Quality & Quantity (2006)
DOI 10.1007/s11135-006-9030-x

Investigating the Number of Non-linear and Multi-modal Relationships Between Observed Variables Measuring Growth-oriented Atmosphere

P. NOKELAINEN[1,⋆], T. SILANDER[2], P. RUOHOTIE[1] and H. TIRRI[3]
[1]*Research Centre for Vocational Education, University of Tampere, P.O. Box 229, 13101 Hämeenlinna, Finland;* [2]*Complex Systems Computation Group, Helsinki Institute for Information Technology, P.O. Box 9800, 02015 HUT, Finland;* [3]*Nokia Group, Nokia Research Center, P.O. Box 407, 00045 Nokia Group, Finland*

Abstract. This study investigates the number of non-linear and multi-modal relationships between observed variables measuring the Growth-oriented Atmosphere. The sample ($N = 726$) represents employees of three vocational high schools in Finland. The first stage of analysis showed that only 22% of all dependencies between variables were purely linear. In the second stage two sub samples of the data were identified as linear and non-linear. Both bivariate correlations and confirmatory factor analysis (CFA) parameter estimates were found to be higher in the linear sub sample. Results showed that some of the highest bivariate correlations in both sub samples were explained via third variable in the non-linear Bayesian dependence modeling (BDM). Finally, the results of CFA and BDM led in different substantive interpretations in two out of four research questions concerning organizational growth.

Key words: categorical data, survey data, non-linear modeling, structural equation modeling, organizational atmosphere

1. Introduction

When an organizational researcher wants to study dependencies between observed and latent variables, for example, the factors of growth-oriented atmosphere (Ruohotie, 1996), the assumptions for the data may become quite challenging in traditional frequentistic statistical analysis.

⋆Author for correspondence: Petri Nokelainen, Research Centre for Vocational Education, University of Tampere, P.O. Box 229, 13101 Hämeenlinna, Finland. E-mail: petri.nokelainen@uta.fi

Few examples of such assumptions are the assumption of continuous measurement level, multivariate normality and linearity of both the data and phenomena under investigation. For example, the result of violating multivariate normality assumption is that chi-square becomes too large (too many models are rejected) and standard errors become too small (significance tests have too much power). In addition, normal distribution analysis sets minimum requirements for the number of observations.

As a first response to the continuous measurement level and multivariate normality assumptions, Muthén and Kaplan (1985) suggest that in treating the ordinal variables as continuous does produce viable results as long as the frequency distributions are unimodal with an internal mode. Johnson and Creech (1983) have studied with simulation studies the categorization error that occurs when continuous variables are measured by indicators with only a few categories. The results indicated that while categorization error does produce distortions in multiple indicator models, under most conditions explored the bias was not sufficient to alter substantive interpretations with a large simulation sample ($N = 5000$). However, authors warranted caution in the use of two-, three- or four-category ordinal indicators, particularly when the sample size is small (in that study ten sub samples of 500 cases were examined). They were also worried about situation when it is not certain that a normal distribution accurately reflects the true distribution of many underlying variables.

The second solution that is presented theoretically by Muthén (1983, 1984, 1989) and applied in practice in LISREL by Jöreskog (2003) is to estimate tetrachoric (for binary variables) or polychoric (for categorical variables) correlations among the ordinal variables and use these correlations to estimate the model using asymptotic distribution free function (ADF) by Browne (1984). Amos (Arbuckle, 1999) software package uses Browne's original name, but EQS by Bentler (1995) describes it as arbitrary generalized least squares (AGLS) and LISREL (Jöreskog and Sörbom, 1998) together with Mplus (Muthén and Muthén, 2001) call it weighted least squares (WLS). The main advantage is that the estimator is not dependent on multivariate normality. However, a limited number of variables (recommendation is below 20) and demand for very large samples that are needed to produce good estimates are the cons of this approach. For example, a simulation study of Yung and Bentler (1994) suggests more than 2000 observations. In addition, Olsson et al. (2000) show that ADF estimation performs poorly when the model is misspecified.

The third approach to address modeling problems with ordinal nonnormal data is the categorical variable model (CVM) developed by Muthén (1993). The model, that is implemented in the Mplus program, uses the general ADF function but without aforementioned limitations. We consider this a viable frequentistic approach to the analysis of ordinal variables in organizational research.

However, none of the techniques described earlier address the problem of non-linear dependencies between observed variables. In this paper, we argue that the Bayesian modeling approach (see, e.g., Bernardo and Smith 2000; Myllymäki et al. 2002), named after English reverend Thomas Bayes (1702–1761) for his contributions (Bayes, 1763), is a viable alternative to frequentistic statistical techniques addressing all the above mentioned modeling problems. Although the subject domain of this paper is organizational atmosphere (or climate) research, we believe that this discussion fully applies to all the other research areas of social sciences where the measurement level of the indicators is categorical.

The essential benefits of Bayesian modeling are summarized in Congdon (2001). Next we discuss more thoroughly the three most important features for this study.

1.1. THE RESEARCHER IS CAPABLE OF INPUTTING A PRIORI INFORMATION TO THE MODEL

The source of subjective information could be, for example, an interview with an expert of a certain topic, or previously collected data. For example, an adaptive online questionnaire is able to profile respondents with Bayesian probabilistic modeling and thus personalize the total number of questions asked from each person. In this application field a priori profile information, that is applied in the profiling stage when considering the closest match profile for the current person is gained from earlier responses of similar population. In this paper, a priori information is needed only when Bayesian networks are learned from data, i.e. in the part where we perform Bayesian dependence modeling (BDM, results are presented in Tables II and III). Prior beliefs are quantified by calculating the *equivalent sample size* (ESS) (see, e.g., Heckerman et al., 1995) parameter value with Equation (1)

$$\mathrm{ESS} = \frac{|V_1| + |V_2| + \cdots + |V_n|}{2N}, \tag{1}$$

where $|V_i|$ denotes the number of values of each variable V_i in the model and N denotes the number of variables V_i in the model (Myllymäki et al., 2002). ESS is a parameter that regulates our behavior when new data is entered, that is, how we update our beliefs when new evidence is presented. If the ESS value is small, new evidence has greater impact to our beliefs than if the parameter value is large. Equation (1) is applied in this paper for two reasons: First, equal priors are set to all variables in the model, as we have no reason to favor any single variable. Second, the ESS parameter calculation is related to so-called Jeffrey's prior that is commonly used as a non-informative prior in Bayesian analysis.

