TOTAL SURVEY ERROR WITH BIOMEASURES

BY BRAD EDWARDS, ALEXANDRU CERNAT, COLTER MITCHELL, AND JOSEPH SAKSHAUG

PRESENTED AT THE INTERNATIONAL TOTAL SURVEY ERROR WORKSHOP, WASHINGTON, DC

SEPTEMBER 18, 2024

AGENDA

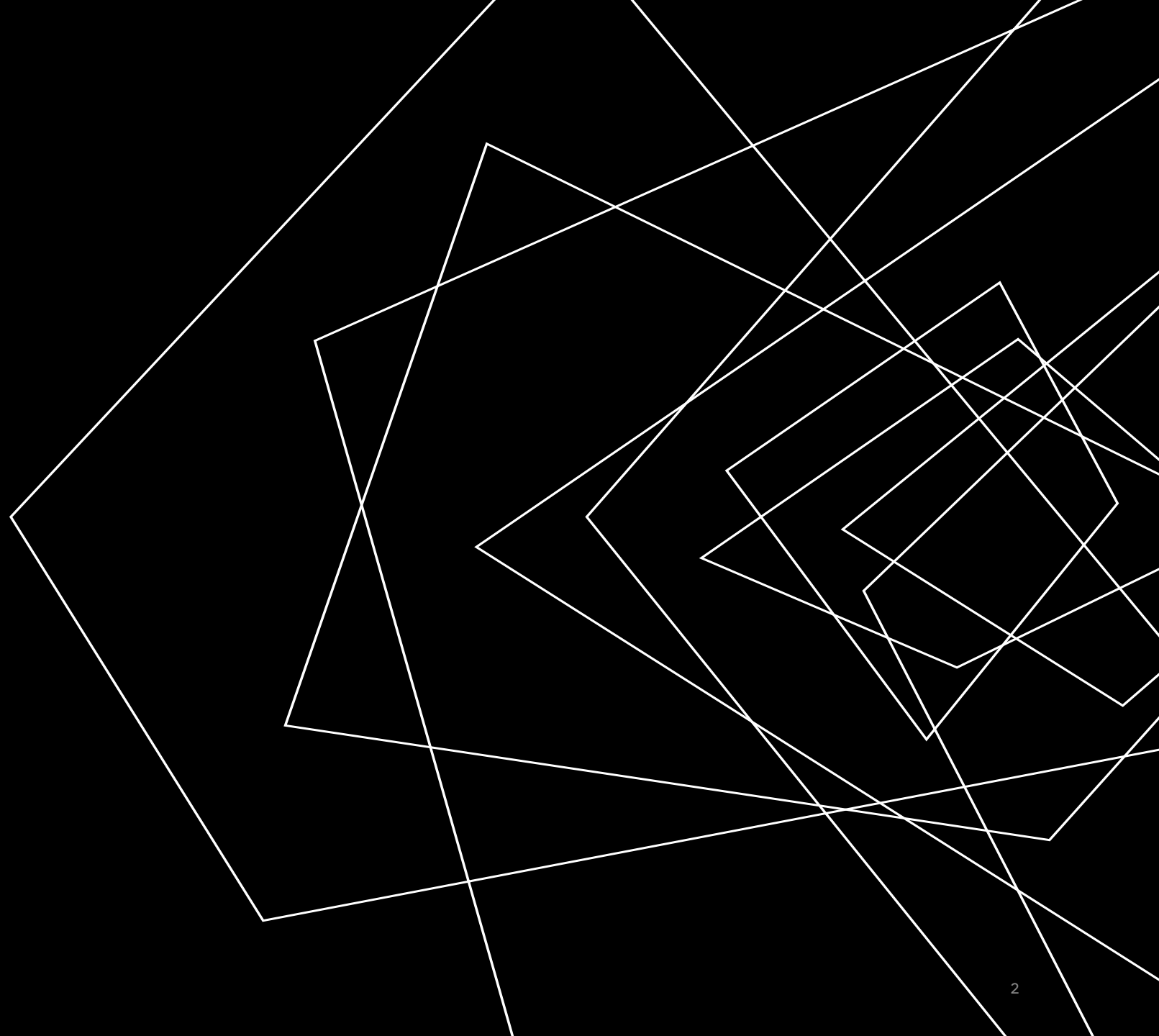
Overview

Errors of Representation

Errors of Measurement

Trade-offs

Conclusions



OVERVIEW

DYNAMOMETER FOR MEASURING GRIP STRENGTH, NATIONAL HEALTH AND AGING TRENDS STUDY (NHATS)

WHAT ARE BIOMEASURES?

- Biological, physical, mental
- Functioning of an organism and its systems
- Focus on individual, not environment
- Some causal mechanisms, some predictive indicators
- Organic vs designed
- Designed in clinical settings with narrow populations

TYPOLOGY

Body and brain

- Anthropomorphic, physiological, cognitive function, molecular, saliva, blood, hair, urine, tissue

Other dimensions

- Body/brain function
- Collection: instrument/method/who/where/timing
- Errors and cost
- Potential for future use

Blood Spot Kit, NHATS

WHY ADAPT TSE TO BIOMEASURES?

- Increasing value and use of biomeasures in social surveys
- New error sources
- Adapted from clinical settings
- Different actors
- Logistics and Tools

ERRORS OF REPRESENTATION

FROM TARGET POPULATION TO SAMPLING FRAME (COVERAGE ERROR)

- Probability sample that starts with survey, followed by biomeasures: coverage error similar to other surveys
- But maybe additional error if likelihood of people being in the frame related to biomeasures
- Nonprobability sample: difficult to separate coverage error from nonresponse leading to selection bias
- Defining population of interest is important:
  for some biomeasures some people have no chance of being selected

FROM SAMPLING FRAME TO SAMPLE (SAMPLING ERROR)

- Probability sample selection process: sampling error similar to other surveys
- Unless sampling procedure influenced by biomeasures (e.g., health or bio data used in stratification or clustering)
- Nonprobability sample selection process: trade-off between sampling error and systematic selection bias due to coverage and nonresponse (UK Biobank, All of Us)
  - Under-representation of key subgroups
  - Reason for increasing popularity of combining probability samples with biomeasures

FROM SAMPLE TO RESPONDENTS (NONRESPONSE OR CONSENT ERROR)

Besides the usual survey nonresponse mechanisms, biomeasure nonresponse can be related to ...

