

## PROGRAM DEĞERLENDİRME TASARILARI

Öğr. Göç. Tansu KARAGÜL

Tasarı, değerlendirme sürecinde kimlerden ne zaman veri toplanacağını gösteren plan olarak ele alınabilir. Tasarı kullanımının en önemli nedeni, değerlendirmede doğru kişilerin doğru zamanlarda yer almasını sağlamaktır. Tasarı yoluyla elde edilen karşılaştırmalı bilgi ile program hakkında ulaşılan kararlar daha objektif olacaktır. Tasarılar, genellikle sonuç değerlendirme yaklaşımı ile ele alınırlar. Sonuç değerlendirmeler olabildiğince deneysel olarak yapılmalıdır.

Deneysel tasarıların, kontrol gruplarının kullanılmadığı sonuç değerlendirmelerde programın başarı veya başarısızlığı hakkında herhangi bir yargıya ulaşmak güçtür.

Program değerlendirilmede biçimlendirici tasarılar kullanılması, değerlendirme ve programla ilgili kişilere programın veya programın bazı öğelerinin yeterliliği konusunda bilgi verir. Değerlendirme tasarısını özel eğitim ve zorunlu eğitim programlarında uygularken, kontrol grubuna özel program uygulamak mümkün değildir. Bu durumda ancak okulda iki farklı program uygulanıyorsa kontrol grubu kullanılabilir. Bunun dışında, eşolmayan kontrol grubu tasarımı kullanmak, küçük çaplı ön-son test uygulamaları yapmak; birbirinin tersi programları aile ve öğrencilerin hoşnutsuzluğunu göz önüne alıp karşılaştırmak, programın sonuçlarını amaçlar ile yerel kriterlerle karşılaştırmak, veya teorilere göre değerlendirmek gibi yollar izlenebilir.

Değerlendirmede, *kimin ne zaman ölçüleceğini* gösteren plan olarak ele alınabilecek, tasarımın, iki temel elementi vardır. Bunlar;

1- Gruplar,

2- Ölçüm zamanları'dır.

Grup kavramı çerçevesinde *kontrol grubu* ve *deneysel grup*'tan söz etmek mümkündür:

\* Gibbon, F. T. C.: Morris, L. L. How to Design a Program Evaluation. London: Sage Pub., 1978. adlı eserden derlenmiştir.

Değerlendirme sürecini planlarken, genellikle bir deney grubunun ölçülmesi planlanır. Bu grup, değerlendirilecek programı alan bir öğrenci, sınıf veya okullardan oluşabilir. Yalnızca deneysel grubun ölçüldüğü çalışmalarda, sonuçların yorumu çok güç ve pek inandırıcı değildir. Bu nedenle, çalışmalarda karşılaştırma yapmak üzere kontrol grubu kullanılması önemlidir.

Kontrol grubu deneysel grubun öğrencilerine olabildiğince benzeyen, deney grubuyla aynı zamanlarda ölçülen fakat, deneysel programı almayıp başka bir program alan gruptur:

Kontrol grubu, yansız seçim yolu ile eşlendirilmiş grup ve yansız seçim uygulanmamış eşit olmayan grup olmak üzere iki şekilde oluşturulabilir.

Grupları yansız olarak seçmek, sonuçları etkileyebilecek çok çeşitli faktörlerden arındırmak, elde edilen sonuçları deney ve kontrol grubuna verilen farklı uygulamalara bağlamak bakımından verimli bir uygulamadır. Kontrol grubunun yansız olarak seçilmesi, hem sonuç hem de biçimsel değerlendirmelerde en iyi sonuçlara ulaşılmasını sağlar. Ancak, yansız seçim yapılmasını engelleyen durumlar olabilir.

Eş olmayan kontrol grubu, deneysel gruba benzer olduğu için seçilmiş bir gruptur. Yansız olarak seçilmemiştir. Bu gruplarla ilgili şu üç nokta gözönünde bulundurulmalıdır:

1- Deneysel grup özel bir yöntemle seçilmişse, kontrol grubu da bu yöntemle seçilmelidir.

