

PROGRAM DEĞERLENDİRME ALANINDA YAYINLANMIŞ DOKTORA TEZLERİNİN META DEĞERLENDİRMESİ (2015-2018)

META-EVALUATION OF PROGRAM EVALUATION DOCTORAL DISSERTATIONS
(2015-2018)

Saadet Aylin YAĞAN¹

Öz

Bu araştırmada, program değerlendirme niteliği taşıyan doktora tezlerinin meta değerlendirmesini yapmak amaçlanmıştır. Meta değerlendirme, değerlendirmenin değerlendirilmesidir. Meta değerlendirmeyle değerlendirmenin güçlü ve zayıf yönleri ortaya konulabilmektedir. Araştırmanın çalışma grubu ölçüt örnekleme yöntemiyle belirlenmiştir. Ölçütler, tezlerin son beş yıl (2015-2019) içinde tamamlanmış, eğitim programları ve öğretim bilim dalında yazılmış ve YÖK veri tabanından erişilebiliyor olmasıdır. Bu şartlara uyan dokuz program değerlendirme tezi üzerinde çalışılmıştır. Tezler, Stufflebeam'in (2012) meta değerlendirme kontrol listesinden dört boyut, 26 standart ve 206 alt standart kullanılarak değerlendirilmiştir. Program değerlendirme tezlerinin genel olarak; değerlendirici güvenilirliği, değerlere saygı, süreç ve ürünler, proje yönetimi, yasal çerçeve, katılımcı hakları, geçerli bilgi ve bilgi yönetimi standartlarını karşıladıkları görülmüştür. Paydaş odaklılık, amaçlar, verilerin niteliği, iletişim, sonuçlar ve etki, uygulama süreçleri, bağlamsal sürdürülebilirlik, kaynak kullanımı, oryantasyon, açıklık ve yalınlık, şeffaflık ve bilgilendirme, çıkar çatışması, sonuç ve kararlarda doğruluk, güvenilirlik, programı betimleme, değerlendirme deseni ve analiz, istatistiksel kanıtlar ile raporlama standartları ise istenen düzeyde karşılanamamıştır.

Anahtar Kelimeler: Meta değerlendirme, Program değerlendirme, Eğitim programları ve öğretim.

Abstract

This study was carried out to meta evaluate doctoral theses which have evaluated a teaching program or an educational system. Meta evaluation is an evaluation of an evaluation. Meta evaluation studies aim to specify weaknesses and strengths of program evaluation studies. Study group was determined by using criterion sampling method. In accordance with the criteria doctoral theses; completed in last five years (2015-2019), that can be reached on YÖK database, and written in the field of curriculum and instruction have been selected. Nine theses have met those criteria. The meta evaluation was conducted by using document review technique considering four dimensions, 26 standards and 206 sub-standards from Stufflebeam's (2012) meta evaluation check list. In conclusion, the standards which theses met highly were; evaluator credibility, values, process and products, project management, formal agreements, participant rights and respect, valid information and information management. Attention to stakeholders, purposes, quality of information, communicating, consequences and influence, practical procedures, contextual viability, resource use, orientation, clarity and fairness, transparency and disclosure, conflicts of interests, justified conclusions and decisions, reliability, program description, evaluation design and analysis, statistical reasoning, and reporting standards have not been met on desirable level.

Keywords: Meta evaluation, Program evaluation, Curriculum and instruction.

1. GİRİŞ

Eğitim-öğretim faaliyetlerinin düzenli şekilde yürütülebilmesi için hazırlanan programların değişen koşullara göre zamanla güncellenmesi veya tamamen değiştirilmesi gerekebilmektedir. Programlardaki yenilenmenin sağlanabilmesi için eksikliklerin saptanması

¹ Dr. Arş. Gör., Gaziosmanpaşa Üniversitesi Eğitim Fakültesi, yagan_a@hotmail.com

ORCID ID: <https://orcid.org/0000-0001-7429-7415>

ve ihtiyaçlara ne ölçüde cevap verilebildiğinin belirlenmesi lazımdır. Bu da program değerlendirme çalışmaları yoluyla yapılabilmektedir. Program değerlendirme, değerlendirme standartlarının belirlenmesi ve programın bu standartlara uygunluğunun tespit edilmesi yoluyla programın değer ve kalitesinin ortaya konması sürecidir (Fitzpatrick, Sanders ve Worthen, 2004: 5). Tanımda sözü edilen standartlar, değerlendirme yürütücüsünden değerlendirmenin amacına, kullanılacak yöntem, teknik ve süreçlerden değerlendirmeden etkilenecek paydaşlara kadar pek çok değişken göz önünde bulundurularak oluşturulmaktadır (Yüksel, 2010: 3). Bu standartların en önemlilerinden biri de değerlendirmenin, değerlendiricinin ya da değerlendirme sisteminin değerlendirilmesidir. Meta değerlendirme kavramında karşılık bulan bu değerlendirme türü ilk olarak Scriven (1969) tarafından dile getirilmiş ve tanımlanmıştır (Scriven, 2009: 3; Stufflebeam, 2001: 185).

Meta değerlendirme, değerlendirmede kullanılan yaklaşım ve uygulamaların kalitesi hakkında güvence sağlayan bir sistemdir. Operasyonel tanımıyla meta değerlendirme, değerlendirmenin ve değerlendirme raporunun yararlık, uygulanabilirlik, yerindelik ve doğruluğu hakkında betimsel ve eleştirel veri edinme ve işleme sürecidir (Stufflebeam, 2001: 185). İyi bir değerlendirme yapabilmek için meta değerlendirmeye ihtiyaç vardır. Değerlendirme esnasında pek çok aksaklık yaşanabilir. Bu nedenle çalışmaların yanlışlık, teknik hatalar, yönetsel zorluklar ve yanlış kullanım gibi konularda kontrol edilmesi gerekmektedir (Stufflebeam, 2011: 99). Meta değerlendirme çalışmaları hem devam eden değerlendirmelere katkı sunmak hem de tamamlanmış değerlendirmelerin kalitesini, değerini ortaya koymak açısından önem taşımaktadır (Stufflebeam, 2011: 99).

Meta değerlendirme, tüm değerlendirme çalışmalarında olduğu gibi belirli bir değerlendirme sistemine ihtiyaç duyar. Bu sistemin en önemli yardımcısı meta değerlendirme standartlarıdır. Literatürde, farklı araştırmacı ya da kuruluşların meta değerlendirme için hazırladığı standartlardan oluşan kontrol listeleri mevcuttur (bkz. Joint Committee on Standards for Educational Evaluation (JCSEE) Program Evaluation Standards, 2011; Scriven Key Evaluation Checklist, 2013; Stufflebeam CIPP Evaluation Model Checklist, 2007; Stufflebeam Program Evaluations Metaevaluation Checklist, 2012). Kapsamlı şekilde hazırlanmış bu kontrol listeleri, meta değerlendirme araştırmaları için son derece faydalıdır. Sistematik bir meta değerlendirme yapmak isteyen araştırmacılar, program değerlendirme raporlarında neler yer almalı ve nasıl yer almalı sorularına kontrol listelerindeki maddeleri kullanarak cevap bulabileceklerdir (Scriven, 2009: 6).

Scriven (2005: 3), değerlendirme çalışması yaparken kontrol listesi kullanmanın avantajlarını şöyle açıklamıştır:

- Kontrol edilmesi gereken önemli bir özelliğin unutulmasını engeller.
- Paydaşların değerlendirme sürecini kolay anlayabilmelerini sağlar.
- Pek çok istatistiksel analiz ve teori uygulamasına göre çok daha geçerli bir yöntemdir.
- Genellikle değerlendiricinin bir özelliğe olduğundan fazla değer biçmesini engeller çünkü değerlendiriciyi her bir boyutu ayrı ayrı değerlendirmeye zorlar.
- Her bir boyutu ayrı ayrı değerlendirmeye ve bu değerlendirmeler üzerinden yargıda bulunmaya yönlendirdiği için değerlendirici görmek istediklerine yoğunlaşmaktan ziyade objektif bir duruş sergiler.
- Geçerli bir kontrol listesi kullanılması gelişigüzel bir listeye kıyasla daha doğru sonuçlar verir.

Meta değerlendirme çalışmalarına uluslararası literatürde özellikle tezlerde (Blake, 2013; Gibbs, 2012; Shanahan, 2014; Torres, 2005) rastlanırken Türkiye’de araştırma sayısı henüz çok yetersizdir. Literatürde geçen kriterlerin kullanıldığı tek bir meta değerlendirme çalışmasına (Yüksel ve Akın, 2013) rastlanmıştır. Eğitim programları ve öğretim alanında

program değerlendirme çalışmaları oldukça yaygın olduğu (Kozikoğlu ve Senemoğlu, 2016: 29) halde meta değerlendirmenin yapılmaması, program değerlendirme çalışmalarının uygunluğunu ve değerini ortaya koyma noktasında eksiklikler olduğunu göstermektedir. Bu araştırmada, eğitim programları ve öğretim alanında son beş yılda (2015-2019) program değerlendirme çalışması yapan doktora tezlerinin meta değerlendirmesinin yapılması amaçlanmaktadır. Böylece değerlendirme çalışmaları sistematiklik, paydaşlara yeterli bilgi sunma düzeyi, bütünlük ve doğruluk, paydaşların ve katılımcıların haklarına riayet etme düzeyi ve tüm topluma sağladığı fayda (American Evaluation Association, 2013) konularında ikinci bir değerlendirmeye daha tabi tutulmuş olacaktır.

Araştırma, Joint Committee on Standards for Educational Evaluation (JCSEE) (2011), tarafından belirlenen ve Stufflebeam'in (2012) detaylandığı standartlar üzerinden yürütülmüştür. Bu standartların yer aldığı beş temel boyut; yararlık, uygulanabilirlik, yerindelik, doğruluk ve hesap verilebilirliktir. Yararlık boyutu, değerlendirme süreç ve sonucunun kullanıcı ihtiyaçlarını ne ölçüde karşıladığı ile ilgilidir. Uygulanabilirlik standartları, değerlendirmenin etkililiğini ve verimini arttırmayı amaçlar. İzlenen prosedür pratik ve anlamlı olmalı, kaynaklar etkili kullanılmalı, politik ve kültürel ilgiler göz önüne alınmalıdır. Yerindelik standartları, değerlendirmenin etik, yasal, doğru, adil ve makul biçimde yerine getirilmesine odaklanır. Değerlendirmede yer alan ve değerlendirmeden etkileyecek kişilerin her türlü hakkı gözetilmeli, şeffaf bir süreç izlenmelidir. Doğruluk standartları, değerlendirmenin geçerlik ve güvenilirliği ile ilgilidir. Değerlendirmede güvenilir, gerçek sonuçlar ile geçerli yorumlara, altı doldurulmuş yargılar ve değerlendirmelere gerek duyulur. Hesap verilebilirlik standartları ise dokümantasyon ve meta değerlendirme perspektifine yoğunlaşarak değerlendirme süreç ve çıktılarının hesap verilebilirliğini arttırmayı hedefler.

2. ARAŞTIRMA MODELİ

Bu bir meta değerlendirme araştırmasıdır. Meta değerlendirme, değerlendirme çalışmalarının ve dolaylı olarak değerlendiricilerin etik ve bilimsel açıdan belirli standartlara göre sistematik olarak incelenmesidir (Scriven, 1991: 228). Meta değerlendirmede amaç değerlendirme sürecini ya da aynı konu üzerine yapılacak gelecek değerlendirmeleri iyileştirmek, değerlendirmenin güçlü ve zayıf yönlerini ortaya koymak ve değerlendirmeden yararlanacak kimselere değerlendirmenin niteliğiyle ilgili bilgi sağlamaktır (Cooksy ve Caracelli, 2009: 2).

2.1. Çalışma Grubu

Araştırmanın çalışma grubunu eğitim programları ve öğretim bilim dalında herhangi bir öğretim programı ya da eğitim sistemini değerlendiren doktora tezleri oluşturmuştur. Ölçüt örnekleme yönteminin işe koşulmuş ve bu çerçevede tezlerin son beş yıl (2015-2019) içinde tamamlanmış olması ile YÖK veri tabanından erişilebiliyor olması ölçütleri kullanılmıştır. 2019 yılı içerisinde ölçütleri taşıyan teze rastlanmadığı için zaman aralığı 2015-2018 olarak güncellenmiştir. Bu şartlara uyan dokuz teze ulaşılmıştır (2015 n=2, 2016 n=5, 2017 n=1, 2018 n=1).

