

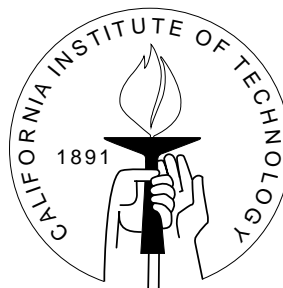
DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

# **CALIFORNIA INSTITUTE OF TECHNOLOGY**

PASADENA, CALIFORNIA 91125

## A BAYESIAN MODEL TO INCORPORATE JOINTLY DISTRIBUTED GENERALIZED PRIOR INFORMATION ON MEANS AND LOADING IN FACTOR ANALYSIS

Daniel B. Rowe



**SOCIAL SCIENCE WORKING PAPER 1110**

February 2001

# A Bayesian Model to Incorporate Jointly Distributed Generalized Prior Information on Means and Loadings in Factor Analysis

Daniel B. Rowe<sup>1</sup>

## Abstract

A Bayesian factor analysis model is outlined in which prior knowledge regarding the model parameters is quantified using prior distributions and incorporated into the inferences along with the data. Recent work (Rowe, 2000a; Rowe, 2000b; and Rowe, 2000c) has focused on the population mean and considered vague, conjugate and generalized conjugate distributions when it was taken to be independent of the factor loadings. More recent work (Rowe, 2001) has taken the population mean and factor loadings to be jointly distributed and used a conjugate prior distribution. In this paper, the population mean vector and the factor loadings are taken to be jointly distributed and a generalized conjugate distribution is used. As mentioned in Press (1982), Rothenburg (1963) pointed out that with a conjugate prior distribution, the elements in the covariance matrices are constrained and may not be rich enough to permit complete freedom of assessment. The generalized conjugate distribution permits complete freedom of assessment. Parameters are estimated by Gibbs sampling and iterated conditional modes algorithms.

---

<sup>1</sup>Requests for reprints should be addressed to Daniel B. Rowe, Division of Humanities and Social Sciences, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA (e-mail: drowe@hss.caltech.edu).

# 1 Introduction

The Bayesian approach to statistics quantifies available prior knowledge regarding the model parameters in the form of prior distributions. Information as to how likely parameter values are is contained in the prior distributions. This is true in Bayesian factor analysis. The prior information regarding the parameters in the form of prior distributions is incorporated into the inferences along with the data.

Recent Bayesian factor analysis work has focused on quantifying and incorporating available prior knowledge regarding the population mean. This recent work (Rowe, 2000a; Rowe, 2000b; and Rowe, 2000c) has considered vague, conjugate normal and generalized conjugate normal distributions for the population mean when it was taken to be independent of the factor loadings. More recent work (Rowe, 2001) has taken the population mean and factor loadings to be jointly distributed and quantified available knowledge regarding their values using a conjugate normal distribution.

As is mentioned in Press (1982), Rothenburg (1963) pointed out that with a conjugate normal prior distribution, the elements in the covariance matrices are constrained and thus may not be rich enough to permit complete freedom of assessment. This is the motivation for the current Bayesian factor analysis model. In this paper, the population mean vector and the factor loading matrix are taken to be jointly distributed and available prior knowledge regarding their values is quantified using a generalized conjugate distribution. The generalized conjugate normal distribution permits complete freedom of assessment.

## 2 Model

### 2.1 Likelihood Function

The Bayesian factor analysis model is:

$$\begin{matrix} (x_j | \mu, \Lambda, f_j) \\ (p \times 1) \end{matrix} = \begin{matrix} \mu \\ (p \times 1) \end{matrix} + \begin{matrix} \Lambda \\ (p \times m) \end{matrix} \begin{matrix} f_j \\ (m \times 1) \end{matrix} + \begin{matrix} \epsilon_j \\ (p \times 1) \end{matrix}, \quad m < p, \quad (2.1)$$

for  $j = 1, \dots, n$ , where  $x_j$  is the  $j^{th}$  observation for subject  $j$ ,  $\mu$  is the overall population mean,  $\Lambda$  is a matrix of constants “common” to all subjects called the factor loading matrix;  $f_j$  is the factor score vector “specific” to each subject  $j$ ; and the  $\epsilon_j$ ’s are observation errors assumed to be mutually uncorrelated and normally distributed  $N(0, \Psi)$  variables as in the traditional model.

