The Total Survey Error Paradigm and Comparison Error: A Component-Level Evaluation

Tom W. Smith

NORC at the University of Chicago

International Workshop on Comparative Survey Design and Implementation

Warsaw, March, 2019

Figure 1: TSE

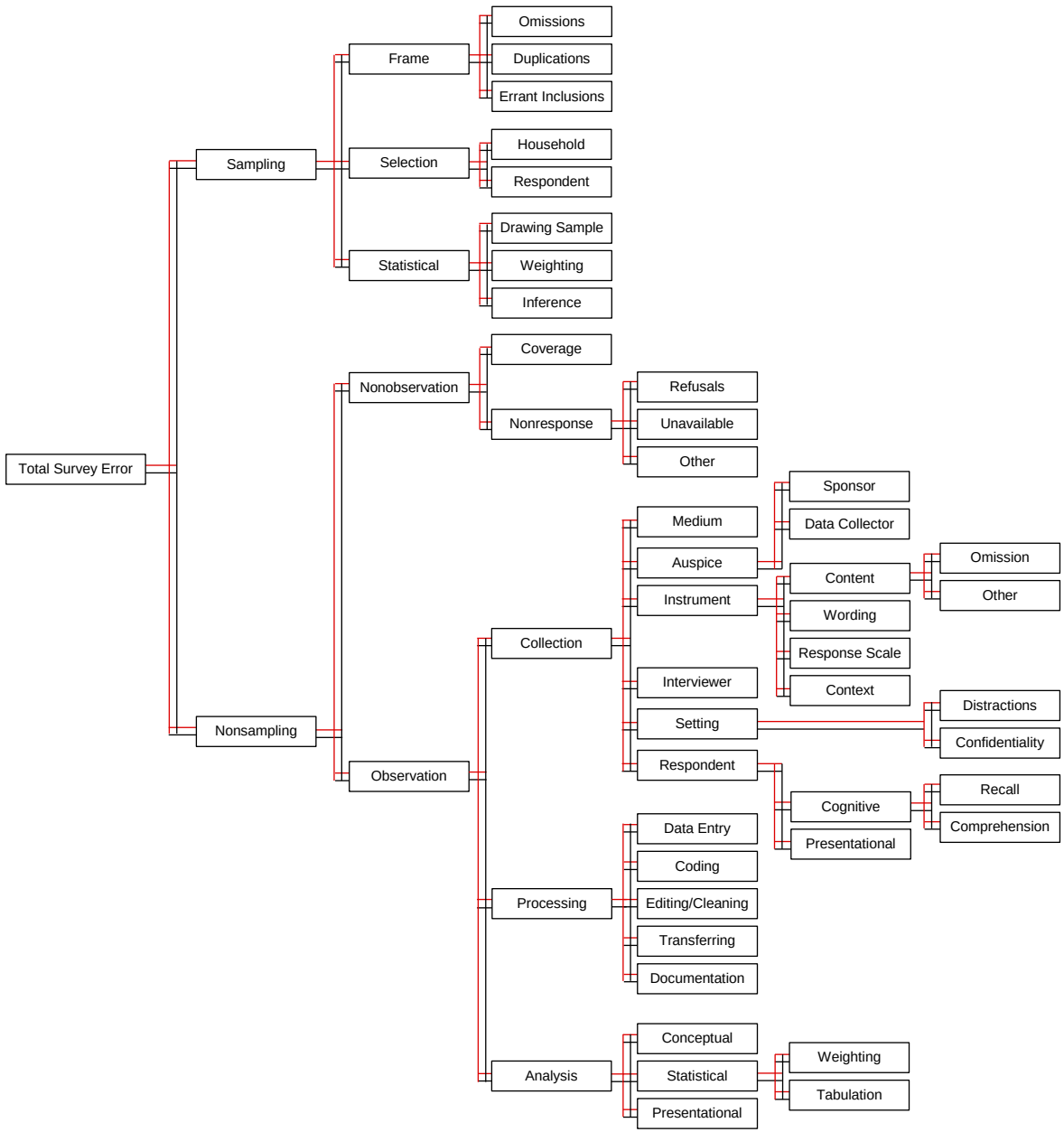
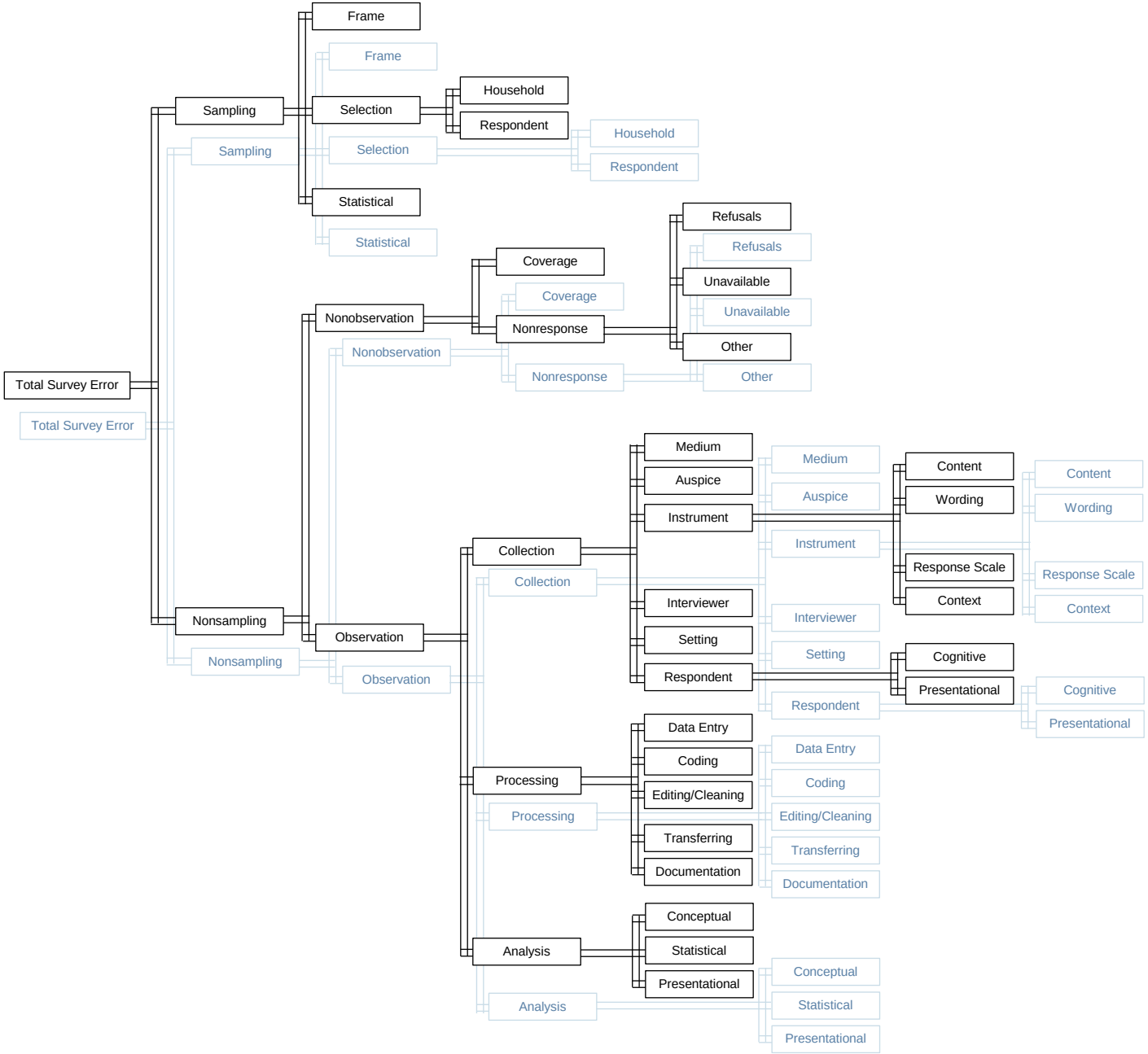