2.2. Verilerin Toplanması ve Analizi

Tezler, Stufflebeam'in (2012) program değerlendirme çalışmaları için oluşturduğu meta değerlendirme kontrol listesi kullanılarak doküman inceleme tekniği ile değerlendirilmiştir. Kontrol listesi beş ana boyut ve 30 standarttan oluşmaktadır. Her bir standardın altında alt standartlar yer almaktadır. Toplam 297 alt standart vardır. Bu araştırmada Stufflebeam'in kontrol listesindeki dört boyut, 26 standart ve 206 alt standart kullanılarak meta değerlendirme

gerçekleştirilmiştir. Hesap verilebilirlik boyutu ile bu boyutun altındaki üç standart (dokümantasyon, iç meta değerlendirme, dış meta değerlendirme), doğruluk boyutunda yer alan mali yükümlülük standardı ve diğer boyutlardaki bazı alt standartlar tezler için uygulanabilir olmadığından kapsam dışı bırakılmıştır. Kimi alt standartlar ise birden fazla standartla ilişkili olduğundan tekrar etmiştir. Standartların puanlanması; 0-2: Zayıf, 3-4: Vasat, 5-6: İyi, 7-8: Çok iyi, 9-10: Mükemmel şeklindedir.

3. BULGULAR

Bu bölümde 26 değerlendirme standardı değerlendirme tezleri bağlamında ayrı ayrı incelenmiştir. Her bir standart incelenirken öncelikle bu standardın gerektirdiği alt standartlara yer verilmiş sonrasında ise tezlerin bu alt standartları ne ölçüde karşıladığı ortaya konmuştur. Tablo 1’de standartlar, alt standartlar ve tezlerin standartlardan aldıkları puanlar yer almaktadır.

Tablo 1. Tezlerin Meta Değerlendirme Standartlarına Göre Puanlaması

| Puan | 0-2: Zayıf | 3-4: Vasat | 5-6: İyi | 7-8: Çok İyi (Ç. İyi) | 9-10: Mükemmel (Mük.) | | | | | |
|------------------|--------------------------------|------------|----------|-----------------------|-----------------------|--------|--------|--------|--------|--------|
| | Standartlar | Tez 1 | Tez 2 | Tez 3 | Tez 4 | Tez 5 | Tez 6 | Tez 7 | Tez 8 | Tez 9 |
| Yararlık | Değerlendirici güvenilirliği | İyi | İyi | Mük. | Mük. | Mük. | Ç. İyi | Ç. İyi | Ç. İyi | Mük. |
| | Paydaş odaklılık | Vasat | İyi | Mük. | Mük. | İyi | İyi | Ç. İyi | Ç. İyi | Ç. İyi |
| | Amaçlar | Vasat | Zayıf | Ç. İyi | İyi | Vasat | İyi | İyi | İyi | Ç. İyi |
| | Değerlere saygı | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | İyi | Vasat | Ç. İyi |
| | Verilerin niteliği | İyi | İyi | Ç. İyi | Mük. | Vasat | İyi | Ç. İyi | Ç. İyi | Ç. İyi |
| | Süreç ve ürünler | Ç. İyi | Ç. İyi | Ç. İyi | Mük. | İyi | Ç. İyi | Ç. İyi | Mük. | Mük. |
| | İletişim | Ç. İyi | İyi | Ç. İyi | Ç. İyi | İyi | Vasat | İyi | Ç. İyi | İyi |
| Sonuçlar ve etki | Vasat | Zayıf | İyi | İyi | Vasat | İyi | İyi | Ç. İyi | Ç. İyi | |
| Uygulanabilirlik | Proje yönetimi | Ç. İyi | Ç. İyi | Ç. İyi | Mük. | Ç. İyi | Mük. | Ç. İyi | Ç. İyi | Mük. |
| | Uygulama süreçleri | Ç. İyi | Mük. | Ç. İyi | Ç. İyi | Vasat | Vasat | Mük. | Mük. | Ç. İyi |
| | Bağlamsal sürdürülebilirlik | İyi | Ç. İyi | İyi | Ç. İyi | İyi | Mük. | Ç. İyi | İyi | İyi |
| | Kaynak kullanımı | Vasat | İyi | İyi | Ç. İyi | Vasat | Vasat | İyi | İyi | İyi |
| Yerindelik | Oryantasyon | Vasat | Vasat | Ç. İyi | Mük. | Vasat | İyi | İyi | Ç. İyi | Ç. İyi |
| | Yasal çerçeve | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | İyi | İyi | Ç. İyi | Ç. İyi |
| | Katılımcı hakları | Ç. İyi | Ç. İyi | İyi | Ç. İyi | İyi | Ç. İyi | Mük. | Mük. | Mük. |
| | Açıklık ve yalınlık | Vasat | Vasat | İyi | Ç. İyi | Vasat | Ç. İyi | Vasat | Vasat | İyi |
| | Şeffaflık ve bilgilendirme | İyi | İyi | Ç. İyi | Mük. | İyi | Ç. İyi | Vasat | Vasat | İyi |
| | Çıkar çatışması | Zayıf | Zayıf | Zayıf | Zayıf | Zayıf | Zayıf | Zayıf | Zayıf | Zayıf |
| Doğruluk | Sonuç ve kararlarda doğruluk | Vasat | Vasat | Vasat | Mük. | Ç. İyi | Vasat | Vasat | Zayıf | İyi |
| | Geçerli bilgi | Vasat | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi | Ç. İyi |
| | Güvenirlik | Vasat | İyi | Vasat | Mük. | Ç. İyi | İyi | İyi | Vasat | Vasat |
| | Programı betimleme | Ç. İyi | İyi | Ç. İyi | Ç. İyi | Vasat | Vasat | Zayıf | Vasat | İyi |
| | Bilgi yönetimi | İyi | Ç. İyi | Ç. İyi | Ç. İyi | İyi | İyi | Ç. İyi | Ç. İyi | Ç. İyi |
| | Değerlendirme deseni ve analiz | Vasat | Vasat | İyi | Ç. İyi | İyi | İyi | Vasat | Vasat | İyi |
| | İstatistiksel kanıtlar | Vasat | Vasat | İyi | İyi | Ç. İyi | İyi | Vasat | Zayıf | Vasat |
| Raporlama | İyi | İyi | İyi | İyi | İyi | Vasat | İyi | Vasat | Vasat | |

3.1. Yararlık Standartları

3.1.1. Değerlendirici güvenirligi: Bu standart, değerlendiricinin konuyla ilgili alan bilgisinin ve tecrübesinin olması, tüm paydaş ve katılımcılara güven vermesi ile ilgilidir. Değerlendirici; araştırma planı, değerlendirmenin teknik kalitesi ve uygulanabilirliği hakkında katılımcıları bilgilendirmiş olmalıdır. Ayrıca katılımcılardan gelen öneri ve eleştirilere yer verilmelidir. Katılımcı önceliklerine ve endişelerine hitap edilmeli, çalışılan grubun sosyo-kültürel ve ekonomik farklılıklarına vakıf olunmalıdır. İncelenen tezler, değerlendirici güvenirligi standardını genellikle karşılamaktadır. Tüm değerlendiriciler EPÖ alanında doktora eğitimi aldıklarından program değerlendirme konusunda temel bilgilere sahiptirler. Tamamı mezun oldukları ya da çalıştıkları alanlarla ilgili araştırmalar yapmışlardır. Dolayısıyla güven veren ve katılımcıların hassasiyetlerini anlayabilecek yapıdadırlar. Öte yandan, değerlendirme planlarının hazırlanmasında paydaşların beklenti veya endişelerinin hesaba katılmadığı görülmektedir. Katılımcı ve paydaşların değerlendirme planından haberdar olması konusunda herhangi bir gayret yoktur.

3.1.2. Paydaş odaklılık: Değerlendirme çalışmalarında öncelikle potansiyel paydaşlarla görüşmeler yapılmalı ve onların ne tür bilgilere ihtiyaç duydukları saptanmalıdır. Hangi paydaşların öncelikli muhatap olduğu uzman görüşü alınarak belirlenmeli ya da değerlendirici, hangi paydaş grubunun neden öncelikli olduğunu açıklamalıdır. Katılımcıların neye göre ve nasıl seçildiği belirtilmeli, paydaş ya da katılımcılar değerlendirme sürecine dâhil edilmelidir. Yeterli sayıda ve farklı gruplardan katılımcıya ulaşılması beklenmektedir. Değerlendirme sonuçlarından faydalanabilecek kurum ya da kişiler açıkça tanımlanmalı ve bu kişilerin ihtiyaçlarını karşılamaya dönük çıkarımlarda bulunulmalıdır. Bu standartlar, tezlerde tam anlamıyla karşılanamamaktadır. Özellikle katılımcılar dışında paydaşların tanımlanması ve paydaşların ihtiyaçlarının belirlenerek tüm değerlendirme sürecine dâhil edilmeleri hususunda sıkıntılar vardır. Bazı tezlerde değerlendirme öncesinde ihtiyaç analizi yapılmıştır. Bu tezler paydaş odaklılık standardını daha yüksek düzeyde karşılamaktadırlar. Katılımcıların nasıl seçildiği, yeterli sayıda katılımcıya ulaşılmış olması ve hangi katılımcıların öncelikli olduğu konusunda uzman görüşüne başvurulması gibi alt standartlar, genellikle yüksek düzeyde karşılanmıştır. Geriye kalan alt standartlar ise kısmen karşılanmıştır.

3.1.3. Amaçlar: Değerlendirmenin amaçları açık bir şekilde belirlenmeli ve bu süreçte paydaşların ihtiyaçları göz önünde bulundurulmalıdır. Amaçların hangi bağlam ve zaman aralığı içerisinde geçerli olduğu ve zamanla ne gibi değişimlere uğrayabileceği tartışılmalıdır. Değerlendirme süreci amaçlandığı şekliyle tezde açıklanmalı, süreç boyunca bir değişiklik yaşanıp yaşanmadığı aktarılmalıdır. Değerlendirmenin başında belirlenen amaçların ne kadarına ulaşılabildiği araştırmacı tarafından tartışılmalı, ulaşılamayan amaçlar ve nedenleri açıklanmalıdır. Araştırmanın amaçları belirlenirken uzman ya da paydaş görüşüne başvurulması bir artıdır. Araştırmanın amaçları göz önünde bulundurularak katılımcıların uzlaştıkları ya da ters düştükleri noktalar belirtilmeli ve bulgular yorumlanırken farklı grupların değerlendirilmesinin amaçlarından farklı şekillerde yararlanacakları göz önüne alınmalıdır. İncelenen tezlerde genellikle değerlendirilmesinin amaçları açık şekilde belirlenmiş ve değerlendirme süreci de açıklanmıştır. Katılımcıların uzlaştıkları ya da ters düştükleri noktalar bazı tezlerde ele alınmışsa da tezlerin tamamında bu alt standart yerine getirilmemiştir. Amaçlar belirlenirken tez danışmanı ile birlikte çalışıldığı ve tez izleme jürisinden fikir alındığı için uzman görüşü alma alt standardının karşılandığı düşünülmektedir. Diğer alt standartlar ise çok az tezde kısmen karşılanmıştır. Amaçların hangi bağlam ve zaman aralığında geçerli olacağı tartışılmamış ve aynı zamanda ulaşılamayan amaçlara değinilmemiştir.

3.1.4. Değerlere saygı: Bu standart, değerlendirilen programla ilgili farklı bakış açıları tartışılmasını gerekli kılmaktadır. Değerlendirmenin sosyal ve politik bağı kurulmalıdır. Değerlendirilen konuya ilişkin yasal mevzuattan, paydaşların bilmesi gereken resmi bilgilerden bahsedilmeli ve farklı kurumların bakış açıları, politikaları ele alınmalıdır. Katılımcıların görüş bildirirken ne gibi değer yargılarını benimsedikleri tartışılmalıdır. Bulgular yorumlanırken farklı değerler veya bakış açıları göz önünde tutularak çatışan ve örtüşen değerler ile bunlara dair yorumlara yer verilmelidir. Değerlendirme tezlerinde değerlere saygı standardı önemli ölçüde sağlanmaktadır. Tezlerin kavramsal çerçeve bölümlerinde değerlendirilen programlar ayrıntılı bir biçimde incelenirken farklı bakış açıları, konuyla ilgili sosyal ve politik bağlam, ilgili yasal mevzuatlar ve eğitim politikaları gerektiği şekilde açıklanmaktadır. Bununla birlikte, katılımcı görüşlerinin o görüşleri besleyen değer yargıları bağlamında analiz edildiğini söylemek zordur.

