

## Latent State-Trait Theory and Research in Personality and Individual Differences

ROLF STEYER<sup>1\*</sup>, MANFRED SCHMITT<sup>2</sup> and MICHAEL EID<sup>3</sup>

<sup>1</sup>University of Jena, Germany

<sup>2</sup>University of Magdeburg, Germany

<sup>3</sup>University of Trier, Germany

### Abstract

*Latent state-trait (LST) theory is a generalization of classical test theory designed to take account of the fact that psychological measurement does not take place in a situational vacuum. The basic concepts of latent state-trait theory (LST theory) are introduced. The core of LST theory consists of two decompositions: (a) the decomposition of any observed score into latent state and measurement error, and (b) the decomposition of any latent state into latent trait and latent state residual representing situational and/or interaction effects. Latent states and latent traits are defined as special conditional expectations. A score on a latent state variable is defined as the expectation of an observable variable  $Y_{ik}$  given a person in a situation whereas a score on a latent trait variable is the expectation of  $Y_{ik}$  given a person. The theory also comprises the definition of consistency, occasion specificity, reliability, and stability coefficients. An overview of different models of LST theory is given. It is shown how different research questions of personality psychology can be and have been analysed within the LST framework and why research in personality and individual differences can profit from LST theory and methodology. Copyright © 1999 John Wiley & Sons, Ltd.*

### INTRODUCTION

Although the distinction between states and traits appears to be as old as thinking about human mind and behaviour (Eysenck, 1983), and provides a conceptual basis in everyday language for the description of personality and individual differences (Chaplin, John and Goldberg, 1988), the concept of *states* was introduced to personality research only during the sixties and seventies (e.g. Cattell, 1966, 1979; Cattell and Scheier, 1961; Nesselroade and Bartsch, 1977; Spielberger, 1972). By contrast, the concept of *traits* has guided personality research from its very beginning (Allport, 1937; Carr and Kingsbury, 1938; Johnson, McClearn, Yuen, Nagoshi,

\*Correspondence to: Rolf Steyer, Friedrich-Schiller-Universität Jena, Institut für Psychologie, Am Steiger 3, Haus 1, D-07743 Jena, Germany.

Ahern and Cole, 1985). In the twenties and thirties of this century, several important research projects investigated the stability of behaviour and the consistency of behaviour across situations. Temporal stability and cross-situational consistency are considered prerequisites for the description of individuals in trait terms. Allport and Vernon (1933) investigated the consistency of expressive movements, Hartshorne and May (1928) the consistency of moral conduct, Dudycha (1936) the consistency of punctuality, and Newcomb (1929) the consistency of introvert-extrovert behaviour. At about the same time, Thurstone (1931; Thurstone and Chave, 1929) suggested conceptualizing social attitudes as behavioural dispositions or traits. After World War II, various attempts were made to develop comprehensive personality trait models based on factor analytic research (Cattell, 1946; Eysenck, 1947; Guilford, 1959). Up until today, the trait approach continues to be the predominant conceptual framework for the description of human personality and the prediction of behaviour (Matthews and Deary, 1998).

However, the validity and usefulness of the trait model have been questioned. For example, Hartshorne and May (1928) found an average correlation of 0.23 among moral behaviours in different situations and concluded that behaviours depend more on characteristics of the specific situation than on traitlike dispositions of the individual. Similarly, Allport and Vernon (1933), Dudycha (1936), and Newcomb (1929) found rather low correlations among behaviours observed in different situations. Based on results from this and related research, Mischel (1968) and others argued that traits are not useful for the description of personality because behaviour depends mostly on the situation and thus cannot be predicted with sufficient accuracy from trait measures. Mischel's critique was countered by several authors who advanced various solutions for the consistency problem.

- (1) Some authors recommended giving up or at least supplementing the trait approach and elaborating cognitive and motivational theories of action, instead (e.g. Alston, 1975; Mischel, 1973; Wright and Mischel, 1987).
- (2) Others, in particular Epstein (e.g. 1983, 1984), defended the trait approach by arguing that a fair test of the consistency assumption must go beyond the level of single situations, because the concept of traits refers to behavioural manifestations in broad classes of situations. Epstein (1979, 1980) proposed aggregating behavioural measures across situations and showed that high correlations can be found among aggregated behavioural measures. Epstein and O'Brien (1985) reanalysed data from previous studies (Allport and Vernon, 1933; Hartshorne and May, 1928; Dudycha, 1929; Newcomb, 1929) and concluded that these data were in perfect agreement with the trait concept if properly aggregated.
- (3) The moderator variable approach was a third reaction to Mischel's critique of the trait approach. Bem and Allen (1974) suggested that individuals may differ in their behavioural consistency across situations. Some individuals may be more traited than others and therefore more predictable. Since the contribution of Bem and Allen (1974), various explanations for individual differences in consistency have been offered and investigated (Ajzen, 1987; Schmitt and Borkenau, 1992; Eid and Diener, 1999; Snyder and Ickes, 1985).
- (4) *Modern interactionism* was a fourth reaction to the consistency debate (Endler and Magnusson, 1976). Just like aptitude-treatment-interaction research (Cronbach and Snow, 1977), this paradigm employed factorial ANOVA designs

to show that, aside from individual differences in behaviour and main effects of situations, behavioural variability is substantially due to interactions between persons and situations (e.g. Bowers, 1973; Endler and Hunt, 1966; Sarason, Smith and Diener, 1975).

Basically, latent state-trait theory (LST theory) presented in this article is another reaction to the consistency controversy. Just like modern interactionism, it aims at taking into account the fact that not only persons but also situations and the interaction between persons and situations are important sources of variance in psychological measurement. However, in contrast to manipulating situation characteristics experimentally and crossing situation factors with personality factors in one design, as was done in most interactionist studies (cf. Bowers, 1973; Endler and Hunt, 1966; Sarason *et al.*, 1975), LST theory strives to consider these sources of variances also in observational (nonexperimental) research and to represent traits and situation and/or interaction effects in structural equation models. The basic difference from Epstein's approach is that LST theory does not try to eliminate situational specificity by aggregation but aims at incorporating it into its models as integral parts. States and traits as well as situation and/or interaction effects and measurement errors are explicit components of the models of data analysis utilizing modern developments in psychometrics: structural equation modelling with latent variables.

Over the last 12 years considerable progress has been made in the development of LST models for different types of data and research questions. Moreover, models of LST theory have been applied in many areas of personality and social psychology. The aim of this article is to review the basic principles of LST theory and to provide a guide through the various questions that can and have been pursued by latent state-trait modelling. The organization of this article is as follows. First, we give a brief conceptual introduction to the basic ideas of LST theory. Second, we provide a formal outline of LST theory and its basic variables and define the reliability, consistency, and occasion specificity coefficients. Third, we describe one model of LST theory in detail, illustrate it by a small example, and give a brief overview of other models of LST theory. Fourth, we describe research questions that can be analysed with LST models, and review previous applications of LST theory. Finally, we discuss LST theory in comparison to other approaches to the state-trait distinction. Readers who are more interested in practical applications of LST theory might read the section on questions that can be pursued by latent state-trait research before the section on the formal outline of LST theory.

