Evaluating Mode Effects for an In-Person Panel Survey that Transitioned to a Mixed Web Mode

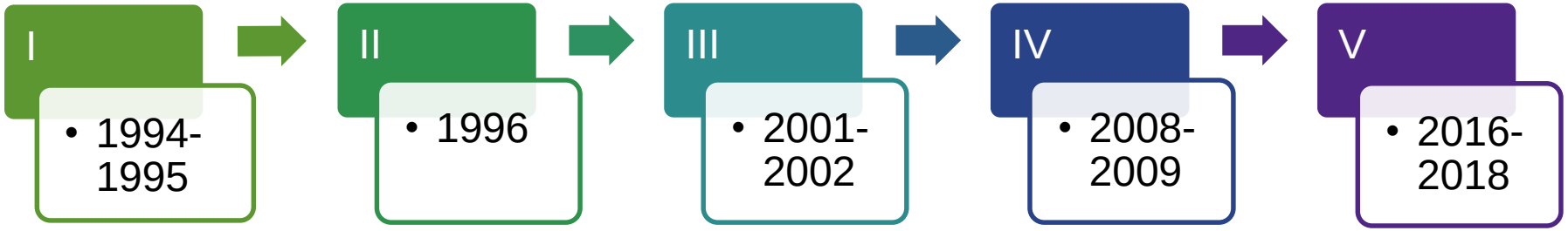
Paul P. Biemer
RTI International and the University of North Carolina

International Total Survey Error Workshop
Bergamo, Italy
June 11, 2019

National Longitudinal Study of Adolescent to Adult Health (AddHealth): Wave V

Purpose: To gather longitudinal data at a national level on health and health behaviors of adolescents and young adults.

- Conducted by UNC Carolina Population Center
- Now in its 5th wave to provide a 24-year longitudinal dataset (1994-2018)
- First administered in 1994 to random sample of students in grades 7–12
- Participants are now 32 to 42 years old
- Sample size = 19,831 (eligible)
- Response rates ~80% per wave
- Face to face interviewing through Wave IV



Goals and Challenges for Wave V

Goals

- Transition from in-person interviewing to web/paper mixed mode
  - Two-phase design with in-person nonresponse followup (NRFU) in 2nd phase
- Reduce costs and maintain data quality

Challenges due to mode change

- Control nonresponse bias
  - Respondents accustomed to in-person mode used in Waves I, II, III, IV
- Collect sufficient information from respondents
  - 90 minute in-person interview in Waves I - IV
  - Now conducted by 50 minute web/paper survey
- Control measurement errors
  - Minimize mode effects in cross-sectional and longitudinal estimates
  - Evaluate the effects of the transition on data quality

