Survey research in the digital age:
Online Nonprobability Surveys and Post-Stratification

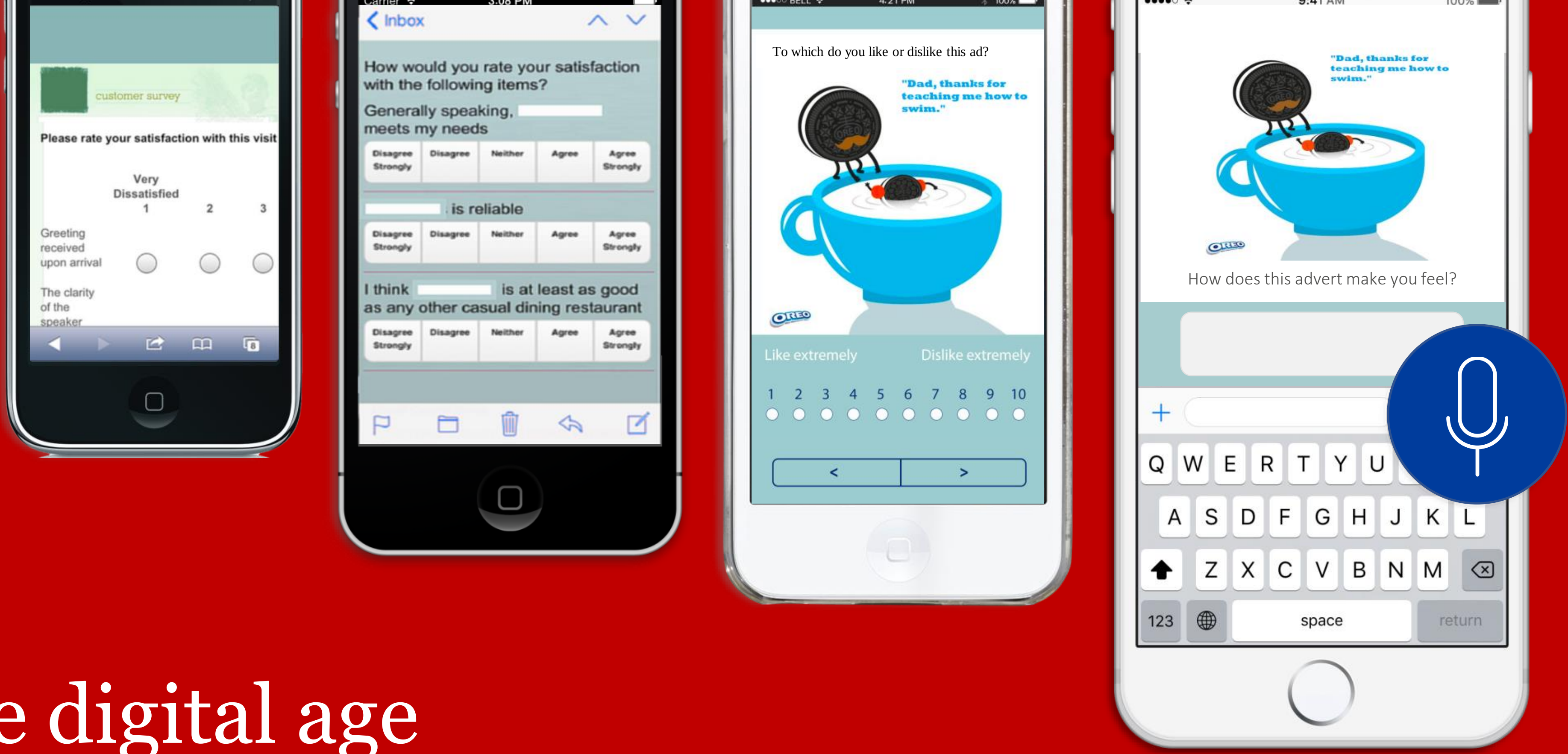
Oriol J. Bosch | Department of Methodology, LSE & RECSM

o.bosch-jover@lse.ac.uk
orioljbosch
https://orioljbosch.com/

Funding: This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 849165; PI: Melanie Revilla); the Spanish Ministry of Science and Innovation under the "R+D+i projects" programme (grant number PID2019-106867RB-I00 /AEI/10.13039/501100011033 (2020-2024), PI: Mariano Torcal); and the BBVA foundation under their grant scheme to scientific research teams in economy and digital society, 2019 (PI: Mariano Torcal).

Who am I?

- PhD Candidate at the **Methodology Department, LSE**
- Non-resident research fellow at the **Research and Expertise Centre for Survey Methodology, UPF**
- MSc in Survey Methods for Social Research from the **University of Essex**
- Worked for the **University of Southampton, Institute for Social and Economic Research, ESS and Netquest**
- Consultant for **The Alan Turing Institute, Wellcome Trust, Social Care Institute for Excellence and MoneyHelper**



Surveys in the digital age

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
    1. Surveys are some of the most frequently used method for collecting data

Table 3. Different types of quantitative data by discipline, 2014–2015.

Discipline	Survey	Admin	Census	Big data	n
Sociology	51%	42%	16%	4%	277
Political Sciences	41%	58%	9%	4%	308
Economics	32%	74%	19%	3%	374
Social Psychology	69%	5%	0%	2%	235
Public Opinion	86%	16%	3%	5%	81
TOTAL	**49%**	**47%**	**11%**	**3%**	**1275**

Sturgis, P., & Luff, R. (2021). **The demise of the survey? A research note on trends in the use of survey data in the social sciences**, 1939 to 2015. *International Journal of Social Research Methodology*, *24*(6), 691-696.

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
  1. Surveys are some of the most frequently used method for collecting data
  2. But they look significantly different than before:

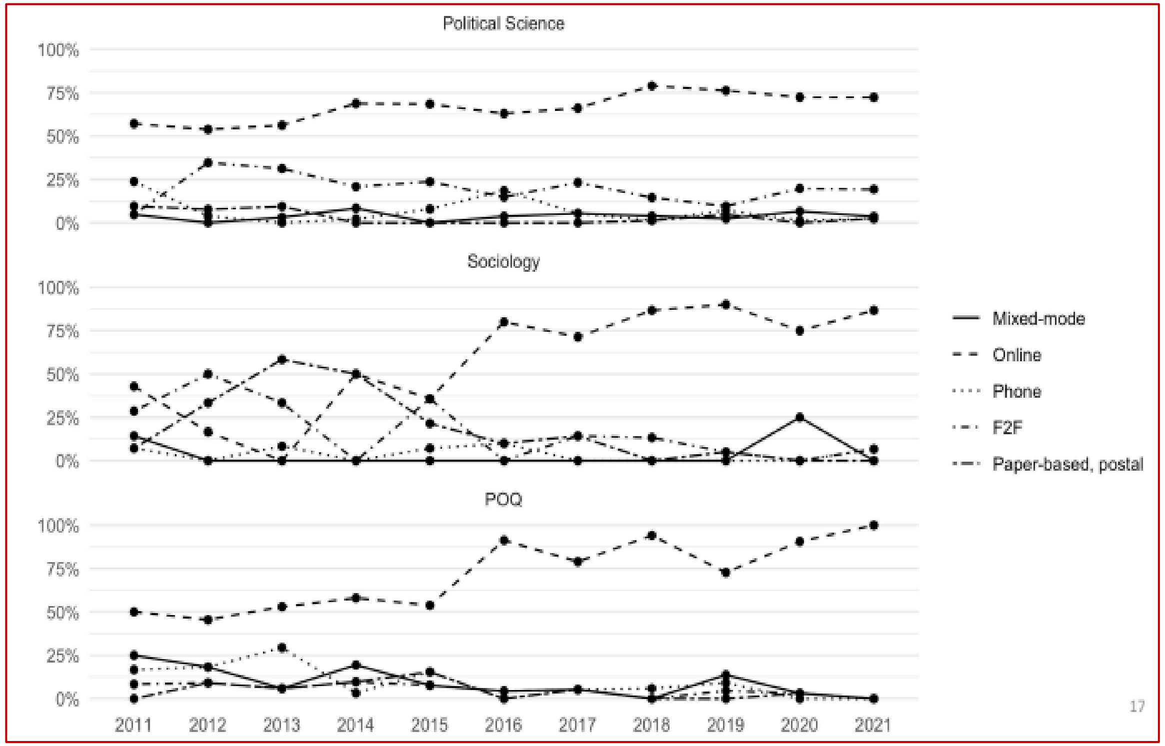
	Sampling	Interviews
1st era	Area probability	Face-to-face
2nd era	Random digital dial probability	Telephone
3rd era	Non-probability	Computer-administered

Groves, R. M. (2011). **Three eras of survey research**. *Public opinion quarterly*, *75*(5), 861-871.

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
  1. Surveys are some of the most frequently used method for collecting data
  2. But they look significantly different than before: **online**

	Sampling	Interviews
1st era	Area probability	Face-to-face
2nd era	Random digital dial probability	Telephone
3rd era	Non-probability	Computer-administered

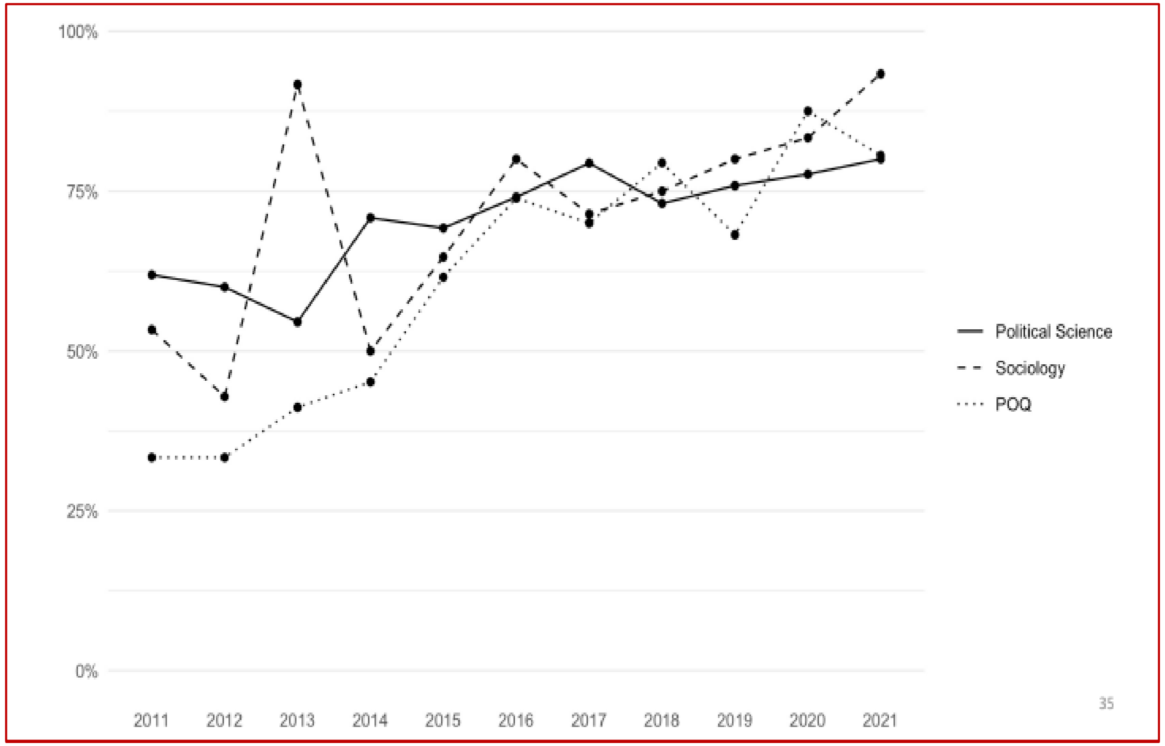


Stefkovics, Harrison, Skinnion and Eichorst (2022). **Survey Use in Substantive Research: Trends Across Disciplines**. *Paper presented at the 77th Annual Conference of the American Association for Public Opinion Research. Available at:* https://drive.google.com/file/d/12x3ZbSp1e1HseCPjX9AJcG4BQp_7Nt64/view

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
  1. Surveys are some of the most frequently used method for collecting data
  2. But they look significantly different than before: **online, nonprobability**

	Sampling	Interviews
1st era	Area probability	Face-to-face
2nd era	Random digital dial probability	Telephone
3rd era	Non-probability	Computer-administered



Stefkovics, Harrison, Skinnion and Eichorst (2022). **Survey Use in Substantive Research: Trends Across Disciplines**. *Paper presented at the 77th Annual Conference of the American Association for Public Opinion Research. Available at:* https://drive.google.com/file/d/12x3ZbSp1e1HseCPjX9AJcG4BQp_7Nt64/view

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
  1. Surveys are some of the most frequently used method for collecting data
  2. But they look significantly different than before: **online, nonprobability**

	Sampling	Interviews
1st era	Area probability	Face-to-face
2nd era	Random digital dial probability	Telephone
3rd era	Non-probability	Computer-administered

Online Nonprobability Surveys

Surveys are (still) relevant

- **A highly relevant but ever changing tool**
  1. Surveys are some of the most frequently used method for collecting data
  2. But they look significantly different than before: **online, nonprobability & linked**

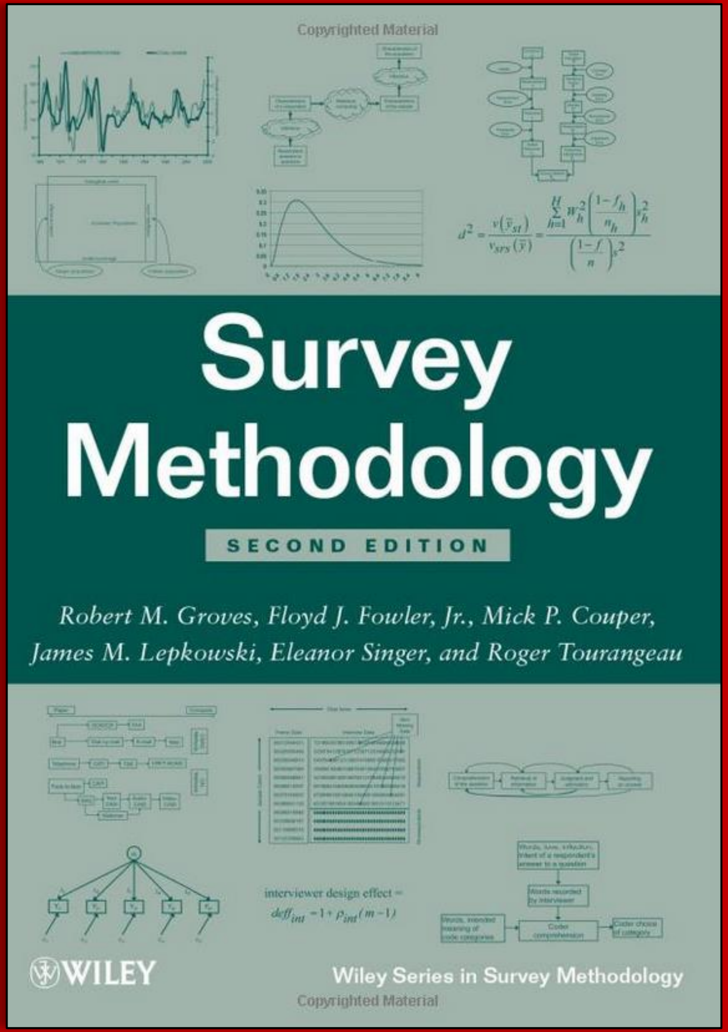
	Sampling	Interviews	Data environment
1st era	Area probability	Face-to-face	Stand-alone
2nd era	Random digital dial probability	Telephone	Stand-alone
3rd era	Non-probability	Computer-administered	Linked

Online Nonprobability Surveys

Smart Surveys / Enhanced online surveys

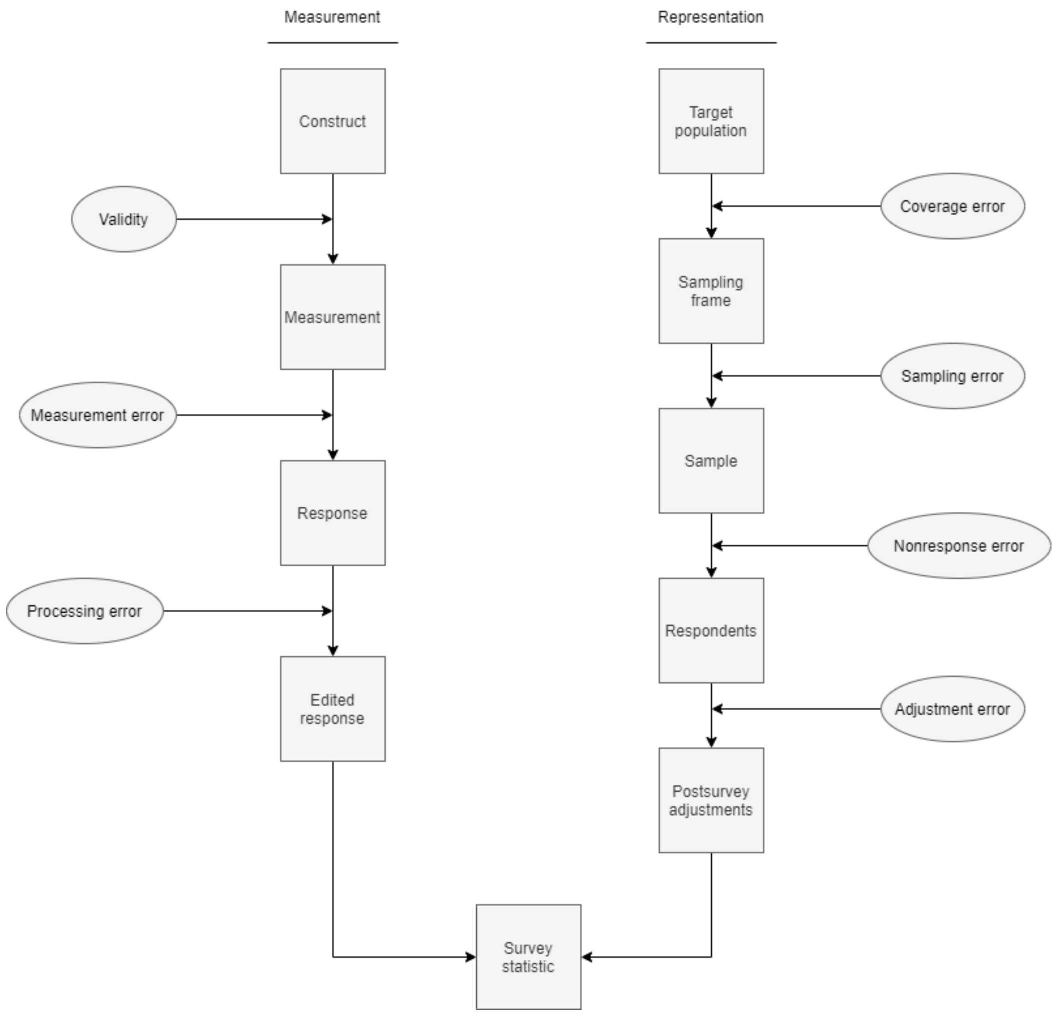
Ricciato, F., Wirthmann, A., & Hahn, M. (2020). **Trusted Smart Statistics: How new data will change official statistics**. *Data & Policy*, *2*.

