Methodika
1989, Vol. III, pp. 25–60

Models of Classical Psychometric Test Theory as Stochastic Measurement Models: Representation, Uniqueness, Meaningfulness, Identifiability, and Testability*)

Rolf Steyer

University of Trier, Department of Psychology

Abstract: Representation, uniqueness, meaningfulness, identifiability, and testability theorems are proven for (a) the model of essentially τ-equivalent variables and (b) for the model of τ-congeneric variables, both of which are shown to be derived measurement models. The set of observational units, the set of variables considered, and the conditional expectations (given the units) of the observable random variables are the crucial components of the derived measurement system. A number of fallacies that oftentimes appeared in the literature are discussed. These fallacies concern the properties of the error variables, the relationship between classical models and probabilistic models for categorical response variables, and the empirical testability of the models. Aside from the well-known implications for the structure of the covariance matrix of the observed variables, other empirically testable implications concern subpopulations. The model of essentially τ-equivalent variables implies the equality of the differences between the expectations of the observed variables in different subpopulations, whereas the model of τ-congeneric variables implies the equality of the factor loadings in different subpopulations. An example from state and trait anxiety research illustrates some of the theoretical results.

More than twenty years have passed since Novick's seminal paper on *The Axioms and Principal Results of Classical Test Theory* (1966) and exactly 20 years since Lord and Novick's fundamental book on *Statistical Theories of Mental Test Scores* (1968). Although the models and procedures proposed by these authors are widely used in empirical psychology as measurement models (see, e.g., the title of Allen & Yen's (1979) book *Introduction to Measurement Theory)*, mathematical psychologists have widely ignored this class of models following the verdict of Suppes and Zinnes (1963) according to which psychological tests are "pseudopointer

*) I thank Angela Auer and Prof. Dr. Peter Schwenkmezger for their cooperation in the study reported in Section 3. I also gratefully acknowledge helpful comments on an earlier draft by Hans Müller, University of Frankfurt/Main and Bernhard Baltes-Götz, University of Trier. This research was funded by the Deutsche Forschungsgemeinschaft (STE 411/1–1).

instruments", the readings of which "do not correspond to any *known* fundamental or derived numerical assignment" (p. 21; italics in the original). Although this statement does not hold true any more with respect to models for categorical response variables (see, e.g., Hamerle, 1979; 1982), it has still been valid with respect to models of classical psychometric test theory.

The present article is intended to fill this gap of knowledge indicated by Suppes and Zinnes and to serve as a basis for a rational discussion of the relationship between models of classical psychometric test theory, deterministic models of measurement, and the probabilistic models of measurement for categorical response variables. Two classical models are treated in detail: (a) the model of essentially τ-equivalent and (b) the model of τ-congeneric variables. Both are shown to be models of derived measurement based on the conditional expectations of the observable (e.g., test-score) variables given the observational units.

The organization of this paper is as follows: The introduction of the formal framework is succeeded by a section on the representation, uniqueness, meaningfulness, identifiability, and testability theorems for the model of essentially τ-equivalent variables. Then an example with new data from state and trait anxiety research illustrates several procedures of empirical tests of this model. In the next section, the model of τ-congeneric variables is treated, again dealing with the theorems mentioned above. The discussion focusses on the properties of the residuals, the relationship to probabilistic models for categorical response variables, and the empirical testability of the models.

1. The formal framework

1.1. The set of possible outcomes

Classical psychometric test theory (or theory of mental tests; Gulliksen, 1950; Lord & Novick, 1968) in its modern version presented by Zimmerman (1975, 1976) considers the following type of random experiments (see Appendix, Note 1): A unit u (e.g., a person or a person-in-a-situation) is sampled from a set U of observational units and the values of u with respect to m manifest properties (e.g., test performances) are registered. Hence, the set of possible outcomes of this kind of random experiment is of the type

$$\Omega = U \times A. \tag{1}$$

If, for example, $A = \Omega_1 \times \cdots \times \Omega_m$, each Ω_i, $i \in I := \{1, \ldots, m\}$, may be the set of possible values with respect to the ith manifest property (e.g., test or item) to be observed. In this case, an outcome $\omega = (u, \omega_1, \ldots, \omega_m) \in \Omega$ consists of an observational unit u and the values $\omega_1, \ldots, \omega_m$ of the unit with respect to the m manifest or observable properties. These values might still be qualitative in nature or already numbers such as test scores. Note that this way of defining the set Ω will enable us to conceptualize persons and their properties (such as their sex or their aptitude) as random events or random variables (see, e.g., Eq. 3 below).

Examples. A person u is sampled from a set U of persons and his or her raw test scores on two parallel forms of a personality questionaire are registered. In this case, $\Omega = U \times \mathbb{N}_0 \times \mathbb{N}_0$, where $\mathbb{N}_0$ denotes the set of natural numbers including 0. The natural numbers are the possible raw scores. As another example consider an ability test consisting of 10 problems that can be solved (+) or not (−). In this case $\Omega = U \times \{+, -\}^{10}$, where $\{+, -\}^{10} := \{+, -\} \times \cdots \times \{+, -\}$ contains the 2^{10} possible combinations such as $\langle +, +, -, -, +, +, -, -, -, + \rangle$.

1.2. The random variables

Next we consider the random variables $Y_i: \Omega \to \overline{\mathbb{R}}$, $i \in I$, (e.g., the test-score variables) that map the possibly qualitative values (such as "+" or "−") of the m manifest properties considered into the set $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$, where $\mathbb{R}$ denotes the set of real numbers. It is assumed that the Y_i have finite and positive variances.

Note that we only require fixed rules of assigning numbers to the possible outcomes. Especially, no measurement model is necessary *at this first step*. Stochastic measurement models use the Y-variables as their *empirical basis*. Whether or not this basis is well-chosen, is not determined a priori but by (a) the validity of the model built on the Y-variables and by (b) its usefulness within a larger theory and its intended applications. Of course, substantive heuristic theories may help to choose useful Y-variables.

Next, we consider the projection (see Appendix, Note 2) $p_U: \Omega \to U$, the values of which are the units $u \in U$, that is,

$$p_U(\omega) = p_U((u, a)) = u, \quad \text{for each } \omega \in \Omega. \tag{2}$$

Note that the mapping p_U is supposed to be a random variable, although its values are not numbers but qualitative elements, namely the observational units such as persons. (See, e.g., Bauer, 1978 or 1981, for the general concepts of measure and probability theory such as nonnumerical random variables.)

1.3. *The true-score variables, the residuals, and their properties*

We now define the variables τ_i to be the conditional expectations (or synonymously, regressions; see Appendix, Note 3) of Y_i given p_U:

$$\tau_i := E(Y_i | p_U), \quad i \in I. \tag{3}$$

If, in a given application, the units are persons, we may call the variable τ_i the *true-score variable* (or *person-regression*) of the variable Y_i. The values of such a person-regression τ_i are the conditional expectations (or *true scores*) $E(Y_i | p_U = u)$ of Y_i given a person u. Note again that the τ_i are random variables, the values of which are real numbers, the "true scores" of the units $u \in U$.

The general properties of the *residuals*

$$\varepsilon_i := Y_i - E(Y_i | p_U), \quad i \in I. \tag{4}$$

have been discussed in some detail by Steyer (1988); they may be written as follows:

$$Y_i = \tau_i + \varepsilon_i, \quad i \in I, \tag{5}$$

$$E(\varepsilon_i | f(p_U)) = 0, \quad i \in I, \quad \text{for every } p_U\text{-measurable mapping } f(p_U) \text{ (see Appendix, Note 4)}, \tag{6}$$

$$E\varepsilon_i) = 0, \quad i \in I, \tag{7}$$

$$E(\varepsilon_i | \tau_j) = 0, \quad i, j \in I, \tag{8}$$

$$Cov(\varepsilon_i, f(p_U)) = 0, \quad i \in I, \quad \text{for every } p_U\text{-measurable numerical function } f(p_U) \text{ (see Appendix, Note 4)}, \tag{9}$$

$$Cov(\varepsilon_i, \tau_j) = 0, \quad i, j \in I, \tag{10}$$

$$Var(Y_i) = Var(\tau_i) + Var(\varepsilon_i), \quad i \in I, \tag{11}$$

where Var( ) and Cov(,) denote the variance and covariance, respectively.

Comments. Note that Equation 8 is a special case of 6, and Equation 10 is a special case of 9. Equation 6 is the crucial one, because it implies the Equations 7 to 11 (see Appendix, Note 5). In order to understand the logical nature of the Equations 5 to 11, it is important to notice that none of these Equations is an axiom. Instead, each of it is an immediate consequence of defining the variables τ_i and ε_i by Equations 3 and 4 (for proofs, see, e.g., Steyer, 1988; Tack, 1980; or Zimmerman, 1975). All equations above are *always true*, that is, they are mathematical tautologies. They also hold if the variables Y_i are dichotomous (see, e.g., Fischer, 1974; Lord, 1980; Weiss, 1983).

Another noteworthy point is that noncorrelation of the residuals (or "errors"), $Cov(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$, $i, j \in I$, which has often been postulated as an axiom in classical psychometric test theory, is *not* a consequence of Equations 3 and 4 (see, e.g., Tack, 1980; Zimmerman, 1975; or Zimmerman & Williams, 1977). Hence, residuals can in fact be correlated among each other in empirical applications, the consequences of which have been discussed in detail by Zimmerman & Williams (1977).

It should also be noticed that the definition of the coefficient of determination

$$\begin{aligned} Rel(Y_i) := R^2_{Y_i | p_U} &:= Var(E(Y_i | p_U))/Var(Y_i) \\ &= Var(\tau_i)/Var(Y_i), \quad \text{if } 0 < Var(Y_i) < \infty, \end{aligned} \tag{12}$$

which is called the *reliability coefficient* in the context of classical psychometric test theory, is *neither* based on the assumption of uncorrelated errors, *nor* on any assumption other than that the variance of the Y-variables are finite and greater than zero. $Rel(Y_i) = 0$ indicates that the variable Y_i is not determined at all by the observational units. If Y_i is dichotomous with values 0 and 1, $Rel(Y_i) = 0$ means that the probability for $Y_i = 1$ does not depend on the observational unit, that is, $P(Y_i = 1 | p_U = u) = P(Y_i = 1)$ for all $u \in U$.

2. *Essentially* τ-equivalent variables

Although the mathematical framework above is sufficient to properly *define* the variables τ_i, ε_i, their variances and the concept of reliability, it is *not* sufficient to derive formulas for the determination of the variances of the τ_i, ε_i or of the coefficient of reliability from the variances, covariances, correlations, or other characteristics of the common distribution of the observable variables Y_i. We will therefore introduce some assumptions defining a model which (a) requires the Y-variables to be essentially τ-equivalent (e.g., resulting from parallel forms of a personality inventory) and (b) that you register at least $m = 2$ of those variables. Later on we will turn to a less restrictive set of assumptions, known as the model of τ-congeneric variables (Jöreskog, 1971), which, in general, presupposes at least $m = 3$ observed variables.

Although the model of essentially τ-equivalent variables is well-known (see, e.g., Lord & Novick, 1968; Fischer, 1974), it is often not well understood. Many authors maintain that models of classical psychometric test theory cannot be tested empirically. This is wrong as will be shown in this section.

The Covariance Structure Theorem, for example, will reveal that the model of essentially τ-equivalent variables with conditional regressive independence, implies a one-factor model such as depicted in Figure 1. The readers familiar with confirmatory factor analysis (Lawley & Maxwell, 1971; Jöreskog, 1967, 1969) are well aware of the testability of this kind of models.