Total Survey Error
Sampling
Frame
Omissions
Duplications
Errant Inclusions
Selection
Household
Respondent
Statistical
Drawing Sample
Weighting
Inference
Nonsampling
Nonobservation
Coverage
Nonresponse
Refusals
Unavailable
Other
Observation
Collection
Medium
Auspice
Sponsor
Data Collector
Instrument
Content
Omission
Other
Wording
Response Scale
Context
Interviewer
Setting
Distractions
Confidentiality
Respondent
Cognitive
Recall
Comprehension
Presentational
Processing
Data Entry
Coding
Editing/Cleaning
Transferring
Documentation
Analysis
Conceptual
Statistical
Weighting
Tabulation
Presentational

Figure 2: TSE and Multiple Surveys



Uses of TSE in Comparative Perspective

The TSE paradigm is a valuable approach for comparative studies for several reasons.

First, it is a blueprint for designing studies. Each component of error can be considered with the object of minimizing comparison error.

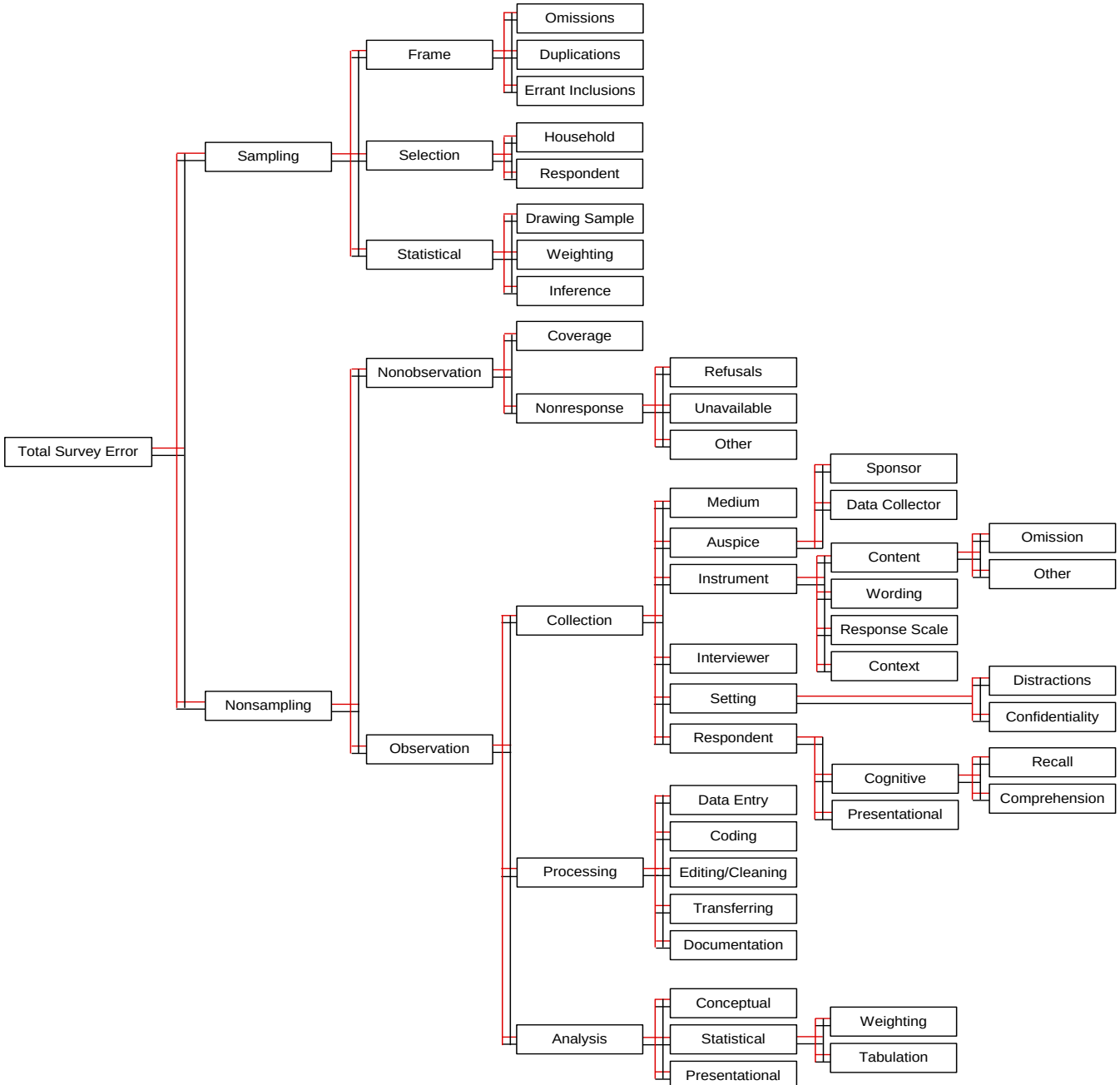
Second, it is a guide for evaluating error after the surveys have been conducted. One can go through each component and assess the level and comparability of the error structures.

