WELCOME TO:

Nonprobability Online Samples: Promises & Pitfalls

AAPOR offers webinar sponsorship opportunities to organizations that would like to support AAPOR's online education program.

In addition, AAPOR has an institutional subscription for greater access to the education webinar recording library.

For more information, contact;

Lailah Johnson, AAPOR Program Manager

ljohnson@aapor.org

Join us for the Next Webinar in the 2023 Series:

Communicating Survey Methods & Results to the Press

Presenters: Emily Swanson & Trevor Tompson

November 13, 2023

1:00 PM – 2:00 PM ET

Please enter your questions in the Q&A box at the bottom of your screen.

Please complete the webinar survey immediately following the session.

TODAY'S PRESENTERS

Carina Cornesse

Olga Maslovskaya

Nonprobability Online Survey Samples

Promises and Pitfalls

Dr. Carina Cornesse
Dr. Olga Maslovskaya

Agenda

- Definition & delineation
- Sampling & recruitment
- Evaluation frameworks
- Empirical evidence
- Adjustment & estimation
- Concluding remarks

Definition & delineation

Definitions

- Nonprobability online samples result from non-random sampling and recruitment procedures on the internet

- Probability samples rely on sampling frames (e.g. list of all email addresses of the students at a university)
  - ➢ Using a random selection procedure, individuals are drawn from the list and invited to an online study
  - ➢ Error can be assessed & controlled along the Total Survey Error framework

- ➢ Nonprobability samples rely on volunteers self-selecting into the study (e.g. post advertisement on university websites)

Definitions

- Problem in terms of representation
  - Some target population members may have unequal chances of self-selecting into the sample (e.g. some people rarely log into their social media, others are on there constantly)
  - For probability samples, unequal selection probabilities can be calculated and controlled for using design weights
  - Members of the target population may have no chance of inclusion at all (e.g. if they do not use social media)

- Problem in terms of verification
  - Some actors may have an interest in a particular outcome
    - Use survey bots to manipulate a study
    - Advertise the study among partisan groups
    - Pay people to volunteer

➢Researchers / fieldwork agencies have limited insights & control into the data generating process

What are nonprobability online samples?

Probability sampling	**Nonprobability sampling**
Relies on proven mathematical principles underlying statistical sampling theory	Relies on modeling assumptions that need empirical proof, which may never be available
Ensures that every unit has a known non-zero chance of being selected	Units have unknown chances of being selected; chances may even be zero

What are nonprobability online samples?

- First nonprobability sampling approach:

The "Representative Method"

(Anders Kiaer, 1895)

- First probability sampling approach:

Sampling "at random"

(Arthur L. Bowley, 1906)

- Kruskal, W., & Mosteller, F. (1980). Representative sampling, IV: The history of the concept in statistics, 1895-1939. *International Statistical Review*, 48(2), 169-195. https://doi.org/10.2307/1403151

What are nonprobability online samples used for?

- Surveys
- Special populations (e.g., hard-to-sample populations; see AAPOR webinar by Mariel McKone Leonard)
- Experiments
- Passive measurement (e.g., web tracking)
- Combinations (data integrative approaches)

Sampling & Recruitment

Sampling & recruitment: Typical procedures

- Sampling (i.e., how to select participants)
  - Convenience
  - Purposive
  - Quota

- Recruitment (i.e., where do the participants come from)
  - Social Media
  - River Sampling
  - Snowball
  - Respondent-driven sampling
  - Online panels & consolidators

Sampling & recruitment: Typical procedures

- Sampling (i.e., how to select participants)
  - **Convenience** – conveniently available to researcher, easy to recruit
  - **Purposive** - researchers rely on their own judgment when choosing members of the population to participate in their surveys – only few individuals or units are usually chosen (e.g., expert panel)
  - **Quota** – respondents are selected according to certain categories (age, gender, region) to mimic population distribution widely used in market research and opinion polling

Quota Sampling 1

- Draw units purposively to match known population characteristics: Mimic population characteristics from census data - This may correct biases or not.
- Approach is widely used in market research and opinion polling
- E.g., you know population totals for age, sex, and region from the census
- E.g., If 10% of population is unemployed, recruit a sample with 10% unemployed persons
- These can be used to form ‘quotas’
- You select your sample until the quotas are filled
- Your sample will match the population on the quota variables

Quota Sampling 2

- Quota sampling does not select population elements at random
- This means we cannot use mathematical properties of sampling distributions for inference from sample to population
- Instead, we must rely on ‘model-based’ inference (design-based inference relies on the known probability of selection in sampling, and model-based inference is based on models of human recruitment behavior and the social context within which sampling is conducted)
- If the variables used for the quotas are strongly predictive of the survey variable, the estimate will be accurate
- Much stronger assumption than for random sampling

Sampling & recruitment: Typical procedures

- Recruitment (i.e., where do the participants come from)
  - Social Media (targeted) advertisement
  - River sampling
  - Snowball sampling
  - Respondent-driven sampling
  - Online panels & consolidators

Targeted Advertisement / Social media (SM)

- Platforms such as FB, Insta, Twitter have millions of users and represent potential respondents (57.6% of global population is represented on at least one SM platform)
- They have powerful targeting capabilities that can be used to recruit hard-to-reach populations
- SM targeting tools allow researchers to track and reach users with specific demographic characteristics and interests based on their behaviour on SM sites and third-party websites the users interact through their SM accounts
- Two main strategies: paid – higher number of recruited individuals (SM platforms provide a sophisticated advertising targeting system that allows specific audiences to be identified based on multiple parameters, parameters can be used to customise ads to reach specific populations: can be placed in newsfeed, at the edge of the screen, between stories etc.) and unpaid – less effective (invitations through groups, use profile pages in SM networks, videos could be created to invite people)
- Main issue: no information about those who actively decides not to participate - selection bias
- Advantage: wide reach and ability to reach audiences when compared to offline strategies
- Effective at reaching rare populations

Zindel, Zaza (2023). Social Media Recruitment in Online Survey Research: A Systematic Literature Review. methods, data, analysis (mda), 17(2), 207 – 248.

Sampling & recruitment: Typical procedures

- **River sampling**
  - Intuition: People are fish swimming in the stream of the internet. You catch them while they are busy going about their swimming business (e.g. reading a news article). You ask them a couple of questions, unhook them from your survey rod, and release them back to the stream, usually never to see them again.
  - invite website visitors to immediate surveys while they are doing some other online activities (e.g., via pop-up windows, online banners; also called “real-time sampling”). All or random sample of site visitors (e.g., every 10th visitor will get pop-up window).
  - Respondent’s characteristics are often estimated, e.g. based on website usage behavior
  - Newer alternative: interactive survey feature embedded in other website content (e.g., news article on a politician includes a topic-related “surveytainment” question), which may lead to further questions & even panel recruitment → no more releasing the fish

Sampling & recruitment: Typical procedures

- **Snowball sampling** – use the respondents to contact new respondents: referral program ("refer a friend" campaigns, rare characteristics, hard-to-reach groups)

- **Respondent-driven sampling** - combines snowball sampling with *mathematical model* that weights the sample to compensate for the fact that the sample was collected in a non-random way; link-tracing sampling and inference methods for studying hidden and hard-to-reach populations; respondents are selected from a social network of existing members of the sample (minority groups, hard-to-reach groups) and *not* from a sampling frame. Usually involves a dual-incentive system

Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, *44*(2), 174–199.

Online Panels

Online panels are

“[online] sample database[s] of potential respondents who declare that they will cooperate for future data collection if selected”

(International Organization for Standardization, 2012)