1.2. BAYESIAN MODELING IS DESIGNED TO ANALYZE DISCRETE CATEGORICAL VARIABLES

Organizational researchers still collect some of their data with paper and pencil or web-based online surveys. The most typical question types in survey research are dichotomous and multiple-choice questions. In both cases, the categories are discrete (e.g., have no overlap and are mutually exclusive) and exhaust the possible range of responses. One of the major differences between traditional Gaussian and Bayesian models lies in the fact that the latter does not require a multivariate normal distribution of the indicators (e.g., observed variables) or underlying phenomena. This feature is especially useful for a researcher who collects her data with, for example, Likert -scale (DeVellis, 2003, pp. 78–80) questions as the response options from 1 to 7 produce data that is more qualitative than quantitative in nature. Measurement level of such item is ordinal and it is not advisable to model it with traditional statistical analysis that rely on the concept of normal distribution, and require the calculation of mean and standard deviation.

1.3. BAYESIAN MODELING IS ABLE TO ANALYZE BOTH LINEAR AND NON-LINEAR DEPENDENCIES BETWEEN VARIABLES

Phenomena under investigation are seldom purely linear or continuous in nature. Unfortunately most commonly applied traditional linear Gaussian models (e.g., regression and factor analysis) are statistically inadequate for understanding non-linear dependencies between variables. Bayesian dependence models for discrete data allow the description of non-linearities as Bayesian theory gives a simple criterion, the probability of the model, to select among such models.

A major drawback with the Bayesian approach at this moment is that only a few applications are capable of analyzing latent variable models. Examples of such software are BUGS[1] that studies Bayesian inference using Gibbs sampling (see Congdon, 2001, 2003) and DSIGoM[2] that is based on the grade of membership analysis. As the major goal of this paper is to investigate the number of non-linear and multi-modal relationships in real-life organizational research data sets, we limit the investigation into dependencies between observed variables and study bivariate correlations (both r_p and r_s) and Bayesian dependence modeling (B-Course[3]) to reach that goal.

The research questions in this study are: (1) What kind of and how many non-linearities are captured by discrete Bayesian networks? (2) Is there difference between the results of linear bivariate correlations and

Bayesian dependence modeling? (3) Does an empirical sample containing pure linear dependencies have better overall fit indices in CFA than a sample containing less linear dependencies? (4) Does an empirical sample containing pure linear dependencies have higher CFA parameter estimates than a sample containing less linear dependencies? (5) Is there difference between the substantive interpretations of the results of CFA and BDM with linear and non-linear samples?

2. Theoretical Background

2.1. GROWTH-ORIENTED ATMOSPHERE

Ongoing learning and self-development by employees are critical to the mission of any modern organization. In order to be successful, educational organizations must provide effective professional development programs for employees over the entire course of their careers (Lawler, 1994). Professional development includes all developmental functions that are directed at the maintenance and enhancement of professional competency. In the modern world updating is ideally a continual, lifelong process that addresses such goals as the acquisition of new and up-to-date information, the development of skills and techniques and the elevation of one's personal esteem. The maintenance and enhancement of competency is subject to the combined effect of many factors, ranging from personal traits to the salient features of the work environment (Fishbein and Stasson, 1990).

Research has shown that important factors in the development of growth orientation are support and rewards from the management, the incentive value of the job itself, the operational capacity of the team and work related stress (Argyris, 1990; Dubin, 1990; Hall, 1990; Kaufman, 1990; Ruohotie, 1996; Nokelainen and Ruohotie, 2003).

Management and leaders face such challenges as how to develop and reward learning, how to empower people, how to support t he development of professional identity, create careers based on interaction, set goals for learning and how to plan development, evaluate learning and its development and how to create commitment to the job and the organization.

The incentive value of the job depends on the opportunities it offers for learning, i.e. developmental challenges, the employees' chances to influence, opportunities to learn collaboratively and the dignity of the job.

The operational capacity of a team or a group can be defined by its members' capability to operate and learn together, by the work group co-operation and by the reputation for effectiveness.

Work related stress might become an obstacle to professional growth as a too heavy mental load and demand for continual alterations may stress people and suppress organizational growth and development.

Ruohotie and Nokelainen (2000) examined the theoretical dimensions of a growth-oriented atmosphere (GOA) in a Finnish vocational education high school. The organization consisted of ten geographically separate units. The sample size was 318 employees, 66% out of the survey population of 479 employees. The target population was Finnish vocational high-school personnel in 1998 ($N = 7,958$).

The instrument utilized in the study contained 80 statements. The response options in a 5-point Likert scale varied from 1 ("Strongly disagree") to 5 ("Strongly agree").

Ruohotie and Nokelainen (2000) constructed fourteen summated scales (Hair et al., 1995, p. 9) to represent the theoretical dimensions of GOA. The scales were formed on the basis of both theoretical aspects of growth-orientation (Ruohotie, 1996) and the results of exploratory factor analysis (Maximum likelihood with Varimax rotation). The 14-factor solution was the most parsimonious representing 67% of the variance within 80 items. Eigenvalues were between 1.05 and 23.98. Respondents indicated only moderate differences in preferences for various dimensions as mean ratings ranged between 3.2 and 3.8. Internal consistency for each factor was estimated with Cronbach's alpha coefficient (1970, pp. 160–161). The alpha values ranged from 0.77 to 0.93.

Although authors report continuous parameters such as mean and alpha on items measured with the non-metric ordinal scale, we consider the results plausible as the underlying phenomenon, a GOA is continuous by nature. The sample size to the number of the observed variables ratio scale in the Ruohotie and Nokelainen study (2000) was acceptable according to empirical and simulation studies (e.g., Cattell, 1978; Gorusch, 1983; MacCallum et al., 1999).