2- Kontrol grubuna deney grubunun aldığı tüm testler verilmelidir. Gruplar testleri aynı zamanda almalıdırlar.

3- Kontrol ve deneysel grup arasındaki farklılıkları açıklayacak veriler toplanmalıdır. Bulguların değeri, deney ve kontrol gruplarının aldıkları program dışında olabildiğince benzer olduklarının gösterilmesine bağlıdır.

Deney grubu deneysel programa katılırken, kontrol grubu için sırasıyla şu seçenekler söz konusudur.

1- Deneysel programa en yakın programı almak,

2- Eski programı veya daha önce denenmemiş bir programı almak,

3- Program almamak. Bu durumda, deneysel grubun aldığı programı karşılaştıracak bir program yoktur. Bu seçenek, deneysel grubun aldığı programın, hiç program almamasından daha iyi veya kötü olduğunu gösterecektir.

**Ölçümler yapıldıkları zamana göre:** Ön testler, son testler, ara testler, hatırlatma testleri ve zaman aralıklı testler olarak ele alınabilir.

1- **Ön Testler:** Ön test, program veya deneysel uygulama başlamadan önce verilen testtir. Ön testler; *öğrenci seçmek, programı planlarken yapılan varsayımları kontrol etmek, programın kazandırdıklarını kontrol etmek ve programın etkilerini belirlemek üzere kullanılabilir.*

a. *Öğrenci seçmek amacıyla kullanılan öntestler:* Test yüksek-düşük puanlara göre programa öğrenci seçmek amacıyla kullanılıyorsa, ikinci testi öntest olarak vermek daha iyidir. İkinci testte değerler daha normal bir dağılım gösterecektir.

b. *Grupların karşılaştırabilirliğini kontrol eden öntestler:* İki grubun son test sonuçları karşılaştırılacaksa, program başlamadan önce grupların benzer olup, olmadıklarının belirlenmesi gerekmektedir. Gruplar yansız olarak seçilmiş ve öğrenci sayısı çoksa bu işleme gerek yoktur. Grupların küçük olduğu, yansız seçim yapılmadığı durumlarda bu işleme gerek duyulur.

c. *Programın kazandırdıklarını kontrol eden öntestler:* Başlangıç noktası belirlenmeden ilerlemeyi belirlemek mümkün olmaz. Bu kapsamdaki öntestler başlangıç düzeyini belirlemeye dönüktür.

d. *Programı test etmek, duyarlılığını arttırmak üzere kullanılan öntestler:* Bazı durumlarda programın sonuçlarda büyük değişimler yaratamadığı düşünülebilir. Son teste çok benzer bir öntest uygulamak, tasarımın gücünü arttıracaktır.

Ön test kullanımının sontest sonuçlarını etkileyebileceği durumlarda öntest kullanılmamalıdır. Bu durum, tutumların ölçülmesi halinde söz konusu olabilir. Programdan önce kişinin öntesti cevaplamaya zorlanması, tutumlarını olumsuz etkileyebilir. Grubun o konuda hiç bir bilgi sahibi olmadığı bilindiği durumlarda da, öntest uygulanması gereksizdir.

Program devam etmekte, ve daha önce öntest uygulanmamış ise arada test uygulayarak, deney ve kontrol grupları arasındaki farklılığın programın başında da var olduğu varsayılabilir. Öntest büyük bir zaman dilimini kapsıyor ve maddi olarak büyük imkanlar gerektiriyorsa, kullanılmayabilir.

Öntest bir *tutum testi, başarı testi veya yetenek testi* olabilir. Tutum testi, genellikle program tutum değişikliğini hedefliyorsa kullanılır.

Öntest kullanımı son testi etkileyebileceği için, öntestin grubun yarısına uygulanması, son testinde öntesti almamış gruba uygulanması tavsiye edilebilir.

Son test başarı testi ise, başarıyı ölçen bir öntest kullanımı anlamlı olacaktır. Başarıyı ölçen bir öntest, grupların eş olup olmadıklarını göstermek amacıyla kullanılabilir.

Eş olmayan kontrol gruplarının kullanıldığı ve grubun küçük olduğu (15 kişiden az) durumlarda grupların başarı bakımından benzerlikleri kontrol edilmelidir.