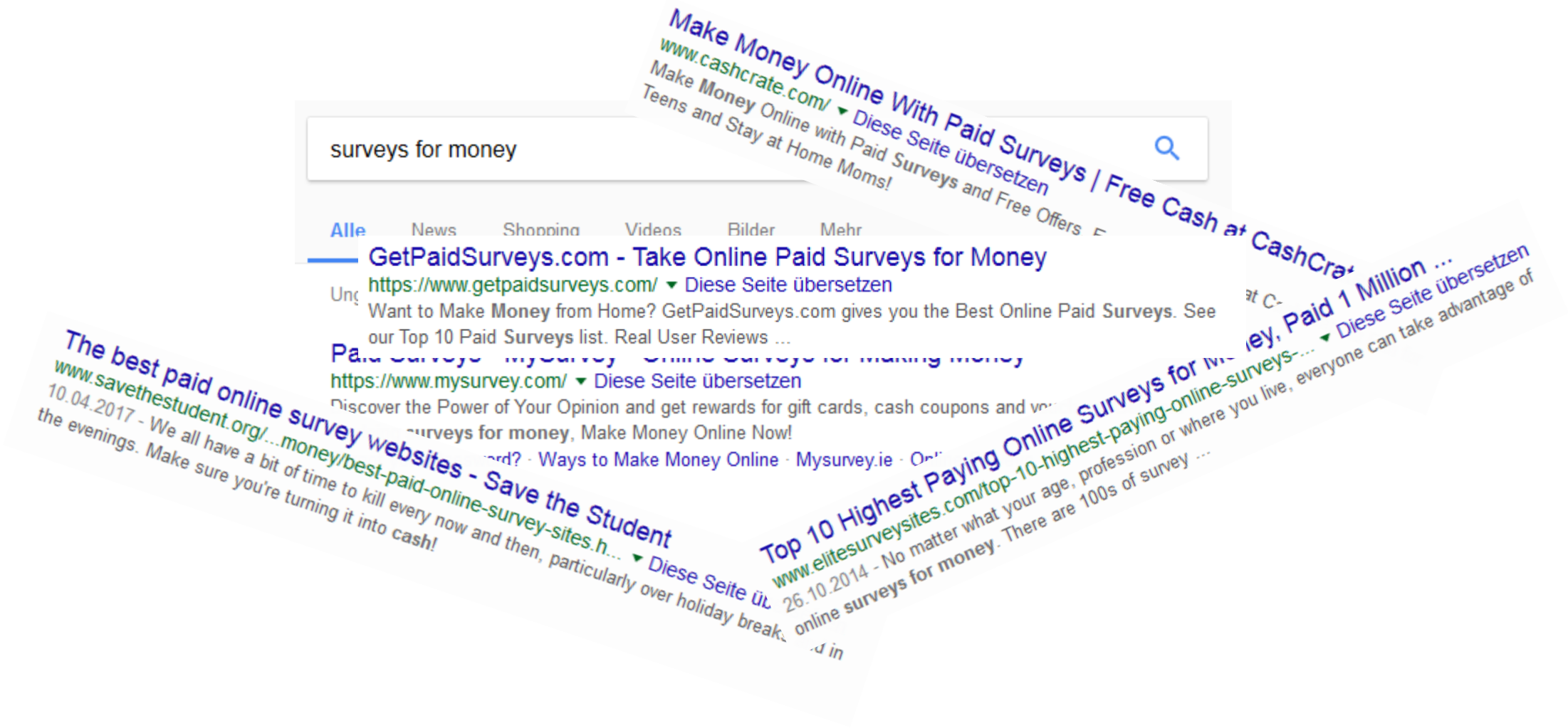
Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (2014). Online panel research: History, concepts, applications and a look at the future. In Callegaro et al. (Eds). Online Panel Research: A data quality perspective (pp. 1-22). Wiley.

Online Panels – Different Types

- Commercial vs. academic
- Online recruited vs. offline recruited
- Monetary vs. non-monetary (e.g., lottery win) incentives
- Regular (e.g., once a month) vs. irregular panel invitations
- …
- **Probability** (e.g. list of addresses, draw sample randomly, contact offline, recruit to panel) **vs. nonprobability sample (recruit volunteers on the internet)**

For more information on probability online panels, see AAPOR webinar by David Dutwin and Ipek Bilgen!

Nonprobability Advertisement

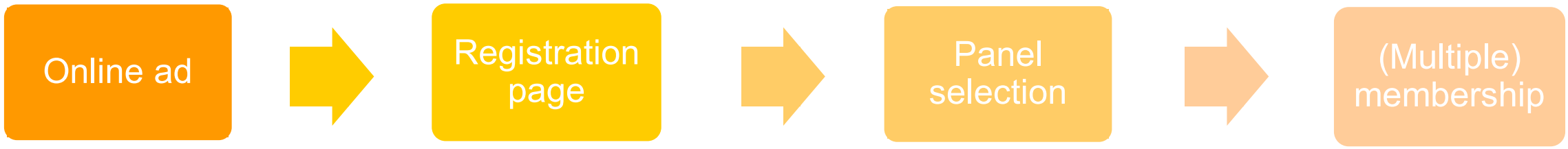


Nonprobability Online Panel Recruitment

- Banner ads on websites / social media
- Invitations via newsgroups / mailing lists
- Search engine ads
- Pop-up windows / “river sampling”
- **…**
- **Online panel consolidators** (collaborative effort in advertising)

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., & Lavrakas, P. J. (2014). Online panel research: History, concepts, applications and a look at the future. In Callegaro et al. (Eds). Online Panel Research: A data quality perspective (pp. 1-22). Wiley.

Nonprobability Panel Consolidators



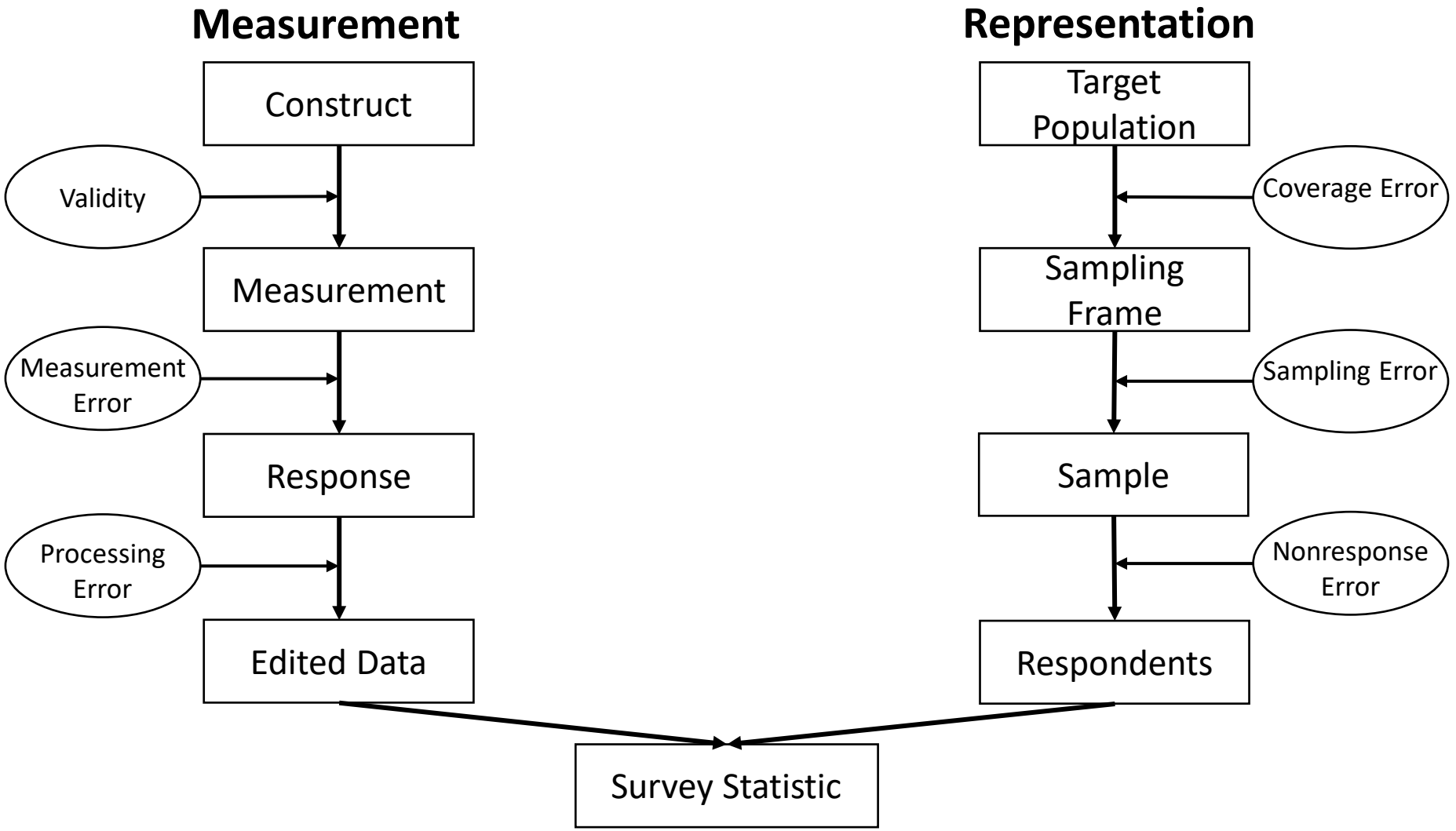
Evaluation frameworks

TSE adaptations

Fitness-for-purpose

Causal inference

Total Survey Error



Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, *74*(5), 849-879. https://doi.org/10.1093/poq/nfq065

Adaptations of the TSE

Unangst et al. (2020). A process for decomposing total survey error in probability and nonprobability surveys: A case study comparing health statistics in US Internet panels. *Journal of Survey Statistics and Methodology*, *8*(1), 62-88. https://doi.org/10.1093/jssam/smz040

- ***Qualitative assessment*** of methods used by the panel survey agency to mitigate error – can be done for both probability and nonprobability samples

- *Quantitative evaluation* of error/bias – can only be done for probability samples because it is a typical TSE error decomposition into coverage error and nonresponse error

Adaptations of the TSE

Unangst et al. (2020). A process for decomposing total survey error in probability and nonprobability surveys: A case study comparing health statistics in US Internet panels. *Journal of Survey Statistics and Methodology*, *8*(1), 62-88. https://doi.org/10.1093/jssam/smz040

Qualitative assessment (***checklists*** that are applied along the typical TSE error components on the representation side of the framework) of methods used by the panel survey agency to mitigate error:

- Coverage error
- Nonresponse error
- Sampling and adjustment error

Unangst et al. 2020, p.75

Table 4. Qualitative Assessment of Methods Used by the Internet Panels to Mitigate Error

Methods used to reduce error	Probability panels		Nonprobability panels		
	A	B	C	D	E
Coverage error					
Uses probability methods	✓	✓			
Covers offline population	✓	✓			
Refreshes on a rolling basis	✓	✓		✓	✓
Offers provisions for disability	✓	✓	✓		
Nonresponse error					
Oversamples groups less likely to respond	✓	✓	✓	✓	✓
Allows proxy reporting, by design	✓	✓	✓	✓	✓
Allows proxy reporting, ad-hoc	✓				
Sampling and adjustment errors					
Avoids routing	✓	✓			
Provides statistically valid weights	✓	✓			✓
Employs quota sampling			✓	✓	

Panels A and B (probability ones) are deemed the most likely to provide reliable estimates of disability among the panels evaluated as they employed the greatest breadth of strategies to reduce error (p. 74).

