuları ele almak için, değerlendirme yerine araştırma çalışmalarını tercih etmelidirler.

Değerlendirmeciye uygun olan tasarı seçenekleri, deneysel/yarı deneysel, ilişkisel, tarihsel / niteliksel olmak üzere üç grupta toplanabilir (Windham ve Chapman,1990, s.46).

Gerçek deneysel tasarılarla ilişkin değerlendirme etkinliğinde deneklerin yansız olarak dağıtılması ve bir kontrol grubunun kullanılması söz konusudur. Yarı deneysel tasarılar ise genelde yanlı atama vardır. Aynı zamanda da kontrol grubu bulunmamaktadır (Balcı, 1995, s. 250). Yarı deneysel tasarılar, bir deneyin koşullarını karşılamazlar. Yine de bunlar, gerçek bir deneyin yapılamayacağı ortamlar için pratik bir seçim olurlar. Yarı deneysel tasarılar ile gerçek deneyler yapılmaya da, bunlardan elde edilebilecek bulgular, neden - sonuç iddiaları için temel işlevi görebilir.

İlişkisel tasarılar, ölçülen değişkenler arasında, kapsamın uyumluluğunun belirlenmesine dayanır. İlişki, rastlantı olarak yorumlanamaz. Kavramsal bir çatıda yorumlandığında ise, programla ilgili kararlara temel olacak yararlı yapıları ortaya çıkarır. Bu tasarılar, çoğu kez, örneklemdaki kişilerin rastlantısal olarak atanması olasılığı olmadığında kullanılır. Hedef grubun katılımı, rastlantısal seçim yerine ölçütlere dayandığında veya gönüllü olarak gerçekleştiğinde, eğitim programlarında bu duruma sıkça rastlanır.

Tarihsel/niteliksel tasarılar, anlam belirlemesini vurgular. Kuramsal ölçümden çok gelişime odaklanır. Bununla birlikte, niteliksel veri toplama tekniklerinin, görüşmelerde olduğu gibi, yukarıda ele alınan tasarı seçeneklerinin bir dizisiyle kullanılabilirliğini belirtmek gerekir.

Her tasarımın güçlü ve zayıf yönleri vardır. Çalışmanın amacı, uygun bir tasarı seçilmesine temel oluşturur. Bu temelden hareketle, tasarımın seçiminde, şu altı konu göz önünde bulundurulmalıdır (Windham ve Chapman,1990, s.46):

- Yansızlık
- İç geçerlilik
- Dış geçerlilik
- İlişkisel
- Uygunluk
- Gerçekleştirilebilirlik

Yansızlık, çalışma işlemlerinin uygun olarak izlenme derecesiyle ilgilidir. İç geçerlilik etkinliğin gözlenen etkilere ne ölçüde yol açtığı ile ilgilidir. Dış geçerlilik bulguların, başka ortamlara gruplara ve bireylere ne ölçüde genellenebilirliğini gösterir (Balcı, 1995, s.246). İlişkisel, tasarımın çalışma amaçlarına uyum derecesini ve değerlendirme sorularının hedeflenen sonuçları verme derecesini gösterir. Uygunluk tasarımın, sosyal ve ekonomik çevreye uyma derecesiyle ilgilidir. Gerçekleştirilebilirlik, önerilen değerlendirme etkinliğinin, belirlenen ortamda uygulanıp uygulanmayacağı ile ilgilidir.

Tasarımın seçimi çoğunlukla; yansızlık, iç geçerlilik, dış geçerlilik, ilişkisel, uygunluk ve gerçekleştirilebilirlik arasındaki dinamik bağlantıları içerir. Bu konulardan birisinin niteliğinde olan değişim diğer konuların niteliğini de etkiler.

