INTRODUCTION

In an ideal survey or test situation, everyone answers all questions and there is no nonresponse. However, life is far from ideal and nonresponse does occur. There are various forms of nonresponse: unit nonresponse, where a whole unit fails to provide data and item nonresponse, where data on particular items are missing. This special issue of KM is devoted to item-nonresponse. For a state-of-the art discussion of unit nonresponse, see the special (unit) nonresponse issue of the Journal of Official Statistics (JOS, 1999:2).

In the case of item-nonresponse, the unit has participated and has provided data, but data for certain items are unavailable for analysis: there are 'blanks' in the data matrix. This can pose serious problems in the analysis and may lead to biased results. On October 7, 1998, the Social and Economical Sciences chapters of the Netherlands Society for Statistics and Operations Research (VVS) organized a one day conference on 'Handling multivariate missing data.' At this meeting, experts discussed solutions to the problems posed by item non-response. A selection of the reviewed papers is now published in this special issue of KM.

In the first article, Edith de Leeuw argues that prevention of missing data should have priority in questionnaire design. She gives an overview of sources of missing data, and a detailed description of methods to reduce the occurrence of item-nonresponse. However, missing data can never be prevented totally, and in the second article, Mark Huisman gives an overview of simple but effective methods to treat missing data. He describes imputation models that are relatively easy to implement. He then evaluates these methods based on simulation studies, and identifies the conditions under which each method works best.

Els Goetghebeur, Geert Molenberghs, and Michael Kenward take it a step further, and discuss the uncertainty in inferred statistical conclusions due from unintended incompleteness of the data. In their article, they identify two sources of uncertainty in statistical inference: imprecision due to sampling (the classical statistical inference problem), and ignorance due to incompleteness of the data set. They illustrate their approach with two data sets.

The next two articles discuss the treatment of missing data in special designs. Gino Verleye focuses on missing data in linear modeling, and presents and overview of techniques to handle missing data in the context of structural equation modeling (SEM). These techniques are evaluated by simulation. The final preferred method is direct Maximum Likelihood estimation, a technique that is becoming increasingly popular in modern SEM software.

Fiona O'Callaghan concentrates on multiple imputation. She discusses the advantages of multiple imputation above single imputation and describes how nonparametric multiple imputation is implemented in the program SOLAS.

Joop Hox ends with a software overview. He describes and evaluates available software for imputation. This article both discusses issues of user friendliness and examines the results of analyzing a small data set.

For the reader's convenience, we have added a short glossary with much used missing-data terminology.

Finally, this special issue could not have been produced without the help of several dedicated persons. Thanks are due to our colleagues, who acted as advisors and reviewers, and to the social sciences and economical chapters of the VVS, who initiated the conference. Finally, we gratefully acknowledge the support of Margo Jansen, chief editor of *KM*.

Joop Hox and Edith de Leeuw