

Intervention and Statistical Underdetermination in Factor Analysis

Abstract

We consider the use of intervention data for resolving underdetermination problems in statistical modelling. The leading example is factor analysis, a major statistical tool in the social sciences. We first relate indeterminacy in factor analysis to the problem of underdetermination. Then we draw a parallel between factor analysis models and Bayesian networks with hidden nodes, which allows us to clarify the use of intervention data for dealing with indeterminacy. We show that in some cases, the indeterminacy can be resolved by an intervention. The upshot is that intervention data can replace the theoretical criteria that are typically employed to resolve underdetermination.

1 Introduction

The problem of underdetermination occurs when rival theories fit the empirical facts equally well, so that we cannot choose among the theories on the basis of empirical fact alone (Douven, 2008). One kind of underdetermination is widespread in the social sciences, in particular where these sciences employ statistical modelling. Within a given experimental setup or population study, it may so happen that the statistical model includes hypotheses that have the same likelihood function. The problem of underdetermination then appears as the problem that the best fitting hypothesis may have a number of equally well fitting rivals. One standard response to this is to look for theoretical criteria, such as simplicity or explanatory force, that force a choice between the rivals.¹ Underdetermination is then resolved by an appeal to theoretical considerations.

¹In factor analysis, in particular, researchers use theoretical criteria pertaining to the variation among the estimations of the statistical parameters, such as “varimax”. See, e.g., Lawley and Maxwell (1971).

In this paper we investigate a different response to such underdetermination in statistical modelling, which makes use of intervention data. The notion of underdetermination is thereby made relative to what is taken as observable in a given experimental setup. Relative to a setup and background theory, several hypotheses under consideration may say the very same things about what observations to expect, i.e., they have exactly the same likelihood functions and thus perform equally well on the observation data. The model consisting of these hypotheses is called ‘unidentifiable’. However, the hypotheses need not be altogether equivalent. We can consider specific changes to the experimental setup, or interventions for short, such that the background theory determines different likelihood functions over the additional results. So, relative to another experimental setup, the hypotheses can be told apart.

Our primary objectives are philosophical. First, we illustrate that a resolution of underdetermination can partly be driven by empirical fact, and so need not only be driven by theoretical considerations. Second, with our discussion we bring to the fore an important and undervalued aspect of scientific confirmation, namely the use of intervention data following experimentation. We believe that insights from the philosophy of experiment (e.g. Hacking (1980); Gooding (1990)) can come to fruition in confirmation theory and we hope to make a modest start with that here. A further objective is methodological: we hope to stimulate the uptake of statistical tools for modelling interventions in social science. Despite the availability of statistical theories and methodological tools for exploiting intervention data (e.g. Spirtes *et al.* (1993); Eberhardt *et al* (2010); Hyttinen *et al* (2012)), scientists are often not aware of the potential of intervention data over and above the use of observational data. It is hoped that this paper will contribute to a better understanding of the benefits of interventions, and hence will stimulate the use of the available statistical tools.

The setting for illustrating these ideas is exploratory factor analysis. As it happens, factor analysis has already made an appearance in the philosophy of science in another context. In Haig (2005) and Schurz (2008), for example, factor analysis is proposed as a model for abductive inference, and thus as a tool for generating new theory. We emphasize that in this paper we take a different perspective. We employ exploratory factor analysis as an illustration of the more general problem concerning statistical underde-

termination, and we focus on the role of interventions in resolving statistical underdetermination. This is of course different from taking the technique of exploratory factor analysis itself as a model for abductive inference, and hence as a tool for theory choice, as Haig and Schurz do.

The paper is set up in the following way. In §2 we describe two distinct problems of indeterminacy in factor analysis. We show in §4 that factor analysis is essentially identical to estimating parameters in a Bayesian network with hidden nodes. Like causal Bayesian networks, the models in factor analysis therefore allow for incorporating intervention data. We argue that in specific cases, intervention data can be used to resolve the underdetermination problem. In §6, finally, we briefly suggest how the model for intervention may prove useful to the philosophy of experiment, and more generally, to scientific methodology.

2 Underdetermination in statistics

In what follows we characterize the problem of statistical underdetermination and make it precise for factor analysis, a well-known statistical technique in psychometrics. Factor analysis is routinely used to interpret psychological test data, and it is a live problem to working psychologists that the data do not allow for a complete determination of the underlying factors. Importantly, the specific focus of this paper is not on the underdetermination of causal structure by data, as is usual in the context of studying interventions, but rather on the underdetermination of parameter values.

2.1 Underdetermination in statistics

Consider a simple statistical problem, in which we estimate the chances of events in independent and identical trials, e.g., results in psychological tests. An observation at time t is denoted by the variable Q_t , with possible values q_t^0 or q_t^1 , for failing and passing the test. We denote a series of t observations or test results by the variable S_t , and the event that earlier results were some ordered series $\langle q_1^0 q_2^1 q_3^0 \dots q_t^1 \rangle$ by $s^{(010\dots 1)}$, or s_t for short. Denoting the hypothesis that the chance of observing q_t^1 is θ with h_θ , we have

$$P(q_{t+1}^1 | h_\theta \cap s_t) = \theta \tag{1}$$

for each trial $t + 1$, an expression² often called the likelihood function of h_θ .

The chance θ of the event q_{t+1}^1 may be any value in $[0, 1]$, so we have a whole continuum of hypotheses h_θ gathered in what we call a statistical model, denoted \mathcal{H} . On the basis of some series of events S_t , we can provide an estimation of θ . We can do so either by defining a prior $P(h_\theta)$ and then computing a posterior by Bayesian conditioning, or by defining an estimator function over the event space, typically the so-called observed relative frequency

$$\hat{\theta}(S_t) = \sum_{i=1}^t I^1(Q_i)/t,$$

in which the indicator $I^1(Q_i) = 1$ if Q_i takes the value q_i^1 and 0 otherwise.

The above estimation problem is completely unproblematic. The observations have a different bearing on each of the hypotheses in the model, i.e. each member of the set of hypotheses. If there is indeed a true hypothesis in the set, then according to well-known convergence theorems (cf. Earman (1992), pp. 141–149), the probability of assigning a probability 1 to this hypothesis will tend to one. In the limit, we can therefore almost always, in the technical sense of this expression, tell the statistical hypotheses apart.³

This situation is different if we take a slightly different set of statistical hypotheses G_ξ , characterized as follows:

$$P(q_{t+1}^1 | g_\xi \cap s_t) = \xi^2, \quad \xi \in [-1, 1].$$

This set of hypotheses is essentially the same as before, only it is labeled in a peculiar way. The hypotheses g_ξ and $g_{-\xi}$ are indistinguishable, because they both assign exactly the same probability to all the observations: $P(q_{t+1}^1 | g_\xi \cap s_t) = P(q_{t+1}^1 | g_{-\xi} \cap s_t)$. In such a case, we speak of an unidentifiable model.