In order to incorporate jointly distributed prior knowledge regarding the mean vector and factor loading matrix, the model is rewritten as:

$$\begin{matrix} (x_j | C, f_j) \\ (p \times 1) \end{matrix} = \begin{matrix} C \\ p \times (m+1) \end{matrix} \begin{matrix} g_j \\ (m+1) \times 1 \end{matrix} + \begin{matrix} \epsilon_j, \\ p \times 1 \end{matrix} \quad m < p, \quad (2.2)$$

where  $C = (\mu, \Lambda)$  and  $g_j' = (1, f_j')$ .

It is assumed that  $C$ , the  $f_i$ ’s, and  $\Psi$  are unobservable and that the distribution of each  $x_j$  can be written as

$$p(x_j | C, f_j, \Psi) = (2\pi)^{-\frac{p}{2}} |\Psi|^{-\frac{1}{2}} e^{-\frac{1}{2}(x_j - Cg_j)' \Psi^{-1} (x_j - Cg_j)}. \quad (2.3)$$

The observation vectors can be arranged into a matrix and the model written as

$$\begin{matrix} (X | C, F) \\ (n \times p) \end{matrix} = \begin{matrix} G \\ n \times p \end{matrix} \begin{matrix} C' \\ (m+1) \times p \end{matrix} + \begin{matrix} E, \\ n \times p \end{matrix} \quad (2.4)$$

where the p-variate observation vectors on  $n$  subjects are  $X' = (x_1, \dots, x_n)$ , the factor scores are contained in  $G' = (g_1, \dots, g_n)$ , and the errors of observation are  $E' = (\epsilon_1, \dots, \epsilon_n)$ .

If proportionality is denoted by “ $\propto$ ” then the likelihood for  $(C, F, \Psi)$  is

$$p(X|C, F, \Psi) \propto |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2} \text{tr} \Psi^{-1} (X - GC')'(X - GC')} \quad (2.5)$$

where the notation  $p(\cdot)$  will generically denote a probability distribution which is distinguished by its argument whose proportionality constant does not depend on its argument.

## 2.2 Priors

Prior distributions are specified for the unknown parameters to quantify available prior information. The joint prior distribution for the parameters is:

$$p(c, F, \Psi) \propto p(c)p(F)p(\Psi), \quad (2.6)$$

where

$$p(c) \propto |\Delta|^{-\frac{1}{2}} e^{-\frac{1}{2} (c - c_0)' \Delta^{-1} (c - c_0)}, \quad (2.7)$$

$$p(F) \propto e^{-\frac{1}{2} \text{tr} F' F} \quad (2.8)$$

$$p(\Psi) \propto |\Psi|^{-\frac{\nu}{2}} e^{-\frac{1}{2} \text{tr} \Psi^{-1} Q}, \quad \nu > 2p, \quad (2.9)$$

with  $\Delta$ ,  $Q$ , and  $\Psi$  positive definite matrices. A generalized conjugate normal distribution is specified for the joint distribution of the population

mean and factor loadings. The vector  $c = \text{vec}(C')$  is specified to have the generalized conjugate normal distribution with mean and covariance hyperparameters  $c_0 = \text{vec}(C'_0)$  and  $\Delta$ . The factor score vectors, the  $f_j$ 's have been specified to be normally distributed with mean zero and identity covariance matrix as in the traditional model. The matrix  $\Psi$  follows an inverted Wishart distribution, with hyperparameters  $(\nu, Q)$  which are to be assessed. It is assumed that  $E(\Psi)$  is a priori diagonal and thus  $Q$  is diagonal, in order to represent traditional psychometric views of the factor model containing “common” and “specific” factors.

## 2.3 Joint Posterior

Using Bayes rule, Equations (2.5)–(2.9) are combined to obtain the joint posterior distribution of the parameters

$$p(c, F, \Psi | X) \propto e^{-\frac{1}{2}\text{tr} F' F} |\Delta|^{-\frac{1}{2}} e^{-\frac{1}{2}(c - c_0)' \Delta^{-1} (c - c_0)} \\ \times |\Psi|^{-\frac{(n+\nu)}{2}} e^{-\frac{1}{2}\text{tr} \Psi^{-1} [(X - GC')'(X - GC') + Q]} \quad (2.10)$$

where the variables are as previously defined. Posterior estimates of the population mean, factor loading matrix, the factor score matrix, and the disturbance covariance are to be determined.

## 3 Estimation

### 3.1 Conditional Posterior Densities

Both the Gibbs sampling and ICM procedures of determining values for the model parameters require the posterior conditional distributions.

