

Güvenirlilik Analizleri ve SPSS UYGULAMALARI

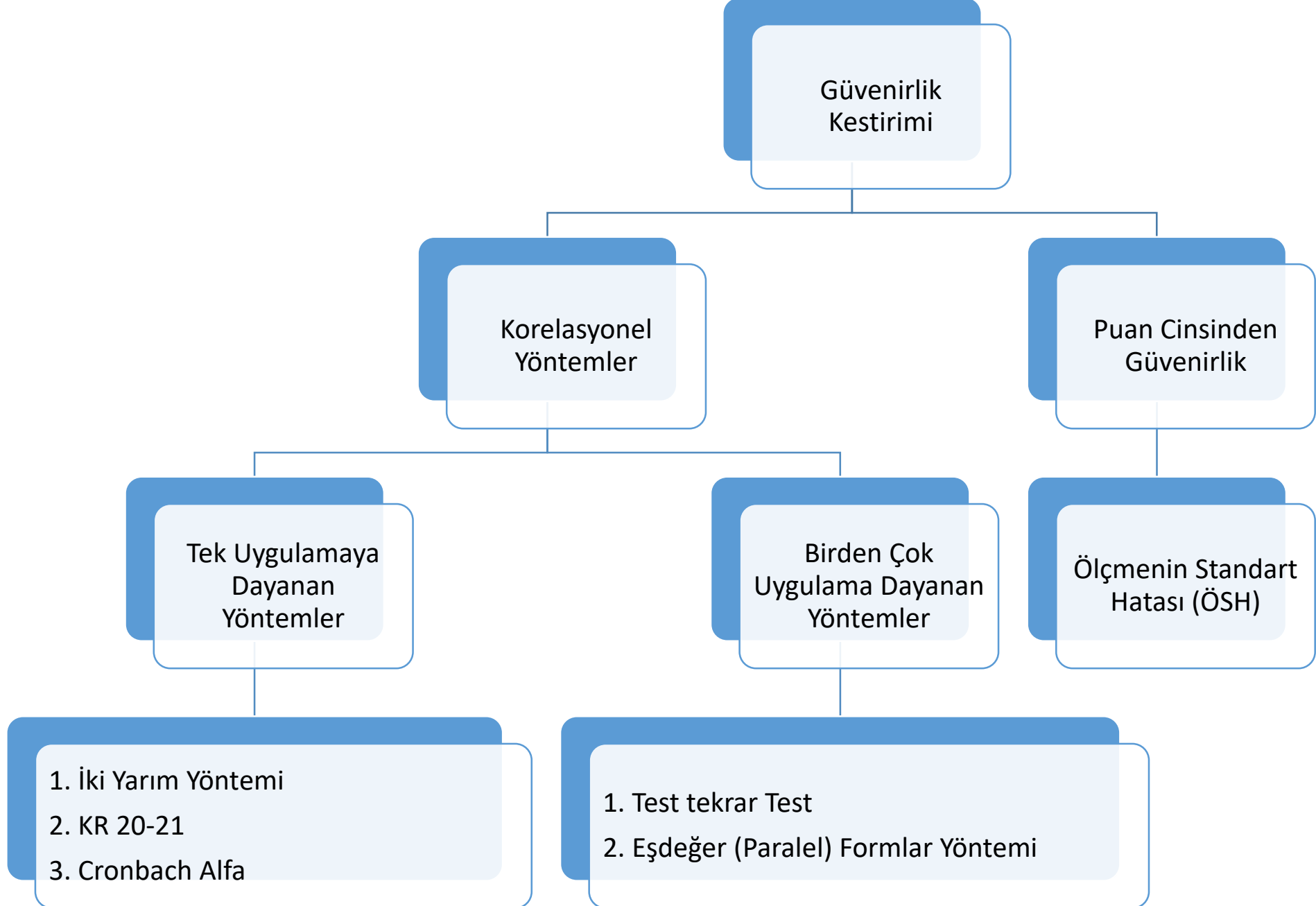
Doç. Dr. Seher Yalçın

GÜVENİRLİK

- Aynı bireyler üzerinde yapılan bir niteliğe ait ölçmelerin benzer şartlar altında tekrar edilebilirliğidir (Köse, 2012).
- Ölçme sonuçlarının tesadüfi hatalardan arınlık derecesi
- Ölçme aracının ölçülmek istenen özelliğe olan duyarlılık derecesi (Atılgan ve diğ., 2011).