#### CONCEPTUAL OUTLINE OF LATENT STATE-TRAIT THEORY

Like modern interactionism, LST theory starts from the premise that human cognition, emotion, and behaviour depend systematically on characteristics of the person (traits), characteristics of the situation, and the interaction between person and situation. Together, these factors create a psychological state which varies across time points to the extent that the situation to which the person is exposed changes. In experimental research, individuals are exposed to systematically different situations such that individual differences and situation differences are unconfounded. In natural life, however, individual differences and situation differences are confounded at any given time point. If we observe several individuals at the same time point, each

of them will most likely be in a different psychological (psychosocial and psychobiological) situation. If we are interested in anxiety, for instance, one person may just have been told by the doctor that s/he needed surgery to remove cancerous skin tissue while another person may just have been offered a job which means the end of worries about unemployment. These two individuals will, at the occasion of measurement (time point), differ in their anxiety *state*. LST theory assumes that this difference reflects both the difference between the situations in which these individuals are, and the difference between their trait anxiety.

If we extend this idea to psychological assessment in general, it means that every test, questionnaire, or other indicator never measures only individual trait differences but also systematic effects of the situation in which the person is when the questionnaire or test is administered. Measurement does not take place in a situational vacuum and we always measure persons in situations. In other words, our instruments always measure individual *state* differences and these state differences are due to *trait* differences, *situation* differences, and *person-situation interactions*.

This idea is not new but a general assumption in experimental personality research. However, LST theory differs from the experimental paradigm in that the situations in which measurement takes place do not have to be known and do not have to be observed. The natural variation of situations between occasions and the natural differences between subjects (persons in situations) within each occasion of measurement is sufficient. It is sufficient to assume that from occasion to occasion, different psychological situations such as the ones mentioned may be given and have an effect on the person's feeling, thinking, and behaving. Finally, situations do not have to be described in some objective sense in order to determine the proportion of variance accounted for by situations and/or interactions. However, if variables describing specific characteristics of situations are present, it will be possible to estimate their effects as well.

### FORMAL OUTLINE OF LATENT STATE-TRAIT THEORY

LST theory provides a formal representation of these ideas by decomposing any observed variable (e.g. a test score variable) into a latent state variable and an error variable. The latent state variable is decomposed into a latent trait variable and an occasion-specific residual which represents the effect of the situation in which the person is when assessment is performed and the person-situation interaction. In the following, the basic variables of LST theory will be described in detail. We restrict our presentation to the *population level*; that means no problems of estimation and testing will be discussed. Of course, in empirical applications, we will also need sample models, estimation, and testing procedures. Techniques for testing models and estimating parameters are described in sufficient detail in the literature on structural equation modelling (Arbuckle, 1997; Bentler, 1995; Bollen, 1989; Jöreskog and Sörbom, 1993; Muthén and Muthén, 1998).

#### Observable Random Variables

Figure 1 represents the decomposition of *observable random variables*  $Y_{ik}$ . Such observable variables may be defined, for instance, by the scoring rule of a psycho-

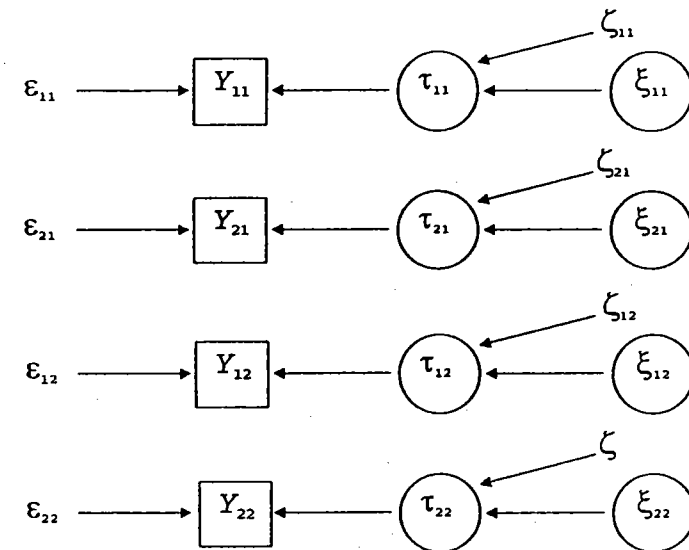


Figure 1. The decomposition of the observable variables of latent state-trait theory.

logical test, by the scoring rule of an instrument for systematic observation, self-ratings, peer ratings, or by any other measurement device, for instance, for hormone levels (see, e.g. Kirschbaum, Steyer, Eid, Patalla, Schwenkmezger and Hellhammer, 1990) or cortical activation (e.g. Hagemann, Naumann, Lürken and Bartussek, in press). In LST theory, we usually need two indices for the observable variables  $Y_{ik}$ : the second index,  $k$ , refers to the  $k$ th occasion of measurement, whereas the first index,  $i$ , refers to the  $i$ th measurement of the person considered within occasion  $k$ . Usually, the scores of the variables  $Y_{ik}$  and  $Y_{il}$  having identical first indices are obtained by the *same* measurement instrument (e.g. a psychological test) used on both occasions  $k$  and  $l$ .

The variables  $Y_{ik}$  are considered *random variables* pertaining to the random sampling of a single person from a larger group of persons (referred to as the 'population') and taking several measurements of the sampled person on each of several occasions. Even if there were no measurement error, the scores of the variables  $Y_{ik}$  would be random, because the person considered is randomly sampled from the population.

#### Latent State Variables and Measurement Errors

Analogously to classical test theory (CTT; Lord and Novick, 1968; Novick, 1966; Steyer, 1989, in press; Zimmerman, 1975, 1976), an observable variable  $Y_{ik}$  is decomposed into the sum of the *true score variable*  $\tau_{ik}$  and the *error variable*  $\epsilon_{ik}$  (part B of Table 1). However, in contrast to CTT, a true score (with respect to the measurement  $i$  on the occasion  $k$  of measurement) is the *expectation* (true mean) of the distribution of

Table 1. Basic variables of latent state-trait theory

A. Notation	
$\tau_{ik}$	latent state variables
$\xi_{ik}$	latent trait variables
$\zeta_{ik} = \tau_{ik} - \xi_{ik}$	latent state residuals
$\epsilon_{ik} = Y_{ik} - \tau_{ik}$	measurement error variables
B. Decomposition of variables	
$Y_{ik} = \tau_{ik} + \epsilon_{ik}$	
$\tau_{ik} = \xi_{ik} + \zeta_{ik}$	
C. Decomposition of variances	
$\text{Var}(Y_{ik}) = \text{Var}(\tau_{ik}) + \text{Var}(\epsilon_{ik})$	
$\text{Var}(\tau_{ik}) = \text{Var}(\xi_{ik}) + \text{Var}(\zeta_{ik})$	
D. Properties of expectations	
$E(\epsilon_{ik}) = E(\zeta_{ik}) = 0$	
E. Properties of covariances	
$\text{Cov}(\epsilon_{ik}, \zeta_{jk}) = \text{Cov}(\epsilon_{ik}, \tau_{jk}) = \text{Cov}(\epsilon_{ik}, \xi_{jk}) = \text{Cov}(\zeta_{ik}, \xi_{jk}) = 0$	

Note: All properties listed in this table (including the decompositions of variances) follow from the definitions of  $\tau_{ik}$ ,  $\xi_{ik}$ , and  $\zeta_{ik}$  (see Steyer and Schmitt, 1990a). There are no assumptions other than that the variances of the  $Y_{ik}$  variables are finite and greater than zero. All equations displayed hold for all  $i, j \in I_k := \{1, \dots, m_k\}$  and  $k, l \in K := \{1, \dots, n\}$ , where  $m_k$  is the number of variables measured on occasion  $k$  and  $n$  is the number of occasions of measurement.

$Y_{ik}$  conditional on the person in the situation (cf. Anastasi, 1983) rather than the person. In LST theory the true score variable  $\tau_{ik}$  is also called the latent state variable.