Verilerin Toplanması

Değerlendirme çalışmaları; veri toplama araçları ve teknikleri ile bir dizi işlemlere dayanır. Bunlar aşağıda verilmiştir:

- Anketler
- Testler
- Puan katsayıları
- Q Sıralamaları
- Toplama sistemleri
- Gözlem sistemleri
- Katılımcı gözlemleri
- Görüşmeler
- Kısa öykü kayıtları
- Arşiv kayıtları
- Dolaylı ölçümler

Bu işlemler ve tekniklerle, çoğu zaman, aynı ölçütleri ele alan veriler toplanabilmektedir. Aynı şekilde, bir işlem ya da teknik farklı ölçütlere ilişkin veriler de toplanabilmektedir. Veri toplama geniş beceri dağılımına sahip olan değerlendirilmeciler, başarılı olarak kabul edilebilir. Verilerin toplanmasında bazı sorunlarla karşılaşabilmektedir. Özellikle gelişmekte olan ülkelerde, veri toplama araçlarının seçimi sorun olmaktadır. Bu sorunu çözümünde ise, daha çok, birden fazla veri toplama aracı ve ölçüt işe koşulmaktadır. Bir çalışmada farklı veri toplama araçlarının ve ölçütlerinin kullanımını, diğer bir deyişle çoğul ölçümü destekleyen belli başlı dört neden vardır:

1. Araçlar, ölçtükları bir konunun belli boyutlarında farklı olabilir. Çoğul ölçümlerin kullanımı, aynı konunun farklı boyutlardan ele alınmasını sağlar.
2. Araçlar, psikometrik özellikleriyle sınırlıdır. Çoğul ölçümlerin kullanılması, bir aracın geçerlilik ve güvenilirliği üzerindeki sınırlamaları azaltmaya yardımcı olur.
3. Anahtar tarafların, proje konusunda önemli olduğuna inandıkları konular ve değerlendirmenin değerini yargılamada kullandıkları ölçütler farklıdır. Çoğul ölçümler, değerlendirmede hedeflenen ölçümler dizisini artırma yolu verir.
4. Anahtar tarafların inandıkları veri türleri ve farklı ölçümlere verdikleri değerler farklıdır. Çoğul ölçümler, tarafların verilerin değeri konusundaki farklı inançlarına yanıt verir. Çoğul ölçüm kullanımı, farklı çıkar gruplarının, çalışma sonuçlarına daha fazla değer verme olasılığını artırır.

Değerlendirmeci, çoğul ölçüm kullanmaya karar verdiğinde, bu ölçümlerden hangilerini seçeceğini belirlemelidir. Bu seçimde ise, şu dört ölçüt, kılavuzluk eder:

1. Geçerlik: Veri toplama aracının geçerliği, ölçülmesi isteneni ne derecede ölçtüğüdür. Geçerliğin değişik türleri, ölçümün belli amaçlara ulaşmada, ne kadar iyi olduğunu gösterir. Yapı geçerliği, ölçüme dayanak olan kuramların geçerliği ile, başka bir deyişle önceden kabul edilen olası neden-sonuç ilişkilerinin belirlenmesi ile ilgilidir (Karasar, 1984, s.160; Balcı, 1995, s.123). Kapsam geçerliği, ölçümün bir programın amaçlarını ne kadar iyi kapsadığına işaret eder. Ayırt etme geçerliği, değerlendirmede kullanılan ölçme aracının, diğer ölçme araçlarıyla olan ilişkilerin derecesi ile ilgilidir. Yordama geçerliği, programdaki katılımcıların ilerdeki davranışlarının önceden kestirilebilmesi amacıyla hizmet eder. Görünüş geçerliği ise, ölçümde kullanılan aracın gerçekten neyi ölçtüğü değil neyi ölçüyor görüldüğü ile ilgilidir.

