Chapter 4

# Evaluating Uncertainty of Model Acceptability in Empirical Applications: A Replacement Approach

Luigi Lombardi[1], Massimiliano Pastore[2] and Massimo Nucci[3]

[1] *Università di Trento, Dipartimento di Scienze della Cognizione e della Formazione. Trento, Italia.*
[2] *Università di Cagliari, Dipartimento di Psicologia. Cagliari, Italia.*
[3] *Università di Padova, Dipartimento di Psicologia. Padova, Italia.*

## 4.1 Introduction

A major problem in psychological measurements is that in some circumtances there is no basis to assume that subjects are responding honestly. Some individuals actually tend to distort their responses in order to reach specific goals. For example, in personnel selection some subjects are likely to fake a personality questionnaire to match the ideal candidate's profile (positive impression management). Similarly, in the administration of diagnostic tests individuals often attempt to malinger posttraumatic stress disorder (PTSD) in order to secure financial gain and/or treatment, or to avoid being charged with a crime.

Possible fake data confront the researcher with a crucial question: If data included fake datapoints, would the answer to the research question be different from what it is? Even in the clearest case -- that is, randomly fake data -- the answer is not necessarily obvious, as even the random perturbation of data constitutes a biased information which decreases the efficiency of parameter estimates and weakens the accuracy of statistical results.

A case of particular empirical interest is the situation in which a researcher wants to evaluate the uncertainty associated to the acceptability of a given model as a result of propagation through the model of errors in input data. For example, within a factorial modeling approach, we might be interested in studying how acceptability of a target model varies as a function of different error levels in the original data set. In that case the crucial question can be rewritten as: If the data contained $q\%$ fake datapoints, what would the probability be that the model is still an acceptable one? A fake-scenario analysis may be considered as a supplementary analysis a researcher can run in order to broaden the sources of information she/he is interested in. However it looks clear that the usefulness of this approach becomes evident only if a researcher feels pretty confident about the consistency of the target model. That is to say, we need a model that adequately reproduces the underlying process of interest.

The issue of perturbations in real data has been substantially neglected in evaluating the uncertainty of model acceptability in covariance structure modeling. In this paper we attempt to contribute to the modeling of methods of treating possible fake data in structural equation models. In particular, our study examines the uncertainty associated with the acceptability of a simple well fitting factorial model. A new approach, called SGR (Sample Generation by Replacements), is developed in order to provide a perturbation model and a sampling procedure to generate a structured collection of perturbations.

Section 2 of this paper will first outline the basic principles of the new replacement approach. Section 3 will then present an illustrative application of the SGR approach to a simple factorial model. Finally, Section 4 will discuss the relation of the SGR method with Monte Carlo simulation studies. At the end of the section some possible extensions of the SGR approach are also outlined.

## 7.2 The Method of Replacements

SGR is a combinatorial method that can be applied to discrete data with a restricted number of values (e.g. Likert scale) and consists of two different components:

1 a perturbation model,

2 a sampling procedure to generate perturbed samples from a given real data set.

### 7.2.1 Basic Elements

**Perturbed data matrices** In many social and psychological surveys the resulted dataset often includes incomplete records (missing data) and/or fake records (fake data). In particular, as regards the fake-data problem, we think of the full dataset as being represented by an $n \times m$ matrix $\mathbf{D}$ (that is, $n$ observations, each containing $m$ elements), of which a certain portion $\mathbf{D}^f$ is actually fake-data. The fake-portion $\mathbf{D}^f$ of $\mathbf{D}$ together with the uncorrupted portion $\mathbf{D}^u$ of $\mathbf{D}$, constitutes the full data set, that is to say $\mathbf{D} = \mathbf{D}^f \cup \mathbf{D}^u$. The exact fake-portion $\mathbf{D}^f$ of $\mathbf{D}$ is assumed to be an unknown parameter and only the number $\varrho$ of fake data points in $\mathbf{D}$ is supposed to be known. The general idea is the following: in order to analyze the data and provide an uncertainty analysis of some statistic of interest we replace some portions $\mathbf{D}_1, \ldots, \mathbf{D}_s$ of $\mathbf{D}$, each of which contains exactly $\varrho$ elements, with new components $\mathbf{X}_1^r, \ldots, \mathbf{X}_s^r$ in such a way that for all $h = 1, \ldots, s$, all the corresponding elements in $\mathbf{X}_h^r$ and $\mathbf{D}_h$ are different. In the SGR approach these new components are generated from an appropriate population, and, therefore, the complete datasets $\mathbf{X}_1, \ldots, \mathbf{X}_s$ (with $\mathbf{X}_h = \mathbf{X}_h^r \cup \mathbf{D}_h^u$; $h = 1, \ldots, s$), are analyzed. We call the data matrices $\mathbf{X}_h$ and $\mathbf{X}_h^r$ the $h^{th}$-perturbed matrix of $\mathbf{D}$ and the $h^{th}$-replaced portion of $\mathbf{D}$, respectively.