3.1.5. Verilerin niteliği: Yapılan değerlendirmenin neden gerekli olduğu ayrıntılı biçimde açıklanmalı ve paydaş gereksinimleri net biçimde ortaya konmalıdır. Paydaşlar, katılımcılar, değerlendirme soruları ve programı değerlendirebilmek için ihtiyaç duyulan bilgi açıkça tanımlanmalıdır. Farklı bakış açılarını yakalamak adına paydaşlarla görüşmeler yapılması avantajlı olacaktır. En önemli paydaşlara ve en önemli sorulara öncelik tanınmalı, paydaşların birincil ihtiyaçlarına yönelik veri edinilmelidir. Değerlendirme sırasında sorular ihtiyaca göre güncellenmeli, programın değerini ve kalitesini ölçebilecek nitelikte yeterli veri toplanmalıdır. Ayrıca, farklı gruplardaki farklı düşüncelerde paydaşlar çalışmaya dâhil edilmelidir. Tezlerde genellikle yalnızca katılımcılar tanımlanmakta, programın diğer paydaşlarından söz edilmemektedir. Değerlendirme öncesinde ya da sırasında paydaşlarla görüşmeler yapılan tezler olsa da bunların sayısı çok azdır. Genellikle en önemli paydaşlara katılımcı olarak yer verilmekle birlikte bazı tezlerde yalnızca tek bir paydaş kitlesiyle çalışılmıştır. Değerlendirme sorularının başlangıçtaki haliyle mi kaldığı yoksa süreç esnasında değişiklikler mi olduğu belirtilmemiştir. Değerlendirme ihtiyaçları ve değerlendirmenin önemi genellikle başarılı şekilde açıklanmıştır ancak çoğunda paydaş ihtiyaçları belirlenmediğinden paydaş gereksinimlerinden çok araştırmacıların odaklandıkları sorular üzerinde durulmuştur. Tezlerin çoğunluğunda ölçek geliştirmeye daha çok çaba harcandığından programın niteliği ve değerini belirleme istenen düzeyde gerçekleşmemiştir.

3.1.6. Süreç ve ürünler: Tüm tez boyunca kısa, öz, net ve doğrudan bir anlatım kullanılmalı, program ve bağlamı tanımlanmalıdır. Değerlendiricinin amacı, bulguları ve izlediği süreç net olarak verilmelidir. Temel sonuçlara açıkça yer verilmeli ve sonuçlar, araştırma sorularına uygun olarak hazırlanmalıdır. Ayrıca, sonuçlar ve öneriler doyurucu olmalıdır. Sonuçlar ve öneriler, bulgular ve literatürle desteklenerek yazılmalıdır. Teknik terminolojiden kaçınmak ve tüm paydaşların rahatlıkla takip edebileceği bir çalışma hazırlamak gerekmektedir. Araştırma hakkında genel bilgi edinmek isteyen araştırmacılar için özet kısmında gerekli bilgilere yer verilmesi de beklenir. İncelenen tezler bu standart açısından oldukça başarılıdır. Sonuç ve önerilerin doyurucu olması, bulgularla desteklenmesi ve izlenen prosedürün açıkça ortaya konması alt kriterleri zayıf bulunmuştur. Diğer alt kriterler yüksek düzeyde karşılanmaktadır.

3.1.7. İletişim: Değerlendirme öncesinde ya da sırasında programın yöneticileriyle, yürütücüleriyle ve programdan etkilenen paydaş kitlesiyle bilgi/görüş alışverişinde bulunulmalıdır. Program değerlendirmesi tüm paydaş ve araştırmacılara açık olmalı, paydaşlar diledikleri an değerlendirme çalışmasına ulaşabilmelidirler. Farklı katılımcılara erişebilmek için sosyal medya, internet siteleri, katılımcılardan diğer katılımcılara ulaşma gibi değişik iletişim yolları kullanılmalıdır. Tezlerde, program yöneticileriyle görüşme alt standardı

karşılanmamaktadır. Tezlerin çoğu program yürütücüleri ve programdan etkilenenleri katılımcı olarak seçtiklerinden bilgi/görüş alışverişi genellikle sağlanmıştır. Ancak, sadece yürütücü ya da sadece programdan etkilenenlerle çalışan tezler de vardır. Tüm tezler YÖK veri tabanında yer almaktadır ve isteyen herkese açıktır. Katılımcılar belirlenirken genellikle kolayca erişilebilecek gruplar seçilmektedir. Katılımcıların seçiminde farklı yollar tercih eden tezler olsa da sayısı çok değildir.

3.1.8. Sonuçlar ve etki: Değerlendirmenin farklı aşamalarında katılımcılarla görüşülmelidir. Somut sonuçlara ve uygulanabilir önerilere yer verilmelidir. Programdaki aksaklıklar ortaya konup çözüm önerileri sunulmalıdır. Sonuçlardan nerelerde faydalanılabileceği açıklanmalı, sonuçların nasıl kullanılabileceğine yönelik örnekler sunulmalıdır. Tezlerde veri toplama dışındaki aşamalarda katılımcılarla birlikte çalışmaya pek rastlanmamaktadır. Sonuçların nasıl kullanılabileceğine yönelik örneklerin de sunulmadığı anlaşılmaktadır. Programdaki aksaklıkların tespit edilip çözüm önerilerinin sunulması ise kısmen karşılanmaktadır.

3.2. Uygulanabilirlik Standartları

3.2.1. Proje yönetimi: Katılımcılarla değerlendirme boyunca profesyonel seviyede ilişki kurularak etkili bir iletişim sürdürülmelidir. Enstitü protokollerine uyulmalı ve süreç, katılımcılara söz verildiği şekilde yürütülmelidir. Katılımcıların günlük işleri çok fazla aksatılmamalı, mahremiyetlerine saygı duyulmalı ve daha evvel vaat edilen süre sınırları aşılmamalıdır. Katılımcıların değerlendirmeyle ilgili çekincelerine cevap verilmeli, kültürel farklılık ya da değerlere özen gösterilmelidir. Veri toplanan tüm katılımcıların görüşleri sunulmalıdır. Herhangi bir katılımcının etik olmayan davranışları söz konusuysa gereken müdahale yapılmalıdır. Proje yönetimi tezlerde üst düzeyde karşılanan bir standarttır. Genellikle olumsuz bir durum belirtilmemiş ve her şey planlandığı şekliyle uygulanmıştır. Mahremiyetle ilgili birtakım önlemler alınmış kişi ya da kurum adları iki tez hariç paylaşılmamıştır. Ancak katılımcıların değerlendirmelerle ilgili çekincelerinden bahsedilmemiştir.

3.2.2. Uygulama süreçleri: Ölçme araçları ve kullanılan yöntemler, ihtiyaç duyulan bilgileri toplamaya elverişli olmalıdır. Veri yükü en aza indirilmeli, ham veri önemsiz verilerden ayrıştırılmalıdır. Gerçekçi bir çalışma takvimi belirlenerek bilinen kısıtlılıklara uygun bir süreç izlenmelidir. Değerlendirmenin yürütüleceği ortamlar belirlenmelidir. İmkân varsa değerlendirme, süreklilik gösteren döngüsel bir sistem haline getirilmelidir. Tezlerin çoğu, uygulama süreçleri standardını yüksek oranda karşılamaktadır. Bazı tezlerde önemsiz bilgilerin ayıklanması, bilgi tekrarının olmaması ve gerçekçi bir takvim belirlenmesi gibi alt standartlar istenen düzeyde değildir. Değerlendirmenin sürekli bir sistem haline getirilmiş olması alt standardını ise ancak iki tez belirli ölçüde karşılamaktadır.

3.2.3. Bağlamsal sürdürülebilirlik: Farklı ilgi gruplarının farklı pozisyonlarda olacağı bilinmeli ve değerlendirme buna göre yürütülmelidir. Yanlılık ihtimaline ya da bulguların yanlış kullanımına karşı önlemler alınmalıdır. Farklı görüşlere de yer verilmelidir. Tezi okuyan herkesin bilmesi gereken bilgi ve detaylar okuyucuya sunulmalıdır. Tez, yayımlama kurallarına uygun yazılmalıdır. İncelenen tezlerde farklı ilgi gruplarından farklı veriler elde edileceği düşünüldüğünden genellikle birden fazla katılımcı grubuyla çalışılmıştır ancak tek bir katılımcı grubundan veri toplanan tezler de vardır. Bağlamsal sürdürülebilirlik standardını yüksek düzeyde karşılayan tezler, yanlılık ve bulguların yanlış kullanımıyla ilgili olarak bulguların tutarlılığı ve bütünlüğünün gözden geçirilmesi, kodlayıcılar arası uyum geçerliliğinin hesaplanması ve araştırmacı rolünün açıklanması gibi yöntemler kullanmışlardır. Farklı

görüşlere yer verme, bilinmesi gereken tüm detayları sunma ve yayım kurallarına uygunluk gösterme tezlerin tamamının başarılı olduğu alt standartlardır.

3.2.4. Kaynak kullanımı: Kaynaklar etkili kullanılmalıdır. Hizmet odaklı ya da mal odaklı harcamalar varsa belirtilmelidir. Harcamalara degecek nitelikte bilgi edinilmelidir. Hesap verilebilir, denetime açık bir süreçle değerlendirme gerçekleştirilmelidir. Programın gelişimine katkıda bulunulmalı, yeni bakış açıları oluşturulmalı ve etkili uygulamalar öne sürülmelidir. Son olarak, mümkün olan en kısa zamanda değerlendirme tamamlanmalıdır. Kaynak kullanımı, tezlerin istenen düzeyde karşılayamadığı bir standarttır. Değerlendirme sürecinde herhangi bir hizmet alımı var mı, harcamalar yüksek mi ve nasıl karşılanmış açıklanmamıştır. Yalnızca bir tez yazarı, TÜBİTAK bursu alarak harcamalarını finanse etmiştir. Harcamalara ya da emeğe degecek bilgi üretimi tam olarak yapılmamıştır. En büyük sıkıntı sonuç ve öneriler bölümlerinde yaşanmaktadır. Değerlendirilen programların gelişimine belirli ölçülerde katkı sağlanırken yeni bakış açıları oluşturma ve etkili uygulamalar sunmada eksikler göze çarpmaktadır. Tezler, denetlenebilir aşamalardan geçmektedir. Tez izleme toplantıları bu alt standardı sağlamada önemli süreçlerdir. Tüm tezler mümkün olan en iyi zaman aralıklarında tamamlanmıştır.

3.3. Yerindelik Standartları

3.3.1. Oryantasyon: Değerlendirilen programın çıktıları ile ihtiyaçlar karşılaştırılmalıdır. Değerlendirmenin neden yapıldığı iyi anlatılmalı ve değerlendirme doğru tanıtılmalıdır. Literatürden faydalanarak programın güçlü ve zayıf yönleri ortaya konmalı olumlu-olumsuz çıktılar belirlenmelidir. Tezlerin çoğunda ihtiyaç belirleme alt standardı zayıftır. Çalışma öncesi paydaş gruplarıyla bir araya gelme, nasıl bir değerlendirme yapılacağını tartışma, pilot görüşmeler yapma ya da ihtiyaç analizi gibi alt standartlar çoğu tezde karşılanamamıştır. Sistemli bir ihtiyaç-çıktı karşılaştırmasından da söz edilememektedir. Katılımcılara değerlendirmenin hedeflerine ilişkin bilgiler ya hiç verilmemiş ya da veri toplama aşamasında yüzeysel olarak aktarılmıştır. Literatürden faydalanarak programın güçlü ve zayıf yönleri ile olumlu ve olumsuz çıktıları ortaya koyma alt standartları kısmen karşılanmaktadır. Programın tüm boyutları istenen manada ortaya konmasa da genel tanıtım yapılmaktadır.