The basics of survey research



The Total Survey Error (TSE) framework

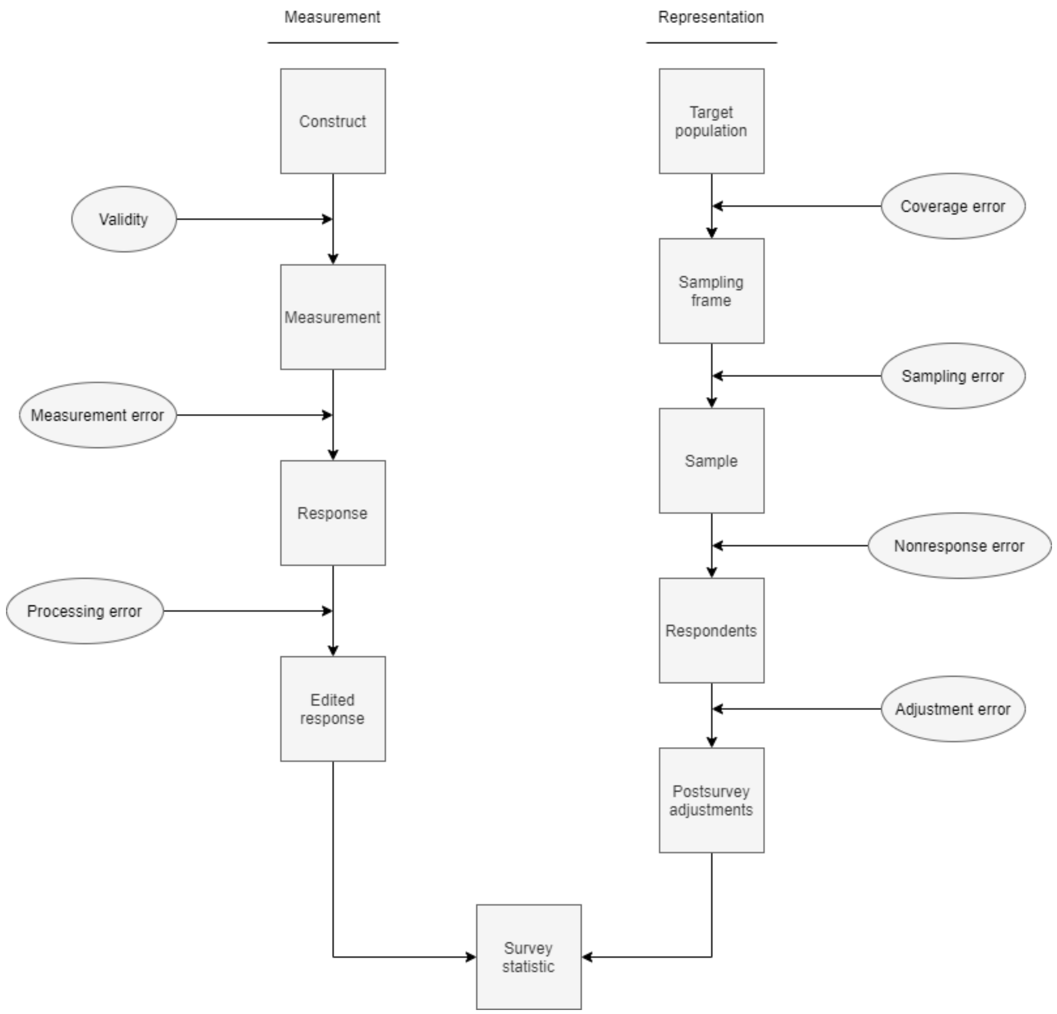
- In general, surveys are used to **make inferences** about a **concept of interest** for a given **population**



The Total Survey Error (TSE) framework

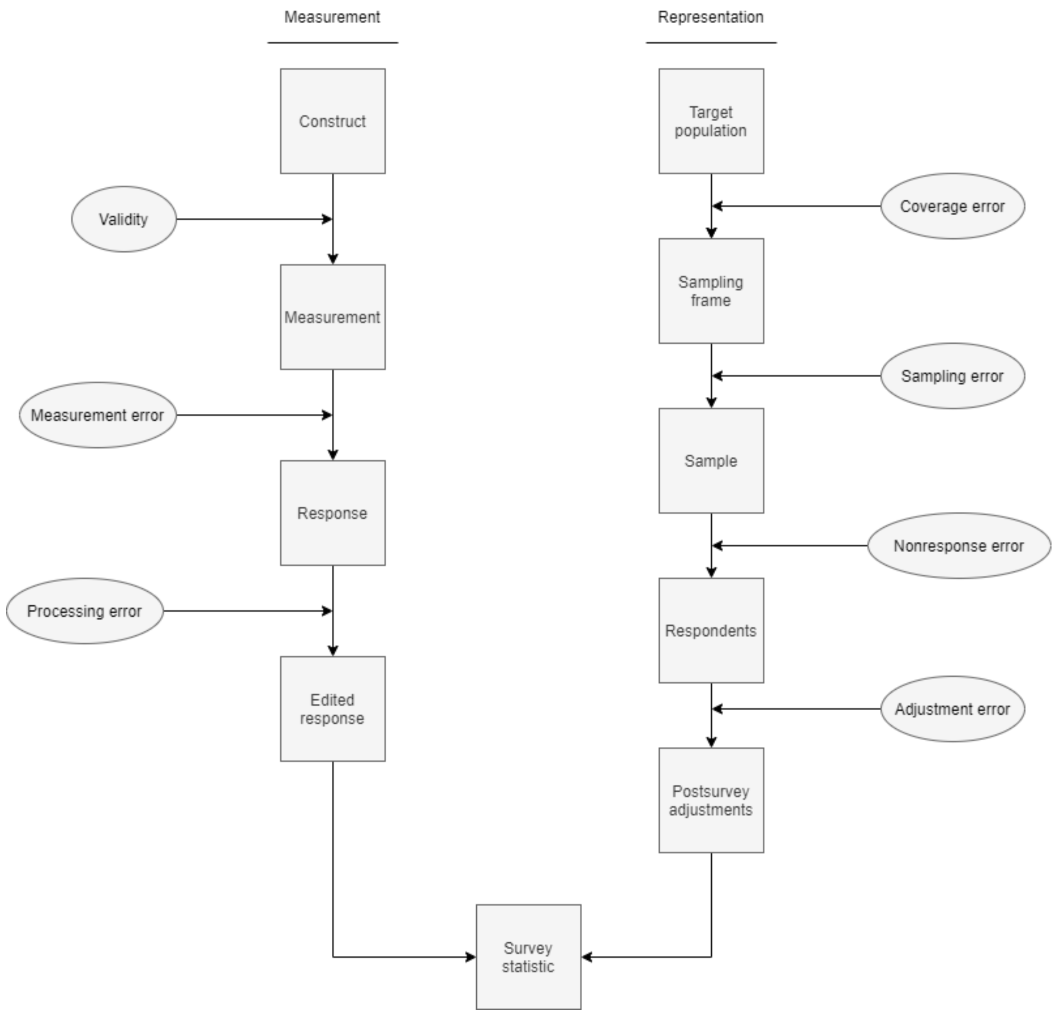
- In general, surveys are used to **make inferences** about a **concept of interest** for a given **population**

- Two parallel processes: **measurement** and **representation**



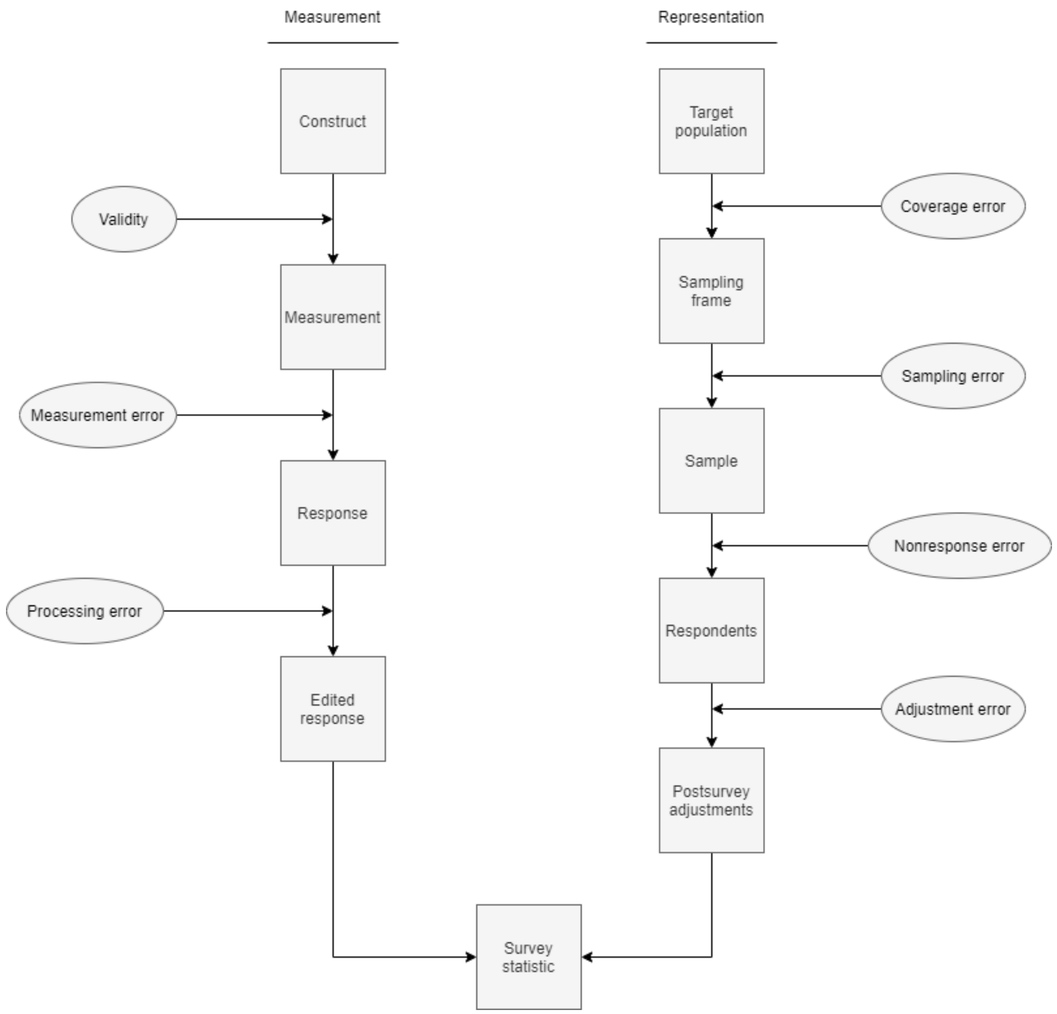
The Total Survey Error (TSE) framework

- In general, surveys are used to **make inferences** about a **concept of interest** for a given **population**
- Two parallel processes: **measurement** and **representation**
- Errors can happen in both sides



The Total Survey Error (TSE) framework

- In general, surveys are used to **make inferences** about a **concept of interest** for a given **population**
- Two parallel processes: **measurement** and **representation**
- Errors can happen in both sides
- The goal is to, within the project's **time** and **budget** constraints, **reduce as much as possible** the errors

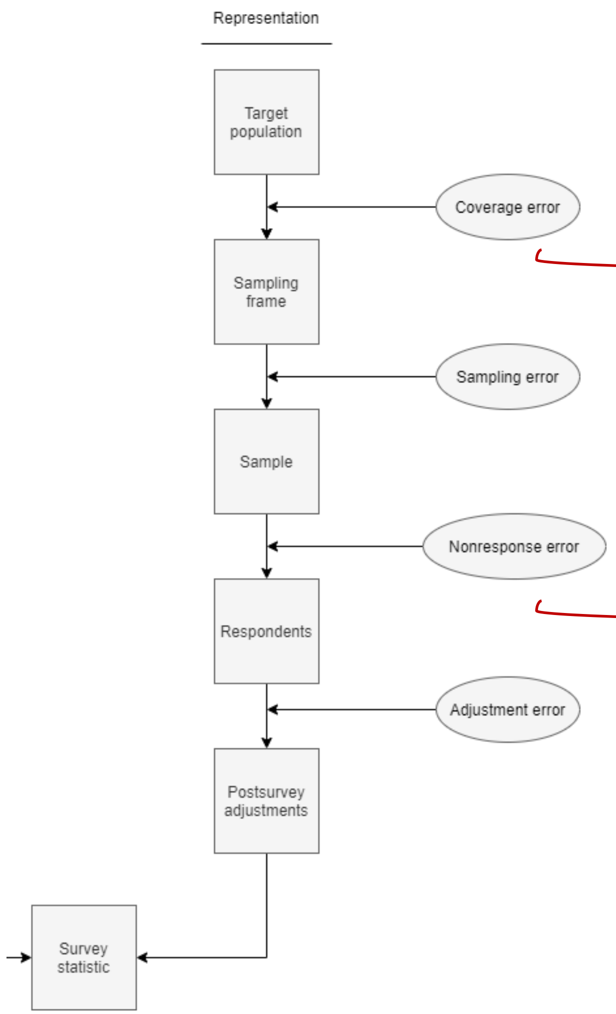


What are online nonprobability surveys?

Two main characteristics

Two main characteristics

- First, they are **online surveys**
  - No interviewer
  - Answered through connected devices
  - Visual instead of aural
  - Easier to design and faster to field
  - Etc.



No offline population

Higher nonresponse for less internet-savvy people?

Two main characteristics

- First, they are **online surveys**
  - No interviewer
  - Answered through connected devices
  - Visual instead of aural
  - Easier to design and faster to field
  - Etc.

- And secondly, they are **nonprobability**

Two main characteristics

- First, they are **online surveys**
  - No interviewer
  - Answered through connected devices
  - Visual instead of aural
  - Easier to design and faster to field
  - Etc.

- And secondly, they are **nonprobability**

Our main interest

Probability VS Nonprobability

webdataopp

- **Probability sampling** ➡ Every unit from a frame population has a **known and non-zero probability** of inclusion

Probability VS Nonprobability

- **Probability sampling** ➡ Every unit from a frame population has a **known and non-zero probability** of inclusion

$\neq$ The sample is selected "at random"

$\neq$ The sample is "representative"

$=$ we understand the selection process

$=$ we know the probability of being in the sample

Probability VS Nonprobability

- Nonprobability sampling ➡ The **selection probabilities are unknown** and, for some people, **zero**.

Probability VS Nonprobability

- Nonprobability sampling ➡ The **selection probabilities are unknown** and, for some people, **zero**.

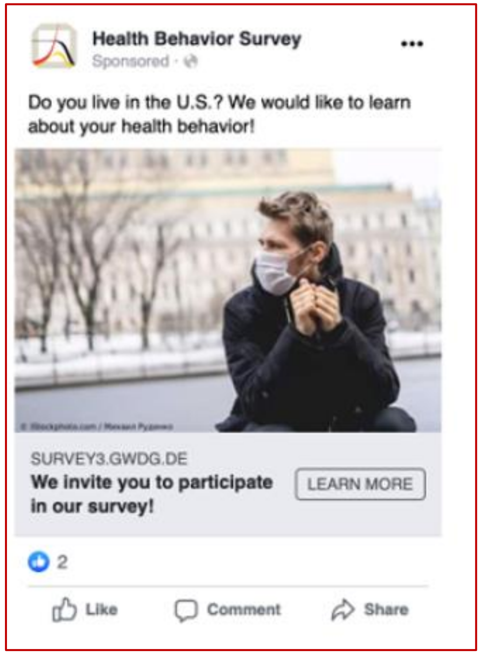
- With online surveys, this is mostly due to two reasons:

Probability VS Nonprobability

- Nonprobability sampling ➡ The **selection probabilities are unknown** and, for some people, **zero**.

- With online surveys, this is mostly due to two reasons:
    1. There is no **frame** to use

Users of a platform are prompted with a link



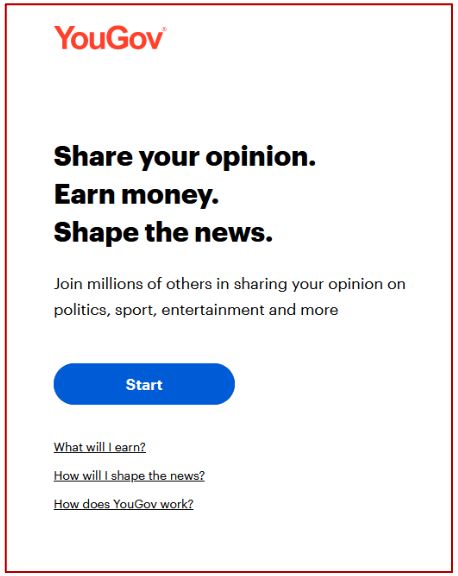
Those who click and answer are the sample

Sample

Probability VS Nonprobability

- Nonprobability sampling → The **selection probabilities are unknown** and, for some people, **zero**.