Unidentifiable models are statistically underdetermined by the observations. Importantly, statistical underdetermination is not definitive: it is not ruled out that there are experiments or additional observations that allow us to disentangle the statistical hypotheses. This paper shows how additional experiments can achieve this.

²We are writing the probability of data according to a particular hypothesis as $P(\cdot | h_\theta)$, and not as $p_{h_\theta}(\cdot)$ or $p_\theta(\cdot)$, thereby following the Bayesian idea that hypotheses h_θ can serve as arguments of the probability function.

³Any infinitely long series of results is in principle consistent with any of the hypotheses H_θ , and in that sense we are encountering an underdetermination problem in the estimation. However, here we will not consider this type of underdetermination.

2.2 Factor analysis

The above example of statistical underdetermination is rather contrived. No reason is given for distinguishing between the regions $\xi > 0$ and $\xi < 0$. However, there are cases in which it makes perfect sense to introduce distinctions between hypotheses that do not differ in their likelihood functions. This subsection is devoted to presenting one of these cases, involving so-called exploratory factor analysis. The exposition is partly borrowed from [omitted for purpose of blind review].

Exploratory factor analysis posits a statistical model of hidden, or latent, random variables on the basis of an analysis of the correlational structure of observed, or manifest, random variables.⁴ Say that in some experiment we observe the levels of fear F and loathing L in a number of individuals indexed i , and we find a positive correlation between these two variables,

$$P(F_i, L_i) > P(F_i)P(L_i).$$

One way of accounting for the correlation is by positing a statistical model over the variables in which fear and loathing may be related directly.

We may feel that it is neither the loathing that instills fear in people, nor the fear that invites loathing. Instead we might think that both feelings are correlated because of the presence or absence of a happiness drug, denoted E . We might for example posit a negative correlation between the drug and the fear, and similarly a negative correlation between the drug and the loathing. Conditional on a certain drug dosage, fear and loathing can be taken as uncorrelated:

$$P(E_i, F_i, L_i) = P(E_i)P(F_i|E_i)P(L_i|E_i).$$

We can then say that the drug dosage is the so-called common factor to the observable, or manifest, variables of fear and loathing. The correlations between drug usage and fear and loathing respectively we call the factor loadings.

⁴See Lawley and Maxwell (1971) for a classical statistical overview, Mulaik (1985) for a philosophically-minded discussion, and Bartholomew and Knott (1999) for a very insightful introduction from a Bayesian perspective. All these treatises introduce exploratory factor analysis next to the much less problematic statistical tool of confirmatory factor analysis. In most of the following we concentrate on the former, and simply call it factor analysis.

Factor analysis has a number of standard applications, which are usually subdivided according to whether the manifest and latent variables are categorical or continuous. In this paper we discuss one of the most straightforward applications of factor analysis, in which both the manifest and latent variables are binary. Our reason is that we are making a conceptual point about interventions and underdetermination. For this purpose the simplest format of factor analysis suffices.

In terms of the example, the drug is either present in subject i , e_i^1 , or absent, e_i^0 , and similarly for fear and loathing. We assume that the probabilistic relations between the variables are independent and identically distributed. Out of the many possible probabilistic dependencies between F_i , L_i and E_i , we thus confine ourselves to

$$P(f_i^1 | e_i^j) = \phi_j, \quad (2)$$

$$P(l_i^1 | e_i^j) = \lambda_j, \quad (3)$$

for $j = 0, 1$, a conditional version of the Bernoulli model of Equation (1). Similarly for the variables E_i ,

$$P(e_i^1) = \epsilon \quad (4)$$

The probability over the variables E_i , L_i and F_i is thus given by five Bernoulli distributions, each characterized independently by a single chance parameter.

In an experimental setting, we can often observe the common factor. For example, we can check whether the drug was taken or not. But in situations in which the causal or mechanistic story behind the correlations is unknown, we may nevertheless want to posit such an underlying story. For example, recurring feelings of fear and loathing may be two of a large number of emotions used to describe individuals in a general population whose constitution is otherwise unknown. It may be that we can account for the correlations in a statistical model by positing the presence or absence of a mental condition, depression D . Exploratory factor analysis is a technique for arriving at such common factors in a systematic way. When given a set of correlations among manifest variables, it produces a statistical model of latent common factors that accounts for exactly these correlations.⁵

⁵Seeing that exploratory factor analysis generates a structure that explains the observed correlations, it is rather natural that Haig (2005) and Schurz (2008) present it as a formal model of abduction.

Unsurprisingly, applications of factor analysis suffer from problems of underdetermination. After all, factor analysis posits the theoretical structure of unobservable common causes, over and above the observed correlations between observable variables. There will generally be many latent common factor models and many different causal structures that fit the data. But even if all modelling choices have been made, statistical underdetermination may appear. In what follows we focus specifically on this restricted version of statistical underdetermination.

2.3 Underdetermination in factor analysis

Consider the factor model of Equations (5), but replace the drug variable E with the depression variable D :

$$\begin{aligned} P(f_i^1 | d_i^j) &= \phi_j, \\ P(l_i^1 | d_i^j) &= \lambda_j, \\ P(d_i^1) &= \delta \end{aligned} \tag{5}$$

Now focus on the dimensions of this model. We count 5 parameters, namely δ , and ϕ_j and λ_j for $j = 0, 1$. On the other hand, we have the binary observations F_i and L_i that can be used to determine these parameters. But because we are using Bernoulli hypotheses, only the observed relative frequencies of the possible combinations of F_i and L_i matter. And because we have 4 possible combinations of F_i and L_i , whose relative frequencies must add up to 1, we have effectively 3 frequencies to determine the 5 parameters in the model. After having used the observations in the determination of the parameters, therefore, we still have 2 degrees of freedom left. Hence the values of the parameters in the model cannot be determined by the observations uniquely.

