Recall: A test is **valid** if it measures what it purports to measure. There are three types of validity to consider:

1. Content validity - subjective rating of items by content experts
2. Criterion-related validity - how well do test scores correlate with some criterion?
3. **Construct validity** (this lecture)

Definition: A test's *construct validity* (Cronbach & Meehl, 1955) is the degree to which it measures a theoretical construct or trait that it was designed to measure.

- establishing construct validity is an ongoing process
- based on current theory about trait being measured, make predictions about how test scores should behave in certain situations
- if predictions are supported, construct validity is enhanced

Two main ways to index construct validity:

1. Multitrait-multimethod matrix (Campbell & Fiske, 1959)
2. Factor analysis

Multitrait-multimethod matrix

- "multiple traits measured by multiple methods"
- steps:
  - think of two (or more) ways to measure construct of interest
  - identify other, distinctly different constructs which can be measured by these methods
  - test one sample, measuring each construct via each method
  - construct matrix of correlations

Example: Mosher (1968) proposed three distinct constructs related to guilt, which were measured by three different methods:

TABLE 10.2. Illustrative Table of Multitrait-Multimethod Matrix Data from a Study by Mosher

	Method 1			Method 2			Method 3		
	Trait A	B	C	A	B	C	A	B	C
1. True-False									
A. Sex-Guilt	(.95)								
B. Hostility-Guilt	.28	(.86)							
C. Morality-Conscience	.58	.39	(.92)						
2. Forced Choice									
A. Sex-Guilt	.86	.32	.57	(.95)					
B. Hostility-Guilt	.30	.90	.40	.39	(.76)				
C. Morality-Conscience	.52	.31	.86	.55	.26	(.84)			
3. Incomplete Sentences									
A. Sex-Guilt	.73	.10	.43	.64	.17	.37	(.48)		
B. Hostility-Guilt	.10	.63	.17	.22	.67	.19	.15	(.41)	
C. Morality-Conscience	.35	.16	.52	.31	.17	.56	.41	.30	(.58)

$N = 62$.

From D. L. Mosher, Measurement of guilt by self-report inventories, *Journal of Consulting and Clinical Psychology*, 32, 690–695. Copyright 1968 by the American Psychological Association. Adapted by permission of the author.

Three types of correlations:

- reliability coefficients (in parentheses) - should be large
- convergent validity coefficients (underlined) - correlations between measures of same construct using *different* measurement methods – should be large
  - *monotrait-heteromethod correlations*
- discriminant validity coefficients (not underlined) - should be smaller
  - correlations between measures of different constructs using same measurement methods (*heterotrait-monomethod*),
  - correlations between different constructs using different measurement methods (*heterotrait-heteromethod*)

Example: Consider the following multitrait-multimethod matrix for the constructs of *introversion* and *neuroticism*, measured by true/false (TF) and multiple choice (MC)

	Introversion TF	Neuroticism TF	Introversion MC	Neuroticism MC
Introversion TF	.80	.20	.75	.17
Neuroticism TF	.20	.85	.15	.71
Introversion MC	.75	.15	.83	.21
Neuroticism MC	.17	.71	.21	.91

Reliability coefficients:

Convergent validity coefficients:

Divergent validity coefficients:

- Heterotrait - monomethod:

- Heterotrait - heteromethod:

Overall validity assessment

Factor analysis

- mathematical technique for analyzing interrelationships among a set of variables and explaining these relationships in terms of a *reduced* number of variables, called **factors**
- in this lecture, we'll learn to do "eyeball" factor analysis
- in the next lecture, we'll learn to do it in JASP

Example 1:

	Test 1	Test 2	Test 3
Test 1	1.00	.98	.95
Test 2	.98	1.00	.97
Test 3	.95	.97	1.00

Example 2:

	Test 1	Test 2	Test 3	Test 4
Test 1	1.00	.90	0	0
Test 2	.90	1.00	0	0
Test 3	0	0	1.00	.90
Test 4	0	0	.90	1.00

Example 3:

	Test 1	Test 2	Test 3	Test 4
Test 1	1.00	.90	.30	.20
Test 2	.90	1.00	.30	.10
Test 3	.20	.30	1.00	.90
Test 4	.20	0.10	.90	1.00