- With online surveys, this is mostly due to two reasons:
  1. There is no **frame** to use
  2. There is a **“frame”**, but it is **unclear how people have been selected** to be part of it (not in a prob. way)



→ **Panel (frame)** → **Sample**

People opt-in into the panel

Panellists are invited to participate, not at random

How can we run online nonprobability surveys?

The 3 key steps

1. Identify **from where you will obtain participants**

2. Prepare the **sampling design**

3. Create an **adjustment approach**

First step: Obtaining participants

Without a proper frame, we need other sources of participants:

First step: Obtaining participants

Without a proper frame, we need other sources of participants:

1. **Social media**

More on this later!

Perrotta et al. *EPJ Data Science* (2021) 10:17
https://doi.org/10.1140/epjds/s13688-021-00270-1

EPJ.org

EPJ Data Science
a SpringerOpen Journal

REGULAR ARTICLE Open Access

Behaviours and attitudes in response to the COVID-19 pandemic: insights from a cross-national Facebook survey

Daniela Perrotta[1*†], André Grow[1†], Francesco Rampazzo[2], Jorge Cimentada[1], Emanuele Del Fava[1], Sofia Gil-Clavel[1] and Emilio Zagheni[1]

[*]Correspondence:
perrotta@demogr.mpg.de
[1]Max Planck Institute for Demographic Research, Konrad-Zuse-Straße 1, Rostock, Germany
Full list of author information is available at the end of the article
[†]Equal contributors

Abstract

Background: In the absence of medical treatment and vaccination, individual behaviours are key to curbing the spread of COVID-19. Here we describe efforts to collect attitudinal and behavioural data and disseminate insights to increase situational awareness and inform interventions.

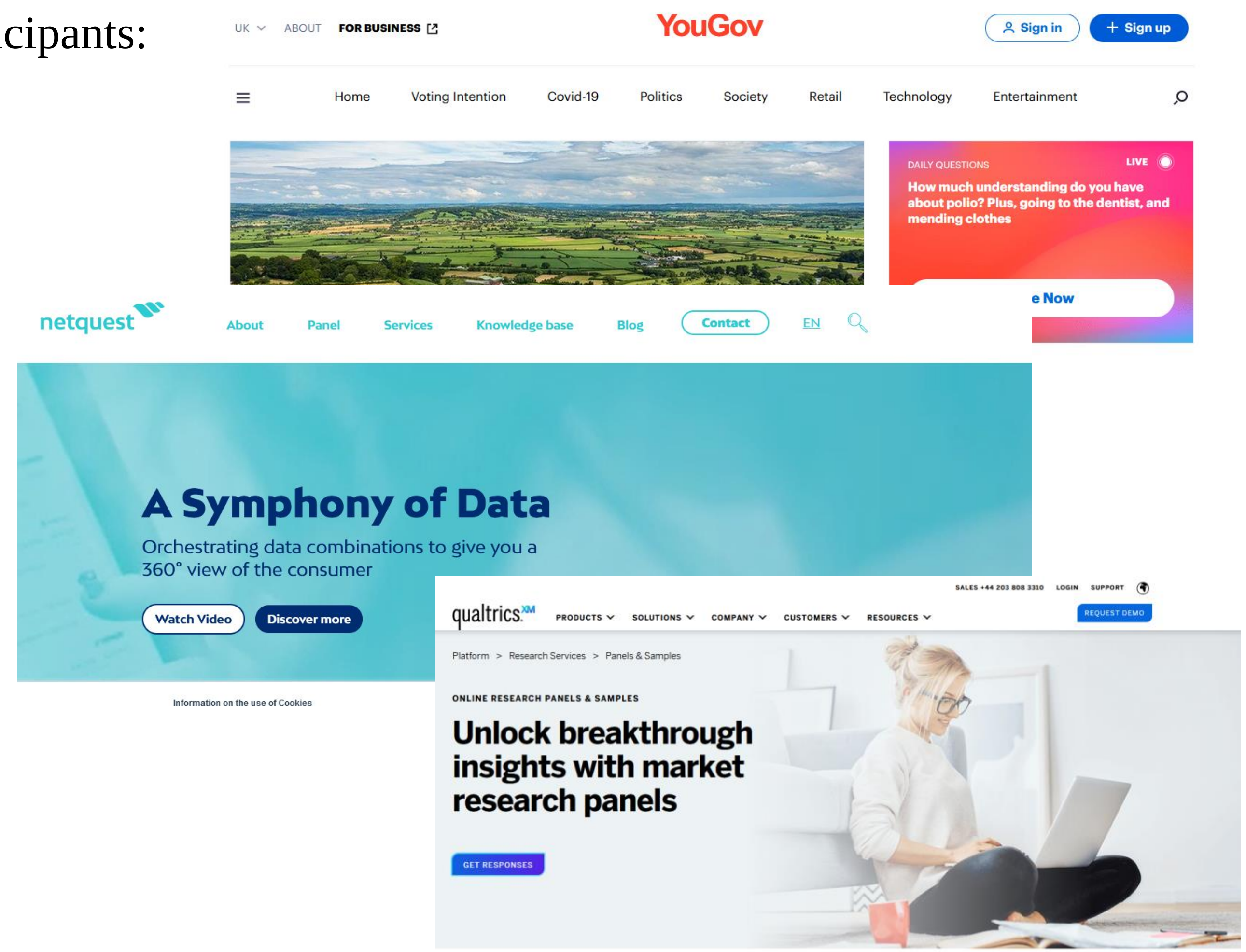
Methods: We developed a rapid data collection and monitoring system based on a cross-national online survey, the "COVID-19 Health Behavior Survey". Respondent recruitment occurred via targeted Facebook advertisements in Belgium, France, Germany, Italy, the Netherlands, Spain, the United Kingdom, and the United States. We investigated how the threat perceptions of COVID-19, the confidence in the preparedness of organisations to deal with the pandemic, and the adoption of preventive and social distancing behaviours are associated with respondents' demographic characteristics.

Results: We analysed 71,612 questionnaires collected between March 13-April 19, 2020. We found substantial spatio-temporal heterogeneity across countries at different stages of the pandemic and with different control strategies in place. Respondents rapidly adopted the use of face masks when they were not yet mandatory. We observed a clear pattern in threat perceptions, sharply increasing from a personal level to national and global levels. Although personal threat perceptions were comparatively low, all respondents significantly increased hand hygiene. We found gender-specific patterns: women showed higher threat perceptions, lower confidence in the healthcare system, and were more likely to adopt preventive

First step: Obtaining participants

Without a proper frame, we need other sources of participants:

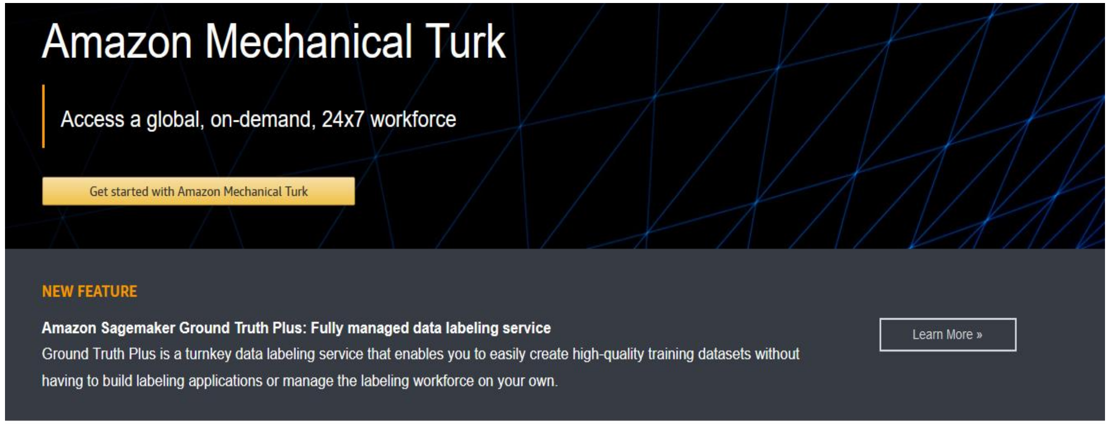
1. **Social media**
2. **Opt-in online panels**

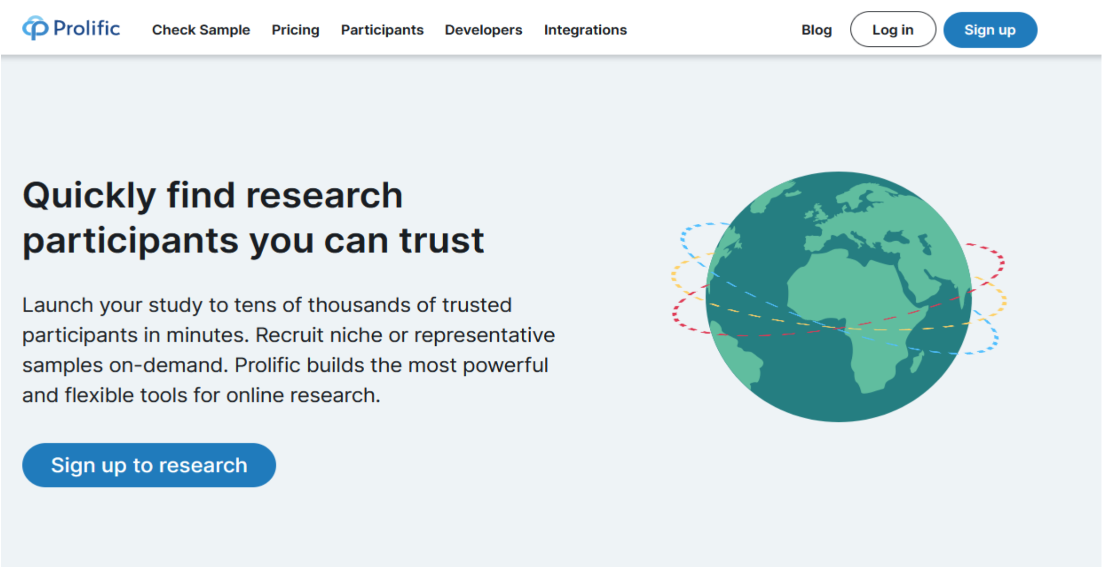


First step: Obtaining participants

Without a proper frame, we need other sources of participants:

1. **Social media**
2. **Opt-in online panels**
3. **Crowdsourcing / participants market places**

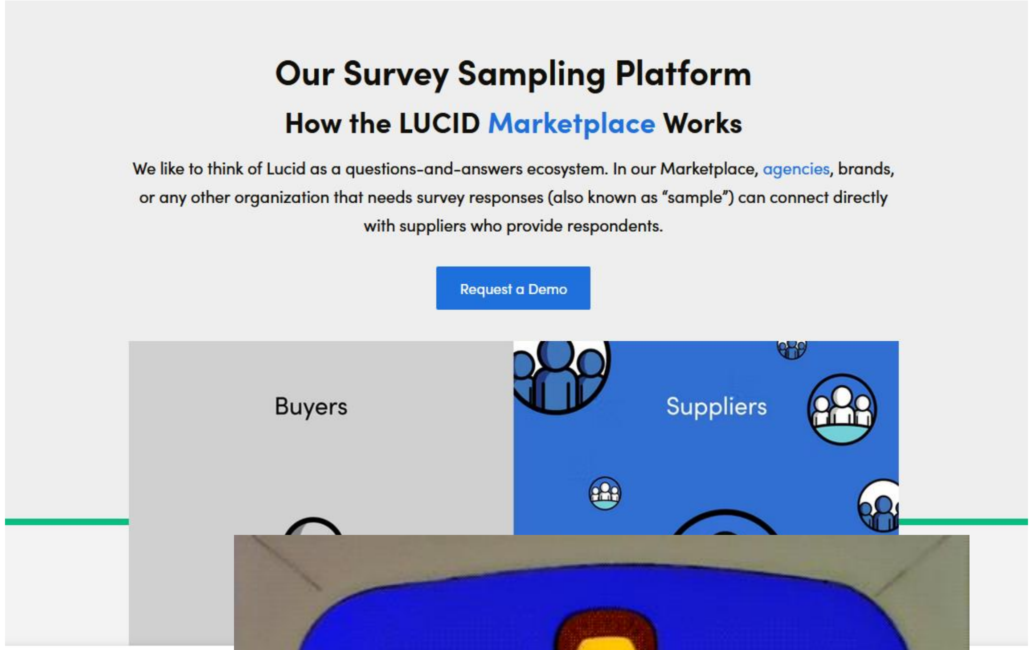




First step: Obtaining participants

Without a proper frame, we need other sources of participants:

1. **Social media**
2. **Opt-in online panels**
3. **Crowdsourcing / participants market places**
4. **Respondent aggregators**



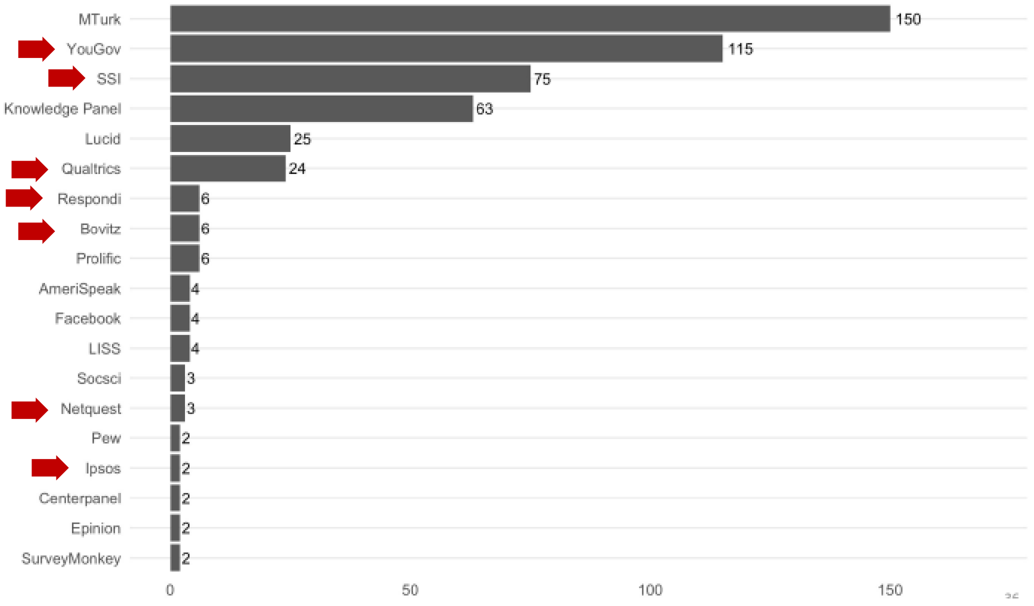
First step: Obtaining participants

Without a proper frame, we need other sources of participants:

1. **Social media**
2. **Opt-in online panels**
3. **Crowdsourcing / participants market places**
4. **Respondent aggregators**

Main focus today

Nº of surveys for platform



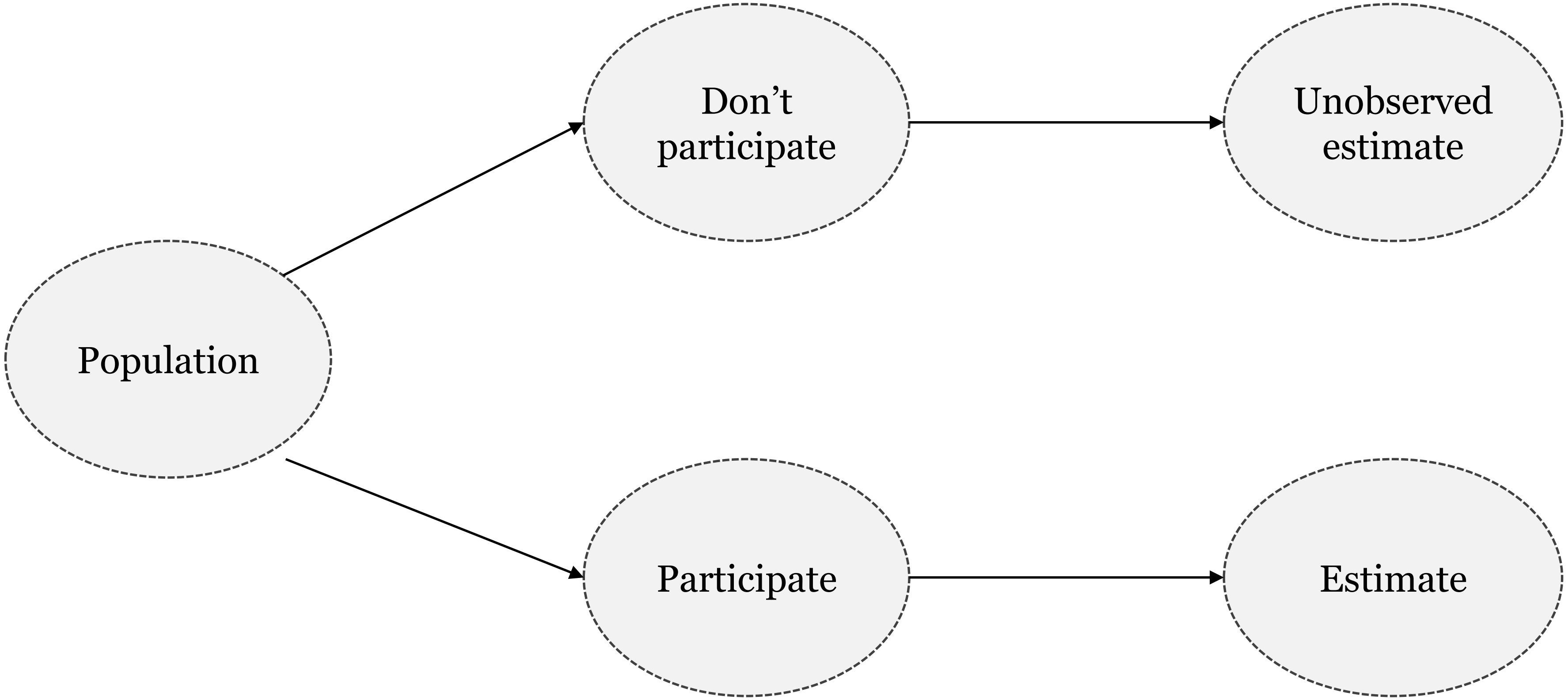
Stefkovics, Harrison, Skinnion and Eichorst (2022). **Survey Use in Substantive Research: Trends Across Disciplines**. *Paper presented at the 77th Annual Conference of the American Association for Public Opinion Research. Available at:* https://drive.google.com/file/d/12x3ZbSp1e1HseCPjX9AJcG4BQp_7Nt64/view