We can state this problem in more detail by looking at the likelihoods for the observations of possible combinations of F_i and L_i . We write $\theta = \langle \delta, \phi_0, \phi_1, \lambda_0, \lambda_1 \rangle$. Further, the observations of individuals i are $f_i^j \wedge l_i^k$, which may be summarized as q_i^u with $u = 2j + k$. The sequences s_t are again observations of individuals $s^{u_1 u_2 \dots u_t}$. Finally, we abbreviate:

$$\begin{aligned} \eta_{01} &\equiv P(f_i^0 \wedge l_i^1 | h_\theta) = \delta(1 - \phi_1)\lambda_1 + (1 - \delta)(1 - \phi_0)\lambda_0, \\ \eta_{10} &\equiv P(f_i^1 \wedge l_i^0 | h_\theta) = \delta\phi_1(1 - \lambda_1) + (1 - \delta)\phi_0(1 - \lambda_0), \\ \eta_{11} &\equiv P(f_i^1 \wedge l_i^1 | h_\theta) = \delta\phi_1\lambda_1 + (1 - \delta)\phi_0\lambda_0, \end{aligned} \tag{6}$$

The fourth likelihood, $P(f_i^0 \wedge l_i^0 | h_\theta)$, can be derived from these expressions. The salient point is that the system of equations resulting from filling in particular values for the likelihoods η_{jk} has infinitely many solutions in terms of the components of θ : for any value of the likelihoods η_{jk} , the space of solutions in θ has 2 dimensions. Hence different hypotheses h_θ will have the same set of likelihoods η_{jk} for the observations. The statistical model is unidentifiable.

Let us briefly elaborate on the unidentifiability of the model. It means that the likelihood function over the model does not have a unique maximum, and so that the maximum-likelihood estimator does not point to a uniquely best hypothesis.⁶ Say that we observe the relative frequencies r_{jk}/t with

$$r_{jk} = \sum_{i=1}^t I^j(F_i, s_t) I^k(L_i, s_t), \quad (7)$$

the number of occurrences of $f_i^j \wedge l_i^k$ in s_t , and with the indicators $I^j(F_i, s_t) = 1$ if $s_t \subset f_i^j$ and 0 otherwise, and $I^k(L_i)$ analogously. By the likelihoods of Equations (6) we can then construct a likelihood function for s_t :

$$P(s_t | h_\theta) = \prod_{jk} \eta_{jk}^{r_{jk}}. \quad (8)$$

The likelihood $P(s_t | h_\theta)$ is maximal at the observed relative frequency, $\eta_{jk} = r_{jk}/t$. But as said, there are infinitely many hypotheses h_θ that have these particular values for the likelihoods. Consequently, there is no unique hypothesis h_θ that has maximal overall likelihood $P(s_t | h_\theta)$.

For future reference we note that, by means of the likelihoods given in Equations (8), we can determine a posterior probability for the hypotheses in the model, $P(h_\theta | s_t)$. And from the posterior distribution over the hypotheses we can generate the expectation value of the parameter θ of the model \mathcal{H} , according to

$$E[\theta] = \int_{\mathcal{H}} \theta P(h_\theta | s_t) d\theta. \quad (9)$$

Here θ runs over $[0, 1]^5$ because the model concerns five independent chances. Like the posterior, the estimations will suffer from the fact that the hypotheses cannot be told apart: they will depend on the prior probability over the

⁶Whether we approach the statistical problem in a classical or a Bayesian fashion, the likelihood function will occupy a central place. The exposition will focus on properties of this function.

hypotheses. Of course, this is usually the case in a Bayesian analysis. What is troublesome is that no amount of additional data can eliminate this dependence of the estimations on the prior.

One reaction is to downplay the underdetermination problem and say that it only concerns the values of these abstract parameters and not the empirical consequences. But because the estimations and expectations are not fully determined, the causal, nomic and conceptual structure of the factors underlying the observed variables is not determined either. Different values for the parameters ϕ_j and λ_j entail different systematic relations between depression, fear and loathing, and ultimately this reflects back on our understanding of the posited notion of depression itself. In the statistical underdetermination exemplified here, we find back the well-known underdetermination of theory by data: we cannot pin down the theoretical superstructure on the basis of data alone.

3 Underdetermination in multivariate linear regression

We are well aware that the statistical model considered in the foregoing is much simpler than what is typical in factor analysis. In this section we argue that the problem outlined above also shows up in more realistic uses of factor analysis. Furthermore, we will reveal that there are actually two problems of statistical underdetermination in factor analysis. The first one, illustrated in §2.3, is made more concrete in the first subsection. The second type is briefly mentioned in the second, mostly because it has been hotly debated in psychological methodology, but also because the present paper can offer a specific angle on it.

3.1 The rotation problem

In many actual applications of factor analysis, the variables are not binary but continuous, the probabilistic relations between the variables are linear regressions with normal errors, and the latent variable is assumed to be governed by some continuous distribution as well. In our example we may write $F_i = f$ for the event that the level of fear is $f \in \mathbb{R}$, and similarly for

depression $D_i = d$. Then the relation between F_i and D_i , for example, is

$$P(F_i = f | D_i = d) = N(\lambda_F d, \sigma_F) \quad (10)$$

in which $N(\lambda x, \sigma)$ is a normal distribution over the values f of F_i . So the relation between the variables D_i and F_i is characterized by a richer family of distributions, parameterized by a regression parameter λ_F and an error of size σ_F .

Despite these differences, the same kind of underdetermination also occurs in the more complicated statistical models. But in such models it takes a slightly different shape. Note first that we can extend factor models like the one above to include any number of common factors. However, once a model includes more than one common factor, we find that the factor loadings are not completely determined. Say, for example, that we analyze fear F , loathing L , and sleeplessness S in terms of two common factors, one of them depression D , but next to that the latent variable C . Every individual is supposed to occupy a specific position in the $C \times D$ surface. We might feel that a more natural way of understanding the surface of latent variables is by labeling the states in this surface differently, for example by introducing variables A and B , both of which are linear combinations of C and D . The factors in a model may be linearly combined or, in more spatial terms, rotated to form any new pair of factors.⁷

The underdetermination problem with this is that, if we allow the latent factors to be correlated, any rotation of factors, e.g., from $\{C, D\}$ to some $\{A, B\}$, will perform equally well on the estimation criterion, be it maximum likelihood, generalized least squares, or similar. This problem is appropriately known as the problem of the rotation of factor scores. Neither the estimation criteria, often maximum likelihood, nor Bayesian methods of incorporating the data lead to a single best hypothesis in the factor model. The result is rather a collection of such models, meaning that the factor model is again unidentifiable, with all the attached problems listed above.