3.3.2. Yasal çerçeve: Değerlendirme çalışmalarında, değerlendirmenin amacı ve soruları, değerlendirmeden etkilenecek kişi ve gruplar, değerlendirme raporu, düzeltmeler, raporun yayınlanması, değerlendirme süreci ve takvimi, mahremiyet / katılımcıların anonimliği, meta değerlendirme ve değerlendirme kaynakları hususlarında değerlendiricinin paydaş ya da uzman görüşüne dayalı resmi bir uzlaşma sağlanması beklenir. Doktora tezleri yazılırken gerek danışman görüşü gerekse tez izleme sınavları ve tez savunmasındaki öğretim üyelerinin katkıları bu standardın yüksek düzeyde karşılanmasını sağlamaktadır. Amaç, sorular, değerlendirme süreci, etikle ilgili yapılanlar tez izleme komitesinin uzlaşısıyla şekillenmektedir. Tez izlemelerin sonunda düzeltme ve öneriler de resmi olarak sunulmaktadır. Ayrıca tezin nasıl yayınlanacağını kuralları bellidir. Tezlerin çoğunda katılımcılar dışında değerlendirmeden etkilenecek kişi ve grupların kim olduğu ve kaynak kullanımı açıklanmamıştır. Çalışma takvimine yer vermeyenler de vardır. Son olarak tez savunmasının kısmi bir meta değerlendirme olduğu düşünülebilir.

3.3.3. Katılımcı hakları: Katılımcılara açık bir biçimde haklarının korunacağı açıklanmalı, değerlendirme sonuçlarının hangi amaçla ve nerelerde kullanılacağı net olarak ortaya konmalıdır. Araştırma için gerekli resmi izinler alınmalı, hukuk kurallarını çiğneyecek bir duruma meydan verilmemelidir. Katılımcıların değer yargıları anlaşılabilir farklı fikirlere saygı duyulmalıdır. Gizliliğe özen gösterilmelidir. Değerlendirme sonucunda kimsenin zarar

görmediğinden emin olunmalıdır. Tezlerde yüksek düzeyde karşılanan standartlardan biri katılımcı haklarıdır. Az sayıda tezde gizlilik ilkesi kısmen ihlal edilerek çalışılan okul isimleri verilmiştir. Bununla ilgili izin alınıp alınmadığı belirtilmemektedir. Yine az sayıda tezde tek bir grupta çalışıldığı için farklı grupların fikirlerine yer verilmediği saptanmıştır. Katılımcıların değer yargılarının anlaşılması kısmen gerçekleştirilmiştir. Bunlar dışındaki alt standartlar başarılı şekilde karşılanmakla birlikte ayrı bir başlık altında etik ile ilgili tüm protokolü açıklayan tez sayısı çok azdır. Katılımcılara değerlendirmenin amaçlarının ve kullanım yerinin açıklandığı ve gizlilikle ilgili taahhütlerin verildiği, ekler bölümünde yer alan ölçme araçlarının başındaki açıklayıcı paragraftan anlaşılmaktadır.

3.3.4. Açıklık ve yalınlık: Programın güçlü ve zayıf yönleri ölçülerek rapor edilmelidir. Amaçlanan ve amaçlanmayan çıktılar belirtilmelidir. Değerlendirme süreci net biçimde ortaya konmalıdır. Programın üstün yönlerinin zayıf yönlerini geliştirmede nasıl kullanılabileceği açıklanmalıdır. Tez, bitirilmeden önce uzmanlarca ve paydaşlarca incelenip eleştirilere göre düzenlemeler yapılmalıdır. Değerlendirmenin sınırlılıkları ve hangi bağlamda kullanılabileceği açıklanmalı, sınırlılıkların değerlendirmeyi nasıl etkileyebileceğinin üzerinde durulmalıdır. İncelenen tezlerde daha çok programın zayıf yönleri üzerinde durulmakta, bazı tezlerde güçlü yönlerinden bahsedilmemektedir. Programın güçlü taraflarının zayıf yönleri kapatmada nasıl kullanılabileceği tartışılmamaktadır. Amaçlanan ya da amaçlanmayan çıktılar tam anlamıyla topluca sunulmayıp genellikle katılımcı görüşlerine göre programdaki zayıf yönler ortaya konmaktadır. Değerlendirme sürecini tam ve sistemli biçimde ortaya koyma zayıf kalan bir alt standarttır. Tezler tamamlanmadan önce uzman incelemesinden geçmekte ancak paydaş ya da katılımcılar tarafından incelenmemektedir. Değerlendirmenin sınırlılıkları tezlerdeki sınırlılıklar bölümünde çoğunlukla veri toplama süreç ve zaman aralığıyla ilgili olarak bildirilmiştir. Değerlendirmeye dair sınırlılıklar ve bunların sonuçlara etkisi tartışılmamaktadır.

3.3.5. Şeffaflık ve bilgilendirme: Değerlendirilen programla ilgili okuyucuların bilmesi gereken kanun ve kurallar açıklanmalıdır. Değerlendirmenin amacı ve içeriği okuyuculara açık olmalıdır. İncelenen öğretim programını hem destekleyen hem de eleştiren görüşlere yer verilmelidir. Veriler doğrudan, açık ve tam olarak sunulmalıdır. Sonuç ve öneriler dengeli ve doyurucu bir şekilde yazılmalı, literatürden ve bulgulardan yeterli derecede kanıtlarla desteklenmelidir. Değerlendirmenin sınırlılık ve bağlamı açıklanmalıdır. İncelenen tezlerde, değerlendirilen programla ilgili bilinmesi gereken kanun ve kurallar literatür bölümünde açıklanmaktadır. Tezlerin tamamı YÖK veri tabanında yer almaktadır ve okuyuculara açıktır. Bulgular genellikle yazılı olarak sunulmuştur ancak bazı analizlerin yapıldığı belirtilmesine rağmen ilgili çıktı ve bulgulara yer verilmediği de olmuştur. Programı destekleyen ve eleştiren görüşlere yer verme alt standardı çoğunlukla karşılanmaktadır. Verilerin sunumunda okuyucu dostu olmayan karmaşık ve ağdalı bir anlatımın kullanıldığı tezlere de rastlanmıştır. Verilerin tamlığı ve doğruluğu noktasında eksiklikler vardır. Özellikle istatistiksel analizlerde bazı hatalar göze çarptığından bulgular yanlış yorumlanabilmiştir (ölçeklerde kararsızım ifadesine üç puan verilmesi, varsayım analizlerinin yapılmaması, geçerlik ve güvenilirlik çalışmalarının yapılmaması gibi). Değerlendirmenin sınırlılıklarından yeterli derecede bahsedilmemiştir. Sonuç ve tartışma kısımları özellikle zayıftır. En önemli çıkarımlarda bulunulması gereken bu bölümler oldukça kısa tutulmuş ve kimi tezlerde programın etkililiği ile ilgili yeterli veriyi sağlamamıştır. Sonuç ve önerileri destekleyici kanıtlar gösterme alt standardı bazı tezlerde karşılanmamaktadır.

3.3.6. Çıkar çatışması: Değerlendirme çalışmalarında olası çıkar çatışmaları açıklanmalı ve bunlara yönelik ne gibi önlemler alındığı belirtilmelidir. Ayrıca, bağımsız uzmanlar ya da farklı paydaş gruplarınca değerlendirmede çıkar çatışması olup olmadığı

incelenmelidir. Tezlerde bu standart karşılanmamaktadır. Örneğin, polis akademisinde, meslek lisesinde ya da özel bir şirkette yönetici olarak görev yapan değerlendiriciler aynı zamanda paydaş olmalarının etkisini açıklamamışlardır. Veri toplama aracında katılımcıların zor durumda kalacaklarını düşündükleri, yöneticilerden ya da kurumdan çekindikleri sorular olup olmadığı ve bunlara ne ölçüde objektif yanıt verebildikleri, değerlendirme sürecinde objektifliği etkileyen bir durumun yaşanıp yaşanmadığı belirtilmemiştir. Değerlendirmenin tamamı uzman ya da paydaş gruplarınca çıkar çatışması yönünden incelenmemiştir.

3.4. Doğruluk Standartları

3.4.1. Sonuç ve kararlarda doğruluk: Bu standart, değerlendirme neticesinde varılan sonuçlar ve alınan kararların doğruluğuyla ve inandırıcılığıyla ilgilendirir. Başka bir ifadeyle, değerlendirme sonrası verilen kararların mantıksal temelleri ele alınır. Bu çerçevede değerlendirmenin sonuç bölümü, değerlendirme sorularına yanıt verecek şekilde düzenlenmiş olmalıdır. Değerlendirme esnasında yapılan iş ve işlemler ile elde edilen bulgularının eksiksiz biçimde paylaşılması gerekmektedir. Değerlendirme sonuçlarının hangi koşullarda anlam ifade ettiği de ayrıntılı biçimde açıklanmalıdır. Değerlendirme sonuçlarının daha sağlam bir zemine oturtulması amacıyla destekleyici nitelikteki bilgi, bulgu ve bilimsel eserlere atıf yapılması önemlidir. Değerlendirme sürecinde incelenen programın olumsuz tarafları varsa bunların da rapor edilmesi beklenir. Bulguların yorumlanmasında ve değerlendirilen programla ilgili kararların alınmasında çok yönlü ve objektif bir yaklaşım önerilmektedir. İncelenen tezlerin üçte ikisinde bu standartlar karşılanamamıştır. Tezlerin sonuç bölümü genel olarak değerlendirme sorularına yanıt verecek biçimde tasarlanmıştır. Benzer şekilde, program değerlendirme çalışmasının sınırlılıklarının yeterli olmasa dahi verildiği gözlenmiştir. Böyle olmakla birlikte, değerlendirme sonuçlarını güçlendirebilmek için literatürdeki çalışmalardan yeterli biçimde yararlanılmadığı söylenebilir. Değerlendirilen programların olumlu ve olumsuz yönlerini objektif biçimde ele alan çalışma sayısı çok azdır. Eğer araştırmacı kendi geliştirdiği programı değerlendirmekteyse olumsuz taraflardan hemen hiç bahsetmemektedir ki tezlerin bir kısmı bu şekildedir. Tezlerin hiçbirinde alternatif bakış açılarına ve değer yargılarına göre farklı yorumlar geliştirilmemiştir. Ayrıca, değerlendirme sonuçlarının yanlış yorumlanmaması için bir tedbir alındığı veya herhangi bir açıklamaya yer verildiği söylenemez.

3.4.2. Geçerli bilgi: Bu standart, değerlendirme sürecinin değerlendirme amaçlarına uygun biçimde yürütülmesine ve netice itibarıyla amaca uygun/geçerli bilgilere ulaşılmasına işaret etmektedir. Bu kapsamda, değerlendirme süreci bir bütün olarak değerlendirme ya da araştırma sorularına odaklı biçimde yürütülmelidir. Koşullar elverdiğince ölçme yönteminde çeşitliliğe gidilmelidir. Geçerli bilginin elde edilebilmesi için ölçülmek istenen psikolojik yapı veya kavramın çok iyi tanınması ve akabinde okuyucuya tanıtılması şarttır. Değerlendirme sürecinde yapılan iş, işlem ve uygulamaların hangi araştırma sorusuna yönelik olduğu da açıklığa kavuşturulmalıdır. Verilerin toplanması, kodlanması, puanlanması, analizi ve temaların belirlenmesi gibi aşamalar ayrıntılı biçimde anlatılmalıdır. Ortaya çıkan bulgu ve sonuçların başlangıçta belirlenen sorulara ne ölçüde yanıt verdiği samimi biçimde tartışılmalıdır. İfade edilen bu koşulların yalnızca tek bir tez tarafından karşılanmadığı, diğer bütün program değerlendirme çalışmalarının standardı karşılama açısından tatmin edici durumda olduğu belirtilebilir. Standardın gerektirdiği niteliklere ulaşamayan tezi ele aldığımızda öncelikli olarak araştırma sorularının bir program değerlendirme çalışması açısından uygun olmadığını söyleyebiliriz. Aynı şekilde, ölçme yönteminde çeşitliliğe gidilmediği görülmektedir. Verilerin toplanması ve analizi de tüm çıplaklığı ile açıklanmamıştır. Standardı karşılayan tezlerde ise her şeyden önce ölçülmek istenen kavramsal yapının derinlemesine incelendiği anlaşılmaktadır. Birçok çalışmada veri toplama araç ve

yöntemleri çeşitlendirilmiştir. Hangi araştırma sorusu için hangi tür bir sürecin izlendiği ve sonuçta nelerin elde edildiği rapor edilmiştir. Verilerin nasıl bir ortamda, hangi zaman zarfında toplandığı açıklanmıştır. Benzer şekilde, verilerin puanlanması, kodlanması ve temaların belirlenmesi gibi aşamalar tatmin edici biçimde açıklanmıştır. Standardı karşılayan çalışmalarda araştırma sorularının ne düzeyde cevaplandığına dair yorumlamada bulunulmamıştır. Her bir araştırma sorusunun cevaplanabilmesi için ayrı bir süreç izlenmiş ve veriler toplanmış ancak hedefe ne ölçüde ulaşıldığı nesnel biçimde tartışılmamıştır.