As outlined already, the term *situation*, as used in this context, does not only refer to observable or experimentally manipulable conditions. Those could be dealt with in the usual ANOVA design or by introducing additional observable variables representing the situation. Rather, the term 'situation' refers to the unobservable psychological conditions that might be relevant for the measurement of the construct considered.

Consider two more examples in addition to the anxiety example given earlier for clarifying this point. Suppose we want to measure the *ability of concentration*. It is to be expected that we would not obtain the same results for John after a night with two hours sleep as after a night with eight hours sleep. Similarly, Bill's score on scale measuring *attitude towards abortion* might depend on whether or not he saw a report on abortion on TV the day before. These examples make clear that psychological measurement does not take place in a situational vacuum, even if measurement of a trait is intended.

The fact that at best only a few aspects of the relevant situation may be known does not mean that situational effects should be considered part of the measurement error. Having repeated measurements of the same construct within the same occasion of measurement and comparing it to repeated measurements of the same construct on different occasions will help to separate trait from situational effects and will allow to estimate the impact of situational effects (in terms of variance components) even if the situation itself remains unknown in part or even completely.

Given the fact that the true scores mentioned above do not characterize the person but the person in the situation, we call the variables  $\tau_{ik}$  *latent state variables*. Since a score of a latent state variable characterizes the person in the situation, the score is composed of three components: (a) a component of the person, (b) a component of the situation, and (c) a component of the interaction between person and situation (Steyer and Schmitt, 1990a, p. 70).

Once the latent state variables are introduced, we may define the *measurement error variables*  $\epsilon_{ik}$  as the difference  $Y_{ik} - \tau_{ik}$ . Hence, the  $\epsilon_{ik}$  represent that part of the observable random variables  $Y_{ik}$  that is not determined by the person, nor by the situation, nor by the interaction between person and situation. The definition of the measurement error variables implies that their expectation is zero and that they are uncorrelated with a number of other latent variables and with the latent state variables pertaining to the same occasion of measurement (for details, see Table 1). These zero correlations are logical consequences of the definitions of the concepts (Steyer and Schmitt, 1990a). Hence, they are not empirically falsifiable, just in the same way as 'being unmarried' cannot be falsified as a property of a bachelor.

### The Latent Traits and Latent State Residuals

It is reasonable for theoretical reasons and useful for substantive applications of LST theory to decompose the true score variables  $\tau_{ik}$  (the latent state variables) into the sum of a trait component  $\xi_{ik}$  and an occasion-specific residual  $\zeta_{ik}$  (part B of Table 1). Whereas the latent state variable  $\tau_{ik}$  characterizes the person in the situation, the latent trait variable  $\xi_{ik}$  characterizes the person itself. The occasion-specific residual  $\zeta_{ik}$  represents situational and/or interactional effects.

The reader may wonder why index  $k$  is needed for the latent trait. The reason is that although trait scores, by definition, are not affected by the situation present on the occasion of measurement, traits may still be subject to change between occasions of measurement being due, for instance, to genetic programmes, learning, or critical life events (cf. Hertzog and Nesselrode, 1987). Our trait concept does not imply that traits are innate or unchangeable, although trait changes will, in general, occur very slowly and will not be the primary reason for intraindividual state variability. Rather, intraindividual state variability will, in general, be a consequence of the fact that the person is in different situations on different occasions of measurement.

Analogous to the error variables, the *latent state residuals*  $\zeta_{ik}$  may be defined as difference variables, namely the difference between the latent state variable  $\tau_{ik}$  and the latent trait variable  $\xi_{ik}$ . According to this definition, the  $\zeta_{ik}$  contain effects of the situation in which the person is measured and/or effects due to the interaction between person and situation. Since, in most applications, the situation is assumed to be constant within an occasion of measurement for a given person, we may also say that the variables  $\zeta_{ik}$  consist of *occasion-specific effects*. Just like the measurement error variables, the definitions of the latent state residuals  $\zeta_{ik}$  imply that they have zero expectation and zero correlation with all latent trait variables  $\xi_{ij}$ . Furthermore, the  $\zeta_{ik}$  and the  $\epsilon_{jk}$  are uncorrelated within the same occasion  $k$  (Table 1). Again, these zero correlations are logical consequences of the definitions of the concepts (Steyer and Schmitt, 1990a) and, therefore, cannot be empirically falsified. Latent states and traits, as defined here, are *not* uncorrelated (cf. Buse and Pawlik, 1991), whereas there is a zero correlation between latent traits  $\xi_{ij}$  and occasion-specific effects  $\zeta_{ik}$ .

Table 2. The most important parameters of LST theory

$\text{Rel}(Y_{ik}) = \text{Var}(\tau_{ik})/\text{Var}(Y_{ik})$ $= \text{Con}(Y_{ik}) + \text{Spe}(Y_{ik})$	Reliability
where	
$\text{Con}(Y_{ik}) = \text{Var}(\xi_{ik})/\text{Var}(Y_{ik})$	Consistency
$\text{Spe}(Y_{ik}) = \text{Var}(\zeta_{ik})/\text{Var}(Y_{ik})$	Occasion specificity
$\text{Cor}(\tau_{ik}, \tau_{il})$	Stability of states
$\text{Cor}(\xi_{ik}, \xi_{il})$	Stability of traits
$\text{Cor}(\xi_{ik}, \xi_{jk})$	Homogeneity of traits

Note:  $\text{Cor}(\xi_{ik}, \xi_{il})$  denotes the correlation between two latent trait variables pertaining to two different occasions but with identical first indices.  $\text{Cor}(\xi_{ik}, \xi_{jk})$  denotes the correlation between two latent trait variables pertaining to the same occasion of measurement but with different first indices.

### Consistency, Occasion Specificity, Reliability, Stability, and Homogeneity

In this article, the latent variables occurring in Figure 1 were introduced in a rather informal way. The more formal definitions show, however, that the variances of the variables  $\xi_{ik}$ ,  $\zeta_{ik}$ , and  $\epsilon_{ik}$  add up to the variance of  $Y_{ik}$ , without any additional assumptions other than that the variances of the variables  $Y_{ik}$  are finite (Steyer and Schmitt, 1990a). Similarly, the variances of the latent state variables  $\tau_{ik}$  can be decomposed additively into the variances of the trait and of the latent state residual (part C of Table 1). Hence, if  $\text{Var}(Y_{ik})$  is greater than zero, the last two equations allow us to define the coefficients of *consistency*, *occasion specificity*, and *reliability* as shown in Table 2.

The coefficient  $\text{Con}(Y_{ik})$  is the proportion of variance of a variable  $Y_{ik}$  reflecting the impact of interindividual differences that are due to the latent trait. The coefficient  $\text{Spe}(Y_{ik})$  is the proportion of variance of a variable  $Y_{ik}$  due to the situation and/or the person-situation interaction on occasion  $k$ . The coefficient  $\text{Rel}(Y_{ik})$  is the proportion of variance of a variable  $Y_{ik}$  determined by the differences between the persons, situations, and the interaction between persons and situations. Because of the additivity of the variances of the latent variables, the consistency and occasion specificity coefficients add up to the reliability coefficient. If one intends to measure a trait, the consistency should be high and the occasion specificity coefficient should be low. If one would like to assess a mood, it should be the other way around. In both cases, however, reliability should be high.