Gibbs sampling requires the conditionals for the generation of random variates for stochastic integration in order to compute marginal mean estimates, while ICM requires them for the determining of modes in order to compute maximum a posteriori estimates.

The conditional posterior distribution of the vector containing the population mean/factor loadings is

$$\begin{aligned}
p(c|F, \Psi, X) &\propto p(c)p(X|F, C, \Psi) \\
&\propto |\Psi|^{-\frac{m+1}{2}} e^{-\frac{1}{2}(c-c_0)'\Delta^{-1}(c-c_0)} \\
&\quad \times |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}\text{tr}\Psi^{-1}(X-GC')'(X-GC')} \quad (3.1)
\end{aligned}$$

which after some algebra becomes

$$p(c|F, \Psi, X) \propto e^{-\frac{1}{2}(c-\tilde{c})'[\Delta^{-1}+\Psi^{-1}\otimes G'G](c-\tilde{c})} \quad (3.2)$$

where

$$\tilde{c} = [\Delta^{-1} + \Psi^{-1} \otimes G'G]^{-1}[\Delta^{-1}c_0 + (\Psi^{-1} \otimes G'G)\hat{c}] \quad (3.3)$$

and

$$\hat{c} = \text{vec}[(G'G)^{-1}G'X]. \quad (3.4)$$

The conditional posterior distribution of the population mean/factor loading vector given the factor scores, the disturbance covariance matrix, and the data is normally distributed.

The conditional posterior distribution of the disturbance covariance matrix is

$$\begin{aligned}
p(\Psi|F, C, X) &\propto p(\Psi)p(X|F, C, \Psi) \\
&\propto |\Psi|^{-\frac{\nu}{2}} e^{-\frac{1}{2}\text{tr}\Psi^{-1}Q} |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}\text{tr}\Psi^{-1}(X-GC')'(X-GC')} \quad (3.5)
\end{aligned}$$

$$\propto |\Psi|^{-\frac{(n+\nu)}{2}} e^{-\frac{1}{2}tr\Psi^{-1}[(X-GC')'(X-GC')+Q]}. \quad (3.6)$$

That is, the conditional distribution of the disturbance covariance matrix given the factor scores, the population mean/factor loadings, and the data follows an inverted Wishart distribution.

The conditional posterior distribution of the factor scores is:

$$\begin{aligned} p(F|\mu, \Lambda, \Psi, X) &\propto p(F)p(X|\mu, F, \Lambda, \Psi) \\ &\propto e^{-\frac{1}{2}tr F'F} |\Psi|^{-\frac{n}{2}} e^{-\frac{1}{2}tr\Psi^{-1}(X-e_n\mu'-F\Lambda')'(X-e_n\mu'-F\Lambda')} \\ &\propto e^{-\frac{1}{2}tr F'F} e^{-\frac{1}{2}tr(X-e_n\mu'-F\Lambda')\Psi^{-1}(X-e_n\mu'-F\Lambda')} \end{aligned}$$

which after some algebra can be written as

$$p(F|\mu, \Lambda, \Psi, X) \propto e^{-\frac{1}{2}tr(F-\tilde{F})(I_m+\Lambda'\Psi^{-1}\Lambda)(F-\tilde{F})'} \quad (3.7)$$

where  $\tilde{F} \equiv (X - e_n\mu')\Psi^{-1}\Lambda(I_m + \Lambda'\Psi^{-1}\Lambda)^{-1}$ .

That is, the factor scores given the population mean/factor loadings, the disturbance covariance matrix, and the data is normally distributed.

The modes of these conditional distributions are  $\tilde{F}$ ,  $\tilde{c}$  (as defined above), and

$$\tilde{\Psi} = \frac{(X - GC')'(X - GC') + Q}{n + \nu}. \quad (3.8)$$

### 3.2 The Gibbs Sampling Algorithm

In order to estimate the parameters of the model from the posterior distribution by Gibbs sampling, start with initial values for  $F$  and  $\Psi$  say  $\bar{F}_{(0)}$  and  $\bar{\Psi}_{(0)}$ . Then cycle through

$$\bar{c}_{(l+1)} = \text{a random variate from } p(c|\bar{F}_{(l)}, \bar{\Psi}_{(l)}, X)$$

$$\begin{aligned}\bar{\Psi}_{(l+1)} &= \text{a random variate from } p(\Psi|\bar{F}_{(l)}, \bar{c}_{(l+1)}, X) \\ \bar{F}_{(l+1)} &= \text{a random variate from } p(F|\bar{c}_{(l+1)}, \bar{\Psi}_{(l+1)}, X).\end{aligned}$$

