Models of Classical Test Theory

Prof. Dr. Rolf Steyer
Chair of Methodology and Evaluation Research - University of Jena

April 3, 2009

Model of Essentially τ-Equivalent Tests and Model of τ-Congeneric Tests

Model of Essentially τ-Equivalent Tests — Assumptions

Definition: Assumptions (a$_2$) and (b)

- (a$_2$) essential τ-equivalence: $\tau_i = \lambda_{ij} + \tau_j$
- (b) uncorrelated errors $Cov(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j$

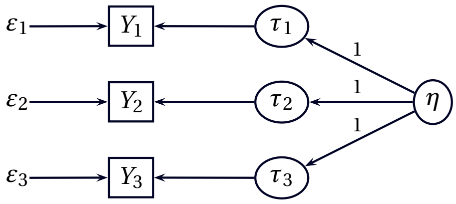
First implication:
(a$_2$) implies the existence of a latent variable η such that:

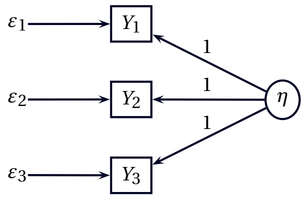
$$\tau_i = \lambda_i + \eta \quad \text{and} \quad Y_i = \lambda_i + \eta + \varepsilon_i \tag{1}$$

for all $i = 1, \ldots, m$

Proof:
Defining $\eta := \tau_1$ and $\lambda_i := \lambda_{i1}$, Equation (1) follows from (a_2).

Path diagram





Implied Covariance Structure

$$\begin{aligned}
Cov(Y_i, Y_j) &= Cov(\lambda_1 + \eta + \varepsilon_1,\ \lambda_2 + \eta + \varepsilon_2) \\
&= Cov(\eta, \eta) + Cov(\eta, \varepsilon_2) + Cov(\varepsilon_1, \eta) + Cov(\varepsilon_1, \varepsilon_2) \\
&= Var(\eta) \\
&= \sigma^2_\eta \\
\\
Var(Y_i) &= Var(\eta) + Var(\varepsilon_i) \\
&= \sigma^2_\eta + \sigma^2_{\varepsilon_i}
\end{aligned}$$

Implied Covariance Matrix

Implied covariance matrix for 3 essentially τ-equivalent tests:

$$\begin{bmatrix} \sigma_\eta^2 + \sigma_{\varepsilon_1}^2 & \sigma_\eta^2 & \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\eta^2 + \sigma_{\varepsilon_2}^2 & \sigma_\eta^2 \\ \sigma_\eta^2 & \sigma_\eta^2 & \sigma_\eta^2 + \sigma_{\varepsilon_3}^2 \end{bmatrix}$$

Testability

Testability in the total population:

$$Cov(Y_i, Y_j) = \sigma_\eta^2, \quad i \neq j$$

Testability in each subpopulation s:

$$\begin{aligned} Cov^{(s)}(Y_i, Y_j) &= \sigma_\eta^{(s)2}, \quad i \neq j \\ E^{(s)}(Y_i) - E^{(s)}(Y_j) &= E^{(s)}(Y_i - Y_j) \\ &= \lambda_{ij} \end{aligned}$$

The constant λ_{ij} is the same in all subpopulations.

Fixing Scales and Identification

Identification

$$\begin{aligned} Var(\eta) &= Cov(Y_i, Y_j), \quad i \neq j \\ Var(\varepsilon_i) &= Var(Y_i) - Cov(Y_i, Y_j), \quad i \neq j \\ Rel(Y_i) &= \frac{Cov(Y_i, Y_j)}{Var(Y_i)}, \quad i \neq j \end{aligned}$$

Fixing the scale of η

First possibility:

Fixing $E(\eta) = 0$ implies $\lambda_i = E(Y_i)$

Second possibility:

Fixing $\lambda_1 = 0$ implies $E(\eta) = E(Y_1)$ and $\lambda_i = E(Y_i) - E(Y_1)$

Uniqueness of Theoretical Concepts

η and the coefficients λ_i are defined by the assumptions uniquely *up to translations*.

