DOCUMENT RESUME

ED 055 101	TM 000 818
AUTHOR	Werts, Charles E.; And Others
TITLE	A Congeneric Model for Platonic True Scores.
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	May 71
GRANT	OEG-2-700033(509)
NOTE	14p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Correlation; Hypothesis Testing; *Mathematical Models; *Mathematics; Probability Theory; *Statistical Analysis; *True Scores

ABSTRACT

To resolve a recent controversy between Klein and Cleary and Levy, a model for dichotomous congeneric items is presented which has mean errors of zero, dichotomous true scores that are uncorrelated with errors, and errors that are mutually uncorrelated. (Author)

RESEARCH BULLETIN

RB-71-22

A CONGENERIC MODEL FOR PLATONIC TRUE SCORES

Charles E. Werts, Robert L. Linn
and Karl Jöreskog

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the authors. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
May 1971

A CONGENERIC MODEL FOR PLATONIC TRUE SCORES

Charles E. Werts, Robert L. Linn, and Karl Jöreskog

Abstract

To resolve a recent controversy between Klein and Cleary and Levy, a model for dichotomous congeneric items is presented which has mean errors of zero, dichotomous true scores that are uncorrelated with errors, and errors that are mutually uncorrelated.

A CONGENERIC MODEL FOR PLATONIC TRUE SCORES[1]

Charles E. Werts, Robert L. Linn, and Karl Jöreskog

In a discussion of platonic true scores, Klein and Cleary (1967) state that the use of platonic true scores makes the assumptions of classical test theory generally untenable. They illustrate their argument with dichotomous items and a dichotomous true score and show that: "The classical test theory formulation $\sigma_X^2 = \sigma_T^2 + \sigma_E^2$, can only be true if the mean error is not zero" (Klein & Cleary, 1967, p. 78). This statement is based on the following definitions of observed (X), true (T), and error (E) scores:

$$T = \begin{cases} 1 \text{ if phenomenon is present} \\ 0 \text{ if phenomenon is absent} \end{cases} ,$$

$$X = \begin{cases} 1 \text{ if phenomenon is rated as present} \\ 0 \text{ if phenomenon is rated as absent} \end{cases} ,$$

and $E = X - T$. Klein and Cleary go on to consider two parallel dichotomous items, X_1 and X_2, and show that the covariance between E_1 and E_2 is positive when the errors, E_1 and E_2, have zero means. With correlated error scores, the correlation between two parallel items overestimates the item reliabilities. In response, Levy (1969) argued that the classical assumptions can be shown to hold for a dichotomous item if

$$X = \begin{cases} a \text{ if phenomenon is rated as present} \\ b \text{ if phenomenon is rated as absent} \end{cases} ,$$

true scores (T) are defined as above and $E = X - T$ as before. This modification will indeed make it possible for the mean error to be zero and the covariance between T and E to be zero. As Klein and Cleary (1969) note, however, Levy does not provide a means of solving for "a" and "b" without knowledge of T. In any practical application, T would be unknown and therefore "a" and "b" would be unknown. Also, no way of obtaining item reliabilities is presented. The purpose of this paper is to provide an alternative formulation which allows for the model parameters to be determined given the structural specification of zero mean error and no correlation among errors for different items or between errors and true scores. Our approach is drawn from latent structure analysis (Anderson, 1959) for the special case of dichotomous latent variables.

I. A Congeneric Model for Dichotomous Items

The equation for congeneric tests is given by Jöreskog (1968, 1970, 1971) as

$$X_{ij} = B_{jT}T_i + I_j + E_{ij} \quad , \tag{1}$$

where T_i is the true score for person i,

X_{ij} is the observed score on item j for person i,

B_{jT} is the slope of the X on T_i regression line,

I_j is the intercept of this regression line, and

E_{ij} is the error for person i on item j.

