

Video 3 of 5: Survey Data Quality

Introduction to Phone Surveys



THE WORLD BANK

MANNHEIM
BUSINESS SCHOOL

Definitions of (Survey) Data Quality

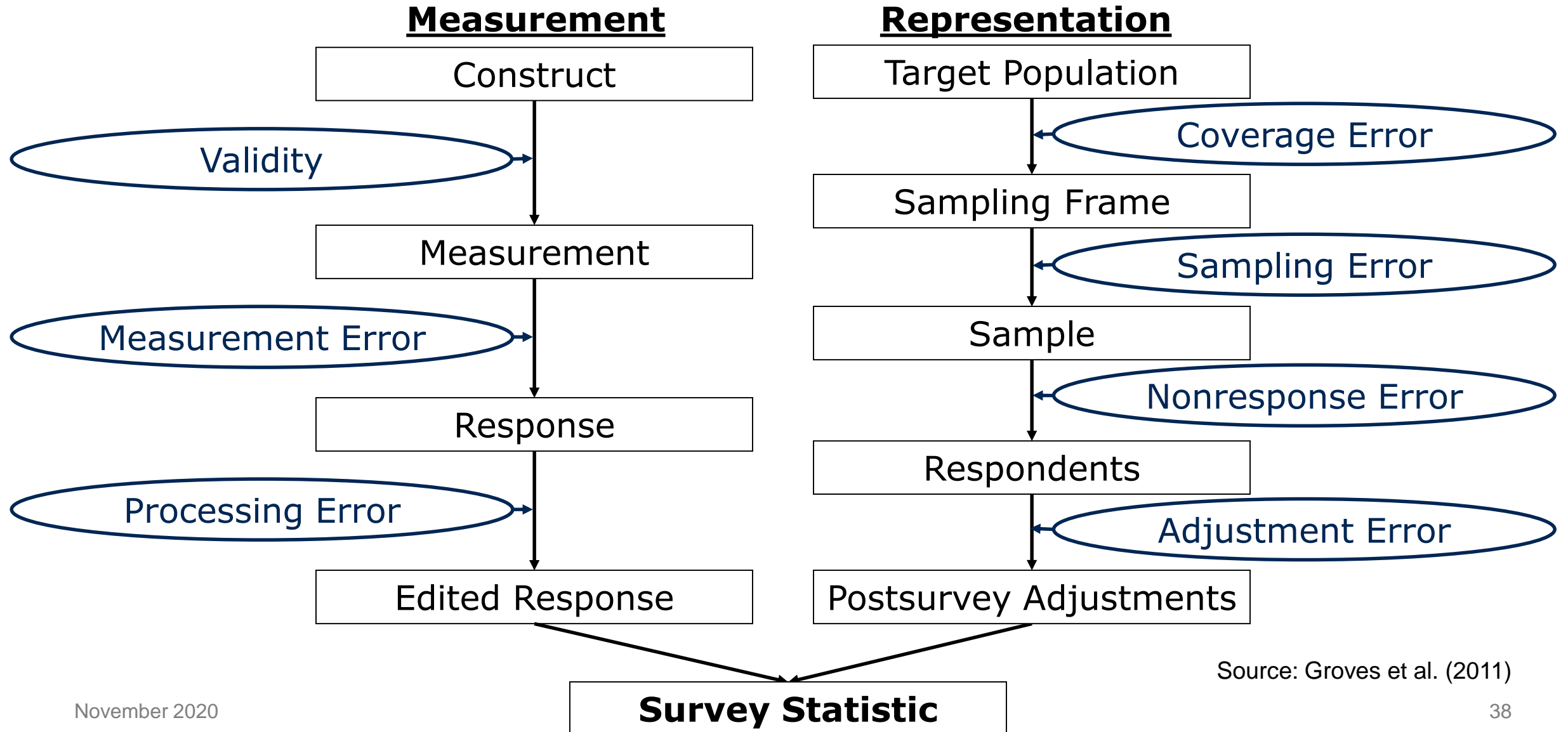
- **Fitness for use** (Juran & Gryna 1980)
 - Freedom from deficiencies
 - Responsiveness to users' needs
- **Different survey quality frameworks**
 - Overall quality dimensions
 - Used by most national statistical offices