Ruohotie and Nokelainen (2000) found that GOA generates togetherness and reflects on developing leadership. Multidimensional scaling provided evidence to conclude that factors representing the incentive value of the job, commitment to work and organization, the clarity of the job and growth motivation are the strongest indicators of GOA. They made the following conclusions based on their research findings: (1) teacher's professional growth-motivation reflects directly with task value on teacher–pupil relationships and on achievement motivation, (2) task value has an effect on GOA, and (3) GOA is the highest in work assignments that offer challenging professional tasks (manager, teacher) and lowest among other workers.

2.2. BAYESIAN MODELING APPROACH

In this paper, we study two different kinds of non-linearities among the items measuring the fourteen factors of the GOA: (1) Non-linear relationships between continuous variables and (2) multi-modal relationships between continuous variables.

The term 'non-linear' is not very informative since it seems to include many different dependence patterns between random variables. As a mathematical concept, linearity (of a mapping) is well defined. In statistics, when describing the relationship between two variables as linear, we usually assume that the mean of the variables is a linear function of the means of some other variables (possibly in some special context, i.e. when certain variables are fixed). However, there are many situations where describing only these linear relationships of variables misses the important aspects of dependencies. The most self-evident shortcoming is that the dependencies between variables may be very non-linear. Moreover, if some variables are measured in the nominal scale, the concept of linearity is not meaningful at all. In ordinal scale linearity (based on means) is also often considered dubious. Linearity may also be a conceptually misleading notion even if the dependencies are mathematically linear. For example, in case of multi-modality the relationship between means may be linear, but the mean of the dependent distribution may lie in a low probability region (i.e. values close to the mean are rare).

Discrete Bayesian networks operate on a nominal finite scale, thus it is trivial that these networks are capable of modeling this type of non-linearity. Any dependence between variables one of which is measured in the nominal scale is non-linear. Consequently, non-linearity due to the scale is not studied in this paper. However, it is worth emphasizing that when data contains nominal scale variables that are not totally independent of all the other variables of the data, Bayesian networks are capable of modeling non-linearities.

Mathematically, linearity is well defined between two sets of continuous variables. However, in this paper, we only study simple non-linear relationships between two variables. In our study, the dependence between variables X and Y is considered non-linear if the mean of the conditional distribution of Y is not a monotonous (i.e. increasing or decreasing) function of X. Similarly, the dependence between variables X and Y is considered multi-modal if the mode of the conditional distribution of Y is not a monotonous function of X.

This study resembles to some extent the work by Hofmann and Tresp (1996, 1998) where they use the method of Parzen windows to allow non-linear dependencies between continuous variables. The emphasis in their work was to demonstrate the possibility to build Bayesian networks that can capture non-linear relationships. By using discretized variables, this possibility comes trivially, but our objective is to find out to what extent this possibility is used, i.e. how many and what kind of non-linearities are captured by discrete Bayesian networks.

Given the identically and independently distributed multivariate data set D over variables V and the prior probability distribution π over Bayesian networks (Pearl, 1988), the Bayesian probability theory allows us to calculate the probability $P(G|D, \pi)$ of any Bayesian network G (Heckerman et al., 1995). Different networks can then be compared by their probability. Finding the most probable Bayesian network for any given data is known to be NP-hard (Non-deterministic Polynomial-time hard) which means that the automatic discovery of the most probable network is a mission impossible (Chickering, 1996). An example of a NP-hard problem is the 'subset sum problem' (e.g., Cormen et al., 1996, p. 951): Given a set of integers, does any subset sum exactly to zero? For example, given the set $\{-5, -3, 1, 2, 9\}$, the answer is YES because the subset $\{-3, 1, 2\}$ sums to zero. Fortunately stochastic search methods have proven to be successful in finding high-probability networks (Chickering et al., 1995). Once the network G has been constructed using data D, we can use it to calculate predictive joint distributions $P(V|G, D)$. Bayesian network structure can be used to effectively calculate the conditional marginals of the predictive joint distribution for single variables, i.e. $P(V_i|A, G, D)$, where A is any subset of the variables of V. In this paper, we only study the marginals, where A is a singleton $\{V_j\}$ and there is either an arrow from V_i to V_j or an arrow from V_j to V_i (we say that V_i and V_j are adjacent in G).

The Bayesian dependence network (Heckerman et al., 1995; Silander and Tirri, 2000; Myllymäki et al., 2002) is a representation of a probability distribution over a set of random variables, consisting of an directed acyclic graph (DAG), where the nodes correspond to domain variables, and the arcs define a set of independence assumptions which allow the joint probability distribution for a data vector to be factorized as a product of simple conditional probabilities. A graphical visualization of the Bayesian network contains two components: (1) observed variables visualized as ellipses and (2) dependences visualized as lines between nodes. Variable is considered as independent of all other variables if there is no line attached to it. Such networks (see Tables II and III) are calculated in this paper using aforementioned B-Course software (Myllymäki et al., 2002). We have shown in our earlier research that Bayesian networks are useful for the explorative analysis of statistical relationships between observed variables (see, e.g., Ruohotie and Nokelainen 2000).

3. Method

3.1. SAMPLE AND PROCEDURE

The sample ($N = 726$) was collected during the year 2001 with a 69-item web-based self-report questionnaire[4] that is a revised version of the previous one (Nokelainen et al., 2002). The instrument had a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The data consists of adult employees from three Finnish vocational high schools (D21, $n=447$; D22, $n=71$; D23, $n=208$). Respondents had three different kinds of job profiles (with 4% missing data, $n=31$): Managers (6%, $n=46$), teachers (61%, $n=462$), and administrative personnel (29%, $n=223$). A respondent's nature of the contract was categorized into three classes (with 3%, $n=26$ missing data): Established (70%, $n=533$), temporary (22%, $n=169$), and part-time (5%, $n=34$) employees.

3.2. MEASURING NON-LINEARITIES

To measure non-linear dependencies captured by Bayesian networks, every variable was tested in each network by conditioning it one by one with its immediate neighbors in the network. It was observed whether the modes and means of the conditional distributions were linear and whether the conditional distributions were unimodal. Linearity of modes and means was tested by recording whether the means and modes were increasing or decreasing functions of conditioning variable. Even clear departures from line like behavior were accepted as linear as long as the direction of correlation (positive, negative) did not change. Therefore, in these experiments, a 'linear' means relationship that can be more or less adequately modeled by line describing how central tendency of the dependent variable varies as a function of the independent variable. In measuring the unimodality of conditional distributions, we judged the dependence to be unimodal if (and only if) none of the conditional distributions $P(Y|X)$ were clearly multimodal. We acknowledge the possible presence of Simpson's Paradox (e.g., Moore and McCabe, 1993, p. 190) when detecting non-linear relationships with sign changes, but prefer to label it as a statistical fact. The reversal of the direction of a comparison or an association when data from several groups are combined to form a single group becomes a paradox only when associated with dubious causal interpretation.

