## Appendix A Technical Details on the Implementation of the Bi-factor Model

The estimation equations and expressions for the covariance matrix of the estimates are easily derived using Fisher's identity (Efron 1977; Louis 1982; Glas 1999). The identity plays an important role in the framework of the EM algorithm, which is an algorithm for finding the maximum of a likelihood marginalized over unobserved data. The principle can be summarized as follows. Let  $L_0(\lambda)$  be the log-likelihood function of parameters  $\lambda$  given observed data  $x_0$ , and let  $L_c(\lambda)$  be the log-likelihood function given both observed data  $x_0$  and unobserved missing data  $x_m$ . The latter is called the complete data log-likelihood. The interest is in finding expressions for the first-order derivatives of  $L_0(\lambda)$ , say, the expressions for  $L'_0(\lambda)$ . Define the first-order derivatives with respect to the complete data log-likelihood as  $L'_m(\lambda)$ . Then Fisher's identity entails that  $L'_0(\lambda)$  is equal to the expectation of  $L'_m(\lambda)$  with respect to the posterior distribution of the missing data given the observed data,  $p(x_m|x_0; \lambda)$ , that is,

$$L_0'(\lambda) = E(L_m'(\lambda)|x_0,\lambda) = \int L_m'(\lambda)p(x_m|x_0,\lambda)dx_m.$$

To apply this framework to IRT, a very general definition of an IRT model is adopted. Assume an IRT model is defined by the probability of a response pattern  $x_n$ , which is a function of a, possibly vector-valued, student parameters  $\theta_n$ , and item parameters *a* and *b*, which are item discrimination and item location parameters of an IRT model. So the IRT model is given by  $p(x_n | \theta_n, a, b)$ . Assume further that the student parameter  $\theta_n$  has a normal density  $N(\theta_n; \mu_{g(n)}, \Sigma_{g(n)})$  where, again, g(n) is the country to which student *n* belongs. The key idea is to view the student parameters  $\theta_n$  as missing data and the item and population parameters *a*, *b*,  $\mu_{g(n)}$ , and  $\Sigma_{g(n)}$  as structural parameters  $\lambda$  to be estimated. Then the complete data log-likelihood for a student *n* is

<sup>©</sup> International Association for the Evaluation of Educational Achievement (IEA) 2016 R. Annemiek Punter et al., Psychometric Framework for Modeling Parental

Involvement and Reading Literacy 2016

R.A. Punter et al., *Psychometric Framework for Modeling Parental Involvement and Reading Literacy*, IEA Research for Education 1, DOI 10.1007/978-3-319-28064-6

$$L_{c(n)}(\lambda) = \log p(\mathbf{x}_n | \theta_{n0}, \theta_{ng}, a, b) + \log N(\theta_{n0}, \theta_{ng}; \mu_{g(n)}, \Sigma_{g(n)})$$

and so the estimation equations are given by

$$\frac{\partial L_0(\lambda)}{\partial a_{i0}} = \sum_n E\left(\theta_{n0}\left(\sum_{j=1}^{m_i} x_{nij} - p_{ij}(\theta_n)\right) \middle| x_n, \lambda\right) = 0$$
$$\frac{\partial L_0(\lambda)}{\partial a_{ig}} = \sum_{n|g(n)=g} E\left(\theta_{ng}\left(\sum_{j=1}^{m_i} x_{nij} - p_{ij}(\theta_n)\right) \middle| x_n, \lambda\right) = 0$$
$$\frac{\partial L_0(\lambda)}{\partial d_{ij}} = \sum_n \left[E\left(p_{ij}(\theta_n) \middle| x_n, \lambda\right) - x_{nij}\right] = 0$$
$$\frac{\partial L_0(\lambda)}{\partial \mu_g} = \sum_{n|g(n)=g} \mu_g - E(\theta_{n0}|x_n, \lambda) = 0$$
$$\frac{\partial L_0(\lambda)}{\partial \sigma_g^2} = \sum_{n|g(n)=g} \sigma_g^2 - E\left(\theta_{n0}^2 - \mu_g^2|x_n, \lambda\right) = 0$$

where all the expectations are relative to the posterior distribution

$$p(\theta_{n0}, \theta_{ng}|x_n, \lambda) \propto p(|\mathbf{x}_n||\theta_{n0}, \theta_{ng}, a, b) N(\theta_{n0}, \theta_{ng}; \mu_g, \Sigma_g).$$

We undertook all calculations using the public domain software package MIRT (Glas 2010). The program uses the EM-algorithm to solve the estimation equations and Gaussian quadrature to evaluate the integrals.

**Open Access** This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, duplication, adaptation, distribution, and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, a link is provided to the Creative Commons license, and any changes made are indicated. The images or other third party material in this chapter are included in the work's Creative Commons license and the respective action is not permitted by statutory regulation, users will need to obtain permission from the license holder to duplicate, adapt, or reproduce the material.

## References

- Efron, B. (1977). Discussion on maximum likelihood from incomplete data via the EM algorithm (by A. Dempster, N. Liard, and D. Rubin). *Journal of the Royal Statistical Society, Series B*, *39*, 1–38.
- Glas, C. A. W. (1999). Modification indices for the 2-PL and the nominal response model. *Psychometrika*, 64(3), 273–294. doi:10.1007/bf02294296.
- Glas, C. A. W. (2010). Multidimensional item response theory (MIRT), manual and computer program. Retrieved from http://www.utwente.nl/gw/omd/Medewerkers/temp\_test/mirt\_ package.zip, http://www.utwente.nl/gw/omd/Medewerkers/temp\_test/mirt-manual.pdf.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society, Series B, 44, 226–233.