To illustrate the application of this definition to the case in which X_{ij} and T_i are both dichotomous (scored 1, 0), consider the case of three items, which is the minimum number of items required to identify model parameters uniquely, given experimentally independent measures. The equations are

$$X_1 = B_{1T}T + I_1 + E_1 \quad , \tag{1a}$$

$$X_2 = B_{2T}T + I_2 + E_2 \quad , \tag{1b}$$

$$X_3 = B_{3T}T + I_3 + E_3 \quad , \tag{1c}$$

where the E 's are mutually uncorrelated and are uncorrelated with T .[2] In the case of dichotomous variables

$$B_{jT} = \frac{P\{X_j = 1, T = 1\} - P\{X_j = 1\}P\{T = 1\}}{P\{T = 1\}P\{T = 0\}} = P\{X_j = 1|T = 1\} - P\{X_j = 1|T = 0\}$$

and

$$I_j = P\{X_j = 1\} - B_{jT}P\{T = 1\} = P\{X_j = 1|T = 0\} \quad .$$

This model is somewhat more complicated than the model considered by Klein and Cleary (1967) where $X = T + E$ with X , T , and E all taking values of 0 or 1. In essence, the congeneric model is equivalent to the model suggested by Levy (1969) if his "a" and "b" are allowed to vary from item to item. For a given item, "a_j" would equal $(1 - I_j)/B_{jT}$, "b_j" would equal $-I_j/B_{jT}$, and Levy's error would equal the error of equations 1, 2, or 3 divided by B_{jT} . To illustrate the point that the congeneric model does allow for the traditional psychometric assumptions in the dichotomous case, consider the following example constructed using the equations provided by Anderson (1959, sec. 2.4).

1. The θ_j (proportion of false negatives, i.e., $P\{X_j = 0|T = 1\} = P\{X_j = 0, T = 1\} \div P_T$), ϕ_j (proportion of false positives, i.e., $P\{X_j = 1|T = 0\} = P\{X_j = 1, T = 0\} \div (1 - P_T)$) , and P_T (the true proportion $P\{T = 1\}$) are given as:

$\theta_1 = .30$, $\theta_2 = .40$, $\theta_3 = .10$;

$\phi_1 = .10$, $\phi_2 = .50$, $\phi_3 = .30$;

$P_T = .60$, and $Q_T = 1 - P_T = .40$.

2. The expected marginal distributions (P_j = Prob $\{X_j = 1\}$) are $P_j = (1 - \theta_j)P_T + \phi_j Q_T$, i.e., $P_1 = .46$, $P_2 = .56$, and $P_3 = .66$.

3. The expected joint probabilities for pairs of items, $P_{jj'} =$ Prob $\{X_j = 1, X_{j'} = 1\} = (1 - \theta_j)(1 - \theta_{j'})P_T + \phi_j\phi_{j'}Q_T \quad (j \neq j')$ are: $P_{12} = .272$, $P_{13} = .390$, and $P_{23} = .384$.

4. The expected joint probability for three items, $P_{jj'j''} =$ Prob $\{X_j = 1, X_{j'} = 1, X_{j''} = 1\} = (1 - \theta_j)(1 - \theta_{j'})(1 - \theta_{j''})P_T + \phi_j\phi_{j'}\phi_{j''}Q_T$ $(j \neq j' \neq j'')$ is $P_{123} = .2328$.

5. The regression weights $(B_{jt} = 1 - \theta_j - \phi_j)$ are $B_{1T} = .60$, $B_{2T} = .10$, and $B_{3T} = .60$.

6. The intercepts $(I_j = P_j - B_{jT}P_T = \phi_j)$ are $I_1 = .10$, $I_2 = .50$, and $I_3 = .30$. The possible events for combinations of the three items and the proportion of people in each event are shown in Table 1. The means of the errors are zero, the true score is uncorrelated with the errors and the errors are uncorrelated with each other.