3.4.3. Güvenirlilik: Program değerlendirme çalışmasıyla ortaya çıkan bilginin güvenilir olması son derece önemlidir. Bu çalışmalarda hangi güvenilirlik testinin niçin kullanıldığı ve elde edilen katsayıların ne olduğu açıkça izah edilmelidir. Çalışmanın konusuna ve niteliğine göre daha önce kullanılmış ve güvenilirlik açısından kendisini kanıtlamış ölçme araçları tercih edilmelidir. Analiz birimi, kafalarda soru işaretine yol açmayacak netlikte belirtilmelidir. Güvenirlilik katsayısını etkileme potansiyeli bulunan katılımcı karakteristiği, değerlendirici yanlılığı, veri toplanan ortamın özellikleri vs. konular tartışılmalıdır. Verilerin puanlanması, analizi, kodlanması ve temaların oluşturulmasında tutarlılığın ne şekilde sağlandığı açıklanmalıdır. Gerekmesi halinde uzman yardımı alınmalıdır. Yeni geliştirilen bir ölçme aracı kullanılacak ise pilot çalışma yapılmalı ve ayrıntılar okuyucu ile paylaşılmalıdır. Son olarak, güvenilirlikle ilgili sınırlılıklar, problemler ve bu problemlerin değerlendirme kararlarına olası yansımaları içtenlikle ele alınmalıdır. İncelenen program değerlendirme tezlerinin çoğunda güvenilir bilginin üretilmediği anlaşılmaktadır. Bu tezlerin tümünde çok temel yöntemsel hatalar tespit edilmiştir. Örneğin, yeni bir ölçek geliştirildiği halde pilot çalışma ve güvenilirlik testlerinin yapılmadığı tezler mevcuttur. Pilot çalışma yapıldığını iddia eden ancak pilot çalışma verilerini paylaşmayan tezlerle de karşılaşmıştır. Güvenirlilik katsayısına etki edebilecek faktörler, birçok tezde göz ardı edilmiştir. Bununla bağlantılı olarak, güvenilirlikle ilgili sınırlılıklar ve bu sınırlılıkların program değerlendirme sürecine nasıl yansıdığı hiçbir tezde tartışılmamıştır. Ayrıca, değerlendirme çalışmalarının özellikle nitel boyutlarında güvenilirlik üzerinde ya çok az durulmuş ya da hiç durulmamıştır. Çalışmaların büyük çoğunluğunda araştırmacıdan kaynaklı yanlılık üzerinde hiç durulmamıştır. Dokuz değerlendirme çalışmasından ikisinde güvenilirlik standartları karşılanmıştır. Bu iki çalışmanın belirgin tek eksiği güvenilirlikle ilgili sınırlılıkları göz ardı etmiş olmalarıdır.

3.4.4. Programı betimleme: Bu standart, değerlendirilecek programın amaçları, içeriği ve işleyişi ile ilgili doyurucu bir analizi gerekli kılmaktadır. Şöyle ki, değerlendirilecek olan programın özelliklerine yönelik farklı veri kaynaklarından etraflıca bilgi toplanır. Bu veri kaynaklarının başında paydaşlar ile resmi evraklar gelir. Programla ilgili tanıtıcı bilgilerin toplanmasını takiben programın işleyişinin derinlemesine analizi gereklidir. Program mevcut haliyle ihtiyaçlara ne ölçüde cevap verebilmektedir? sorusuna cevap aranır. Bu soru için de olabildiğince çeşitli veri kaynaklarına başvurulur. En sonunda ise yazılı program ile gerçekleşen program arasındaki farklılıkların tespit edilmesi ve yorumlanması ile sözü edilen standart karşılanmış olur. İncelenen tezlerin çoğunda değerlendirilen programla ilgili yeterli bilgilendirme yapılmışsa da az sayıda tezde programın işleyişi, felsefi temelleri ve amaçlarına dair yüzeysel bilgi sunulmuştur. Yazılı program ile sahada gerçekleşen programı birebir karşılaştırmak yerine katılımcı görüşleri doğrultusunda uygulamayı aktarma daha fazla tercih edilmiştir. Bununla birlikte programlara temel oluşturan yasal mevzuat ile program tasarımlarını karşılaştıran ve sonraki süreçlerde de program tasarımlarıyla sahadaki uygulanış arasındaki farkları irdeleyen tezler de bulunmaktadır. Standardı karşılayabilen tezlerin başka bir belirgin özelliği de programa dair bilgi toplarken mümkün olduğunca kaynak çeşitliliğine gitmeleridir.

3.4.5. Bilgi yönetimi: Program değerlendirme sürecinde karşılanması beklenen standartlardan birisi de bilginin doğru ve etkili yönetimidir. Bu standart, bilginin toplanmasından raporlama aşamasına kadar bir dizi adımdan ibarettir. Öncelikli olarak, bilgi kaynaklarının çeşitlendirilmesi gerekmektedir. Bunun yanında veri toplama yöntemlerinin de çeşitlendirilmesi önemlidir. Değerlendirilmesi planlanan programla ilgili daha önce yapılmış bir inceleme varsa bu çalışmadan da faydalanılması gerekmektedir. Hangi veri kaynağının niçin seçildiğine dair ikna edici açıklamalar beklenmektedir. Verilerin toplanması kadar verilerin toplandığı grup ya da örneklem de önemlidir. Bu örneklem ya da grubun da okurlara tanıtımı ihmal edilmemelidir. Veri toplama araçları çalışmanın *Ekler* bölümünde veya içeriğinde mutlaka yer almalıdır. Çalışmanın sonuna gelindiğinde verilerin doğruluğuna ilişkin bir şüphe oluştuysa bu durum da dürüstçe rapor edilmelidir. Bütün bu gereklilikler ile incelenen dokuz program değerlendirme çalışmasını kıyasladığımızda toplamda altı çalışmanın bilgi yönetimi standardını karşıladığını söyleyebiliriz. Kalan üç çalışmanın ise ortalamanın biraz üzerinde ancak arzu edilen noktanın da uzağında bulunduğu anlaşılmaktadır. Söz konusu üç çalışma yakından incelendiğinde gerekçelendirme sorununun yaşandığı ortaya çıkmıştır. Hangi veri kaynaklarının, hangi veri toplama yöntem ve tekniklerinin ne amaçla seçildiğine dair açıklamalar ya yoktur ya da doyurucu düzeyde değildir. Hedef evren, ulaşılabilir evren, seçilen örneklem ve gerçekleşen örneklemden bahsedilmemiştir. Bu çalışmalarda ortaya çıkan diğer bir zaaf ise yanlışlık sorunu ile hiç karşılaşmamış gibi davranılmasıdır. Şöyle ki, en üst düzey çalışmalarda bile örneklem seçimi ile verilerin toplanması esnasında öngörülemez bazı olumsuzluklar yaşanmakta ve güvenilirlik katsayısı bundan etkilenmektedir. Ancak, bahsi geçen üç çalışmada bu tür durumlar rapor edilmemiştir. Esasında toplanan verilerin objektifliğine dair şüpheli durumlar, standardı karşılayabilen program değerlendirme çalışmalarında da açıklanmamıştır. Standardı karşılayabilen tezlerin en çok öne çıkan artısı veri toplama araçları ile yöntemlerinde çeşitliliğe gitmiş olmalarıdır. Dört veya beş farklı veri kaynağı kullanan program değerlendirme çalışmalarına dahi rastlanmıştır. Yine bu altı çalışmada görülen olumlu yönlerden birisi de veri toplama araçlarını şeffaf biçimde paylaşmış olmalarıdır. Bu çalışmalarda tıpkı yetersiz olan üç çalışmada olduğu gibi hedef evrenden bahsedilmemiştir. Ancak, seçilen örneklem ve bunların katılım oranları (geri dönüş oranları) belirtilmiştir.

3.4.6. Değerlendirme deseni ve analiz: Bir program değerlendirme çalışmasında verilerin doğru girilip girilmediğinin kontrol edilmesi; kodlama, puanlama ve analiz aşamaları için de ön koşuldur. Veri girişini takiben kodlama ve puanlama aşamaları da belli aralıklarla gözden geçirilmelidir. Program değerlendirme çalışmalarının bir bütün olarak uzman denetimine sokulması gerekmektedir. Gerekmesi halinde istatistiksel sonuç tabloları ile bu tablolara ilişkin yorumlar da uzman denetimine sokulmalıdır. Gözlem, görüşme, anket ve ölçekler yardımıyla toplanan veriler için katılımcılardan teyit alınmalı ve eğer varsa katılımcıların düzeltme veya ekleme yapmasına izin verilmelidir. Değerlendirme sırasında toplanan bilgiler, değerlendirme sorularıyla ilgili olmalıdır. Bulguların yazımında, paydaş görüşlerinin alınması önemlidir. Ayrıca, toplanan verilerin özetlenmesi de yine değerlendirme sorularına yanıt verir nitelikte olmalıdır. İncelenen tezlerin yaklaşık yarısı (n=4), değerlendirme deseni ve analiz standardı açısından ortalamanın üstüne çıkamamıştır. Dört tezin vasatın üzerinde bir performans sergilemekle birlikte standardı tam anlamıyla karşılamamanın uzağında kaldığı görülmüştür. Standartların gerisinde kalan sekiz program değerlendirme çalışmasının hiçbirinde veri girişlerinin kontrol edilip edilmediği bilinmemektedir. Benzer şekilde, puanlama ve kodlamanın da kontrol edilip edilmediğine dair bir bilgilendirme yoktur. Yapılan istatistiksel analizlerin ve bu analizlere bağlı geliştirilen yorumların uzman denetimine tabi tutulmadığı görülmüştür ancak burada belirtmek gerekir ki incelenen dokuz tezin hiçbirinde

uzman denetimini gerektirecek boyutta üst düzey bir istatistik kullanılmamıştır. Aynı şekilde, hiçbir tezde katılımcılardan verilere ilişkin teyit alındığına dair işaret bulunmamıştır. Bulguların yazımı ve yorumlanmasında paydaşlardan görüş alan herhangi bir program değerlendirme tezi bulunmamaktadır. Öte yandan, değerlendirme sürecinde toplanan bilgiler, bu bilgilerin analizi ve özetlenmesinin değerlendirme sorularına yanıt verir nitelikte olduğu söylenebilir.

3.4.7. İstatistiksel kanıtlar: Başta varsayım testleri olmak üzere doğru istatistiksel testlerin seçimi, yorumlanması ve raporlanması, program değerlendirme çalışmalarının kalitesinde önemli rol oynamaktadır. Bu kapsamda nicel analizler için ön koşul niteliğindeki varsayım testleri yapılmalı ve varsayımların karşılanması halinde testlere ya da analizlere devam edilmelidir. Varsayım testlerinin karşılandığına dair ikna edici kanıtların raporlanması da gerekmektedir. Veri setinin toplanmasıyla ortaya çıkan merkezi dağılım ölçüleri ve varsa aykırı değerler saptanmalıdır. Veriler, anlaşılır biçimde ve uygun görseller (tablo, şekil vs.) yardımıyla özetlenmelidir. İstatistiksel anlamlılığın yanı sıra pratikte anlamlılık yani etki büyüklükleri ele hesaplanıp yorumlanmalıdır. Nitel ve nicel analizler sonucunda ortaya çıkan bulguların sınırlılıkları da tartışılmalıdır. Nitel araştırmalarda değerlendirme sorularına cevap veren temalar kullanılmalıdır. Ayrıca hem nitel hem de nicel değerlendirmeler için geçerlik, güvenilirlik çalışmaları yapılmalıdır. İncelenen tezlerin çoğunda istatistiksel kanıtlar açısından geliştirilmesi gereken pek çok nokta vardır. Özellikle, pratikteki anlamlılığın yorumlanmasına hiçbir tezde değinilmemiştir. Etki büyüklüğünün hesaplanmadığı veya yanlış etki büyüklüğü göstergesinin raporlandığı olmuştur. Bir program değerlendirme çalışmasında ise istatistiksel anlamlılık dahi yorumlanmamıştır. Analizler sonucu yapılan yorumların sınırlılıklar açısından değerlendirilmesi, varsayım testlerinin yapılması ve uç değerlerin ele alınması da eksik bırakılan noktalardandır. Fark istatistiğine bakan tezler bulunsa da ortak etkiden söz eden herhangi bir teze rastlanmamıştır. Merkezi dağılım ölçülerinin verilmesi ve uygun görsel kullanımı alt standartları genellikle karşılanmaktadır. Nitel araştırmalarda uygun temalandırma alt standardı karşılanmakla birlikte geçerlik ve güvenilirlik çalışmalarında eksikler gözlenmiştir.