“Fit-For-Purpose” Assessments

- Case-by-case assessments:

“Will the data be able to meet the research objectives?”

- If a dataset allow to answer a research question, then you need to do an assessment whether the dataset you found is fit for the purpose of answering your specific research question.

Dever, J. A., Amaya, A., Srivastav, A., Lu, P.-J., Roycroft, J., Stanley, M., Stringer, M. C., Bostwick, M. G., Greby, S. M., Sanitbanez, T. A., & Williams, W. W. (2020). Fit for purpose in action: Design, implementation, and evaluation of the National Internet Flu Survey. *Journal of Survey Statistics and Methodology*. https://doi.org/10.1093/jssam/smz050

- Data do not GENERALLY have to be accurate and unbiased, they just have to serve the purpose they are intended for,
  - E.g., if I’m interested in attitudes on vaccination, it may not matter if my survey sample overrepresents married people, if being married or not has nothing to do with my variables of interest, i.e., on vaccination attitudes.

“Fit-For-Purpose” Assessments

- Case-by-case assessments:

“Will the data be able to meet the research objectives?”

Dever et al. (2020). Fit for purpose in action: Design, implementation, and evaluation of the National Internet Flu Survey. *Journal of Survey Statistics and Methodology*. https://doi.org/10.1093/jssam/smz050

“[…] there is no universal survey design that will fit all needs. Each team must define the components for their survey, along with their relative importance to meeting the research objectives within practical constraints. Essentially, they must specify the conditions that fit the purpose of their research.” (p. 452)

“Fit-For-Purpose” Assessments

- Case-by-case assessments:

“Will the data be able to meet the research objectives?”

Potential Criteria:

- Timeliness
- Accessibility
- Relevance
- Interpretability
- Accuracy
- Coherence

Statistics Canada (2017). *Statistics Canada's Quality Assurance Framework* (3rd ed.). Available at https://www150.statcan.gc.ca/n1/en/pub/12-586-x/12-586-x2017001-eng.pdf?_st=ZBVUvHOe

Causal Inference Framework

Another way to approach data quality in NP samples is to consider causal inference framework

Mercer et al. (2017) establish a link between causal inference and survey inference and provides insight into the conditions under which nonprobability surveys can be expected to provide estimates free of selection bias and proposes a framework that determine the level of selection bias in survey estimates.

Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: parallels between causal inference and survey inference. *Public Opinion Quarterly*, *81*(S1), 250-271.

Causal Inference Framework

Framework composed of **three elements** that determine the level of selection bias in survey estimates Mercer et al. (2017):

1. **Exchangeability**
    - Are all confounding variables known and measured for all sampled units?
    - E.g. if sample is younger than target population, age needs to be measured in the survey
    - When unobserved confounders are present, it is not possible to isolate the effect of the treatment from the effect of the confounder without additional assumptions.
2. **Positivity**
    - Does the sample include **all** of the necessary kinds of units in the target population, or are certain groups with distinct characteristics missing?
    - E.g. an online survey lacks non-internet users
    - It requires that all subjects have a positive probably of receiving treatment. If certain types of respondents receive only treatment or control, these two groups will have systematic differences that cannot be resolved.
    - Groups that are underrepresented but present can be weighted up but it is not possible to weight up groups that were not surveyed.
3. **Composition**
    - Does the sample distribution match the target population with respect to the confounding variables, or can it be adjusted to match?

Empirical evidence

Research Synthesis 2020

German Online Panel Comparison Study

Australian Sample Comparison Study

UK Understanding Society App Study Comparison

General picture

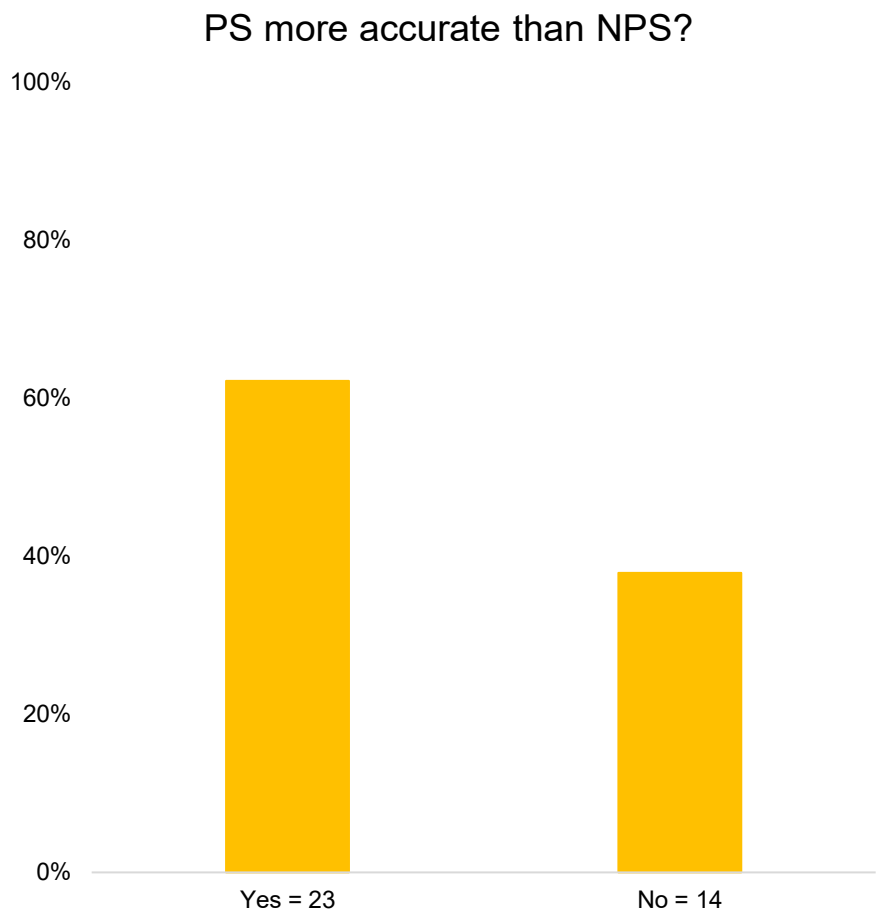
- Research shows that probability-based **surveys** more accurately reflect the general population than nonprobability **online samples** on univariate statistics

- Cornesse, C. et al. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology, 8*(1), 4-36.

- Cornesse, C., & Bosnjak, M. (2018). Is there an association between survey characteristics and representativeness? A meta-analysis. *Survey Research Methods*, 12(1), 1-13.

Research Synthesis

- 25 articles
- Across 9 countries
- Across 15 years
- Including 37 studies
- Comparing more than 100 samples to external benchmarks



Cornesse et al. (2020). A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *Journal of Survey Statistics and Methodology* 8(1), 4–36. https://doi.org/10.1093/jssam/smz041

- Nonprobability samples may perform better than usual in...
  - Bi- and multivariate analyses compared to univariate analyses
  - Studies that focus on election outcomes rather than other topics
  - Examples:
    - Dassonneville et al. (2020). The effects of survey mode and sampling in Belgian election studies: a comparison of a national probability face-to-face survey and a nonprobability Internet survey. *Acta Politica*, *55*(2), 175-198.https://doi.org/10.1057/s41269-018-0110-4
    - Kennedy et al. (2016). Evaluating online nonprobability surveys: Vendor choice matters; Widespread errors found for estimates based on blacks and Hispanics. Retrieved from http://assets.pewresearch.org/wp-content/uploads/sites/12/2016/04/Nonprobability-report-May-2016-FINAL.pdf
    - Ansolabehere, S., & Schaffner, B. (2014). Does survey mode still matter? Findings from a 2010 multi-mode comparison. *Political Analysis, 22*(3), 285-303. https://doi.org/10.1093/pan/mpt025

A Closer Look: Study I

Yeager et al. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75(4), 709-747. https://doi.org/10.1093/poq/nfr020

- 1 probability telephone survey sample
- 1 probability internet survey sample
- 7 nonprobability internet survey samples

➢ Compare each to a gold-standard benchmark (e.g. general social survey)

- Examined variables: Socio-demographics, health measures
- The bias is smallest in the probability telephone survey sample.
- Followed by the probability internet survey sample.
- The bias is highest in the nonprobability internet survey samples.