Sontest başarı testi ise, öntest genellikle yetenek testidir. Bu sayede son test sonuçlarını öğrencinin yetenek düzeyine göre yorumlamak mümkündür olur.

2- **Son Testler:** Bu testler genellikle program veya deneysel uygulamanın bitiminde kullanılır. Sonuçları kısmen programa bağlıdır. Bu testler, programın sonuçlarını ölçtükleri için mutlaka uygulanırlar. Son testin ne zaman uygulanacağı genellikle önceden bellidir. Bütün bir yıl süren programın sonuna dek beklenildiğinde öğrenciler yorgun olabileceği için program hakkında anlamlı bilgi alınamayabilir. Bu durumda, testi daha önce uygulamak daha anlamlı olur.

3- **Ara Testler:** Program devam ederken bazı ölçümler alınabilir. Ara testler, biçimlendirici ve sonuç değerlendirmede kontrol grubunun olmadığı durumlarda, programın etkisinin göstergesi olarak kullanılabilir.

4- **Hatırda Tutmayı Gözleyen Testler:** Bir program yıl sonu testinde diğerine göre çok daha etkili gözükülebilir. Ancak, yaz tatili dönüşü yapılan testte aynı sonuç gözlenmeyebilir. Bu nedenle eğer bir program diğer programdan daha etkiliyse, sonuçlar Eylül ayında sontest verilerek kontrol edilmelidir hala aynı programın sonuçları daha iyiye, programın daha etkili olduğunu düşünülebilir.

5- **Zaman Aralıklı Testler:** Bu testler, program uygulandıktan sonra programın uzun süreli etkilerinin kalıcılığını gözlemek üzere kullanılır. Programdan önce ve sonra eşit aralıklı olarak verilen testler zaman aralıklı testlerdir.

Tasarımın iki temel elementi (kim-ne zaman) ile ilgili sorular çeşitli şekillerde cevaplandırılabilir. Kim ölçülecek sorusuna verilecek cevaplar şunlar olabilir: Yalnız deneysel grup, deneysel grup ve eş kontrol grubu,

deneysel grup ve eş olmayan kontrol grubu. Ne zaman ölçülecek sorusuna verilecek cevaplar şunlar olabilir: Öntest-sontest, sontest, zaman aralıklı testler.

Tasarımın iki temel elementi (kim-ne zaman) çeşitli şekillerde birleşerek farklı değerlendirme tasarılarını oluştururlar.

Tasarı 1: Eş kontrol grubu, ön-son test

Tasarı 2: Eş kontrol grubu, son test,

Tasarı 3: Eş olmayan kontrol grubu,

Tasarı 4: Tek grup, zaman aralıklı testler,

Tasarı 5: Eş olmayan kontrol grubu, zaman aralıklı testler,

Tasarı 6: Önce ve Sonra Tasarımı.

Uygulanacak tasarımın seçilmesi önemli bir basamaktır, ancak tasarımın doğru uygulanması daha önemlidir. Hangi tasarım seçilirse seçilsin bazı sorunlarla karşılaşmak olasıdır. Hastalık, sınıf değiştirme, taşınma gibi sebeplerden dolayı oluşan veri kaybı program sonuçlarını etkileyebilir. Bir grupta oluşan öğretmen değişikliği gibi bir değişim verileri etkileyebilir. Deneysel uygulamanın diğer etkilerden yeterince arındırılmaması da sonuçları etkileyebilir.

Bütün bunlar sonuçları yorumlarken göz önüne alınmalıdır.

### KONTROL GRUBU TASARILARI - TASARI 1-2-3

#### Tasarı 1:

Programa alınacak öğrenci, sınıf/ okulun belirlenmesi	→	Gruplara öntest uygulanması	→	Kontrol ve deney grubunun yansız olarak belirlenmesi	→	Kontrol grubunun alternatif programı; deney grubunun deneysel programı olması
(1)		(2)		(3)		(4)

(5) → Aynı koşullar altında her iki gruba son test uygulanması.