A standard reaction to the rotation problem is to adopt the theoretical criterion that the latent variables must be independent. In that case, we cannot freely rotate the axes in the space of latent variables anymore, because the parameterisation of the space must be such that there are no

⁷Basically this is a coordinate transformation in the space of latent variables, characterizing it in terms of different bases.

correlations between the latent variables. But there are alternative theoretical criteria for choosing the parameterisation of the space of latent variables. For example, it may be interesting to have maximal variation among the regression coefficients which, intuitively, comes down to coupling each latent variable with a distinct subset of manifest variables. The thing to note is that, from the point of view of statistics, the choice for how to parameterize the space of latent variables is underdetermined: we cannot decide between these parameterizations on the basis of the observations alone.

In this paper we will not elaborate the mathematical details of underdetermination in these more complicated models. For present purposes, it suffices to use the simpler factor model of Equations (2) to (4). The crucial characteristic in all of what follows is that there are latent variables explaining the correlational structure among the manifest variables, and that these structures are not fully determined by the correlations among the manifest variables.

3.2 Factor score indeterminacy

There is another problem with factor analysis that can be framed as underdetermination, and which has received considerable attention within statistical psychology.⁸

Say that we have rotated the factors to meet the theoretical criterion of our choice, for instance by simply assuming a single common factor or by fixing the independence of the latent factors. Can we then reconstruct the latent variable itself, that is, can we provide a labeling in which each individual, i.e. each valuation of the observable variables, is assigned a determinate expected latent score? Sadly, the classical statistical answer here is negative. We still have to deal with the so-called indeterminacy of factor scores, meaning that there is a variety of ways in which we can organize the allocation of the individuals on the latent scores, all of them perfectly consistent with the estimations.⁹

⁸See Steiger (1979) for some historical context, Maraun (1996) for a philosophical evaluation, McDonald (1974) for an excellent classical statistical discussion, and Bartholomew and Knott (1999) for a Bayesian account of it.

⁹There are some restrictions to this allocation. For example, as worked out in Ellis and Juncker (1997), if we let the number of manifest variables increase and assume that the latent variable is tail-measurable in terms of these manifest variables, then the factor scores are determined up to a functional transformation.

The type of underdetermination presented by factor score indeterminacy depends on what we take to be the statistical inference underlying factor analysis. In the context of this paper, we take the factor analysis to specify a complete probability assignment over the latent and manifest variables, including a prior probability over the latent variables. As explained in Bartholomew and Knott (1999), factor score indeterminacy is thereby eliminated, as long as there are sufficiently many manifest variables that are related to the latent variables according to distributions of a suitable, namely exponential, form. In this paper we will therefore ignore most of the discussion on factor score indeterminacy.

There is one point at which the problem of factor score indeterminacy enters the present discussion. We will show in the following that intervention data can also be used to choose among a class of priors. But as indicated, the problem of choosing a prior probability is related to the problem of factor score indeterminacy. Therefore the use of intervention data, which resolves the problem of underdetermination discussed above, provides a new perspective on the problem of the indeterminacy of factor scores as well. We will return to this idea in §5.2.

4 Interventions to resolve underdetermination

In the foregoing we have shown that factor analysis suffers from statistical underdetermination. We now explain the underdetermination inherent to factor analysis by identifying analogous problems in the estimation of parameters in Bayesian networks. This leads us to consider a specific solution to the underdetermination problem, namely by means of intervention data. We first introduce Bayesian networks in §4.1, then the notion of intervention in §4.2, and finally its use in resolving underdetermination in §4.3.

4.1 Bayesian networks and factor analysis

In general, a Bayesian network consists of a directed acyclic graph on a finite set of variables $\{F, L, D, E \dots\}$ together with the probability distributions like of each variable conditional on its so-called parents in the graph, for example $P(E \mid Par_E)$. The graph is related to the probability distribution over the variables by an assumption known as the Markov Condition: each variable is probabilistically independent of its non-descendants in the graph,

conditional on its parents, e.g., $E \perp\!\!\!\perp NonDesc_E \mid Par_E$; see Pearl (2000). Under this assumption the network suffices to determine the joint probability distribution over the variables, via the identity

$$P(F, L, \dots) = P(F \mid Par_F) \times P(L \mid Par_L) \times \dots \quad (11)$$

The probability of any valuation on the left hand side of this equation can be computed by filling in these valuations on the right hand side.

It is well-known that Bayesian networks, structural equations modeling, and factor analysis are closely related. Effectively, the introduction of factor analysis for the binary variables $\{F, L, D\}$ was already an introduction to a specific class of Bayesian networks. First, we assume that there are no inter-subject dependencies and that the same probability assignment describes all subjects,

$$P(F_i, L_i, D_i) = P(F_{i'}, L_{i'}, D_{i'}), \quad (12)$$

so that we can omit the subscripts i . For each subject the factor analysis determines a probability function $P(F, L, D)$ that observes a specific symmetry: conditional on the latent depression D there is no correlation between the manifest fear F and loathing L ,

$$P(F, L, D) = P(D)P(F|D)P(L|D). \quad (13)$$

On this basis we can build a network, with the variables F , L and D as nodes. Quite apart from the exact probability values, the probability function determined by factor analysis can thus be represented in a Bayesian network whose graph is depicted in figure 1.

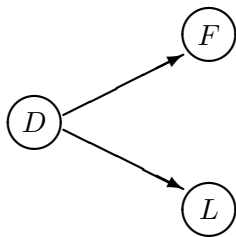


Figure 1: The graphical structure representing the independence relations in a factor analysis of depression, fear and loathing.

There are also differences between the theory of Bayesian networks and factor analysis. For one, factor analysis entails a rather specific network

structure: there are latent parent nodes, observable child nodes, there are typically fewer parents than children, and any child can be connected to any parent. On the other hand, applications of the former are usually restricted to probability functions over finite or at least countable domains. Nodes with continuous domains are not that commonly discussed, although they have been studied in the context of structural equations models, for example in Pearl (2000) and, from the side of latent variable modeling, in von Eye and Clogg (1994). A related difference is that in most applications of factor analysis the probability functions that are considered are restricted to normal distributions over latent nodes, and to linear regressions with normal errors between latent and observable nodes. Applications of Bayesian networks are typically, but not necessarily, restricted to Bernoulli distributions.

In this paper we approach factor analysis more from the angle of Bayesian networks, using the framework for inference over Bayesian networks presented in [omitted for blind reviewing]. Hence the statistical underdetermination presented in §2.3 is framed as a problem to do with determining the posterior probability distribution over the parameters that characterize the Bayesian network of Figure 1. As announced, we are going to resolve this statistical underdetermination by means of intervention data. To this aim we first introduce interventions in the context of Bayesian networks.