**Uncertainty evaluation** Let $\mathcal{M}$ be a well fitting statistical model for the original data set $\mathbf{D}$ (for simplicity the model is assumed to be consistent). More precisely, it is assumed that, conditional upon $\mathbf{D}$, $\mathcal{M}$ satisfies some opportune model acceptability criterion $\phi_{\mathcal{M}}$. We may think of $\phi_{\mathcal{M}}$ as a mapping from the sample space $\mathcal{X}_{n\times m}$ into the Boolean set $\{0, 1\}$, where $\phi_{\mathcal{M}}(\mathbf{X}) = 1$ denotes that $\mathcal{M}$ is an acceptable model for $\mathbf{X}$. For example, within the structural equation modeling approach, we may think of $\mathcal{M}$ and $\phi_{\mathcal{M}}$ as a confirmatory factorial model and as a conjuctive combination of g.o.f. (goodness-of-fit) constraints, respectively. The main goal of a replacement analysis is the evaluation of $\phi_{\mathcal{M}}$ under the perturbed sample space $\mathcal{X}^*_{n\times m} = \{\mathbf{X}_h : h = 1, \ldots, s\} \subset \mathcal{X}_{n\times m}$. This idea is summarized in Figure 4.1

In the next section we will introduce a new simple replacement procedure that comes down to the most elementary model instantiation of the SGR approach.

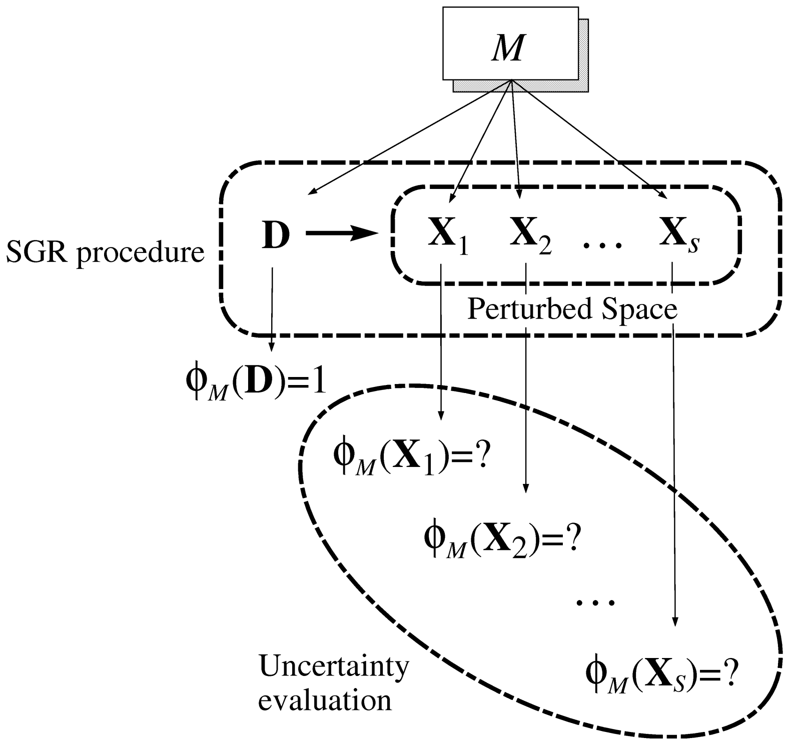


*Figure 4.1.* Overall scheme of a replacement approach.

### 7.2.2 The SGR Model

Let us assume that the entry $d_{ij}$ of $\mathbf{D}$ takes values on a possibly small set $V = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{N}$. Then, the set of all the perturbed matrices $\mathbf{X}$ with exactly $\varrho$ replacements can be derived by means of the following simple procedure. Define the set

$$S_\varrho(\mathbf{D}) = \{\mathbf{X} \in \mathcal{X}_{n\times m} : L_0(\mathbf{D}, \mathbf{X}) = \varrho\} \tag{1}$$

where $\mathcal{X}_{n\times m}$ and $L_0$ denote the family of all possible $n \times m$ matrices in $V$ and the so called $L$-zero norm (Chaturvedi, Green and Carrol, 2001), respectively. $S_\varrho(\mathbf{D})$ is called the *circle* in $\mathcal{X}_{n\times m}$ with center $\mathbf{D}$ and radius $\varrho$. Notice that $L_0$ defines a counting metric in $\mathcal{X}_{n\times m}$ in that in the limiting case as $p \to 0$, the $L_p$

norm-based function

$$L_p(\mathbf{D},\mathbf{X}) = \sum_{i=1}^{n}\sum_{j=1}^{m} |d_{ij} - x_{ij}|^p \tag{2}$$

simply counts the number of mismatches in the matrices $\mathbf{D}$ and $\mathbf{X}$. More precisely, the $L_p$ norm approaches a counting metric as $p \to 0$; the $L_0$ metric is then defined as this limiting case.

