

# **2021 International Total Survey Error Workshop**

## **TSE for Data Collection about COVID**

**24 September 2021**

**Discussant:** Paul J. Lavrakas, Ph.D.

# Background

- Sources of Information Upon Which to Base Comments
  - Presenter's advance slide sets
    - Additional detail in Notes sections
  - Presenter's supplementary materials/links/etc.
  - Additional internet searches
  - Follow-up questions to presenters
  - Today's oral presentations

# Total Survey Quality (TSQ)

- Is a Survey Fit for Purpose? (cf. Santos' 2014 AAPOR Presidential Address; *POQ*, 787(3), 769-777)
- TSQ Framework (cf. Biemer and Lyberg, 2003, *Introduction to Survey Quality*, Wiley; Devers et al., 2019 *JSSAM* article)
  - **Timeliness** – Are the estimates for the time period that is needed, and available when needed?
  - **Completeness** – Are estimates available for all necessary constructs?
  - **Accessibility** – Are data/findings readily available to those who seek/need them?
  - **Interpretability** – Are findings easily understood by relevant users?
  - **Relevance** – Are data sufficient for the necessary analyses to be carried out?
  - **Coherence** – Are estimates consistent with other likely-to-be-valid sources?
  - **Accuracy** – Do estimates describe the target population within acceptable levels of error (Total Survey Error)

# Total Survey Error (TSE)

- All Major Sources of Bias and Variance (cf. Groves, *Survey Errors and Survey Costs*, Wiley, 1989)
  - Representation
    - **Coverage** – how well does the frame represent the target population
    - **Sampling** – what sampling design is used and how well does the initial sample represent the frame and what level of precision does it provide
    - **Nonresponse** – how well does the final/responding sample represent the target population
    - **Adjustment** – to what extent does weighting reduce bias on key statistics without excessively increasing variance
  - Measurement
    - **Specification** – are all key constructs included for measurement
    - **Measurement/Response**
      - **Questionnaire** – how well are constructs operationalized
      - **Respondents** – to what extent are respondents able and willing to provide reliable and valid data
      - **Interviewers** – to what extent do interviewer increase or decrease bias and variance in the data they gather/generate
      - **Mode of Data Collection** – does this mode affect the quality of the data that respondents provide
    - **Processing** – does the processing of the raw data increase and/or decrease biases and variance in the final data used for analyses

# TSQ Evaluation

	<b>Fields USA</b>	<b>Cornesse Germany</b>	<b>Phelps UK</b>	<b>Moore UK</b>
<b>Timeliness</b>	Yes	Unknown, but assume Yes	Results for the Prime Minister and others	Unknown, but assume Yes
<b>Completeness</b>	Unknown , but assume at least Partial	Unknown, but assume least Partial	Acknowledged to be Partial	Unknown, but assume at least Partial
<b>Accessibility</b>	Unknown, but assume Yes	Unknown, but assume Yes	Yes	Unknown, but assume Yes
<b>Interpretability</b>	Unknown, but assume Yes	Unknown, but assume Yes	Yes	Unknown, but assume Yes
<b>Relevance</b>	Unknown, but assume Yes	Unknown, but assume Yes	Partial	Unknown, but assume Yes
<b>Coherence</b>	Nothing to compare against	Nothing to compare against	Nothing to compare against	Nothing to compare against
<b>Accuracy</b>	<b>ADDRESSED ON NEXT SLIDE</b>			

# TSE Evaluation

	<b>Fields USA</b>	<b>Cornesse Germany</b>	<b>Phelps UK</b>	<b>Moore UK</b>
<b>Coverage</b>	Full coverage	Unknown, and unknown coverage biases	High coverage	Infer high coverage
<b>Sampling</b>	Systematic probability sample	From existing panel	Multiple sample sources	From existing panel
<b>Nonresponse</b>	Partial contact info; diff. NR, likely biases; 8% RR	High response, but Unknown NR biases	Raise RRs vs. combat diff. NR; some EXPs	Bias prevention; Bias adjustment
<b>Adjustment</b>	NR adjustments	Unknown	Weighting vs. Modeling	Considerable effort here, two approaches
<b>Specification</b>	Unknown	Unknown	Rushed	Unknown
<b>Questionnaire</b>	Unknown	Unknown	Rushed creation, barriers for input	Unknown
<b>Respondents</b>	Unknown	Unknown	Over-burdened panelists; False Ps & Ns	Unknown
<b>Interviewers</b>	NA	NA	Not enough time for proper training	NA
<b>Data Mode</b>	Unknown	Unknown	Online & phone; no mode effects testing	Unknown
<b>Processing</b>	Unknown	Unknown	Seek highly skilled/reliable people	Unknown

# Possible Additional Comments

- From today's oral presentations
  - Fields et al.
  - Cornesse et al.
  - Phelps et al.
  - Moore et al.

# High-Level Take-Away Observations

- Driven by a “Something is Better than Nothing” mindset
- The importance of the Personnel who are planning and carrying out the research and their own personal commitment to quality
- The value of using both a TSQ mindset and a TSE mindset in planning these studies
  - Due to the timing imperative and constraints for planning and implementation
  - Due to limit funding, thus the need to make cost tradeoff decisions to minimize TSE
- Value of adding Experimentation as part of Adaptive/Responsive Design efforts
- Given the “Panel” nature of the various samples, why not study NR biases using previous data known about each sampled case, which includes the nonresponding sampled cases; e.g., in the USA, use most recent ACS and/or Census data (even if only at the block group level)
- Maximize the Leveraging of other auxiliary frame data to help with (1) Understanding Noncoverage, (2) Sampling Design, (3) Reducing Nonresponse, (4) Imputing Missing Values, and (5) Weighting



**Thank You!**

***pjlavrakas@comcast.net***