Güvenirlilik

- Güvenirlilik kavramı, geçerlik kavramında olduğu gibi ölçme araçları için değil, o ölçme araçlarından elde edilen puanlarla veya sonuçlarla ilgili bir kavramdır (Reynolds, Livingston ve Wilson, 2006; Akt. Köse, 2012).



Tek Uygulamaya Dayanan Yöntemler

İki Yarıya Bölme Yöntemi

- Tesadüfi, Tek-çift, konu alanı vb.
- Tutarlılık
- İlişkinin düşük olması,
 - Testin gelişigüzel yanıtlanmasından
 - İki yarının farklı davranışları ölçmesinden kaynaklanabilir.

Spearman Brown (Düzeltilme Formülü)

r_{11} : İki alt teste ait güvenilirlik katsayısı

r_X : Testin tümüne ait güvenilirlik katsayısı

- Testin içindeki alt test sayısı 2'den fazla ve k tane ise

$$r_X = \frac{k(r_{11})}{1 + (k - 1)r_{11}}$$

Tek Uygulamaya Dayanan Yöntemler

• KR 20-21

- Testteki her maddenin aynı değişkeni ölçtüğü sayılıtısına dayanır
- Testteki tüm maddeler arasındaki tutarlılığı (Maddelerin birbirleriyle ve testin tümüyle) (İç tutarlılık)