3.4.8. Raporlama: Program değerlendirme tezlerinin yazımında yani raporlaştırma sürecinde raporun adil, tarafsız ve objektif olması gerekmektedir. Çarpıtmalardan ve öznel ifadelerden sakınılmalıdır. Tez izleme komitesi başta olmak üzere uzman görüşlerine göre raporda gerekli düzenlemeler yapılmalıdır. Ayrıca, sonuç ve tartışma kısımlarında paydaşların değer yargıları dikkate alınmalıdır. Program değerlendirme raporuyla ilgili olarak bağımsız uzman görüşlerine başvurulması da öneriler arasındadır. İncelenen tezler raporlaştırma standardını istenen düzeyde karşılayamamaktadır. Tezlerin son halleri araştırmacı ve tez komisyonu dışında üçüncü bir göz tarafından ele alınmamıştır. Yanlılık ihtimaline karşı herhangi bir kontrol mekanizmasından bahsedilmemektedir. Standardı karşılayan ve karşılamayan tezlerde gerçek dışı ve çarpıtma bir yorum yoktur. Tezlerin doğası gereği paydaşlara ait farklı değer yargılarının dikkate alınması çok gerekmemiştir. Yine tez çalışmalarının doğal süreci içerisinde tez izleme komisyonlarının denetiminden geçtiği rahatlıkla ifade edilebilir.

4. SONUÇ ve TARTIŞMA

Son beş yılda yayınlanmış dokuz program değerlendirme tezi, Stufflebeam'in meta değerlendirme kontrol listesindeki dört boyut ve 26 standarda göre incelenmiş ve tezlerin bu standartları karşılamada gösterdikleri ortalama performanslar bu bölümde tartışılmıştır.

Meta değerlendirme kontrol listesinin **yararlılık** boyutunda bulunan sekiz standart, program değerlendirme tezlerinin başlangıçta belirtilen amaçlara ne ölçüde uygun olduklarını

ve pratikte ne ölçüde fayda sağladıklarını tespit etmeye yöneliktir. Bu standartlardan “değerlendirici güvenilirliği”, “değerlere saygı” ile “süreç ve ürünler” incelenen tezler tarafından büyük oranda karşılanmıştır. Söz konusu üç standardın büyük oranda karşılanması, öncelikle program değerlendirme çalışmalarının alan bilgisi ve tecrübesine sahip kişilerce yürütüldüğünü göstermektedir. Tez yazarlarının tümü, eğitim programları ve öğretim bilim dalında doktora eğitimi almışlardır. Üzerinde çalışılan grupların sosyal, kültürel ve ekonomik özelliklerinin iyi analiz edildiği ve dolayısıyla tez yazarlarıyla katılımcılar arasında güven ilişkisinin tesis edilebildiği görülmektedir. İkinci olarak, program değerlendirme çalışmasına temel oluşturabilecek politik ve sosyal bağlam ile ilgili mevzuatın genellikle iyi irdelendiği anlaşılmaktadır. Değerlendirme aşamasına geçilmeden eğitim programlarıyla ilgili tüm yasal çerçeve ile programları uygulayan veya programdan etkilenen kurumların bakış açıları ele alınmıştır. Böylelikle, ortaya çıkan bulguların yorumlanmasında farklı görüş, grup ve kurumların değer yargılarını hesaba katmak mümkün olmuştur. Üçüncü olarak, değerlendirilen programların; hedef, içerik, eğitim durumları ve sınav durumları başta olmak üzere tüm boyutlarıyla ve sade bir dil kullanılarak tanıtıldığı söylenebilir. Benzer şekilde, değerlendirme çalışmasının amaçları, takip edilen aşamaları ve bulguları da şeffaf biçimde paylaşılmıştır.

Yararlık boyutunda yer alan “amaçlar”, “iletişim” ve “sonuçlar ve etki” standartlarının program değerlendirme tezleri tarafından karşılanmadığı görülmektedir. Bu düşük performansın çeşitli nedenleri vardır. Bulgular kısmında da belirtildiği üzere ‘amaçlar’ standardı, değerlendirme hedeflerinin farklı pencerelerden tanımlanması esasına dayanır. Paydaşlar, hedef kitle ve ilgili kurumların değerlendirme sürecinden ne beklediği ve değerlendirme sonuçlarını nasıl kullanacağı izah edilmelidir. Bulguların yorumlanmasında da farklı paydaş görüşlerinin hesaba katılması önerilmektedir. Ayrıca, program değerlendirme için öngörülen süreç ve işlemlerin ne ölçüde gerçekleştirilebildiği de dürüst biçimde rapor edilmelidir. Program değerlendirme tezlerinin bahsi geçen bu şartların çok gerisinde kaldığı görülmektedir. ‘Problem durumları’ veya ‘araştırmanın önemi’ gibi başlıklarda paydaşlar ve ilgili kurumların değerlendirme sürecine yönelik beklentilerinden söz edilmesi gerekmektedir. İncelenen program değerlendirme tezlerinde ise böyle bir bilgi ve yoruma rastlanmamıştır. Değerlendirme süreci öncesinde hedef kitle, paydaşlar ve ilgili kurumlara yönelik bir ihtiyaç analizinin yapılmadığı görülmektedir. Farklı kesimlerinin bakış açıları öğrenilemediği için bulguların yorumlanmasında alternatifler göz ardı edilmiştir. Değerlendirme öncesinde planlanan süreç ve işlemlerin ne oranda gerçekleştirilebildiğine yönelik şeffaf davranan tez sayısı da sınırlı kalmıştır. “Sonuçlar ve etki” standardı, değerlendirme sonuçlarının etkili kullanımına yönelik olarak geliştirilmiştir. Değerlendirme sonuçlarının etkili kullanılabilmesi ise günlük eğitim yaşantılarına hitap eden somut anlatımlarla ve örneklerle mümkündür. Oysaki incelenen program değerlendirme çalışmalarının büyük çoğunluğunda (n=7) son derece soyut ifadelerden oluşan anlatımlar bulunmaktadır. Yine tezlerin öneri kısımlarında eğitimci veya öğrencilere hitaben bir yönlendirme söz konusu olmamıştır. Daha ziyade, hayata geçirilmesi zor olan bazı genel tavsiyelere yer verilmiştir. İletişim standardında da benzer bir durum gözlenmiştir. Program değerlendirme tezlerinin yarısından çoğu (n=5), bu standartla ortaya konulan beklentilerin uzağında kalmışlardır. Değerlendirme esnasında katılımcılarla görüş alışverişinin sınırlı tutulduğu ve süreklilik arz etmediği görülmektedir. Aynı şekilde, paydaşlar ve programı yürüten kurumlarla kurulan iletişim de sınırlı kalmıştır. İletişimdeki kopukluk nedeniyle araştırmacı dışında herhangi bir kişi ya da grubun değerlendirme sürecine gerçek manada bir katkısından söz edilememektedir.

Yararlık boyutunda hiç karşılanamayan ve büyük oranda karşılanan standartların yanı sıra kısmen karşılanan iki standart bulunmaktadır: “*paydaş odaklılık*” ile “*verilerin niteliği*”. Söz konusu standartlar, dokuz program değerlendirme tezinden beşi tarafından ‘çok iyi’ veya ‘mükemmel’ düzeyde karşılanmıştır. Diğer taraftan, dört program değerlendirme tezinin standartların gerektirdiği niteliklere sahip olamadıkları görülmüştür. ‘Paydaş odaklılık’ standardı, her şeyden önce kimlerin paydaş olduğunun doğru biçimde saptanmasına dayanır. Sonrasında, paydaşların işine yarayacak değerlendirme bulgularının neler olabileceği hesaba katılmalıdır. Ayrıca, değerlendirme sürecinin başından sonuna kadar paydaş görüşlerinden istifade edilmesi de gerekmektedir. Tezlere bakıldığında paydaşların belirlenmesiyle ilgili genel bir sıkıntıya rastlanmamıştır. Program değerlendirme tezlerinin hitap ettikleri kesimler, kurumlar, paydaşlar ve diğer hedef kitle açısından kapsayıcı nitelikte oldukları da söylenebilir. Ancak, paydaşların ne tür değerlendirme bulgularına ihtiyaç duydukları birçok tezde incelenmemiştir. Başka bir anlatımla, paydaşların yürütülmekte olan eğitim programıyla ilgili hangi bilgiye ulaşmak istedikleri sorulmamış bunun yerine tez yazarının ulaşmak istediği bilgilere odaklanılmıştır. ‘Verilerin niteliği’ standardı, değerlendirilen programın genel kalitesini ortaya çıkarabilecek yeterlilikteki bilgilere işaret etmektedir. Yeterli ve derinlikli bilgilerin elde edilebilmesi için ve dolayısıyla, üzerinde çalışılan programın daha iyi anlaşılabilmesi için değerlendirme soruları genişletilmeli veya kimi durumlarda güncellenmelidir. Program değerlendirme tezlerine genel olarak bakıldığında yarı yapılandırılmış görüşme formlarının kullanıldığı anlaşılmaktadır. Ayrıca, birçok tezde ölçek geliştirme çalışmalarının da yapıldığı tespit edilmiştir. Ölçek geliştirme süreci ile yarı yapılandırılmış formların özellikleri dikkate alındığında değerlendirme soruları üzerinde gerekli düzenleme ve eklemelerin yapıldığı söylenebilir. Öte yandan, eğitim programlarının genel kalite, verimlilik ve önemini ortaya çıkarabilecek; sahada bulunan paydaş ve uygulayıcıların ihtiyaçlarına cevap verebilecek nitelikte kapsamlı çalışmaların yapıldığını söylemek zordur.

Meta değerlendirme kontrol listesinin **uygulanabilirlik** boyutunda bulunan “*proje yönetimi*” standardı, program değerlendirme tezlerinin tamamı tarafından çok iyi veya mükemmel derecede karşılanmıştır. Proje yönetimi, katılımcılarla kurulan ilişki biçimlerinden mahrem bilgilerin gizli tutulmasına veya sosyal-kültürel hassasiyetlerin gözetilmesine kadar bir dizi konuda profesyonel bir tutum sergilenmesini zorunlu kılmaktadır. Program değerlendirme tezleri, yapısal özelliklerine bakıldığında inançsal, kültürel ve sosyal açıdan riskli konular içermemektedir. Kişilerin özel hayatlarına dair konuların gündeme gelmediği ve dolayısıyla bu konularda sorun yaşanmadığı anlaşılmaktadır. Katılımcılardan yalnızca bazı kişisel bilgiler istenmiş ve bu tür bilgilerin gizliliği tüm tezlerde sağlanmıştır. Katılımcı ve diğer paydaşlarla profesyonelliğin dışına taşan bir ilişki kurulmadığı rahatlıkla ifade edilebilir. Uygulanabilirlik boyutunda tezlerin başarı sağladığı ikinci standart “*uygulama süreçleri*” olmuştur. Uygulama süreçleri standardında başarı sağlayabilmenin yolu program değerlendirme için seçilen metot, işlem basamakları ve çalışma takviminin makul ve inandırıcı olmasından geçmektedir. Tezlerde anlatılan uygulama adım ve takvimleri gerçekleştirilebilir gözükmektedir. Veri toplama araç ve yöntemlerinde araştırmanın gidişatına ve toplanan verilerin özelliklerine göre zaman zaman bazı düzenlemelere gidilmiştir. Öte yandan, program değerlendirme tezlerinden ikisi söz konusu standartların son derece gerisinde kalmışlardır. Bunun temel nedeni, veri toplama yöntemleriyle ilgili proje, kurum veya ekip desteği olmaksızın gerçekleştirilemeyecek hayali bir tablodan söz etmeleridir.