Next the proportion

$$\pi_\varrho = \frac{\#\{\mathbf{X} \in S_\varrho(\mathbf{D}) : \phi_{\mathcal{M}}(\mathbf{X}) = 1\}}{\#S_\varrho(\mathbf{D})} \tag{3}$$

is evaluated, which represents a measure of the uncertainty of the acceptability criterion $\phi_{\mathcal{M}}(\mathbf{D})$ under $\varrho$ fake datapoints in $\mathbf{D}$.

It is straightforward to verify that the cardinality $\#S_\varrho(\mathbf{D})$ (also called the *perimeter* of the circle $S_\varrho(\mathbf{D})$) is a function of the binomial coefficient in that

$$\#S_\varrho(\mathbf{D}) = \binom{nm}{\varrho}(k-1)^\varrho \tag{4}$$

where $k$ denotes the number of elements in the value set $V$.

SGR grounds on two basic assumptions: ($\mathcal{A}_1$) the *principle of indifference* and ($\mathcal{A}_2$) the *number $\varrho$ of perturbed units* in the data sample $\mathbf{D}$. The first assumption reflects the fact that in the absence of further knowledge (a) all entries in $\mathbf{D}$ are assumed to be equally likely in the process of replacement (b) given an entry $d_{ij} = v \in V$ of $\mathbf{D}$ (with $V$ being the admissible value set for $\mathbf{D}$), the probability $p(v')$ for an element $v' \in V \setminus \{v\}$ to replace $v$ in the $(ij)$-cell of $\mathbf{D}$ is assumed equal to $\frac{1}{\#V-1}$. In other words, SGR assumes the random world model described in Bacchus et al. (1994). Therefore, SGR can be used whenever we deal with randomly fake-data. The second assumption pertains to the number $\varrho$ of perturbed units to represent in the model. The choice of the amount obviously depends on the availability of external knowledge about process faking. For example, in a personnel selection context this quantity could be represented by the supposed maximal number of fake answers in a personality questionnaire. In order to stress the importance of the $L_0$ metric we have renamed the procedure $L_0$-SGR (Sample Generation by Replacements under $L_0$ metric).

**Sampling approximation of $S_\varrho(\mathbf{D})$** The reader may notice that the number of perturbed matrices $\mathbf{X}$ in $S_\varrho(\mathbf{D})$ can be very large, depending on the size of the original dataset $\mathbf{D}$, the assumed quantity $\varrho$ and the size $k$ of the value set $V$. For this reason, in the evaluation of the proportion $\pi_\varrho$, rather than using all possible matrices in $S_\varrho(\mathbf{D})$, we resort to generating random samples of $S_\varrho(\mathbf{D})$. More specifically we use a pseudorandom number generator to repeatedly replace the original $h$ $(h = 1, \ldots, \varrho)$ entries of $\mathbf{D}$ with alternative values in $V$. In such a way, given a random sample $R_\varrho$ of $S_\varrho(\mathbf{D})$ with a sufficiently large size $s$, the sample proportion $\pi^*_\varrho$ might boil down to a satisfactory approximation of the population proportion $\pi_\varrho$. In particular, an approximate confidence interval for $\pi_\varrho$ based on a normal distribution is given by $\pi^*_\varrho \pm t\sqrt{\text{var}^*(\pi^*_\varrho)}$ where $t$ is the upper $\alpha/2$ point of the $t$-distribution with $s - 1$ degrees of freedom and $\text{var}^*(\pi^*_\varrho)$ is an unbiased estimator of the variance of $\pi_\varrho$ (see, for example, Thompson 2002, 39-40). To obtain an estimator $\pi^*_\varrho$ having probability at least $1-\alpha$ of being no further than a selected value $d$ from the population proportion $\pi_\varrho$, the sample size formula based on the normal approximation gives

$$s = \frac{0.5N}{0.25 + (N+1)(d^2/z^2)},$$

where $N = \#S_\varrho(\mathbf{D})$ and $z$ is the upper $\alpha/2$ point of the normal distribution.