Second Step: Sampling designs

We **cannot** rely on **randomization** techniques anymore

Second Step: Sampling designs

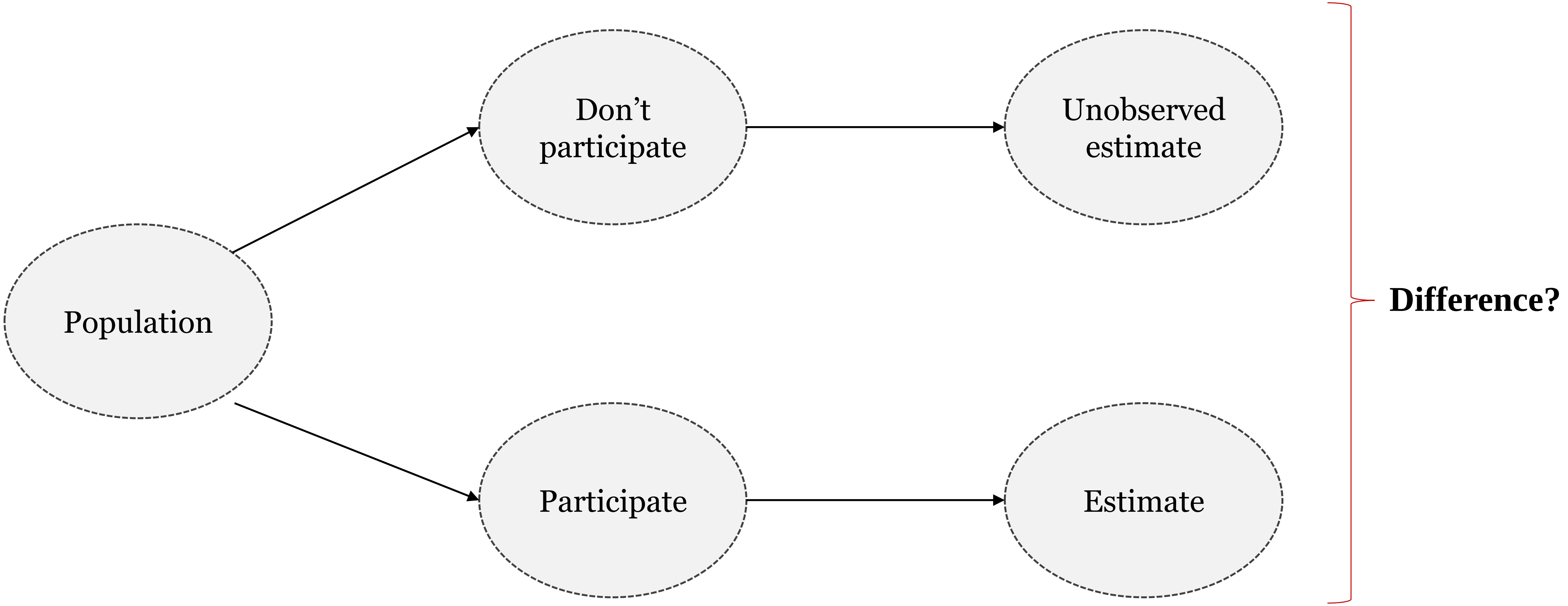
Change of paradigm: **sampling design through the lenses of causal inference**



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.

Second Step: Sampling designs

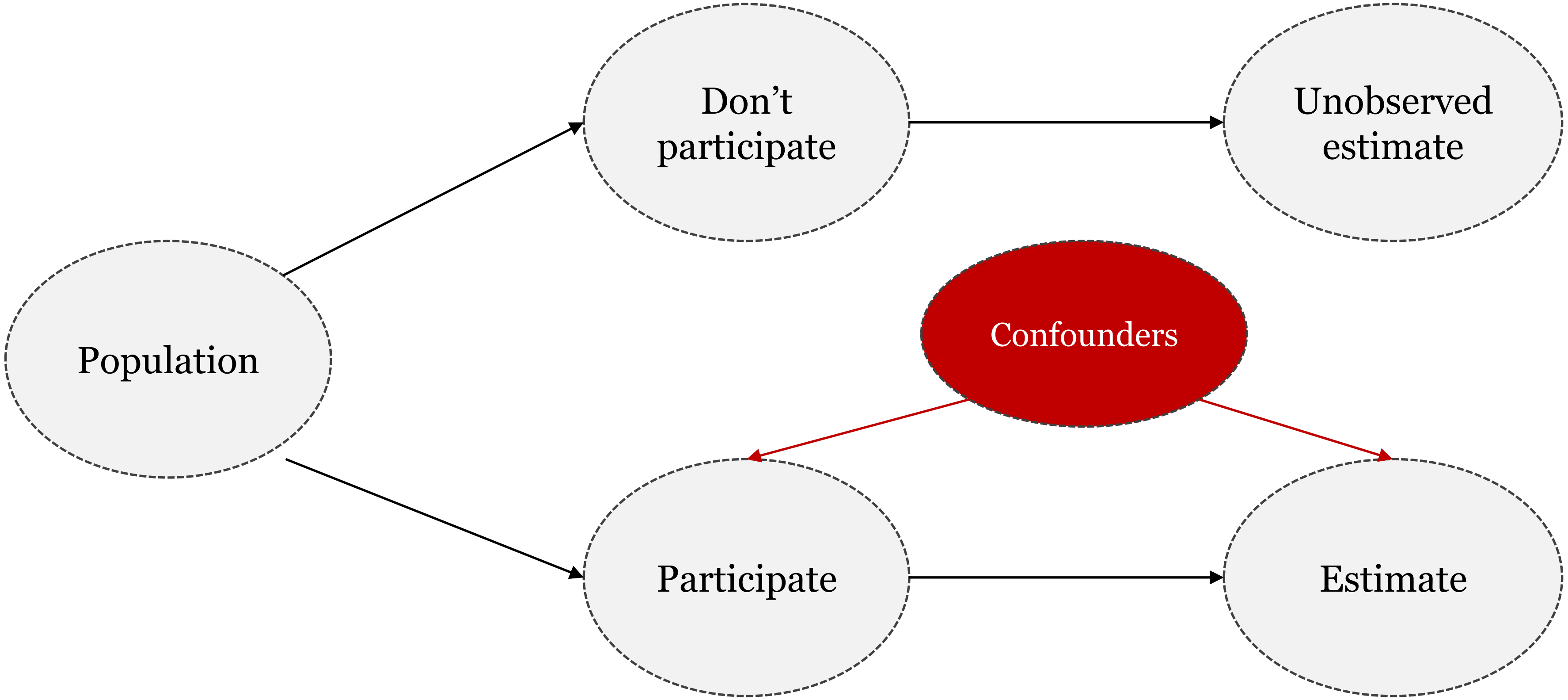
Change of paradigm: **sampling design through the lenses of causal inference**



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.

Second Step: Sampling designs

Change of paradigm: **sampling design through the lenses of causal inference**

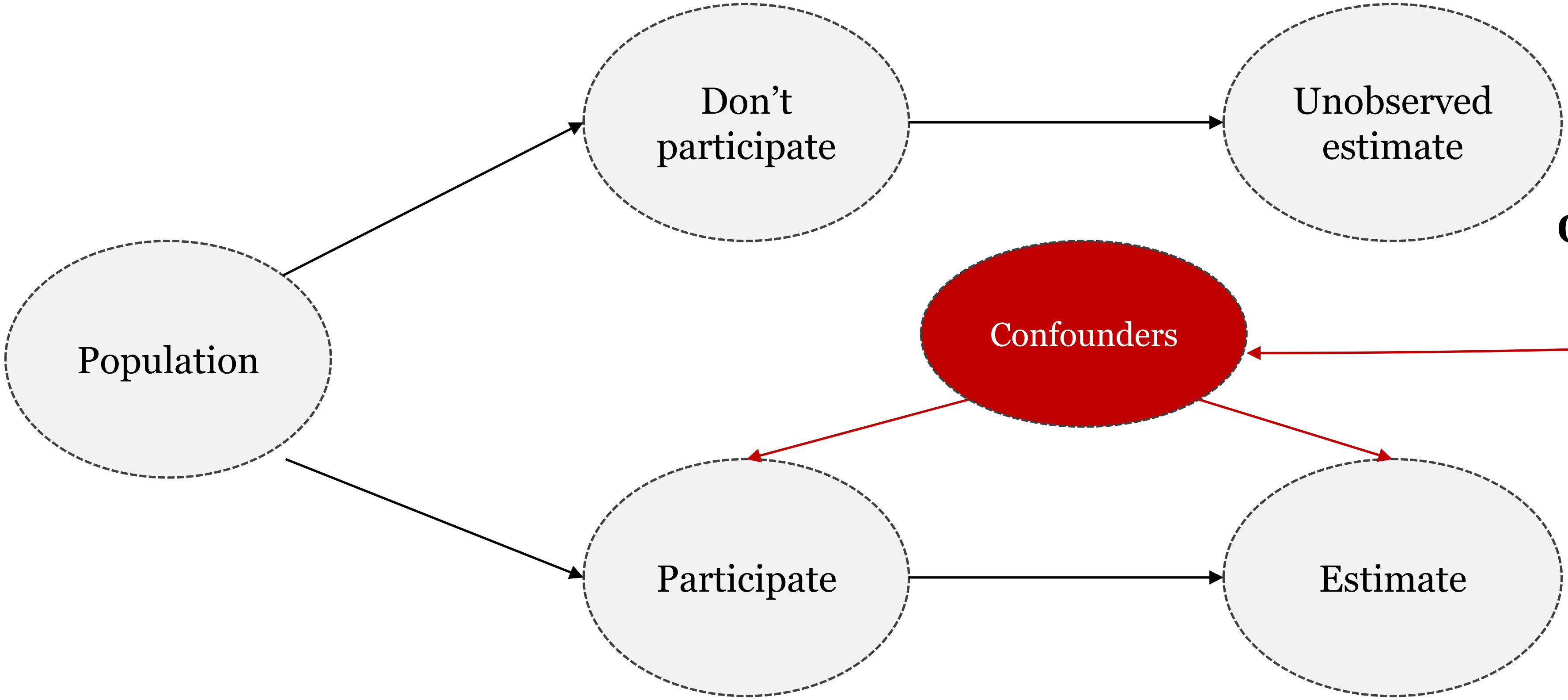


Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). **Theory and practice in nonprobability surveys: parallels between causal inference and survey inference**. *Public Opinion Quarterly*, *81*(S1), 250-271.

Second Step: Sampling designs

web data opp

Change of paradigm: **sampling design through the lenses of causal inference**



Our design needs to account for this

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.

Second Step: Sampling designs

Fit for purpose design

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., ... & Wenz, A. (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.

Second Step: Sampling designs

Fit for purpose design

- The sampling design **does not need to produce a snapshot** of the population
- It only needs to **mitigate any bias that the confounders** might introduce

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., ... & Wenz, A. (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.

Second Step: Sampling designs

Fit for purpose design

- The sampling design **does not need to produce a snapshot** of the population
- It only needs to **mitigate any bias that the confounders** might introduce

Most common approach: **Quota sampling**

Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., De Leeuw, E. D., Legleye, S., ... & Wenz, A. (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.

Second Step: Sampling designs

Quota sampling → Sampling approach that **matches the distribution of a given variable in the sample** with the actual **population distribution**

Second Step: Sampling designs

Quota sampling → Sampling approach that **matches the distribution of a given variable in the sample** with the actual **population distribution**

Desired sample size: 1,000
Distribution of gender in population: 50% male, 48% female, 2% other
Gender in the sample: 500 males, 480 females, 20 others

Second Step: Sampling designs

web data opp

Cross quota sampling → This can get more complex when quotas are crossed

	Male	**Female**
White	350 (35%)	300 (30%)
Non-white	200 (20%)	150 (15%)

Not only about marginal distributions, but also about how the individuals in the subgroups represents the population subgroups

Third Step: Adjustment approach

Even the **best design might not be able to correct** for **all** the biases: we need some modelling

Third Step: Adjustment approach

Even the **best design might not be able to correct** for **all** the biases: we need some modelling

The general logic: statistical models to correct the estimates through weights that re-balance the estimates towards the population (in terms of the confounders).

Third Step: Adjustment approach

Example: **Raking**

Third Step: Adjustment approach

Example: **Raking**

- Choose a set of variables where **the population distribution is known**

Third Step: Adjustment approach

Example: **Raking**

- Choose a set of variables where **the population distribution is known**
- Iteratively **adjust the weight for each case** until the **sample distribution aligns with the population** for those variables

Third Step: Adjustment approach

Example: **Raking**

- Choose a set of variables where **the population distribution is known**
- Iteratively **adjust the weight for each case** until the **sample distribution aligns with the population** for those variables

- **Sample should be:** 48% male and 52% female, and 40% with a high school education or less, 31% who have completed some college, and 29% college graduates
- **First step:** adjust the weights so that gender ratio for the weighted survey sample matches the desired population distribution
- **Second step:** weights are adjusted so that the education groups are in the correct proportion
- **Third step:** If the adjustment for education pushes the sex distribution out of alignment, then the weights are adjusted
- **Etc** until he weighted distribution of all of the **weighting variables matches their specified targets**.

Third Step: Adjustment approach

- Many more complex methods exist, like **propensity weighting and matching**, which can even be **combined**!

Third Step: Adjustment approach

- Many more complex methods exist, like **propensity weighting and matching**, which can even be **combined**!

- But the evidence seems to suggest that the **key part is to choose the right variables**, with **simple models performing similarly** as more complex ones!


Pew Research Center

FOR RELEASE JANUARY 26, 2018

For Weighting Online Opt-In Samples, What Matters Most?

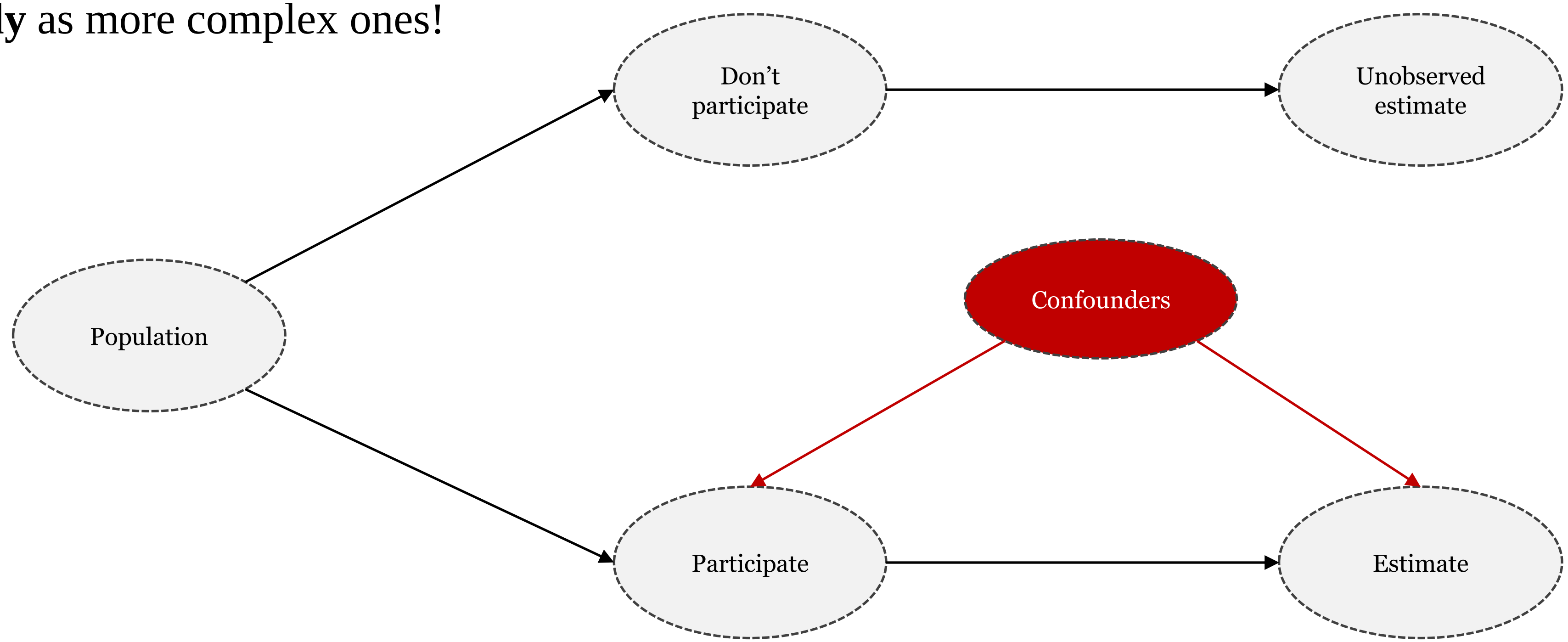
The right variables make a big difference for accuracy. Complex statistical methods, not so much

BY *Andrew Mercer, Arnold Lau, and Courtney Kennedy*

Third Step: Adjustment approach

- Many more complex methods exist, like **propensity weighting and matching**, which can even be **combined**!

- But the evidence seems to suggest that the **key part is to choose the right variables**, with **simple models performing similarly** as more complex ones!



Third Step: Adjustment approach

- Many more complex methods exist, like **propensity weighting and matching**, which can even be **combined**!

- But the evidence seems to suggest that the **key part is to choose the right variables**, with **simple models performing similarly** as more complex ones!