Additionally, we may define the correlation  $\text{Cor}(\tau_{ik}, \tau_{il})$  to be the *stability of the latent state variables* and  $\text{Cor}(\xi_{ik}, \xi_{il})$  to be the *stability of the latent trait variables* between occasions  $k$  and  $l$  of measurement. The use of the term 'stability' in this context presumes that  $\tau_{ik}$  and  $\tau_{il}$  ( $\xi_{ik}$  and  $\xi_{il}$ ) represent the same states (same traits) in some sense to be specified. For instance, using the same index  $i$  might indicate that the same test is given on the two occasions  $k$  and  $l$ . However, in some applications, a parallel form might do as well. If this assumption does not hold, one should simply use the term 'correlation' (e.g. between  $\tau_{ik}$  and  $\tau_{il}$ ) instead of the term 'stability'. Furthermore, the correlation  $\text{Cor}(\tau_{ik}, \tau_{jk})$  between two latent state variables  $\tau_{ik}$  and  $\tau_{jk}$  differing in their first index indicates the homogeneity or similarity of state on the same occasion of measurement. Analogously, the correlation  $\text{Cor}(\xi_{ik}, \xi_{jk})$  between two latent trait variables  $\xi_{ik}$  and  $\xi_{jk}$  differing in their first index reflects the

homogeneity or similarity of traits. The notions of state and trait homogeneity or similarity are meaningful from the perspective of LST theory because latent states, latent traits, latent state residuals, and measurement error are defined with respect to and for each manifest variable.

Accordingly, all latent variables and all coefficients introduced above have two indices,  $i$  and  $k$ ; that means, they are defined with respect to a method (or measurement device)  $i$  and an occasion  $k$  of measurement. Hence, each observable variable  $Y_{ik}$  has its own latent state component  $\tau_{ik}$ , its own latent trait component  $\xi_{ik}$ , its own latent state residual  $\zeta_{ik}$ , and its own measurement error variable  $\epsilon_{ik}$  (Figure 1). Obviously, it is not possible to identify the variances and covariances of the latent variables and to estimate individual state, trait, and residual scores unless (restrictive) assumptions are introduced. Different assumptions define special *models* of LST theory.

### MODELS OF LST THEORY

Models of LST theory are defined by assumptions about the basic variables of LST theory. These assumptions can differ in their complexity. The assumptions defining specific models serve to identify the theoretical parameters such as reliability, consistency, and occasion specificity. Many of these models have already been described in some detail by Steyer, Ferring and Schmitt (1992). Another, rather simple model of LST theory for two measurement instruments ( $i = 1, 2$ ), two occasions of measurement ( $k = 1, 2$ ), and mean-corrected variables is depicted in Figure 2. This model, called the *multistate-multitrait model* (cf. Eid and Diener, 1999), is defined by the following five assumptions:

- (1) All latent trait variables  $\xi_{ik}$  belonging to the same measurement instrument  $i$ , but different occasions of measurement  $k$ , are identical:  $\xi_{i1} = \xi_{i2}$ . This assumption implies that in this specific model the latent traits are perfectly stable and that a common latent trait variable can be defined by  $\xi_i = \xi_{i1} = \xi_{i2}$ . The latent traits, however, may differ between the different measurements taking differences between measurement instruments into account. The correlation of the common latent trait variables indicate the homogeneity or similarity of the two different measurements considered.
- (2) All latent state residuals  $\zeta_{ik}$  belonging to the same occasion of measurement  $k$ , but different measurements  $i$ , are identical:  $\zeta_{1k} = \zeta_{2k}$ . This assumption means that the situation and/or interaction effects are identical for the two measurements under consideration. This assumption implies that a common latent state residual can be defined by  $\zeta_k = \zeta_{1k} = \zeta_{2k}$ .
- (3) All latent state residuals  $\zeta_{ik}$  belonging to different occasions of measurement are uncorrelated. This means that the state residuals measure pure occasion-specific influences that are not shared with other occasions of measurement.
- (4) All error variables are uncorrelated implying that all correlations between the observed variables are explained by the latent traits and latent state residuals.
- (5) All error variables  $\epsilon_{ik}$  of an occasion  $k$  are uncorrelated with the common state residuals  $\zeta_l$  of another occasion  $l$ . The uncorrelatedness of the error variables and the common state residuals within the same occasion of measurement, however, is not an assumption, but a logical consequence of the definition of the state residuals and the error variables (see part E of Table 1).

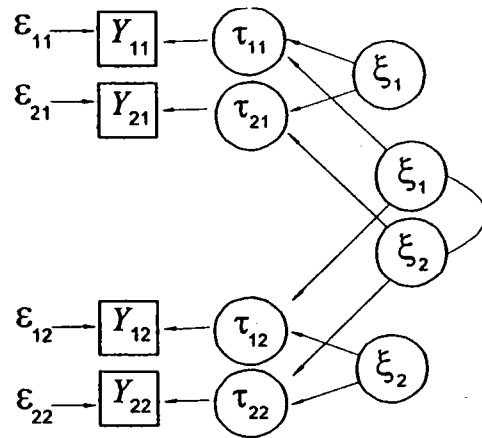


Figure 2. A multistate-multitrait model for two measurement instruments and two occasions of measurement. Because the latent state variables are perfectly determined by the latent traits and the state residuals (see part B of Table 1), there are no residuals for the variables  $\tau_{ik}$ .

All five assumptions can be wrong in an empirical application and are subject to empirical tests, whereas all properties of LST variables summarized in Table 1 cannot be wrong since they are logical consequences of the concepts. Computer programs for structural equation modelling such as AMOS (Arbuckle, 1997), EQS (Bentler, 1995), LISREL (Jöreskog and Sörbom, 1993), or Mplus (Muthén and Muthén, 1998) can be used to test some implications for the covariance structure of the observed variables that are implied by these assumptions. Moreover, these programs can be used to estimate the theoretical parameters of this model, that is, the variances and the covariances of the latent variables.

In Table 3 a small application of this model is summarized. The observed variables are two test halves of a German mood scale assessing the mood state of *awakeness versus sleepiness* (Steyer, Schwenkmezger, Notz and Eid, 1997). The covariance matrix in Table 3 represents the covariances of the two test halves administered on two occasions of measurement three weeks apart from each other. The sample size is  $N = 503$ . In addition to the model restrictions described above, it is assumed that the variances of the two latent trait variables are identical, that the variances of the two latent state residuals are identical and that all error variances are identical. As a first consequence of these assumptions, all observed variables have the same reliability, consistency, and occasion specificity (Table 3). As a second consequence, the covariance matrix implied by the model contains only four *different* values (Table 3). This reflects that four parameters of the model were estimated: (1) the variances of the two latent trait variables, (2) their covariance, (3) the variance of the two latent state residuals, and (4) the variance of the four measurement error variables. This model fits the data well ( $\chi^2 = 5.85$ ,  $df = 6$ ,  $p = 0.44$ ). The reliability coefficient reveals that the assessment of current mood is imposed by a small measurement error only. The small consistency and large specificity coefficients as well as the relatively small stability coefficients of the states indicate that there are strong situation and/or interaction influences on mood. Interindividual differences in mood are more variable

Table 3. Results of an empirical application of the multistate-multitrait model

A. Covariance matrix of the observed variables				
$Y_{11}$	17.06			
$Y_{21}$	15.27	16.16		
$Y_{12}$	4.80	4.50	17.22	
$Y_{22}$	4.80	4.98	15.94	17.06
B. Fitted covariance matrix				
$Y_{11}$	16.87			
$Y_{21}$	15.61	16.87		
$Y_{12}$	4.89	4.64	16.87	
$Y_{22}$	4.64	4.89	15.61	16.87

C. Estimated parameters of LST theory ( $i, k = 1, 2$ )

$Rel(Y_{ik}) = 0.94$	Reliability
$Con(Y_{ik}) = 0.29$	Consistency
$Spe(Y_{ik}) = 0.65$	Occasion specificity
$Cor(\tau_{11}, \tau_{12}) = Cor(\tau_{21}, \tau_{22}) = 0.31$	Stability of states
$Cor(\xi_{ik}, \xi_{il}) = 1$	Stability of traits
$Cor(\xi_{1k}, \xi_{2k}) = 0.95$	Homogeneity of traits

Note: The stability coefficients of the traits are not estimated. The perfect trait stability is a property of the model analysed. It is implied by the assumption that the common latent trait variables do not change over time.

than stable. The large correlations of the two common latent state and the common latent trait variables show that both test halves are very homogeneous. In sum, this small empirical example shows that models of LST theory are suitable for measuring persons in situations and for estimating the impact of situation and/or interaction effects.