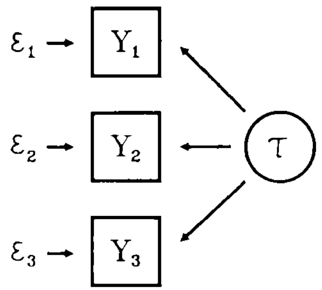

Figure 1: Arrow diagram of the factor model implied by the model of $m = 3$ essentially τ-equivalent variables with conditional regressive independence.

Less well-known and seldomly applied are tests of the equality of the differences between the expectations of the Y-variables in subpopulations. The hypothesis of equal differences between the expectations of the Y-variables in subpopulations is equivalent to the hypothesis of equal coefficients λ_i in subpopulations (see Theorem 2.2 and Corollary 2.11 below). This test has been proposed by Müller (1980). (See also Moosbrugger und Müller, 1982 and Moosbrugger, 1982). These tests do *not* rely on a conditional independence assumption such as the one needed for the derivation of the covariance structure.

In the following definition we refer to a probability space $(\Omega, \mathfrak{A}, P)$ representing a random experiment characterized by Equation 1 in which at least $m \geqq 2$ Y-variables are observed. Later, we show that in such an experiment the covariance of essentially τ-equivalent and conditionally regressively independent Y-variables yields the true score variance, and that an *empirically testable consequence* is the equality of the covariances of the Y-variables. However, such a test requires $m \geqq 3$ Y-variables unless further restrictions are introduced.

While these points are well-known, the treatment of the *existence* (or *representation*), *uniqueness*, and *meaningfulness* theorems seem to be new. These theorems not only clarify the logical structure of the model but also bridge the gap between classical theory of psychometric tests and representation theory of measurement.

2.1. *Definition (essentially τ-equivalent variables)*

The random variables $Y_1, \ldots, Y_m$ on a probability space $(\Omega, \mathfrak{A}, P)$ are called *essentially τ-equivalent* if and only if the following conditions hold:

(a) $(\Omega, \mathfrak{A}, P)$ is a probability space such that $\Omega = U \times A$.
(b) The projection $p_U: \Omega \to U$ is a random variable on $(\Omega, \mathfrak{A}, P)$.
(c) $Y_i, i \in I = \{1, \ldots, m\}$, are numerical random variables on $(\Omega, \mathfrak{A}, P)$ with finite and positive variances and covariances and $E(Y_i \mid p_U)$ denotes the p_U-conditional expectation of Y_i, $i \in I$.
(d) For each pair $(i, j) \in I \times I$, there is a real number λ_{ij} such that

$$E(Y_i \mid p_U) = \lambda_{ij} + E(Y_j \mid p_U). \tag{13}$$

Comments. According to Condition 2.1 (d), the variables $Y_1, \ldots, Y_m$ have regressions $E(Y_i \mid p_U) = \tau_i$ that are essentially equivalent, that is, the true score variables τ_i are equivalent up to a translation (i.e., up to an addition of a real constant; see Appendix, Note 6). In the following theorem (see Appendix, Note 7), a condition is formulated which is equivalent to Condition 2.1 (d).

2.2. *Theorem (existence)*

The random variables $Y_1, \ldots, Y_m$ are essentially τ-equivalent, if and only if Conditions 2.1 (a) to (c) hold as well as

(d′) there exist a (i.e., at least one) numerical random variable τ on $(\Omega, \mathfrak{A}, P)$ and real numbers λ_i such that

$$E(Y_i \mid p_U) = \lambda_i + \tau, \quad \text{for each} \quad i \in I. \tag{14}$$

Remarks.

(i) The variance of τ is finite because it is the same as the variance of $E(Y_i \mid p_U)$. The latter is finite because of Equation 11 and the assumption that the variances of the Y-variables are finite.

(ii) Condition (d′) is equivalent to:

(d″) there exist a numerical random variable τ on $(\Omega, \mathfrak{A}, P)$ and real numbers λ_i such that

$$Y_i = \lambda_i + \tau + \varepsilon_i, \tag{15}$$
$$\text{where} \quad \varepsilon_i := Y_i - E(Y_i \mid p_U), \quad \text{for each} \quad i \in I.$$

(iii) Condition (d′) is also equivalent to:

(d‴) There exist a function $\varphi: U \to \overline{\mathbb{R}}$ and real numbers λ_i such that

$$E(Y_i \mid p_U = u) = \lambda_i + \varphi(u), \quad \text{for each} \quad i \in I. \tag{16}$$

Comments. The function φ directly assigns a number to each observational unit $u \in U$. This number is identical with $\tau(\omega)$, if $\omega = (u, a)$. Hence, both τ and φ characterize the observational units (e.g., persons), whereas λ_i describes a property of the variable Y_i. The latent variable τ has the formal advantage over φ to be a random variable on $(\Omega, \mathfrak{A}, P)$, the same probability space on which Y_i and ε_i are random variables. Also observe that τ is the composition of φ with p_U, i.e. $\tau = \varphi(p_U)$.

Before we deal with the uniqueness problem, consider the following definition.

2.3. *Definition (model of essentially τ-equivalent variables)*

$\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is called a *model of essentially τ-equivalent variables* if and only if

$$\mathbf{E}(\mathbf{y} \mid p_U) := (E(Y_1 \mid p_U), \ldots, E(Y_m \mid p_U)), \tag{17}$$

Conditions 2.1 (a) to (c), and Condition 2.2 (d') hold.

Comments. In this definition, no new assumption are introduced. It just contains a convenient way to summarize the assumptions already discussed in the paragraphs above. The next theorem shows that neither τ (and therefore also neither φ) nor the coefficients λ_i are uniquely defined (see Appendix, Note 8).

2.4. *Theorem (uniqueness)*

(i) If $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables, and if for $\alpha \in \mathbb{R}$:

$$\tau' := \tau + \alpha \tag{18}$$

$$\boldsymbol{\lambda}' := (\lambda_1 - \alpha, \ldots, \lambda_m - \alpha), \tag{19}$$

then $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'\rangle$ is a model of essentially τ-equivalent variables, too.

(ii) If both $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ and $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'\rangle$ are models of essentially τ-equivalent variables, then there is an $\alpha \in \mathbb{R}$, such that Equations 18 and 19 hold.

Comments. The results above imply that the *differences* between the values of τ (and therefore between the values of φ) as well as between the coefficients λ_i are meaningful, that is, they are invariant under the admissible transformations, namely the translations. Hence, the variances σ_τ^2 and the reliabilities of the Y-variables are meaningful, too. In terms of Suppes and Zinnes (1963), the paragraphs above deal with a derived *difference scale* in the *narrow* sense (see Appendix, Note 9).

In the following corollary we also consider the functions $\varphi, \varphi' : U \to \overline{\mathbb{R}}$, which are related to τ and τ' by $\tau = \varphi(p_U)$ and $\tau' = \varphi'(p_U)$, respectively. Specifically, if $E(Y_i \mid p_U) = \lambda_i' + \tau'$, then

$$E(Y_i \mid p_U = u) = \lambda_i' + \varphi'(u) \tag{20}$$

holds for φ'.

2.5. *Corollary (meaningfulness)*

If $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ and $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'\rangle$ are models of essentially τ-equivalent variables, then

$$\tau(\omega_1) - \tau(\omega_2) = \tau'(\omega_1) - \tau'(\omega_2), \quad \omega_1, \omega_2 \in \Omega, \tag{21}$$

$$\varphi(u_1) - \varphi(u_2) = \varphi'(u_1) - \varphi'(u_2), \quad u_1, u_2 \in U, \tag{22}$$

$$\lambda_i - \lambda_j = \lambda_i' - \lambda_j', \quad i, j \in I, \tag{23}$$

$$\sigma_\tau^2 = \sigma_{\tau'}^2, \tag{24}$$

$$\sigma_\tau^2 / \sigma_{Y_i}^2 = \sigma_{\tau'}^2 / \sigma_{Y_i}^2, \quad i \in I. \tag{25}$$

Comments. The proof of this corollary is simple and left to the reader. Note that this is not a complete list of all meaningful terms. In order to prove the theorems above, we do not need any assumption on the independence of the errors among each other. However, such an additional assumption will now be introduced in order to derive the next theorem on the structure of the covariance matrix implied by the model. This theorem is also the basis for solving the problem of the identification of the variances σ_τ^2. Note that all variables τ_i, $i \in I$, have identical variances if Condition 2.2 (d') holds, that is,

$$\mathrm{Var}(\tau_i) = \sigma_\tau^2, \quad i \in I. \tag{26}$$

Also remember that τ and its associated vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ are not uniquely defined. According to Theorem 2.3, there is a whole family of pairs $(\tau, \boldsymbol{\lambda})$, each pair of members being related to each other by a translation. In the following sections, $(\tau, \boldsymbol{\lambda})$ denotes an arbitrary member of that family. Before we formulate the Covariance Structure Theorem, let us introduce the assumption of conditional regressive independence.

2.6. *Definition (model of essentially τ-equivalent variables with conditional regressive independence)*

$\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is called a *model of essentially τ-equivalent variables with conditional regressive independence* if and only if Conditions 2.1 (a) to (d) hold and

$$E(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) = E(Y_i \mid p_U), \quad i \in I. \tag{27}$$

Comments. Equation 27 implies, among other things, uncorrelated residuals $\varepsilon_i := Y_i - E(Y_i \mid p_U)$. In fact, this is all we need in the present context.

Hence we might replace this equation by the weaker assumption of uncorrelated residuals. However, Equation 27 gives more intuitive insight (see Appendix, Note 10). Some implications of Definition 2.6 are collected in the following theorem (see Appendix, Note 11).

2.7. Theorem (covariance structure)

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables with conditional regressive independence, then, for $i, j \in I$:

$$\operatorname{Cov}(Y_i, Y_j) = \begin{cases} \sigma_\tau^2, & i \neq j, \\ \sigma_\tau^2 + \sigma_{\varepsilon_i}^2, & i = j, \end{cases} \tag{28}$$

$$\operatorname{Cov}(\tau, \varepsilon_i) = 0\,, \tag{29}$$

$$\operatorname{Cov}(\varepsilon_i, \varepsilon_j) = 0, \qquad i \neq j. \tag{30}$$

Comments. Formulated in terms of factor analysis, Definition 2.6 implies a one-factor model with factor loadings equal to 1 (see Figure 1). Note that this theorem is a first proposition dealing with the empirical testability of the model. If there are $m \geqq 3$ variables, the model of essentially τ-equivalent variables with conditional regressive independence can be tested empirically because Equation 28 then implies the equality of the covariances $\operatorname{Cov}(Y_i, Y_j)$, $i \neq j$. If, for instance, $m = 3$, the covariance matrix Σ has the following structure:

$$\Sigma = \begin{pmatrix} \sigma_\tau^2 + \sigma_{\varepsilon_1}^2 & \sigma_\tau^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 + \sigma_{\varepsilon_2}^2 & \sigma_\tau^2 \\ \sigma_\tau^2 & \sigma_\tau^2 & \sigma_\tau^2 + \sigma_{\varepsilon_3}^2 \end{pmatrix}. \tag{30}$$

Statistical tests of this kind of hypotheses are implemented in well-known computer programs such as LISREL (Jöreskog & Sörbom, 1984) or EQS (Bentler, 1984).

The next theorem will show *that* and *how* the variances σ_τ^2 and $\sigma_{\varepsilon_i}^2$ can be determined by the variances and covariances of the variables Y_i. A proof that the unknown parameters of a model can be computed from known or estimable parameters such as variances and covariances of the observable variables means to solve the *problem of identifiability*. Its solution is the prerequisite for dealing with statistical problems such as estimation.

2.8. Corollary (identification of variances)

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables with conditional regressive independence, then, for $i, j \in I$:

$$\sigma_\tau^2 = \operatorname{Cov}(Y_i, Y_j), \qquad i \neq j \tag{32}$$

$$\sigma_{\varepsilon_i}^2 = \operatorname{Var}(Y_i) - \operatorname{Cov}(Y_i, Y_j), \quad i \neq j. \tag{33}$$

Comments. This corollary (see Appendix, Note 12) is a constructive one in the sense that it tells us how to determine the theoretical variances. Equation 32 shows that the variance of τ and the variance of the error variables (see Equation 33) can be determined by the variances and covariances of essentially τ-equivalent variables which are conditionally regressively independent.