After the first random variates called the “burn in” are discarded compute from the next  $L$  samples

$$\bar{F} = \frac{1}{L} \sum_{l=1}^L \bar{F}_{(l)} \quad \bar{c} = \frac{1}{L} \sum_{l=1}^L \bar{c}_{(l)} \quad \bar{\Psi} = \frac{1}{L} \sum_{l=1}^L \bar{\Psi}_{(l)}$$

which are the sampling based marginal posterior mean estimates of the parameters.

### 3.3 The ICM Algorithm

In order to estimate the parameters of the model from the posterior distribution by ICM, start with initial values for  $\tilde{F}$ , and  $\Psi$  say  $\tilde{F}_{(0)}$ , and  $\tilde{\Psi}_{(0)}$ , form  $G_{(0)} = (e_n, F_{(0)})$ , then cycle through

$$\begin{aligned}\hat{c}_{(l)} &= \text{vec}[(\tilde{G}'_{(l)}\tilde{G}_{(l)})^{-1}\tilde{G}'_{(l)}X] \\ \tilde{c}_{(l+1)} &= [\Delta^{-1} + \tilde{\Psi}_{(l)}^{-1} \otimes \tilde{G}'_{(l)}\tilde{G}_{(l)}]^{-1}[\Delta^{-1}c_0 + (\tilde{\Psi}_{(l)}^{-1} \otimes \tilde{G}'_{(l)}\tilde{G}_{(l)})\hat{c}_{(l)}] \\ \tilde{\Psi}_{(l+1)} &= \frac{(X - \tilde{G}_{(l)}\tilde{C}'_{(l+1)})'(X - \tilde{G}_{(l)}\tilde{C}'_{(l+1)}) + Q}{n + \nu} \\ \tilde{F}_{(l+1)} &= (X - e_n\tilde{\mu}'_{(l+1)})\tilde{\Psi}_{(l+1)}^{-1}\tilde{\Lambda}_{(l+1)}(I_m + \tilde{\Lambda}'_{(l+1)}\tilde{\Psi}_{(l+1)}^{-1}\tilde{\Lambda}_{(l+1)})^{-1}\end{aligned}$$

where  $\tilde{G}_{(l)} = (e_n, \tilde{F}_{(l)})$ , until convergence is reached with the joint modal (maximum a posteriori) estimator for the unknown parameters  $(\tilde{c}, \tilde{F}, \tilde{\Psi})$ .

## 4 Example

In this section the Gibbs sampling and the ICM procedures for estimating the parameters of the Bayesian factor analysis model are implemented

and the resulting estimators are presented. The data is extracted from an example in Kendall 1980, p.53. The problem as originally stated (Press & Shigemasu, 1989) and in subsequent Bayesian factor analysis papers is the following.

There are 48 applicants for a certain job, and they have been scored on 15 variables regarding their acceptability. They are:

- |                                |                       |
|--------------------------------|-----------------------|
| (1) Form of letter application | (9) Experience        |
| (2) Appearance                 | (10) Drive            |
| (3) Academic ability           | (11) Ambition         |
| (4) Likeability                | (12) Grasp            |
| (5) Self-confidence            | (13) Potential        |
| (6) Lucidity                   | (14) Keenness to join |
| (7) Honesty                    | (15) Suitability      |
| (8) Salesmanship               |                       |

The raw scores of the applicants on these 15 variables, measured on the same scale, are presented in Table 1. The question is, Is there an underlying subset of factors that explain the variation observed in the scores? If so, then the applicants could be compared more easily.

The underlying structure is postulated (Press & Shigemasu, 1989) as in previous work, a model with 4 factors. This choice is based upon a principal components analysis which found that 4 factors accounted for 81.5% of the variance. Based upon underlying theory the prior factor loading matrix

$$\Lambda'_0 = \begin{pmatrix} 0 & 0 & 0 & 0 & .7 & .7 & 0 & .7 & 0 & .7 & 0 & .7 & .7 & .7 & 0 & 0 \\ 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 & .7 \\ 0 & 0 & 0 & .7 & 0 & 0 & .7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

was assessed.

Table 1: Raw scores of 48 applicants scaled on 15 variables.