## 7.3 Empirical Data Example

In this exploratory study we tested the new procedure on data from a study in the personality domain. The current section is divided into three subsections: the first introduces the empirical data set and the factorial model; the second discusses the use of the $L_0$-SGR model to generate the family of perturbed datasets; and the third evaluates the acceptability criterion with respect to percentage of perturbed entries in the original dataset.

### 7.3.1 Original Dataset and CFA Model

We illustrate the entire procedure using data collected by Vidotto and Marchesini (2000) on the interrelation between personality and student learning[2]. Partecipants were 351 undergraduate students at the University of Modena (Italy). Ages ranged from 18 to 31, with a mean of 21.01 and a standard deviation of 2.28. Data consisted of responses to 4 of the 155 items of the Modena Resources Personality Inventory (MRPI) (Vidotto and Marchesini, 2000) scored on a 5-point *agree-disagree* scale. The four items in Table 4.1 were used as operational indicators of the theoretical construct *emotional instability*. This psychological construct was validated in a series of factorial studies that showed the plausibility of the factorial model depicted in Figure 4.2 (Vidotto and Marchesini, 2000).

[2]We are grateful to Giulio Vidotto and Cristina Marchesini for providing us with such a data set.

*Table 4.1.* Operational indicators of the theoretical construct *emotional instability* (Vidotto and Marchesini, 2000).

| *Item* | *Description* |
|---|---|
| 1 | Several times I gave up because what I wanted to reach was too difficult. |
| 2 | I am never really relaxed. |
| 3 | When I think of my future I have negative feelings. |
| 4 | I have difficulties in sleeping because I cannot stop thinking of my problems. |

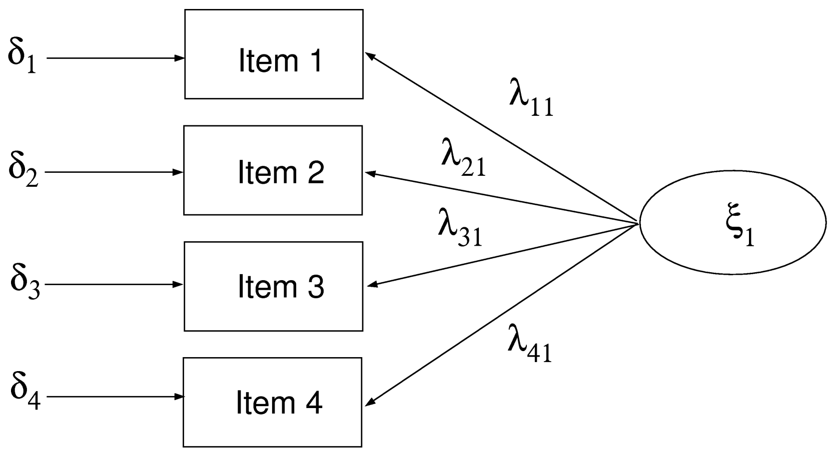


*Figure 4.2.* Factorial model.

The resulting $(351 \times 4)$ data matrix $\mathbf{D}$ was subjected to a confirmatory factor analysis (CFA) based on the above model. Results of the CFA were as follows: $\chi^2(2) = 2.295$, CFI=0.988, NNFI=0.996 and SRMR=0.02, respectively. The chi-square test statistic was not significant ($p = 0.318$). According to commonly accepted cutoff values, all considered fit indices indicated a satisfactory fit for the CFA model.

### 7.3.2 SGR Analysis

**Variables and acceptability criteria** Two different variables were considered in our analysis

1 the proportion $\epsilon$ of *perturbed units* in $\mathbf{D}$ with $40$ different levels: $\frac{1}{40}$, $\frac{2}{40}$, $\frac{3}{40}, \ldots, \frac{40}{40}$ as the independent variable,

2 the proportion $\pi_\varrho$ of *acceptable models* in $R_\varrho$ as the dependent variable with $\varrho$ being the integer approximation of $\epsilon(351 \times 4)$.

Unfortunately, since different definitions of model acceptability have been proposed within the context of structural equation modeling (Bentler and Bonnet, 1980; Hu and Bentler, 1999; Mulaik et al, 1989), there are no consistent standards for what is considered an acceptable model. In this study we considered three different acceptability criteria. The first criterion ($\phi^{(1)}_{\text{CFA}}$) was simply the $\chi^2$ statistic which provided a baseline fit criterion for our analysis. The remaining two criteria ($\phi^{(2)}_{\text{CFA}}$ and $\phi^{(3)}_{\text{CFA}}$) were based on Hu and Bentler's two-index presentation strategy. This strategy has proven to retain relatively acceptable proportions of simple and complex models and reject reasonable proportions of various types of misspecified models in most conditions (Hu and bentler, 1999). More formally the criteria were defined as follows