The *proportion* of variance of Y_i determined by τ_i, the reliability coefficient, is another meaningful parameter that can be determined from the variances and covariances of the Y-variables. The relevant formulas are displayed in the following corollary.

2.9. Corollary (identification of the reliability coefficient)

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables with conditional regressive independence, then, for $i, j \in I$:

$$\operatorname{Rel}(Y_i) = \frac{\operatorname{Var}(\tau_i)}{\operatorname{Var}(Y_i)} = \frac{\operatorname{Cov}(Y_i, Y_j)}{\operatorname{Var}(Y_i)}, \quad i \neq j. \tag{34}$$

Furthermore, if $\operatorname{Var}(Y_i) = \sigma_Y^2$, for all $i \in I$, then:

$$\operatorname{Rel}(Y_i) = \frac{\operatorname{Var}(\tau_i)}{\operatorname{Var}(Y_i)} = \operatorname{Cor}(Y_i, Y_j), \quad i \neq j,\ i, j \in I, \tag{35}$$

where Cor(,) denotes the correlation.

Comments. Using Corollary 2.8, the proof is simple and left to the reader. If, in an application, the coefficient $\operatorname{Rel}(Y_i)$ is greater than zero, one might talk about the nontrivial existence of a latent variable.

We now turn to another empirical testable consequence of the model of essentially τ-equivalent variables, namely the equality of the differences between the expectations of the Y-variables in different subpopulations. This consequence does not rely on any independence assumption.

To be more precise, consider the event

$$\Omega^{(1)}: = \{\omega \in \Omega: p_U(\omega) = u \in U^{(1)}\}, \quad U^{(1)} \subset U,\ P(\Omega^{(1)}) > 0\,, \tag{36}$$

that the observational unit sampled belongs to subpopulation $U^{(1)}$. For such an event there always exists the conditional probability measure $P^{(1)}: \mathfrak{A} \to [0, 1]$ on $(\Omega, \mathfrak{A})$ defined by

$$P^{(1)}(A): = P(A \mid \Omega^{(1)}) \qquad \text{for all } A \in \mathfrak{A}. \tag{37}$$

Passing from the probability space $(\Omega, \mathfrak{A}, P)$ to $(\Omega, \mathfrak{A}, P^{(1)})$ simply means to pass from the total population U to the subpopulation $U^{(1)}$. Keeping

$(\Omega, \mathfrak{A})$ as a measurable space has mathematical advantages, because the functions $Y_i: \Omega \to \overline{\mathbb{R}}$, $\tau: \Omega \to \mathbb{R}$ etc. are random variables both on the probability space $(\Omega, \mathfrak{A}, P^{(1)})$ *and on* $(\Omega, \mathfrak{A}, P)$. However, the functions Y_i etc. have different distributions with respect to the probability measures P and $P^{(1)}$.

2.10. Theorem (one subpopulation)

(i) If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables and $P^{(1)}$ is defined by Equation 37, then $\mathcal{M}^{(1)}: = \langle(\Omega, \mathfrak{A}, P^{(1)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent, variables too.

(ii) If $\mathcal{M}$ is a model of essentially τ-equivalent variables with conditional regressive independence, then $\mathcal{M}^{(1)}$ is a model of essentially τ-equivalent variables with conditional regressive independence, too.

Comments. (See Appendix, Note 13). Using identical symbols, especially $\mathbf{y}: = (Y_1, \ldots, Y_m)$ and τ in both $\mathcal{M}$ and $\mathcal{M}^{(1)}$ means that $\mathbf{y}$ and τ are identical $\mathfrak{A}$-measurable functions in the two models $\mathcal{M}$ and $\mathcal{M}^{(1)}$. (Note again, however, that their distributions are not identical in the two models.) Correspondingly, the real vector $\boldsymbol{\lambda}$ is identical in both models.

The crucial proposition of Theorem 2.10 is that (i) the τ-*values* (and therefore the φ-*values*) as well as the λ-constants are the same in the total population and in a subpopulation, and that (ii) in the model with conditional regressive independence, the equality restrictions on the covariances (see Eqs. 28 and 31) still hold in the subpopulation. Again, this implication is empirically testable. Note, however, that the covariances in the subpopulations do not have to be identical to the covariances in the total population.

Theorem 2.10 implies that the vector $\boldsymbol{\lambda}$ will be identical in two different subpopulations $U^{(1)}$ and $U^{(2)} \subset U$. If $P^{(2)}$ denotes the probability measure pertaining to the second subpopulation, we may formulate the following corollary:

2.11. Corollary (two subpopulations)

(i) If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ is a model of essentially τ-equivalent variables, then $\mathcal{M}^{(1)}: = \langle(\Omega, \mathfrak{A}, P^{(1)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ and $\mathcal{M}^{(2)}: = \langle(\Omega, \mathfrak{A}, P^{(2)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}\rangle$ are models of essentially τ-equivalent variables, too, and for $i, j \in I$:

$$E^{(1)}(Y_i) - E^{(1)}(Y_j) = E^{(2)}(Y_i) - E^{(2)}(Y_j), \tag{38}$$

$$E^{(1)}(Y_i - Y_j) = E^{(2)}(Y_i - Y_j). \tag{39}$$

(ii) If $\mathcal{M}$ is a model of essentially τ-equivalent variables with conditional regressive independence, then $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ are models of essentially τ-equivalent variables with conditional regressive independence, too.

Comments. The proof (see Appendix, Note 14) is essentially based on the additive decomposition of the Y-variables (see Equation 15). Equation 39 – which is equivalent with Equation 38 – is a hypothesis about $m - 1$ independent equalities of the expectations of (difference-) variables. If $m = 2$, and the additional assumptions of normality and homogeneity of variances can be made, the statistical standard procedure for such a hypothesis is the t-test for independent groups. If $m > 2$, the multivariate t-test applies, which can be computed, for instance, by the SPSS-procedure MANOVA.

Hence, there are several procedures for the empirical test of the model of essentially τ-equivalent variables: (a) The test fo the covariance structure in the total population. (b) The test of the covariance structure in the subpopulations. (c) The test of the equality of the differences of the expectations of the Y-variables in subpopulations. Each of these tests can lead to a rejection of the model. Note, however, that points (a) and (b) rest on the additional assumption of conditional regressive independence (or at least, of uncorrelated error variables).

3. Example: Essentially τ-equivalent state anxiety test halves

The statistical test based on the theorems above will now be illustrated by an example from state-trait anxiety research. 179 students of the University of Trier were assessed twice with the State-Trait Anxiety Inventory (STAI; Laux, Glanzmann, Schaffner, & Spielberger, 1981) with two months between the two occasions of measurement. Table 1 displays the sample means, the covariance and correlation matrices between the test halves of the State Anxiety Inventory. Note that in the following paragraphs, "test halves" is an abbreviation for those random variables, the values of which are the sum scores of the persons on the items defining the test halves.

Table 1: Covariances, correlations and means of the state anxiety test halves in the total sample.

	SA_{11}	SA_{21}	SA_{12}	SA_{22}
SA_{11}	24.670	0.879	0.399	0.439
SA_{21}	21.895	25.135	0.406	0.471
SA_{12}	10.353	10.624	27.239	0.904
SA_{22}	11.665	12.636	25.258	28.683
$\overline{SA}_{ik}$	20.296	22.011	21.475	22.860

Note. SA_{ik}: State anxiety, ith test half, kth occasion. $\overline{SA}_{ik}$: means of the variables. The first test half consists of the items 2, 3, 4, 5, 7, 8, 12, 14, 16, 17, the second of the items 1, 6, 9, 10, 11, 13, 15, 18, 19, 20. (See Appendix, Note 15.)

3.1. Testing the covariance structure within one occasion

First, we investigate whether or not the state anxiety test halves SA_{11} and SA_{21} assessed at the first occasion of measurement, can be assumed to be essentially τ-equivalent, to be conditionally regressively independent, and to have equal variances. As shown above, the assumptions of essentially τ-equivalent variables with conditional regressive independence imply a special factor model (Lawley & Maxwell, 1971). Hence, the program LISREL VI (Jöreskog & Sörbom, 1984) can be used for data analysis. Of course, additional assumptions concerning the sampling model (including distributional assumptions) have to be made if maximum likelihood (ML) estimation and the likelihood ratio test are to be applied. Details can be found in the LISREL manual. The results of the analysis using ML-estimation are displayed in Figure 2.

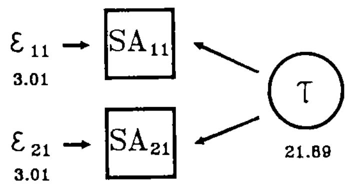

Figure 2: Arrow diagram of the factor model implied by the model of essentially τ-equivalent state anxiety test halves with conditional regressive independence for the first occasion of measurement. The latent variable τ is referred to as the latent state anxiety variable pertaining to the first occasion. The numbers are the estimated variances of the corresponding variables. The variances of the error variables are restricted to be equal. Criteria of model fit: (according to LISREL VI): $\chi_1^2 = 0.07$, $p = 0.79$; goodness of fit (adj.) = 0.99, absolutely largest standardized residual = 0.09.

The test criteria given in Figure 2 reveal that there is no reason to reject the model of essentially τ-equivalent state anxiety test halves with conditional regressive independence and equal error variances for the first occasion. Furthermore, if we compare the variances and covariances estimated under the model assumptions

$$\hat{\mathrm{V}}\mathrm{ar}(SA_{11}) = \hat{\mathrm{V}}\mathrm{ar}(SA_{21}) = 24.902 \text{ und } \hat{\mathrm{C}}\mathrm{ov}(SA_{11}, SA_{21}) = 21.895$$

with the empirical variances and covariances (see Table 1), we find a very good model fit. Hence the reliability of the test halves may be estimated by

$$\hat{\mathrm{C}}\mathrm{ov}(SA_{11}, SA_{21}) / \hat{\mathrm{V}}\mathrm{ar}(SA_{11}) = 21.895/24.902 \approx 0.88 .$$

The latent state anxiety variable determines about 88 % of the variance of a state anxiety test half. The remaining 12 % are the proportion of the error variances.

The test of significance above should be interpreted with some caution. First, not rejecting the null hypothesis should not be confused with accepting the null hypothesis. Second, the test is – in this case – only a test of the equality of the variances of the two variables. A more powerful test of such a model needs more than two variables which are assumed to be essentially τ-equivalent (see Section 3.3).

3.2. Testing the equality of the differences of the expectations of the Y-variables in two subpopulations

According to theorem 2.11, the model of essentially τ-equivalent variables implies the equality of the expectations

$$E^{(1)}(Y_i - Y_j) = E^{(2)}(Y_i - Y_j), \quad i, j \in I,$$

in two subpopulations $U^{(1)}$ and $U^{(2)}$ of U. If we test this hypothesis for the two state anxiety test halves assessed at the first occasion of measurement, then $m = 2$, and we test the equality of the expectations

$$E^{(1)}(SA_{11} - SA_{21}) = E^{(2)}(SA_{11} - SA_{21}) \tag{40}$$

in two subpopulations $U^{(1)}$ and $U^{(2)}$. Possible subpopulations are *men* vs. *women* or two *different fields of study* (i.e., psychology vs. others). Table 2 displays means, covariance, and correlation matrices of the state anxiety test halves in the male and femal subsamples.

Table 2: Covariances, correlations, and means of the state anxiety test halves in the two subsamples.