	ZAMAN	
	1 (Öntest)	2 (Sontest)
Deney grubu	R 0	X 0
Kontrol grubu	R 0	0

R: Yansız seçim.

0 Deneysel uygulama

X Değerlendirilecek program

Tasarı 1 ile ilgili yargıda bulunabilmek üzere tasarı hakkında yürütme ile ilgili veriler ve çıktı verilerinin rapor edilmesi gerekir. Tasarının ne kadar başarılı uygulandığı programın uygulanması, kontrol grubunun deneysel uygulamadan soyutlanması, programdaki değişimler, ve deneysel grup ve kontrol grubundaki veri kayıpları ile ilgilidir.

Tasarı 1'de veriler tablo ve grafikler kullanılarak sunulabilir. Eğer öntest ve sontestler aynı test ise sonuçları grafik üzerinde göstermek uygun olacaktır. Eğer ayrı testlerse, yalnızca son test sonuçlarının gösterilmesi daha uygundur. Grafik kullanımı kolay anlaşılır ve yorumlanabilir olması nedeniyle tavsiye edilmektedir.

Bu tasarıda öntestten önce yansız seçilmiş büyük bir grup varsa, grupların ortalama test değerlerinin anlamlı derecede farklı olmaması beklenilir. Bu, grupların birbirine benzer olduğunun göstergesidir.

Öntest sonuçları arasındaki fark anlamlıysa grupları yeniden oluşturmak, gruplarda aşırı yüksek-düşük puanların olup, olmadığını kontrol etmek, varsa bunları değerlendirme dışı tutmak gerekebilir.

Bir başka seçenekte bu tasarıyı, eş olmayan kontrol grubu tasarımı olarak ele almaktır.

Tasarı 1'de deney grubunun sontest ortalaması kontrol grubununkinden yüksek ve t-testi sonucu anlamlıysa "Deneysel program, kontrol grubundan anlamlı derecede yüksek sontest değerleri yarattı" ifadesini kullanmak mümkündür.

Tasarı 1'in sonuçlarını yorumlarken bir grubun sontest ortalamasının diğer grubun sontest ortalamasından anlamlı derecede farklı olup, olmadığı belirlenmelidir. Her iki grubun öntest ve sontestlerden kazandıkları arasındaki farklılığa bakmak yanıltıcı olabilir. Kazançları karşılaştırmak öntest sonuçları yüksek olan grubun aleyhine olabilir. Bu nedenle, sontest sonuçlarının karşılaştırılması gerekir.

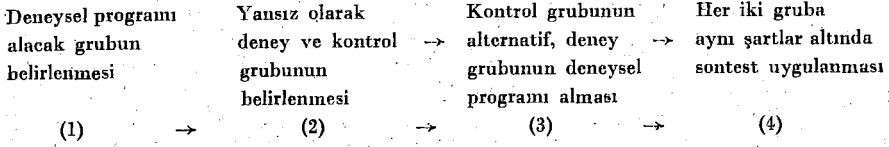
Deney grubunun sontest ortalaması öntest grubununkinden yüksek, fakat anlamlı derecede farklı değilse, ANOVA veya ANCOVA gibi çeşitler istatistiksel testler uygulanabilir.

Deney grubunun sontest ortalaması kontrol grubununkinden anlamlı derecede düşükse, bu deneysel programın sonuç ölçümleri üzerinde olumsuz etki yaptığı anlamına gelir.

Tasarı 1 ile ilgili sonuçları rapor ederken, tasarının yeterince uygulanıp, uygulanmadığı belirtilmeli, eksik veriler varsa değişiklikler not

edilmelidir. Sonuçlar, hem istatistiksel hem de eğitimsel anlamı bakımından tartışılmalıdır.

