How can consumer research platforms enhance or extent (mobile) web survey data?

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

o.bosch-jover@lse.ac.uk

orioljbosch

https://orioljbosch.com/

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE ■

upf. **Universitat Pompeu Fabra** *Barcelona*

RECSM Research and Expertise Centre for Survey Methodology

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**
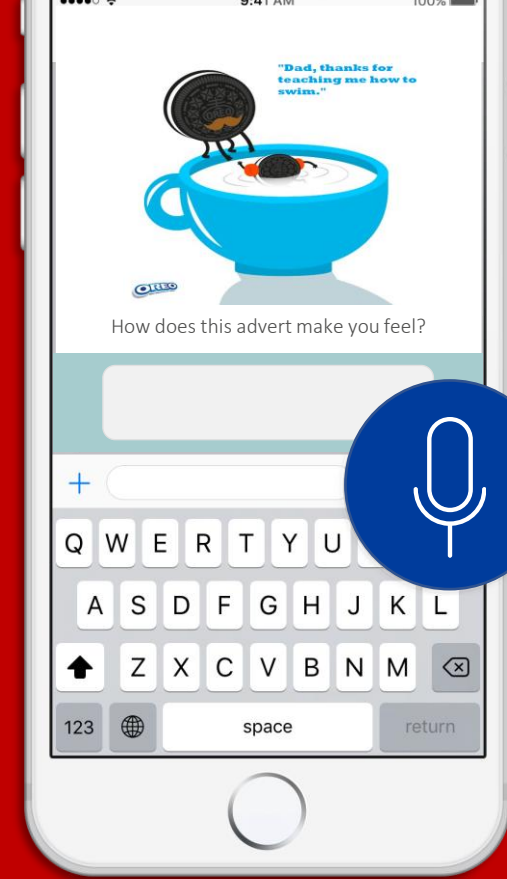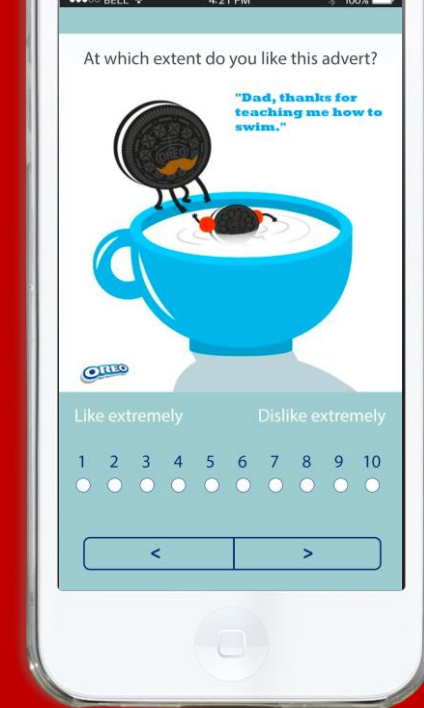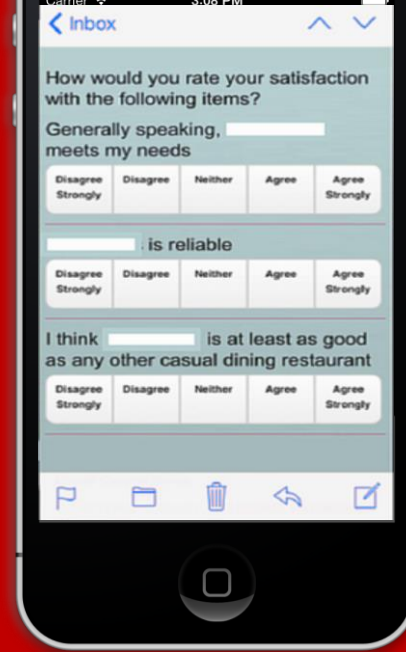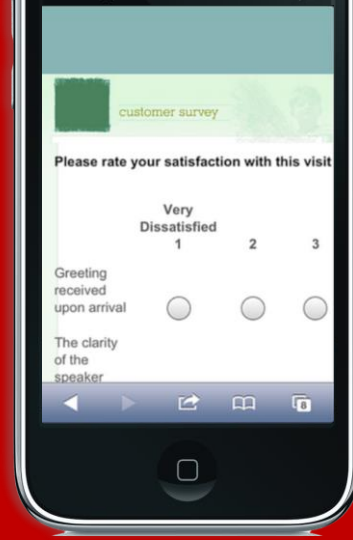
How could we enhance or extend (mobile) web survey data?