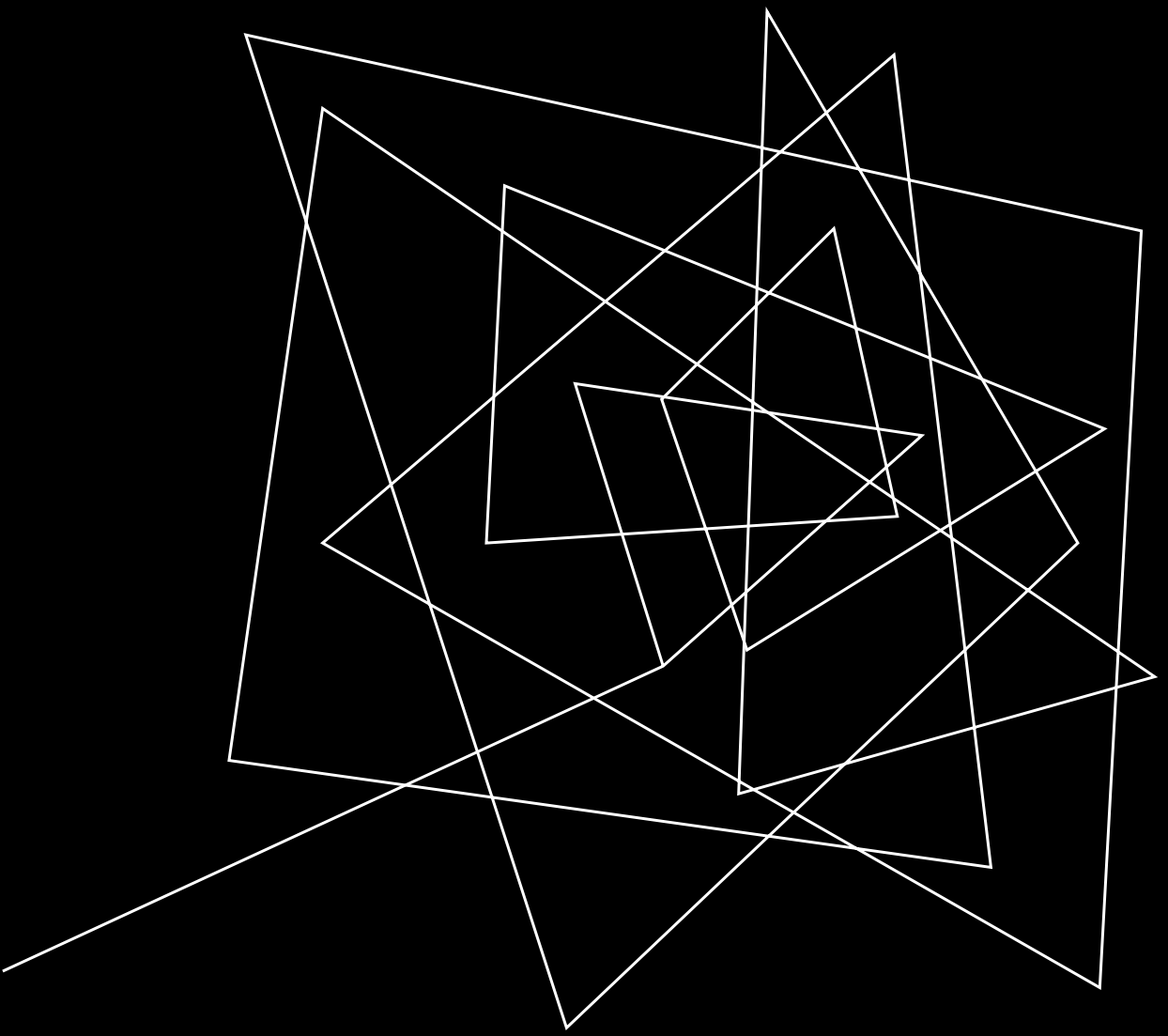
1. Biomeasure collection mode
2. Biomeasure collection in stages over time
3. Consent requirements for biomeasure collection
4. Consent requirements for biomeasure archiving or sharing
5. Impossibility of biomeasure collection for certain respondents
6. Engagement and collection skills of collector

FROM RESPONDENTS TO POST-SURVEY ADJUSTMENT (ADJUSTMENT ERROR)

- Weights often not present or misused

Strategies to correct for nonresponse

- Separate by stage, create conditional models for each and develop weight for each mechanism (Cernat et al, 2021)
- Heckman selection models (Clarke & Houle, 2014)
- Multiple imputation (Hannemann et al, 2020)



ERRORS OF MEASUREMENT

FROM CONSTRUCT TO MEASUREMENT (VALIDITY)

Validity highly dependent on biomeasure

- Match to clinical gold standard, or
- Predictive capacity of health or behavioral outcome

If clinical comparison unavailable, criterion validity

- Examine correlation with known covariates

Novel measures or those with large number of potential algorithms (e.g., genetics): few validity analyses exist

Tension: Is deviation from gold standard a sign of lower validity or more real-life measurement?