Insert Table 1 about here

II. Identification

In an actual problem the situation would be reversed from the example shown in section I, i.e., the probabilities P_1, P_2, P_3, P_{12}, P_{13}, P_{23}, and P_{123} correspond to observed scores, and it would be desirable to identify the seven parameters, θ_1, θ_2, θ_3, ϕ_1, ϕ_2, ϕ_3, and P_T . In principle, one could solve the seven equations for this purpose:

$$P_1 = (1 - \theta_1)P_T + \phi_1Q_T \ , \tag{2a}$$

$$P_2 = (1 - \theta_2)P_T + \phi_2Q_T \ , \tag{2b}$$

$$P_3 = (1 - \theta_3)P_T + \phi_3Q_T \ , \tag{2c}$$

$$P_{12} = (1 - \theta_1)(1 - \theta_2)P_T + \phi_1\phi_2Q_T \quad , \tag{2d}$$

$$P_{13} = (1 - \theta_1)(1 - \theta_3)P_T + \phi_1\phi_3Q_T \quad , \tag{2e}$$

$$P_{23} = (1 - \theta_2)(1 - \theta_3)P_T + \phi_2\phi_3Q_T \quad , \tag{2f}$$

$$P_{123} = (1 - \theta_1)(1 - \theta_2)(1 - \theta_3)P_T + \phi_1\phi_2\phi_3Q_T \quad . \tag{2g}$$

The solution to these equations is facilitated by noting that in the congeneric model the expected covariance $(C_{jj'})$ between two items is given by

$$C_{jj'} = B_{jT}B_{j'T}V_T \quad ,$$

where V_T is the variance of T. Translating into probabilities:

$$(P_{jj'} - P_jP_{j'}) = (1 - \theta_j - \phi_j)(1 - \theta_{j'} - \phi_{j'})P_TQ_T \quad . \tag{3}$$

This means that

$$C_{12} = P_{12} - P_1P_2 = (1 - \theta_1 - \phi_1)(1 - \theta_2 - \phi_2)P_TQ_T \quad , \tag{4a}$$

$$C_{13} = P_{13} - P_1P_3 = (1 - \theta_1 - \phi_1)(1 - \theta_3 - \phi_3)P_TQ_T \quad , \tag{4b}$$

$$C_{23} = P_{23} - P_2P_3 = (1 - \theta_2 - \phi_2)(1 - \theta_3 - \phi_3)P_TQ_T \quad . \tag{4c}$$

These equations can be solved for

$$(1 - \theta_1 - \phi_1)^2\, P_TQ_T = \frac{C_{12}C_{13}}{C_{23}} = B^2_{1T}P_TQ_T \quad , \tag{5a}$$

$$(1 - \theta_2 - \phi_2)^2\, P_TQ_T = \frac{C_{12}C_{23}}{C_{13}} = B^2_{2T}P_TQ_T \quad , \tag{5b}$$

$$(1 - \theta_3 - \phi_3)^2\, P_TQ_T = \frac{C_{13}C_{23}}{C_{12}} = B^2_{3T}P_TQ_T \quad . \tag{5c}$$

The triple covariance C_{123} is defined (Boudon, 1968, p. 226) as the expectation of the products of the deviations of all three variables simultaneously, which is equal in the dichotomous case to $C_{123} = P_{123} - P_1(P_{23} - P_2P_3) - P_2(P_{13} - P_1P_3) - P_3(P_{12} - P_1P_2) - P_1P_2P_3$. (6)

Using equations (2a, b, c, d, e, f) equation (6) may be translated to $C_{123} = B_{1T}B_{2T}B_{3T}P_TQ_T(Q_T^2 - P_T^2)$ and from equations (5a, b, c) we obtain

$$C_{123} = \frac{\sqrt{C_{12}C_{13}C_{23}}(Q_T^2 - P_T^2)}{\sqrt{P_TQ_T}} . \qquad (7)$$

Applying these equations to our example,

1. Compute covariances by equations (4a, b, c):

   $C_{12} = .0144$, $C_{13} = .0864$, and $C_{23} = .0144$.