$$KR_{20} = \frac{K}{K-1} \left[1 - \frac{\sum pq}{S_x^2} \right]$$

K = Testin soru sayısı

p = Madde güçlüğü

$q = 1 - p$

S_x^2 = Testin varyansı

Tek Uygulamaya Dayanan Yöntemler

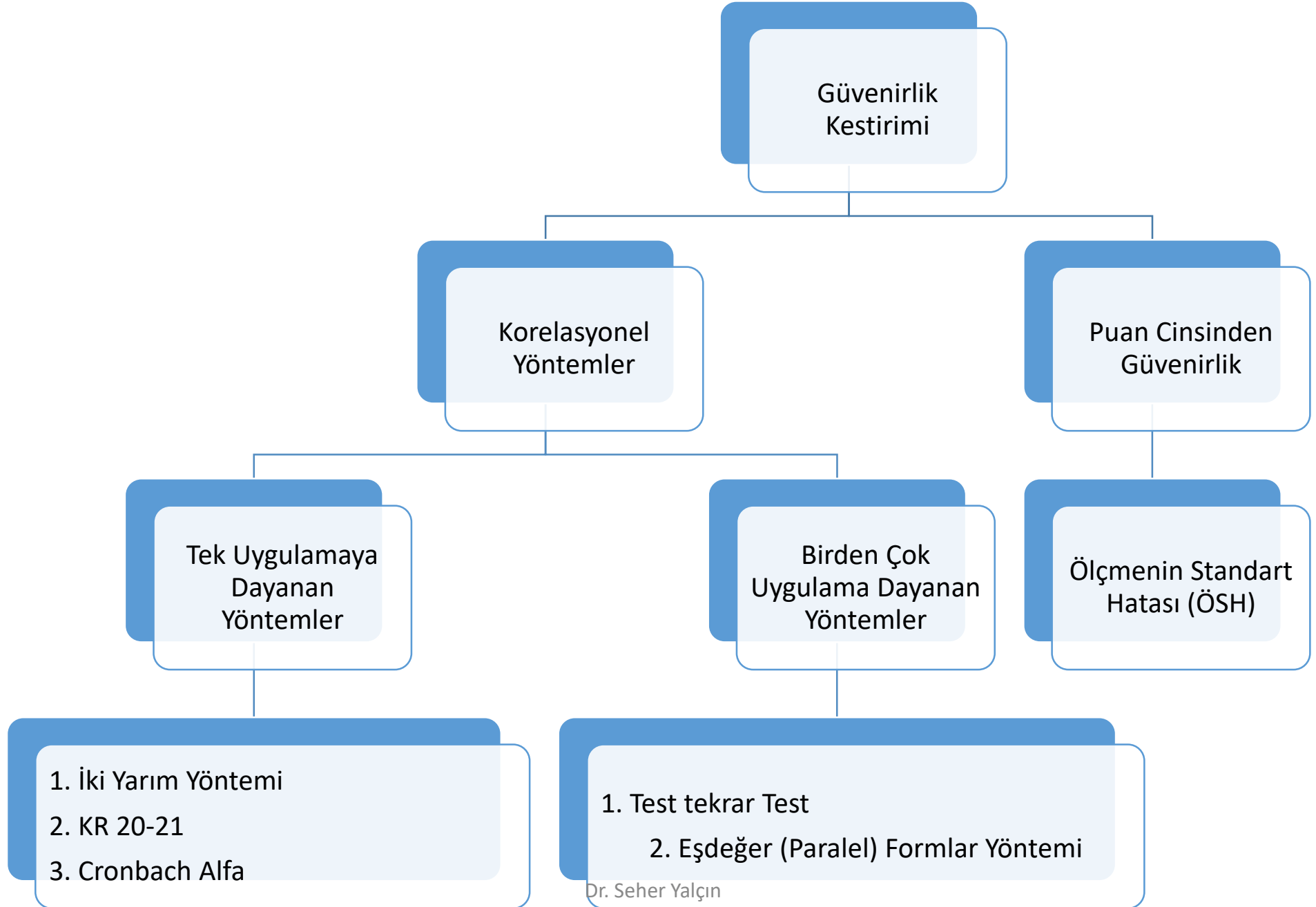
Cronbach-Alfa

Genellikle ağırlıklı puanlama veya dereceleme yöntemiyle puanlama uygulandığında

İç tutarlılık

$$\alpha = \frac{K}{K-1} \left[1 - \frac{\sum s_j^2}{S_x^2} \right]$$

K = Testin soru sayısı
 S_j^2 = j maddesinin varyansı
 S_x^2 = Testin varyansı



Birden fazla Uygulamaya Dayanan Yöntemler

- **Test Tekrar Test Yöntemi** (Tekin, 1996)
- Kararlılık (Uzun süreli), grup içindeki sırada tutarlılık
- Yüksek güvenirlik katsayısı, testin kararlı olduğunu ve ölçme aracının uygulamadan kaynaklı tesadüfi hatalardan da arınık olduğunu gösterir.

Birden fazla Uygulamaya Dayanan Yöntemler

- **Test Tekrar Test Yöntemi** (Tekin, 1996)
- Ölçülen nitelik sürekli değişkenlik gösteriyorsa, uygun değil
- Süre hatırlanmayacak kadar uzun, öğrenme olmayacağı kadar kısa olmalı

Birden fazla Uygulamaya Dayanan Yöntemler

Paralel Testler (Eşdeğer Formlar) Yöntemi

- Ölçülen davranışlar, soru sayısı, ortalama vb. eşit
- Tutarlılık (Kısa süreli)
- Yüksek güvenirlik katsayısı,
 - iki eşdeğer testten elde edilen puanların tutarlı olduğunu,
 - iki testin aynı davranışı ölçtüğünü gösterir.

Güvenirliğin Puan Olarak Kestirimi: Ölçmenin Standart Hatası ve Testlerde Kullanımı (Köse, 2012)

- Güvenirlik, hatasızlığın ölçüsü ise hatanın ölçüsü de ölçmenin standart hatasıdır. Ölçmenin standart hatası;

$$\text{Standart Hata} = ss. \sqrt{1 - \text{Güvenirlik}}$$

Güvenirlilik Katsayısını Etkileyen Koşullar (Baykul, 2000;

Köse, 2012; Tekin, 1996)

- **Puanlamanın Nesnelliği**
- **Puanlayıcı Güvenirliliği**
- **Test Edilen Grubun Değişkenliği (Homejen- heterojen olması)**
- **Testin uzunluğu**
- **Testin Güçlük Düzeyi**

Karşılaştırma

- Geçerlik tüm hata türlerinden etkilenirken, güvenirlik sadece **TESADÜFİ HATA**lardan etkilenir.
- Güvenirlik 0 ile 1 aralığında değer alırken
- Geçerlik -1 ile +1 aralığında değer alır

Karşılaştırma

- Bir ölçme aracının güvenilir olması geçerli olduğu anlamına gelmez.
- Bir test geçerli olmasa da güvenilir olabilir. Yani güvenilirlik geçerlik için gerekli fakat yeterli değildir.