Data Collection was Spread Over 4 Subsamples and Three Years

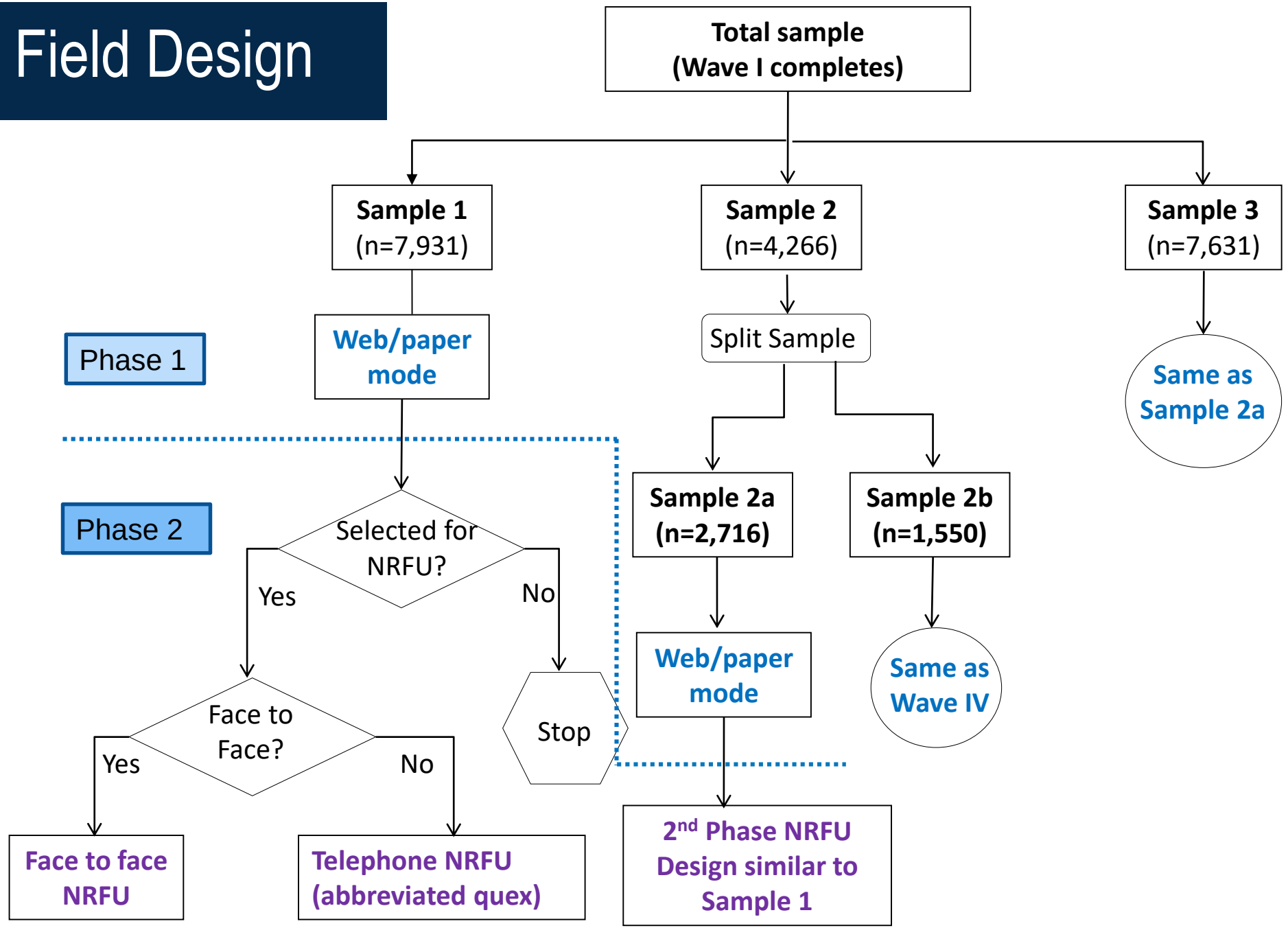
Mixed Mode Samples (Web/Paper with In-Person NRFU)

- **Sample 1 (2016-2017)**
  - Experiment with mode, questionnaire design and incentives in 1st data collection year (*n* = 7,931)
- **Sample 2a (2017-2018)**
  - Additional experiments (incentives/prenotice letters) (*n* = 2,716)
- **Sample 3 (late 2017-2018)**
  - Ultimately combined with Sample 2a (n = 7,631)