FROM MEASUREMENT TO RESPONSE (MEASUREMENT ERROR)

Overall measurement quality is good, and well-researched

Data collector effects can be moderate to large, depending on measure

- SHARE: variance partition coefficients .05-.28 (peak air flow) for physical functioning measures; cross-national differences (Waldmann et al, 2023)
- Consent varies by experience, but no consistent trend in biomeasures across waves

Respondent effects with self-administration

Lab effects can be much larger

Olfactory Test

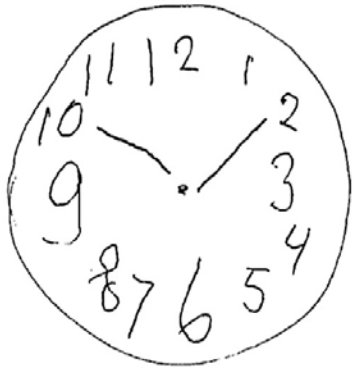
National Social Life, Health & Aging Project (NSHAP)

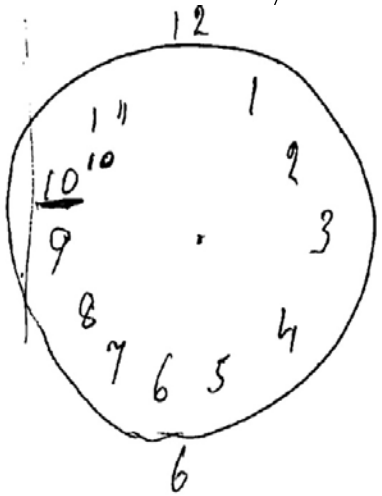
FROM RESPONSE TO EDITED DATA (PROCESSING ERROR)

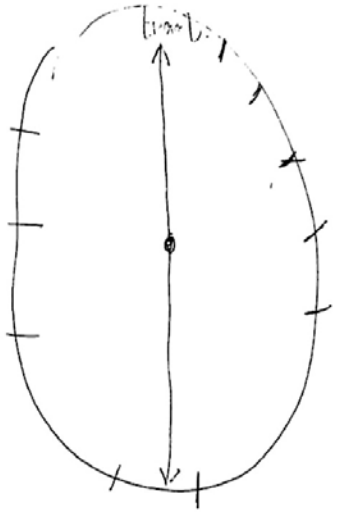
Unique to biomeasures

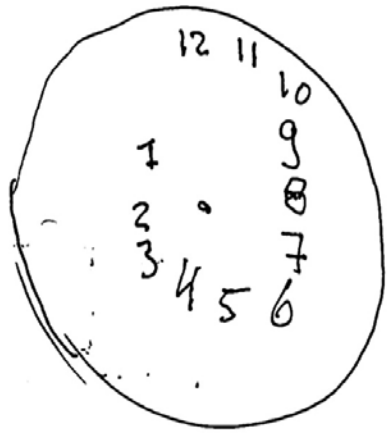
- Shipping (time, temperature, humidity)
- Receipt, batch effects
- Processing, clinical cut-offs (?), assays
- QC assessments, choices; clinical cut-offs might be different
- Few benchmarks for correcting scores (e.g., time of day)