Kullanışlılık

- Ekonomiklik (zaman, emek, para)
- Uygulanabilirlik (formatı, anlaşılabilirliği)
- Puanlanabilirlik (güvenilir, kolay)

Güvenirlilik Analizi

SPSS Statistics Data Editor

Transform Analyze Direct Marketing Graphs Utilities Add-ons Window

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests

ş meslekideney açık
im

ş	meslekideney im	açık
1,00	24,00	6,00
16,00	3,00	6,00
16,00	9,00	4,00
12,00	4,00	7,00
1,00	40,00	2,00
4,00	5,00	11,00
16,00	17,00	3,00
11,00	3,00	2,00
16,00	15,00	3,00

Reliability Analysis...
Multidimensional Unfolding (I

Güvenirlilik Analizi- Cronbach Alfa Yöntemi

The image shows the SPSS Reliability Analysis dialog box and its Statistics sub-dialog box. The main dialog box has the 'Model' dropdown set to 'Alpha' and the 'Items' list containing m1 through m9. The Statistics sub-dialog box has the 'Scale if item deleted' checkbox checked. The 'Model' dropdown in the main dialog is circled in red, and the 'Scale if item deleted' checkbox in the Statistics sub-dialog is also circled in red.

Reliability Analysis: Statistics

Descriptives for

- Item
- Scale
- Scale if item deleted

Inter-Item

- Correlations
- Covariances

Summaries

- Means
- Variances
- Covariances
- Correlations

ANOVA Table

- None
- F test
- Friedman chi-square
- Cochran chi-square

Hotelling's T-square

- Hotelling's T-square
- Tukey's test of additivity

Intraclass correlation coefficient

Model: Two-Way Mixed Type: Consistency

Confidence interval: 95 % Test value: 0

Continue Cancel Help

Model: Alpha

Scale label:

OK Paste Reset Cancel Help

6,00	10,00	8,00	5,00	6,00	3	9,00
8,00	10,00	5,00	9,00	4,00	3	10,00

Güvenirlilik Analizi- Cronbach Alfa Yöntemi- Çıktılar

Reliability Statistics	
Cronbach's Alpha	N of Items
,761	10

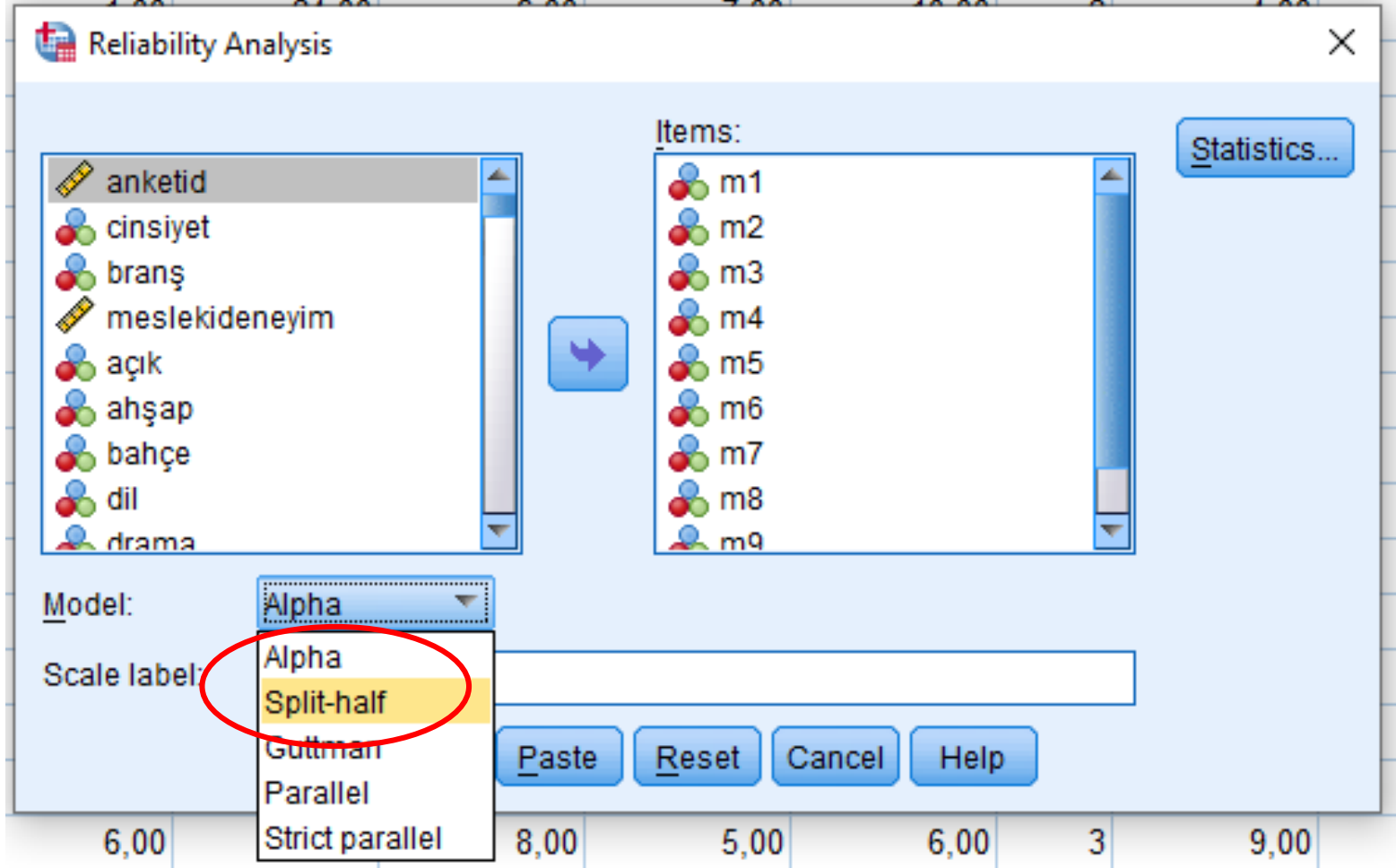
Çıktılarda, ilk tabloda Cronbach alfa değerinin .761 olduğu, 10 maddelik bir ölçek için hesaplandığı görülmektedir.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
m1	37,3837	9,922	,653	,717
m2	37,3488	10,170	,583	,726
m3	37,5698	9,475	,663	,709
m4	37,4128	9,928	,605	,721
m5	37,6163	9,489	,653	,710
m6	37,4360	10,002	,560	,725
m7	40,0814	13,795	-,364	,873
m8	37,8372	9,599	,435	,741
m9	37,8779	9,254	,574	,718
m10	37,5930	9,845	,609	,719

İkinci tabloda ise, eğer ilgili maddeler sıra ile ölçekten çıkarılırsa ölçek ortalaması, varyansı, düzeltilmiş madde-toplam korelasyonu ve Cronbach alfa değerinin ne olacağı görülmektedir.

Güvenirlilik Analizi – İki Yarıya Bölme Yöntemi



Güvenirlilik Analizi – İki Yarıya Bölme Yöntemi

Reliability Statistics

Cronbach's Alpha	Part 1	Value	,848
		N of Items	5 ^a
	Part 2	Value	,316
		N of Items	5 ^b
Total N of Items		10	
Correlation Between Forms			,593
Spearman-Brown Coefficient	Equal Length		,744
	Unequal Length		,744
Guttman Split-Half Coefficient			,738

a. The items are: m1, m2, m3, m4, m5.

b. The items are: m6, m7, m8, m9, m10.

İlk tabloda, 10 maddeyi tek ve çiftler olarak ikiye böldüğünü ve bu iki yarı için ayrı ayrı Cronbach Alfa katsayısı hesapladığı görülmektedir.

Ayrıca, Spearman-Brown katsayısı da tabloda yer almaktadır.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
m1	37,3837	9,922	,653	,717
m2	37,3488	10,170	,583	,726
m3	37,5698	9,475	,663	,709
m4	37,4128	9,928	,605	,721
m5	37,6163	9,489	,653	,710
m6	37,4360	10,002	,560	,725
m7	40,0814	13,795	-,364	,873
m8	37,8372	9,599	,435	,741
m9	37,8779	9,254	,574	,718
m10	37,5930	9,845	,609	,719