#### Other Models

In this article only a very simple model of LST theory has been introduced to demonstrate the basic principles of defining the latent variables of structural equation models as functions of the basic variables of LST theory. Steyer *et al.* (1992) as well as Eid, Schneider and Schwenkmezger (1999) discuss further models of LST theory in which different restrictions are made. In particular, models with general trait factors and method-specific factors can be defined within this framework. Furthermore, models can be restricted in such a way that models without state residuals (perfect consistency) or without latent trait variables (perfect occasion specificity) emerge. Finally, multiconstruct LST models are useful for investigating the correlations among different traits and for exploring whether situation effects and person-situation interactions on occasions of measurement generate correlated latent state residuals of different psychological variables. Examples for this type of research will be given later in this paper (see also Schmitt, in press). Multiconstruct LST models including method specific factors can also be useful for exploring the convergent and discriminant construct validity of assessment instruments. Models of confirmatory factor analysis that are implied by the assumptions of LST models are very similar

and sometimes equivalent to other structural equation models for analysing longitudinal data that have been developed in other theoretical traditions (e.g. Kenny and Zautra, 1995; Marsh and Grayson, 1994; Ormel and Schauffeli, 1991).

Models of LST theory have also been developed for other types of observed and latent variables. The basic principles of LST theory and the LST model described so far are appropriate for metric observed variables, such as test scores or physiological variables, and metric latent variables (latent dimensions). This theory and the models defined on it are not suitable for categorical observed variables, for example, dichotomous response variables with categories *yes* and *no*, or Likert scales with a relatively small number of ordered categories. Moreover, this approach is not suitable for applications in which it is not reasonable to order individuals on latent dimensions, for example, when interindividual differences can be described best by a typological model with qualitative latent variables (latent classes). However, special LST models have been developed for categorical observed variables and for typological latent structures. Eid (1996) expanded LST theory for categorical observed variables with ordered categories and metrical latent variables and showed how item response models can be defined within this theoretical framework. These models can be applied to analyse the consistency and specificity of single item responses. Therefore, these models can be used for item selection and test construction. For instance, items with large consistency coefficients can be selected for personality questionnaires whereas items with large specificity coefficients can be selected for questionnaires assessing variable states such as moods or emotions (Eid, 1997, 1998). Latent state-trait models have also been developed for the measurement of latent state and trait typologies. Eid and Langeheine (1999) demonstrated how log-linear models with latent variables can be applied to define LST models with categorical observed and categorical latent variables (latent classes).

#### QUESTIONS TO BE PURSUED BY LATENT STATE-TRAIT RESEARCH

Models of LST theory are a useful methodological tool for answering different research questions of personality psychology. There are at least six important research questions that can be analysed within this methodological framework.

- (1) What are the *proportions of variance* of observable variables that are determined by (a) *trait effects (consistency)*, (b) *situation and/or interaction effects (occasion specificity)*, and (c) *measurement error (unreliability)*? This is the most typical research question for which models of LST theory have been applied. Answers to these questions tell us to what degree measurement instruments assess (a) traitlike versus (b) statelike attributes, and (c) how reliable the measurements are. There are many kinds of observed variables for which this question is important. For instance, LST models have been used to validate scales for the assessment of mood and emotions (Dumenci and Windle, 1996; Eid, Notz, Steyer and Schwenkmezger, 1994; Steyer, Majcen, Schwenkmezger and Buchner, 1989; Steyer, Schwenkmezger and Auer, 1990), attitudes (Steyer and Schmitt, 1990b), values (Schmitt, Schwartz, Steyer and Schmitt, 1993), motives (Schmitt and Steyer, 1993), personality traits (Deinzer *et al.*, 1995), and school-related constructs (Tanzer, 1998). But, not only questionnaire scores can be analysed by LST models. Kirschbaum *et al.* (1990)

estimated the consistency and occasion specificity of repeatedly measured cortisol in saliva to explore whether there are stable interindividual differences in cortisol reactivity. Hagemann *et al.* (in press) examined the consistency of EEG indicators of resting cortical activation asymmetry. Eid and Diener (1999) calculated the intraindividual standard deviations of daily affect ratings within seven different weeks and applied a LST model to the week-specific intraindividual standard deviations to scrutinize whether intraindividual variability in affect is a stable personality trait. Thus, LST models can be applied to quite different kinds of measurement. The only assumption to be met is that the basic decomposition of LST theory is reasonable for the measurements under consideration.

- (2) How do *different traits correlate*? In models of LST theory, a latent state variable is decomposed into a latent trait variable and an occasion-specific residual. The latent trait variable characterizes the person across different occasions of measurement. This trait variable is not affected by the occasion-specific fluctuations that can usually not be avoided in empirical studies. As Deinzer *et al.* (1995) pointed out, even the assessment of personality dimensions is often affected by situational and/or interactional influences. The latent trait variables of LST models are freed from situational fluctuations and are, therefore, more accurate representations of personality constructs. In extended LST models, the latent variables of different constructs can be correlated. Steyer *et al.* (1990) used these multiconstruct models to validate Spielberger's (1983) State-Trait Anxiety Inventory. In particular, they showed that the common latent trait variable of repeatedly measured anxiety states was strongly correlated with the common latent trait variable of the repeatedly administered trait scale. Schmitt (in press) investigated the correlation among the latent trait variables of questionnaire measures for two facets of family climate, family cohesion during childhood and parent-child attachment during adulthood. Schmitt and Steyer (1992) explored the correlation among the latent trait variables of measures for general life satisfaction and for job related life satisfaction. Other applications in which this method of validating state and trait scales has been employed are provided by Eid *et al.* (1994).
- (3) How do *different latent state residuals correlate*? In multiconstruct LST models, the associations between the latent state residuals of different constructs can also be analysed. These residuals measure the pure occasion-specific effects, and they can be considered as states that are freed from stable individual differences. The correlations between different state residuals indicate how strongly the situational and/or interactional influences are related across different constructs. In the two multi-construct studies mentioned in the last paragraph (Schmitt, in press; Schmitt and Steyer, 1992), significant and substantial correlations among the latent state residuals were found. In the family climate study by Schmitt (1999), the latent state residual of current attachment between adult daughters, who were the subjects in this study, and their aging mothers was correlated positively within occasions of measurement with the latent state residual of retrospective judgments of family cohesion during the daughter's childhood. This means that the situational and interactional effects which influenced the daughter's judgment of current attachment to her mother at the occasion of measurement also influenced her retrospective judgment of family cohesion. Significant and substantial within occasion correlations among latent-state residuals were also found in the life

satisfaction study (Schmitt and Steyer, 1992). It seems reasonable to assume that situation and interaction effects which influenced the person's job satisfaction affected the person's rating of general life satisfaction in the same direction. Eid and Diener (in press) discussed how the correlations of different residuals can be used to test the ecological validity of the mood-as-information paradigm (Schwarz and Strack, 1999) in nonexperimental studies. In particular, Eid and Diener (in press) showed that there are only small correlations between the state residuals of a repeatedly administered mood state scale, on the one hand, and the state residuals of repeatedly administered personality and life-satisfaction scales, on the other hand.