4.2 Interventions

A causally interpreted Bayesian network, or causal net for short, is a Bayesian network where the graph is interpreted as a causal graph. That is, each arrow in the graph is interpreted as denoting a direct causal relationship from the parent variable to the child variable. Under this interpretation, the Markov Condition is called the Causal Markov Condition. It says that each variable is probabilistically independent of its non-effects conditional on its direct causes. It is often assumed that the Causal Markov Condition is bound to hold if the graph in the net is correct and is closed under common causes (i.e., any common causes of variables in the net are also included in the net). While there are situations in which the condition is implausible, it can be justified as a default assumption [omitted for blind reviewing], and we shall take it for granted here.

Causal nets are helpful for predicting the effects of interventions. When an experimenter intervenes to fix the value of a variable, she interrupts the

normal course of affairs and sets the variable exogenously. The usual mechanisms, according to which the variable is determined, are thereby replaced with new mechanisms, according to which the variable is determined only by the experimenter. An ‘ideal’ or ‘divine’ intervention is one in which the intervention only changes the intended variable, without changing other variables under consideration and without changing other causal relationships under consideration. By means of Equation (11) we can determine the probability P' that some variable F takes value f^1 after an ideal intervention has been performed that sets E to e^1 . Note that the causal net determines two different probability distributions, P before and P' after intervention. While P and P' will coincide on the non-descendants of E , their probabilities for the variables downstream from E will be different.

Causal nets can also handle a more general notion of intervention that is central to our concerns. We might set the probability to a new value $P'(e^1) = \epsilon' \neq P(e^1)$ while leaving the rest of the network intact. In other words, we transform the causal net by eliminating arrows into E , set its unconditional distribution to $P'(e^1) = \epsilon'$, and then determine the new probabilities for other variables. This kind of intervention is sometimes called an ‘imperfect’ or ‘stochastic’ intervention, to distinguish it from the divine interventions considered above. A stochastic intervention is itself a special case of another kind of intervention—called a parametric intervention—where, instead of intervening to fix the effect variable, one intervenes to change how the causes impact on the effect variable. See Korb *et al* (2004) and Eberhardt and Scheines (2007) for discussion of these kinds of intervention.

Interventions can help with underdetermination in two ways. First, they can help with underdetermination of causal structure, as described extensively in Spirtes *et al.* (1993). If more than one causal structure is compatible with evidence, one can intervene, collect more evidence, and use this new evidence to decide between the causal structures. To take the example presented in the foregoing, suppose variables F , L and D are all measured, and that the resulting data shows that F and L are probabilistically independent conditional on D , written $F \perp\!\!\!\perp L \mid D$. This evidence is compatible with the causal graph of Figure 1, but equally with Figures 2 and 3. The evidence can be used to fill in the conditional probability distributions on these causal models, but cannot decide between them. An intervention can decide between them, however. If, after intervening to change the distribution of D ,

the distribution of F and L are changed, then that favours Figure 1. Otherwise if only the distribution of L is changed after intervention, then Figure 2 is supported, and if only the distribution of F is changed then Figure 3 is supported.



Figure 2: A chain of fear F causing depression D , which causes loathing L .



Figure 3: A chain of loathing L causing depression D , which causes fear F .

More important for our concerns is that interventions can be used to resolve the statistical underdetermination of the parameters. Suppose that the causal structure is known and that data is collected which helps to estimate the probability distributions of some variables conditional on their parents, but which does not determine conditional distributions that attach to other variables. By carrying out an ideal intervention, an experimenter effectively changes the conditional distribution of one variable without changing the distributions of other variables. The data obtained after the intervention can then be used in conjunction with the old data to further constrain the values of the underdetermined distributions.

4.3 Interventions and underdetermination

In this section we show how interventions can be used to resolve the statistical underdetermination introduced in §2.3. We consider the example of depression, fear, and loathing. In the next section we sketch how the idea can be extended to factor analysis more generally.

Let us briefly explain the general idea of using interventions to resolve underdetermination. We need to assume that the factor model is more than a convenient way of representing the probability functions involved. The arrows in the factor model need to be interpreted causally, that is, the common factors must be taken as the causes of the observed variables. With this causal assumption in place, an intervention on the subjects will indeed

change the distribution over the latent variables of the subjects, and not the probabilistic relations between the latent and the manifest variables. After the intervention we obtain an entirely new estimation problem for the parameters in the Bayesian network. However, because the data are obtained by intervention, we can assume that the parameters associated with the relations between latent and manifest variables do not change. To accommodate the intervention data, we therefore have a smaller space of parameters available. In the following we show that, depending on the model, intervention data can thus be used to select a unique best estimate for the parameter values in the factor model.

Consider again the model characterized by Equations (5), (12) and (13). As explained in the foregoing, an intervention is an external shift to the probability assignment. In this particular case, we intervene on the node D , giving all the subjects a treatment intended to change the probability for depression. In terms of the foregoing, we change the probability of depression, $P(d^1) = \delta$, to a new value,

$$P'(d_i^1) = \delta',$$

which is supposed to be lower than δ . The relations of the depression variable to the variables of fear and loathing, given by $P'(f_i^1|d_i^1) = \phi_j$ and $P'(l_i^1|d_i^1) = \lambda_j$, are not changed by the intervention: the treatment is supposed to change the probability for depression but not how depression, whether absent or present, affects feelings of fear and loathing. Finally, after the intervention we record the observations s'_t in the same set of t individuals. In analogy to Equation (7), we observe the numbers of the occurrences in the new sequence of observations s'_t ,

$$r'_{jk} = \sum_{i=1}^t I^j(F_i, s'_t) I^k(L_i, s'_t).$$

So r'_{jk}/t are the relative frequencies of the variables F and L as observed after the intervention.

To get the point of this across quickly, we focus again on the dimensions of the model. This time we count a number of 6 parameters, namely δ , ϕ_j and λ_j for $j = 0, 1$, and finally δ' . On the other hand, we have a richer set of observations that can be used to determine these parameters. Specifically, we have 3 observed relative frequencies of $f_i^j \wedge l_i^k$ before intervention, and 3 of them after intervention, so six in total. Whereas previously we had two

degrees of freedom left after the incorporation of the data, we can now fill in all the parameter values of the factor model.