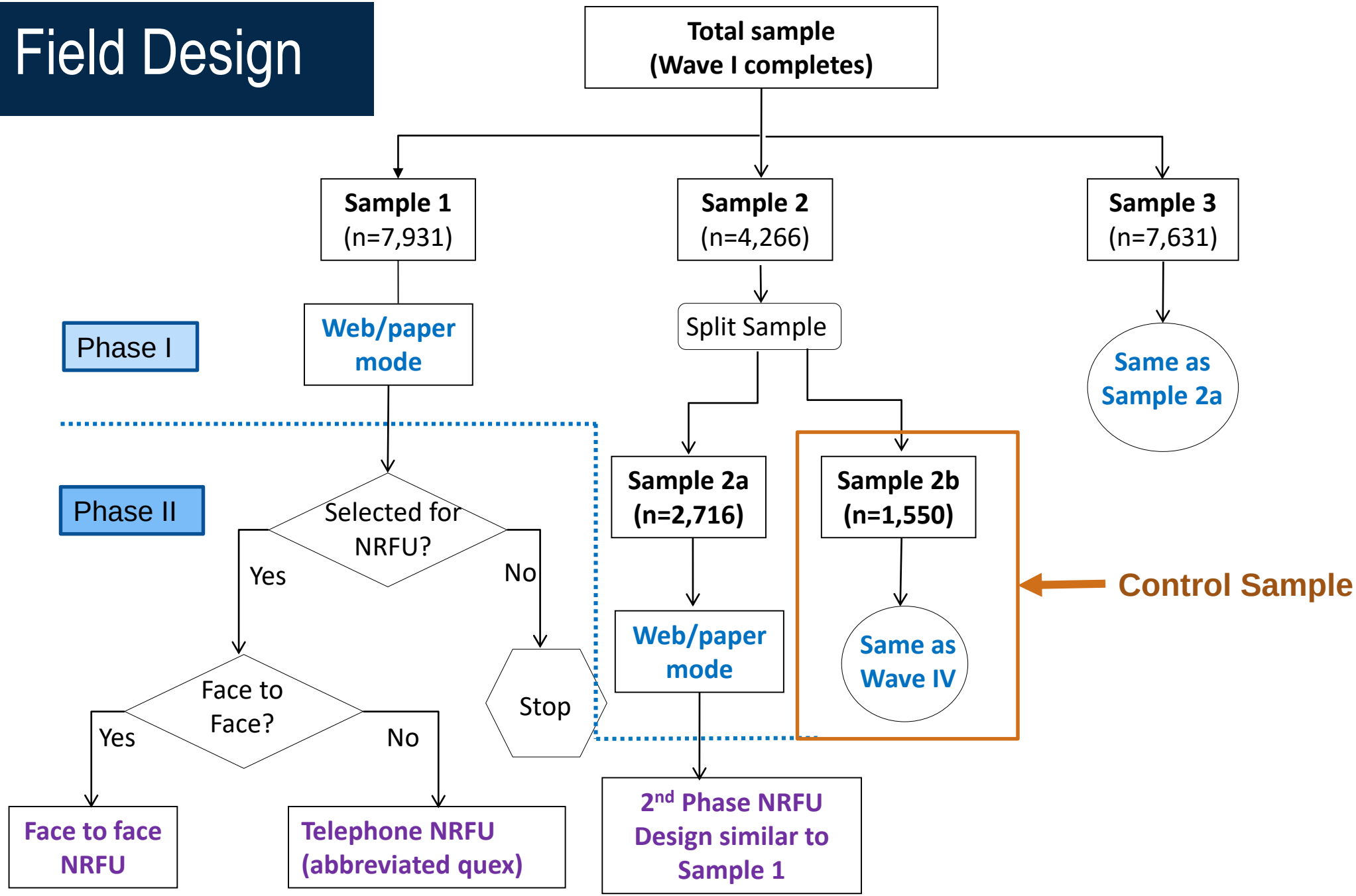
Control Sample

- **Sample 2b (2017-2018)**
  - Control sample replicating in-person design used in prior waves (*n* = 1,550)

Wave V – Field Design



Wave V – Field Design



Definitions and Notation

- MM = mixed mode (web/paper/in-person NRFU)
- CM = control mode (in-person)
- Samples 1, 2a and 3 are interviewed by the MM
- Sample 2b is interviewed by the CM
- Differential mode effect (DME) is defined

$$\text{DME} = \bar{y}_{\text{MM}} - \bar{y}_{\text{CM}}$$

mean of Samples 1, 2a and 3

mean of Sample 2b

- DME reflects both nonresponse (NR) error and measurement error (ME)