4. Results

4.1. RESEARCH QUESTION 1: WHAT KIND OF (AND HOW MANY) NON-LINEARITIES ARE CAPTURED BY DISCRETE BAYESIAN NETWORKS?

Research evidence based on Bayesian network modeling of six independent empirical data describing the dimensions of the GOA showed that only 22% of all dependencies were purely linear, i.e. linear mode, linear mean, unimodal. This is the best data for traditional linear analysis as no information is lost due to non-linearity. The total number of non-linear dependencies on the data was 57%. Sixteen percent of dependencies were purely non-linear (non-linear mode, non-linear mean, multimodal). Multimodality was the most common violation of linearity in all samples. The results show that Bayesian networks capture non-linear dependencies on the data, as 46% of the pairwise (unconditional) dependencies of the models are vaguely, and 23% severely, non-linear.

4.2. RESEARCH QUESTION 2: IS THERE DIFFERENCE BETWEEN THE RESULTS OF LINEAR BIVARIATE CORRELATIONS AND BAYESIAN DEPENDENCE MODELING?

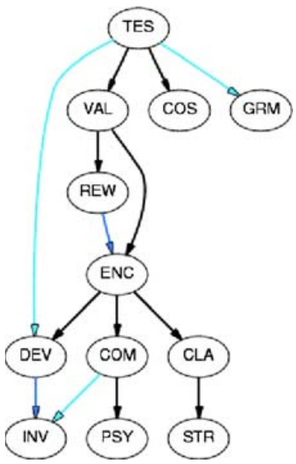
To answer the second research question, we compared the results of linear correlational analysis and non-linear Bayesian network modeling. Our hypothesis was that the results of both linear and non-linear analysis should be the same if the level of non-linearity in the sample has no practical effect. We used in this stage of the analysis only D21 ($n = 447$) and D23 ($n = 208$) samples. Those two samples were chosen as the sample sizes of the D21 and D23 data are more appropriate for the analysis than D22 ($n = 71$). The D21 data represents in this analysis linear sample, as 24% of its dependencies are unimodal and have a linear mode and mean. The D23 data is more non-linear in nature as it has 5% less similar pure linear dependencies.

We begin by analyzing the 14 GOA factors sum correlations of the linear D21 data with the Spearman rank order method (r_s) and mean correlations with Pearson's product moment method (r_p). As the results of both correlational analyzes proved to be alike, we study only the Pearson product moment correlations here.

Comparison of the BDM and correlation solutions is presented in Table I. On the left-hand side shows the visualization of the network where nodes represent variables and arches represent dependencies between them. Strength of each dependence on the model is indicated with a color; a darker color indicates a stronger statistical relationship between the two variables. Importance ranking corresponding to the color of the arcs in the final model is presented in the middle part of the table. The column on the right hand side contains the results of the correlation matrix. BDM shows nine strong and five weaker relationships between the GOA factors.

Team spirit (TES) is the most important variable in the model as it has a direct statistical relationship to the Growth motivation (GRM), Community spirit (COS), Valuation of the job (VAL) and Developing of know-how (DEV) factors. The role of the VAL factor is also important as it is a connecting node to the Rewarding of know-how (REW) factor, that in turn is connected to the Encouraging leadership (ENC) factor. If leaders encourage their subordinates, they feel more commitment to work and

Table I. Bayesian network model of the dimensions of growth-oriented atmosphere (linear sample, $n = 447$)



Network model	Dependence	Probability ratio	r_p
	TES–>COS	1:1.000.000	0.712**
	ENC–>CLA		0.709**
	ENC–>COM		0.645**
	TES–>VAL		0.608**
	VAL–>REW		0.598**
	VAL–>ENC		0.754**
	ENC–>DEV		0.764**
	CLA–>STR		0.488**
	COM–>PSY		−0.480**
	REW–> ENC	1:58863	0.639**
	DEV–> INV	1:13112	0.674**
	COM–>INV	1:23	0.645**
	TES–>DEV	1:4.67	0.541**
	TES–>GRM	1:4.59	0.232**

ENC denotes encouraging leadership, STR strategic leadership, REW know–how rewarding, DEV know–how developing, INV incentive value of the job, CLA clarity of the job, VAL valuation of the job, COS community spirit, TES team spirit, PSY psychic stress of the job, COM commitment to work and organization, and GRM growth motivation.
**Correlation is significant at the 0.01 level (2–tailed).

organization, their incentive value of the job increases and they feel less psychic stress.

The Bayesian dependence network model is generally congruent with the correlation matrix as both methods found the same two factors, namely the Build-up of work requirements (BUI) and the Students attitudes towards the teacher (STA), independent of all the other factors. Third factor not belonging to the model is the GRM as it has only a weak relationship to the TES factor. Non-linear modeling found nine strong dependencies between the factors as the correlational analysis found five. However, the results were almost identical, as all but two of the high-correlation dependencies on the matrix were also present in the Bayesian model. The missing dependencies were between the VAL and DEV factors, $r(447) = 0.718$, $p < 0.01$, and the REW and DEV factors, $r(447) = 0.650$, $p < 0.01$. However, BDM suggests that the dependence between the two factors and the DEV factor is mediated by the ENC factor. This finding, as it is repeated later in this paper with the non-linear sample, could be interpreted as superior enabling her subordinates' motivation and commitment to work. The non-linear model provides new information by revealing the relationship between the TES, VAL, and REW factors. The Bayesian model indicates that ENC has stronger effect on the clarity of the job factor than on the strategic leadership factor. The correlation matrix supports this finding as it shows higher positive correlation between the ENC and Clarity of the job factors, $r(447) = 0.709$, $p < 0.01$.