Web surveys bring new opportunities

Web surveys bring new opportunities

- Web surveys are essentially **multi-device**

Web surveys bring new opportunities

- Web 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.
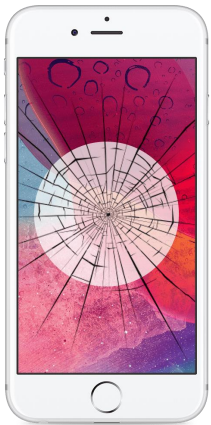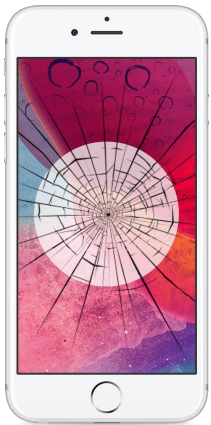
Web surveys bring new opportunities

- Web surveys are essentially **multi-device**

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

How does this compare with your experience?

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.
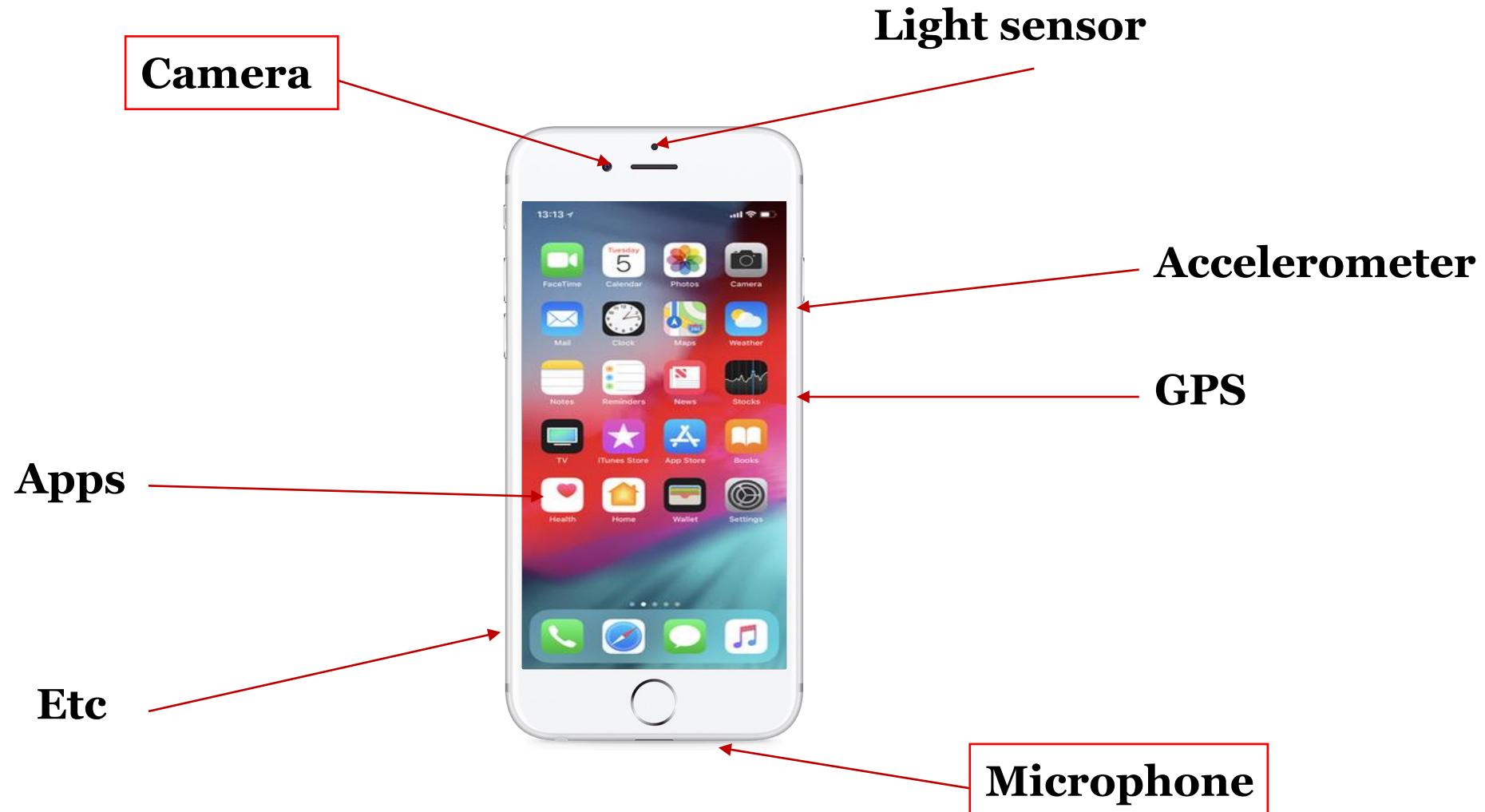
Modern devices are packed with technology tha[...]