$$\text{E(DME)} = (B_{NR}^{MM} - B_{NR}^{CM}) + (B_{ME}^{MM} - B_{ME}^{CM})$$

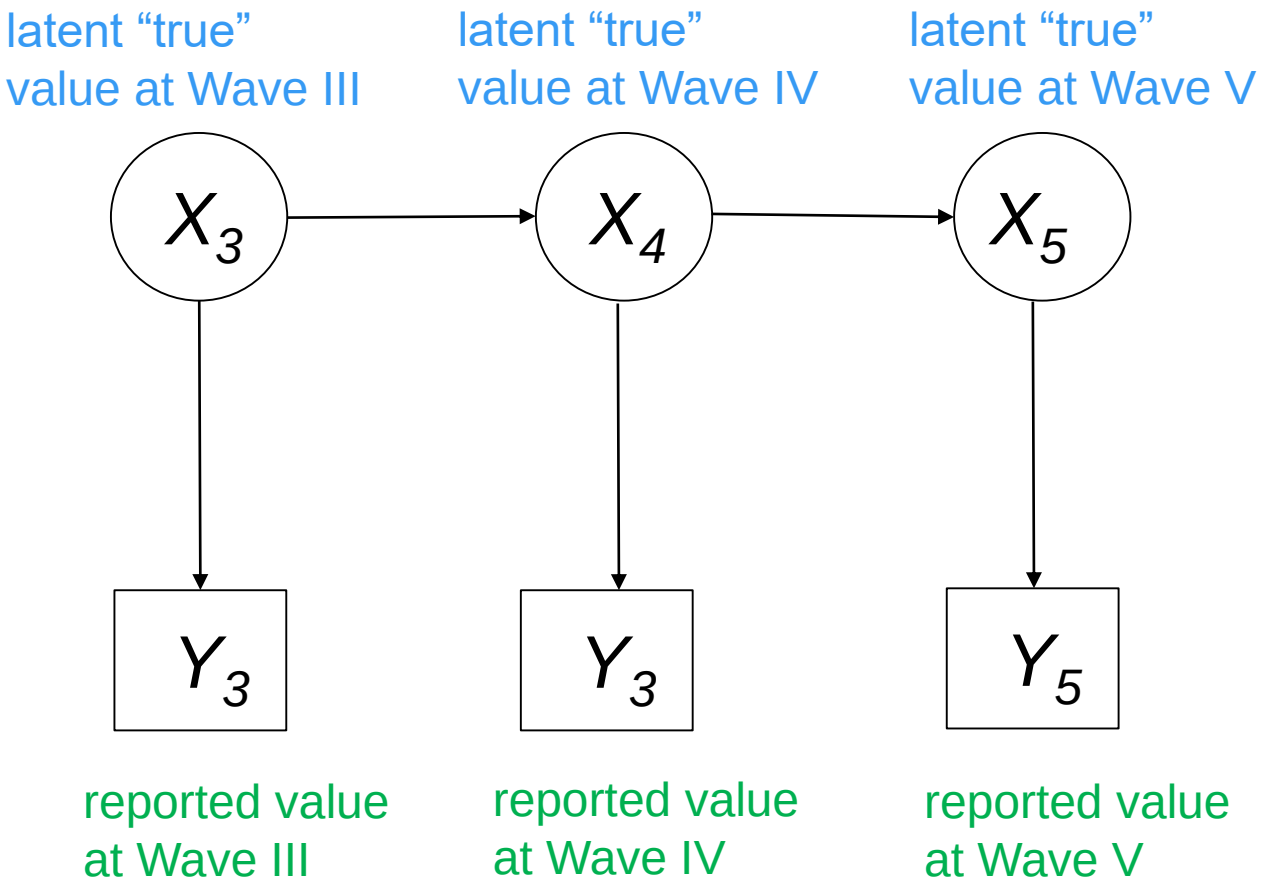
- Both weighted and unweighted analysis will be explored