Next we compare the correlations and BDM of the non-linear data *D*23. The commitment to work and organization (COM) is the central factor in the model presented in Table II. The factor has direct connections to the Psychic stress of the job (PSY), the Incentive value of the job (INV), the Clarity of the work role (CLA), TES and ENC factors. Closer examination of the frequency distributions (not presented in Table II) shows, that employees of the vocational high school have high commitment to work and organization, TES and VAL. Employee's responses show that they are disappointed to the REW and their work roles are not explicit. As with the linear data D21, the correlation matrix shows high correlation, $r(208) = 0.699$, $p < 0.01$, between the VAL and DEV and REW factors. Again, the BDM shows that the dependency between the two variables is mediated by the ENC factor.

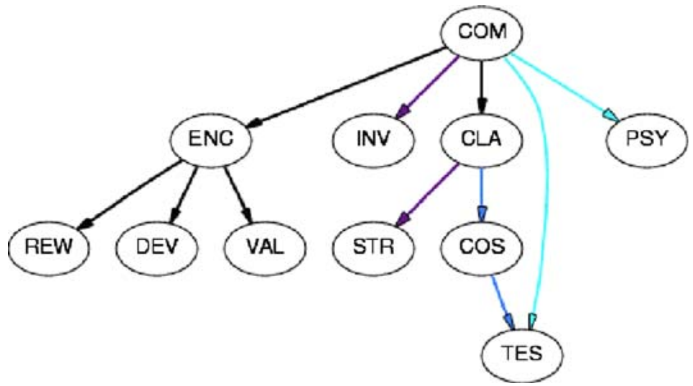
4.3. RESEARCH QUESTION 3: DOES AN EMPIRICAL SAMPLE CONTAINING PURE LINEAR DEPENDENCIES HAVE BETTER OVERALL FIT INDICES IN CFA THAN SAMPLE CONTAINING LESS LINEAR DEPENDENCIES?

We tested the GOA model fit to both samples, linear D21 and non-linear D23, with confirmatory (restricted) factor analysis. Our hypothesis was, that linear data should fit the model better than the data that contains more non-linear dependencies.

Earlier in this study both correlation and Bayesian network analysis indicated, that three dimensions of the GOA, namely "10. Students' attitudes toward teacher", "12. Build-up of work requirements" and "14. Growth motivation", are not present in two empirical data. Thus we conducted confirmatory factor analysis with the remaining 11 factors: "1. Encouraging leadership", "2. Strategic leadership", "3. Know-how rewarding", "4. Know-how developing", "5. Incentive value of the job", "6. Clarity of the job", "7. Valuation of the job", "8. Community spirit", "9. Team spirit", "11. Psychic stress of the job" and "13. Commitment to work and organization".

Table III shows the model fit indices of the confirmatory factor analysis. First section in the table presents measures of absolute fit that determine the degree to which the model predicts the observed correlation matrix (Hair et al., 1995, p. 683). The root mean square error of approximation (RMSEA) is designed to evaluate the approximate fit of the model in the

Table II. Bayesian network of the dimensions of growth-oriented atmosphere (non-linear sample, $n = 208$)



Network model	Dependence	Probability ratio	r_p
	ENC–>DEV	1:Inf.	0.742**
	ENC–>VAL		0.723**
	COM–>CLA		0.611**
	ENC–>REW		0.665**
	COM–>ENC		0.636**
	CLA–>STR	1:1.000.000	0.582**
	COM–>INV		0.693**
	CLA–>COS	1:58948	0.500**
	COS–>TES	1:7695	0.565**
	COM–>TES	1:34	0.586**
	COM–>PSY	1:1.7	−0.373**

ENC denotes encouraging leadership, STR strategic leadership, REW know–how rewarding, DEV know–how developing, INV incentive value of the job, CLA clarity of the job, VAL valuation of the job, COS community spirit, TES team spirit, PSY psychic stress of the job, and COM commitment to work and organization.

**Correlation is significant at the 0.01 level (2–tailed).

Table III. Model fit indices of the growth-oriented atmosphere model

	Absolute fit measures							Incremental fit measures	
Sample	χ^2	df	χ^2/df	p	RMSEA	$C.I.$ (90)	SRMR	TLI	CFI
D21[a]	4448.113	1375	3.235	0.000	0.071	0.068 0.073	0.071	0.835	0.853
D23[b]	2897.437	1375	2.107	0.000	0.073	0.069 0.077	0.077	0.810	0.831

RMSEA denotes root mean square error of approximation with 90% confidence interval, TLI tucker–Lewis coefficient, CFI comparative fit index.
[a]Linear sample D21 $n = 447$.
[b]Non–linear sample D23 $n = 208$.

population (Kaplan, 2000, p. 112). The estimate was in both samples below the fair fit level of 0.08 (Hair et al., 1995, p. 685), indicating good fit (Browne and Cudeck, 1993). The upper limit of the 90% confidence interval was also above the cutoff value in both samples. The standardized root mean square residuals (SRMR) help the investigator to examine how well the aspects of the data are captured by the model (Loehlin, 2004, p. 70). SRMRs were in both samples below a cut-off value of 0.08 (Hu and Bentler, 1999).

The second section in Table III presents incremental fit measures that compare the proposed model to a baseline model that all other models should be expected to exceed (Hair et al., 1995, p. 685). The Tucker–Lewis index (TLI), a.k.a. the Non-normed Fit Index (NNFI), and a similar measure, the comparative fit index (CFI), were both slightly below the recommended level of 0.90 (Tucker and Lewis, 1973) in both samples.

The results indicate that the GOA model was performing slightly better with the linear D21 sample than non-linear D23 sample. We also tested the model with CFA implemented in Mplus that uses ADF function (weighted least squares) instead of ML and allows categorical indicators (Muthén and Muthén, 2001). The fit indices for the linear sample were also with this method slightly better than for the non-linear sample.

4.4. Research question 4: Does an empirical sample containing pure linear dependencies have higher CFA parameter estimates than sample containing less linear dependencies?

The factor covariances of the GOA model presented in Table IV show that the linear sample (D21) has higher overall parameter estimates and smaller error variances than the non-linear sample. Next we will discuss the differences in strength of dependencies between the two samples in more detailed manner as we interpreted the results via four different theoretical aspects.