- (4) What are the correlates of *true state change*? What explains interindividual differences in intraindividual state change between two occasions of measurement? In LST models, state change is represented by occasion-specific influences. One way to explore the conditions for state change is to explain the occasion-specific deviations by exogenous variables within multiconstruct models as it has been outlined above. Another way is to define latent difference variables (e.g.  $\tau_{ik} - \tau_{i(k-1)}$ ) and to explain interindividual differences in these state change variables by other variables. Based on the concepts of LST theory, Steyer, Eid and Schwenkmezger (1997) developed a measurement model in which the true intraindividual state change scores are the values of a latent variable that can be explained by other variables (for other applications of this approach, see Steyer, Partchev and Shanahan, in press). Hence, these models are a viable alternative to growth curve models.
- (5) What are the correlates of *true trait change*? What explains interindividual differences in intraindividual trait change between two occasions of measurement? Models of LST theory can be defined to consider two different kinds of change, state variability and trait change. Eid and Hoffmann (1998) showed how LST models can be extended to measure trait change in addition to state variability. To apply a model with state variability and trait change, it is necessary to have at least two different periods of time between which trait change occurs. Furthermore, it is necessary to have at least two different occasions of measurement within each period of time. Then, an LST model without trait change can be defined within each period of time to identify state variability. For the (changing) latent trait variables of the two different time periods a latent autoregressive change process can be assumed and variables explaining trait change can be included in the analysis. Alternatively, extending the approach of Steyer *et al.* (1997a), true trait change variables may be explicitly included in the model as well as other variables that may explain the interindividual differences in true trait change.
- (6) Are there differences between subpopulations with respect to questions (1) to (5)? All questions described so far can also be analysed in a multigroup study. A typical application is the analysis of gender differences with respect to stability in mood (Steyer, 1989). Another application is to compare groups differing in the stability of living conditions or in the number of critical life events in a certain period of time. Consider as an example East and West Germans as groups. After the German reunification, East Germans were submitted to substantial changes in their economic and social living conditions, while the living conditions of West Germans remained rather stable. Smaller state and trait stabilities can therefore be expected for East Germans than for West Germans in many psychological

domains, for example, social and political attitudes, values, living styles, life satisfaction, and perhaps even mental health (Schmitt and Maes, 1998). In multigroup structural equation modelling, hypotheses about the equality of LST parameters across different (mutually exclusive) subgroups can be tested. For example, it can be investigated whether or not the variances of the state residuals (indicating the degree of occasion-specific influences) are equal across subpopulations differing in their mood reactivity (Eid *et al.*, 1994).

## DISCUSSION

The distinction between states and traits has been an issue of controversial debates (Allen and Potkay, 1981, 1983; Zuckerman, 1983), and is closely related to the fundamental fact that situations and/or interactions between persons and situations influence observable psychological measures to a considerable degree. The goal of our contribution has been to present a coherent theory relating observable variables (e.g. obtained by psychological test procedures) to the state and trait constructs.

In general, psychological tests and other observations do not measure *only* states or *only* traits as has been suggested by Allen and Potkay (1981). According to LST theory, latent states and latent traits are not incompatible with each other. Instead, a latent trait is one of the components constituting the latent state, i.e. *every* observable variable is more or less traitlike: the *proportion* of variance of the latent state that is determined by the latent trait may range from zero to one; its size may be estimated in empirical applications via simultaneous equation modelling. For this purpose we need repeated measurements on at least two occasions with at least two instruments measuring the same states.

The situations in which measurement takes place do not have to be known and do not have to be observed. The natural variation of situations between occasions and the natural differences between subjects (persons in situations) within each occasion of measurement is sufficient. This does not mean that it would not be possible to introduce additional situational variability by experimental manipulation if this is desired in order to explore the range of situational variability of some attribute.

It should be noted that the proportions of variance represented by the reliability, consistency, and occasion-specificity coefficients refer to the observational setting chosen. These coefficients will change, of course, if additional situational variation is introduced in another empirical study. This problem, which might be called *the relativism problem*, is due to the stochastic nature of all basic variables of LST theory. Latent state and trait variables have been defined as *random variables*. Hence, they are defined only with respect to the random experiment considered. This has important substantive implications. Suppose we study state and trait anxiety as measured by Spielberger's (1983) State-Trait Anxiety Inventory applying these scales on several occasions at the end of a university lecture (setting 1) (Steyer *et al.*, 1989). Obviously, the range of situations in which the students will answer to the items of the scale would be greater if the setting for assessing the scales were less homogeneous allowing various day times and various experiences before the STAI is applied (setting 2). The latent state variable, the latent state residual, and the latent trait variable, however, refer to the random experiment chosen (i.e. the observational setting 1). Hence, the

occasion-specificity coefficient, for instance, will be smaller in setting 1 than in setting 2.

Even the latent trait variable would have different meanings in the two settings mentioned above, because its scores are defined to be the conditional expected values given the person. Consequently, the trait scores are, by definition, averages across different distributions of situations. In other words, the true trait scores of Fritz Meier might differ between the two settings. This seems to be in conflict with a naive trait concept according to which one person has one and only one trait score at a given time. Therefore, it might be worthwhile to elaborate this point.

Suppose there are only three situations in setting 1 in which a person could be, each with probability 1/3, and that the state anxiety scores were 10, 20, and 30, respectively. According to our definition, the trait score would then be  $(1/3)(10 + 20 + 30) = 20$ . Obviously, we would obtain another trait score for setting 2, if there are four situations with equal probability and state scores 10, 20, 30, and 40 for the same individual. The trait score would then be 25. Thus, the intraindividual distribution of the situations is a part of the definition of the individual trait score. The naive idea that each person has one and only one trait score at a given time, and that this score is independent from the setting considered, rests on the assumption that a trait score can be measured without situational effects. However, psychological measurement does not take place in a situational vacuum, and situations have some effect on almost every measure. Even if you try to measure trait anxiety, the situation in which the items are answered will influence the result to some degree. Hence, the trait score can be defined only by some aggregation procedure such as taking the expectation across the different situations in which the individual might be at the occasion of measurement. Only if the situational effects were actually zero, would the mean be identical with the score being constant for the different situations, and the relativism problem would be nonexistent at this point.

Whereas the basic decomposition of LST theory is not restricted at all, models of LST theory are defined by restrictive and, therefore, empirically testable assumptions. Models of LST theory are needed to identify the basic parameters of LST theory in an empirical application. Beside this more technical aspect, models of LST theory are important for testing psychological theories. Is there one common latent trait variable for all constructs considered? Is there trait change in addition to state variability? Such questions and many more can be scrutinized on the basis of appropriately specified LST models.

## REFERENCES

- Ajzen, I. (1987). 'Attitudes, traits, and actions: dispositional prediction of behavior in personality and social psychology'. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, Vol. 20, pp. 1-63, Academic Press, New York.
- Allen, B. P. and Potkay, C. R. (1981). 'On the arbitrary distinction between states and traits', *Journal of Personality and Social Psychology*, 41: 916-928.
- Allen, B. P. and Potkay, C. R. (1983). 'Just as arbitrary as ever: comments on Zuckerman's rejoinder', *Journal of Personality and Social Psychology*, 44: 1087-1089.
- Allport, G. W. (1937). *Personality, a Psychological Interpretation*, Holt, New York.
- Allport, G. W. and Vernon, P. E. (1933). *Studies in Expressive Movement*, Macmillan, New York.