Madde toplam puan ve alt ölçek korelasyonları

SPSS Statistics Data Editor

Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation

	f5ortlm	toplaml	var	var	var	var	var	var
	4,00							
	3,50							
	3,00							
		145,0						
	5,00	142,0						
	4,50	144,0						
	4,00							
	3,50							
	4,50							
	3,00	135,0						
	4,00							
	3,00							
	5,00	140,0						
	4,00							

Bivariate Correlations

Variables:

- Zscore(m38) [ZS...
- Zscore(m39) [ZS...
- Zscore(m40) [ZS...
- Zscore(m41) [ZS...
- f1
- f2
- f3
- f4
- f5
- m6
- m7
- m8
- toplaml
- f1ortlm
- f2ortlm
- f3ortlm
- f4ortlm
- f5ortlm

Correlation Coefficients

Pearson Kendall's tau-b Spearman

Test of Significance

Two-tailed One-tailed

Flag significant correlations

OK Paste Reset Cancel Help

Madde toplam puan ve alt ölçek korelasyonları

Correlations

	m1	m2	m3	m4	m5	m6	m7	m8	toplam	f1 ortlm	f2ortlm	f3ortlm	f4ortlm	f5ortlm
m1	1	,727**	,584**	,472**	,427**	,454**	-,342**	,412**	,673**	,322**	,357**	,305**	,220**	,196**
Sig. (2-tailed)		,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	,003	,008
N	183	183	181	183	183	178	181	183	156	183	183	183	183	183
m2	,727**	1	,537**	,387**	,397**	,417**	-,217**	,363**	,574**	,309**	,277**	,317**	,203**	,187**
Sig. (2-tailed)	,000		,000	,000	,000	,000	,003	,000	,000	,000	,000	,000	,006	,011
N	183	184	182	184	184	179	182	184	156	184	184	184	184	184
m3	,584**	,537**	1	,545**	,583**	,367**	-,270**	,375**	,673**	,337**	,385**	,242**	,220**	,194**
Sig. (2-tailed)	,000	,000		,000	,000	,000	,000	,000	,000	,000	,000	,001	,003	,009
N	181	182	182	182	182	177	180	182	156	182	182	182	182	182
m4	,472**	,387**	,545**	1	,557**	,545**	-,255**	,279**	,583**	,226**	,206**	,289**	,172**	,086
Sig. (2-tailed)	,000	,000	,000		,000	,000	,001	,000	,000	,002	,005	,000	,020	,243
N	183	184	182	184	184	179	182	184	156	184	184	184	184	184
m5	,427**	,397**	,583**	,557**	1	,470**	-,194**	,362**	,632**	,273**	,396**	,210**	,226**	,140
Sig. (2-tailed)	,000	,000	,000	,000		,000	,009	,000	,000	,000	,000	,004	,002	,059
N	183	184	182	184	184	179	182	184	156	184	184	184	184	184

Her bir madde için alt ölçek toplam puan ve tüm ölçek toplam puan ile olan korelasyon sonuçları görülmektedir.

Test tekrar Test ve Paralel Formlar güvenirligi

- Test tekrar Test ve Paralel Formlar güvenirligi katsayilari, SPSS'te Pearson Korelasyon katsayisi ile hesaplanabilir.

Kaynaklar

- Atılgan, H., Dođan, N. & Kan, A. (2011). *Eđitimde ölçme ve deđerlendirme*. Ankara: Anı Yayıncılık.
- Baykul, Y. (2000). *Eđitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara. ÖSYM Yayınları.
- Çokluk, Ö., Şekerciođlu, G. & Büyüköztürk, Ş. (2010). *Sosyal bilimler için çok deđişkenli istatistik*. Ankara: Pegem Akademi.
- Köse, İ. A. (2012). Geçerlik ve Güvenirlik. Demirtaşlı, R. N. *Eđitimde ölçme ve deđerlendirme*. Edge Akademi, Ankara.
- Şencan, H. (2005). Sosyal ve davranışsal ölçümlerde güvenirlik ve geçerlilik. Ankara: Seçkin Yayıncılık.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York Harper Collins.
- Tavşancıl, E. (2010). *Tutumların ölçülmesi ve SPSS ile veri analizi* (4. baskı). Ankara: Nobel Yayın Dađıtım.
- Tekin, H. (1996). *Eđitimde ölçme ve deđerlendirme*. Ankara: Yargı Yayınları.