This means: the latent variable η and the coefficients λ_i are measured on a *difference scale*.

Meaningfulness

Under translations of η and the coefficients λ_i, the following propositions have invariant truth values ('true' or 'false'):

- $\eta(\omega_1) - \eta(\omega_2)$, for all $\omega_1, \omega_2 \in \Omega$
- $\lambda_i - \lambda_j$
- $Var(\eta)$
- $Rel(Y_i)$

Test Lengthening

Spearman-Brown Formula and Cronbach's α

Test Lengthening $S = Y_1 + \ldots + Y_m$

Assumptions (a_2), (b) and (c) imply:

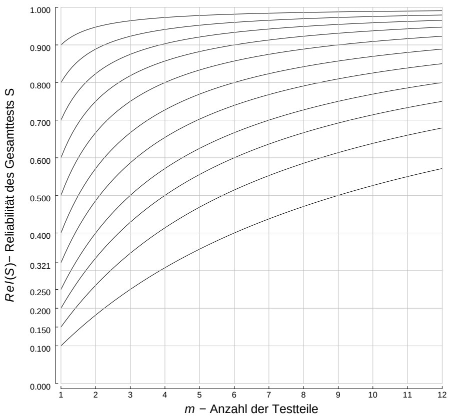
Spearman-Brown Formula: $$Rel(S) = \frac{m \cdot Rel(Y)}{1 + (m-1) \cdot Rel(Y)}$$

Assumption (b) (uncorrelated errors) implies that α is a *lower bound for the reliability* of $S = Y_1 + \ldots + Y_m$:

$$\textit{Cronbach's}\ \alpha := \frac{m}{m-1} \cdot \left[1 - \frac{\sum_{i=1}^{m} Var(Y_i)}{Var(S)} \right] \tag{2}$$

Reliability

Reliability of the sum scores $S := \sum_{i=1}^{m} Y_i$ as a function of test lengthening by factor m and of the reliability of the original variables Y_i, under the assumptions (a_1) [or (a_2)], (b) and (c).



Grafische Darstellung der Spearman−Brown−Formel

Model of τ-Congeneric Tests

Model of τ-Congeneric Tests — Assumptions

Assumptions:

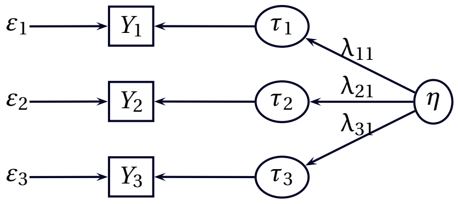
- *(a_3)* τ-congenericity
  $$\tau_i = \lambda_{ij0} + \lambda_{ij1}\, \tau_j, \quad \lambda_{ij0}, \lambda_{ij1} \in \mathbb{R}, \quad \lambda_{ij1} > 0$$
- *(b)* uncorrelated errors: $Cov(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j$

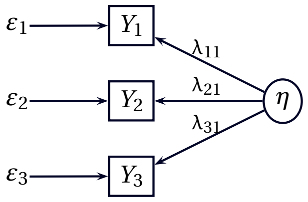
(a_3) implies the existence of a latent variable η and coefficients λ_{i0} and λ_{i1} such that:

$$\tau_i = \lambda_{i0} + \lambda_{i1}\, \eta \quad \text{and, therefore,} \quad Y_i = \lambda_{i0} + \lambda_{i1}\, \eta + \varepsilon_i$$

Proof:
Just fix $j = 1$ in (a_3) and define:
$\eta := \tau_1, \ \lambda_{i0} := \lambda_{i10}$ and $\lambda_{i1} := \lambda_{i11}$

Model of τ-Congeneric Tests — Path Diagram





Illustration

Tabelle. Illustration of the relationship between the manifest and latent variables in the model of τ-congeneric tests