$$
\begin{aligned}
\phi^{(1)}_{\text{CFA}}(\mathbf{X}) &= \begin{cases} 1 & \text{if} \quad \chi^2(2) < 5.99 \\ 0 & \text{if} \quad \chi^2(2) \geq 5.99 \end{cases} \\
\phi^{(2)}_{\text{CFA}}(\mathbf{X}) &= \begin{cases} 1 & \text{if} \quad \text{CFI} \geq 0.96 \quad \text{and} \quad \text{SRMR} < 0.09 \\ 0 & \text{if} \quad \text{CFI} < 0.96 \quad \text{and} \quad \text{SRMR} \geq 0.09 \end{cases} \\
\phi^{(3)}_{\text{CFA}}(\mathbf{X}) &= \begin{cases} 1 & \text{if} \quad \text{NNFI} \geq 0.95 \quad \text{and} \quad \text{SRMR} < 0.09 \\ 0 & \text{if} \quad \text{NNFI} < 0.95 \quad \text{and} \quad \text{SRMR} \geq 0.09 \end{cases}
\end{aligned}
$$

In its basic form, a large value of the chi-square statistic, relative to its degrees of freedom, is evidence that the model does not give a very good description of the data, whereas a small chi-square is evidence that the model is a good one for the data. In particular, in our study the CFA model was rejected whenever the statistic exceeded the value of 5.99, that is the $(1 - \alpha)$-percentile of the chi-square distribution with two degrees of freedom and $\alpha = 0.05$.

Unlike the chi-square test statistic, the *Comparative Fit Index* (CFI) and the *Nonnormed Fit Index* (NNFI, Bentler and Bonnet, 1980) offer a way to quantify the degree of fit along a continuum. In particular, they are incremental fit indices that measure the proportionate improvement in fit by comparing a target model with a more restricted nested baseline model. The CFI measures the improvement in noncentrality in going from the target model to the baseline model. Likewise the NNFI index can be used to compare either alternative models or a proposed model against a null model. Finally the *Standardized Root Mean Square Residual* (SRMR) is an absolute-fit index that directly assesses how well an *a priori* model reproduces the sample data. Hu and Bentler (1999) showed that a cutoff value close to .96 for CFI (resp. to .95 for NNFI) in combination with SRMR $>$ .09 resulted in the least sum of Type I and Type II error rates.

**Proportion estimates** In order to compute the estimate $\pi_{\varrho}^{*(i)}$ $(\forall i = 1, 2, 3)$ we resort to generating a family $R_{\varrho} \subset S_{\varrho}(\mathbf{D})$ of 3000 different perturbed matrices $\mathbf{X}$ with exactly $\varrho$ replacements in accordance to the procedure described in Section 2.1. Hence, $3000 \times 40 = 120000$ different perturbed matrices were constructed in the complete replacement design. Next the estimate $\pi_{\varrho}^{*(i)}$ was computed by

$$\pi_{\varrho}^{*(i)} = \frac{\#\{\mathbf{X} \in R_{(\varrho)} : \phi_{\text{CFA}}^{(i)}(\mathbf{X}) = 1\}}{3000}. \tag{5}$$

Notice that a sample size = 3000 would be sufficient to guarantee an estimate $\pi_{\varrho}^{*(i)}$ within distance $d = 0.05$ from the true proportion $\pi_{\varrho}^{(i)}$ with probability $0.95$ ($\alpha = 0.05$).

**Results** Figure 4.3 shows $\pi_{\varrho}^{*(1)}$ (proportion of acceptable models under $\phi^{(1)}$) as a function of $\epsilon$ (proportion of replacements). The upper bound for $\pi_{\varrho}^{*(1)}$ was considered the proportion of *acceptable solutions* (that is solutions for which not improper parameter estimates occur). As espectd, the percentage of acceptable models decreased with larger percentage of replaced elements in $\mathbf{D}$. In particular, $\pi_{\varrho}^{*(1)}$ converged towards an approximate value of 0.67 as $\epsilon \to 0.50$. Moreover, the difference between the percentage of acceptable solutions and the percentage of acceptable models also decreased with larger percentage of replaced elements. Interestingly, $\pi_{\varrho}^{*(1)}$ resulted larger than 0.90 as $\epsilon \leq 0.25$. That is to say that more than 90% of the models resulted to be acceptable when a maximum of 25% of randomly fake data were generated (see also Tab 1.2, second column).