### Tasarı 2:



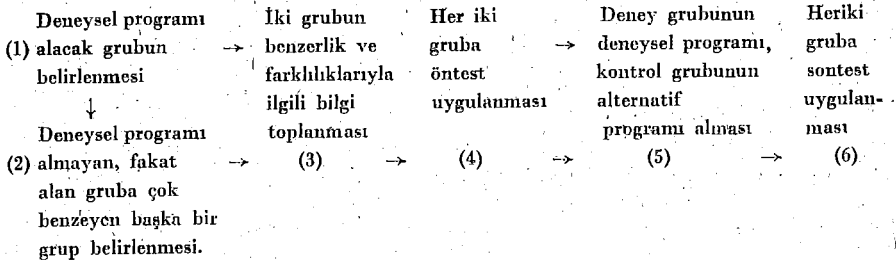
(sontest)		
Deney grubu	R	X
Kontrol grubu	R	0

Tasarı 2'nin uygulanması sonunda tasarımın yürütülmesi ile ilgili deney ve kontrol gruplarının son ölçümleri ile ilgili bilgi sunulması gereklidir.

Sonuçları rapor ederken daha önce belirtilen noktalara değinilmesi (uygulama, beklenmeyen değişimler vb.) gerekir. Öntest sonuçları olmadığı için yansız yapılan seçimin ayrıntılı olarak açıklanması gerekir.

Deney ve kontrol gruplarının son test sonuçları arasındaki farkın anlamlılığının güvenilirlik sınırları içinde sunulması gerekir. İstatistiksel olarak anlamlı bir farklılık varsa, bunun eğitimsel anlamı tartışılmalıdır.

### Tasarı 3:



ZAMAN		
	(1) (Öntest)	(2) (Sontest)
Deney grubu	0	X
Kontrol grubu	0	0

Bu tasarı, tasarı 1'e benzerlik gösterir. Aradaki farklılık grupların yansız olarak seçilmemiş olmalarından kaynaklanır. Bu tasarı 3'ün değerini Tasarı 1'e göre azaltan bir faktördür. Ancak deney ve kontrol gruplarının başlangıçtaki benzerliklerini gösteren bulgulara ulaşarak Tasarı 3'ün değerini arttırmak mümkün olabilir. Bu tasarıdaki öntest, programın değiştirmeye çalıştığı özellikler bakımından grupların karşılaştırılabilir olduğunu belirlemeye yöneliktir.

Deney ve kontrol gruplarının öntest değerleri arasında farklılık olmaması, tasarımın değerini artırır. Grupların başlangıçta benzer oldukları ortaya konulmuş olur. Öntest sonuçları birbirinden anlamlı derecede farklıysa, çeşitli yollar izlenebilir. En kolay yol, bu sorundan söz edip, başka bir şey yapmamaktır. Ayrıca, sonuçlar grafiğe aktarılıp eğilim incelenebilir, aynı konunun yansız seçilmiş bir grupta tekrarlanması tavsiye edilebilir.

Sontest sonuçları anlamlı derecede farklı, öntest değerleri anlamlı derecede farklı değilse, programın iyi sonuçlar ürettiği söylenebilir. Ancak, bulguların eğitimsel anlamlılığına karar verilmesi gerekir.

Öntest ve sontest sonuçları anlamlı derecede farklı değilse, programın farklılık yaratmadığı konusunda bir yargıya ulaşılabilir. Bunu ANCOVA, ANOVA gibi istatistiksel tekniklerle sınamak mümkündür.

Tasarı 3'te de tasarımın uygulanması ve son ölçümlerin sonuçları ile ilgili bilgilerin sunulması gerekir.

#### ZAMAN ARALIKLI TASARILAR - TASARI 4-5

##### Tasarı 4:

Tekrar, tekrar kullanılacak ölçülebilen bir özellik seçilmesi.	Deney grubunun nasıl oluşturulacağına karar verilmesi. Grubun çeşitli kereler ölçülen aynı grup mu, farklı öğrenicilerden oluşan ayrı bir grup mu olacağının belirlenmesi.	Deneysel programın uygulanmasından önce düzenli aralıklarla en az üç ölçüm alınması.	Deneysel programın yürütülmesinin kontrol edilmesi.	Deneysel programın uygulanması bittikten sonra zaman aralıklı ölçümler alınması.
(1)	→ (2)	→ (3)	→ (4)	→ (5)