	Men (N = 89)				Women (N = 90)			
	SA_{11}	SA_{21}	SA_{12}	SA_{22}	SA_{11}	SA_{21}	SA_{12}	SA_{22}
SA_{11}	26.864	0.889	0.519	0.585	22.771	0.875	0.264	0.278
SA_{21}	22.782	24.440	0.547	0.634	21.220	25.813	0.274	0.319
SA_{12}	14.562	14.647	29.306	0.900	6.345	7.002	25.309	0.908
SA_{22}	16.575	17.144	26.638	29.919	6.971	8.513	24.021	27.256
$\bar{SA}_{ik}$	20.236	21.629	21.786	23.112	20.355	22.389	21.667	22.611

Note. See the note of Table 1.

Note that Equation 40 corresponds to the comparison of the differences

$$20.236 - 21.629 = -1.393 \text{ und } 20.355 - 22.389 = -2.034$$

(see Table 2). The corresponding t-value with 177 degrees of freedom is 1.756 ($p = 0.081$) and is not significant on the 5%-level. Hence, this test, too, gives no reason to reject the hypothesis that the test halves SA_{11} and SA_{21} are essentially τ-equivalent.

This kind of model tests could be applied in different subsamples. Also, the test of the covariance structure can be carried through in each of the two subsamples (see Theorem 2.11). The data necessary for such a test are displayed in Table 2. Instead of these tests, we now investigate if the model of essentially τ-equivalent variables with conditional regressive independence can be postulated for all four test halves.

3.3. *Testing the covariance structure across the two occasions*

Essentially τ-equivalent test variables may not only be constructed by parallel test forms. They may also arise from repeated assessments of the subjects with the same test. However, this requires temporal stability in the sense that the latent variable does not change between occasions of measurement, or at least that it only changes by a constant which is identical for all persons in the population.

We now test the assumption that the model of essentially τ-equivalent variables with conditional regressive independence holds for all four state anxiety test halves (see Figure 3).

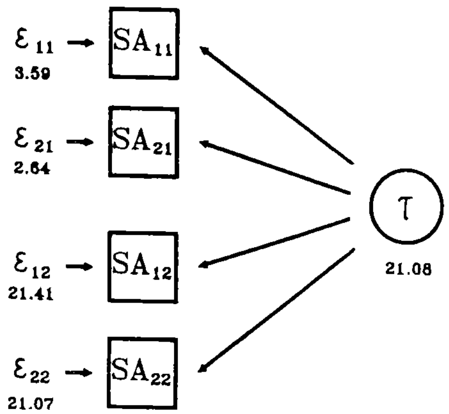

Figure 3: Arrow diagram of the factor model implied by the model of essentially τ-equivalent state anxiety test halves with conditional regressive independence for the two occasions of measurement. The numbers are the estimated variances of the corresponding variables. The variances of the error variables are not restricted to be equal. Criteria of model fit: (according to LISREL VI): $\chi^2_5 = 315.34$, $p = 0.00$; goodness of fit (adj.) = 0.20, absolutely largest standardized residual = 3.00.

The χ^2-value and the other criteria reveal that this model does *not* fit at all, and therefore it has to be rejected. Hence, it was demonstrated how to test the model of essentially τ-equivalent variables (with conditional regressive independence). While the tests within the first occasion of measurement did not give any reason for a rejection of the model, the analysis of both occasions lead to a rejection of the hypothesis that the model of essentially τ-equivalent variables with conditional regressive independence holds for all four test halves. This is in accordance with the assumption that the test halves measure *states* which can change between different occasions of measurement.

A better model for these data is displayed in Figure 4. This model assumes essentially τ-equivalent test halves within each occasion of measurement and uncorrelated errors between all four test halves considered. According to this model, there is a separate latent state anxiety variable for each occasion of measurement. Hence, this model may be called a model of essentially τ_k-equivalent variables (with conditional regressive independence), where k indicates an occasion of measurement. These models are natural generalizations of the models treated in this paper (see Steyer, 1987; Majcen, Steyer, & Schwenkmezger, 1988; Steyer, Majcen, Schwenkmezger, & Buchner, 1988).

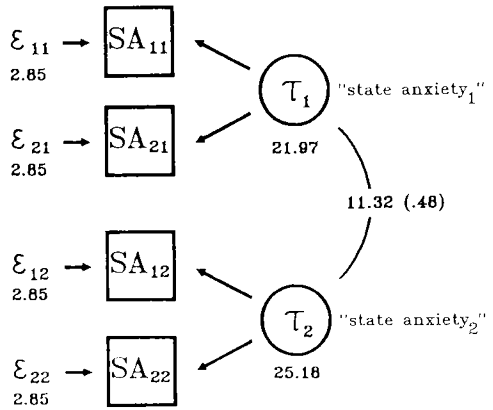

Figure 4: Arrow diagram of the factor model implied by the model of essentially τ_k-equivalent state anxiety test halves with conditional regressive independence. The estimated correlation between the two latent state anxiety variables is 0.48. The variances of the error variables are fixed to be equal. Criteria of model fit (according to LISREL VI): $\chi^2_6 = 7.54$; $p = 0.27$; goodness of fit (adj.) = 0.97; absolutely largest standardized residual = 0.61.

4. Congeneric variables

We now replace Condition 2.1 (d) by a less restrictive condition resulting in the model of τ-congeneric variables (cf. Jöreskog, 1971). Again, we treat the representation, uniqueness, meaningfulness, identifiability, and testability theorems. Figure 5 shows the factor model implied by the assumptions of this model if conditional regressive independence holds (see the Covariance Structure Theorem).

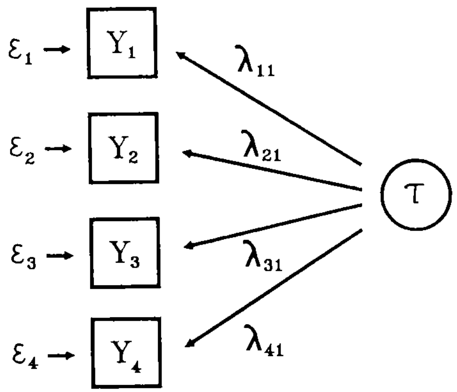

Figure 5: Arrow diagram of the factor model implied by the model of $m = 4$ τ-congeneric variables with conditional regressive independence.

While the identification of the parameters requires only $m \geqq 3$ Y-variables, a test of the covariance structure needs $m \geqq 4$ of them, unless further restrictions (such as the equality of some error variances or the equality of some loadings) are imposed. Recall that the model of essentially τ-equivalent variables with conditional regressive independence only needed $m \geqq 2$ Y-variables for parameter identification and $m \geqq 3$ variables for a test of the covariance structure. Obviously, stronger assumptions require fewer observable variables and vice versa.

4.1. Definition (τ-congeneric variables)

The random variables $Y_1, \ldots, Y_m$ are called *τ-congeneric* if and only if Conditions 2.1 (a) to (c) hold and:

(d) For each pair $(i, j) \in I \times I$, there are $\lambda_{ij0}, \lambda_{ij1} \in \mathbb{R}$, $\lambda_{ij1} > 0$, such that

$$E(Y_i \mid p_U) = \lambda_{ij0} + \lambda_{ij1} \cdot E(Y_j \mid p_U). \tag{41}$$

Comments. According to Condition 4.1 (d), the true-score variables $E(Y_i \mid p_U)$ are positive linear functions of each other. This implies the existence of a random variable τ and of real numbers $\lambda_{i0}, \lambda_{i1}$ such that $E(Y_i \mid p_U) = \lambda_{i0} + \lambda_{i1}\tau$, where $\lambda_{i1} > 0$, for each $i \in I$. This variable τ is uniquely defined up to a positive linear transformation, that is, if τ fulfills the requirements then also $\tau' := \alpha + \beta\tau$ will fulfill it, where $\alpha, \beta \in \mathbb{R}$, $\beta > 0$. Also, the coefficients λ_{i0} and λ_{i1} are not uniquely defined. This will be formulated more precisely in the following theorems (see Appendix, Note 16).

4.2. Theorem (existence)

The random variables $Y_1, \ldots, Y_m$ are called *τ-congeneric* if and only if Conditions 2.1 (a) to (c) hold as well as:

(d') there exist a numerical random variable τ on $(\Omega, \mathfrak{A}, P)$ and $\lambda_{i0}, \lambda_{i1} \in \mathbb{R}$, $\lambda_{i1} > 0$, such that

$$E(Y_i \mid p_U) = \lambda_{i0} + \lambda_{i1}\tau, \quad i \in I. \tag{42}$$

Remarks. The variance of τ is finite following from Equation 11 and the assumption that the variances of the Y-variables are finite. According to Equation 42, the model of τ-congeneric variables additively decomposes a true score variable $\tau_i := E(Y_i \mid p_U)$ into a variable $\lambda_{i1}\tau$ and a parameter λ_{i0}. Equation 42 implies that there is a function $\varphi: U \to \mathbb{R}$ such that $\tau = \varphi(p_U)$. Hence, a value of the variable τ is a parameter characterizing the *observational unit*, whereas the constants λ_{i0} and λ_{i1} characterize the *variable* Y_i. This may be considered an example of *nonadditive conjoint* *measurement;* in this model, we simultaneously assign a real number to each observational unit u and two real numbers to each variable Y_i. Before dealing with the problem of the uniqueness of these assignments, consider the following definition, which justifies the name *τ-congenerity*.

4.3. *Definition (model of τ-congeneric variables)*

$\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is called a *model of τ-congeneric variables* if and only if Conditions 2.1. (a) to (c) and Condition 4.2 (d') hold, and $\boldsymbol{\lambda}_0 := (\lambda_{10}, \ldots, \lambda_{m0})$, $\boldsymbol{\lambda}_1 := (\lambda_{11}, \ldots, \lambda_{m1})$.

Comments. As already mentioned above, neither τ, nor the coefficients λ_{i0} and λ_{i1} are uniquely defined in this model. There is a whole family of triples $(\tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1)$ fulfilling the requirements of Definition 4.3, and $(\tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1)$ will denote *any* member of this family. In other words, the triple $(\tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1)$ may be arbitrarily chosen; the only requirement is that Equation 42 holds for its components. The next theorem deals with the degree of uniqueness of the latent variable τ and the coefficients λ_{i0} and λ_{i1}.

4.4. *Theorem (uniqueness)*

(i) If $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables and for $\alpha, \beta \in \mathbb{R}$, $\beta > 0$:

$$\tau' := \alpha + \beta \cdot \tau, \tag{43}$$

$$\boldsymbol{\lambda}'_0 := (\lambda_{10} - \lambda_{11}\alpha/\beta, \ldots, \lambda_{m0} - \lambda_{m1}\alpha/\beta), \tag{44}$$

$$\boldsymbol{\lambda}'_1 := (\lambda_{11}/\beta, \ldots, \lambda_{m1}/\beta), \tag{45}$$

then $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'_0, \boldsymbol{\lambda}'_1\rangle$ is a model of τ-congeneric variables, too.

(ii) If both $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ and $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'_0, \boldsymbol{\lambda}'_1\rangle$ are models of τ-congeneric variables then there are $\alpha, \beta \in \mathbb{R}$, $\beta > 0$, such that Equations 43 to 45 hold.

Admissible transformations. According to this theorem (see Appendix, Note 17), the latent variable τ is uniquely defined up to a *positive linear transformation* and the coefficients λ_{i1} are uniquely defined up to a *similarity transformation* (i.e., a multiplication by a positive real number). Hence, the model of τ-congeneric variables defines a family of triples $(\tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1)$. If both $(\tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1)$ and $(\tau', \boldsymbol{\lambda}'_0, \boldsymbol{\lambda}'_1)$ are members of this family then the components of these triples are related by the transformations described above.