Likewise $\pi_{\varrho}^{*(2)}$ was negatively related with percentage of replacements (see Figure 4.4). The second criterion showed a stronger difference between the

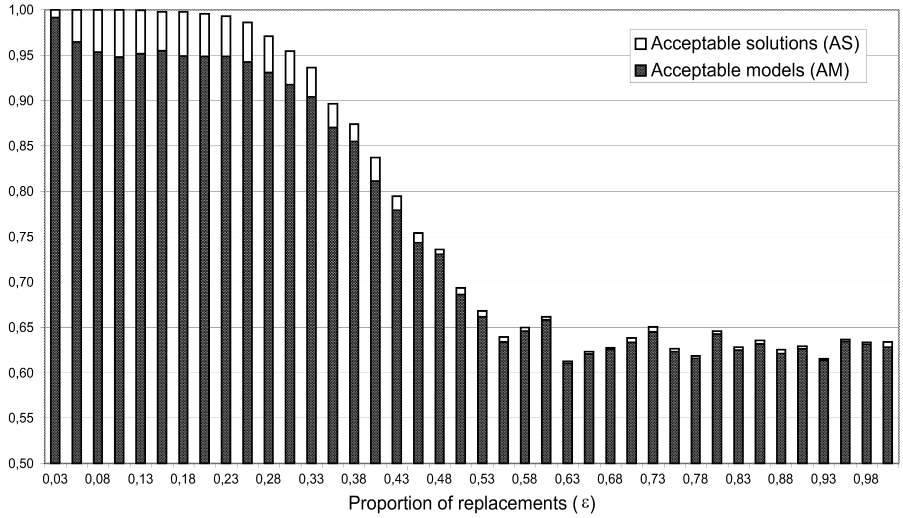


*Figure 4.3.* Parameter $\pi_{\varrho}^{*(1)}$ as a function of percentage of replacements - criterion $\phi^{(1)}$. The percentage of acceptable solutions is the upper bound for $\pi_{\varrho}^{*(1)}$ (superimposed bar-charts).

proportion of acceptable solutions and the proportion of acceptable models (left part of Figure 1.4, $\epsilon \leq 0.30$). Therefore $\phi^{(2)}$ resulted in a more conservative criterion than the Chi-square statistic. This was in accordance with Hu and Bentler's results (Hu and Bentler, 1999) (see also Tab. 1.2, third column). Finally $\phi^{(3)}$ was shown to be the most conservative of the three criteria: $\pi_{\varrho}^{*(3)} \in [0.72 - 0.82]$ with $\epsilon \leq 0.25$ (see Figure 4.5 and Table 4.2, fourth column).

Figure 4.6 shows the acceptability patterns associated with $\chi^2$, CFI, NNFI and SRMR. A dominance relation can be read from Figure 4.6 as follows

$$\text{NNFI} \succ \text{CFI} \succ \chi^2 \succ \text{SRMR} \sim \text{AS},$$

where $X \succ Y$ denotes that $X$ is more conservative than $Y$. Notice that SRMR and AS (proportion of acceptable solutions) share the same pattern. Hence $\phi^{(2)}$ and $\phi^{(3)}$ come down to CFI and NNFI, respectively. This turns out to the linear order $\phi^{(3)} \succ \phi^{(2)} \succ \phi^{(1)}$.

Overall our results suggested that the performance of the CFA model was sensitive to perturbed data sets. This effect was stronger in the third criterion as it showed a clear replacement effect. In general, in SGR we recommend to choose more conservative criteria in order to better evaluate the effect of eventual fake data.

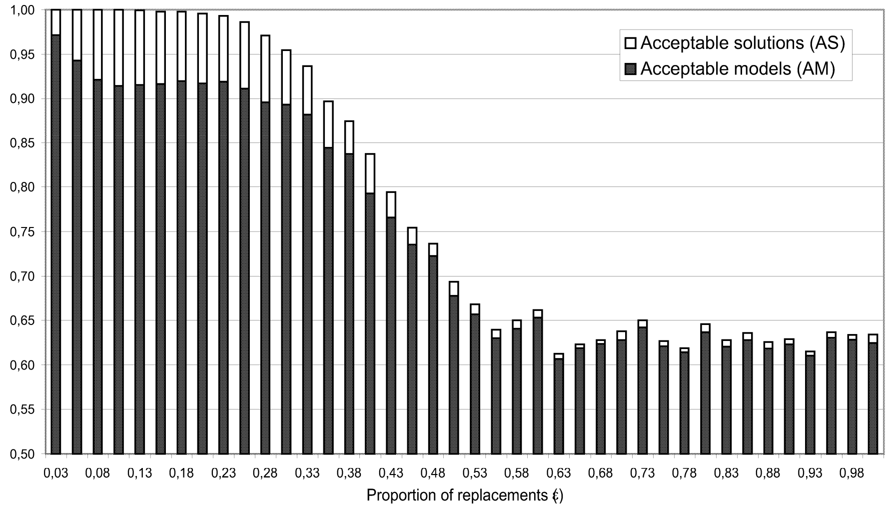


*Figure 4.4.* Parameter $\pi_{\varrho}^{*(2)}$ as a function of percentage of replacements - criterion $\phi^{(2)}$. The percentage of acceptable solutions is the upper bound for $\pi_{\varrho}^{*(2)}$ (superimposed bar-charts).