A Closer Look: Study II

MacInnis et al. (2018). The accuracy of measurements with probability and nonprobability survey samples: Replication and extension. *Public Opinion Quarterly*, 82(4), 707-744. https://doi.org/10.1093/poq/nfy038

- 1 probability internet survey sample
- 6 nonprobability internet survey samples
- 2 internet survey samples that combine probability and nonprobability approaches

- The bias is smallest in the probability internet survey sample.
- Followed by the internet survey samples that combine probability and nonprobability approaches.
- The bias is highest in the nonprobability internet survey samples.

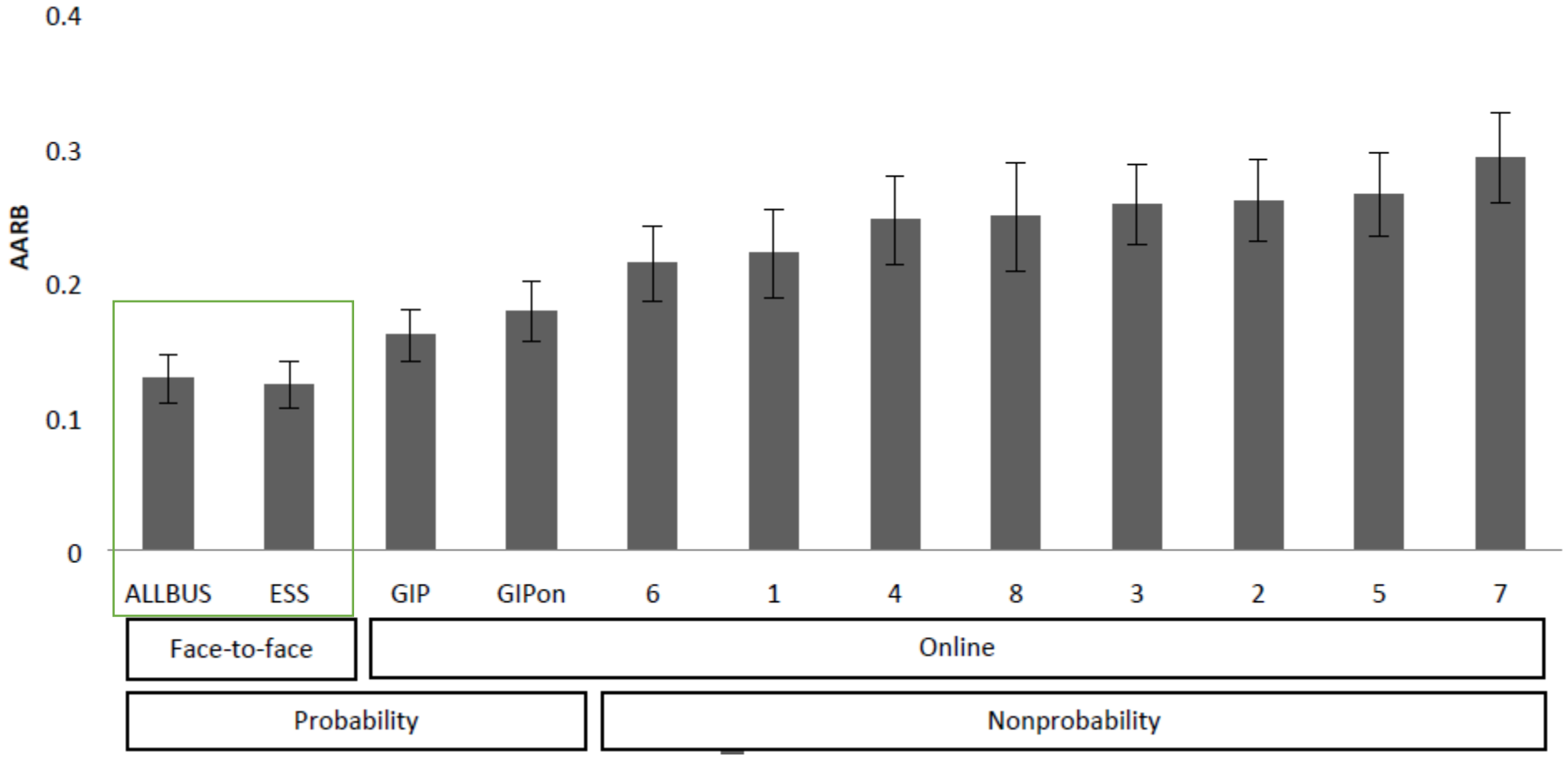
A Closer Look: Study III

Dutwin, D., & Buskirk, T. D. (2017). Apples to oranges or gala versus golden delicious? Comparing data quality of nonprobability internet samples to low response rate probability samples. *Public Opinion Quarterly*, 81(S1), 213-239. https://doi.org/10.1093/poq/nfw061

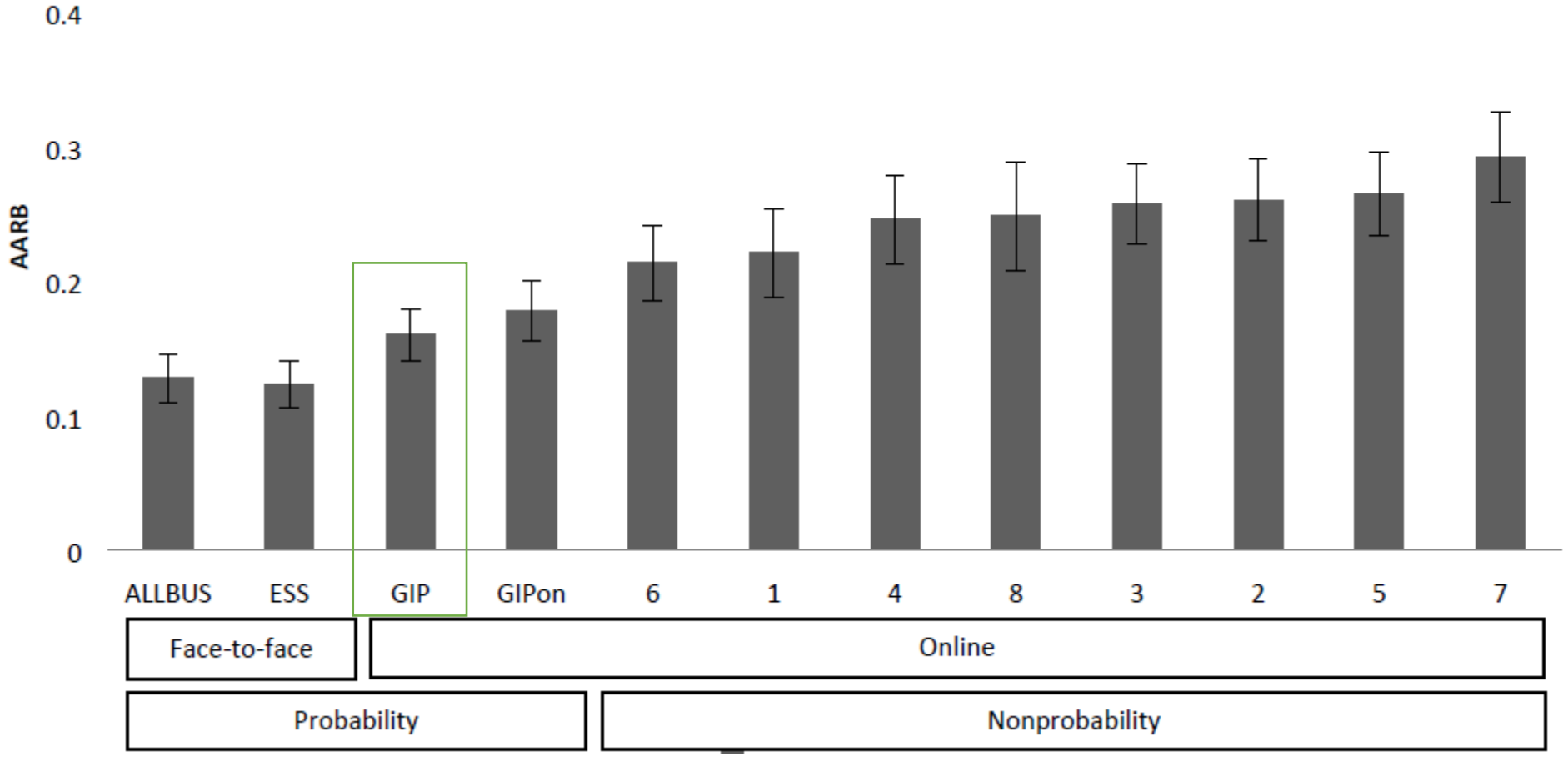
- 1 probability face-to-face survey sample
- 2 probability telephone survey samples
- 2 nonprobability online survey samples

- The probability face-to-face survey sample has the lowest bias.
- Followed by the probability telephone survey samples.
- Bias is highest in the nonprobability online survey samples.