Note that the transformations are closely interconnected. Transforming τ to $\tau' = \alpha + \beta \cdot \tau$ implies a transformation of both $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_1$ (see Eqs. 44 and 45). Transforming τ to $\tau' = 0 + \beta \cdot \tau$ only implies a transformation of $\boldsymbol{\lambda}_1$. Transforming τ to $\tau' = \alpha + 1 \cdot \tau$ only implies a transformation of $\boldsymbol{\lambda}_0$.

In the following corollary we also consider the functions $\varphi, \varphi' : U \to \mathbb{R}$, which are related to τ and τ' by

$$\tau = \varphi(p_U) \quad \text{and} \quad \tau' = \varphi'(p_U), \tag{46}$$

respectively. Specifically, if $E(Y_i \mid p_U) = \lambda'_{i0} + \lambda'_{i1}\tau'$, then

$$E(Y_i \mid p_U = U) = \lambda'_{i0} + \lambda'_{i1}\varphi'(u) \tag{47}$$

holds for φ'. (The proof is simple algebra and left to the reader.)

4.5. *Corollary (meaningfulness)*

If $\mathcal{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ and $\mathcal{M}' := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau', \boldsymbol{\lambda}'_0, \boldsymbol{\lambda}'_1\rangle$ are models of τ-congeneric variables, then

$$\frac{\tau(\omega_1) - \tau(\omega_2)}{\tau(\omega_3) - \tau(\omega_4)} = \frac{\tau'(\omega_1) - \tau'(\omega_2)}{\tau'(\omega_3) - \tau'(\omega_4)}, \quad \omega_1, \ldots, \omega_4 \in \Omega, \tag{48}$$

$$\frac{\varphi(u_1) - \varphi(u_2)}{\varphi(u_3) - \varphi(u_4)} = \frac{\varphi'(u_1) - \varphi'(u_2)}{\varphi'(u_3) - \varphi'(u_4)}, \quad u_1, \ldots, u_4 \in U, \tag{49}$$

$$\frac{\lambda_{i1}}{\lambda_{j1}} = \frac{\lambda'_{i1}}{\lambda'_{j1}}, \quad i, j \in I, \tag{50}$$

$$\lambda_{i1}^2 \sigma_\tau^2 = \lambda_{i1}'^2 \sigma_{\tau'}^2, \tag{51}$$

$$\lambda_{i1}^2 \sigma_\tau^2 / \sigma_{Y_i}^2 = \lambda_{i1}'^2 \sigma_{\tau'}^2 / \sigma_{Y_i}^2, \quad i \in I. \tag{52}$$

Comments. According to this corollary, ratios of differences between values of τ and φ are invariant under the admissible (i.e., positive linear) transformations. Because of

$$\mathrm{Var}(E(Y_i \mid p_U)) = \mathrm{Var}(\tau_i) = \lambda_{i1}^2 \sigma_\tau^2, \quad i \in I, \tag{53}$$

it follows from Equation 51, that the variances of the regressions $E(Y_i \mid p_U)$ are meaningful, and because of

$$\mathrm{Rel}(Y_i) = \lambda_{i1}^2 \sigma_\tau^2 / \sigma_{Y_i}^2, \quad i \in I, \tag{54}$$

it follows from Equation 52 that the reliability coefficients $\mathrm{Rel}(Y_i)$ are meaningful parameters. Note that this is not a complete list of all meaningful terms.

Our next theorem will formulate one of the empirically testable consequences of the model of τ-congeneric variables. However, this implication requires the additional assumption of conditional regressive independence which was not necessary to derive the propositions above.

4.6. *Definition (model of τ-congeneric variables with conditional regressive independence)*

$\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is called a *model of τ-congeneric variables with conditional regressive independence* if and only if Conditions 2.1 (a) to (c), Condition 4.2 (d′), and Equation 27 hold.

Remarks. Note that this definition implies that the covariances between the Y-variables are positive as well as the coefficients λ_{i1}. The implications of the model defined above for the structure of the covariance matrix are treated in the following theorem (see Appendix, Note 18).

4.7. *Theorem (covariance structure)*

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables with conditional regressive independence, then, for $i, j \in I$:

$$\mathrm{Cov}(Y_i, Y_j) = \begin{cases} \lambda_{i1}\sigma_\tau^2\lambda_{j1}, & i \neq j, \\ \lambda_{i1}\sigma_\tau^2\lambda_{j1} + \sigma_{\varepsilon_i}^2, & i = j, \end{cases} \tag{55}$$

$$\mathrm{Cov}(\tau, \varepsilon_i) = 0\,, \tag{56}$$

$$\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0, \qquad i \neq j. \tag{57}$$

Comments. Formulated in terms of factor analysis, Definition 4.6 implies a one-factor model with possibly unequal factor loadings (see Figure 5). Since τ is determined uniquely only up to a positive linear transformation, the variance σ_τ^2 is not a meaningful parameter per se. Also, the coefficients λ_{i1} are no meaningful parameters per se, because these coefficients are uniquely defined only up to a similarity transformation. However, the products $\lambda_{i1}^2\sigma_\tau^2$ are meaningful, because they are invariant under the admissible transformations of the coefficients λ_{i0}, λ_{i1}, and the latent variable τ. The next corollary deals with the identification of the products $\lambda_{i1}^2\sigma_\tau^2$ (see Appendix, Note 19.

4.8. *Theorem (identification of the products $\lambda_{i1}^2\sigma_\tau^2$ and the error variances)*

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables with conditional regressive independence, then, for $i, j, k \in I$:

$$\lambda_{i1}^2\sigma_\tau^2 = \frac{\mathrm{Cov}(Y_i, Y_j)\cdot\mathrm{Cov}(Y_i, Y_k)}{\mathrm{Cov}(Y_j, Y_k)}, \quad i \neq j, i \neq k, j \neq k, \tag{58}$$

$$\sigma_{\varepsilon_i}^2 = \mathrm{Var}(Y_i) - \mathrm{Var}(\tau_i) = \mathrm{Var}(Y_i) - \lambda_{i1}^2\sigma_\tau^2\,. \tag{59}$$

Comments. Equation 58 shows that, for given variables $Y_1, \ldots, Y_m$ (with a given common distribution), the product $\lambda_{i1}^2\sigma_\tau^2$ is a given positive real number, because Definition 4.6 implies that the covariances of the Y-variables and the coefficients λ_{i1} are positive. Hence, it follows from Equation 58 that if σ_τ^2 is fixed to an arbitrary positive real number (e.g., $\sigma_\tau^2 = 1$), then each coefficient λ_{i1}, $i \in I$, may be computed from Equation 58. Vice versa, if one of the coefficients λ_{i1} is fixed to an arbitrary positive real number (e.g., $\lambda_{11} = 1$), then also σ_τ^2 may be computed from Equation 58). In the case $\lambda_{11} = 1$, the other coefficients λ_{j1}, $1 \neq j \in I$, may be computed from

$$\lambda_{j1} = +\left(\frac{\mathrm{Cov}(Y_j, Y_k)\cdot\mathrm{Cov}(Y_j, Y_l)}{\mathrm{Cov}(Y_k, Y_l)\cdot\sigma_\tau^2}\right)^{1/2}, \quad j \neq k, j \neq l, k \neq l, \ j, k, l \neq 1\,. \tag{60}$$

Comments. As already mentioned above, the variance of τ is not meaningful and not of interest per se. However, the variances of the variables $\tau_i = E(Y_i \mid p_U)$ *are* meaningful (see Eqs. 51 and 53). Furthermore, the proportion of variance of Y_i determined by τ_i, the reliability coefficient, is another meaningful parameter that does not depend on the specific choice of τ (see Eqs. 52 and 54). The relevant formulas for the identification of the reliability coefficient are given in the following theorem.

4.9. *Theorem (identification of the reliability coefficient)*

If $\mathcal{M}: = \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables with conditional regressive independence, then, for $i, j, k \in I$:

$$\mathrm{Rel}(Y_i) = \frac{\mathrm{Var}(\tau_i)}{\mathrm{Var}(Y_i)} = \frac{\mathrm{Cov}(Y_i, Y_j)\cdot\mathrm{Cov}(Y_i, Y_k)}{\mathrm{Cov}(Y_j, Y_k)\cdot\mathrm{Var}(Y_i)}, \quad i \neq j, i \neq k, j \neq k\,. \tag{61}$$

Furthermore, if $\mathrm{Var}(Y_i) = \sigma_Y^2$, for all $i \in I$, then, for $i, j, k \in I$:

$$\mathrm{Rel}(Y_i) = \frac{\mathrm{Var}(\tau_i)}{\mathrm{Var}(Y_i)} = \frac{\mathrm{Cor}(Y_i, Y_j)\cdot\mathrm{Cor}(Y_i, Y_k)}{\mathrm{Cor}(Y_j, Y_k)}, \quad i \neq j, i \neq k, j \neq k\,. \tag{62}$$

Comments. The proof is well-known (see, e.g., Lord & Novick, 1968, p. 218) and simple if Equation 55 and the rules of computation for covariances are used. Obviously, we need at least 3 Y-variables to identify the reliability coefficient in the model of τ-congeneric variables with conditional regressive independence. In the model of essentially τ-equivalent variables with conditional regressive independence 2 Y-variables were sufficient.

Testability. Recall that the model of essentially τ-equivalent variables with conditional regressive independence implies the equality of the covariance $\mathrm{Cov}(Y_i, Y_j)$ which is empirically testable if $m \geq 3$ (see Equations 28 and 31). This property does not hold any more for the model of τ-congeneric variables. However, Equation 55 puts a weaker restriction on the covariances if there are at least $m \geqq 4$ variables. According to Equation 55, if there are $m = 4$ variables Y_i, their $m \cdot (m-1)/2 = 6$ covariances can be computed from four parameters. If σ_τ^2 is fixed to 1 (see last three paragraphs), these four parameters are the factor loadings λ_{11} to λ_{41}. If, however, λ_{11} is fixed to 1, these four parameters are the factor loadings λ_{21} to λ_{41} and σ_τ^2. Tests of this kind of restrictions on the covariance matrix are implemented in the programs on simultaneous equation models mentioned before. If, however, there are only $m = 3$ variables Y_i, then there are only $m \cdot (m-1)/2 = 3$ covariances to be computed from 3 factor loadings (if σ_τ^2 is fixed to 1). In this case the model of τ-congeneric variables with conditional regressive independence does not impose any empirically testable restrictions on the covariance matrix.

Implications concerning subpopulations. We now turn to other empirical testable consequences of the model of τ-congeneric variables. Unfortunately, there seems to be no analogue to the equality of the differences between the expectations of the Y-variables in different subpopulations, which has been shown to be implied by the model of essentially τ-equivalent variables (see Cor. 2.11). In the following theorems, we make use of the notation and concepts introduced in the corresponding parts of Section 2 (see Appendix, Note 20).

4.10. *Theorem (one subpopulation)*

(i) If $\mathscr{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables and $P^{(1)}$ is defined by Equation 36, then $\mathscr{M}^{(1)} := \langle(\Omega, \mathfrak{A}, P^{(1)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables, too.

(ii) If $\mathscr{M}$ is a model of τ-congeneric variables with conditional regressive independence, then $\mathscr{M}^{(1)}$ is a model of τ-congeneric variables with conditional regressive independence, too.

Comments. Again, using identical symbols, such as $\mathbf{y} := (Y_1, \ldots, Y_m)$ and τ in both $\mathscr{M}$ and $\mathscr{M}^{(1)}$ means that $\mathbf{y}$ and τ are identical $\mathfrak{A}$-measurable functions in the two models $\mathscr{M}$ and $\mathscr{M}^{(1)}$. (Note again, however, that their distributions are not identical in the two models.) Correspondingly, the real vectors $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_1$ are identical in both models. Furthermore, the *values* of τ (and therefore the *values* of φ) in the total population and in a subpopulation are identical. According to proposition (ii), the conditional regressive independence also carries over to a subpopulation.