Uygulanabilirlik boyutunda yer alan “*kaynak kullanımı*” standardı, dokuz program değerlendirme tezinden sekizi tarafından karşılanamamıştır. Bu standart, kaynakların verimli

ve etkili kullanılması gerektiğine işaret eder. Program değerlendirme çalışmalarından beklenen şey, uygulayıcılarının işlerini kolaylaştıracak pratik ve somut katkılardır. Ayrıca programla ilgili kararlara (programa devam, programın revize edilmesi veya programın sonlandırılması) da objektif verilerle katkı sağlanmış olmalıdır. İncelenen program değerlendirme çalışmaları, alanda görev yapan eğitimcilere veya diğer personele herhangi bir pratik katkı vadedememiştir. Mevcut programların devamını, güncellenmesini veya sonlandırılmasını etkileyebilecek ölçüde ve kalitede bir program değerlendirme tezine rastlanamamıştır. Yeni bakış açılarının oluşturulmadığı tezlerde haliyle de tartışma ve sonuç kısımları zayıf kalmıştır. Program değerlendirme tezlerinin önemli sorunlarından biri de pratikte neler yapılması gerektiğini açıklayamamalarıdır. Bu nedenle, yeni program geliştirme çalışmalarına öncülük edemedikleri söylenebilir.

Meta değerlendirme kontrol listesinin *yerindelik* boyutunda yer alan toplam altı standardın program değerlendirme tezleri tarafından ne ölçüde karşılanabildiği incelenmiştir. Program değerlendirme tezlerinin büyük oranda karşılayabildiği iki yerindelik standardı tespit edilmiştir. Bunlar, “yasal çerçeve” ile “katılımcı hakları” standartlarıdır. *Yasal çerçeve* standardı, adından da anlaşılacağı üzere program değerlendirme sırasındaki tüm iş ve işlemlerin yasal bir zemine oturtulmasını amaçlar. Program değerlendirmede; amaçlar, araştırma soruları, dönüt ve düzeltmeler, iş akış takvimi, kişisel bilgilerin korunması ve meta değerlendirme gibi konularda her şeyin ilgili mevzuat çerçevesinde yürümesi gerekmektedir. İncelenen program değerlendirme tezleri bu kurala büyük ölçüde uyum sağlamıştır zira tez izleme komisyonlarının varlığı, bu konularda hem uzlaşmanın sağlanması hem de uzlaşılan konuların takibini zorunlu kılmaktadır. *Katılımcı hakları* standardı; kişisel bilgilerin gizliliğini sağlamak, değerlendirmenin amacını bulanık bir nokta bırakmaksızın katılımcılara açıklamak, katılımcıların değerlerine saygılı olmak gibi temel insan hakları ve etik ilkelere işaret eder. Değerlendirme sürecinin amaçları birçok çalışmada net biçimde açıklanmıştır. Kişisel bilgilerinin gizli kalacağına dair yazılı bilgilendirmenin yapıldığı, tez eklerinde yer alan veri toplama araçlarından anlaşılmaktadır. Katılımcıların kültürel ve sosyal farklılıklarının gündeme geleceği veya hassasiyet oluşturacağı herhangi bir değerlendirme çalışması yoktur. Bu nedenlerle, genel anlamda, *katılımcı hakları* standardının getirdiği şartlara uyulmuştur.

Yerindelik boyutunda bulunan standartlardan üçü, program değerlendirme tezlerinin büyük çoğunluğu tarafından karşılanamamıştır. Bunlar; *oryantasyon*, *açıklık ve yansızlık* ile *şeffaflık ve bilgilendirme*. Oryantasyon standardı, değerlendirilecek programın güçlü ve zayıf yönlerinin analiz edilmesiyle ilgilidir. Burada, çıktılar ile paydaş ihtiyaçlarının karşılaştırılması önem arz etmektedir. Program değerlendirme tezlerinden beş tanesi oryantasyon standardını karşılayamamıştır. Bu tezlerde paydaş ihtiyaçları belirlenmediği için daha sonra ihtiyaç-çıktı karşılaştırması da mümkün olamamıştır. Paydaş ihtiyaçlarının saptanması bir kenara program çıktıları dahi araştırmayan tezler bulunmaktadır. Program değerlendirme tezlerinden dört tanesi, oryantasyon standardını “çok iyi” veya “mükemmel” düzeyde karşılamıştır. Bu program değerlendirme tezlerinden bir kısmı aynı zamanda program geliştirme tezidir. Dolayısıyla, görece kapsamlı ihtiyaç analizleri yapılmıştır. “Çok iyi” şeklinde puanlanan tezlerde ise ihtiyaç analizleri zayıf kalmış ancak programın güçlü ve zayıf yönleri ciddi biçimde ve program çıktıları temel alınarak tartışılmıştır. *Açıklık ve yalınlık*, değerlendirme raporunun hazırlanmasındaki samimiyet ve şeffaflıkla ilgili bir standarttır. Raporunda, arzu edilen çıktıların yanı sıra arzu edilmeyen değerlendirme çıktılarına da yer verilmesi beklenir. Değerlendirme sürecine yönelik eleştiriler de açıkça paylaşılmalıdır. Sınırlılıklar ve bu sınırlılıkların değerlendirme çalışmasındaki kaliteyi nasıl etkilediği de belirtilmelidir. Dokuz program değerlendirme tezinden yedi tanesi, söz konusu ilkelerin son

derece gerisinde kalmışlardır. Sınırlılıklar ve bunların değerlendirme sürecindeki genel niteliğe yansımaları hiçbir tezde tartışılmamıştır. Amaçlanmayan çıktıların rapora yansıtıldığını söylemek zordur. Tezler, mevzuat gereği komisyon denetiminden geçtikleri için komisyonlarda yöneltilen eleştirilere veya önerilere göre birtakım düzenlemelerin yapıldığı varsayılabilir. Son olarak, değerlendirilen programların sadece aksaklıklarından bahsedilmiş olup iyi işleyen, paydaşların ve katılımcıların memnun olduğu noktalar genellikle görmezden gelinmiştir. Şeffaflık ve bilgilendirme standardı, birçok alt standart açısından birbirini tamamlar niteliktedir. Şeffaflık ve bilgilendirme, adından da anlaşılabilceği üzere tüm değerlendirme sürecine değil, daha ziyade raporlama aşamasına odaklanmaktadır. Üç program değerlendirme tezi, şeffaflık standardını karşılarken altı tez ise şeffaflığı sağlayamamıştır. Tezler, YÖK veri tabanından herkese açıktır. Bu, artı bir özelliktir. Öte yandan, değerlendirilen programlarla ilgili olarak hem destekler nitelikte hem de eleştiri niteliğinde görüşlere birlikte yer verilmesi gerekirken genelde eleştiriler yansıtılmıştır. İstatistiksel analizlerin sonuçları, doyurucu biçimde paylaşılmadığı gibi ölçek hazırlama süreçlerinde yapılan bazı temel hatalar ('kararsızım' ifadesine üç puan verilmesi gibi), bütün bulguları ve bulgulara bağlı yorumları sakatlamıştır. Benzer şekilde, varsayım testleri ile geçerlilik ve güvenilirlik testlerinin yapılmadığı görülmüştür.

Yerindelik boyutunda bulunan *çıkar çatışması*, yukarıda bahsedilen standartlardan negatif olarak ayrılmaktadır. Bu standart, dokuz program değerlendirme tezinin hiçbirisi tarafından karşılanamamıştır. Tezlerin tümü için, 10 puanlık derecelendirme sisteminde 0-2 puana tekabül eden "zayıf" nitelendirmesi yapılmıştır. Tezlerin bu standardı karşılayamamasında belli dinamikler rol oynamıştır. İyi bir program değerlendirme çalışmasında potansiyel çıkar çatışmalarının önceden saptanması ve gerekli tedbirlerin alınması gerekir. Çıkar çatışmalarının değerlendirme sürecine olan etkisini asgariye indirmek adına farklı değerlendirme uzmanları işe koşulmalı yahut farklı paydaş gruplarından görüş alınmalıdır. Dokuz program değerlendirme tezinin hiçbir tanesinde çıkar çatışmasına yönelik bir bulgu ya da yoruma rastlanmamıştır. Örneğin, kendi çalıştığı kurumda program geliştiren, programı yürüten ve sonra da bu programı uygulayan tez yazarları bulunmaktadır. Hem yönetici hem paydaş ve hem de değerlendirici olmanın objektifliği zedeleyip zedelemeyeceği konusuna girilmemiştir. Programın uygulayıcılarını, yöneticilerini, hedef kitleyi veya politika yapımcıları rahatsız edecek soruların kullanılıp kullanılmadığı, eğer kullanıldıysa ne gibi dönütler alındığı net biçimde rapor edilmemiştir.

Meta değerlendirme kontrol listesinin **doğruluk** boyutunda toplam sekiz standart bulunmaktadır: *sonuç ve kararlarda doğruluk, geçerli bilgi, güvenilirlik, programı betimleme, bilgi yönetimi, değerlendirme deseni ve analizler, istatistiksel kanıtlar, iletişim ve raporlama*. Bunlardan 'geçerli bilgi' ve 'bilgi yönetimi', program değerlendirme tezlerinin büyük çoğunluğu tarafından karşılanmıştır. *Geçerli bilgi* standardı, değerlendirme sürecinin bir bütün olarak başlangıçta belirlenen amaçlara uygun biçimde yürütülmesini esas alır. Program değerlendirme tezlerinden sekizi, geçerli bilgi standardını tatmin edici düzeyde sağlamaktadır. Tezler, ölçme yöntemlerinde çeşitlilik sağlayarak daha doğru sonuçlara ulaşabilmişlerdir. Ölçmek istenen psikolojik yapı, değişken ve kavramların ayrıntılı biçimde tanımlandığı söylenebilir. Verilerin toplanması, kodlanması, yorumlanması ve sonuçların yazımında takip edilen yollar, ayrıntılı biçimde açıklanmıştır. Standardı karşılayamayan tek doktora tezinde ise araştırma sorularının doğru seçilememesinden başlayarak değerlendirme sürecinin bütününe yansıyan aksaklıklar saptanmıştır. Program değerlendirme tezlerinin büyük oranda başarı sağladıkları diğer bir standart, *bilgi yönetimidir*. Bilgi yönetimi ile kastedilen, bilginin doğru kaynaklardan doğru veri toplama araçlarıyla toplanması ve sonrasında yapılan tüm işlemlerin

ikna edici biçimde açıklanmasıdır. Tezlerden altısı, bilgi yönetimi standardını ‘çok iyi’ düzeyde karşılamışlardır. Bu tezlerin göze çarpan özelliği, veri toplama araçlarını çeşitlendirmiş olmalarıdır. Yine bu tezlerde çalışma grupları ya da örnekleme ile ilgili yeterli düzeyde bilgiler verilmiştir. Üç tez ise standardı karşılama noktasında zayıf kalmıştır. Tezlerdeki genel sorun, seçilen bilgi kaynakları ile veri toplama yöntemlerini gerekçelendirmemiş olmalarıdır. Ayrıca, örnekleme deseni ile yanıtlama oranlarına dair bilgilerin verilmemiş olduğu anlaşılmaktadır.

