



## UvA-DARE (Digital Academic Repository)

### Measurement bias detection through Bayesian factor analysis

Barendse, M.T.; Albers, C.J.; Oort, F.J.; Timmerman, M.E.

**DOI**

[10.3389/fpsyg.2014.01087](https://doi.org/10.3389/fpsyg.2014.01087)

**Publication date**

2014

**Document Version**

Other version

**Published in**

Frontiers in Psychology

[Link to publication](#)

**Citation for published version (APA):**

Barendse, M. T., Albers, C. J., Oort, F. J., & Timmerman, M. E. (2014). Measurement bias detection through Bayesian factor analysis. *Frontiers in Psychology*, 5, Article 1087. <https://doi.org/10.3389/fpsyg.2014.01087>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## APPENDICES

### 1 COMPUTATION OF $\sigma^2$ IN TABLE 1

The computation of  $\text{Var}(X) = \sigma^2(X)$  is done via the formula

$$\sigma^2(X) = a^2 + b^2 + c^2(1 + \rho^2) + d^2 + 2ab\rho.$$

In this appendix we outline how to arrive at this formula. For simplicity, we write  $X$  for  $X_1$ .

Consider the Structural Equation Model in Figure 1 where  $T, V$  are both standard normally distributed with covariance  $\rho$  and product  $T \times V = TV^*$ . Apart from  $T, V, TV^*$  we have a loading  $d$  of another standard normal variable on  $X$ , say  $W$  (denoted by the arrow at the bottom of the figure). The loadings on  $X$  by  $T, V$  and  $TV^*$  are denoted by  $a, b$  and  $c$ , respectively. Basic probability dictates that

$$\begin{aligned} \text{Var}(X) &= a^2\text{Var}(T) + b^2\text{Var}(V) + c^2\text{Var}(TV^*) + d^2\text{Var}(W) \\ &+ 2ab\text{Cov}(T, V) + 2ac\text{Cov}(T, TV^*) + 2bc\text{Cov}(V, TV^*). \end{aligned}$$

Since  $W$  is uncorrelated with  $T, V$  and  $TV^*$ , its covariances are eliminated from above formula. Subsequently, we fill in  $\text{Var}(T) = \text{Var}(V) = \text{Var}(W) = 1$  and  $\text{Cov}(T, V) = \rho$  to obtain

$$\begin{aligned} \text{Var}(X) &= a^2 + b^2 + c^2\text{Var}(TV^*) + d^2 + 2ab\rho + 2ac\text{Cov}(T, TV^*) + 2bc\text{Cov}(V, TV^*) \\ &= a^2 + b^2 + c^2\text{Var}(TV^*) + d^2 + 2ab\rho + 2(a + b)c\text{Cov}(T, TV^*), \end{aligned} \quad (6)$$

where in the second step we use that  $\text{Cov}(T, TV^*) = \text{Cov}(V, TV^*)$  because of symmetry arguments. To proceed further, we need to establish both  $\text{Var}(TV^*)$  and  $\text{Cov}(T, TV^*)$ .

That  $\text{Var}(TV^*) = \text{Var}(T \times V) = 1 + \rho^2$  follows immediately from Bohrnstedt & Goldberger (1969, Eq. (6)).

$$\begin{aligned} \text{Var}(T \times V) &= E^2(T)\text{Var}(V) + E^2(V)\text{Var}(T) + 2E(T)E(V)\text{Cov}(T, V) + \\ &\quad \text{Var}(T)\text{Var}(V) + \text{Cov}^2(T, V) \\ &= 0 + 0 + 0 + 1 \times 1 + \rho^2. \end{aligned} \quad (7)$$

We establish  $\text{Cov}(T, TV^*) = 0$  using Bohrnstedt & Goldberger (1969, Eq. (13)) and supplementary variable  $U \sim N(1, 0)$  (thus,  $U \equiv 1$ ):

$$\begin{aligned} \text{Cov}(T, TV^*) &= \text{Cov}(T \times V, U \times V) \\ &= \text{Cov}(T, U)\text{Cov}(T, V) + \text{Cov}(T, T)\text{Cov}(V, U) \\ &= 0 \times \rho + 1 \times 0 = 0. \end{aligned} \quad (8)$$

That the covariances of  $T$  and  $V$  with  $U$  are zero is obvious:  $U$  does not vary, thus also not covary.

Plugging in (7) and (8) into (6), we obtain

$$\text{Var}(X) = a^2 + b^2 + c^2(1 + \rho^2) + d^2 + 2ab\rho.$$

The last column of Table 1 is computed with this formula.

## 2 SPECIFICATION OF PRIOR INFORMATION

In this appendix, we outline how the prior distributions behind the model are specified. To stay in line with the main part of the manuscript we use  $N(\mu, \sigma^2)$  as notation for a normal distribution with variance  $\sigma^2$ . We employ the notation introduced in Equation (2).

The following parts are in common for the models with latent violator, continuous observed violator and dichotomized violator:

$$\begin{aligned}
 u &\sim N(0, 1) \\
 a &\sim N(0, \tau^{-2}) \\
 b &\sim \begin{cases} N(0, \tau^{-2}) & \text{for items without uniform bias} \\ N(0.4, \tau^{-2}) & \text{for items with uniform bias} \end{cases} \\
 c &\sim \begin{cases} N(0, \tau^{-2}) & \text{for items without nonuniform bias} \\ N(0.4, \tau^{-2}) & \text{for items with nonuniform bias} \end{cases} \\
 d &= \tau^{-2} \\
 \begin{pmatrix} \text{Var}(T) & \text{Cov}(T, V) \\ \text{Cov}(T, V) & \text{Var}(V) \end{pmatrix} &\sim \text{Wishart} \left( \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, 4 \right),
 \end{aligned}$$

where  $\rho$  is either equal to 0 or 0.5 (Table 2). Each of the precisions,  $\tau^2$  has the Gamma(9, 4) distribution as hyperprior.

In case there is a continuous observed violator (conditions 9–16 in Table 4), we adjust the hyperprior for the precision of the observed violator to  $\tau^2 - 1000 \sim \text{Gamma}(4, 9)$ . This way, we ensure that the precision is extremely large and, thus, the variance is extremely small, making the violator practically static.

In case there is a dichotomized observed violator (conditions 17–24 in Table 4), we adjust value  $u$  of the violator to  $u \sim N(0, \tau^{-2} = 1/63)$ .