Table IV. Factor covariances of the growth-oriented atmosphere model

Data D21 ($n = 447$)		f01_enc	f02_str	f03_rew	f04_dev	f05_inv	f06_cla	f07_val	f08_cos	f09_tes	f11_psy	f13_com
f01_enc	Encouraging leadership	1.000										
f02_str	Strategic leadership	0.417	1.000									
f03_rew	Know−how rewarding	0.752	0.500	1.000								
f4_dev	Know−how developing	0.871	0.365	0.651	1.000							
f05_inv	Incentive value of the job	0.543	0.212	0.377	0.537	1.000						
f06_cla	Clarity of the job	0.886	0.490	0.647	0.707	0.423	1.000					
f07_val	Valuation of the job	1.116	0.391	0.720	0.827	0.585	0.799	1.000				
f08_cos	Community spirit	0.510	0.288	0.384	0.506	0.323	0.501	0.552	1.000			
f09_tes	Team spirit	0.541	0.274	0.392	0.484	0.333	0.498	0.586	0.691	1.000		
f11_psy	Psychic stress of the job	−0.361	−0.226	−0.277	−0.316	−0.387	−0.363	−0.434	−0.219	−0.240	1.000	
f13_com	Commitment to work and organization	0.686	0.360	0.495	0.570	0.575	0.560	0.725	0.366	0.370	−0.547	1.000
	Data D23 ($n = 208$)											
f01_enc	Encouraging leadership	1.000										
f02_str	Strategic leadership	0.503	1.000									
f03_rew	Know−how rewarding	0.632	0.491	1.000								
f04_dev	Know−how developing	0.736	0.509	0.677	1.000							
f05_inv	Incentive value of the job	0.462	0.243	0.373	0.518	1.000						
f06_cla	Clarity of the job	0.684	0.600	0.587	0.633	0.366	1.000					
f07_val	Valuation of the job	0.798	0.388	0.661	0.736	0.537	0.563	1.000				
f08_cos	Community spirit	0.367	0.369	0.388	0.460	0.207	0.585	0.419	1.000			
f09_tes	Team spirit	0.351	0.269	0.306	0.384	0.268	0.383	0.402	0.534	1.000		
f11_psy	Psychic stress of the job	−0.218	−0.184	−0.186	−0.179	−0.138	−0.341	−0.221	−0.304	−0.250	1.000	
f13_com	Commitment to work and organization	0.569	0.383	0.487	0.613	0.602	0.586	0.604	0.390	0.440	−0.319	1.000

Data D21 error variances range between 0.040 and 0.091. All p-values were $p < 0.001$. Data D23 error variances range between 0.055 and 0.106. All p-values range between $p < 0.001 - 0.022$.

4.5. Research question 5: Is there difference between substantive interpretations of the results of CFA and BDM with linear and non-linear samples?

Bollen (1989, p. 281) states “nonsense results for individual parameters can occur in conjunction with good overall fit measures ... ”. We examined both Bayesian dependence models (see Tables II and III) and component fit measures (see Table IV) to see if the estimates between factors make sense according to the GOA model. We will focus on the following four aspects of the model: (1) Support and rewards from the management have a positive influence on how employer experiences her know-how is being rewarded and developed and her job is valuated; (2) The incentive value of the job has a positive influence on the know-how developing and valuation of the job; (3) The operational capacity of the team and team spirit correlate positively and (4) The work-related stress hinders the development of all the other factors of the GOA.

The *first aspect* is that support and rewards from the management are in essential role in the development of growth orientation (Ruohotie, 1996). The correlational analysis shows that both the VAL and DEV factors are connected to the REW factor. The relationship is also present in corresponding Bayesian models (see Tables II and III), but in both samples the models include the ENC as a mediating component between the two factors. The CFA covariance matrix presented in Table IV shows strong positive covariances between the ENC, REW, DEV, VAL and CLA factors. The parameter estimates range between 0.752 and 1.116 (linear sample D21), and 0.634 and 0.798 (non-linear sample D23). Also Bayesian models for both data contain the aforementioned relationships.

The *second aspect* is that INC depends on the opportunities it offers for learning, i.e. developmental challenges and valuation of the job (Ruohotie, 1996). Factor covariances presented in Table III support the GOA theory in both samples (D21 $\phi_{\mathrm{INV\_DEV}} = 0.537$; D21 $\phi_{\mathrm{INV\_VAL}} = 0.585$; D23 $\phi_{\mathrm{INV\_DEV}} = 0.518$; D23 $\phi_{\mathrm{INV\_VAL}} = 0.537$). The Bayesian model for the linear sample (D21) shows only partial support for the second theoretical assumption as there is a connection between the INV and the DEV factors, but no direct connection exists between the VAL factor and the two other factors (Table I). The non-linear sample (D23) does not show any evidence that supports the GOA theory (Table II). In both Bayesian models the ENC factor acts as a mediator between the variables under investigation.

The *third aspect* is the relationship between the COS and TES factors. According to Argyris (1992) and Ruohotie (1996), community members should discuss about developing components in their work and learn from each other. Factor covariances support the GOA theory in both samples:

Table V. Summary of the results based on theoretical assumptions

Theoretical assumptions		D21[a] BDM[c]	D21[a] CFA[d]	D23[b] BDM[c]	D23[b] CFA[d]
1. Support and rewards from the management	ENC(+) → REW, VAL,DEV	●[e]	◙[g]	◙[g]	◙[g]
2. The incentive value of the job	INV (+) → DEV, VAL	○[f]	●[e]	—[h]	●[e]
3. Operational capacity of the team	COS (+) ↔ (+) TES	◙[g]	◙[g]	●[e]	●[e]
4. Work–related stress	PSY (–) → All the other factors	○[f]	◙[g]	○[f]	●[e]

Note. ENC denotes Encouraging leadership, REW know-how rewarding, DEV know-how developing, INV incentive value of the job, VAL valuation of the job, COS community spirit, TES team spirit, and PSY psychic stress of the job.
[a]Linear sample D21 $n = 447$. [b]Non–linear sample D23 $n = 208$. [c]Bayesian Dependence Model. [d]Confirmatory Factor Analysis. [e]Research evidence supports the theoretical assumption. [f]Research evidence supports the theoretical assumption partially. [g]Research evidence supports the theoretical assumption strongly. [h]Research evidence does not support the theoretical assumption.

