

# Missing values in large-scale assessment studies

## *Guest editorial*

*Steffi Pohl<sup>1</sup> & Christian Aßmann<sup>2</sup>*

Within the last decades, the increased availability of data from large scale assessment studies triggered tremendous research on educational issues, for example, international comparisons, competence development, and the role of education for competence development. Despite immense efforts in field work, missing values are inevitably occurring in data collected in large scale assessment studies. Their occurrence requires careful adjustment of existing estimation procedures to provide useful and valid estimates of parameters typically derived from large-scale studies. These adjustments are challenging due to the typical features of large scale assessment data. Often, large scale assessment data exhibits a hierarchical structure with individual measurements clustered within educational institutions or regional units. Furthermore, the data sets are high dimensional with regard to surveyed variables as well as with regard to competence domains.

This special issue on *Missing values in large-scale assessment studies* is interested in bringing together papers dealing with different methodological approaches of handling missing values in large scale studies. The papers stem from different disciplines, address different specific aspects of missing values, use different methodological perspectives, and apply their approaches to different large-scale assessment studies. By bringing them together in one issue, we hope to give an overview of the broad research field and to enrich the view of the reader. The papers in the special issue follow two different general approaches of dealing with missing values. The first one uses model-based methods accounting for item nonresponse in competence test items modeled within Item Response Theory (IRT) models. These methods have proven to be very successful and the papers in the special issue further elaborate on them. The other general approach is imputation, which is one of the state of the art methods for dealing with item nonresponse in variables not modeled within a latent variable framework. The papers in this special

---

<sup>1</sup> *Correspondence concerning this article should be addressed to:* Prof. Dr. Steffi Pohl, Methods and Evaluation / Quality Management, Free University Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany; email: [steffi.pohl@fu-berlin.de](mailto:steffi.pohl@fu-berlin.de)

<sup>2</sup> Dr. Christian Aßmann, Chair of Statistics and Econometrics, Otto-Friedrich-University Bamberg and Leibniz Institute for Educational Trajectories, 96045 Bamberg, Germany; email: [christian.assmann@uni-bamberg.de](mailto:christian.assmann@uni-bamberg.de)

issue dealing with imputation extend that approach, making it applicable to the complex data structure of large-scale assessments.

Focusing on missing values arising in competence data, Rose, von Davier, and Nagengast (2015) compare different approaches for dealing with missing responses in IRT models. They delineate the specific assumptions and restrictions of the models, derive how the models adjust for missing responses, and provide guidelines for the choice of an appropriate method in applications. As the choice of an approach accounting for missing responses within competence test items depends on the specific missing mechanism, Köhler, Pohl, and Carstensen (2015) thoroughly investigate the occurrence of missing values in competence test items. By analyzing data sets from different competence domains and age groups, and considering a large number of explaining variables, they were able to get a thorough picture on the missing data process. Glas, Pimentel, and Lamers (2015) suggest an extension of the model-based approaches that allow a) for polytomous responses, and b) for the incorporation of further explaining variables for the missing process. They derive the model, evaluate its performance in a simulation study, and convincingly show its applicability in an empirical study. Vidotto, Kaptein, and Vermunt (2015) present an imputation model for categorical data enriched with a latent class structure. By consideration of latent classes, the model offers inherent flexibility to account for nonlinearities and hierarchical structures, which are typically present in large-scale assessment data. To account for hierarchical structures explicitly in imputation models without presupposing a specific parametric dependence structure, Vink, Lazendic, and van Buuren (2015) suggest an extension of the semi-parametric multiple imputation approaches. The paper by Abmann, Gaasch, Pohl, and Carstensen (2015) provides a Bayesian strategy to cope with missing values in survey variables, which are used as explaining variables for latent competencies. The suggested approach uses the devices of data augmentation to allow for efficient parameter estimation in a given parametric IRT model.

This special issue, thus, provides a comprehensive overview of current approaches to analyze large scale assessment data and brings together different aspects of it. As treatment of missing values has many facets, not all are covered within each paper or within the special issue. However the suggested strategies and approaches can be considered as complementary, allowing for a comprehensive analysis of large scale assessment data.

## References

- Abmann, C., Gaasch, C., Pohl, S., & Carstensen, C. H. (2015). Bayesian estimation in IRT models with missing values in background variables. *Psychological Test and Assessment Modeling*, 57 (4), 595-618.
- Glas, C. A. W., Pimentel, J. L., & Lamers, S. M. A. (2015). Nonignorable data in IRT models: Polytomous responses and response propensity models with covariates. *Psychological Test and Assessment Modeling*, 57 (4), 523-541.

- Köhler, C., Pohl, S., & Carstensen, C. H. (2015). Investigating mechanisms for missing responses in competence tests. *Psychological Test and Assessment Modeling*, 57 (4), 499-522.
- Rose, N., Von Davier, M., & Nagengast, B. (2015). Commonalities and differences in IRT-based methods for nonignorable item nonresponses. *Psychological Test and Assessment Modeling*, 57 (4), 472-498.
- Vidotto, D., Kaptein, M. C., & Vermunt, J. K. (2015). Multiple imputation of missing categorical data using latent class models: state of the art. *Psychological Test and Assessment Modeling*, 57 (4), 542-576.
- Vink, G., Lazendic, G., & Van Buuren, S. (2015). Partitioned predictive mean matching as a multilevel imputation technique. *Psychological Test and Assessment Modeling*, 57 (4), 577-594.