2. Güvenirlik: Bir aracın güvenilirliği, zaman uyumlu bir ölçümü, ne derecede verdiğidir. Güvenirlik, çelişkili yönergeler, zayıf şekilde ifade edilen sorular, ölçüm ortamındaki değişim gibi etmenlerle azalabilir. Farklı hata kaynakları, güvenilirlik katsayılarını, farklı yöntemlerle belirlemeyi gerektirir. İç uyum güvenilirliği, bir testteki soruların türdeşliğinin(homojenliğinin) kestirilmesidir. Diğer bir deyişle bu, bir testteki soruların aynı şeyi ne ölçüde ölçebileceğinin, olası yanıtıdır. Test-tekrar test yöntemiyle güvenilirliğin belirlenmesi, zaman içinde bir ölçümün uyumunun ölçülmesidir. Paralel formlarla güvenilirliğin belirlenmesi, aynı ölçümün iki formu arasındaki ilişkiyi ortaya koyar.

3. Uygunluk: Ölçme araçlarının değerlendirilecek programla ne kadar ilgili olduğunu ve ne kadar kabul edilebileceğini gösterir. Bu, hem lojistik, hem de kültürel duyarlılıkla ilgilidir. Söz gelimi, yazılı anketler genellikle okur yazar olmayan yanıtlayıcılara, okul öncesi eğitim çağındaki çocuklara uygun değildir.

4. Uygulama koşulları ve zamanı: Geçerli, güvenilir ve uygun ölçme araçları, uygun zaman ve koşullarda kullanılmaya işaret eder. Daha da ötesi, verilere ilişkin puanlamaya ve puanlamaya dayalı yorumlara işaret eder. Ölçme araçlarının uygulanma zamanı ve koşulları, çalışma gereksinimlerini karşılamayacak şekilde olursa, verilerin niteliği olumsuz yönde etkilenir. Bu da, değerlendirmenin değeri konusunda kuşkuyla neden olur.

Özellikle gelişmekte olan ülkelerde düzenlenen değerlendirme çalışmalarının değerine yönelik eleştiriler olabilmektedir. Bu eleştirilerin en önemlilerinden biri, elde edilen verilerin niteliğinin düşük olduğudur(Windham ve Chapman,1990, s.49): Verilerin niteliğinin düşük olması, belli başlı üç kaynaktan ortaya çıkar:

1. Araçlar uygun olmayabilir. Veri toplama araçlarının uygun olmaması durumunda, yeni araçlar tasarlanıp geliştirilmelidir. Veri toplama araçları, zamanın kısıtlı olması ve kaynakların yetersiz olması nedeniyle, yeterli psikometrik nitelikten yoksun olarak hazırlanmış olabilir. Başka bir deyişle araçlar, iyi bir şekilde geliştirilmemiş olabilir. Veri toplama araçlarını geliştirmek ve standart hale getirebilmek için gereken zaman, değerlendirme çalışmasında göz önünde bulundurulmalıdır. Hatta bunlar, değerlendirmede bütçeleme ve planlama konusu olmalıdır.

2. Raporlar ve belgeler gibi veri kaynaklarının yöresel dile ya da değerlendirmecinin diline çevrilmesi işlemlerine yeterince ilgi olmayabilir.

Değerlendirme ekibince kullanılan çevirmenler, ilgili dilde çok iyi olmayabilir ya da bir dilden diğerine çevirmede teknik terminolojiyi iyi kullanamayabilir. Değerlendirmecinin, program hakkında ileri sürdüğü iddialar, yanlış anlamalardan veya çeviri hatalarından kaynaklanıyor olabilir.

3. Değerlendirmeciler, potansiyel yanıtlayıcıların kültürel duyarlılığını ve/veya ortamın lojistik gerçeklerini anlamada başarısız olabilir ve uygun olmayan veri toplama tekniklerini kullanabilir(Windham ve Chapman,1990, s.46):

Değerlendirmenin, tasarlanma, veri toplama araçları geliştirme ve örneklemi oluşturulma aşamalarında, değerlendirmecilerin, önemli çabaları vardır. Değerlendirmeci, bu çabalarını düşünerek, kültürel önyargıları, araçların duyarlılığını veya yeni verileri toplamayı göz ardı edebilir ya da elde edilen verilerin niteliğine yönelik eleştirileri ilgililere bildirmede gönülsüz olabilir. Ayrıca, değerlendirmeci, çalışmada yapılan zaman ve kaynak yatırımını göz önüne alıp, eleştirilerin "önyargı" ya da "kabullenilebilir" olduğuna karar verebilir ve bunu yaparken ne kadar önyargının kabul edilebileceğini tam olarak anlamayabilir.