Person	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	6	7	2	5	8	7	8	8	3	8	9	7	5	7	10
2	9	10	5	8	10	9	9	10	5	9	9	8	8	8	10
3	7	8	3	6	9	8	9	7	4	9	9	8	6	8	10
4	5	6	8	5	6	5	9	2	8	4	5	8	7	6	5
5	6	8	8	8	4	5	9	2	8	5	5	8	8	7	7
6	7	7	7	6	8	7	10	5	9	6	5	8	6	6	6
7	9	9	8	8	8	8	8	8	10	8	10	8	9	8	10
8	9	9	9	8	9	9	8	8	10	9	10	9	9	9	10
9	9	9	7	8	8	8	8	5	9	8	9	8	8	8	10
10	4	7	10	2	10	10	7	10	3	10	10	10	9	3	10
11	4	7	10	0	10	8	3	9	5	9	10	8	10	2	5
12	4	7	10	4	10	10	7	8	2	8	8	10	10	3	7
13	6	9	8	10	5	4	9	4	4	4	5	4	7	6	8
14	8	9	8	9	6	3	8	2	5	2	6	6	7	5	6
15	4	8	8	7	5	4	10	2	7	5	3	6	6	4	6
16	6	9	6	7	8	9	8	9	8	8	7	6	8	6	10
17	8	7	7	7	9	5	8	6	6	7	8	6	6	7	8
18	6	8	8	4	8	8	6	4	3	3	6	7	2	6	4
19	6	7	8	4	7	8	5	4	4	2	6	8	3	5	4
20	4	8	7	8	8	9	10	5	2	6	7	9	8	8	9
21	3	8	6	8	8	8	10	5	3	6	7	8	8	5	8
22	9	8	7	8	9	10	10	10	3	10	8	10	8	10	8
23	7	10	7	9	9	9	10	10	3	9	9	10	9	10	8
24	9	8	7	10	8	10	10	10	2	9	7	9	9	10	8
25	6	9	7	7	4	5	9	3	2	4	4	4	4	5	4
26	7	8	7	8	5	4	8	2	3	4	5	6	5	5	6
27	2	10	7	9	8	9	10	5	3	5	6	7	6	4	5
28	6	3	5	3	5	3	5	0	0	3	3	0	0	5	0
29	4	3	4	3	3	0	0	0	0	4	4	0	0	5	0
30	4	6	5	6	9	4	10	3	1	3	3	2	2	7	3
31	5	5	4	7	8	4	10	3	2	5	5	3	4	8	3
32	3	3	5	7	7	9	10	3	2	5	3	7	5	5	2
33	2	3	5	7	7	9	10	3	2	2	3	6	4	5	2
34	3	4	6	4	3	3	8	1	1	3	3	3	2	5	2
35	6	7	4	3	3	0	9	0	1	0	2	3	1	5	3
36	9	8	5	5	6	6	8	2	2	2	4	5	6	6	3
37	4	9	6	4	10	8	8	9	1	3	9	7	5	3	2
38	4	9	6	6	9	9	7	9	1	2	10	8	5	5	2
39	10	6	9	10	9	10	10	10	10	10	8	10	10	10	10
40	10	6	9	10	9	10	10	10	10	10	10	10	10	10	10
41	10	7	8	0	2	1	2	0	10	2	0	3	0	0	10
42	10	3	8	0	1	1	0	0	10	0	0	0	0	0	10
43	3	4	9	8	2	4	5	3	6	2	1	3	3	3	8
44	7	7	7	6	9	8	8	6	8	8	10	8	8	6	5
45	9	6	10	9	7	7	10	2	1	5	5	7	8	4	5
46	9	8	10	10	7	9	10	3	1	5	7	9	9	4	4
47	0	7	10	3	5	0	10	0	0	2	2	0	0	0	0
48	0	6	10	1	5	0	10	0	0	2	2	0	0	0	0

The hyperparameters were assessed as  $\mu_0 = 7.5e_{15}$ ,  $\Delta = \delta_0 I_{75}$  where  $\delta_0 = 1/100$ ,  $Q = 0.2I_{15}$ , and  $\nu = 33$ . The 15 dimensional unit vector has been denoted by  $e_{15}$ . The population mean, factor loadings, factor scores,

and disturbance covariance matrix may now be estimated. It was found that a burn in period of 5, 000 samples worked well, so then the next 25, 000 samples were taken for the Gibbs estimates.