2. Using equation (6) compute $C_{123} = -.001728$.
3. From equation (7),

$$\frac{C_{123}}{\sqrt{C_{12}C_{13}C_{23}}} \approx -.4082 = \frac{(Q_T^2 - P_T^2)}{\sqrt{P_TQ_T}} .$$

4. Solving for $P_T = 1 - Q_T$ we obtain $P_T = .60$.
5. From equations (5a, b, c) and substituting in this value of P_T ,

   $B_{1T} \approx .60$,

   $B_{2T} \approx .10$,

   $B_{3T} \approx .60$.

6. It can be shown (equations 2a, b, c) that $\phi_j = P_j - B_{jT}P_T$ permitting calculation of $\phi_j = I_j$:

$$I_1 = \phi_1 = .10 \quad ,$$

$$I_2 = \phi_2 = .50 \quad ,$$

$$I_3 = \phi_3 = .30 \quad .$$

7. Since $\theta_j = 1 - B_{jT} - \phi_j$,

$$\theta_1 = .30 \quad ,$$

$$\theta_2 = .40 \quad ,$$

$$\theta_3 = .10 \quad .$$

8. Item reliabilities R_{jj} are $R_{jj} = B^2_{jT}P_TQ_T/P_jQ_j$, i.e.,

$$R_{11} = .3478 \quad ,$$

$$R_{22} = .0097 \quad ,$$

$$R_{33} = .3850 \quad .$$

In the case of three congeneric items the model parameters are just identified, i.e., there are seven equations in seven unknowns, which is the reason that the parameters may be obtained as an exact function of the observed probabilities. In the case of overidentified models one of the estimating procedures discussed by Anderson (1959) can be used. One procedure minimizes a χ^2 function of the observed probabilities (P_0) and the expected probabilities (P_E) generated as a function of the parameter estimates (Cochran, 1968; Mote & Anderson, 1965). In the general case of J items there will be $(2^J - 1)$ independent observed probabilities in the cross-tabulation table from which $(2J + 1)$ parameters are to be estimated. In the special case of two items of equal accuracy the reliability is the correlation between these items, but the model parameters cannot be identified

(Cochran, 1968, sec. 6) since $P_E\{X_1 = 1, X_2 = 0\} = P_E\{X_1 = 0, X_2 = 1\}$, i.e., there are only two independent probabilities to estimate three parameters (θ, ϕ, P_T) .

III. Variations

It is sometimes the case that three items with errors that are uncorrelated with true scores or errors of other items are available but one of these measures another variable, i.e.,

$$X_1 = B_1T_1 + I_1 + E_1 \quad ,$$

$$X_2 = B_2T_1 + I_2 + E_2 \quad , \qquad (8)$$

$$X_3 = B_3T_2 + I_3 + E_3 \quad .$$

In econometrics X_3 is called an "instrumental" variable (Johnston, 1963, p. 165). The equation for X_3 can be transformed into

$$X_3 = B_3^*T_1 + I_3^* + E_3^* \quad , \qquad (8a)$$

where

$$B_3^* = B_{T_2T_1}B_3 \quad .$$

B_3^* is identified but $B_{T_2T_1}$ and B_3 are not. In the case of dichotomous variables, therefore, the true proportion P_{T_1} may be estimated as shown in section II by treating X_3 as a congeneric measure of T_1 and $B_3^* = (1 - \theta_3 - \phi_3)(1 - \theta_{T_1} - \phi_{T_1})$, where $\theta_{T_1} = P\{T_2 = 0 | T_1 = 1\}$ and $\phi_{T_1} = P\{T_2 = 1 | T_1 = 0\}$. The validity of such an analysis is dependent on the correctness of the independence assumption.