We can ask participants to perform new tasks…

Camera

Light s[...]

Article

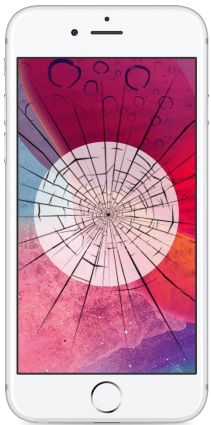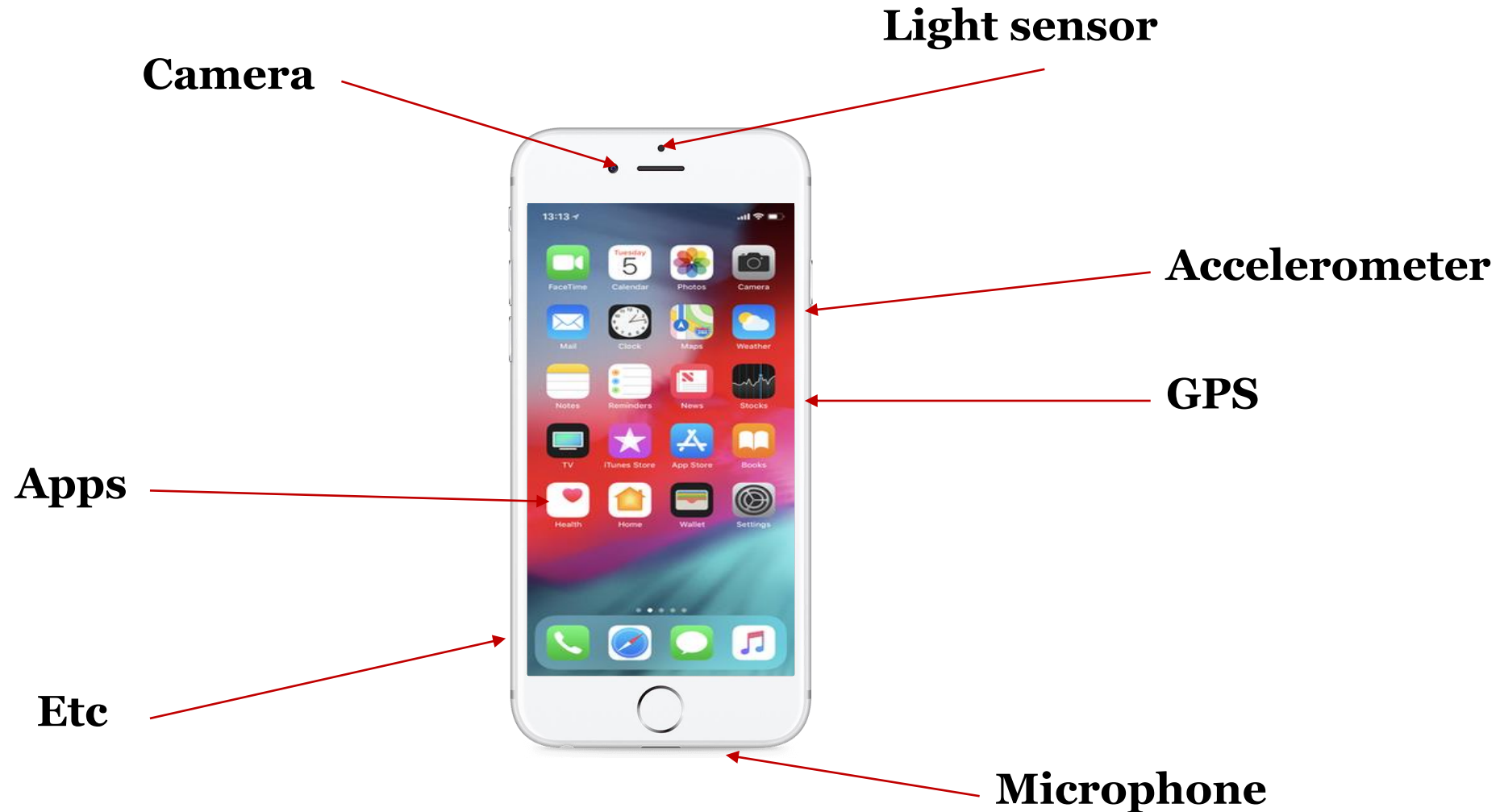**Answering Mobile Surveys With Images: An Exploration Using a Computer Vision API**