However, new advancements are proposed every year…maybe more complex methods will make a difference eventually

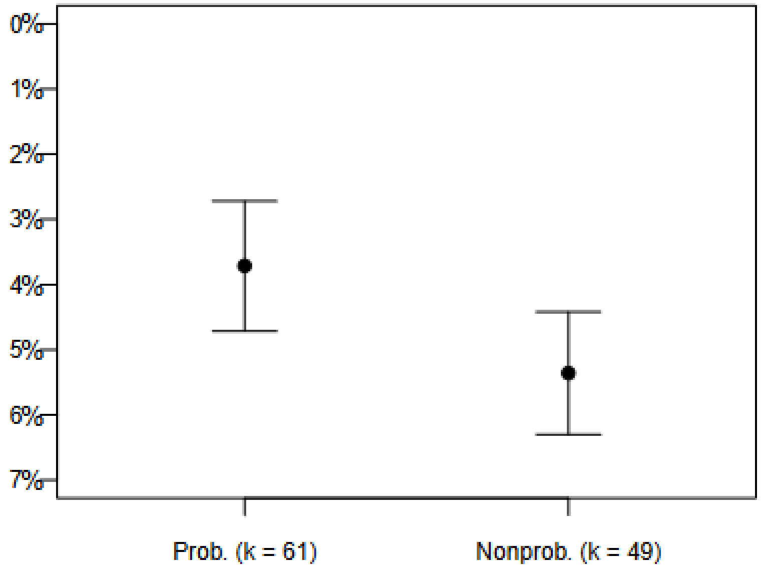
Food for thought: combining small probability samples with big nonprobability ones

So are online nonprobability surveys any good?

Online nonprobability surveys are generally less representative

webdataopp

Most research has found that **probability-based online surveys are more representative**



Mean absolute bias subgroup comparison results by probability versus nonprobability surveys as moderator

Cornesse, C., & Bosnjak, M. (2018, April). **Is there an association between survey characteristics and representativeness? A meta-analysis.** In *Survey Research Methods* (Vol. 12, No. 1, pp. 1-13).

Online nonprobability surveys are generally less representative

Most research has found that **probability-based online surveys are more representative**

And that weighting does not solve this

A REVIEW OF CONCEPTUAL APPROACHES AND EMPIRICAL EVIDENCE ON PROBABILITY AND NONPROBABILITY SAMPLE SURVEY RESEARCH

CARINA CORNESSE*
ANNELIES G. BLOM
DAVID DUTWIN
JON A. KROSNICK
EDITH D. DE LEEUW
STÉPHANE LEGLEYE
JOSH PASEK
DARREN PENNAY
BENJAMIN PHILLIPS
JOSEPH W. SAKSHAUG
BELLA STRUMINSKAYA
ALEXANDER WENZ

Online nonprobability surveys are generally less representative

But results **vary** a lot **depending on the type of survey** (source + sampling + weighting)

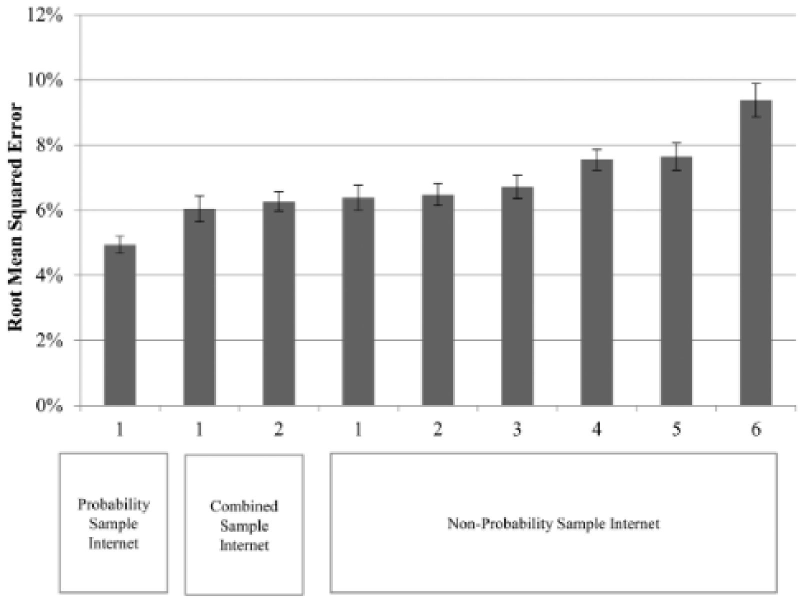


Figure 1. Root mean squared errors for the probability internet sample, the probability plus nonprobability combined samples, and the nonprobability samples across secondary demographics and nondemographics, with our poststratification.

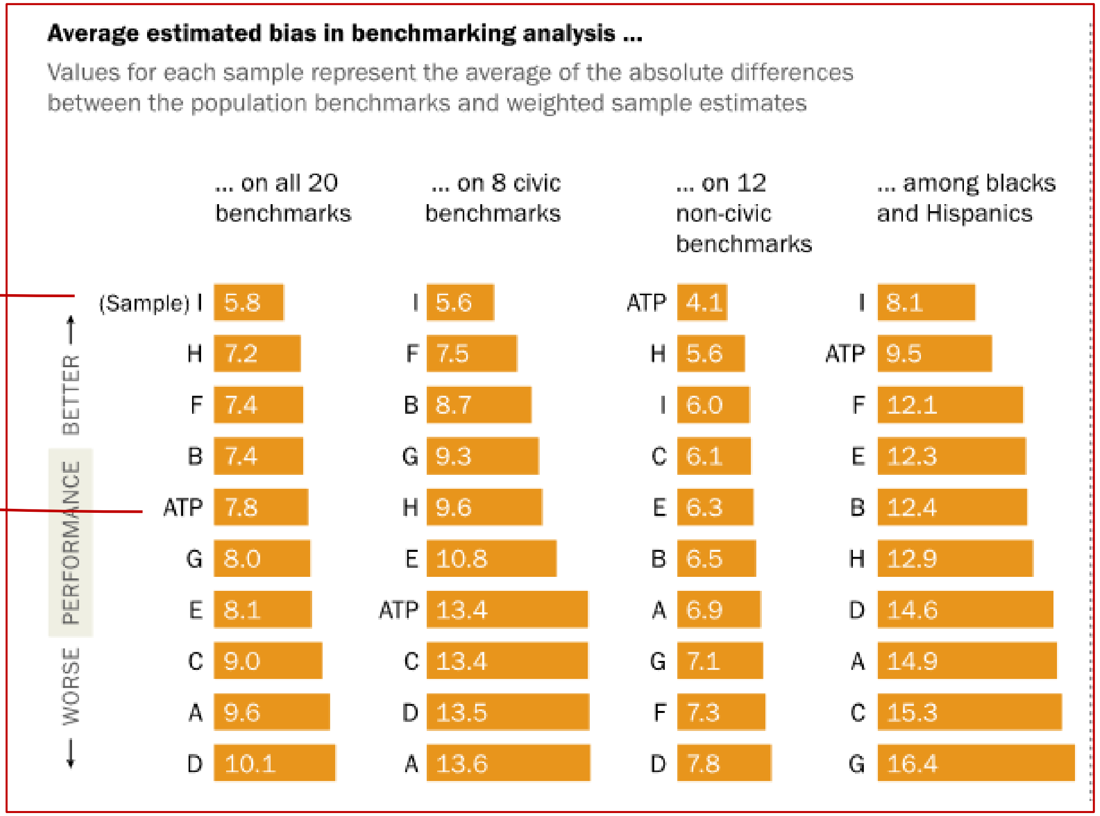
MacInnis, B., Krosnick, J. A., Ho, A. S., & Cho, M. J. (2018). **The accuracy of measurements with probability and nonprobability survey samples: Replication and extension.** *Public Opinion Quarterly*, *82*(4), 707-744.

Online nonprobability surveys are generally less representative

Even sometimes being **better than probability-based online panels**!

YouGov

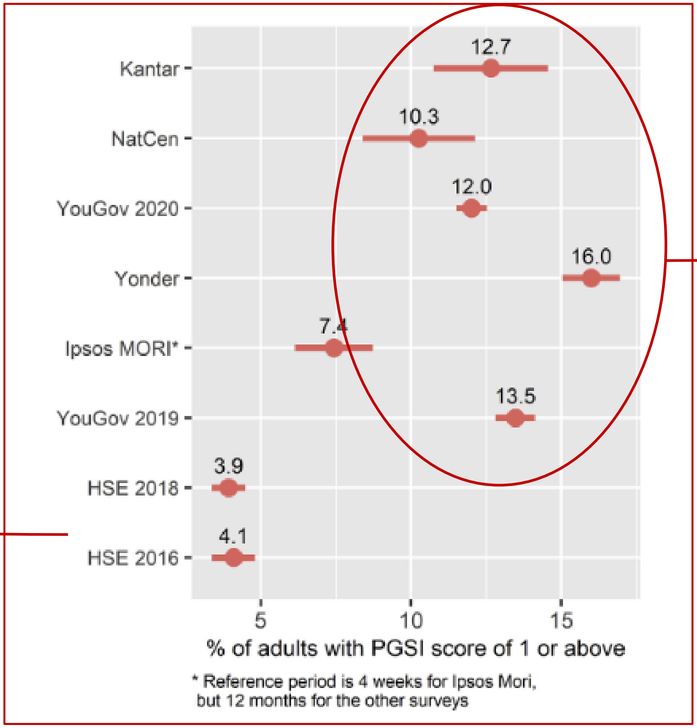
Probability based



Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., & Gimenez, A. (2016). **Evaluating online nonprobability surveys**. *Pew Research Center*, *61*.

Online nonprobability surveys are generally less representative

Sometimes the problem will simply be that…**they are online**, not non-probability!



Prob and nonprob panels

Prob F2F

Sturgis, P., & Kuha, J. (2022). **How survey mode affects estimates of the prevalence of gambling harm: a multisurvey study.** *Public Health*, *204*, 63-69.

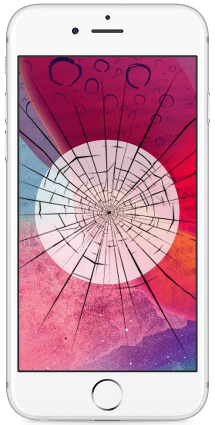
Enhancing nonprobability online surveys

(my research area!)

HOW COULD WE ENHANCE?

Online surveys bring new opportunities

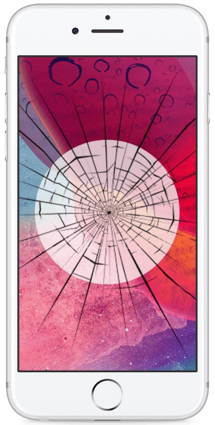
- Online surveys are essentially **multi-device**



Online surveys bring new opportunities

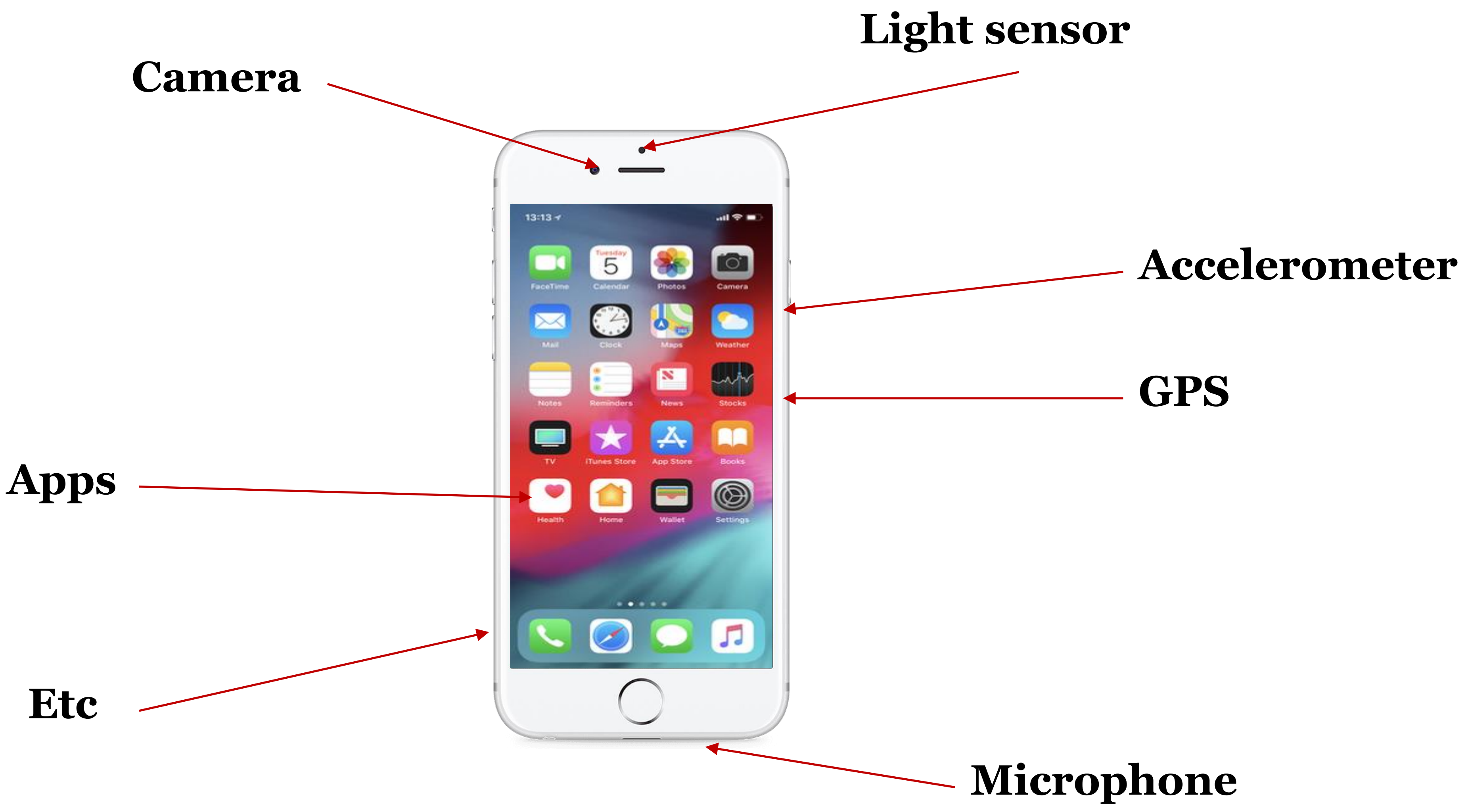
- Online surveys are essentially **multi-device**

- Smartphone usage to answer web surveys:
  - Millennials: 78.8%
  - Boomers: 36.2 %



Bosch, O. J., Revilla, M., & Paura, E. (2019). **Do Millennials differ in terms of survey participation**?. *International Journal of Market Research*, *61*(4), 359-365.

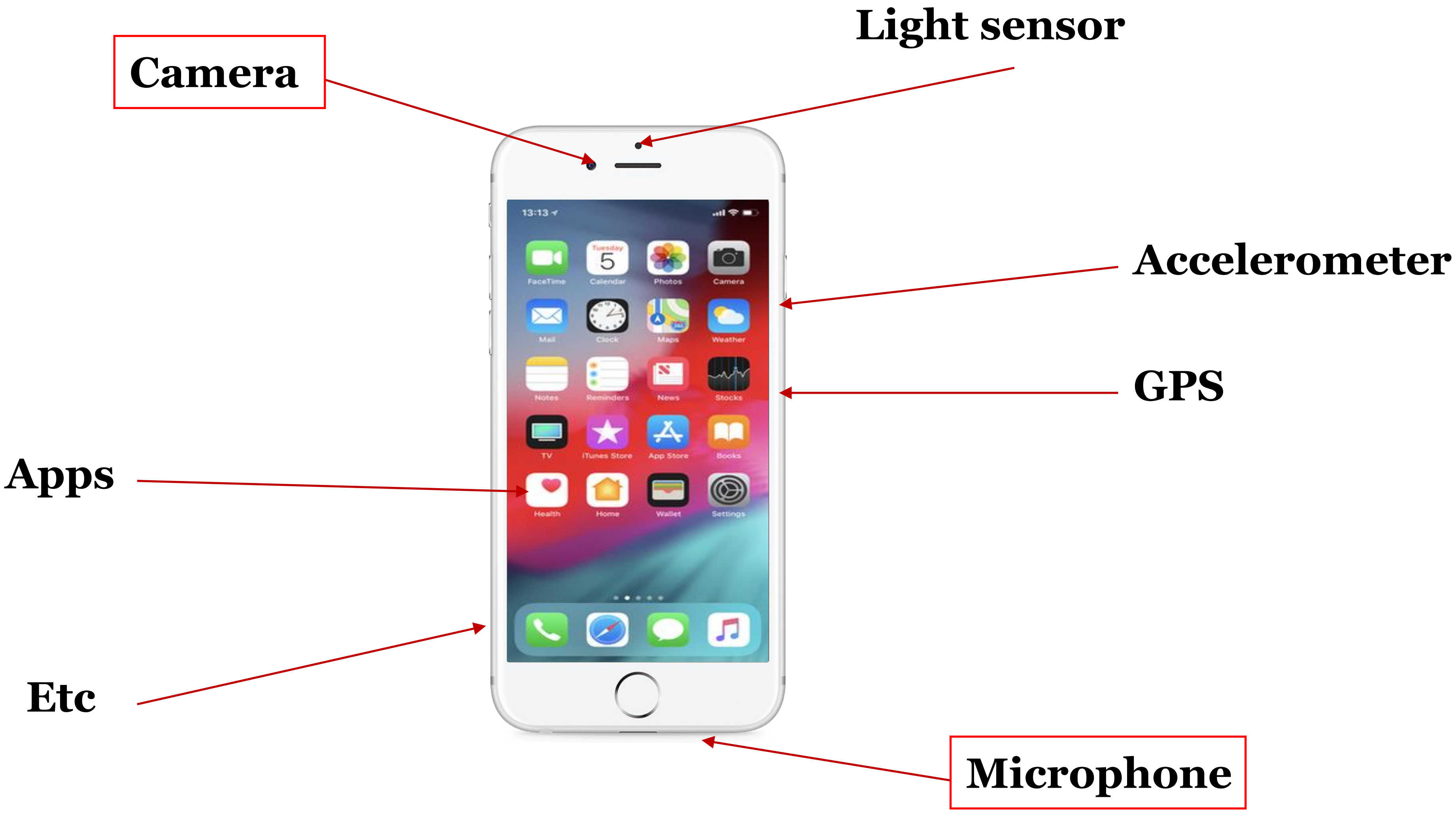
Modern devices are packed with technology that we can use



HOW COULD WE ENHANCE?