ZAMAN						
	1	2	3	4	5	6
Deney Grubu	0	0	0	X	0	0



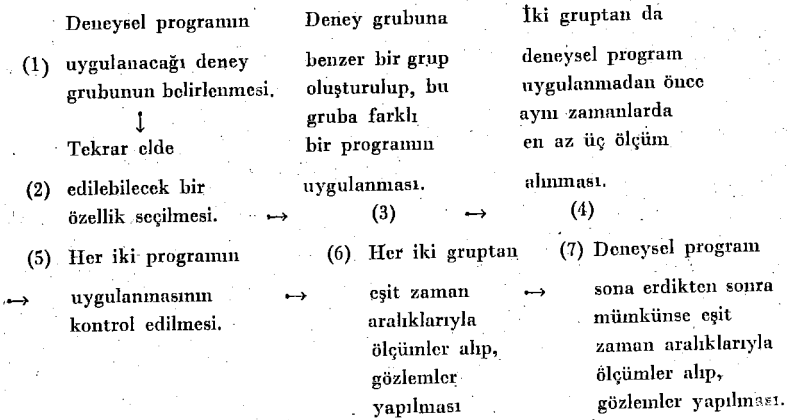
Tasarı istenilen şekilde uygulanmış ve sonuçları bir yargıya ulaşabilmesini sağlıyorsa bu durumda verileri değerlendirip, istatistiksel bir teknik kullanmak yerine, sonuçlar zaman aralıklı grafik kullanarak açıklanabilir. Verilerin zaman aralıklı grafikler kullanarak ifade edilmesi daha kolay yorumlanabilmelerini sağlar.

Zaman aralıklı verileri yorumlarken, olumlu gelişmeler sadece programa bağlanabilir. Ancak, aşağıdaki noktalar gözardı edilmemelidir:

- Program öncesi ve sonrasında program dışında bir değişiklik oldu mu?
- Grubun oluşumu aynı kaldı mı?
- Ölçme tekniği aynı kaldı mı?
- Grubu etkileyecek başka olaylar oldu mu?
- Bulunan doğruların kullanımı teknik olarak uygun muydu?
- Sonuçlar belli bir yargıya ulaşılacak kadar dengeli miydi?

Bazı durumlarda programın uygulandığı grubu analiz bakımından kendi içinde alt gruplara ayırmak yararlı olabilir. (Kızlar, erkekler gibi).

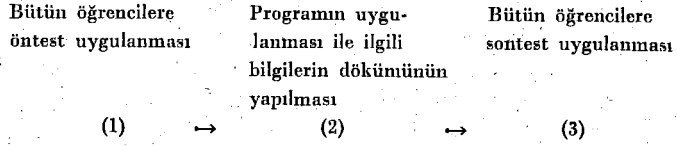
#### Tasarı 5:



	ZAMAN					
	1	2	3	4	5	6
Deney Grubu	0	0	0	0	0	0
Eş Olmayan Kontrol Grubu	0	0	0	0	0	0

Tasarı 5 tasarı 4'e benzerlik göstermektedir. Tasarı 5'de eş olmayan kontrol grubu ölçümleri yer alır. Gerçek bir kontrol grubu oluşturulabilirse sonuçlar çok daha gerçekçi olacaktır.

### Tasarı 6:



ZAMAN			
	1	2	
Deney Grubu	0	X	0

Tasarı 6 gerek biçimsel gerekse sonuç ve değerlendirme açısından en az uygun olanıdır. Bu tasarıda karşılaştırma yapacak bir kontrol grubu yoktur. Bu tasarı "sonuçlar ne kadar iyi", "bu sonuçları program mı doğurdu" sorularına cevap vermede yetersizdir. Tasarı 6 kullanılacak en son seçenek olmalıdır.

Bu tasarı uygulanış biçiminden dolayı bazı ciddi sorunlar doğurur. Kontrol grubunun olmaması elde edilebilecek olan küçük fakat önemli sonuçların görülmesini engeller. Bu tasarı, yalnızca uygulamanın tasarı sırasında planlanan amaçlara ne kadar yaklaştığını gösterir.

Bu tür değerlendirmeler programın gerçek yeterliliğini göstermekte başarılı sayılmazlar.