Two-Part Strategy for Estimating Mode Effects

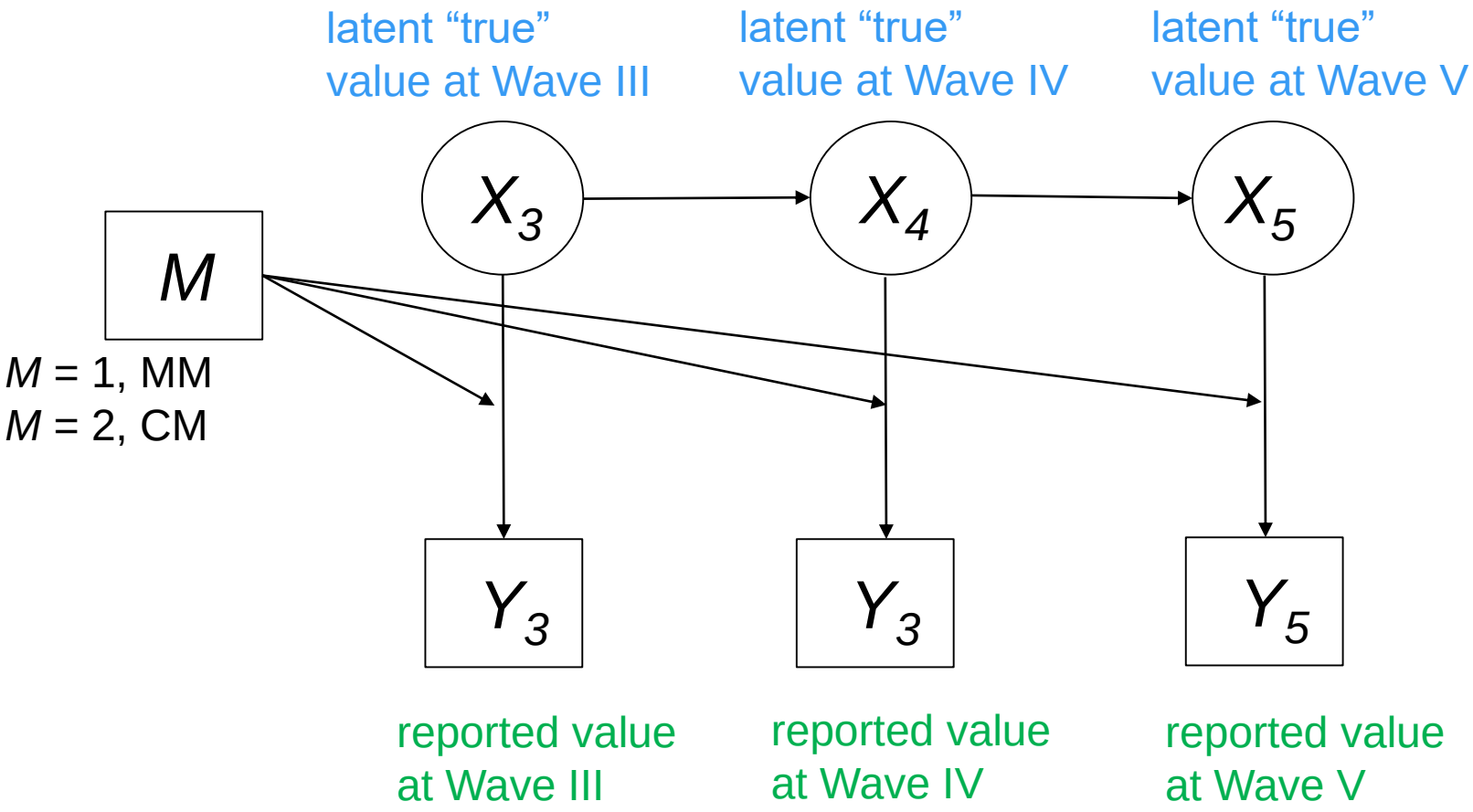
- Part 1 – Estimation of differential mode effects (DMEs)
  - Purpose: To identify questionnaire items having the larges DMEs
  - Restricted to items that are common across three most recent waves: Waves III, IV and V

- Part 2 – Estimation of Mode Bias
  - Purpose: For the items with largest DMEs, separately evaluate the bias for the mixed mode (MM), the control (in-person) mode (CM) and the combined sample
  - Markov Latent Class Analysis is used to separately estimate measurement bias for MM and CM