Table 2 displays the Gibbs sampling and ICM estimates of the population mean along with the prior and sample means.

Table 2: Gibbs Sampling and ICM estimates of the mean.

p	Gibbs Mean	ICM Mean	Sample Mean	Prior Mean
1	7.4428	7.5036	6.0000	7.5000
2	7.4227	7.4281	7.0833	7.5000
3	7.3676	7.3823	7.0833	7.5000
4	7.0546	7.0445	6.1458	7.5000
5	7.7018	7.6952	6.9375	7.5000
6	7.6340	7.6397	6.3333	7.5000
7	7.8716	7.8640	8.0417	7.5000
8	6.7194	6.7076	4.7917	7.5000
9	6.5418	6.6060	4.2292	7.5000
10	7.0916	7.0962	5.3125	7.5000
11	7.4439	7.4354	5.9792	7.5000
12	7.6207	7.6459	6.2500	7.5000
13	7.3572	7.3746	5.6875	7.5000
14	6.9244	6.9135	5.5625	7.5000
15	7.8685	7.9574	5.9583	7.5000

Table 3 displays the Gibbs sampling and ICM estimates of the factor loadings. For enhanced interpretability, the rows of the factor loading matrices have been rearranged. It is seen that factor 1 loads heavily for variables 5, 6, 8, 10, 11, 12, and 13; factor 2 heavily on variable 3; factor 3 heavily on variables 1, 9, and 15; while factor 4 loads heavily on variables 4 and 7. These four factors in terms of the original variables are factor 1: Self-confidence, Lucidity, Salesmanship, Drive, Ambition, Grasp, Potential;

Table 3: Gibbs (left) and ICM (right) Estimates of Factor Loadings.

p	1	2	3	4	1	2	3	4
5	.7749	-.0474	-.1603	.0084	.8320	-.0651	-.1984	-.0119
6	.7748	-.0059	-.0294	.0715	.8185	-.0006	-.0515	.0625
8	.7860	-.0436	.0867	-.0677	.8323	-.0611	.0854	-.0971
10	.7232	.0131	.1875	.0120	.7569	-.0043	.1945	-.0035
11	.7958	-.0354	.0095	-.0982	.8533	-.0567	.0034	-.1255
12	.7139	.0682	.0924	.0909	.7508	.0962	.0863	.0944
13	.6754	.1602	.1438	.1892	.7043	.2000	.1446	.2048
3	.0971	.8536	.0726	.0102	.1085	.9838	.0612	.0201
1	.0489	-.0016	.7593	-.0171	.0243	-.0628	.8284	-.0010
9	.0725	.0930	.8645	-.0402	.0318	.1383	.9349	-.0586
15	.2291	-.0403	.7206	.0171	.2082	-.0143	.7894	.0061
4	.1098	.0177	.1407	.7270	.0852	.0056	.1668	.8262
7	.0545	-.0043	-.1957	.7696	.0385	.0145	-.2202	.8705
2	.1905	-.0013	.1001	.1748	.2049	.0365	.1273	.2119
14	.3259	-.2561	.2430	.3312	.3262	-.3665	.2793	.3842

factor 2: Academic ability; factor 3: Form of letter application, Experience, Suitability; and factor 4: Likeability, Honesty. These factors may be loosely interpreted as factor 1 being personality, factor 2 being academic ability, factor 3 being position match, and factor 4 being charisma.

In Table 4, the Gibbs sampling and ICM estimates of the factor scores are presented. Note the similarity of most of the values for the two estimation methods. Factor scores may now be interpreted. For example, if an employer wished to choose a person for the position with a “good” personality and academic ability but was not necessarily a “good” position match or charismatic, person 10 could be selected.

Table 5 displays the Gibbs sampling and ICM estimates of the disturbance covariance matrix. Note the similarity between the two matrices. The variances along the diagonal in the covariance matrix are uniformly smaller for the ICM estimation procedure than for the Gibbs sampling procedure.

Table 4: Gibbs (left) and ICM (right) Estimates of the Factor Scores.