Program değerlendirme için verileri toplamada karşılaşılan sorunların büyük bir bölümü, belli başlı üç kaynaktan ortaya çıkar.

Bunlardan birincisi, veri toplama alt yapısının yeterli olmayışıdır. Eğer veri toplama işlemlerinin posta aracılığıyla gerçekleşmesi gerekiyorsa ve ülkede posta sistemi iyi işlemezse bu önemli bir sorundur. Benzer şekilde, verilerin telefonla toplanması gerekiyorsa, evreni ve örneklemi oluşturanların telefonunun olması ve telefon sisteminin iyi işlemesi gereği vardır. Veri toplama araçlarından bazılarının, örneğin anketlerin, insanların çoğunluğunun okur yazar olmadığı yerlerde sınırlı kullanımı vardır. Ayrıca programı uygulayan bazı kurumlarda, katılımcıların devam durumu ve benzeri konularda, kayıt formlarının düzenli kullanılmaması ya da hiç kullanılmaması elde edilecek verilerin niteliğini sınırlamaktadır.

İkincisi, veri toplamada bireysel katılımı destekleyen tutumlar gelişmemiş olabilir. Yanıtlayıcılar, veri toplama araçlarını doldurmayı kabul etmeyebilir veya görüşmelerde katılmak istemeyebilir ya da katılırsa içtenlikle yanıt vermeyebilir. Yanıtlayıcılar, verilerin kendilerine karşı kullanılacağı kaygısıyla gerekli bilgileri vermeyebilir.

Üçüncüsü, kullanılan veri toplama teknikleri görelidir. Yanıtlayıcıların, belirtilmesinin uygun olmadığını düşündüğü sorulara -bir öğretmen ya da program yöneticisine ilişkin görüş ve düşünceleri yansıtan sorulara- ölçme araçlarında ve görüşmelerde yer vermiş olabilir. Ayrıca, görüşmeciler soruları yavaş ve dolaylı olarak yanıtlamak için kültürel tercihleri anlamada başarısız olabilir. Böylece görüşmeciler, tartışılan konular yeterince açıklanmadan görüşmeyi bitirir.

Bazı değerlendirme ortamlarında, verilerin niteliğine ilişkin sorunların kabul edilir bir çözümü yoktur. Seçilen veri toplama stratejileri çok iyi olsa bile, bu stratejiler belli kültürel ortamlar için uygun ve işler olmayabilir. Bu, ele alınacak değerlendirme sorularının yapısını şekillendirir. Veri toplama işlemlerine ilişkin kararlarla karşılaşan değerlendirmeciye üç ilke yön gösterir.

1. Değerlendirmeciler, sonuçları destekleyen verilerin niteliği hakkında açık olmalıdır. Bütün değerlendirme raporları, verilerin niteliğine ait sınırların yeterince sergilendiği bir bölümü içermelidir.

2. Belli ölçüte ilişkin veriler, belli veri toplama teknikleriyle elde edilebilir. Uluslararası ortamlardaki değerlendirme raporları aynı konuları ele almak için geniş bir dizi veri toplama tekniklerini kullanabilmelidir. Çoğul tekniklerin kullanımında uzmanlık, gelişmekte olan ülkelerde çalışan değerlendirme raporları için şarttır.

3. Değerlendirmeciler, çoğul ölçümler yapabilmeyi yanı sıra, bir tekniği diğeriyle değiştirme yeteneğine de sahip olmalıdır.

Verilerin Çözümlemesi

Özellikle az gelişmiş ülkelerde, değerlendirme raporları verilerin çözümlemesi bakımından belli başlı iki temel sorunla karşılaşılır. Bunlardan birincisi, gelişmekte olan ülkelerde elde edilen verileri çözümlemek için gerekli yerel gücünün olmayışı, diğeri ise anahtar tarafların değerlendirme sonuçlarını anlamada, genelde, çok az eğitime ve deneyime sahip olmasıdır.