Person	True scores		Latent scores	Observed scores		error scores		
u	τ_1	τ_2	η	Y_1	Y_2	ε_1	ε_2	$P(Y_i = y_i \mid U = u)$
				10	20	-2	-3	1/3
1	12	23	34	12	24	0	1	1/3
				14	25	2	2	1/3
				7	15	-3	-5	1/3
2	10	20	30	9	22	-1	2	1/3
				14	23	4	3	1/3
				3	14	-5	-3	1/3
3	8	17	26	10	15	2	-2	1/3
				11	22	3	5	1/3

Anmerkung: Fictive numbers. Each of the three Persons has its own (intra-individual) distribution of the Y-variables, but only a single score on each of the latent variables τ_i and η.

Uniqueness

(a_3) implies the existence of a latent variable η and coefficients λ_{ij0} and λ_{ij1} such that:

(1) $\tau_i = \lambda_{i0} + \lambda_{i1}\,\eta$ and (2) $Y_i = \lambda_{i0} + \lambda_{i1}\,\eta + \varepsilon_i$

The latent variable η and the coefficients λ_{i0} and λ_{i1} are uniquely defined only up to *linear transformations*. Hence, the equations above also hold for each variable $\eta' := \alpha + \beta\,\eta$, although with new coefficients $\lambda'_{i0} = \lambda_{i0} - \frac{\lambda_{i1}\,\alpha}{\beta}$ and $\lambda'_{i1} = \frac{\lambda_{i1}}{\beta}$.

Proof:
Inserting $\eta' := \alpha + \beta\,\eta$ as well as λ'_{i0} and λ'_{i1} into Equation (1) yields

$$\begin{aligned}\tau_i &= \lambda'_{i0} + \lambda'_{i1}\,\eta' \\ &= (\lambda_{i0} - \frac{\lambda_{i1}}{\beta}\,\alpha) + \frac{\lambda_{i1}}{\beta}\,(\alpha + \beta\,\eta) \\ &= \lambda_{i0} + \lambda_{i1}\,\eta\end{aligned}$$

Fixing the Scale

In order to have a uniquely defined scale, we have to fix location and units of measurement, for instance:

either by: $E(\eta) = 0$ and $Var(\eta) = 1$

or by: $\lambda_{10} = 0$ and $\lambda_{11} = 1$

Implied Covariance Structure

$$\begin{aligned}Cov(Y_i, Y_j) &= Cov(\lambda_{i0} + \lambda_{i1}\,\eta + \varepsilon_i,\ \lambda_{j0} + \lambda_{j1}\,\eta + \varepsilon_j) \\ &= \lambda_{i1}\lambda_{j1}\,Var(\eta) \qquad \text{for } i \neq j\end{aligned}$$

$$\begin{aligned}Var(Y_i) &= Var(\lambda_{i0} + \lambda_{i1}\,\eta + \varepsilon_i) \\ &= \lambda_{i1}^2\,Var(\eta) + Var(\varepsilon_i)\end{aligned}$$

Implied Covariance Matrix for Three τ-Congeneric Tests

$$\begin{bmatrix}\lambda_{11}^2\sigma_\eta^2 + \sigma_{\varepsilon_1}^2 & & \\ \lambda_{11}\lambda_{21}\sigma_\eta^2 & \lambda_{21}^2\sigma_\eta^2 + \sigma_{\varepsilon_2}^2 & \\ \lambda_{11}\lambda_{31}\sigma_\eta^2 & \lambda_{21}\lambda_{31}\sigma_\eta^2 & \lambda_{31}^2\sigma_\eta^2 + \sigma_{\varepsilon_3}^2\end{bmatrix}$$