The Standard Markov Latent Class Model



Modification for Estimating Mode Effects



Parameters of the Basic Model for Binary Variables

- **Transition Probabilities**

$$\Pr(X_3 = x_3)$$

$$\Pr(X_4 = x_4 \mid X_3 = x_3)$$

$$\Pr(X_5 = x_5 \mid X_4 = x_4)$$

- **Error Probabilities**

Constrained

$$\Pr(Y_3 = y_3 \mid X_3 = x_3, M = m) =$$

$$\Pr(Y_4 = y_4 \mid X_4 = x_4, M = m) =$$

$$\Pr(Y_5 = y_5 \mid X_5 = x_5, M = 2) \text{ for } m = 1, 2$$

Unconstrained

$$\Pr(Y_5 = y_5 \mid X_5 = x_5, M = 1)$$

Parameters of the Basic Model for Binary Variables

- **Transition Probabilities**

$$\Pr(X_3 = x_3)$$

$$\Pr(X_4 = x_4 \mid X_3 = x_3)$$

$$\Pr(X_5 = x_5 \mid X_4 = x_4)$$

- **Error Probabilities**

Constrained

$$\Pr(Y_3 = y_3 \mid X_3 = x_3, M = m) =$$

$$\Pr(Y_4 = y_4 \mid X_4 = x_4, M = m) =$$

$$\Pr(Y_5 = y_5 \mid X_5 = x_5, M = 2) \text{ for } m = 1, 2$$

Unconstrained

$$\Pr(Y_5 = y_5 \mid X_5 = x_5, M = 1)$$

- **Model Assumptions/Constraints**
  - *Transitions do not depend upon mode, M*

- *Number of parameters = 5*

- **Model Assumptions/Constraints**
  - *MM (i.e., M=1) only affects Wave V*
  - *In-person error probs are equal across waves*
  - *Errors are independent across waves*

- *Number of parameters = 4*

Number of cells in data matrix:	16
Number of parameters:	9+2
Model degrees of freedom:	5

LEM Model

```
*LEM is a free software for LCA
lat 3
man 4
dim  2  2  2   2  2  2  2
lab  M3 M4 M5  S Y3 Y4 Y5
mod M3
M4|M3 {M3.M4}
M5|M4 {M4.M5}
Y3|M3.S eq2
Y4|M4.S eq2
Y5|M5.S eq2
rec 16
rco
sta Y3|M3.S [.9 .1 .9 .1 .1 .9 .1
.9]
des [1 0 1 0 2 0 2 0
     1 0 1 0 2 0 2 0
     1 0 3 0 2 0 4 0]
* Data goes here
```

Illustration for Binary Variables

M=1 (MM Sample)

	$Y_5 = 1$		$Y_5 = 2$	
	$Y_3=1$	$Y_3=2$	$Y_3=1$	$Y_3=2$
$Y_4=1$	10,358	154	247	273
$Y_4=2$	387	262	117	6,480

M=2 (CM Sample)

	$Y_5 = 1$		$Y_5 = 2$	
	$Y_3=1$	$Y_3=2$	$Y_3=1$	$Y_3=2$
$Y_4=1$	888	23	31	33
$Y_4=2$	43	32	20	560

Illustration for Binary Variables (unweighted analysis)

M=1 (MM Sample)

	$Y_5 = 1$		$Y_5=2$	
	$Y_3=1$	$Y_3=2$	$Y_3=1$	$Y_3=2$
$Y_4=1$	10,358	154	247	273
$Y_4=2$	387	262	117	6,480

M=2 (CM Sample)

	$Y_5 = 1$		$Y_5=2$	
	$Y_3=1$	$Y_3=2$	$Y_3=1$	$Y_3=2$
$Y_4=1$	888	23	31	33
$Y_4=2$	43	32	20	560

Mixed Mode (Web/Paper/NRFU)	
False Positives	0.0445
False Negatives	0.0334
True proportion (est.)	0.60
RelBias in the sample prop.	1.77%
Control Mode (In-Person)	
False Positives	0.0295
False Negatives	0.0247
True proportion (est.)	0.60
RelBias in the sample prop.	0.82%

Extensions to be Explored in the Analysis

- Compare unweighted, selection weighted and fully weighted estimates
- Add grouping variables to the model like age, race, sex to reduce heterogeneity as well as to investigate sources/causes of bias
- Use full information maximum likelihood (FIML) to compensate for nonresponse as well as to evaluate nonresponse bias
- In addition to error rates and relative bias, estimate and compare other quality metrics such as:
  - Project/compare MSEs of the weighted, full sample (1, 2a, 2b and 3) under the MM protocol and CM
  - Evaluate the effects of mode on longitudinal change estimates

Some Questions for Discussion

- Suggestions for improving the general approach
- Other important research questions to explore in the analysis
- Potential issues with the MLCA approach and how they can be addressed