Modern devices are packed with technology that we can use

We can ask participants to perform new tasks…



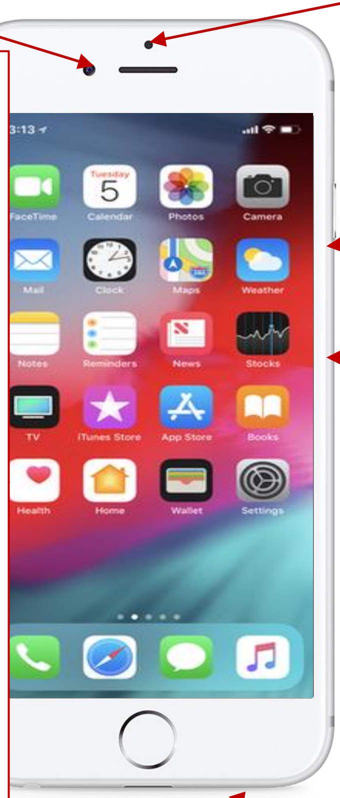
HOW COULD WE ENHANCE?

Modern devices are packed with technology tha

We can ask participants to perform new tasks…

Light s

Camera



Microphone

Article

Social Science Computer Review
2019, Vol. 37(5) 669-683

DOI: 10.1177/0894439318791515
journals.sagepub.com/home/ssc
SAGE

Answering Mobile Surveys With Images: An Exploration Using a Computer Vision API

Oriol J. Bosch[1], Melanie Revilla[1], and Ezequiel Paura[2]

Abstract

Most mobile devices nowadays have a camera. Besides, posting and sharing images have been found as one of the most frequent and engaging Internet activities. However, to our knowledge, no research has explored the feasibility of asking respondents of online surveys to upload images to answer survey questions. The main goal of this article is to investigate the viability of asking respondents of an online opt-in panel to upload during a mobile web survey: First, a photo taken in the moment, and second, an image already saved on their smartphone. In addition, we want to test to what extent the Google Vision application programming interface (API), which can label images into categories, produces similar tags than a human coder. Overall, results from a survey conducted among millennials in Spain and Mexico (N = 1,614) show that more than half of the respondents uploaded an image. Of those, 77.3% and 83.4%, respectively, complied with what the question asked. Moreover, respectively, 52.4% and 65.0% of the images were similarly codified by the Google Vision API and the human coder. In addition, the API codified 1,818 images in less than 5 min, whereas the human coder spent nearly 35 hours to complete the same task.

Keywords

mobile web survey, image recognition, computer vision, API, smartphone, new data types

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Royal Statistical Society
Statistics in Society
Series A

ORIGINAL ARTICLE | Open Access

A new experiment on the use of images to answer web survey questions

Oriol J. Bosch, Melanie Revilla, Danish Daniel Qureshi, Jan Karem Höhne

First published: 20 May 2022 | https://doi.org/10.1111/rssa.12856

Funding information: German Science Foundation, through the Collaborative Research Center 884 "Political Economyof Reforms", 139943784; European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme, 849165

SECTIONS PDF TOOLS SHARE

Abstract

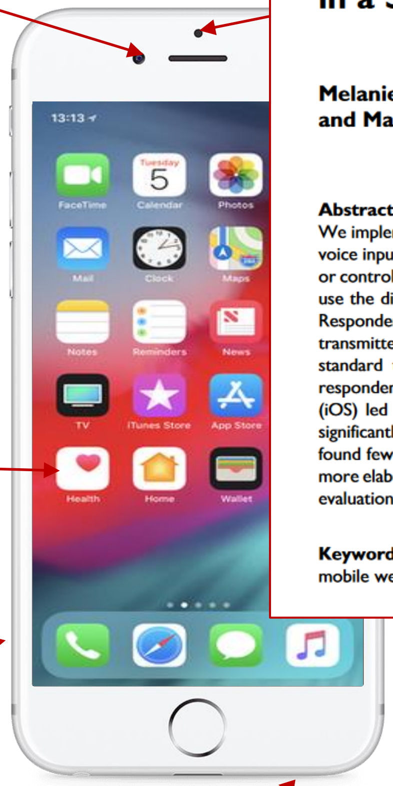
Images might provide richer and more objective information than text answers to open-ended survey questions. Little is known, nonetheless, about the consequences for data quality of asking participants to answer open-ended questions with images. Therefore, this paper addresses three research questions: (1) What is the effect of answering web survey questions with images instead of text on breakoff, noncompliance with the task, completion time and question evaluation? (2) What is the effect of including a motivational message on these four aspects? (3) Does the impact of asking to answer with images instead of text vary across device types? To answer these questions, we implemented a 2 × 3 between-subject web survey experiment (N = 3043) in Germany. Half of the sample was required to answer using PCs and the other half with smartphones. Within each device group, respondents were randomly assigned to (1) a control group answering open-ended questions with text; (2) a treatment group answering open-ended questions with images; and (3) another treatment group answering open-ended questions with images but prompted with a motivational

HOW COULD WE ENHANCE?

web data opp

Modern devices are packed with technology that we can use

We can ask participants to perform new tasks…



Article

Social Science Computer Review
2020, Vol. 38(2) 207-224

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0894439318810715
journals.sagepub.com/home/ssc
SAGE

Testing the Use of Voice Input in a Smartphone Web Survey

Melanie Revilla[1], Mick P. Couper[2], Oriol J. Bosch[1], and Marc Asensio[1]

Abstract

We implemented an experiment within a smartphone web survey to explore the feasibility of using voice input (VI) options. Based on device used, participants were randomly assigned to a treatment or control group. Respondents in the iPhone operating system (iOS) treatment group were asked to use the dictation button, in which the voice was translated automatically into text by the device. Respondents with Android devices were asked to use a VI button which recorded the voice and transmitted the audio file. Both control groups were asked to answer open-ended questions using standard text entry. We found that the use of VI still presents a number of challenges for respondents. Voice recording (Android) led to substantially higher nonresponse, whereas dictation (iOS) led to slightly higher nonresponse, relative to text input. However, completion time was significantly reduced using VI. Among those who provided an answer, when dictation was used, we found fewer valid answers and less information provided, whereas for voice recording, longer and more elaborated answers were obtained. Voice recording (Android) led to significantly lower survey evaluations, but not dictation (iOS).

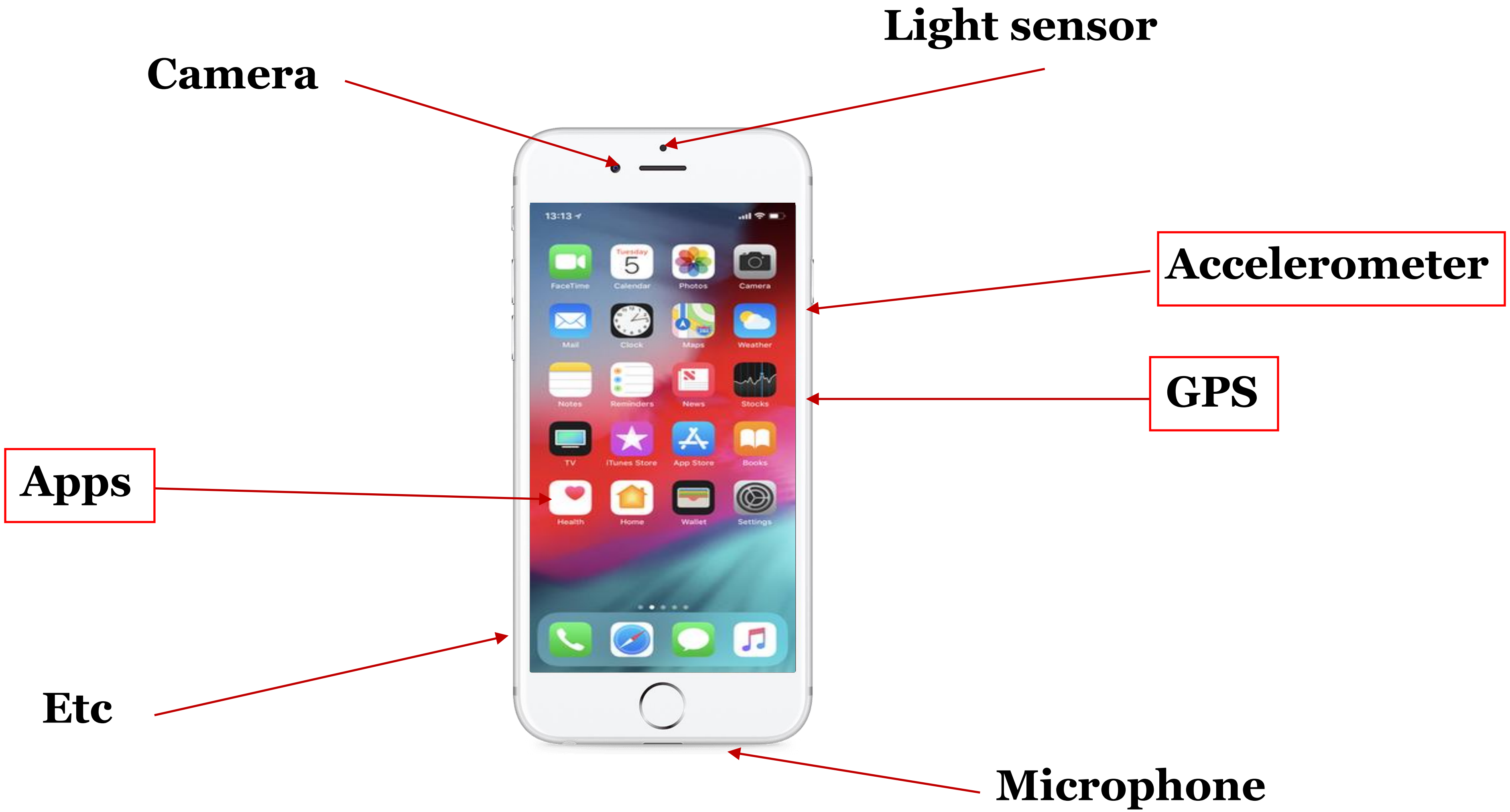
Keywords

mobile web surveys, dictation, voice recording, speech-to-text, data quality

HOW COULD WE ENHANCE?

Modern devices are packed with technology that we can use

…or passively track them



Modern devices are packed with technology that we can use

…or passively track them

Light sensor

Camera

Accelerometer

GPS

Apps

Etc

Microphone

ORIGINAL ARTICLE

Journal of the Royal Statistical Society

When survey science met web tracking: presenting an error framework for metered data

Oriol J. Bosch[1] | Melanie Revilla[2]

[1]The London School of Economics and Political Science

[2]Research and Expertise Centre for Survey Methodology (RECSM), Universitat Pompeu Fabra

Correspondence
Oriol J. Bosch, Department of Methodology, The London School of Economics and Political Science, London, WC2B 4RR, United Kingdom
Email: o.bosch-jover@lse.ac.uk

Funding information
European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 849165)

Metered data, also called "web-tracking data", are generally collected from a sample of participants who willingly install or configure, onto their devices, technologies that track digital traces left when people go online (e.g., URLs visited). Since metered data allow for the observation of online behaviours unobtrusively, it has been proposed as a useful tool to understand what people do online and what impacts this might have on online and offline phenomena. It is crucial, nevertheless, to understand its limitations. Although some research has explored the potential errors of metered data, a systematic categorisation and conceptualisation of these errors are missing. Inspired by the Total Survey Error, we present a Total Error framework for digital traces collected with Meters (TEM). The TEM framework (1) describes the data generation and the analysis process for metered data and (2) documents the sources of bias and variance that may arise in each step of this process. Furthermore, using a case study, we show how the TEM can be applied in real life to identify, quantify and reduce metered data errors. This framework can help improve the quality of both stand-alone metered data research projects, as well as foster the understanding of how and when survey and metered data can be combined.

Why using apps and sensors for survey research?

Researchers

- Reduce measurement issues (e.g. objective)
- Provide new data
- Massive and granular
- Real time

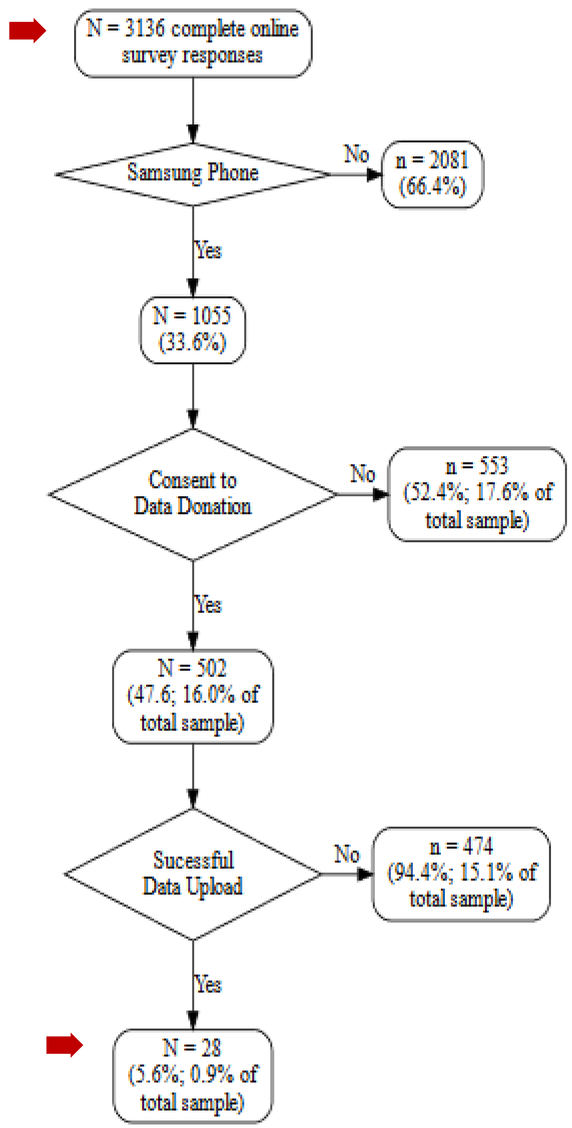
Participants

- Reduce time
- Reduce efforts
- More enjoyable

But expected disadvantages as well

Selection bias in who participates

- Privacy issues
- Technical limitations
- Lack of skills



Silber, H., Keusch, F., Breuer, J., Siegers, P., Beuthner, C., Stier, S., ... & Weiß, B. (2021). **Linking surveys and digital trace data: Insights from two studies on determinants of data sharing behavior.** *SocArXiv Papers*.

But expected disadvantages as well

Selection bias in who participates

- Privacy issues
- Technical limitations
- Lack of skills

New types of errors of measurement

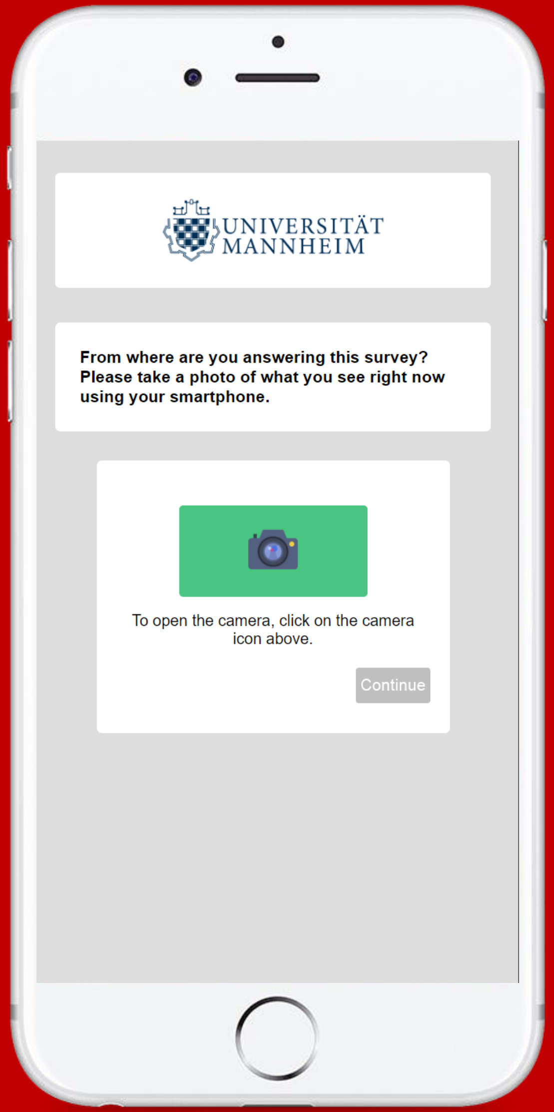
- Technology errors
- Coding and processing errors
- Device-related errors

Error components	Specific error causes
Specification errors	- Defining what qualifies as valid information
	- Measuring concepts with by-design missing data
	- Inferring attitudes and opinions from behaviours
Measurement errors	- Tracking undercoverage
	- Technology limitations
	- Technology errors
	- Hidden behaviours
	- Social desirability
	- Extraction errors
	- Misclassifying non-observations
	- Shared devices
Processing errors	- Coding error
	- Aggregation at the domain level
	- Data anonymisation
Coverage errors	- Non-trackable individuals
Sampling errors	- Same error causes as for surveys
Missing data error	- Non-contact
	- Non-consent
	- Tracking undercoverage
	- Technology limitations
	- Technology errors
	- Hidden behaviours
	- Social desirability
	- Extraction errors
	- Misclassifying non-observations
Adjustment errors	- Same error causes as for surveys