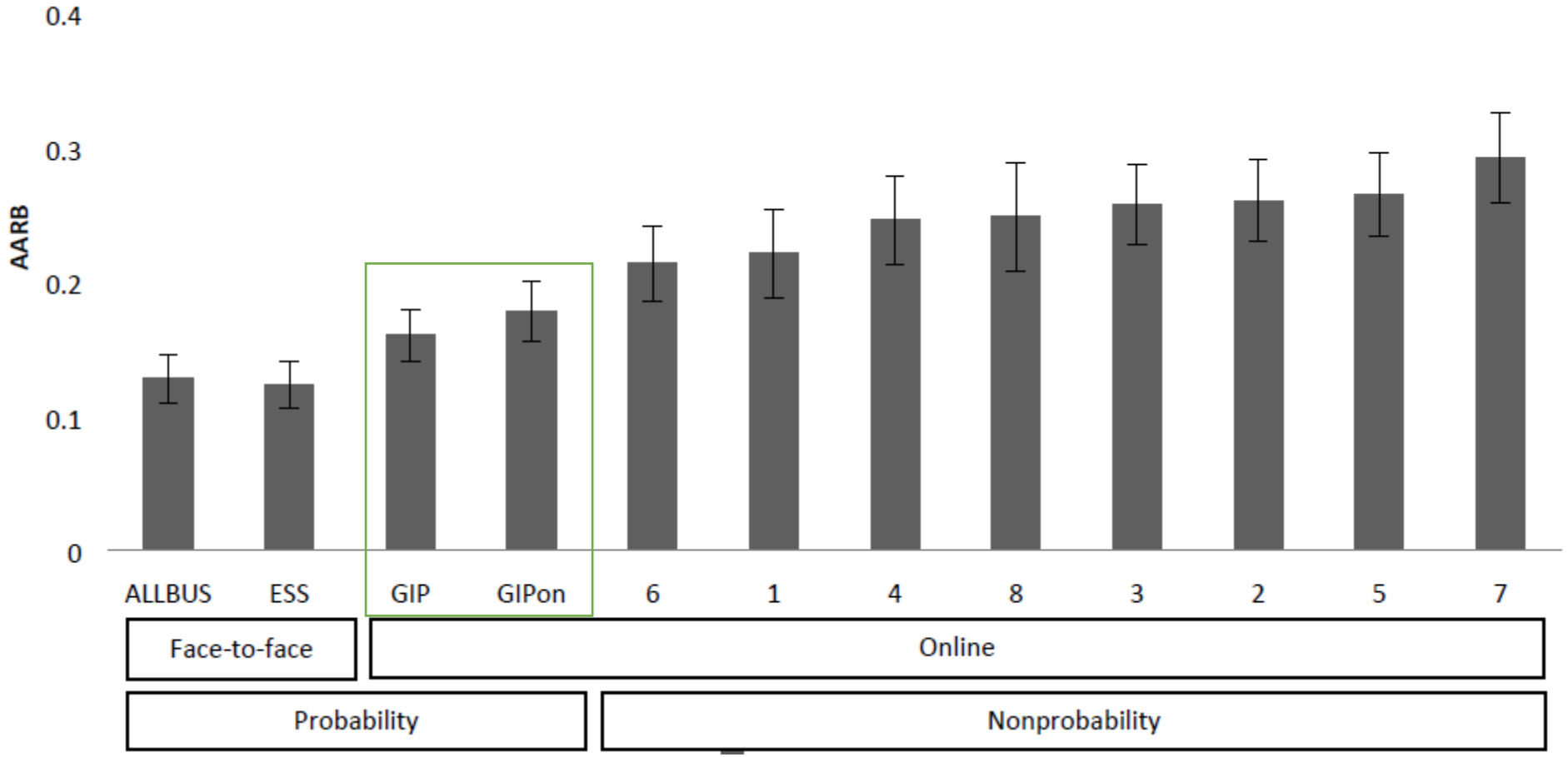
Baseline Bias Per Panel



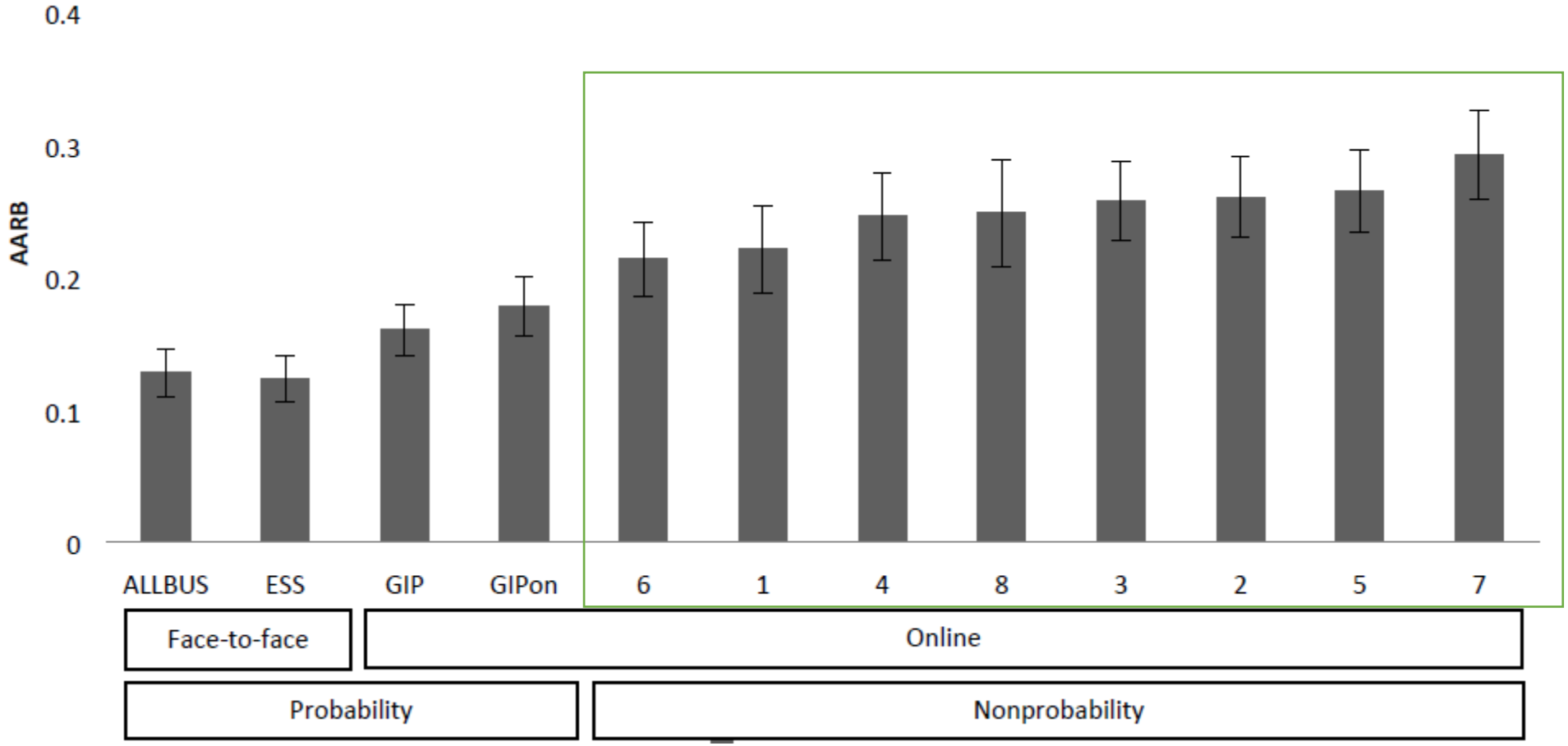
Baseline Bias Per Panel



Baseline Bias Per Panel



Baseline Bias Per Panel



Research Gaps / Current Research Topics

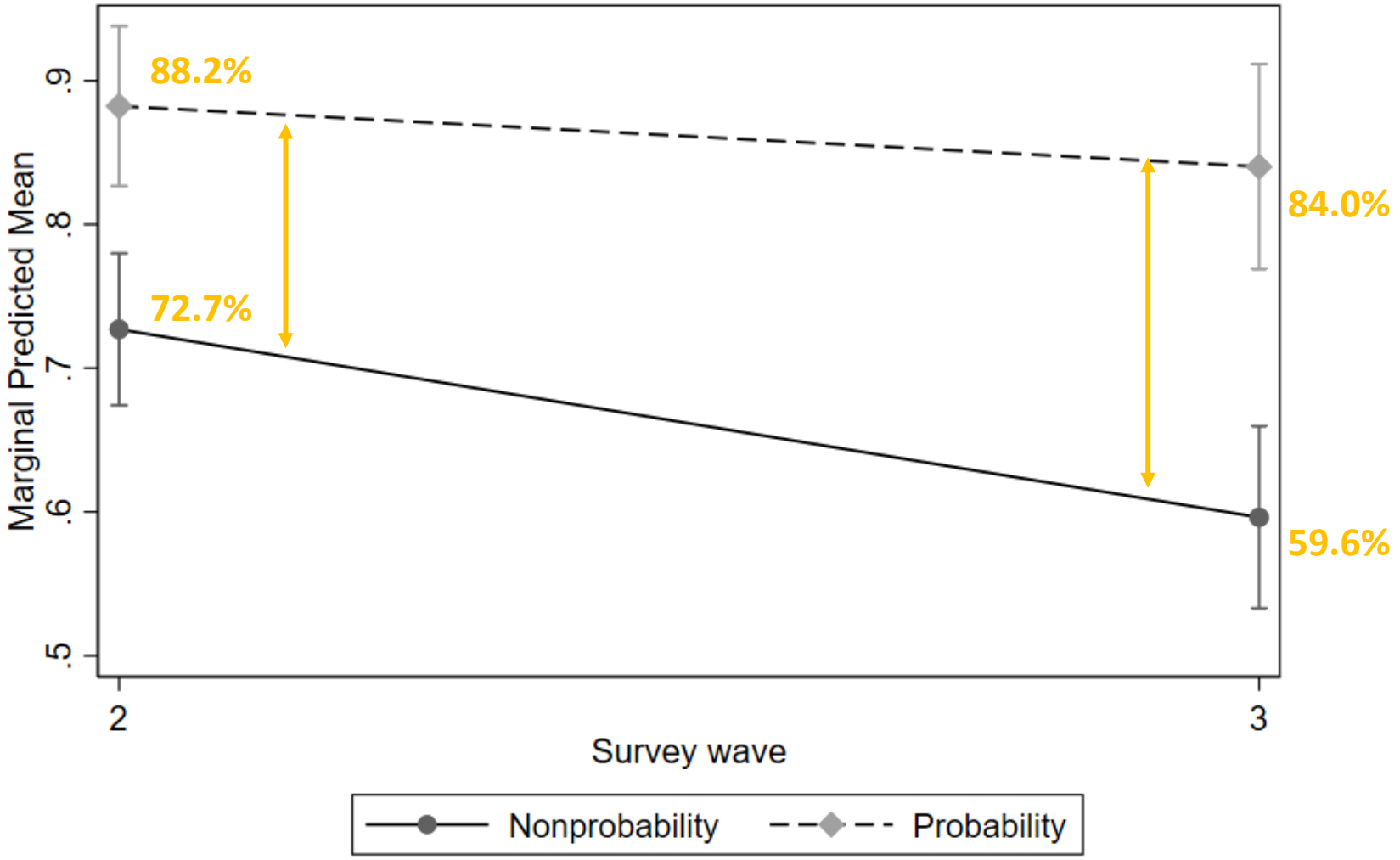
- To what extent …
  - non-probability online samples be used for **longitudinal research**?
  - do results from nonprobability online samples lead to **coherent conclusions** regarding correlations?
  - does the **answer quality** in nonprobability online samples differ from other samples?