Common Dimensions of Survey Quality Frameworks

Dimension	Description
Accuracy	Total survey error minimized
Credibility	Data considered trustworthy by survey community
Comparability	Demographic, spatial, and temporal comparisons valid
Usability/Interpretability	Documentation clear and metadata well-managed
Relevance	Data satisfy users needs
Accessibility	Access to data user friendly
Timeliness/Punctuality	Data deliveries adhere to schedules
Completeness	Data rich enough to satisfy analysis objectives without undue burden on respondents
Coherence	Estimates from different sources can be reliably combined

Source: Biemer (2010)

Survey Lifecycle from a Quality Perspective



Source: Groves et al. (2011)

Total Survey Error

- Concept, way of thinking about various sources of error that may affect survey statistics
- “Error” \neq “mistake”, but rather reflects uncertainty (or lack of confidence) of inference
- Design each component of survey to minimize error inherent to that component
 - For example, design shorter questions for telephone than personal interviews to reduce measurement error
- Assess level of error associated with alternative procedures and choose combination of approaches best suited to problem

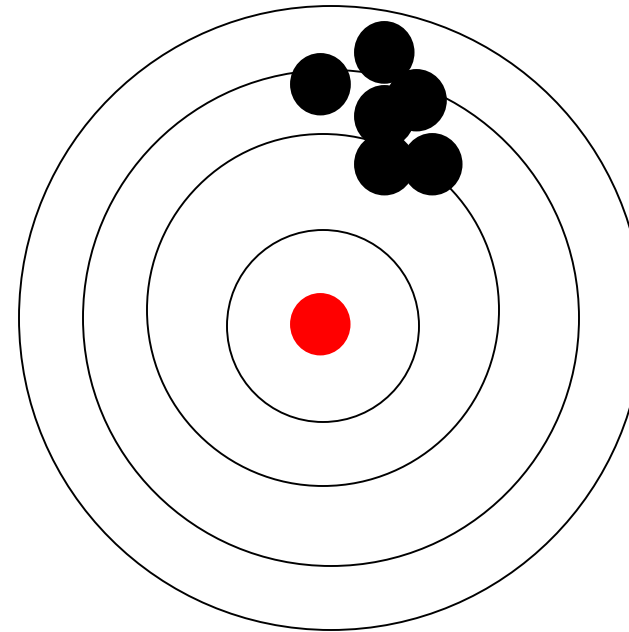
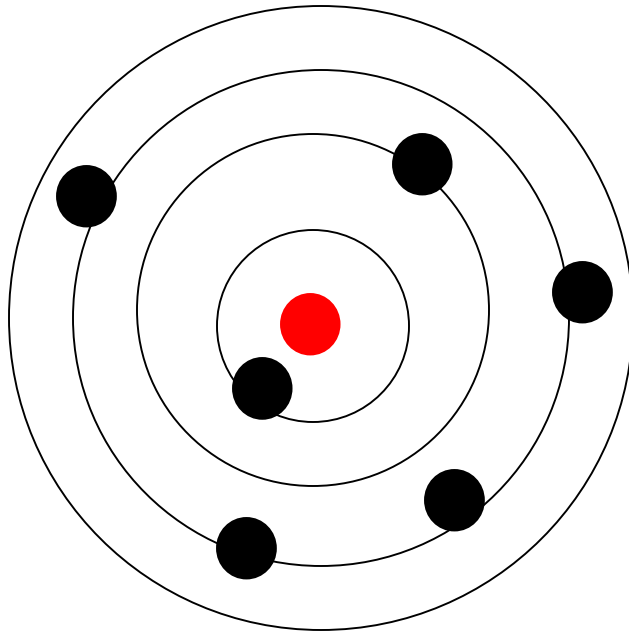
Total Survey Error

- Survey errors can arise from many sources
 - Survey topic, available funding, sampling frame, data collection method, interviewer training, questionnaire design, etc.
- In sum, notion of TSE guides design decisions
 - TSE framework helps understanding potential impact of design decisions on survey errors
 - Together with costs, explicit part of design decisions
- Statistical notion of error expressed as mean square error (MSE)
 - Squared sum of all variable errors and biases
 - Errors specific to certain statistic or estimate
 - In practice MSE rarely fully measured

Variable Error and Bias

- Variable error
 - Variation over replication
 - Often represented by variance of statistic
 - Arises because achieved values vary over different units in design (e.g., interviewers, sample persons, questions)
 - Estimated from sample itself, using replication-based methods
- Bias
 - Systematic deviation from “true value”
 - Directional error (e.g., bigger reports than actually the case)
 - To estimate bias requires external data (“truth”) or assumptions about direction of effects

Variable Error and Bias



END OF VIDEO 3