Person	1	2	3	4	1	2	3	4
1	0.1345	-2.9596	-0.4676	-0.5694	0.1072	-2.5971	-0.4669	-0.5542
2	0.7016	-1.4401	0.1724	0.3778	0.7095	-1.3008	0.1755	0.3918
3	0.3574	-2.4718	-0.2090	-0.0315	0.3423	-2.1888	-0.2228	-0.0210
4	-0.8883	0.5324	-0.3954	0.0500	-0.8795	0.6286	-0.4789	0.0181
5	-0.9955	0.4686	0.1153	0.7956	-0.9367	0.6020	0.0859	0.7458
6	-0.4176	-0.0950	0.0413	0.3413	-0.4210	0.0445	-0.0913	0.2596
7	0.3887	0.1789	0.8792	0.2187	0.4067	0.1965	0.8431	0.2176
8	0.6696	0.6234	0.8508	0.2581	0.6780	0.5145	0.8258	0.2853
9	0.1108	-0.3162	0.7929	0.2963	0.1401	-0.2358	0.7394	0.2988
10	1.1936	1.2864	-1.0712	-1.5635	1.1010	1.3768	-1.0191	-1.4631
11	0.9154	1.5635	-1.2729	-2.9267	0.8123	1.5756	-1.1982	-2.6725
12	0.8646	1.5032	-1.4568	-1.0106	0.7861	1.5336	-1.3927	-0.9282
13	-1.2126	0.4158	-0.3338	1.1044	-1.1114	0.4443	-0.3058	0.9939
14	-1.2786	0.5798	-0.1908	0.5709	-1.1733	0.5823	-0.2130	0.5459
15	-1.3209	0.5503	-0.4894	0.7208	-1.2686	0.8233	-0.5487	0.5767
16	0.1947	-0.8481	0.1832	0.0308	0.1850	-0.5082	0.1612	-0.0393
17	-0.2343	-0.2007	0.0108	-0.0550	-0.2140	-0.3107	-0.0401	-0.0694
18	-0.6579	0.2725	-1.1198	-1.0603	-0.6284	0.1191	-1.1032	-0.8997
19	-0.7011	0.3960	-0.9778	-1.3312	-0.6912	0.3210	-0.9760	-1.1749
20	0.1042	-0.3394	-0.9983	0.9042	0.1125	-0.2212	-0.9310	0.8582
21	-0.0505	-0.6726	-1.1644	0.7482	-0.0577	-0.3015	-1.1239	0.6196
22	0.8388	-0.3434	-0.2551	0.6908	0.8235	-0.5886	-0.2228	0.7523
23	0.8393	-0.4134	-0.5336	0.9815	0.8521	-0.5018	-0.4231	1.0283
24	0.6065	-0.2651	-0.3124	1.1806	0.6082	-0.5252	-0.2565	1.1957
25	-1.4653	0.0061	-1.0224	0.3990	-1.3770	0.0170	-0.9954	0.4007
26	-1.2933	0.0347	-0.5533	0.3259	-1.2148	0.0533	-0.5586	0.3148
27	-0.2851	-0.1433	-1.6354	0.9624	-0.2651	0.2101	-1.5518	0.8242
28	-2.1007	-0.7934	-1.6819	-1.5453	-2.0787	-1.1363	-1.7300	-1.4301
29	-2.2565	-1.2764	-1.7700	-2.7212	-2.2377	-1.5408	-1.6920	-2.4859
30	-1.3667	-1.1555	-1.6845	0.4291	-1.3243	-1.2473	-1.7217	0.3418
31	-1.0563	-1.5698	-1.3601	0.5805	-1.0395	-1.6706	-1.3927	0.4943
32	-0.6625	-0.8294	-1.8608	0.5092	-0.7343	-0.7275	-1.9126	0.3594
33	-0.9335	-0.9113	-2.0361	0.5417	-0.9896	-0.7722	-2.0851	0.3620
34	-1.9457	-0.3723	-1.7604	-0.4894	-1.9205	-0.4368	-1.7633	-0.4920
35	-2.5630	-1.4780	-1.1591	-0.3648	-2.4511	-1.3480	-1.2242	-0.3084
36	-1.2528	-0.9264	-0.7561	-0.2182	-1.1899	-0.9295	-0.8103	-0.0960
37	0.1237	-0.6028	-2.1257	-0.9224	0.1033	-0.4238	-2.0469	-0.8441
38	0.1841	-0.6838	-2.0511	-0.6680	0.1735	-0.5999	-1.9136	-0.5639
39	0.8560	0.7530	1.0114	1.1256	0.8184	0.4856	0.9195	1.0281
40	0.9891	0.7317	0.9656	1.0394	0.9505	0.4498	0.8914	0.9506
41	-2.7503	0.5554	1.3567	-2.7862	-2.6501	0.8159	1.1073	-2.6441
42	-3.0832	0.6785	1.3810	-3.3277	-3.0185	0.7472	1.0952	-3.2206
43	-2.0790	1.0863	-0.4081	-0.3189	-2.0421	1.1542	-0.4480	-0.4829
44	0.3683	-0.0517	-0.3039	-0.3576	0.3287	-0.0239	-0.3416	-0.3410
45	-0.7102	1.8317	-0.8250	0.8932	-0.6754	1.4975	-0.8816	0.8355
46	-0.2473	1.8050	-0.9863	1.0544	-0.2121	1.5549	-0.9756	1.0402
47	-2.4612	1.7657	-2.7212	-0.3284	-2.3767	1.7224	-2.7069	-0.4288
48	-2.4342	1.7660	-2.7512	-0.7582	-2.3687	1.7010	-2.7645	-0.8235