TABLE 1 Specific Error Causes for Metered Data by Error Component

Bosch, O. J., & Revilla, M. (2021). **When survey science met online tracking: presenting an error framework for metered data.** RECSM working paper

VISUAL DATA

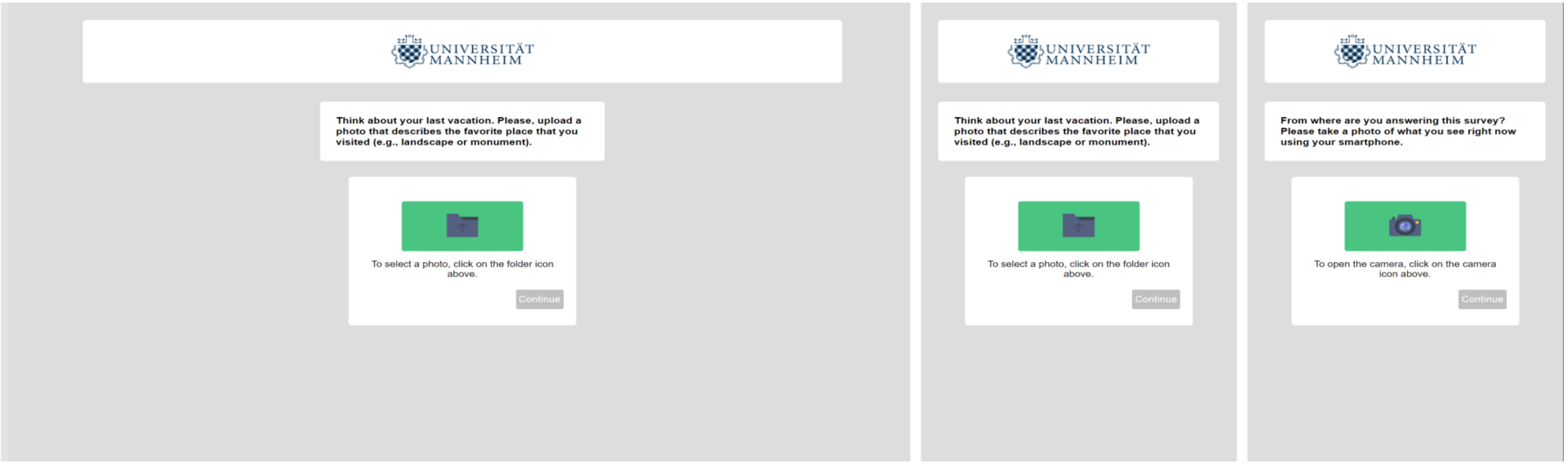

UNIVERSITÄT MANNHEIM
From where are you answering this survey? Please take a photo of what you see right now using your smartphone.
To open the camera, click on the camera icon above.
Continue

Visual data

Bosch, Revilla, Qureshi and Hohne (2022)

Compare

1) Asking to type an answer
2) Asking to send an image
3) Asking to send an image + motivational message



Visual data

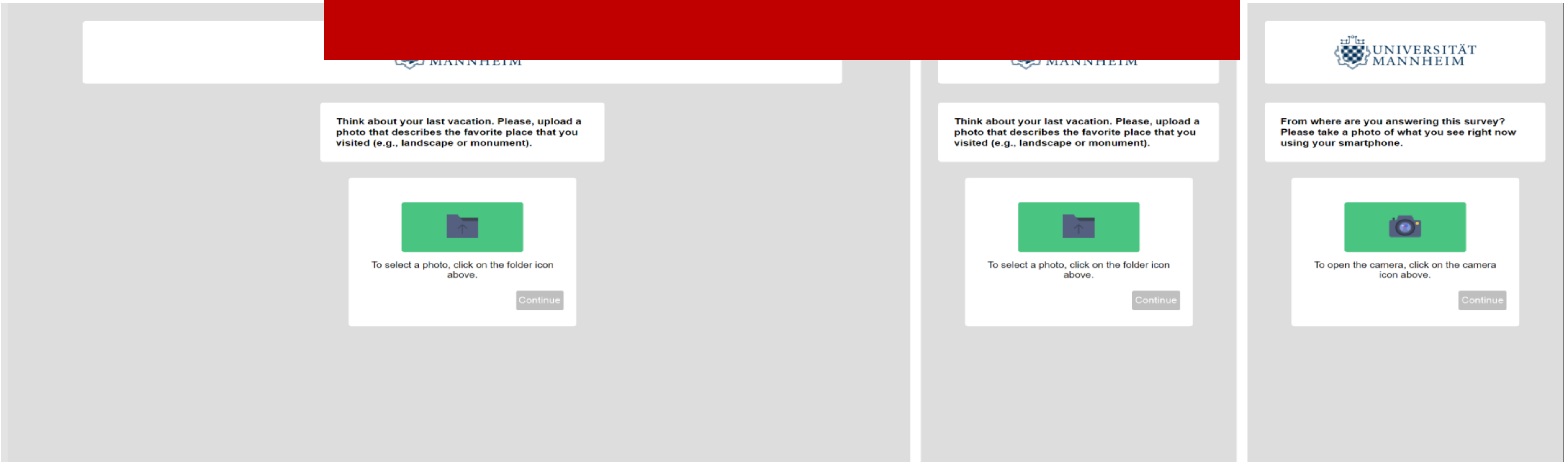
Bosch, Revilla, Qureshi and Hohne (2022)

Compare

1) Asking to type an answ
2) Asking to send an imag
3) Asking to send an imag

What is the impact of asking for images on:

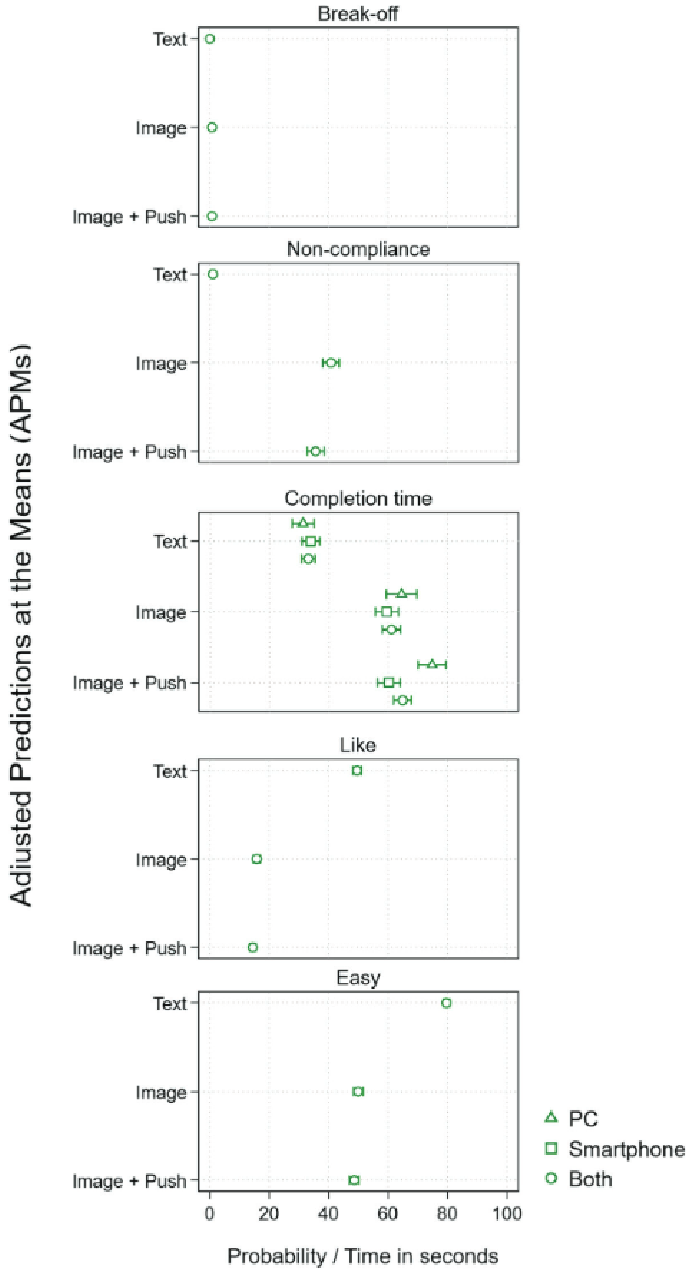
- response rates,
- completion time,
- and question evaluation?



RESULTS

Visual data

Bosch, Revilla, Qureshi and Hohne (2022)



Bosch, Oriol J., et al. **"A new experiment on the use of images to answer web survey questions."** *Journal of the Royal Statistical Society* (2022).

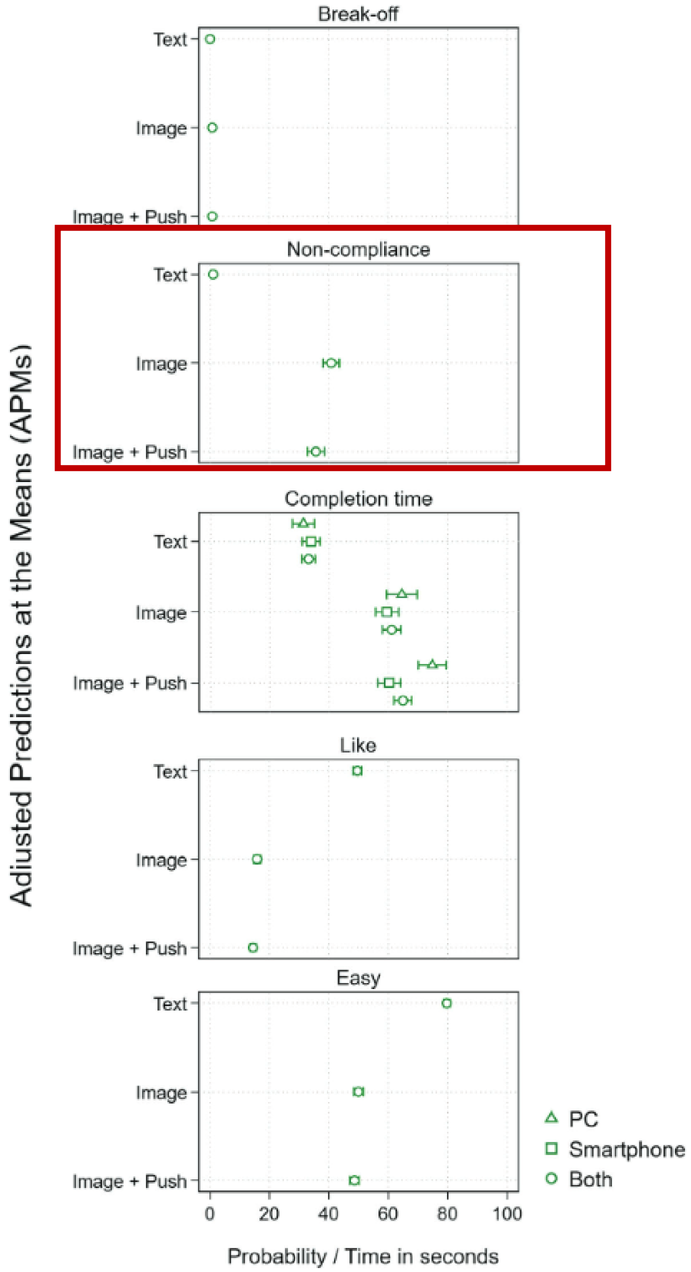
RESULTS

Visual data

Bosch, Revilla, Qureshi and Hohne (2022)

Asking for images:

- **Increases the probability of item nonresponse** (34-39 p.points higher probability)



Bosch, Oriol J., et al. **"A new experiment on the use of images to answer web survey questions."** *Journal of the Royal Statistical Society* (2022).

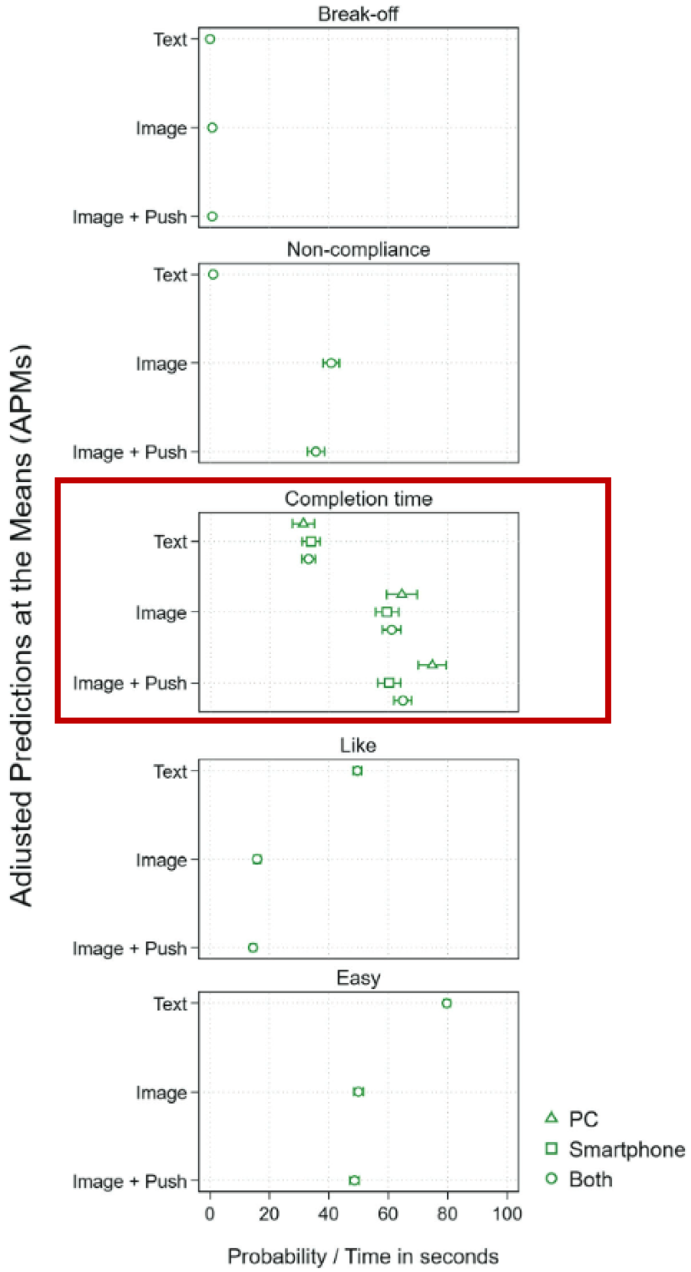
RESULTS

Visual data

Bosch, Revilla, Qureshi and Hohne (2022)

Asking for images:

- **Increases the probability of item nonresponse** (34-39 p.points higher probability)
- **Increases completion times** (25.6 to 43.52 seconds more)



Bosch, Oriol J., et al. **"A new experiment on the use of images to answer web survey questions."** *Journal of the Royal Statistical Society* (2022).

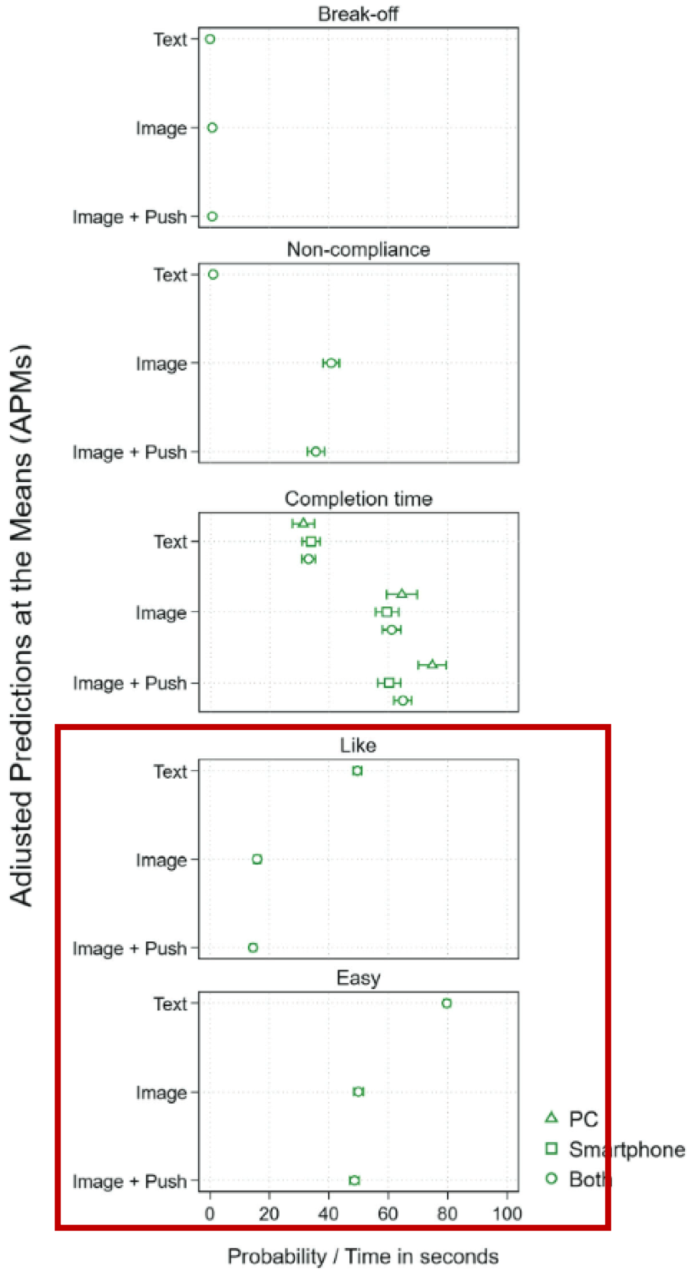
RESULTS

Visual data

Bosch, Revilla, Qureshi and Hohne (2022)

Asking for images:

- **Increases the probability of item nonresponse** (34-39 p.points higher probability)
- **Increases completion times** (25.6 to 43.52 seconds more)
- **Decreases the probability of enjoying and finding questions easy** (~30 p.p lower probability of liking and finding the questions easy)



Bosch, Oriol J., et al. **"A new experiment on the use of images to answer web survey questions."** *Journal of the Royal Statistical Society* (2022).