- Alston, W. P. (1975). 'Traits, consistency, and conceptual alternatives for personality theory', *Journal for the Theory of Social Behavior*, 5: 17-48.
- Anastasi, A. (1983). 'Traits, states, and situations: a comprehensive view'. In: Wainer, H. and Messick, S. (Eds), *Principles of Modern Psychological Measurement*, pp. 345-356, Erlbaum, Hillsdale, NJ.
- Arbuckle, J. L. (1997). *Amos User's Guide*, SPSS, Chicago, IL.
- Bem, D. J. and Allen, A. (1974). 'On predicting some of the people some of the time: the search for cross-situational consistency in behavior', *Psychological Review*, 81: 506-520.
- Bentler, P. M. (1995). *EQS. Structural Equations Program Manual*, Multivariate, Encino, CA.
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*, Wiley, New York.
- Bowers, K. S. (1973). 'Situationism in psychology: an analysis and a critique', *Psychological Review*, 80: 307-336.
- Buse, L. and Pawlik, K. (1991). 'Zur State-Trait Charakteristik verschiedener Meßvariablen der psychophysiologischen Aktivierung, der kognitiven Leistung und der Stimmung in Alltagssituationen' ['On the state-trait characteristic of various variables of psychophysiological activation, cognitive ability, and mood in every day situations'], *Zeitschrift für experimentelle und angewandte Psychologie*, 28: 521-538.
- Car, H. A. and Kingsbury, F. A. (1938). 'The concept of traits', *Psychological Review*, 45: 497-524.
- Cattell, R. B. (1946). *The Description and Measurement of Personality*, World, New York.
- Cattell, R. B. (1966). 'The data box: its ordering of total resources in terms of possible relational systems'. In: Cattell, R.B. (Ed.), *Handbook of Multivariate Experimental Psychology*, pp. 67-128, Rand McNally, Chicago, IL.
- Cattell, R. B. (1979). *Personality and Learning Theory*, Vol. 1, Springer, Berlin.
- Cattell, R. B. and Scheier, I. H. (1961). *The Meaning and Measurement of Neuroticism and Anxiety*, Ronald, New York.
- Chaplin, W. F., John, O. P. and Goldberg, L. R. (1988). 'Conceptions of states and traits: dimensional attributes with ideals as prototypes', *Journal of Personality and Social Psychology*, 54: 541-557.
- Cronbach, L. J. and Snow, R. E. (Eds) (1977). *Aptitudes and Instructional Methods. A Handbook for Research on Interactions*, Irvington, New York.
- Deinzer, R., Steyer, R., Eid, M., Notz, P., Schwenkmezger, P., Ostendorf, F. and Neubauer, A. (1995). 'Situational effects in trait assessment: the FPI, NEOFFI and EPI questionnaires', *European Journal of Personality*, 9: 1-23.
- Dudycha, G. J. (1936). 'An objective study of punctuality in relation to personality and achievement', *Archives of Psychology*, 29: 1-53.
- Dumenci, L. and Windle, M. (1996). 'A latent trait-state model of adolescent depression using the Center for Epidemiologic Studies—Depression Scale', *Multivariate Behavioral Research*, 31: 313-330.
- Eid, M. (1996). 'Longitudinal confirmatory factor analysis for polytomous item responses: model definition and model selection on the basis of stochastic measurement theory', *Methods of Psychological Research*, 1: 65-85 [Available: Online <http://www.pabstpublishers.de/mpr/>].
- Eid, M. (1997). 'Happiness and satisfaction: an application of a latent state-trait model for ordinal variables'. In: Rost, J. and Langeheine, R. (Eds), *Applications of Latent Trait and Latent Class Models in the Social Sciences*, pp. 145-151, Waxmann, Münster.
- Eid, M. (1998). Analyzing individual differences, In: Fischer, A. H. (Ed.), *Proceedings of the 10th ISRE Conference*, pp. 139-142, Faculty of Psychology, Amsterdam.
- Eid, M. and Diener, E. (in press). 'Global judgments of subjective well-being: situational variability and long-term stability'. In: Zumbo, B. (Ed.), *Social Indicators/Quality of Life Research Methods: Methodological Developments and Issues*, Kluwer, Dordrecht.
- Eid, M. and Diener, E. (1999). 'Intraindividual variability in affect: reliability, validity, and personality correlates', *Journal of Personality and Social Psychology*, 76: 662-676.
- Eid, M. and Hoffmann, L. (1998). 'Measuring variability and change with an item response model for polytomous variables', *Journal of Educational and Behavioral Statistics*, 23: 193-215.