$\phi_{COS\_TES} = 0.691$ (D21), and $\phi_{COS\_TES} = 0.534$ (D23) (Table V). The dependence between the two factors is also present in both Bayesian dependence models (Tables II and III).

The *fourth aspect* is that work related PSY hinders people to give their best performance in the work, and may thus become an obstacle to professional growth (see, e.g., Ruohotie, 1996; Edwards and Rothbard, 1999). Table IV shows that factor covariances between the PSY and the other factors are stronger in the linear sample, as the parameter estimates range between -0.219 and -0.547 (linear sample D21), and -0.138 and -0.319 (non-linear sample D23). The highest (negative) covariances in both samples are between PSY and COM. The dependence between PSY and COM is also present in both Bayesian dependence models, but no other support for the fourth theoretical aspect is present (Tables II and III).

The results are summarized in Table V. First notion is that both Bayesian dependency models do not support the second theoretical assumption about the relationship between INV and DEV and VAL. However, INV and VAL have a statistical dependence also in BDM derived from the linear sample (Table I). The second notion is that the fourth theoretical assumption about the negative influence of PSY on all the other factors is only partially supported in both Bayesian models. Finally, the linear sample (D21) has in most cases higher CFA parameter estimates than the non-linear sample (Table IV). This is a natural finding as in CFA only linear factor covariances between variables are estimated.

5. Discussion

This study investigated the number of linear and non-linear dependencies between the items measuring 14 dimensions of the GOA. An empirical sample represented employees of three Finnish vocational high schools ($N = 726$). Bayesian theory was discussed and Bayesian dependence models for discrete data were introduced as a model family capable of describing non-linearities. Next, empirical data was analyzed to find out if non-linear dependencies weaken the robustness of bivariate linear statistical methods (represented by correlation analysis) when compared with non-linear modeling (represented by Bayesian dependency modeling).

Investigation of empirical data ($N = 726$) showed that only 22% of all dependencies between variables were purely linear (linear mode,linear mean, unimodal). Sixteen percent of dependencies were purely non-linear (non-linear mode, non-linear mean, multimodal). Multimodality was the most common reason for the violation of linearity in both data sets.

Investigations were continued with two sub samples of the vocational high school data, namely D21 ($n = 447$) and D23 ($n = 208$). The D21 sample represents in this study linear empirical data with 23.9% of pure linear and 15.0% of pure non-linear dependencies and the D23 sample represents non-linear data with only 16.2% of pure linear dependencies and 18.3% of pure non-linear dependencies.

The subject domain interpretations of linear correlational analysis and non-linear BDM were compared to learn if the results would lead to different subjective interpretations. The results showed that in general Bayesian network models were congruent with the correlation matrixes as both methods found the same variables independent of all the other variables. However, non-linear modeling found with both linear and non-linear samples a greater number of strong dependencies between the GOA factors. Comparison of the correlations and dependencies in Bayesian networks showed, that in both samples linear correlations indicated a direct connection between REW, DEV and VAL, as Bayesian models indicated indirect connections between the variables encouraging leadership acting as a mediator in-between.

Further, we focused on the following four aspects of the GOA theory as our motivation was to investigate if there were differences between the results of linear (CFA) and non-linear (BDM) analysis with the linear and non-linear samples: (1) Support and rewards from the management have a positive influence on how employer experiences her know-how is being rewarded and developed and her job is valuated; (2) The INC has a positive influence on the DEV and VAL; (3) The operational capacity of the team and team spirit correlate positively; (4) The work-related PSY hinders the development of the GOA.

Results showed that the analysis techniques produced similar results for two out of four theoretical aspects, namely the first and third. Different results leading to different substantive interpretations were considered for the second and fourth theoretical aspect as follows. The BDM was able to find only partial support from the linear and no support at all from the non-linear sample for the assumption that the INC would have a positive influence on the DEV and VAL. Both Bayesian dependency models suggested that the components under investigation are not directly related, but instead indirectly connected to each other via encouraging leadership. The fourth theoretical aspect was supported in both linear analyses, as the PSY factor was negatively related all the other factors. However, in both Bayesian dependency models the PSY factor was related only to commitment to work and organization. Finally, linear methods (i.e., bivariate *r* and CFA) found stronger statistical relationships between factors measuring the GOA from a linear than non-linear sample. We fully agree with Grilli and Rampichini (2004, submitted for publication) when they state that use of a proper model is always a desirable feature of the analysis as thus we may expect the resulting inferences to be generally more reliable.

Notes

1. BUGS is available at http://www.mrc-bsu.cam.ac.uk/bugs/
2. DSIGoM is available at http://www.dsisoft.com/grade_of_membership.html
3. B-Course is available at http://b-course.hiit.fi
4. Growth-oriented Atmosphere Questionnaire (GOAQ) is available at http://www.uta.fi/aktkk/goaq/

References

Arbuckle, J. L. (1999). *Amos for Windows. Analysis of Moment Structures*. Chicago, IL: SmallWaters.

Argyris, C. (1992). *On Organizational Learning*. Cambridge, MA: Blackwell Publishers.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society* 53: 370–418.

Bentler, P. M. (1995). *EQS Structural Equations Program Manual*. Encino, CA: Multivariate Software Inc.