Let us make this more precise. As before, we have the likelihoods of Equations (6). But to these expressions we now add the likelihoods of the hypotheses after the intervention:

$$\begin{aligned} P'(f_i^0 \wedge l_i^1 | h_\theta) &= \delta'(1 - \phi_1)\lambda_1 + (1 - \delta')(1 - \phi_0)\lambda_0 \equiv \eta'_{01}, \\ P'(f_i^1 \wedge l_i^0 | h_\theta) &= \delta'\phi_1(1 - \lambda_1) + (1 - \delta')\phi_0(1 - \lambda_0) \equiv \eta'_{10}, \\ P'(f_i^1 \wedge l_i^1 | h_\theta) &= \delta'\phi_1\lambda_1 + (1 - \delta')\phi_0\lambda_0 \equiv \eta'_{11}. \end{aligned} \quad (14)$$

The system of equations that results from equating likelihoods and observed relative frequencies is:

$$\eta_{jk} = \frac{r_{jk}}{t}, \quad \eta'_{jk} = \frac{r'_{jk}}{t}.$$

Each of these two constrains the parameters in θ in a particular way.

The Appendix to this paper shows that if this system of equations has a solution, then the solution is unique up to a transformation of the two values for D . Solutions thus come in mirror-image pairs, differing in the interpretation of the values for the variable D or, in other words, differing in whether the intervention has beneficial or adverse effects on the probability of being depressed. On the assumption that the treatment reduces the probability for depression, every hypothesis h_θ in the model is associated with a unique set of values for the likelihoods η_{jk} and η'_{jk} . The conclusion is that if the data are generated by a chance process specified by a hypothesis h_θ , then we can identify this hypothesis, in the same way as we were able to identify the true h_θ in the model of Equation (1).

As indicated, this does not hold for the entire range of possible values for the observed frequencies. For extremal values there is still an infinity of solutions. Moreover, certain combinations of frequencies simply do not match with any of the statistical hypotheses within the model. In those cases the intervention data overdetermine the factor model, and the factor model fails to fit all the correlations. We must then look for a richer statistical model. It seems rather natural to incorporate this aspect of scientific reasoning in our account, and describe how statistical models are adapted when intervention data yield a bad fit. The idea is that the overdetermination due to intervention may lead to controlled and formally specified changes in the model,

and that this may lead to a formal account of theory change. However, such an account is beyond the scope of the current paper.

The main conclusion for now is that intervention data can indeed be used to resolve the statistical underdetermination, as it was introduced in §2.1. If there are parameter values matching the relative frequencies exactly, then on the assumption that the treatment is beneficial, these values are unique: the likelihood function has a unique maximum after the normal and the intervention data are incorporated. While we have only shown this for a simple example, it is readily seen, and briefly considered in the Appendix, that the example generalizes. The example serves as a proof of principle and supports the central idea of this paper, which is that interventions can be used to adjudicate between previously indistinguishable hypotheses and thereby replace theoretical criteria that fulfil this role.

5 Philosophical and practical implications

We now discuss the philosophical and practical implications of the foregoing. After that we briefly revisit the indeterminacy of factor scores. Intervention data can be used to resolve this indeterminacy, at least in the form it takes within a Bayesian statistical model.

5.1 Interventions replace theoretical criteria

The philosophical upshot of this is that empirical criteria for theory evaluation, based on the targeted acquisition of intervention data, can take the place of the theoretical criteria that normally guide theory choice in the face of underdetermination. Where we had otherwise used a theoretical criterion to choose among the equally well fitting alternative hypotheses, we can now decide on the basis of additional data, obtained after intervention. Within statistics, one might say, the problem of underdetermination has fuzzy edges: it can be resolved by an appeal to theoretical criteria, but it can also be resolved by extending the realm of observations with intervention data. It is worth emphasizing that we do not need to know anything about the exact impact of the intervention. That is, we do not need to know the exact value of δ' . The mere fact that we have changed something to the probability of the latent variable suffices.

Clearly, this is not to say that the use of intervention data requires no assumptions whatsoever. As indicated in the foregoing, the new data can only be taken as pertaining to the same parameters if we assume that the causal structure of the latent and observed variables is, at least roughly, correct. More specifically, we need to assume that the probabilistic relations between the latent and the observed variables, expressed in ϕ_i and λ_i , remain invariant under intervention. So in order to employ the intervention data for a resolution of the statistical underdetermination, we have to make particular causal assumptions. Nevertheless we think that the resolution of underdetermination by causal assumptions and further empirical data is to be preferred over a resolution that employs a theoretical criterion only.

The resolution of underdetermination is of philosophical interest: if there is less empirical underdetermination in factor analysis than commonly thought, then factor analysis, which has long been regarded by some as somewhat speculative (see, e.g., Furfey and Daly (1937)), is put on a firmer footing. The use of interventions to resolve underdetermination in factor analysis is also of methodological and practical interest. Recall the problem of underdetermination due to the rotation of latent variables, as discussed in §2.3. This rotation problem is particularly pressing for the design of clinical and personality tests: how do we relate clusters of test items to specific personality traits? And what traits should we distinguish in the first place? The fact that we can opt for a multitude of latent structures, each associated with a different causal story on how the correlations between observed variables has emerged, presents researchers with a genuine problem. The standard response to this problem is to employ theoretical criteria on the latent variables, for example by supposing that the traits are independent, or by choosing the latent variables such that the regression parameters show maximal variation, thus associating each test with a minimal number of traits.

The idea of the present paper is that these theoretical criteria can be replaced by intervention data. For example, for clinical psychologists working with factor analysis, interventions may constrain the latent structure behind their tests, thereby providing a clearer view of what the tests are measuring. However, it leads us to far away from the line of this paper to explicate an application here or work out the statistical details. For a more extensive discussion on the use of interventions, we refer to Hyttinen *et al*

(2012), which presents a more general treatment of the use of intervention data for identifying causal models.

5.2 Interventions and the indeterminacy of factor scores

We briefly remark on the problem of the indeterminacy of factor scores, as discussed in 3.2. Insofar as there is a problem with factor scores in the Bayesian treatment, intervention data can play an interesting role.

Recall that the expected value $E[\theta]$, given in Equation (9), depends on the posterior probability over the parameter $P(h_\theta|s_t)$, and that according to Equation (8), this posterior depends on the prior probability $P(h_\theta)$. As shown in Bartholomew and Knott (1999), the indeterminacy of factor scores in classical factor analysis derives directly from the fact that a prior probability is not provided. And because in a Bayesian treatment such a prior is assumed, we can say that Bayesian factor analysis is not affected by factor score indeterminacy. However, the prior is assumed, not derived, so a classical statistician may well ask for a motivation of the prior probability assignment.