Identification of Loadings

$$\begin{aligned} \frac{Cov(Y_i, Y_j) \cdot Cov(Y_i, Y_k)}{Cov(Y_j, Y_k)} &= \frac{\lambda_{i1}\lambda_{j1} Var(\eta) \cdot \lambda_{i1}\lambda_{k1} Var(\eta)}{\lambda_{j1}\lambda_{k1} Var(\eta)} \\ &= \lambda_{i1}^2 Var(\eta) \quad i \neq j, \quad i \neq k, \quad j \neq k \end{aligned}$$

If we fix $Var(\eta) = 1$, then:

$$\lambda_{i1} = \sqrt{\frac{Cov(Y_i, Y_j) \cdot Cov(Y_i, Y_k)}{Cov(Y_j, Y_k)}} \quad i \neq j, \quad i \neq k, \quad j \neq k$$

Identification of Other Parameters

If we fix the scale of η by

$$E(\eta) = 0 \quad \text{and} \quad Var(\eta) = 1,$$

then:

$$\begin{aligned} Var(\varepsilon_i) &= Var(Y_i) - \lambda_{i1}^2 \\ Rel(Y_i) &= \frac{\lambda_{i1}^2}{Var(Y_i)} \end{aligned}$$

Testability in the total Population

We can test the model of τ-congeneric tests in the total population by testing the hypothesis:

$$\frac{Cov(Y_i, Y_k)}{Cov(Y_j, Y_k)} = \frac{Cov(Y_i, Y_l)}{Cov(Y_j, Y_l)}$$

This can be seen by inserting the implied covariance structure for each of the four covariances:

$$\frac{\lambda_{i1}\lambda_{k1} Var(\eta)}{\lambda_{j1}\lambda_{k1} Var(\eta)} = \frac{\lambda_{i1}\lambda_{l1} Var(\eta)}{\lambda_{j1}\lambda_{l1} Var(\eta)} \quad i \neq k, \quad i \neq l, \quad j \neq k, \quad j \neq l$$

Testability Using the Implied Mean Structure

The model of τ-congeneric tests implies:

$$\frac{E^{(1)}(Y_i) - E^{(2)}(Y_i)}{E^{(1)}(Y_j) - E^{(2)}(Y_j)} = \frac{E^{(3)}(Y_i) - E^{(4)}(Y_i)}{E^{(3)}(Y_j) - E^{(4)}(Y_j)}$$

This can be seen as follows: Inserting the model equations into $E^{(s)}(Y_i)$ yields:

$$E^{(s)}(\lambda_{i0} + \lambda_{i1}\eta + \varepsilon_i) = \lambda_{i0} + \lambda_{i1}E^{(s)}(\eta) \quad \text{for each subpopulation } s = 1, \ldots, 4,$$

Inserting this result into the first equation yields:

$$\frac{\lambda_{i0} + \lambda_{i1}E^{(1)}(\eta) - [\lambda_{i0} + \lambda_{i1}E^{(2)}(\eta)]}{\lambda_{j0} + \lambda_{j1}E^{(1)}(\eta) - [\lambda_{j0} + \lambda_{j1}E^{(2)}(\eta)]} = \frac{\lambda_{i0} + \lambda_{i1}E^{(3)}(\eta) - [\lambda_{i0} + \lambda_{i1}E^{(4)}(\eta)]}{\lambda_{j0} + \lambda_{j1}E^{(3)}(\eta) - [\lambda_{j0} + \lambda_{j1}E^{(4)}(\eta)]}$$

$$\frac{\lambda_{i1}}{\lambda_{j1}} = \frac{\lambda_{i1}}{\lambda_{j1}}$$

Meaningfulness

For instance, propositions about

- $\frac{\eta(\omega_1) - \eta(\omega_2)}{\eta(\omega_3) - \eta(\omega_4)}$
- $\frac{\lambda_{i1}}{\lambda_{j1}}$
- $\lambda_{i1}^2 \, Var(\eta)$

have *invariant truth values* under the admissible transformations (here: linear transformations). The first equation means that η is measured on an *interval scale*.