METERED DATA

Metered data

- It is becoming vital to better understand what people do online and what impact this has on online and offline phenomena.
- Self-reports might not be best suited for this
- Alternative: directly observe what people do online using digital tracking solutions, or *meters.*
  - **Group of tracking technologies**
  - **Installed on participants devices**.
  - **Collect traces** left by participants when **interacting with their devices online: e.g. URLs or apps visited**
- We call the resulting data: **metered data**.

Metered data

- It is becoming vital to better understand what people do online and what impact this has on online and offline phenomena.

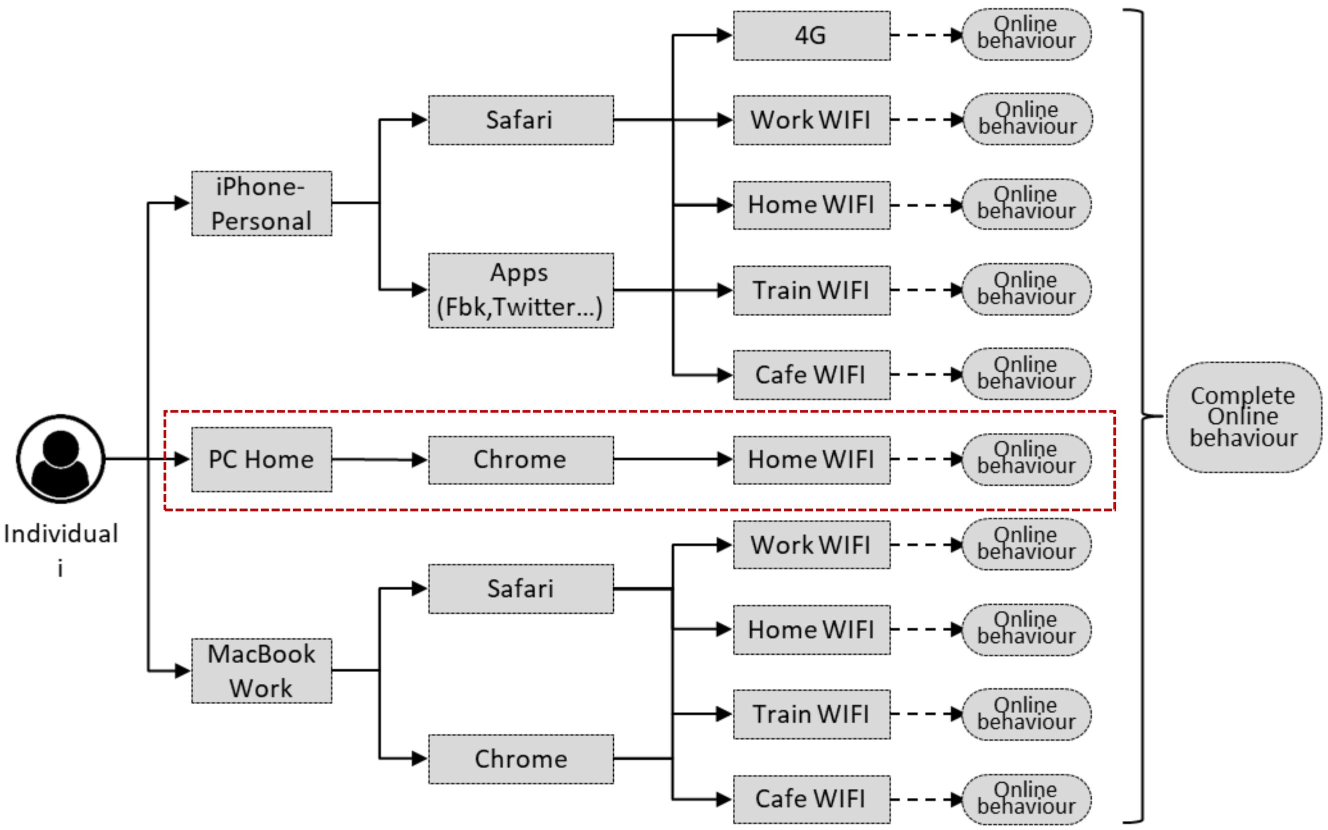
- Self-reports might not be best suited for this

Is metered data actually unbiased?

- Alternative: directly observe what people do online using digital tracking solutions, or *meters*.
  - **Group of tracking technologies**
  - **Installed on participants devices**.
  - **Collect traces** left by participants when **interacting with their devices online: e.g. URLs or apps visited**

- We call the resulting data: **metered data**.

Biases of metered data: tracking undercoverage



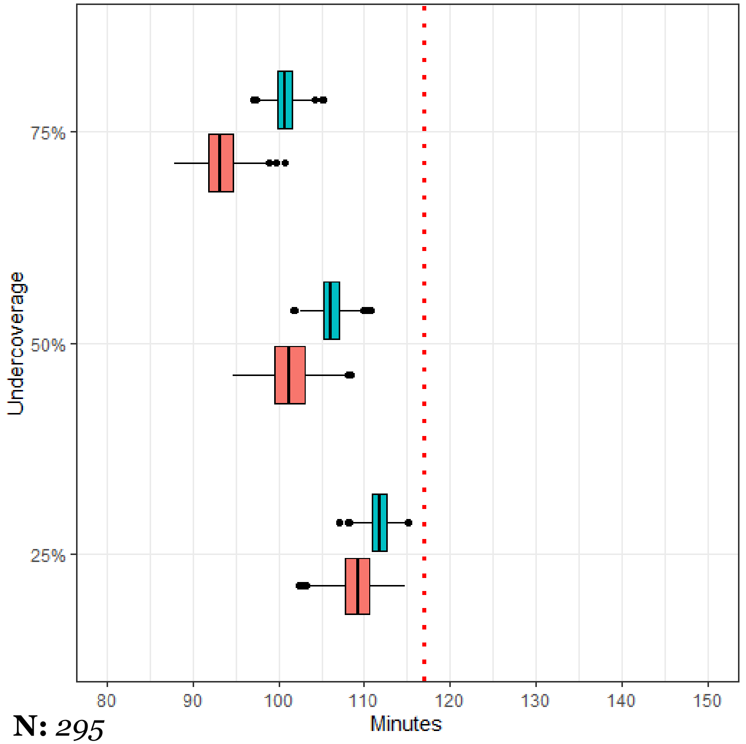
Objective: measuring individuals' behaviours

Reality: vector of those behaviours that individuals' do through all their *targets*

RESULTS

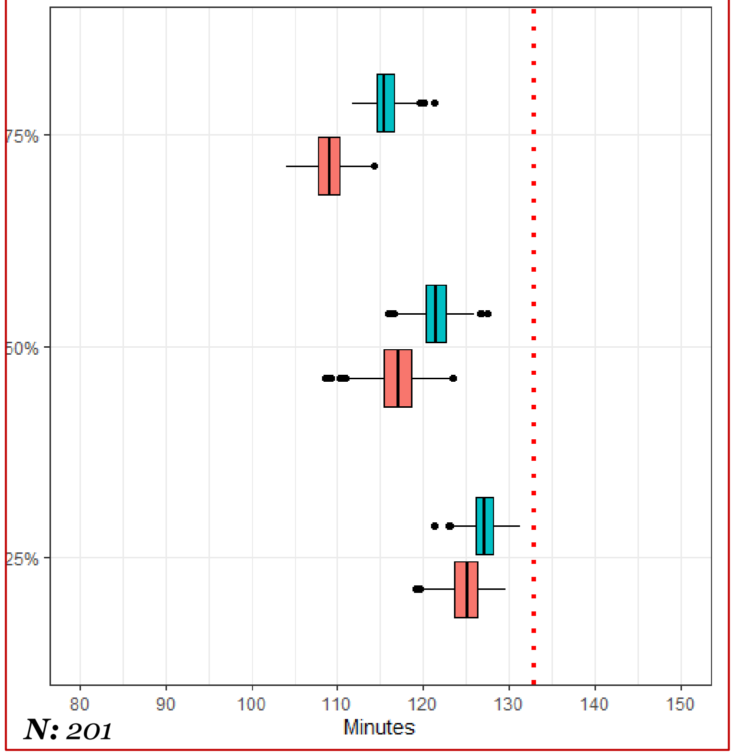
Average time spent on the internet

SPAIN



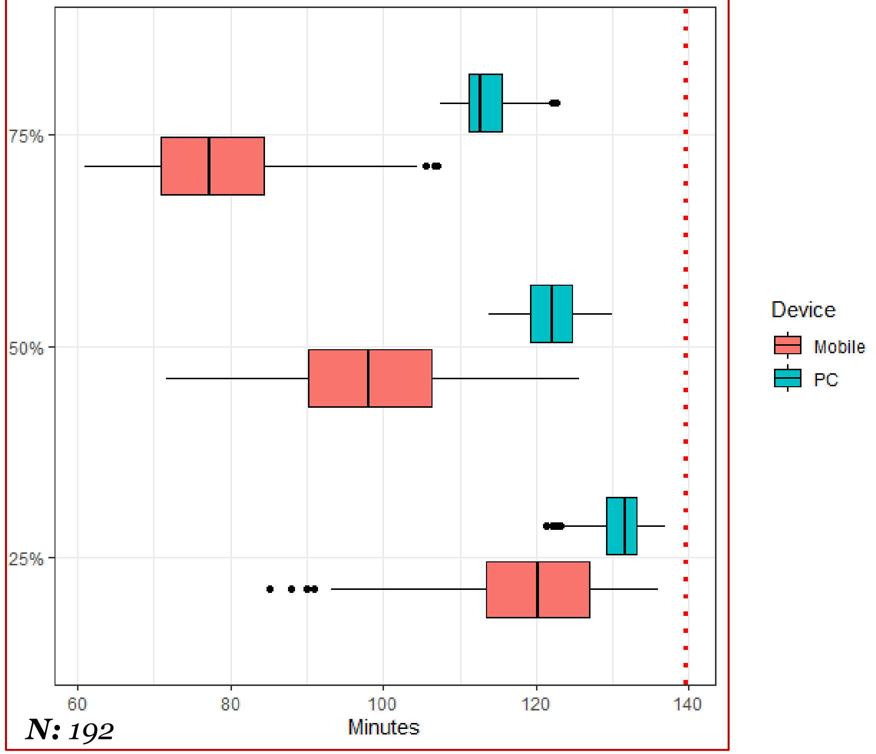
N: *295*

ITALY



N: *201*

PORTUGAL



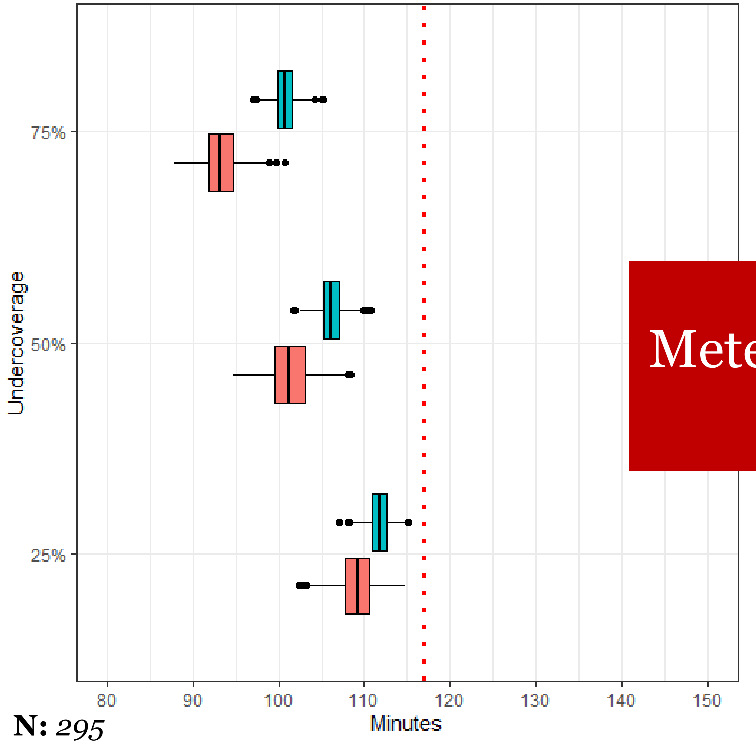
N: *192*

Avg. bias: *5 – 38 minutes* *5 – 23 minutes* *5 – 24 minutes*

RESULTS

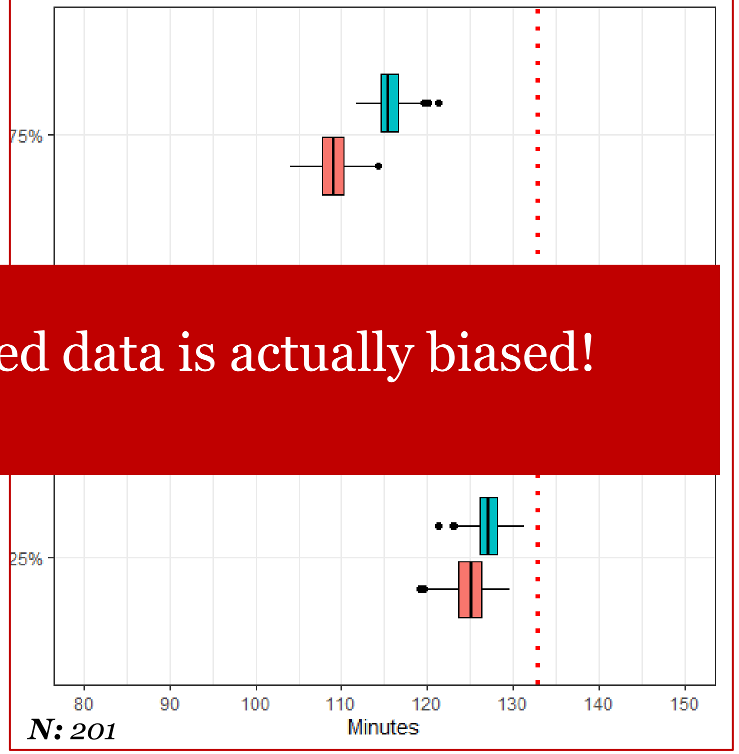
Average time spent on the internet

SPAIN



N: *295*

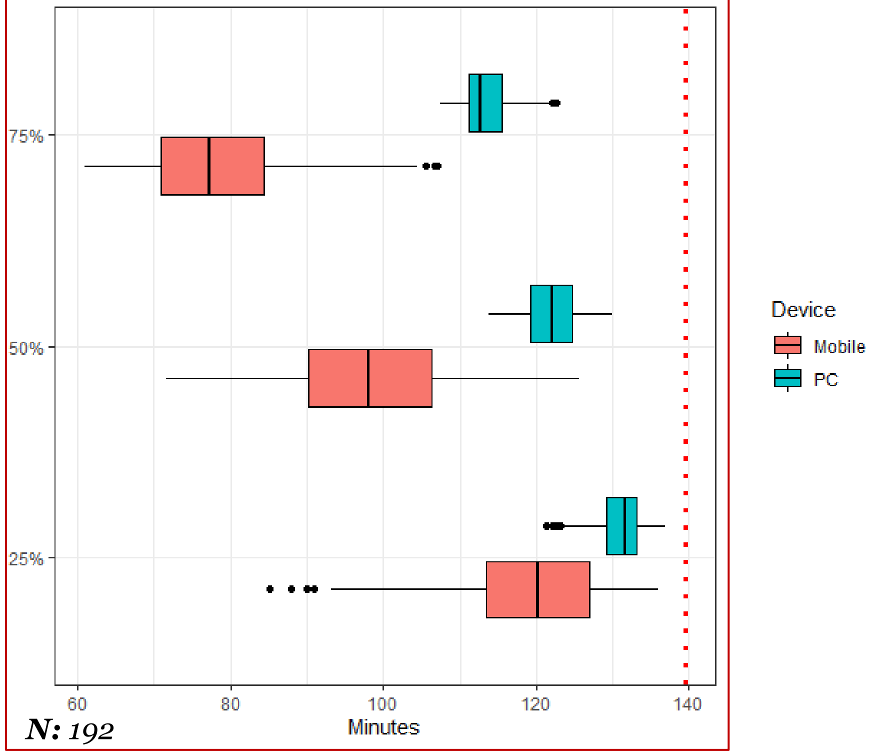
ITALY



N: *201*

Metered data is actually biased!

PORTUGAL



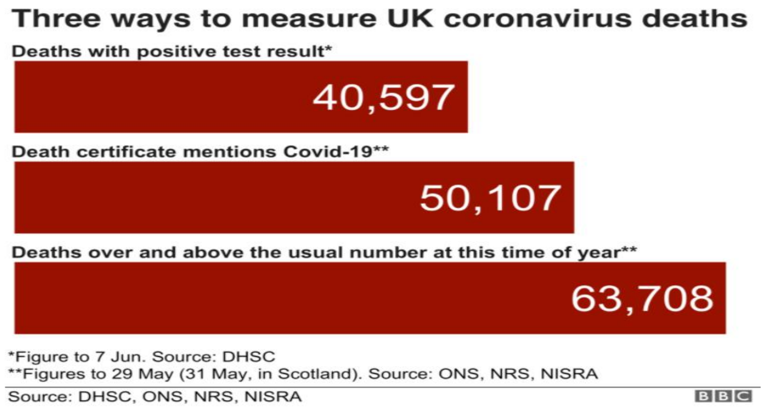
N: *192*

Avg. bias: *5 – 38 minutes* *5 – 23 minutes* *5 – 24 minutes*

Closing remarks

Take-home messages

- Any data collection method suffer from errors
  - This is not just the case of surveys…



- Probably not realistic to aim to perfect measures
  - What we need is to be aware of the errors and their consequences
  - Try to minimize them / correct for them
  - Be careful about not concluding to much!

Thanks!

Questions?

Oriol J. Bosch|PhD Candidate, The London School of Economics

o.bosch-jover@lse.ac.uk

orioljbosch

https://orioljbosch.com/