  - can **adjustment** procedures, **data integration**, advanced **estimation** techniques improve estimates?
  - are results still valid **today**?
  - are new **data collection modes** affected?
  - is there progress on novel **methodologies**?
  - are **non-survey studies** affected?

German Online Panel Comparison Study

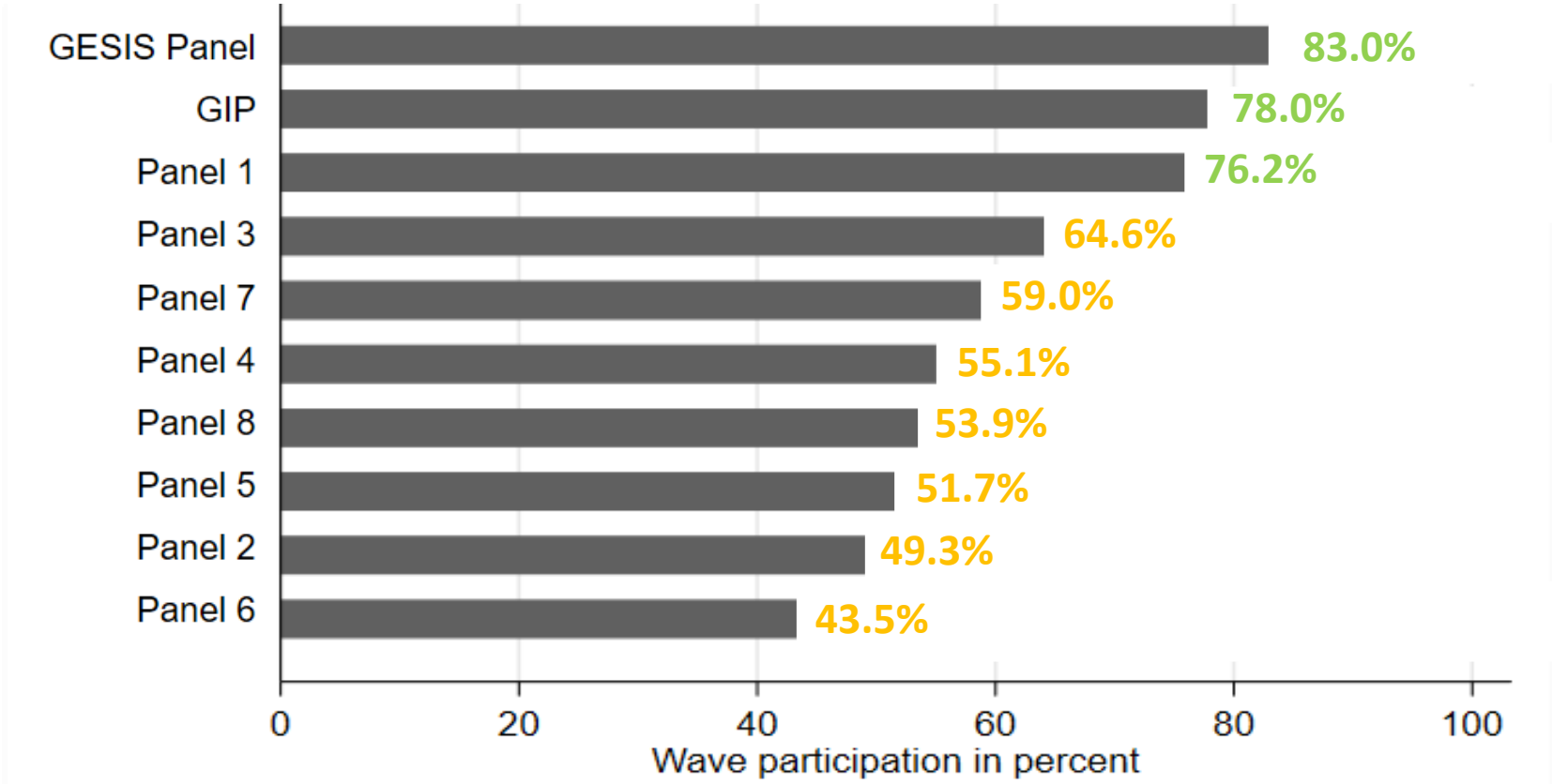
- Large-scale survey comparison
  - 2 probability-based online panels
  - 8 nonprobability online panels*
  - 3 panel waves
  - Identical questionnaires
  - Same field times