- Eid, M. and Langeheine, R. (1999). 'Measuring consistency and occasion specificity with latent class models: a new model and its application to the measurement of affect', *Psychological Methods*, 4: 100-116.
- Eid, M., Notz, P., Steyer, R. and Schwenkmezger, P. (1994). 'Validating scales for the assessment of mood level and variability by latent state-trait analyses', *Personality and Individual Differences*, 16: 63-76.
- Eid, M., Schneider, C. and Schwenkmezger, P. (1999). 'Do you feel better or worse? On the validity of perceived deviations of mood states from mood traits', *European Journal of Personality*, 13: 283-306.
- Endler, N. S. and Hunt, JMcV (1966). 'Sources of behavioral variance as measured by the S-R Inventory of Anxiousness', *Psychological Bulletin*, 65: 336-346.
- Endler, N. S. and Magnusson, D. (Eds) (1976). *Interactional Psychology and Personality*, Wiley, New York.
- Epstein, S. (1979). 'The stability of behavior: I. On predicting most of the people much of the time', *Journal of Personality and Social Psychology*, 37: 1097-1126.
- Epstein, S. (1980). 'The stability of behavior: II. Implications for psychological research', *American Psychologist*, 35: 790-806.
- Epstein, S. (1983). 'A research paradigm for the study of personality and emotions', *Nebraska Symposium on Motivation*, 30: 91-154.
- Epstein, S. (1984). 'The stability of behavior across time and situations'. In: Zucker, R., Aronoff, J. and Rabin, A.I. (Eds), *Personality and the Prediction of Behavior*, pp. 209-268, Academic Press, San Diego, CA.
- Epstein, S. and O'Brien, E. (1985). 'The person-situation debate in historical and current perspective', *Psychological Bulletin*, 98: 513-537.
- Eysenck, H. J. (1947). *Dimensions of Personality*, Routledge and Kegan Paul, London.
- Eysenck, H. J. (1983). 'Cicero and the state-trait theory of anxiety: another case of delayed recognition', *American Psychologist*, 38: 114-115.
- Guilford, J. P. (1959). *Personality*, McGraw-Hill, New York.
- Hagemann, D., Naumann, E., Lürken, A. and Bartussek, D. (in press). 'EEG trait asymmetry and affective style I: the latent state and trait structure of resting asymmetry scores', *39th Annual Meeting of the Society for Psychophysiological Research, Granada, Psychophysiology (Supplement)*.
- Hartshorne, H. and May M. A. (1928). *Studies in the Nature of Character, Vol. 1: Studies in Deceit*, Macmillan, New York.
- Hertzog, C. and Nesselroade, J. R. (1987). 'Beyond autoregressive models: some implications of the trait-state distinction for the structural modeling of developmental change', *Child Development*, 58: 93-109.
- Johnson, R. C., McClearn, G. E., Yuen, S., Nagoshi, C. T., Ahern, F. M. and Cole, R. E. (1985). 'Galton's data a century later', *American Psychologist*, 40: 875-892.
- Jöreskog, K. G. and Sörbom, D. (1993). *LISREL 8. Structural Equation Modeling with the SIMPLIS Command Language*, Scientific Software, Chicago, IL.
- Kenny, D. A. and Zautra, A. (1995). 'The state-trait error model for multiwave data', *Journal of Consulting and Clinical Psychology*, 63: 52-59.
- Kirschbaum, C., Steyer, R., Eid, M., Patalla, U., Schwenkmezger, P. and Hellhammer, D. H. (1990). 'Cortisol and behavior: 2. Application of a latent state-trait model to salivary cortisol', *Psychoneuroendocrinology*, 15: 297-307.
- Lord, F. M. and Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*, Addison-Wesley, Reading, MA.
- Marsh, H. W. and Grayson, D. (1994). 'Longitudinal confirmatory factor analysis: common, time-specific, item-specific, and residual-error components of variance', *Structural Equation Modeling*, 1: 116-145.
- Matthews, G. and Deary, I. J. (1998). *Personality Traits*, Cambridge University Press, Cambridge.
- Mischel, W. (1998). *Personality and Assessment*, Wiley, New York.
- Mischel, W. (1973). 'Toward a cognitive social learning reconceptualization of personality', *Psychological Review*, 80: 252-283.
- Muthén, L. and Muthén, B. (1998). *Mplus User's Guide*, Muthén and Muthén, Los Angeles.
- Nesselroade, J. R. and Bartsch, T. W. (1977). 'Multivariate perspectives on the construct validity of the trait-state distinction'. In: Cattell, R.B. and Dreger, R.M. (Eds), *Handbook of Modern Personality Theory*, pp. 221-238, Hemisphere, Washington, DC.
- Newcomb, T. M. (1929). *Consistency of Certain Extrovert-Introvert Behavior Patterns in 51 Problem Boys*, Columbia University, Teachers College, Bureau of Publications, New York.
- Novick, M. R. (1966). 'The axioms and principal results of classical test theory', *Journal of Mathematical Psychology*, 3: 1-18.
- Ormel, J. and Schaufeli, W. B. (1991). 'Stability and change in psychological distress and their relationship with self-esteem and locus of control: a dynamic equilibrium model', *Journal of Personality and Social Psychology*, 60: 288-299.
- Sarason, I. G., Smith, R. E. and Diener, E. (1975). 'Personality research: components of variance attributable to the person and the situation', *Journal of Personality and Social Psychology*, 32: 199-204.
- Schmitt, M. (in press). 'Mother-daughter attachment and family cohesion: single and multi construct latent state-trait models of current and retrospective perceptions', *European Journal of Psychological Assessment*, in press.
- Schmitt, M. and Borkenau, P. (1992). 'The consistency of personality'. In: Caprara, G.-V. and Van Heck, G.L. (Eds), *Modern Personality Psychology. Critical Reviews and New Directions*, pp. 29-55, Harvester-Wheatsheaf, New York.
- Schmitt, M. and Maes, J. (1998). 'Perceived injustice in unified Germany and mental health', *Social Justice Research*, 11: 59-78.
- Schmitt, M., Schwartz, S. H., Steyer, R. and Schmitt, T. (1993). 'Measurement models for the Schwartz Values Inventory', *European Journal of Psychological Assessment*, 9: 107-121.
- Schmitt, M. and Steyer, R. (1992). 'Self-reports on life satisfaction: to what extent do they reflect a stable trait versus systematic situational effects?' paper presented at the 6th European Conference on Personality, Groningen.
- Schmitt, M. and Steyer, R. (1993). 'A latent state-trait model (not only) for social desirability', *Personality and Individual Differences*, 14: 519-529.
- Schwarz, N. and Strack, F. (1999). 'Reports of subjective well-being: judgemental processes and their methodological implications'. In: Kahneman, D., Diener, E. and Schwarz, N. (Eds), *Foundations of Hedonic Psychology: Scientific Perspectives on Enjoyment and Suffering*, Russell Sage Foundation, New York.
- Snyder, M. and Ickes, W. (1985). 'Personality and social behavior'. In: Lindzey, G. and Aronson, E. (Eds), *Handbook of Social Psychology*, 3rd edn, Vol. 2, pp. 883-947, Random, New York.
- Spielberger, C. D. (1972). 'Anxiety as an emotional state'. In: Spielberger, C.D. (Ed.), *Anxiety: Current Trends in Theory and Research*, Vol. 1, pp. 23-49, Academic Press, New York.
- Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (STAI Form Y)*, Consulting Psychologists, Palo Alto, CA.
- Steyer, R. (1989). 'Models of classical psychometric test theory as stochastic measurement models: representation, uniqueness, meaningfulness, identifiability, and testability', *Methodika*, 3: 25-60.
- Steyer, R. (in press). 'Classical test theory'. In: Ragin, C. and Cook, T. (Eds), *International Encyclopedia of the Social and Behavioral Sciences. Logic of Inquiry and Research Design*, Pergamon, Oxford.
- Steyer, R., Eid, M. and Schwenkmezger, P. (1997a). 'Modeling true intraindividual change: true change as a latent variable', *Methods of Psychological Research—Online*, 2: 21-33 <http://www.pabst-publishers.de/mpr/>.
- Steyer, R., Ferring, D. and Schmitt, M. (1992). 'States and traits in psychological assessment', *European Journal of Psychological Assessment*, 2: 79-98.
- Steyer, R., Majcen, A.-M., Schwenkmezger, P. and Buchner, A. (1989). 'A latent state-trait anxiety model and its application to determine consistency and specificity coefficients', *Anxiety Research*, 1: 281-299.
- Steyer, R., Partchev, I. and Shanahan, M. (1999). 'Modeling true intra-individual change in structural equation models: the case of poverty and children's psychosocial adjustment?' In: Little, T.D., Schnabel, K.U. and Baumert, J. (Eds), *Modeling Longitudinal and Multiple-*

- Group Data: Practical Issues, Applied Approaches, and Specific Examples*, Erlbaum, Hillsdale, NJ.
- Steyer, R. and Schmitt, M. J. (1990a). 'The effects of aggregation across and within occasions on consistency, specificity, and reliability', *Methodika*, 4: 58-94.
- Steyer, R. and Schmitt, M. J. (1990b). 'Latent state-trait models in attitude research', *Quality and Quantity*, 24: 427-445.
- Steyer, R., Schwenkmezger, P. and Auer, A. (1990). 'The emotional and cognitive components of trait anxiety: a latent state-trait model', *Personality and Individual Differences*, 11: 125-134.
- Steyer, R., Schwenkmezger, P., Notz, P. and Eid, M. (1997b). 'Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF) [Multidimensional Mood Questionnaire]', Hogrefe, Göttingen.
- Tanzer, N. K. (1998). Assessment of domain-specificity in school-related Likert-type inventories, unpublished habilitation thesis, University of Graz, Austria.
- Thurstone, L. L. (1931). 'The measurement of social attitudes', *Journal of Abnormal and Social Psychology*, 26: 249-269.
- Thurstone, L. L. and Chave, E. J. (1929). *The Measurement of Attitude*, University of Chicago Press, Chicago.
- Wright, J. C. and Mischel, W. (1987). 'A conditional approach to dispositional constructs: the local predictability of social behavior', *Journal of Personality and Social Psychology*, 53: 1159-1177.
- 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.
- Zuckerman, M. (1983). 'The distinction between trait and state scales is not arbitrary: comment on Allen and Potkay's "On the arbitrary distinction between states and traits"', *Journal of Personality and Social Psychology*, 44: 1083-1086.