Theorem 4.10 implies that the vectors $\boldsymbol{\lambda}_0$ and $\boldsymbol{\lambda}_1$ are identical in two different subpopulations $U^{(1)}$ and $U^{(2)} \subset U$. Under the additional assumption of conditional regressive independence, this is another empirically testable implication of the model. Note that the assumption of conditional regressive independence, or at least the assumption of uncorrelated errors is necessary, because otherwise $\boldsymbol{\lambda}_1$ would not be identified. If $P^{(2)}$ denotes the probability measure pertaining to the second subpopulation, we may formulate the following corollary (see Appendix, Note 21).

4.11. *Corollary (two subpopulations)*

(i) If $\mathscr{M} := \langle(\Omega, \mathfrak{A}, P), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ is a model of τ-congeneric variables, then $\mathscr{M}^{(1)} := \langle(\Omega, \mathfrak{A}, P^{(1)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ and $\mathscr{M}^{(2)} := \langle(\Omega, \mathfrak{A}, P^{(2)}), \mathbf{E}(\mathbf{y} \mid p_U), \tau, \boldsymbol{\lambda}_0, \boldsymbol{\lambda}_1\rangle$ are models of τ-congeneric variables, too.

(ii) If $\mathscr{M}$ is a model of τ-congeneric variables with conditional regressive independence, then $\mathscr{M}^{(1)}$ and $\mathscr{M}^{(2)}$ are models of τ-congeneric variables with conditional regressive independence, too.

Comments. It should be observed that proposition (i) does not imply any empirically testable hypotheses. However, proposition (ii) *is* empirically testable, but this test relies on the conditional independence assumption. Hypotheses of the type

$$H_0: \begin{array}{l} \lambda_0^{(1)} = \lambda_0^{(2)},\ \lambda_1^{(1)} = \lambda_1^{(2)}, \\ \mathrm{Cov}^{(1)}(\varepsilon_i, \varepsilon_j) = \mathrm{Cov}^{(2)}(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j,\ i, j \in I, \end{array} \tag{63}$$

may easily be tested using the programs for simultaneous equation models already mentioned. Note that the corollary above does *not* imply the variances $\mathrm{Var}^{(1)}(\tau)$ and $\mathrm{Var}^{(2)}(\tau)$ to be equal in the two subpopulations. The distributions of τ in the two subpopulations are *not* restricted by the model of τ-congenerity with conditional regressive independence. Also note that the Y-variables usually can *not* be used to define the subpopulations $U^{(1)}$ and $U^{(2)}$ as long as there is a positive variance for each of the error variables ε_i. Examples for criteria that *may* be used to partition U into subpopulations have been given in Section 3.

To summarize, there are two procedures for the empirical test of the model of τ-congeneric variables with conditional regressive independence: (a) The test of the covariance structure in the total population and (b) the test of the covariance structure in the subpopulations with the equality restrictions formulated in Equation 63. Each of these tests can lead to a rejection of the model. Note that no test is available which does *not* rely on the assumption of conditional regressive independence (or at least, of uncorrelated error variables).

5. Discussion

Our discussion will focus on four topics which seem appropriate to organize some of the methodological consequences of this article.

Properties of the error variables. A first point to be discussed are the properties of the error variables defined by Equation 4. For instance, there are criticisms of the theory of classical psychometric tests (CTT), denying the noncorrelation between a residual and a true score variable (see Equation 10), or denying the expectation of zero for the residuals (see Equation 7; see, e.g., Buttgereit, 1980; Hilke, 1980; Zeller & Carmines, 1980, p. 11). As long as the definitions of the variables τ_i and ε_i are accepted, Equations 3 to 12 cannot be wrong in empirical applications. The only thing that may be discussed in this context is whether or not the variables τ_i or ε_i defined by Equations 3 and 4 are of substantive interest.

The relationship between classical and probabilistic models. Since Equations 3 to 12 are definitions and their implications, it should be observed that these Equations do not constitute a model that could be wrong in an application. These equations are shared by models for dichotomous Y-variables (such as the Rasch model) as well. Since these equations are not based on any restrictive assumptions, they are a mathematical background for both types of models, although the models for dichotomous Y-variables do not explicitly use these equations. The Rasch model and the model of essentially τ-equivalent variables, for instance, differ only in their decomposition of the conditional expectations $\tau_i = E(Y_i \mid p_U)$ (see Appendix, Note 6) and in the restriction of the Rasch model to dichotomous Y-variables with values 0 and 1. Since the Rasch model deals with the decomposition of τ_i, it implicitly also deals with the error variable $\varepsilon_i := Y_i - \tau_i$.

Another point of critique of classical psychometric test theory is that it is said to be deterministic; according to those authors (see, e.g., Kubinger, 1986), this is in contrast to the so-called probabilistic models. However, the *probability* $P(Y = 1)$ that a dichotomous variable Y with values 0 and 1 takes on the value 1 *is and expectation*. Recall that the expectation of a discrete variable Y with n different values is defined $E(Y) := \sum_{i=1}^{n} y_i \cdot P(Y = y_i)$. Hence, if Y is dichotomous with values 0 and 1, then $E(Y) = 1 \cdot P(Y = 1) + 0 \cdot P(Y = 0) = P(Y = 1)$. The same kind of argument holds for *conditional* expectations and probabilities. The only difference is that, in some classical models, the true score variable $\tau_i := E(Y_i \mid p_U) = P(Y_i = 1 \mid p_U)$ is *additively* decomposed such that $\tau_i = \lambda_i + \tau$, $\lambda_i \in \mathbb{R}$, (see the model of essentially τ-equivalent variables), whereas in the Rasch model the decomposition is *additive in the logits*, that is, $\ln(\tau_i/(1 - \tau_i)) = \tau^* + \lambda_i^*$, or equivalently,

$$\tau_i = \frac{\exp(\tau^* + \lambda_i^*)}{1 + \exp(\tau^* + \lambda_i^*)}, \quad \lambda_i^* \in \mathbb{R}. \tag{64}$$

Both classes of models are stochastic or probabilistic. Hence, *classical* vs. *probabilistic* is a misleading dichotomy.

Note that the model of essentially τ-equivalent variables will even hold, for instance, if the variables Y_i are dichotomous provided that they are parallel. The variables Y_i, $i \in I$, are called *τ-parallel* if and only if Conditions 2.1 (a) to (d) hold with $\lambda_{ij} = 0$ for each pair of Y-variables and $Var(Y_i \mid p_U) = Var(Y_j \mid p_U)$, $i, j \in I$. However, if the variables Y_i are dichotomous and *not parallel*, they may only be essentially τ-equivalent in very special cases (see Appendix, Note 22). (For an example, see Moosbrugger, 1985.) If there are dichotomous variables Y_i with values 0 and 1, the Rasch model (see, e.g., Anderson, 1980; Fischer, 1974; Rasch, 1960/1980) should be considered instead of the model of essentially τ-equivalent variables.

Empirical testability. Another point to be discussed is the objection that models of classical psychometric test theory cannot be tested empirically. While it is certainly true that they *are* hardly ever tested, it is *not correct* to maintain that they *cannot* be tested, because of their mathematical structure. It was shown above *that* and *how* such tests can be conducted. Hence, if the models presented are considered models of classical test theory, models of classical test theory *can* be tested empirically.

Qualitative vs. quantitative basis of measurement models. Oftentimes, people object basing a measurment model on "arbitrarily" assigned numbers such as test scores. According to their argument, it has to be shown empirically that a relation (with specific properties) on the set of objects to be measured must be given *before* the assignment of numbers is justified. Although this principle might be heuristically fruitful in some research, it should not become a dogma. In the models treated in this article, a similar principle is involved according to which the conditional expectations have to fulfill certain restrictions. Only if these restrictions are met, we may meaningfully introduce the latent variable τ and the function φ which assigns numbers to the observational units. Hence, the choice of the Y-variables is restricted by the assumptions and properties of the model.

Choosing the Y-variables should also be based on *substantive* considerations, and the same is true for the choice of the basic relations on the set of objects in deterministic measurement models. *Logically*, starting with test score variables has the same justification as starting with qualitative observations. The latter, too, are arbitrarily selected. Preferring one over the other can only be justified by *substantive* arguments, not by *logical* ones. For instance, the points raised by Epstein and O'Brien (1985) in the person-situation debate (e.g., broadness of trait to be investigated) may serve as an argument against model building on qualitative item levels. However, this is a *substantive* argument; no general recommendation can be made on *logical* grounds. If the purpose is to count pieces of fruit, it is not necessary to count apples and pears separately, although the latter procedure has is merits if the purpose is to study the diversity of fruit. In other words, whether starting with certain *numerical* random variables is fruitful or not, should be determined (a) by *model tests* and (b) by the *usefulness of the model* for the practical purpose considered.

6. Summary and conclusion

In this paper the basic concepts of classical psychometric test theory were treated following the reformulation presented by Zimmerman (1975). Aside from a more convenient notation, *representation, uniqueness*, and *meaningfulness* theorems were derived which clarify the scale level of the latent variables. These theorems were supplemented by theorems on *identifiability* and *testability*, explicating the logic of measurement via models of classical psychometric test theory. It was emphasized that, aside from the well-known implications for the structure of the covariance matrix of the observed variables, there are other empirically testable implications concerning subpopulations. The model of essentially τ-equivalent variables implies the equality of the differences between the expectations of the observed variables in different subpopulations, whereas the model of τ-congeneric variables implies the equality of the factor loadings in different subpopulations.

References

Allen, M. J., & Yen, W. M. (1979) *Introduction to measurement theory*. Monterey, CA: Brooks/Cole.

Andersen, E. B. (1980) *Discrete statistical models with social science applications*. Amsterdam: North-Holland.

Bauer, H. (1978) *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie* (3rd ed.). Berlin: De Gruyter.

Bauer, H. (1981) *Probability theory and elements of measure theory*. New York: Academic Press.

Bentler, P. M. (1984) *Theory and implementation of EQS, a structural equations program*. Los Angeles: BMDP Statistical Software.

Billingsley, P. (1986) *Probability and measure* (2nd ed.). New York: Wiley.

Breiman, L. (1968) *Probability*. Reading, MA: Addison-Wesley.

Buttgereit, P. (1980) Testdiagnostik (Testdiagnostics). In W. Wittling, (Ed.), *Handbuch der klinischen Psychologie. Band 1: Methoden der klinisch-psychologischen Diagnostik*, 158–233. Hamburg: Hoffman und Campe.

Epstein, S., & O'Brien, E. J. (1985) The person-situation debate in historical and current perspective. *Psychological Bulletin, 98*, 513–537.

Fischer, G. H. (1974) *Einführung in die Theorie psychologischer Tests* (Introduction to the theory of psychological tests). Bern: Huber.

Gulliksen, H. (1950) *Theory of mental tests*. New York: Wiley.

Hamerle, A. (1979) Über die meßtheoretischen Grundlagen von Latent-Trait Modellen (On the measurement theoretical foundations of latent trait models). *Archiv für Psychologie, 132*, 19–39.

Hamerle, A. (1982) *Latent-Trait-Modelle* (Latent trait models). Weinheim: Beltz.

Hilke, R. (1980) Zu den impliziten „anthropologischen" Voraussetzungen der Klassischen Testtheorie (On the anthropological assumptions implicit in classical test theory). *Diagnostica, 26*, 99–111.

Hinderer, K. (1980) *Grundbegriffe der Wahrscheinlichkeitstheorie* (Basic concepts of probability theory). Berlin: Springer.

Jöreskog, K. G. (1967) Some contributions to maximum likelihood factor analysis. *Psychometrika, 32*, 443–482.

Jöreskog, K. G. (1969) A general approach to confirmatory factor analysis. *Psychometrika, 34*, 183–202.