*1 panel is hybrid: RDD recruitment, quota sub-sampling

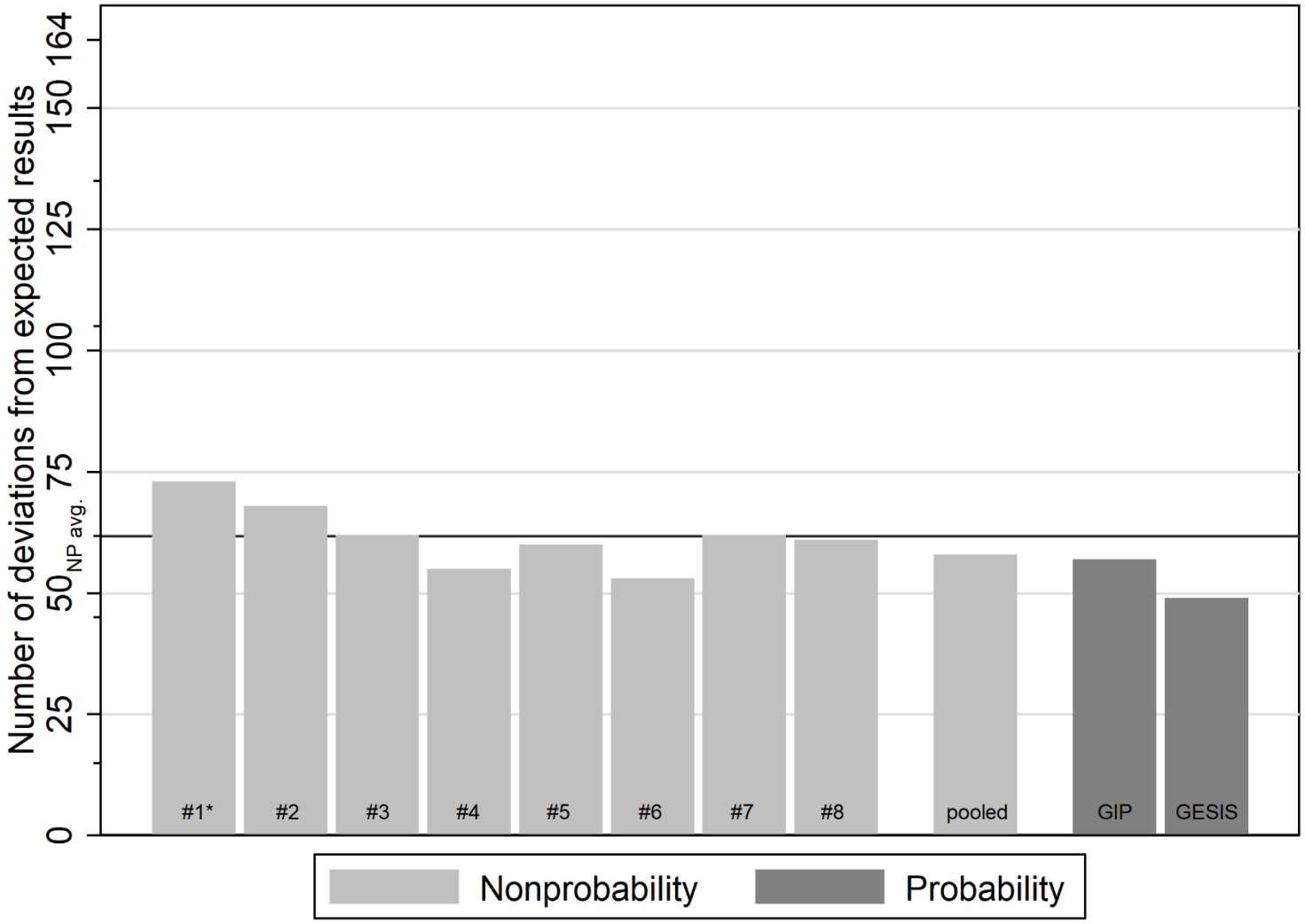
Longitudinal Research: Re-surveying wave 1 respondents



Longitudinal Research: Re-surveying respondents at both wave 2 & 3



Correlations: Deviations from expected results



Answer quality: Different indicators

- Significantly more straight-lining in grid questions in all nonprobability panels than in the probability-based panels
- But…
  - No difference in item nonresponse rates
  - No more midpoint selection tendencies
  - No measurement equivalence across *any* of the examined panel studies

Einarsson et al. (2022). Measurement Equivalence in Probability and Nonprobability Online Panels. International Journal of Market Research, 64(4), 484–505.

Cornesse, Carina and Annelies Blom (2021). Response Quality in Nonprobability and Probability-Based Online Panels. Sociological Methods and Research, https://doi.org/10.1177/0049124120914940.

Conclusions from GOPCS

- On univariate estimates & attrition, probability-based panels clearly perform better.
- In terms of measurement quality, results are more mixed.
- We need to learn more about the data generating process applied to the nonprobability online panels to explore potential causal mechanisms!
- We need to think more in terms of fitness-for-purpose rather than general/universal data quality!

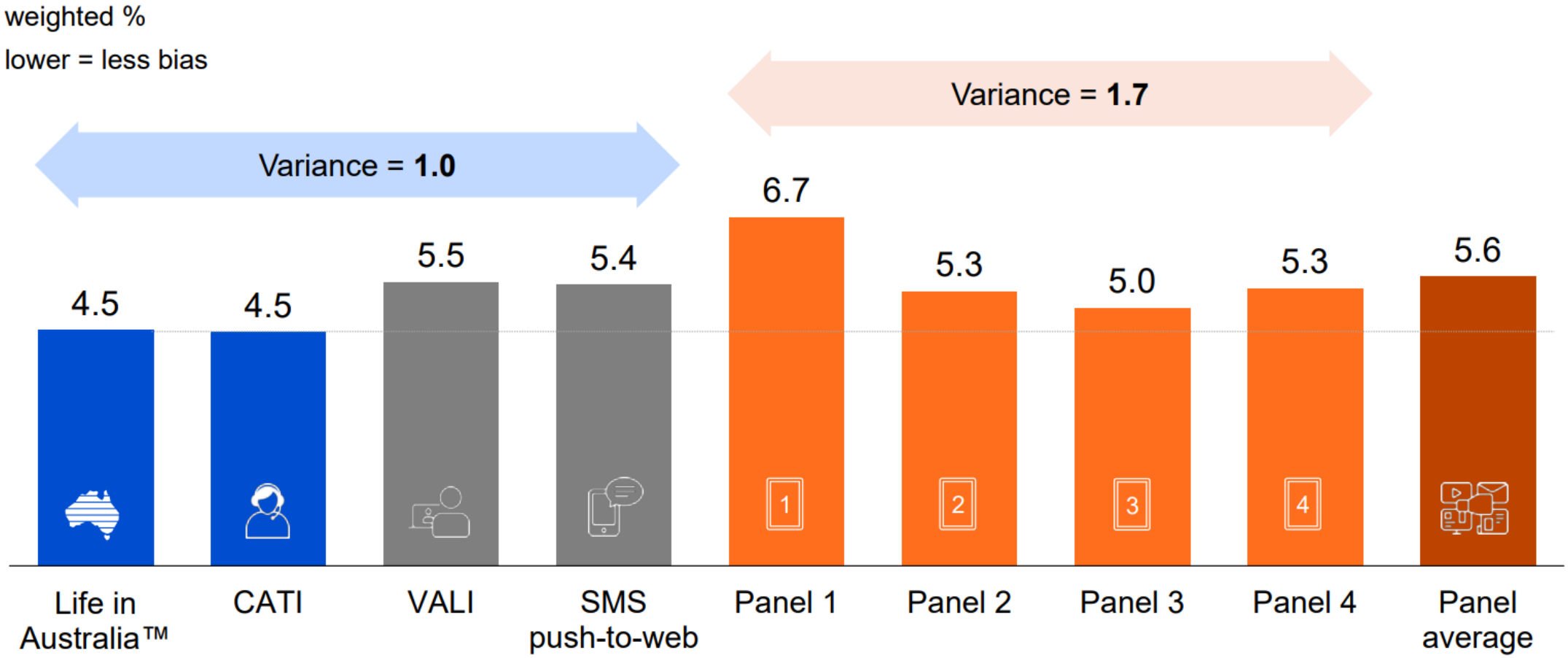
Australian Comparative Study of Survey Methods

- Pennay, D., Phillips, B., Neiger, D., Ward, A., Slamowicz, S., Lethborg, A. (2023). Australian Comparative Study of Survey Methods. https://srcentre.com.au/our-research/methods-research

- 1 probability-based online panel (Life in Australia™)
- 1 CATI survey
- 1 VALI survey
- 1 SMS-push-to-web survey
- 4 nonprobability online panels

Contact:
Dina Neiger: dina.neiger@srcentre.com.au
Benjamin Phillips: benjamin.phillips@srcentre.com.au

Overall accuracy (secondary demographics + substantive)



Pennay, D., Phillips, B., Neiger, D., Ward, A., Slamowicz, S., Lethborg, A. (2023). Australian Comparative Study of Survey Methods. https://srcentre.com.au/our-research/methods-research

Comparison with OPBS+

Compares accuracy of measures in 2015 vs 2022

Same demos and substantive variables

Comparable methods only

- CATI
- Life in Australia™
- 3 x non-prob panels

All five surveys included in this historical comparison produced more accurate measurements in 2022 than 2015

In the case of CATI, this serves as a reminder that response rates are generally a poor predictor of survey accuracy

(Kennedy & Hartig, 2019)

Gap between …		OPBS+ 2015	ACSSM 2022
	Life in Australia™ and the least biased non-prob panel	-1.4	-0.2
	Life in Australia™ and the 3-panel average	-2.0	-0.9
	CATI and the least biased non-prob panel	-1.6	-0.6
	CATI and the three-panel average	-2.2	-1.7

Kaczmirek, Lars, Benjamin Phillips, Darren Pennay and Dina Neiger. 2019. Building a Probability-Based Online Panel: Life in Australia™. CSRM & SRC Methods Paper No. 2/2019. Canberra: ANU Centre for Social Research & Methods, Research School of Social Sciences, College of Arts & Social Sciences, the Australian National University.

Lavrakas, Paul J., Darren Pennay, Dina Neiger and Benjamin Phillips. 2022. Comparing Probability-Based Surveys and Nonprobability Online Panel Surveys in Australia: A Total Survey Error Perspective. Survey Research Methods 16(2):241–66.