## 5 Conclusion

A Bayesian factor analysis model was detailed in which available prior information either from substantive experts or previous experiments can

Table 5: Gibbs (top) and ICM (bottom) Estimates of the Disturbance Covariance Matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	.2079	.0189	-.0180	.0284	.0185	.0105	-.0116	-.0139	-.0973	-.0255	.0227	.0081	-.0010	.0602	-.0694
2		.4561	.0290	-.0174	.0214	-.0461	.0232	.0309	-.0324	-.0628	.0834	.0233	.0091	-.0689	.0466
3			.0515	-.0090	.0003	-.0055	.0027	.0074	.0073	-.0024	.0001	-.0005	-.0048	-.0148	.0145
4				.1101	-.0284	.0193	-.0674	.0050	-.0233	-.0130	.0177	-.0075	.0137	.0359	-.0109
5					.1010	-.0097	.0370	.0011	.0135	-.0138	.0125	-.0265	-.0273	-.0087	-.0192
6						.1070	-.0178	-.0050	-.0021	-.0500	-.0438	.0446	-.0085	-.0210	-.0044
7							.1086	-.0001	.0171	.0129	-.0116	.0025	-.0132	-.0309	.0098
8								.0887	-.0047	.0031	.0003	-.0320	-.0254	.0043	.0245
9									.1521	-.0082	-.0025	-.0010	-.0049	-.0219	-.0184
10										.1410	-.0063	-.0521	.0042	.0321	.0236
11											.0951	-.0127	.0056	.0200	-.0203
12												.1062	.0203	-.0172	-.0011
13													.0841	-.0211	.0046
14														.1821	-.0599
15															.1449
1	.1761	.0047	.0208	.0059	.0190	.0110	-.0156	-.0194	-.0893	-.0297	.0142	.0089	.0005	.0218	-.0751
2		.4072	-.0114	-.0377	.0148	-.0550	.0094	.0198	-.0436	-.0723	.0673	.0021	-.0169	-.0719	.0311
3			.0218	.0030	.0055	-.0087	-.0083	.0092	-.0165	.0080	.0058	-.0151	-.0182	.0364	-.0118
4				.0810	-.0225	.0202	-.0711	.0052	-.0176	-.0122	.0117	-.0106	.0087	.0082	-.0124
5					.0814	-.0138	.0334	-.0049	.0248	-.0146	.0030	-.0270	-.0253	-.0100	-.0112
6						.0968	-.0174	-.0073	.0048	-.0454	-.0458	.0387	-.0093	-.0162	-.0017
7							.0859	.0059	.0317	.0173	-.0108	-.0037	-.0172	-.0323	.0220
8								.0722	-.0027	-.0023	-.0100	-.0321	-.0256	-.0027	.0201
9									.1274	-.0046	-.0022	-.0007	-.0085	-.0047	-.0272
10										.1277	-.0138	-.0487	.0045	.0227	.0184
11											.0750	-.0171	.0001	.0073	-.0238
12												.0896	.0103	-.0107	-.0050
13													.0669	-.0118	-.0039
14														.1095	-.0519
15															.1219

be quantified and incorporated into the inferences along with current data. An added feature of the Bayesian factor analysis model is that there is no need to rotate the factor loading matrix. The rotation is automatically found. In addition, knowledge regarding the parameter values is allowed to accumulate as subsequent data is acquired. Available prior information regarding parameters was incorporated with a joint distribution for the population mean and factor loadings through a generalized conjugate prior distribution which permits complete freedom of assessment and does not suffer from the possible limitation of whether it is sufficiently rich.