Doğruluk boyutunda yer alan diğer altı standart, program değerlendirme tezleri tarafından tam anlamıyla karşılanamamıştır. Tezlerin en zayıf performans sergiledikleri standart, *istatistiksel kanıtlar* olmuştur. İstatistiksel analizlerle ilgili olarak ciddi mesafe kaydedilmesi gerekmektedir. Örneğin, pratikte anlamlılık ifadesi olan etki büyüklüğü ya hesaplanmamış ya da yanlış etki büyüklüğü katsayısı hesaplanmıştır. Verilerin hazır hale getirilmesi, aykırı değerlerin ayıklanması, varsayım testlerinin yapılması gibi noktalarda da bir hayli eksiklik tespit edilmiştir. *Değerlendirme deseni ve analizler*, değerlendirme sırasında toplanan nitel ve nicel verilerin doğru girilmesi; nitel analizler sonucu elde edilen temalar ile istatistiksel analizler sonucu ortaya çıkan tabloların doğru yorumlanması; bütün bir değerlendirme sürecinin değerlendirme soruları ile paralel yürümesi gibi hususları kapsar. Program değerlendirme tezleri yakından incelendiğinde bu standardın bir tez haricinde karşılanmadığı görülmüştür. Nitel ve nicel veri girişlerinin kontrol edilmediği, katılımcılardan bu konuda teyit alınmadığı anlaşılmaktadır. Program değerlendirme desenlerine bakıldığında bütün bir sürecin değerlendirme sorularına yanıt verme odaklı olduğu söylenebilir. Bu, program değerlendirme çalışmaları açısından bir artıdır. Ancak, ortaya çıkan tablo ve temaların yorumlanmasında paydaş, katılımcı ve uzman gibi unsurlardan yardım veya görüş alınmadığını da belirtmekte fayda vardır. *Sonuç ve kararlarda doğruluk* standardı, program değerlendirme tezlerinin başarı sağlayamadığı standartlardan biridir. Değerlendirme sonunda ele alınan eğitim programıyla ilgili olarak doğru ve inandırıcı bir karar verilmesi, standardın karşılanması açısından önemlidir. Program değerlendirme tezlerinin büyük kısmı, inceledikleri programlarla ilgili olarak objektif, çok yönlü ve sağlam kararlar alamamışlardır. Değerlendirme kararlarının hangi kapsam, bağlam ve eğitimsel yaşantılarla sınırlandırıldığı anlaşılamamaktadır. Alternatif bakış açıları ve yorumların karar aşamasında hesaba katılmadığı söylenebilir. Değerlendirme tezlerinde dikkat çeken bir diğer nokta da literatür ile bulguların harmanlanmamış olmasıdır. Literatürde yer alan çalışmaların, değerlendirme bulgularını destekleyip desteklemediği tartışılmamıştır. Program değerlendirme tezlerinin yetersiz kaldığı standartlardan biri de *güvenirliliktir*. Güvenirlik standardının karşılanabilmesi için program değerlendirme esnasında kullanılan tüm güvenirlilik yöntemlerinin mantıksal temelleriyle birlikte açıklanması gerekmektedir. Güvenirlilik testlerinde ortaya çıkan düşük katsayıların nedenleri tartışılmalı ve güvenirlilikle ilgili ortaya çıkan bu durumun değerlendirme kararına olası yansımaları ele alınmalıdır. Bu açıdan bakıldığında, incelenen dokuz program değerlendirme tezinin yedisinde güvenilir bilgilerin üretilmediği ifade edilebilir. Pilot çalışmalarda veya yeni geliştirilen ölçeklerde güvenirlilik katsayısının paylaşılması gibi eksiklikler mevcuttur. Hesaplanan güvenirlilik katsayılarında ise katsayıların değerlendirme sürecine olası etkileri yorumlanmamıştır. *Programı betimleme* standardı, değerlendirilen programın çok yönlü olarak tanıtılmasını gerekli kılmaktadır. Birçok farklı bilgi kaynağından yararlanılması suretiyle programın; amaçları, işleyişi, başardıkları, başaramadıkları vs. hususlarda ayrıntılı bir sunumun yapılması beklenmektedir. Yazılı program ile gerçekleşen program arasındaki farkın doyurucu biçimde analiz edildiği bir program değerlendirme çalışmasına rastlanmamıştır. Ancak, programları yasal mevzuat ve standartlar çerçevesinde tanıtan çalışmalar da bulunmaktadır. Standardı karşılayan az sayıda tezin (n=3) öne çıkan başlıca özelliği, programların nasıl işlediğine dair farklı kaynaklardan veri toplamaları yani veri kaynaklarında çeşitliliğine

gitmeleridir. Program değerlendirme tezlerinin karşılayamadığı son standart ise *iletişim ve raporlaştırmadır*. Standart, adından da anlaşılacağı üzere raporlaştırma sürecinin sağlıklı yürümesi üzerinedir. Program değerlendirme tezleri, mevzuat gereği bir komisyon denetiminden geçmektedir. Bu nedenle, göze batacak düzeyde bir çarpıtmadan bahsetmek zordur. Program değerlendirme raporlarının objektifliği ve tarafsızlığı sağladıklarını belirtebiliriz. Öte yandan, yanlılığa karşı ne gibi tedbirlerin alındığı konusunda bir bilgilendirme yoktur. Ayrıca, daha objektif bir rapor için paydaşlardan veya katılımcılardan herhangi bir destek alınmadığı anlaşılmaktadır.

Meta değerlendirme kontrol listesine yer alan dört boyuta bütüncül bakıldığında, program değerlendirme tezlerinin *uygulanabilirlik* boyutunda nispeten başarı gösterdikleri ancak buna karşın yararlık, yerindelik ve doğruluk boyutlarında standartların büyük ölçüde karşılanamadığı söylenebilir.

5. ÖNERİLER

Bu bölümde daha faydalı, anlaşılır ve işe vuruk program değerlendirme çalışmalarının yürütülebilmesi için dikkat edilmesi gereken noktalar ele alınmıştır.

1. “Problem durumları” ile “araştırmanın önemi” gibi kısımların yazımında paydaşlardan ve ilgili kurumlardan görüş alınması ve böylelikle, değerlendirme sürecinin daha sağlam bir zeminde ilerlemesi son derece önemlidir. Farklı bakış açılarından faydalanılması halinde değerlendirme bulgularının çok yönlü olarak yorumlanabilmesi de mümkün olacaktır.
2. Program değerlendirme tezlerinin sonuç ve öneri kısımlarında eğitimcilerin ve öğrencilerin hayatlarına dokunacak sade bir anlatım tarzı tercih edilmelidir. Öneriler, somut örneklerle desteklenmeli ve gerçekçi olmalıdır.
3. Program değerlendirme sürecinde iletişimin sürekliliği sağlanmalıdır. Değerlendirmenin çeşitli aşamalarında katılımcıların yanı sıra paydaşlar ve ilgili kurum yöneticileriyle bilgi alışverişinde bulunulması önemlidir. Bu sayede, farklı kişi ya da grupların değerlendirme sürecine gerçek anlamda katkı sağlayabilmesinin önü açılacaktır. İletişim kanallarının sürekli açık tutulması halinde değerlendirme sürecindeki aksaklıklara daha rahat çözüm bulunabilecektir.
4. Program değerlendirme tezlerinin paydaş odaklı hazırlanması son derece önemlidir. Program değerlendirme çalışması henüz başlamadan paydaş grupları ve bu grupların ihtiyaçları doğru biçimde saptanmalıdır. Daha sonra, paydaşların ne tür değerlendirme bulgularına ihtiyaç duydukları tespit edilerek araştırma soruları bu kapsamda gözden geçirilmelidir.
5. Program değerlendirme çalışmalarında kaynakların verimli kullanılabilmesi gerekmektedir. Program değerlendirme süreciyle ortaya çıkan bulgular; programın sonlandırılması, revize edilmesi veya aynen sürdürülmesi gibi kararların alınmasına yetecek düzeyde olmalıdır. Program değerlendirme çalışmalarının bilhassa tartışma bölümleri, programı uygulayanlar için yeni ve özgün bakış açıları sunabilmelidir.
6. Program değerlendirme tezleri, üzerinde çalışılan eğitim programını hem zayıf hem de güçlü yönleriyle ele almalıdır. Bunun yapılabilmesi, ihtiyaç analizi ile program

çıktılarının karşılaştırılmasına bağlıdır. Diğer türlü, programların yalnızca aksayan yönleri ele alınacak ve güçlü yönleri göz ardı edilecektir.

7. Program değerlendirme tezinin yazımında başlangıçta öngörülmemiş, arzu edilmeyen ve sürpriz niteliğinde olan bulgular da paylaşılmalıdır. Benzer şekilde, program değerlendirme sürecinde karşılaşılan sınırlılıklar ve bu sınırlılıkların değerlendirme sonuçlarına etkisi açık yüreklilikle tartışılmalıdır.
8. Bir eğitim programının yürütülmesi ve değerlendirilmesi esnasında; değerlendirme uzmanı, program yürütücüsü, öğrenci, program hazırlama ekibi, akademik çevre vs. gruplar arasında çıkar çatışmaları söz konusu olabilmektedir. Program değerlendirme tezlerinin yazımında bu potansiyel çıkar çatışmaları mutlaka hesaba katılmalıdır. Bu çerçevede, programı değerlendiren uzman başta olmak üzere tüm paydaşların aldığı pozisyon net biçimde tarif edilmeli, objektifliği zedeleyebilecek durumlar belirtilmelidir. Çıkar çatışmalarının etkisini en aza indirmek için farklı program değerlendirme uzmanlarından görüş alınması seçeneği de düşünülmelidir.
9. Program değerlendirme tezlerinde istatistiksel kanıtların inandırıcı ve yeterli ölçüde sunulabilmesi değerlendirme sonrası alınacak kararların isabetini de arttıracaktır. İstatistiksel testlerin uygulanmasından önce veri setinde gerekli temizlik yapılmalı, varsayımlar karşılanmalı ve bununla ilgili işlemler rapor edilmelidir. Ayrıca, istatistiksel anlamlılığın yanı sıra etki büyüklüğü değerlerinin de mutlaka paylaşılması lazımdır. Ek olarak, nitel veya nicel veri girişleri kontrol edilmeli, veri girişlerinin doğru olduğu konusunda katılımcılardan teyit alınmalıdır.
10. Program değerlendirme çalışmalarının sonuçları, daha önce benzer konularda yapılmış çalışmaların sonuçlarıyla karşılaştırılmalıdır. Ayrıca sonuçlar, literatürdeki bilgilerle harmanlanarak yazılmalıdır.
11. Program değerlendirme süreci için yeni bir ölçme aracı geliştirildiğinde ölçme aracına ilişkin güvenilirlik kanıtları paylaşılmalıdır.

KAYNAKÇA

- American Evaluation Association (2013). *Guiding principles for evaluators*. [Standards]. Erişim adresi: www.eval.org/p/cm/ld/fid=51
- Blake, A. M. (2013). *An empirical examination and meta-evaluation of the impact evaluation process*. (PhD thesis). Wayne State University, Michigan.
- Cooksy, L. J. & Caracelli, V. J. (2009). Metaevaluation in practice: Selection and application of criteria. *Journal of Multi-Disciplinary Evaluation*, 6(11), 1-15.
- Fitzpatrick, J.L, Sanders, J.R., & Worthen, B.R. (2004). *Program evaluation: Alternative approaches and practical guidelines*. Upper Saddle River, NJ: Pearson.
- Gibbs, M. (2012). *Miblsi program evaluation of participatory elementary schools from 2003-2009*. (PhD thesis). Wayne State University, Michigan.
- Joint Committee on Standards for Educational Evaluation. (2011). *The program evaluation standards*. [Standards]. Erişim adresi: www.jcsee.org/program-evaluation-standards-statements

- Kozikoğlu, İ. ve Senemoğlu, N. (2016). Eğitim programları ve öğretim alanında yapılan doktora tezlerinin içerik analizi (2009-2014). *Eğitim ve Bilim*, 40 (182), 29-41 doi: 10.15390/EB.2015.4784
- Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.
- Scriven, M. (2013). *Key evaluation checklist*. [Standards]. Erişim adresi: www.michaelscriven.info/images/KEC_3.22.2013.pdf
- Scriven, M. (2009). Meta evaluation revisited. *Journal of Multi-Disciplinary Evaluation*, 6(11), iii-viii.
- Scriven, M. (2005). *The logic and methodology of checklists*. [Article]. Erişim adresi: <http://citeseerx.ist.psu.edu/viewdoc/download>
- Shanahan, K. A. (2014). *Meta evaluation of a university teacher education assessment system*. (PhD thesis). University of Louisville, Kentucky.
- Stufflebeam, D. L. (2007). CIPP evaluation model checklist. [Standards]. Erişim adresi: https://www.wmich.edu/sites/default/files/attachments/u350/2014/cippchecklist_mar07.pdf
- Stufflebeam, D. L. (2011). Meta evaluation. *Journal of Multi-Disciplinary Evaluation*, 7(15), 99-158.
- Stufflebeam, D. L. (2012). *Program evaluations meta evaluation checklist*. [Standards]. Erişim adresi: https://wmich.edu/sites/default/files/attachments/u350/2014/program_metaeval_short.pdf
- Stufflebeam, D. L. (2001). The meta evaluation imperative. *American Journal of Evaluation*, 22, 183-209. doi: 10.1177/109821400102200204.
- Torres, E. B. (2005). *A meta evaluation of an evaluation of a second language course*. (Master thesis). Brigham Young University, Utah.
- Yüksel, İ. (2010). *Türkiye için program değerlendirme standartları oluşturma çabası*. (Doktora tezi). Anadolu Üniversitesi, Eskişehir.
- Yüksel, İ. ve Akın, Z. (2013). Öğrenci başarısını belirleme sınavının (ÖBBS) öğrenci değerlendirme standartları kapsamında değerlendirilmesi: Bir meta değerlendirme araştırması. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*, 32(1), 473-495.