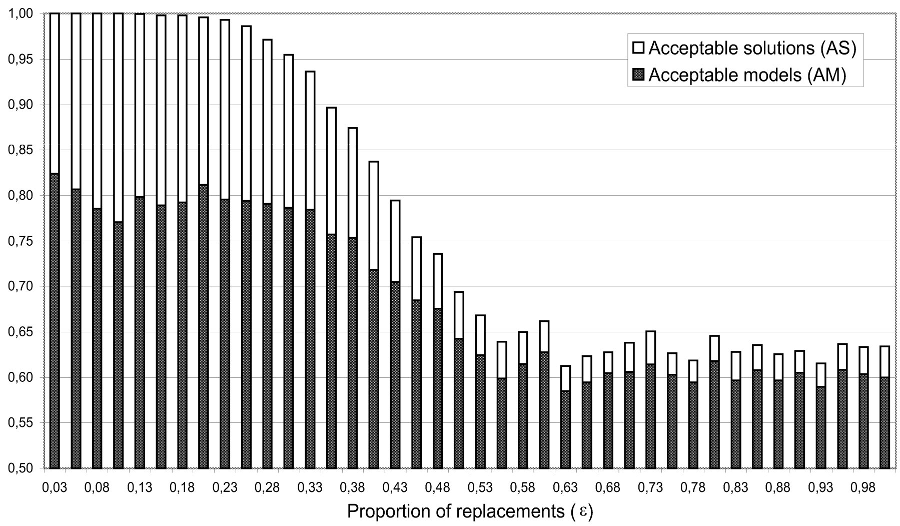


*Figure 4.5.* Parameter $\pi_{\varrho}^{*(3)}$ as a function of percentage of replacements - criterion $\phi^{(3)}$. The percentage of acceptable solutions is the upper bound for $\pi_{\varrho}^{*(3)}$ (superimposed bar-charts).

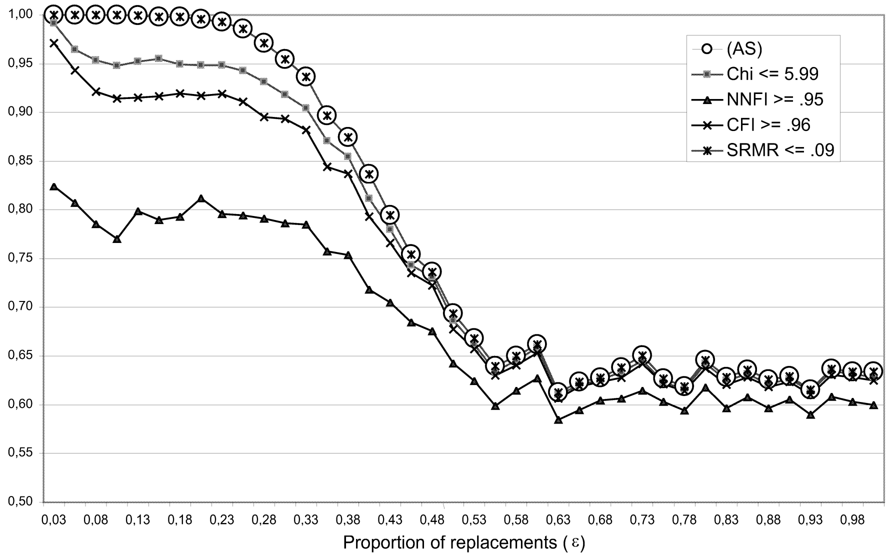


*Figure 4.6.* Model acceptability patterns as a function of percentage of replacements.