Jöreskog, K. G. (1971) Statistical analysis of sets of congeneric tests. *Psychometrika, 36*, 109–133.

Jöreskog, K. G., & Sörbom, D. (1984) *LISREL IV: Analysis of linear structural relationships by the method of maximum likelihood*. Mooresville, In: Scientific Software.

Kolmogorov, A. N. (1977) *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Basic concepts of probability). Berlin: Springer. (Original work published 1933).

Krantz, D. H., Luce, R.D., Suppes, P., & Tversky, A. (1971) *Foundations of measurement. Vol. I*. New York: Academic Press.

Kubinger, K. D. (1986) Adaptive Intelligenzdiagnostik (Adaptive diagnostics of intelligence). *Diagnostica, 32*, 330–344.

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981) *Das State-Trait-Angstinventar. Theoretische Grundlagen und Handanweisung* (State-Trait Anxiety Inventory. Theoretical foundations and manual). Weinheim: Beltz.

Lawley, D. N., & Maxwell, A. E. (1971) *Factor analysis as a statistical method*. London: Butterworths.

Loève, M. (1977) *Probability theory I*. Berlin: Springer.

Loève, M. (1977) *Probability theory II*. Berlin: Springer.

Lord, F. M. (1980) *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum.

Lord, F. M., & Novick, M. R. (1968) *Statistical theories of mental test scores*. Reading, Mass.: Addison-Wesley.

Majcen, A.-M., Steyer, R., & Schwenkmezger, P. (1988) Konsistenz und Spezifität bei State- und Trait-Angst (Consistency and Specificity in state and trait anxiety). *Zeitschrift für Differentielle und Diagnostische Psychologie, 9*, 105–120.

Moosbrugger, H. (1982) Dimensionalitätsuntersuchungen von FPI-Skalen mit dem Klassischen Latent-Additiven Testmodell (KLA-Modell) (Investigations on the dimensionality of FPI scales with the classical latent additiv test model). *Zeitschrift für Differentielle und Diagnostische Psychologie, 3*, 241–264.

Moosbrugger, H. (1985) Zur Angemessenheit des Klassischen Latent-Additiven Testmodells (KLA-Modell) bei variierten Reaktionsmodalitäten (On the appropriateness of the classical latent additiv test model for varying reaction modalities). *Zeitschrift für Differentielle und Diagnostische Psychologie, 6*, 2–9.

Moosbrugger, H., & Müller, H. (1982) A classical latent additive test model (CLA model). *The German Journal of Psychology, 6*, 145–149.

Müller, H. (1980) *Über „probabilistische" und „klassische" Testmodelle mit latent-additiver Struktur: Neue Aspekte der Klassischen Testtheorie* (On "probabilistic" and "classical" test models with a latent additiv structure: New aspects of classical test theory). Unveröffentlichte Diplomarbeit, Institut für Psychologie der JWG-Universität Frankfurt am Main.

Novick, M. R. (1966) The axioms and principal results of classical test theory. *Journal of Mathematical Psychology, 3,* 1–18.

Rasch, G. (1980) *Probabilistic models for some intelligence and attainment tests.* Chicago: University of Chicago Press. (Original work published 1960)

Roussas, G. G. (1973) *A first course in mathematical statistics.* Reading, MA: Addison-Wesley.

Steyer, R. (1987) Konsistenz und Spezifität: Definition zweier zentraler Begriffe der Differentiellen Psychologie und ein einfaches Modell zu ihrer Identifikation (Consistency and specificity: Definition of two central concepts of personality psychology and simple model for their identification). *Zeitschrift für Differentielle und Diagnostische Psychologie, 8,* 245–258.

Steyer, R. (1988) Conditional expectations: An introduction to the concept and its applications in empirical sciences. *Methodika, 2,* 53–78.

Steyer, R., Majcen, A.-M., Schwenkmezger, P., & Buchner, A. (1988) A latent state-trait anxiety model and its application to determine consistency and specificity coefficients. *Anxiety Research, 1,* 281–299.

Suppes, P., & Zinnes, J. L. (1963) Basic measurement theory. In R. D. Luce, R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology. Vol. 1.* 1–76. New York: Wiley.

Tack, W. H. (1980) Zur Theorie psychometrischer Verfahren. Formalisierung der Erfassung von Situationsabhängigkeit und Veränderung (On the theory of psychometric procedures. Formalization of the assessment of situational specificity and change). *Zeitschrift für Differentielle und Diagnostische Psychologie, 1,* 87–108.

Weiss, D. J. (1983) *New horizons in testing: Latent trait test theory and computerized adaptive testing.* New York: Academic Press.

Witting, H. (1985) *Mathematische Statistik I* (Mathematical statistics I). Stuttgart: Teubner.

Zeller, R. A., & Carmines, E. G. (1980) *Measurement in the social sciences.* London: Cambridge University Press.

Zimmerman, D. W. (1975) Probability spaces, Hilbert spaces, and the axioms of test theory. *Psychometrika, 40,* 395–412.

Zimmerman, D. W. (1976) Test theory with minimal assumptions. *Educational and Psychological Measurement, 36,* 85–96.

Zimmerman, D. W., & Williams, R. H. (1977) The theory of test validity and correlated errors of measurement. *Journal of Mathematical Psychology, 16,* 135–152.

Appendix

Note 1: The random experiments considered are not the sampling experiments which would consist of repeating the experiment considered several times. Sampling experiments are of a technical nature; they are necessary for the purposes of estimation and testing. If the parameters were already known we would never conduct sampling experiments. Substantive theory would only be concerned with the random experiments characterized by Equation 1. Statistical models which characterize the sampling process are not treated in this paper.

Note 2: Observe that a random variable need not necessarily be numerical. The projection p_U, for example, is nonnumerical; its values are the observational units $u \in U$. A probability space $(\Omega, \mathfrak{A}, P)$ can always be constructed such that p_U has a distribution. The basic concepts of measure theory and probability theory used in this paper can be found in books such as Bauer (1978, 1981), Billingsley (1986), Breiman (1968), Loève (1977, 1978), or Roussas (1973).

Note 3: The general concept of the regression (or the conditional expectation) $E(Y \mid X)$ of a random variable Y given a possibly nonnumerical random variable X might be new to some readers although it is well-known in the literature since Kolmogorov's (1933/1977) seminal work. For most of this paper it will be sufficient to know that $E(Y \mid X)$ is a random variable the values of which are the conditional expectations $E(Y \mid X = x)$ of Y given that X takes on the value x. Mathematical introductions can be found in the books mentioned in Note 2. An introduction for psychometricians and methodologically interested psychologists is given by Steyer (1988). Note that propositions and equations about conditional expectations hold with probability 1 or almost surely with respect to the probability measure $P (P-a\,s)$. Since this is well-known to mathematicians and might be confusing for other readers, it will only be mentioned in the proofs given in this appendix if we simultaneously deal with two probability spaces (see, e.g., Note 11).

Note 4: See one of the books mentioned in Note 2 for an introduction into the concept of measurability. The reader not interested in the mathematical details may simply read Equation 6 to hold for every mapping $f(p_U)$. Let $Z: \Omega \rightarrow \Omega'$ be a mapping, let $\mathfrak{A}$ be a σ-algebra on Ω, and let $\mathfrak{A}' := \{A' \subset \Omega' : Z^{-1}(A') \in \mathfrak{A}\}$ (see, e.g., Hinderer, 1980, p. 78 and Witting, 1985, p. 407). Z is defined to be p_U-*measurable* if and only if $Z^{-1}(\mathfrak{A}') \subset p_U^{-1}(\mathfrak{F}_U)$, where $Z^{-1}(\mathfrak{A}') := \{Z^{-1}(A') : A' \in \mathfrak{A}'\}$, $p_U^{-1}(\mathfrak{F}_U) := \{p_U^{-1}(F) : F \in \mathfrak{F}_U\}$, and $\mathfrak{F}_U := \{F \subset U : p_U^{-1}(F) \in \mathfrak{A}\}$.

Note 5: It is easily seen that Equation 4 implies $E(\varepsilon_i \mid p_U) = 0$, because $E(\varepsilon_i \mid p_U) = E(Y_i - E(Y_i \mid p_U) \mid p_U) = E(Y_i \mid p_U) - E(E(Y_i \mid p_U) \mid p_U) = E(Y_i \mid p_U) - E(Y_i \mid p_U) = 0$. In order to give some intuitive understanding: if $E(Y_i \mid p_U = u) = 0$ for each value u of p_U, then the property $E(Y_i \mid f(p_U(\omega) = f(u))) = 0$ will also hold for every image f(u). This is the essence of Equation 6. Now, if f is numerical, the regression line is parallel to the axis of the regressor (Eqs. 6 and 8), then the covariance and the correlation between ε_i and its regressor will be zero, too (Eqs. 9 and 10).

Note 6: The Rasch model may be defined by Conditions 2.1 (a) to 2.1 (d) with two changes:

(a) the Y-variables are dichotomous with values 0 and 1;
(b) Equation 13 is replaced by

$$\ln\left(\frac{E(Y_i \mid p_U)}{1 - E(Y_i \mid p_U)}\right) = \lambda_{ij} + \ln\left(\frac{E(Y_j \mid p_U)}{1 - E(Y_j \mid p_U)}\right) \tag{65}$$

(see Hamerle, 1979, p. 35). Hence, in the Rasch model, the logarithms of the odds ratios are translations of each other, whereas in the model of essentially τ-equivalent variables, the conditional expectations are assumed to be translations of each other. Note that point (a) implies $E(Y_i \mid p_U) = P(Y_i = 1 \mid p_U)$.

An alternative, but equivalent way to define the Rasch model is to postulate Conditions 2.1 (a) to 2.1 (c) and instead of Equation 65:

$$E(Y_i \mid p_U) = P(Y_i = 1 \mid p_U) = \frac{\exp(\lambda_i^* - \tau^*)}{1 + \exp(\lambda_i^* - \tau^*)}, \quad \lambda_i^* \in \mathbb{R}, \tag{65'}$$

where τ^* is a real-valued random variable on $(\Omega, \mathfrak{A}, P)$. Of course, the variables Y_i are assumed to be dichotomous with values 0 and 1. Note that Equation 65′ is also equivalent with

$$Y_i = \frac{\exp(\lambda_i^* - \tau^*)}{1 + \exp(\lambda_i^* - \tau^*)} + \varepsilon_i, \quad \lambda_i^* \in \mathbb{R}, \tag{65''}$$

where $\varepsilon_i := Y_i - E(Y_i \mid p_U)$. This shows that the Rasch model, too, implicitly deals with the error variables ε_i. (See Note 10 for a remark on the assumption of conditional regressive independence, which is called "local stochastic independence" in the literature on Rasch models.)

Note 7 (Proof of Theorem 2.2): The proof is trivial: If we define, for instance, $\tau := E(Y_1 \mid p_U)$ and $\lambda_i := \lambda_{i1}$, then Equations 13 and 14 are easily seen to be equivalent. Similarly, if Equation 14 holds, then Equation 13 follows with $\lambda_{ij} = \lambda_i - \lambda_j$.

Note 8 (Proof of Theorem 2.4): Again, the proof of the first part trivial and the second is simple: If $E(Y_i \mid p_U) =: \tau_i = \lambda_i + \tau$ and $\tau_i = \lambda_i' + \tau'$, then $\tau' - \tau = \lambda_i - \lambda_i'$. Hence, the difference $\lambda_i - \lambda_i'$ has to be the *same* real number for *each* $i \in I$, which allows to define $\alpha := \lambda_i - \lambda_i'$.