Verilerin çözümlemesinde, gerekli gücün zayıf olmasının farklı nedenleri vardır. Örneğin, bilgisayar veya çözümler için gerekli olan yazılım yoktur. Ayrıca, bazı ülkelerde gerekli bilgisayara ulaşmak ulusal güvenlik nedeniyle sınırlı olabilmektedir. Bilgisayar kullanımına izin verilmeden önce birçok karmaşık etkinlikler gereklidir. Bunlara ek olarak veri girişi ve istatistiksel çözümler konularında eğitilmiş personel yoktur. Personelin olmayışı, kendi çözümlemesini yapmak isteyen değerlendirme raporları etkilenebilir (Windham ve Chapman, 1990, s.51).

Verilerin çözümlemesi için kısıtlı güç, toplanabilecek ve kullanılacak veri türlerinin azalmasına yol açar. Hatta değerlendirme raporlarını bile, sınırlayabilir. Sonuçta değerlendirme, raporların önemi dikkate alınarak tasarlanacağı halde, verilerin kolayca çözümlemesi için, değerlendirme raporlarının yeteneği ön plana çıkarılarak tasarlanabilir. Ayrıca, veriler temel değerlendirme raporları konusunda toplansa bile, yararlı hale getirilemeyebilir. Bu da, raporların, planlama kararları üzerinde etkisi olmayacağı anlamına gelmektedir. Değerlendirmenin değerlendirme raporlarının ilk aşamalarında verileri nasıl çözümleyeceğini dikkatli şekilde planlaması önemlidir.

Değerlendirmenin anahtar tarafları, genelde, niceliksel sonuçları anlamak bakımından yeterli eğitime ve deneyime sahip değildirler. Çoğu yerel yönetim personeli, yönetim yetenekleriyle üst sorumlu konumuna gelmiştir, fakat, yine de istatistik, ölçme ve değerlendirme, araştırma tasarımı ve araştırma konularında formal eğitimden yoksundur. Rastlantısal örneklem ve genelleme gibi kavramlar, onların bazıları için anlaşılması zor kavramlardır. Bunların çoğu için, çok değişkenli çözümler ve ilişki gibi kavramlar ile bu kavramlar arasındaki fark kolay anlaşılır değildir. Aynı şekilde, çoğu müşteri ve maddi destek sağlayanların da, temel çözümlere ilişkin kavramlar konusunda, yeterli alt yapısı yoktur.

Program değerlendirme raporlarına ilişkin çözümler, daha kolay anlaşılabilir şekilde sunulabilir. Buna, bir bulguyu göstermede, ilişkiler yerine frekans dağılımlarının kullanımı örnek olarak gösterilebilir. Fakat, eğitim ve kalkınmaya ilişkin daha önemli

bazı sorular basit istatistiklere indirgenemez. Böyle durumlarda, aşırı basitleştirme, hiç çalışma yapılmamasının getirdiği zarardan, daha fazla zarar getirebilir.

Sonuçların Yorumu ve Raporlanması

Değerlendirme raporlarının raporlanmasında değerlendirme raporları, hem politik hem de teknik bir dizi konuyla karşılaşır. Eğer değerlendirme raporları dikkatlice düşünerek şekillendirildiği sorulara ait sonuçları anlamlı şekilde sunamazsa, uygun olarak toplanmış olan verilerin değeri kaybolur. Verilerin değerinin kaybolmasını engellemek için değerlendirme raporlarının, çalışmanın bu noktasında şu soruları dikkate alarak rapor hazırlaması gerekir:

- Sonuçlar, anahtar tarafların anlayacağı şekilde nasıl sunulabilir?
- Sonuçlar, anahtar taraflara nasıl ulaştırılabilir?
- Değerlendirmenin önerileri daha kesin olarak nasıl ifade edilebilir?
- Değerlendirmenin, çalışmanın sınırlarını sunmada sorumluluğu nedir.
- Olumsuz sonuçlar nasıl raporlanmalıdır.
- Sonuçlar yanlış yorumlandığında veya sosyal, siyasal avantajlar için kullanıldığında değerlendirme raporları hangi sorumlulukları taşır?