Bernardo, J. & Smith, A. (2000). *Bayesian Theory*. New York: Wiley.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology* 37: 1–21.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In: K. A. Bollen and S. Long (eds.), *Testing Structural Equation Models*. Newbury Park, CA: Sage, pp. 136–162.
Cattell, R. B. (1978). *The Scientific Use of Factor Analysis in Behavioral and Life Sciences*. New York, NY: Plenum Press.
Chickering, D. M. (1996). Learning Bayesian networks is NP–complete. In: D. Fisher and H. J. Lenz (eds.), *Learning from Data: Artificial Intelligence and Statistics*. Berlin: Springer Verlag, pp. 121–130.
Chickering, D. M., Geiger, D. & Heckerman, D. (1995). Learning Bayesian networks: Search methods and experimental results. *Proceedings of Fifth Conference on Artificial Intelligence and Statistics*, Ft. Lauderdale, FL, USA, pp. 112–128.
Congdon, P. (2001). *Bayesian Statistical Modelling*. London: Wiley.
Congdon, P. (2003). *Applied Bayesian Modelling*. London: Wiley.
Cormen, T. H., Leiserson, C. E. & Rivest, R. L. (1996). *Introduction to Algorithms*, 16th edn. Cambridge, MA: MIT Press.
Cronbach, L. J. (1970). *Essentials of Psychological Testing*, 3rd edn. New York, NY: Harper & Row.
DeVellis, R. F. (2003). *Scale Development. Theory and Applications*, 2nd edn. Thousand Oaks, CA: Sage.
Dubin, S. (1990). Maintaining competence through updating. In: S. Willis and S. Dubin (eds.), *Maintaining Professional Competence*. San Francisco: Jossey-Bass, pp. 9–43.
Edwards, J. R. & Rothbard, N. P. (1999). Work and family stress and well-being: an examination of person–environment fit in the work aznd family domains. *Organizational Behavior and Human Decision Processes* 77(2): 85–129.
Fishbein, M. & Stasson, M. (1990). The role of desires, self–predictions, and perceived control in the prediction of training session attendance. *Journal of Applied Social Psychology* 20: 173–198.
Gorusch, R. (1983). *Factor Analysis*, 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
Grilli, L. & Rampichini, C. (2004). Multilevel factor models for ordinal variables. (Submitted for publication).
Hair, J. F., Anderson, R. E., Tatham R. L. & Black, W. C. (1995). *Multivariate Data Analysis*, 4th edn. Englewood Cliffs, NJ: Prentice Hall.
Hall, D. (1990). Career development theory in organizations. In: D. Brown and L. Brooks (eds.), *Career Choice and Development*. San Francisco: Jossey–Bass, pp. 422–454.
Heckerman, D., Geiger, D. & Chickering, D.M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning* 20: 197–243.
Hofmann, R. & Tresp, V. (1996). Discovering structure in continuous variables using Bayesian networks. In: D. S. Touretzky, M. C. Mozer and M. E. Hasselmo (eds.), *Advances in Neural Information Processing Systems*, vol. 8. Cambridge, MA: MITPress, pp. 500–506.
Hofmann, R. & Tresp, V. (1998). Non–linear Markov networks for continuous variables. In: M. I. Jordan, M. S. Kearns and S. A. Solla (eds.), *Advances in Neural Information Processing Systems*, Vol. 10. Cambridge, MA: MITPress, pp. 521–527.
Hu, L. & Bentler, P. (1999). Cut–off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* 6(1): 1–55.
Johnson, D. R. & Creech, J. C. (1983). Ordinal measures in multiple indicator models: a simulation study of categorization error. *American Sociological Review* 48: 398–407.

Jöreskog, K. G. (2003). *Structural Equation Modeling with Ordinal Variables using LISREL*, Retrieved 13.2.2005 from: http://www.ssicentral.com/lisrel/ordinal.htm

Jöreskog, K. G. & Sörbom, D. (1998). *LISREL 8 Program Reference Guide*. Chicago, IL: Scientific Software International.

Kaplan, D. (2000). *Structural Equation Modeling*. Foundations and Extensions. Thousand Oaks: Sage.

Kaufman, H. (1990). Management techniques for maintaining a competent professional work force. In: S. Willis and S. Dubin (eds.), *Maintaining Professional Competence*, San Francisco: Jossey–Bass, pp. 249–261.

Lawler, E. E. (1994). From job–based to competence–based organizations. *Journal of Organizational Behaviour* 15: 3–15.

Loehlin, J. C. (2004). *Latent Variable Models*, 4th edn. Mahwah, NJ: Lawrence Erlbaum Associates.

MacCallum, R. C., Widaman, K. F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods* 4: 84–99.

Moore, D. S. & McCabe, G. P. (1993). *Introduction to the Practice of Statistics*, 2nd edn. New York: Freeman.

Muthén, B. O. (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics* 22: 48–65.

Muthén, B. O. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* 49: 115–132.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika* 54: 557–585.

Muthén, B. O. (1993). Goodness of fit with categorical and other non–normal variables. In: K. A. Bollen and J. S. Long (eds.), *Testing Structural Equation Models*. Newbury Park, CA: Sage, pp. 205–243.

Muthén, B. O. & Kaplan D. (1985). A comparison of some methodologies for the factor analysis of non–normal Likert variables. *British Journal of Mathematical and Statistical Psychology* 38: 171–189.

Muthén, L. K. & Muthén, B. O. (2001). *Mplus User's Guide*, 2nd edn. Los Angeles, CA: Muthén & Muthén.

Myllymäki, P., Silander, T., Tirri, H. & Uronen, P. (2002). B–course: a web–based tool for Bayesian and causal data analysis. *International Journal on Artificial Intelligence Tools* 11(3): 369–387.

Nokelainen, P. & Ruohotie, P. (2003). Modeling the prerequisites of empowerment. In: B. Beairsto, M. Klein and P. Ruohotie (eds.), *Professional Learning and Leadership*. Hämeenlinna, FI: RCVE, pp. 147–176.

Nokelainen, P., Ruohotie, P. & Tirri, H. (2002). *Visualization of Growth–oriented Atmosphere.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Olsson, U.H., Foss, T., Troye, S.V. & Howell, R.D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling* 7(4): 557–595.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. San Mateo, CA: Morgan Kaufmann.

Ruohotie, P. (1996). Professional growth and development. In: K. Leithwood, S. Chapman, D. Carson, P. Hollinger and A. Hart (eds.), *International Handbook of Educational Leadership and Administration*. Dordrecht: Kluwer Academic Publishers, pp. 419–445.

Ruohotie, P. & Nokelainen, P. (2000). Beyond the Growth–oriented Atmosphere. In: B. Beairsto and P. Ruohotie (eds.), *Empowering Teachers as Lifelong Learners*, Hämeenlinna, FI: RCVE, pp. 147–167.

Silander, T. & Tirri, H. (2000). *Model Selection for Bayesian Networks*. Paper presented at the Annual Meeting of American Educational Research Association, New Orleans, USA.

Tucker, T. L. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrica* 38: 1–10.

Yung, Y. F. & Bentler, P. M. (1994). Bootstrap corrected ADF test statistics in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology* 47: 63–84.