



# Discrete vs. Continuous: The Transition in Data Processing Systems

Seventh International Workshop  
On Comparative Survey  
Design and  
Implementation (CSDI)

Université de Lausanne

25-27 March 2010



# PRESENTATION OUTLINE

- GROWTH OF CONTINUOUS INTERVIEWING IN SURVEY RESEARCH
- CONTINUOUS INTERVIEWING AND CAI SYSTEMS
- RESPONSIVE DATA PROCESSING SYSTEMS
- SYSTEM IMPLEMENTATION
- DATA QUALITY IMPLICATIONS

# Growth of Continuous Interviewing

- More common in large government-sponsored data collections
- Advantages:
  - Timeliness
  - Frequency
  - Improved quality
  - Reduced costs
  - Increased flexibility in managing content and in field operations (responsive designs)
- Disadvantages:
  - Less reliable national estimates
  - Greater variation in small area estimates
  - New challenges for secondary users

# Continuous Interviewing and CAI Systems

- CAI data entry in theory supports automated data processing:
  - Built-in edit checks
  - Built-in computed variables
  - Since interviews are sent from the field to the coordinating center/project headquarters daily:
    - Ability to evaluate problematic cases quickly
    - Ability to contact interviewer or respondent to check information soon after the interview is completed

# Responsive Data Processing Systems

- Processing tasks automated to check new interviews daily including:
  - Editing and preliminary consistency checking of completed interviews
  - Creation of recoded variables to facilitate analytic use
  - Disclosure review to decide which variables are appropriate for public release
  - Creation of extensive documentation on the entire life cycle to facilitate use by analysts

# Responsive Data Processing Systems - 2

- Processing tasks which occur after the collection of all interviews:
  - Imputation of item missing values
  - Generation of several types of weights dependent on the survey population
  - Variance estimation
  - Any final changes to the data (e.g., top and bottom coding, swapping, perturbation) before public release

# Implementing the System

- Create clear editing rules and establish a set of checking programs that will carefully evaluate the specific features of the data collection instrument
  - Test the programs with practice cases or during a pretest
- Modify checking programs as necessary if questions/response categories change during continuous interviewing
- Run post-processing recode programs after the data are checked and edited
  - Modify if necessary if any cases fail to meet coding rules

# Implementing the System - 2

- Within bounds of editing rules, check internal consistency of cases. For example, in surveys where major family events over long time periods are reported, it is easy for the respondent to provide forget or incorrectly remember dates.
- Finding a single mistake might avoid extensive case-editing or imputation long after the data collection period ends.
- Daily processing in place of a potential lengthy period of processing after data collection.
- Maintain a data and documentation repository which can record all editing decisions as they occur.
- Post-processing information is available to users much faster than in more traditional systems.



# Data Quality Implications

- Describe data checking and cleaning operations from the start of data collection until the end
- Minimize “over editing” of data
- Maximize the “cost-error optimization” ratio
- Take advantage of knowledge of the interviewer whenever appropriate