Unstructured data

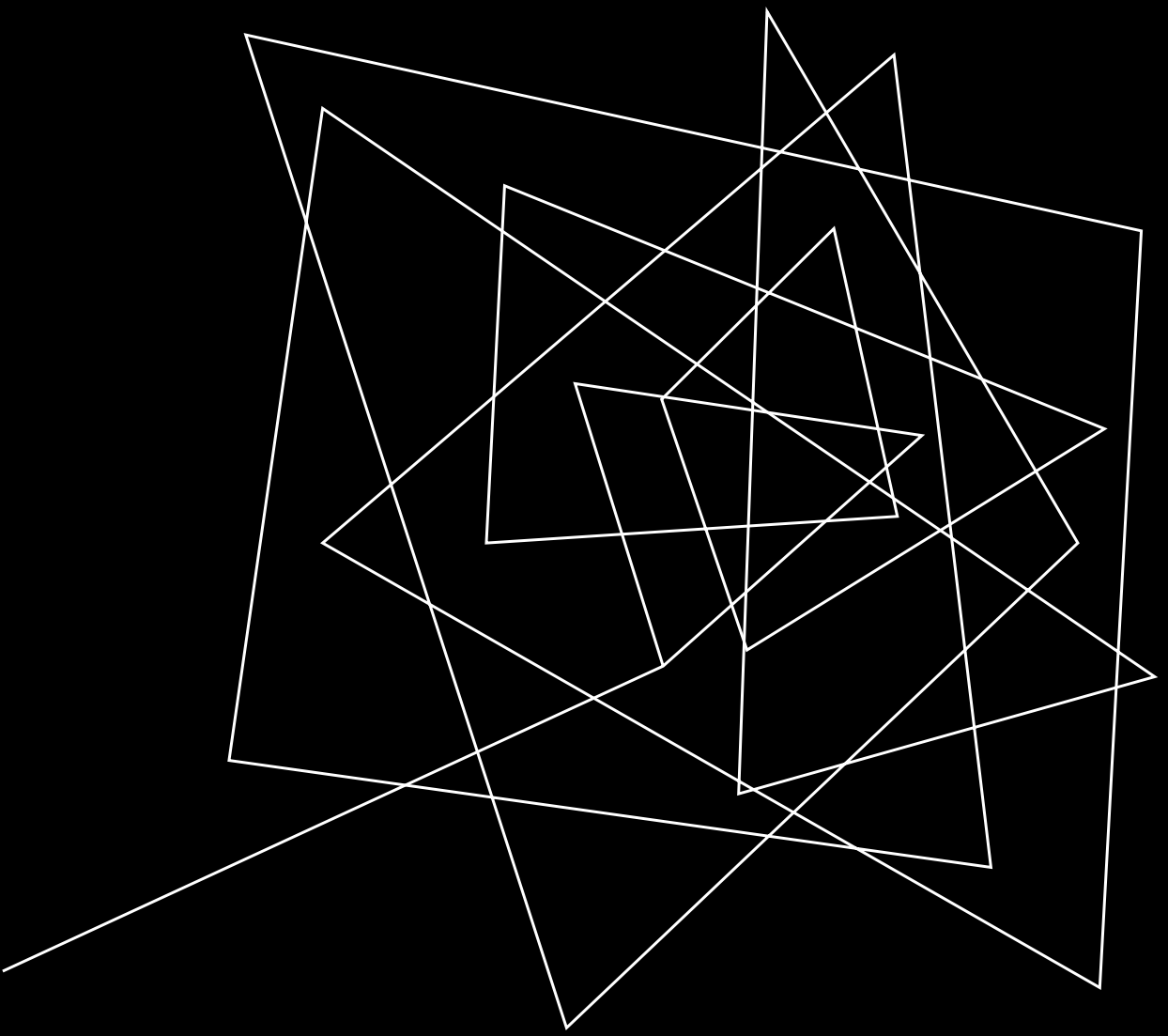








Clock Drawing Test, NHATS



TRADE-OFFS

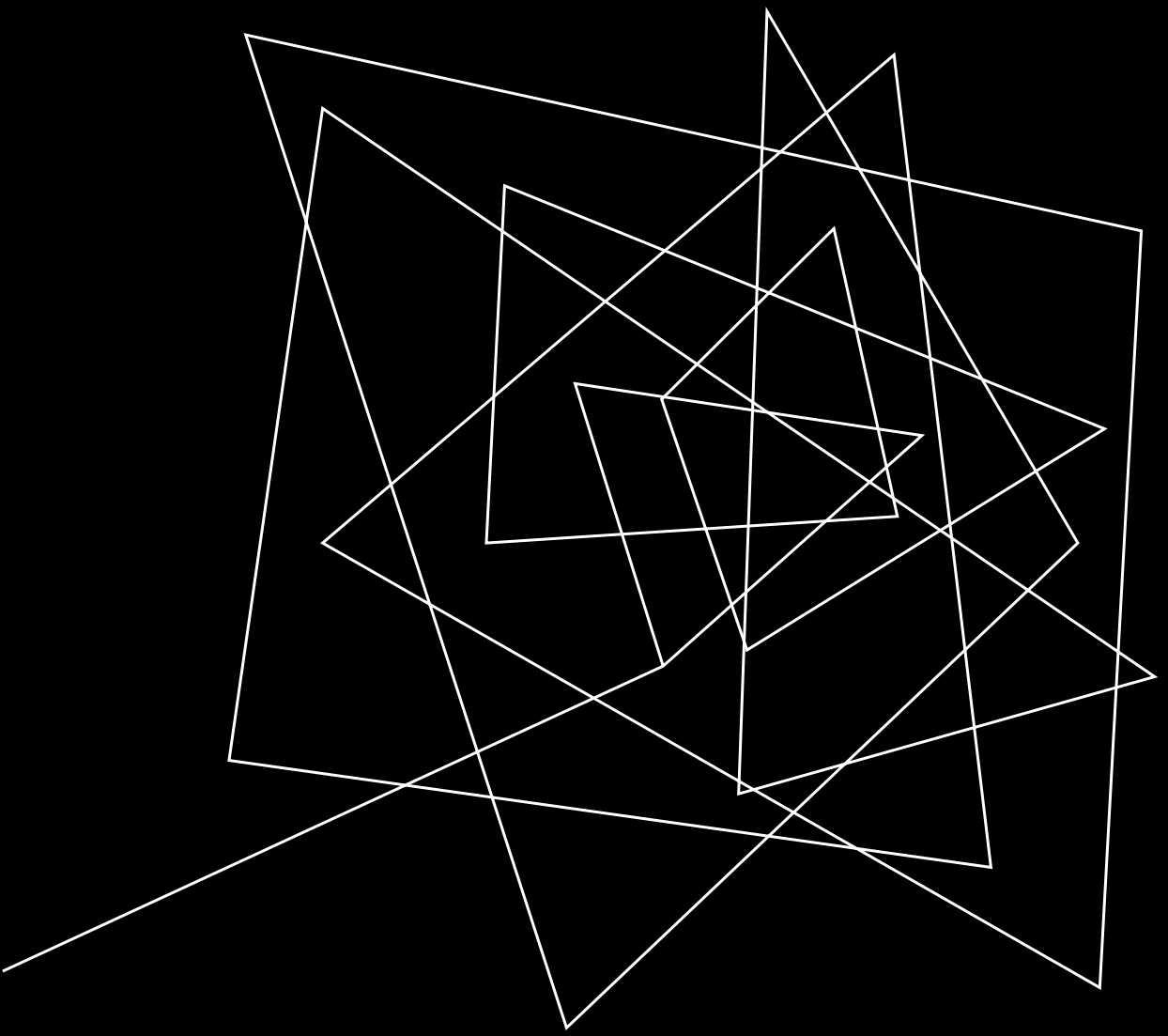
DESIGN TRADE-OFFS

- Sampling error (precision) and selection bias
- Nonresponse and measurement
- Internal and external validity
- Measurement (quality) and cost

Shoebox Audiometry Test, NHATS

Strategies

1. Minimize total error of the survey (but challenging if no gold standard) by focusing on main analyses and research questions
2. Minimize errors most relevant to the research questions, accounting for unintended consequences of other TSE dimensions



CONCLUSIONS

TAKEAWAYS

Importance of probability samples to make biomeasure research reflect the population better

Developing best practices and sharing knowledge essential to advance bio-social surveys, but little has been published

More research needed

- Effects of biomeasures on nonresponse by mode
- Interviewer and nurse effects, mitigation approaches
- Gap in biomeasure adaptation for non-clinical settings

Consistent use of terminology and frameworks like TSE crucial for addressing biomeasure collection challenges

THANK YOU

Brad Edwards

bradedwards@westat.com