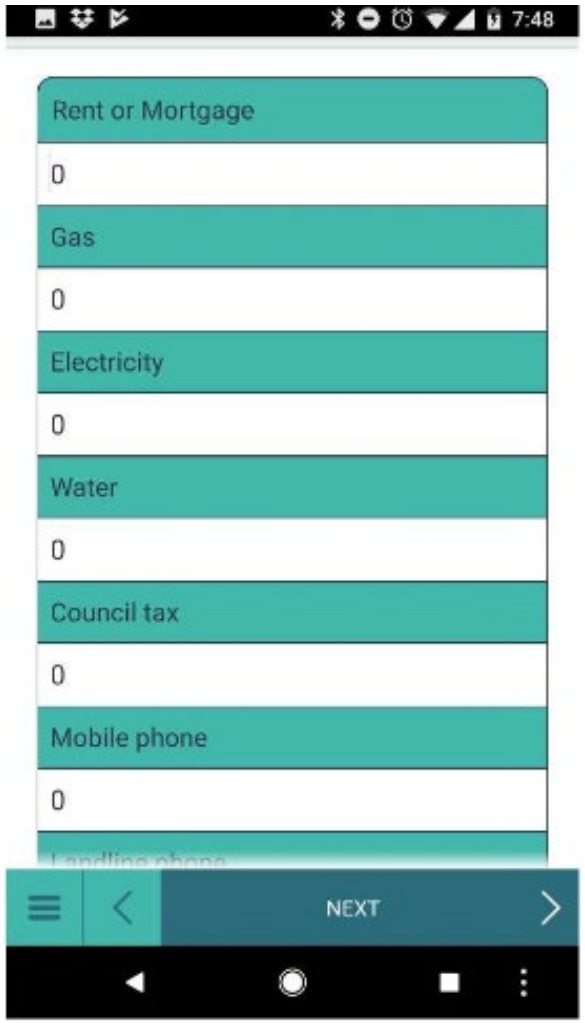
Evidence from the UK

Maslovskaya and Jessop (2021) Evaluating data quality in the UK probability-based and nonprobability-based online panels.

- 1 probability-based face-to-face survey sample – BSA - benchmark
- 1 probability-based online panel (recruited on the back of BSA)
- 3 nonprobability-based online panels
- Variables: demographic and socio-economic as well as substantive questions

- The bias is smallest in the probability-based online panel
- The bias is highest in the nonprobability online panels
- Estimates from probability-based online panel are the closer to BSA as well as bivariate associations

What about non-survey data?



Jäckle, A., Cornesse, C., Wenz, A., Couper, M. (under review). Measuring expenditure with a mobile app: Do probability-based and nonprobability panels differ?

Data: Panels

- ***Understanding Society* Innovation Panel (IP)**
  - stratified and clustered probability sample of households in Great Britain
  - annual interviews of all household members aged 16+
  - mixed-mode survey
- **Lightspeed GMI online access panel**
  - Commercial nonprobability online access panel
  - Same baseline questionnaire as IP

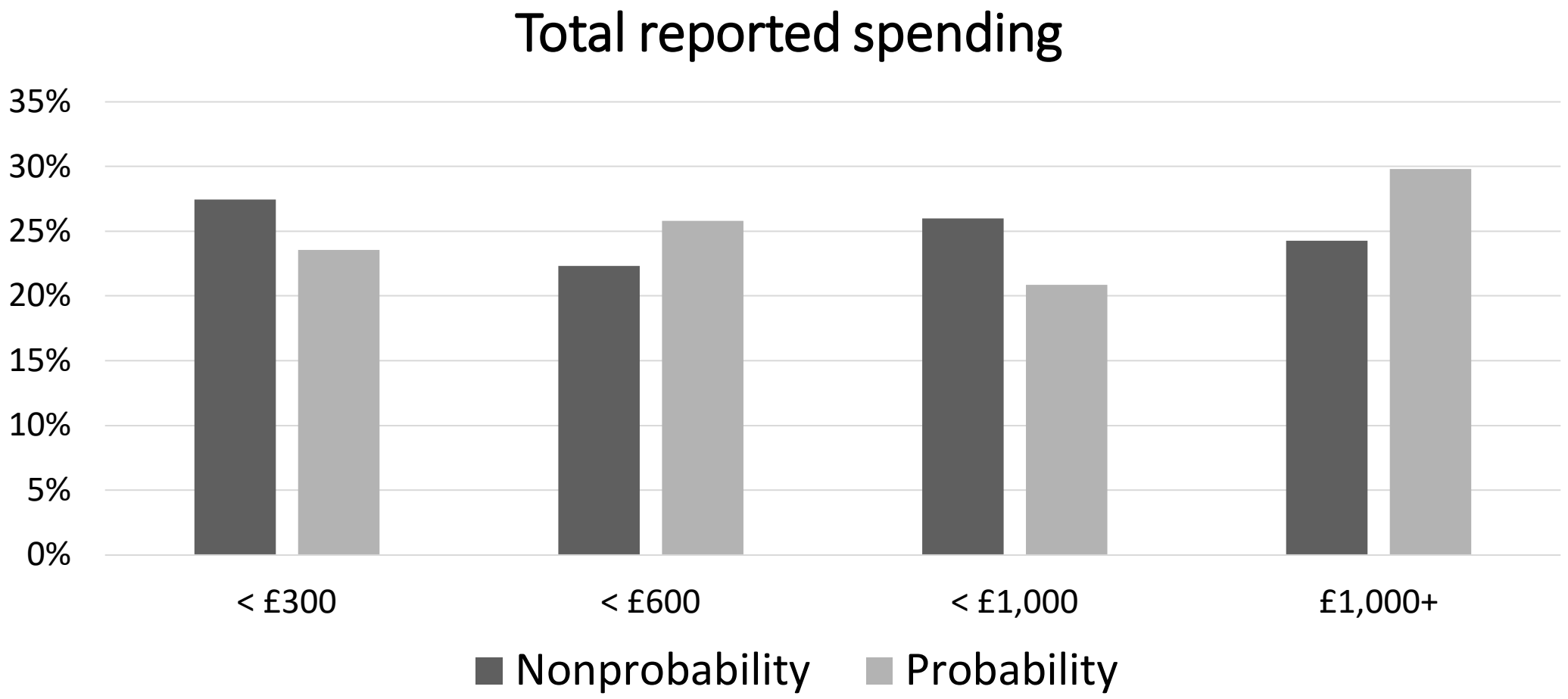
Task	Nonprob. Sample	Prob. Sample
Completed baseline questionnaire	2,878	2,638
Reported at least one purchase in app	408 (14%)	446 (17%)

Do different types of people participate?

Financial behavior (unweighted)

Category	Nonprob. Sample	Prob. Sample
Keep a budget: Yes	76%	42%
Check bank balance: Most days	45%	32%
Check bank balance: app	54%	46%

Do differences between samples matter?



Differences in outcomes of interest persist after controlling for socio-demographics, digital affinity, financial behavior

Adjustment & Estimation

Adjustment and estimation

- Adjustment procedures
  - Raking / poststratification
  - Propensity weights
  - Data integration (e.g. blended calibration)

- Estimation methods
  - Quasi-randomization
  - Super-population modeling
  - Doubly robust estimation
  - Multi-level regression & poststratification

Recommended literature

- Adjustment procedures

- Lee, S. (2006). Propensity score adjustment as a weighting scheme for volunteer panel web surveys. Journal of official statistics, 22(2), 329-349.
- Mercer, A., Lau, A., Kennedy, C. (2018). For Weighting Online Opt-In Samples, What Matters Most? https://www.pewresearch.org/methods/2018/01/26/how-different-weighting-methods-work/
- Wiśniowski, A., Sakshaug, J. W., Perez Ruiz, D. A., & Blom, A. G. (2020). Integrating probability and nonprobability samples for survey inference. Journal of Survey Statistics and Methodology, 8(1), 120-147.

- Estimation methods

- Beaumont, J. F., & Rao, J. N. K. (2021). Pitfalls of making inferences from non-probability samples: Can data integration through probability samples provide remedies? Surv. Stat, 83, 11-22.
- Valliant, R. (2020). Comparing alternatives for estimation from nonprobability samples. Journal of Survey Statistics and Methodology, 8(2), 231-263.

Concluding remarks

Nonprobability online samples

- Nonprobability survey samples are very diverse & recruited in various ways.
- Have become essential for experimental research and surveying hidden populations (see e.g. AAPOR webinar by Mariel McKone Leonard)
- Point of contention: Can findings be generalized to a broader population? When and why?
- Inference relies on assumptions. New/adapted frameworks can help with their evaluation
- Empirically, the literature so far suggests that biases persist and are usually bigger than in probability-based survey samples. There are exceptions.
- It may be that non-survey add-on studies inherit such biases from their parent survey samples. **More research is needed.**

Suggestions for moving forward

- It will be easier to assess and built trust in nonprobability online samples if …
  - they become fully transparent.
  - more in-depth and multi-faceted research can be conducted, aiming to understand data-generating mechanisms.
  - methodologies progress (e.g., data integration, respondent-driven sampling)
- Let's keep pushing the boundaries in this research field!

Feel free to contact us!

Carina Cornesse
ccornesse@diw.de

Olga Maslovskaya
om206@soton.ac.uk