Following the ideas of the foregoing, the prior probability may be determined by means of intervention data. Instead of choosing a single prior, we might consider a whole collection of possible priors over the parameter values. For example, we might consider as priors all so-called symmetric Beta-distributions,

$$P(h_\theta) = \frac{(2n-1)!}{((n-1)!)^2} \delta^{n-1} (1-\delta)^{n-1},$$

parameterized by the natural numbers $n > 0$. Effectively, we thereby increase the dimension of the parameter space by one. But we might know from a different study that the chance of being depressed after the treatment δ' has some particular value, or is functionally related to the chance on depression before treatment. This reduces the number of parameters by one again, because δ' is then fixed, or every δ' is coupled to a unique value δ . The net effect is that we can again estimate all the parameters, namely δ , ϕ_j and λ_j for $j = 0, 1$, and finally the second-order parameter n .¹⁰

¹⁰In the statistical literature, the idea that we can confirm or disconfirm probability distributions over statistical parameters has become known as hierarchical Bayesian modelling. See, for instance, Chapter 5 of Gelman (2004), [omitted for blind reviewing] ,

In other words, just as we can estimate the effects of an intervention, δ' , we can estimate the prior probability assignment that best suits the factor model. Of course, this is just a toy example. We have not said anything about the more realistic continuous case, in which we typically assume a normal distribution over the continuous variable D_i as prior. Moreover, it is unrealistic to suppose that there is a clear and deterministic relation between the parameters governing the distribution over the variables D_i before and after the intervention.

Nevertheless, we maintain that the foregoing illuminates how intervention data can be of use in dealing with the rightful heir of the problem of factor score indeterminacy in Bayesian factor analysis, namely the problem of how to choose a prior.

6 Conclusion

In this paper we have investigated the use of interventions for the problem of statistical underdetermination: if two statistical hypotheses have exactly the same likelihoods for all the possible observations, then how do we choose between them? While an answer to this question often invokes theoretical criteria such as simplicity and explanatory considerations, we have provided a partial answer in terms of empirical criteria. The idea is to use the background theory that generates the hypotheses, namely the causal picture. This theory provides us with a recipe for how to deal with interventions. Together with some assumptions on the causal structure of the latent and observed variables, the intervention data enable us to tell the statistically underdetermined hypotheses apart.

We illustrated the problem of underdetermination by means of factor analysis. That is, we have worked out how interventions can be framed in terms of alterations to the factor model, and how the intervention data can then be employed to resolve the underdetermination of the factor loadings. In this paper we have not applied the same ideas to the more practical setting of factor analysis with normal distributions over continuous variables. But we believe that the underdetermination problem identified in discrete Bayesian networks is in all the relevant respects similar to the underdetermination and the recent philosophical appraisal in Henderson *et al* (2010) and [omitted for blind reviewing].

mination associated with the rotation of factors in the continuous setting, and we are confident that in future work we can present a resolution of this problem of rotation on the basis of intervention data. The aforementioned paper of Hyttinen *et al* (2012) goes a long way in this direction. Our contribution lies in the philosophical reception of these ideas and their connection to foundational problems in methodology generally, and in factor analysis particularly.

We like to mention one specific theme for future research. We suggested that, relative to a given causal picture that links latent and observable variables, intervention data can also guide extensions of the statistical model. The rough idea is that the specifics of the misfit between model and intervention data will suggest how the latent structure might be adapted to repair the fit. Model selection techniques and further considerations of complexity or conservativity might then determine which of these adaptations is most appropriate. The methods and algorithms for putting this idea to work have yet to be determined, but we think that there are many potential applications of the idea. A tool for guiding extensions of statistical models can be of use to experimental scientists, but also to computer scientists working on the automated search of network structures.

Such applications lie within the realm of statistical methodology. However, there may also be a rather different application of the present ideas, within more traditional philosophy of science. The confirmatory practice of scientists has received a lot of attention from formally oriented philosophers of science, often with the aim of explaining or rationalizing science, or to provide scientists with norms that guide the inference from data to theory. Experimental practice, on the other hand, has not been subject to the same scrutiny from the point of view of formal modelling. Experiments have been the subject of science studies, but formal philosophers of science have by and large avoided the subject. We believe experiments should be included among the topics of formal philosophy of science, especially because the tools to describe interventions in mathematical terms are available. We hope that with the present study, we are contributing to the development of a formal philosophy of experiment.

Acknowledgements

Omitted for the purpose of blind reviewing.

Appendix

This appendix substantiates the claim that if the system of Equations (6) and (14) has a solution, then this solution is unique. We are only dealing with the specific example of this paper and do not generalize the result. The actual generalization will bring rather cumbersome algebraic expressions and, we believe, little added insight. The reader may glean the strategy for an analytical investigation of the solution space, and an associated proof strategy for the general case, from what follows.¹¹

We first simplify the expressions to

$$\begin{aligned}\eta_{10} + \eta_{11} &= \delta\phi_1 + (1 - \delta)\phi_0 \equiv f, \\ \eta_{01} + \eta_{11} &= \delta\lambda_1 + (1 - \delta)\lambda_0 \equiv l, \\ \eta'_{10} + \eta'_{11} &= \delta'\phi_1 + (1 - \delta')\phi_0 \equiv f', \\ \eta'_{01} + \eta'_{11} &= \delta'\lambda_1 + (1 - \delta')\lambda_0 \equiv l',\end{aligned}\tag{15}$$

where f and l are the frequencies of fear and loathing respectively. We can now solve for δ as well as δ' and derive the first set of four constraints on the parameters:

$$\begin{aligned}\delta &= \frac{f - \phi_0}{\phi_1 - \phi_0} = \frac{l - \lambda_0}{\lambda_1 - \lambda_0}, \\ \delta' &= \frac{f' - \phi_0}{\phi_1 - \phi_0} = \frac{l' - \lambda_0}{\lambda_1 - \lambda_0}.\end{aligned}\tag{16}$$

The intuitive meaning is that f and f' must both sit in between ϕ_0 and ϕ_1 , and that the relative positions of f and f' within this interval must be equaled by the relative positions of l and l' in between λ_0 and λ_1 . In terms of freedom in the parameter space, there are thus two degrees of freedom left. If, for example, we determine ϕ_0 and ϕ_1 , the values for λ_0 and λ_1 as well as the values for δ and δ' follow.

¹¹With the aid of the solver in Mathematica, we have also investigated this space numerically. Special thanks go to David Atkinson for providing help with this, and for initially presenting us with an alternative, more elegant proof of uniqueness.