Note 9: In terms of Suppes and Zinnes (1963), Definition 2.1 and Theorems 2.2 and 2.4 deal with a derived *difference scale* in the *narrow* sense. Let φ_i, $i = 1, \ldots, m$, be the factorizations (see Bauer, 1978, p. 300) of the conditional expectations $E(Y_i \mid p_U)$, i.e., $\varphi_i(u) := E(Y_i \mid p_U = u)$. Furthermore, define the numerical functions $\varphi_i^* : U \times \{Y_1, \ldots, Y_m\} \to \overline{\mathbb{R}}$, by

$$\varphi_i^*(u, Y_i) := \varphi_i(u), \quad i \in I,$$

define the function $\varphi^* : U \times \{Y_1, \ldots, Y_m\} \to \overline{\mathbb{R}}$, by

$$\varphi^*(u, Y_i) := \varphi(u), \quad i \in I,$$

(see Eq. 16), and define the function $\lambda^* : U \times \{Y_1, \ldots, Y_m\} \to \mathbb{R}$, by

$$\lambda^*(u, Y_i) := \lambda_i.$$

The *derived measurement system* can then be written

$$\mathfrak{B} = \langle U \times \{Y_1, \ldots, Y_m\}, \varphi_1^*, \ldots, \varphi_m^* \rangle.$$

The *representing relations* are given by

$$\varphi_i^*(u, Y_i) = \varphi^*(u, Y_i) + \lambda^*(u, Y_i) = \varphi(u) + \lambda_i, \quad i \in I,$$

(see Eq. 16) and $(\varphi^*, \lambda^*): U \times \{Y_1, \ldots, Y_m\} \to \overline{\mathbb{R}} \times \overline{\mathbb{R}}$ is a *derived numerical assignment* (see Suppes & Zinnes, 1963, p. 18).

Note 10: If the variables $Y_1, \ldots, Y_m$ are nonnegative, then Equation 27 is equivalent with

$$E(Y_1 \cdot \ldots \cdot Y_m \mid p_U) = E(Y_1 \mid p_U) \cdot \ldots \cdot E(Y_m \mid p_U), \quad m \in \mathbb{N},$$

(see, e.g., Bauer, 1978, p. 298). If the variables $Y_1, \ldots, Y_m$ are dichotomous with values 0 and 1, then $P(Y_i = 1 \mid p_U) = E(Y_i \mid p_U)$. In this case, the equation above – and therefore also Equation 27 – define conditional (or "local") stochastic independence. Hence, the only difference between the Rasch model and the model of essentially τ-equivalent variables with conditional regressive independence is that Equation 13 is replaced by Equation 65.

Note 11 (Proof of Theorem 2.7): Equation 28 follows from the fact that a residual $\varepsilon_i = Y_i - E(Y_i \mid p_U)$ is uncorrelated with every numerical p_U-measurable function (see Eq. 9). Using Theorem 2.2, we get

$$\begin{aligned} \mathrm{Cov}(Y_i, Y_j) &= \mathrm{Cov}(\lambda_i + \tau + \varepsilon_i, \lambda_j + \tau + \varepsilon_j) \\ &= \mathrm{Var}(\tau) + \mathrm{Cov}(\varepsilon_i, \varepsilon_j), \end{aligned}$$

which is Equation 28 (note that λ_i is a constant). Noncorrelation of the error variables (Eq. 30) follows from Equation 27, because

$$\begin{aligned} \mathrm{Cov}(\varepsilon_i, \varepsilon_j) = \mathrm{Cov}(Y_i - E(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m), \\ Y_j - E(Y_j \mid p_U)). \end{aligned}$$

This equation shows that ε_j is a function which is measurable with respect to (Y_j, p_U). However, Y_j and p_U are regressors with respect to which ε_i is a residual. Hence, again the general theorem applies that a residual is uncorrelated with every numerical function of its regressors.

Note 12 (Proof of Corollary 2.8): Equation 32 is Equation 28 rewritten for $i \neq j$. Equation 33 follows from Theorem 2.2, Equation 11, and the well-known rules of computation for covariances.

Note 13 (Proof of Theorem 2.10):

(i) Because $\mathcal{M}$ is a model of essentially τ-equivalent variables and the indicator $I_{\Omega^{(1)}}$ (see Eq. 36) is a p_U-measurable function:

$$E(Y_i \mid p_U) = E(Y_i \mid p_U, I_{\Omega^{(1)}}) = \lambda_i + \tau \quad (P-as), \quad i \in I.$$

This equation implies that for $i \in I$

$$E^{(1)}(Y_i \mid p_U) := E_{I_{\Omega^{(1)}}=1}(Y_i \mid p_U) = E(Y_i \mid p_U, I_{\Omega^{(1)}} = 1)$$
$$= \lambda_i + \tau \quad (P^{(1)}-as)$$

is a conditional expectation with respect to the conditional probability measure $P^{(1)}$. The constants λ_i and also the $\mathfrak{A}$-measurable function $\tau: \Omega \to \overline{\mathbb{R}}$ are unchanged when passing from the probability measure P to $P^{(1)}$.

(ii) We have to show that Equation 27 implies:

$$E^{(1)}(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) = E^{(1)}(Y_i \mid p_U) \quad (P^{(1)}-as).$$

This implication may be derived as follows:

$$\begin{aligned}
&E^{(1)}(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m)\\
:= {}& E_{I_{\Omega^{(1)}}=1}(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m)\\
= {}& E(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m, I_{\Omega^{(1)}} = 1)\\
= {}& E(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m, I_{\Omega^{(1)}})\\
= {}& E(Y_i \mid p_U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m)\\
= {}& E(Y_i \mid p_U)\\
= {}& E(Y_i \mid p_U, I_{\Omega^{(1)}})\\
= {}& E(Y_i \mid p_U, I_{\Omega^{(1)}} = 1)\\
= {}& E_{I_{\Omega^{(1)}}=1}(Y_i \mid p_U)\\
:= {}& E^{(1)}(Y_i \mid p_U) \quad (P^{(1)}-as).
\end{aligned}$$

Note 14 (Proof of Corollary 2.11):

(i) The proposition that $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$ are models of essentially τ-equivalent variables is an immediate implication of Theorem 2.10. Equations 38 and 39 can be derived as follows:

$$E^{(1)}(Y_i) = E^{(1)}(\tau) + \lambda_i \quad \text{and} \quad E^{(2)}(Y_i) = E^{(2)}(\tau) + \lambda_i, \quad i \in \mathbb{R},$$

and

$$E^{(1)}(Y_j) = E^{(1)}(\tau) + \lambda_j \quad \text{and} \quad E^{(2)}(Y_j) = E^{(2)}(\tau) + \lambda_j, \quad j \in \mathbb{R},$$

imply:

$$E^{(1)}(Y_i) - E^{(1)}(Y_j) = \lambda_i - \lambda_j \quad \text{and} \quad E^{(2)}(Y_i) - E^{(2)}(Y^j) = \lambda_i - \lambda_j,$$

where $i, j \in I$. However, these equations imply Equations 38 and 39.

(ii) This is an immediate consequence of Theorem 2.10.

Note 15: Each item has four categories. The instruction was to rate how he or she feels right now. The item numbers refer to the following list: (1) Ich bin entspannt. (2) Ich bin nervös. (3) Ich bin bekümmert. (4) Ich fühle mich wohl. (5) Ich bin aufgeregt. (6) Ich bin beunruhigt. (7) Ich fühle mich ausgeruht. (8) Ich bin besorgt, daß etwas schief gehen könnte. (9) Ich bin zufrieden. (10) Ich bin zappelig. (11) Ich fühle mich selbstsicher. (12) Ich fühle mich angespannt. (13) Ich bin ruhig. (14) Ich bin gelöst. (15) Ich fühle mich geborgen. (16) Ich bin verkrampft. (17) Ich bin froh. (18) Ich bin besorgt. (19) Ich bin überreizt. (20) Ich bin vergnügt. These items were embedded in a larger questionnaire.

Note 16 (Proof of Theorem 4.2): Condition 4.1 (d) implies 4.2 (d'): If we define, for instance, $\tau := E(Y_1 \mid p_U)$, $\lambda_{i0} := \lambda_{i10}$, and $\lambda_{i1} := \lambda_{i11}$, then Equations 41 and 42 are easily seen to be equivalent. Similarly, if Equation 42 holds, then Equation 41 follows with $\lambda_{ij0} := \lambda_{i0} - \lambda_{j0}$ and $\lambda_{ij1} := \lambda_{i1} - \lambda_{j1}$.

Note that Conditions 4.1 (d) and 4.2 (d') are both equivalent to:

(d'') there exist a numerical random variable τ on $(\Omega, \mathfrak{A}, P)$ and real numbers $\lambda_{i0}, \lambda_{i1} \in \mathbb{R}$, $\lambda_{i1} > 0$, such that

$$Y_i = \lambda_{i0} + \lambda_{i1}\tau + \varepsilon_i, \tag{66}$$

where $\varepsilon_i := Y_i - E(Y_i \mid p_U)$, for each $i \in I$.

They are also equivalent to:

(d''') There exist a function $\varphi: \Omega \to \overline{\mathbb{R}}$ and real numbers $\lambda_{i0}, \lambda_{i1} \in \mathbb{R}$, $\lambda_{i1} > 0$, such that

$$E(Y_i \mid p_U = u) = \lambda_{i0} + \lambda_{i1}\varphi(u), \quad \text{for each } i \in I. \tag{67}$$

Note 17 (Proof of Theorem 4.4): The first part of this theorem is trivial. In the second part, the assumptions imply

$$E(Y_i \mid p_U) = \lambda_{i0} + \lambda_{i1}\tau = \lambda'_{i0} - \lambda'_{i1}\tau', \quad i \in I. \tag{68}$$

Hence,

$$\tau' = \left(\frac{\lambda_{i0} - \lambda'_{i0}}{\lambda'_{i1}}\right) + \left(\frac{\lambda_{i1}}{\lambda'_{i1}}\right) \cdot \tau$$

and

$$\alpha = \left(\frac{\lambda_{i0} - \lambda'_{i0}}{\lambda'_{i1}}\right), \beta = \left(\frac{\lambda_{i1}}{\lambda'_{i1}}\right).$$

Note that β is a *positive* real number, which proves Equations 43 and 45. Inserting α and β in $\lambda'_{i0} = \lambda_{i0} - (\lambda_{i0} - \lambda'_{i0})$ results in $\lambda'_{i0} = \lambda_{i0} - \lambda_{i1} \cdot \alpha/\beta$. This proves Equation 44.

Note 18 (Proof of Theorem 4.7): The proof follows the same line of argument as that of Theorem 2.7. Only Equation 55 is different. However, using Equations 29 and 30, it follows from Equation 66 and the well-known rules of computation for covariances.

Note 19 (Proof of Theorem 4.8): Equations 58 to 59 immediately follow from Equation 55. Observe that the covariances are positive following from Theorem 4.2 and Definition 4.6.

Note 20 (Proof of Theorem 4.10): The proof follows exactly the same line of argument as in Note 13. Simply replace $\lambda_i + \tau$ by $\lambda_{i0} + \lambda_{i1}\tau$.

Note 21 (Proof of Corollary 4.11): This corollary is an immediate consequence of Theorem 4.10.

Note 22: Suppose the variables were dichotomous, essentially τ-equivalent, but not parallel. Then there are at least two indices i, j such that $P(Y_i = 1 \mid p_U) = P(Y_j = 1 \mid p_U) + \lambda_{ij}$, $0 \neq \lambda_{ij} \in \mathbb{R}$. If $P(Y_j = 1 \mid p_U = u)$ occurs with a probability greater than zero and if it is such that the sum $P(Y_j = 1 \mid p_U = u) + \lambda_{ij}$ is greater than one, then the value of $P(Y_i = 1 \mid p_U)$ would be greater than one. This is a contradiction to the definition of $P(Y_i = 1 \mid p_U)$ to be a conditional probability. Such a contradiction does not occur if the variables Y_i are parallel or if $P(Y_i = 1 \mid p_U)$ takes on its values in an adequately restricted range.