*Table 4.2.* ($\alpha = .05$)-confidence intervals for $\pi_{\varrho}^{*(i)}$ as a function of percentage of replacements ($100\epsilon$).

| | Acceptability criteria | | |
|---|---|---|---|
| $100\epsilon$ | $\pi_{\varrho}^{*(1)}$ | $\pi_{\varrho}^{*(2)}$ | $\pi_{\varrho}^{*(3)}$ |
| 10% | [0.90 - 1.0] | [0.87 - 0.97] | [0.72 - 0.82] |
| 20% | [0.90 - 1.0] | [0.87 - 0.97] | [0.77 - 0.87] |
| 30% | [0.87 - 0.97] | [0.83 - 0.93] | [0.72 - 0.82] |
| 40% | [0.77 - 0.87] | [0.73 - 0.83] | [0.67 - 0.77] |
| 50% | [0.63 - 0.73] | [0.63 - 0.73] | [0.59 - 0.69] |
| 60% | [0.62 - 0.72] | [0.60 - 0.70] | [0.58 - 0.68] |
| 70% | [0.60 - 0.70] | [0.58 - 0.68] | [0.55 - 0.65] |
| 80% | [0.61 - 0.71] | [0.59 - 0.69] | [0.57 - 0.67] |
| 90% | [0.60 - 0.70] | [0.58 - 0.68] | [0.55 - 0.65] |
| 100% | [0.60 - 0.70] | [0.58 - 0.68] | [0.55 - 0.65] |

### 7.4 Discussion and Possible Extensions

In this paper we have focused on the problem of evaluating acceptance criteria when data with possible fake entries are analyzed. The SGR approach is essentially based on a data generation - data analysis double procedure, which is carried out by using a given target model. The model is supposed to be correct, that is to say to adequately simulate the underlying process. However, model mispecifications (e.g. error in model parameters, structures, assumptions and specifications) could be another important source of error. Therefore an evaluation of model adequacy is needed. This could be done by using external information about the model (e.g. theoretical constructs) and/or by considering former evaluations of the model (e.g. results replicated with different empirical data).

The reader may have already noticed some similarities between the approach proposed here and standard Monte Carlo experiments in structural equation modeling (Bentler, 1990; Curran et al., 1996; Hu & Bentler, 1998, 1999; la Du & Tanaka, 1989; Mulaik et al., 1989). For example, the idea of generating new data sets. However, the two approaches are substantially different. Usually a Monte Carlo experiment uses a hypothesized model to generate new data under various conditions. Therefore the simulated data are used to evaluate some characteristics of the model. This, of course, implies that the distribution of the random component in the assumed model must be known, and it must be possible to generate pseudorandom samples from that distribution under the desired conditions planned by the researcher.

Instead of using the hypothesized model structure to generate simulated data sets, our approach uses the original data sample in order to generate a new family of data sets. In particular, these new data sets are obtained by adding structured perturbations in the original data set. In the latter case, each new sample represents an alternative scenario which is directly derived from the original sample. Next, the result of a target criterion can be compared with the ones obtained from the perturbed samples. In this case, of course, the distributional properties of the statistics are not those that hold under a particular model hypothesis (like for Monte Carlo simulation studies); rather they are the properties under a model whose parameters corresponds to values fitted from a structured collection of perturbed samples that are generated from a given real data set.

Several possible extensions of the elementary $L_0$-SGR model may be considered. In the present paper, under the assumption of the principle of indifference, a very simple SGR model has been proposed as a model for Likert-type data. However, the current approach can be straightforwardly extended to categorical data as well as to continuous data. In particular, a SGR model for continuous data would imply a different kind of metric, for example either the city-block distance $(L_1)$ or the standard Euclidean distance $(L_2)$. These new extensions would enlarge the general replacement schema by adding more complex constraints with which we could provide more structured perturbed scenarios.

All in all, although our approach is tentative and more work is needed to better understand its full potentialities, the overall lack of modeling data-error effects in structural models suggests that SGR can be a promising new method to analyze acceptability criteria under empirical data set perturbed with error.

## References

Bacchus, F., Grove, A., Halpern, J.Y., & Koller, D. (1994). Generating new beliefs form old. *Uncertainty in Artificial Intelligence (UAI-94)*, 37-45.

Bentler, P.M. (1990). Comparative fit indexes in structural model. *Psychological Bulletin*, *107*, 238-246.

Bentler, P.M., & Bonnett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Chaturvedi, A., Green, P. E., & Caroll, J. D. (2001). K-modes clustering. *Journal of Classification*, *18*, 35-55.

Curran, P.J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16-29.

Hu, L., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424-453.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit Indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

la Du, T.J., Tanaka, & J.S. (1989). Influence of sample size, estimation method, and model specification on goodness-of-fit assessments in structural equation models. *Journal of Applied Psychology.*, *74*, 625-635.

Mulaik, S.A., James, L.R., Van Alstine, J., Bennet, N., Lind, S., & Stilwell, C.D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin*, *105*, 430-445.

Thompson, S.K. (2002). *Sampling* (2nd ed.). New York: Wiley.

Vidotto, G., & Marchesini, C. (2000). *La realizzazione professionale. Analisi delle risorse personali e dei processi decisionali per l'orientamento scolastico-professionale*. Milano: Franco Angeli.