The above analysis can be extended to the case of four items with mutually uncorrelated errors and no correlation between error and true scores, two of each measuring different variables:

$$\begin{aligned} X_1 &= B_1T_1 + I_1 + E_1 \quad , \\ X_2 &= B_2T_1 + I_2 + E_2 \quad , \\ X_3 &= B_3T_2 + I_3 + E_3 \quad , \\ X_4 &= B_4T_2 + I_3 + E_3 \quad . \end{aligned} \tag{9}$$

Following the above line of reasoning all parameters in this model (P_{T_1}, $P\{T_1 = 1, T_2 = 1\}$, P_{T_2}, θ_1, θ_2, θ_3, θ_4, ϕ_1, ϕ_2, ϕ_3, and ϕ_4) may be identified. There are 15 independent proportions in the cross-tabulation table, so that the minimized χ^2 would have four degrees of freedom. In principle, a measure of the tenability of certain assumptions is obtained from changes in the χ^2. For example, if it were desired to test the hypothesis that X_1 and X_2 were of equal accuracy, increases in the total χ^2 (with two degrees of freedom), resulting from setting $\theta_1 = \theta_2$ and $\phi_1 = \phi_2$, would be an indicator of the tenability of this hypothesis.

References

Anderson, T. W. Some scaling models and estimation procedures in the latent class model. In U. Grenander (Ed.), *Probability and statistics, The Harold Cramér volume*. New York: Wiley, 1959. Pp. 9-38.

Boudon, R. A new look at correlation analysis. In H. M. Blalock & A. B. Blalock (Eds.), *Methodology in social research*. New York: McGraw-Hill, 1968. Pp. 199-235.

Cochran, W. G. Errors of measurement in statistics. *Technometrics*, 1968, *10*, 637-666.

Johnston, J. *Econometric methods*. New York: McGraw-Hill, 1963.

Jöreskog, K. G. Statistical models for congeneric test scores. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 1968, 213-214.

Jöreskog, K. G. A general method for analysis of covariance structures. *Biometrika*, 1970, *57*, 239-251.

Jöreskog, K. G. Statistical analysis of sets of congeneric tests. *Psychometrika*, 1971, *36*, in press.

Klein, D. F., & Cleary, T. A. Platonic true scores and error in psychiatric rating scales. *Psychological Bulletin*, 1967, *68*, 77-80.

Klein, D. F., & Cleary, T. A. Platonic true scores: Further comment. *Psychological Bulletin*, 1969, *71*, 278-280.

Levy, P. Platonic true scores and rating scales: A case of uncorrelated definitions. *Psychological Bulletin*, 1969, *71*, 276-277.

Mote, V. L., & Anderson, R. L. An investigation of the effect of misclassification on the properties of X^2-tests in the analysis of categorical data. *Biometrika*, 1965, *52*, 95-109.

Footnotes

[1] The research reported herein was performed pursuant to Grant No. OEG-2-700033(509) with the United States Department of Health, Education, and Welfare and the Office of Education.

[2] The true scores are not *independent* of the error scores or errors of each other, as is assumed in Anderson's (1959) derivations however, for our purposes the assumption that these variables are *uncorrelated* yields the same formulas.

Table 1

Possible Events for Three Congeneric Dichotomous Items

Proportion of People	T	X_1	X_2	X_3	E_1	E_2	E_3
.2268	1	1	1	1	.3	.4	.1
.0252	1	1	1	0	.3	.4	-.9
.1512	1	1	0	1	.3	-.6	.1
.0168	1	1	0	0	.3	-.6	-.9
.0972	1	0	1	1	-.7	.4	.1
.0108	1	0	1	0	-.7	.4	-.9
.0648	1	0	0	1	-.7	-.6	.1
.0072	1	0	0	0	-.7	-.6	-.9
.0060	0	1	1	1	.9	.5	.7
.0140	0	1	1	0	.9	.5	-.3
.0060	0	1	0	1	.9	-.5	.7
.0140	0	1	0	0	.9	-.5	-.3
.0540	0	0	1	1	-.1	.5	.7
.1260	0	0	1	0	-.1	.5	-.3
.0540	0	0	0	1	-.1	-.5	.7
.1260	0	0	0	0	-.1	-.5	-.3