We now determine these two values by a further set of two equations. We can judiciously substitute terms appearing in Equations (15) into the expressions for fear and loathing occurring together:

$$\begin{aligned}\eta_{11} &= \delta\phi_1\lambda_1 + (1 - \delta)\phi_0\lambda_0 \equiv c, \\ \eta'_{11} &= \delta'\phi_1\lambda_1 + (1 - \delta')\phi_0\lambda_0 \equiv c' .\end{aligned}$$

Here we abbreviate the frequencies of fear and loathing occurring together as c and c' . With some algebraic reformulation the substitution leads to

$$\lambda_0\phi_1 = f\lambda_0 + l\phi_1 - c, \quad (17)$$

$$\lambda_1\phi_0 = f\lambda_1 + l\phi_0 - c. \quad (18)$$

We can derive the analogous expressions for the parameters by using the frequencies after intervention. Combining the equations we get

$$\lambda_0 = \frac{l' - l}{f - f'}\phi_1 - \frac{c' - c}{f - f'}, \quad (19)$$

$$\lambda_1 = \frac{l' - l}{f - f'}\phi_0 - \frac{c' - c}{f - f'}. \quad (20)$$

Together with the constraints of Equation (16) these two linear relations between the λ 's and ϕ 's are sufficient for determining all the values of the parameters.

To solve the equations we fill in the expression for λ_0 of Equation (19) into Equation (17), thereby obtaining a quadratic equation for ϕ_1 :

$$\left(\frac{l' - l}{f - f'}\phi_1 - \frac{c' - c}{f - f'}\right)\phi_1 = f\left(\frac{l' - l}{f - f'}\phi_1 - \frac{c' - c}{f - f'}\right) + l\phi_1 - c.$$

A parallel expression for ϕ_0 can be obtained by filling in λ_1 of Equation (20) into Equation (18), but if soluble within the domain $[0, 1]$, this expression will yield the same two solutions. Once we choose either of the two solutions for ϕ_1 , the parameter ϕ_0 takes on the other value. And once we have solved for ϕ_1 and chosen whether it obtains the higher or the lower of the two values, we thereby fix the values of all the other parameters. Swapping around the two solutions will effectively swap around the ordering among δ and δ' , according to the expressions above.

Going on the interpretation of depression, fear, loathing, and treatment, the normal case will have $f' < f$, $l' < l$ and $c' < c$ so that $\lambda_0 < \lambda_1$, $\phi_0 < \phi_1$, and $\delta' < \delta$. A further investigation of the space of solutions can be

undertaken by identifying of each point in the space of frequencies whether or not the constraints can all be met. However, for present purposes the abstract characterization suffices, alongside the remark that the space of solutions is non-empty.

References

- Barnett, V. (1999). *Comparative Statistical Inference*. John Wiley, New York.
- Bartholomew, D.J. and Knott, M. (1999). *Latent variable models and factor analysis*. Oxford University Press, New York.
- Douven, I. (2008). Underdetermination. In S. Psillos and M. Curd (eds.), *The Routledge Companion to the Philosophy of Science*, London, Routledge, pp. 292–301.
- Earman, J. (1992). *Bayes or Bust*. MIT press, Cambridge (MA).
- Eberhardt, F. and Scheines, R. (2007). Interventions and causal inference. *Philosophy of Science*, 74:981–995.
- Eberhardt, F. , Hoyer, P. and Scheines, R. (2010). Combining Experiments to Discover Linear Cyclic Models. *Journal of Machine Learning Research* 9:185-192.
- Ellis, J. L. and Juncker, B. W. (1997). Tail-measurability in monotone latent variable models. *Psychometrika*, 62(4): 495–523.
- von Eye, A. and Clogg, C. C. (1994). *Latent variables analysis: applications for developmental research*. Sage, Thousand Oaks (CA).
- Furfey, P. H. and Daly, J. F. (1937). A Criticism of Factor Analysis as a Technique of Social Research. *American Sociological Review*, 2(2): 178–186.
- Gelman, A. *et al.* (2004). *Bayesian Data Analysis, Second Edition*. Boca Raton: Chapman and Hall.
- Gooding, D. (1990) *Experiment and the Making of Meaning* Kluwer, Dordrecht
- Hacking, I. (1983) *Representing and intervening* Cambridge University Press, Cambridge
- Haig, B.D. (2005). An abductive theory of scientific method. *Psychological Methods* 10:371–388.

- Henderson, L., N. D. Goodman, J. B. Tenenbaum, J. F. Woodward (2010). The Structure and Dynamics of Scientific Theories: a Hierarchical Bayesian Perspective. *Philosophy of Science*, 77:172–200.
- Hyttinen, A. , Eberhardt, F. and Hoyer, P. (2012) Causal Discovery for linear cyclic models with latent variables. Manuscript.
- Korb, K., Hope, L., Nicholson, A., and Axnick, K. (2004). Varieties of causal intervention. In *Proceedings of the Pacific Rim International Conference on AI*, New York. Springer.
- Klugkist, I., Laudy, O. and Hoijtink, H. (2005) Inequality Constrained Analysis of Variance: A Bayesian approach. *Psychological Methods* 10:477–493.
- Laudan, L. and Leplin, J. (1991) Empirical Equivalence and Underdetermination. *Journal of Philosophy* 88:449–472.
- Lawley, D.N. and Maxwell, A.E. (1971). Factor analysis as a statistical method. Butterworths, London.
- Maraun, M.D. (1996). Metaphor Taken as Math: Indeterminacy in the Factor Analysis Model. *Multivariate Behavioral Research* 31:517–538.
- McDonald, R.P. (1974). The measurement of factor indeterminacy. *Psychometrika* 39:203–222.
- Mulaik, S.M. (1985). Factor analysis and Psychometrika: Major developments. *Psychometrika* 51:23-33
- Neapolitan, R. E. (2003). *Learning Bayesian networks*. Pearson / Prentice Hall, Upper Saddle River NJ.
- Pearl, J. (2000). *Causality*. MIT press, New York.
- Schurz, G. (2008). Common Cause Abduction and the Formation of Theoretical Concepts. TPD preprints, No.2.
- Spirtes, P., Glymour, C., and Scheines, R. (1993). *Causation, Prediction, and Search*. MIT Press, Cambridge MA, second (2000) edition.
- Steiger, J.H. (1979). Factor indeterminacy in the 1930's and the 1970's some interesting parallels. *Psychometrika* 44:157–167
- Waldorp, L. and Wagenmakers, E.J. (2006) Model selection: Theoretical developments and applications (Special issue). *Journal of Mathematical Psychology* 50.
- Williamson, J. (2005). *Bayesian nets and causality: philosophical and computational foundations*. Oxford University Press, Oxford.