Third, it can set a methodological research agenda for study error and for the design of experiments and other studies to fulfill that agenda.

Fourth, it goes beyond examining the separate components of error and provides a framework for the combining of the individual error components into their overall sum.

Fifth, by considering error as an interaction across surveys, it establishes the basis for a statistical model for the handling of error across surveys.

Figure 3: TSE


Total Survey Error
Sampling
Frame
Omissions
Duplications
Errant Inclusions
Selection
Household
Respondent
Statistical
Drawing Sample
Weighting
Inference
Nonsampling
Nonobservation
Coverage
Nonresponse
Refusals
Unavailable
Other
Observation
Collection
Medium
Auspice
Sponsor
Data Collector
Instrument
Content
Omission
Other
Wording
Response Scale
Context
Interviewer
Setting
Distractions
Confidentiality
Respondent
Cognitive
Recall
Comprehension
Presentational
Processing
Data Entry
Coding
Editing/Cleaning
Transferring
Documentation
Analysis
Conceptual
Statistical
Weighting
Tabulation
Presentational

Approaches to Minimizing Comparison Error

- Do it exactly the same way (SW)
- Do it in a equivalent way (EW)
- Do it a different way (DW)
- Combinations (C)

Sampling (EW)

As Kish (1994, p. 173) notes, “Sample designs may be chosen flexibly and there is no need for similarity of sample designs. Flexibility of choice is particularly advisable for multinational comparisons because the sampling resources differ greatly between countries. All this flexibility assumes probability selection methods: known probabilities of selection for all population elements.”

	Design Effect	Aux. Data
Population Register	Low	Person-level
ABS/Multi-stage area probability sample	Moderate+	Area-level
Random Route/Walk	Moderate+	Uncertain

Interview (Validation/Verification) (EW-SW-C)

1. Verification Reinterview
2. Allensbach Trick Question
3. Allensbach Handwriting Comparison
4. Field Supervision with Team of Interviewers
5. CAPI Time Stamps (Length/Proximity)
6. CARI
7. GPS Readings
8. Cross Checking Against Databases
9. Duplication Analysis
10. Other Data Analysis Techniques

Data Entry

DW CAPI vs. PAPI – different data entry => different error structures

CAPI -Skip driven, but not data entry checks, comment capture field
PAPI - Manual following of skips, need to data enter from hardcopy, manual entry of comments in margins

SW Both PAPI - double data entry and reconciliation/two pass verification

Non-Response

SW – AAPOR/WAPOR Standard Definitions; common case management system; callback rules

EW - But the survey climate is likely to be different, the effectiveness of different persuasion messages variable, optimal contact days and times will differ, etc. So procedures to contact and persuade sampled persons to become respondents will vary across countries to achieve the goals of increasing the response rate and reducing non-response bias.

Question Wording

SW – Could employ same translation procedures (e.g. committee translation, rules for written/full translations vs. ad hoc interviewer translations, minority languages to be covered, etc.

EW – Any translation by its very nature is not the same and can only hope to be equivalent.

Data Collector (DW)

House Effects – Data Collector and country usually totally overlap and are indistinguishable analytically