Sonuçların nasıl raporlanacağı, veriler toplanmadan önce belirlenmelidir. Etkili raporlama şartları gerektirir:

1. Çalışma, ilgili değerlendirme raporları etrafında organize edilmelidir.
2. Planlama ve sağlanan destek dikkate alınarak, çalışmanın tamamlanmasıyla ulaşılan sonuçların sunumuna ne zaman gereksinim olduğu belirlenmelidir.
3. Anahtar taraflar açıkça belirlenmeli ve değerlendirme raporları, müşteri ve maddi destek sağlayan, değerlendirme raporunu kimin alacağı konusunda ve bunu diğer taraflarla paylaşmak için sorumluluğu kimin üstleneceği üzerinde anlaşmalıdır.
4. Değerlendirme raporları, değerlendirme raporlarının, temel taraflarla anlaşılabilir anlaşılamayacağını bilmelidir. Böylelikle bulgular, taraflara en uygun şekilde sunulabilir.
5. Sonuçların raporlanması ve önerilerin yapılmasında değerlendirme raporlarının rolü açıkça anlaşılmalıdır.

Sonuç

Program değerlendirme çalışmaları belli bir sistematige uyularak gerçekleştirilmelidir. Program değerlendirmede belli bir sistematige uyulması, çalışmayı hem daha kapsamlı hale getirir hem de değerlendirmecinin alacağı kararlara rehberlik eder.

Belli bir sistematige uyularak gerçekleştirilen program değerlendirme çalışmalarındaki; anlamın kesinleştirilmesi, amacın belirlenmesi, anahtar tarafların belirlenmesi, olanakların ve engellerin belirlenmesi, yanıt aranacak soruların belirlenmesi, tasarının kesinleştirilmesi, verilerin toplanması, verilerin çözümlenmesi ile sonuçların yorumlanması ve ilgililere bildirilmesi işlemleri, çalışmanın sonuçlarını çok yönlü olarak etkileme gücüne sahiptir. Bu işlemlerin ve işlemlere uymada dikkate alınması gerekenlerin, değerlendirmecilerce bilinmesi yapılacak değerlendirme çalışmasının niteliğini artırmada önemli rol oynar.

KAYNAKÇA

- Balcı, A. **Sosyal Bilimlerde Araştırma, Yöntem, Teknik ve İlkeler**. Ankara: 72 TDFO Bilgisayar - Yayıncılık San. Tic. Ltd. Sti., 1995.
- Erden, A. M. **Eğitimde Program Değerlendirme**. Ankara: Personel Eğitim Merkezi Yayın No:21, 1995.
- Gürkan, T. "Eğitim Biliminde Yöntem". **Eğitim Bilimine Giriş** (Ed: F. Varış). Konya: Atlas Kitabevi, 1994.
- Karasar, N. **Bilimsel Araştırma Yöntemi: Kavramlar, İlkeler Teknikler**. Ankara: Bahçelievler PK. 33, 1984.
- Kaya, Z. "Eğitimde Program Değerlendirme Yaklaşımları" (Bildiri Metni). İstanbul: **Marmara Üniversitesi II. Ulusal Eğitim Sempozyumu, 18-20 Eylül, 1996**.
- Saylan, N. **Eğitimde Program Tasarısı Temeller-Premsipler-Kriterler**. Balıkesir: İnce Ofset, 1995.
- Windham, D.M. and D. W. Chapman. **Advances in Educational Productivity**. (Ed: H.J. Walberg). London: Jai Press Inc., 1990.
- YÖK/Dünya Bankası. Endüstriyel Okullar Projesi Dökümanları (Çoğaltma). Ankara: Tarihsiz.