Social Science Computer Review
2019, Vol. 37(5) 669-683
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0894439318791515
journals.sagepub.com/home/ssc
⊗SAGE

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 | (cc) (i)

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
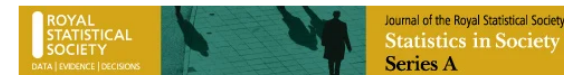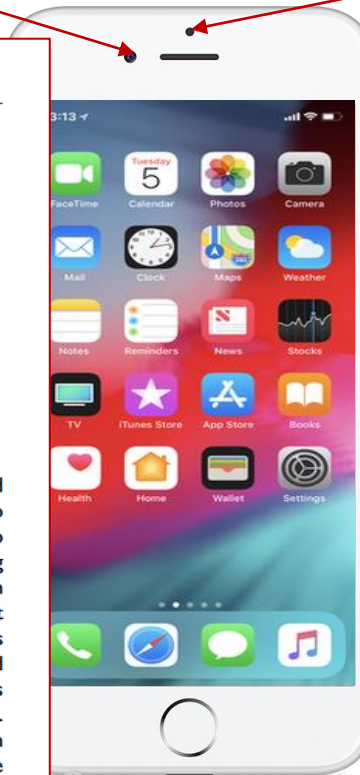
Microphone

Modern devices are packed with technology that we can use

We can ask participants to perform new tasks…

Camera

Apps

Etc

Microphone

eter

Article

Testing the Use of Voice Input in a Smartphone Web Survey

Social Science Computer Review
2020, Vol. 38(2) 207-224
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0894439318810715
journals.sagepub.com/home/ssc

⑤SAGE

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

Modern devices are packed with technology that we can use

…or passively track them

Light sensor

Camera

Accelerometer

GPS

Apps

Etc

Microphone

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.
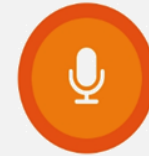
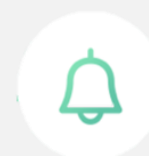# Focus on 4 new types of data collection opportunities
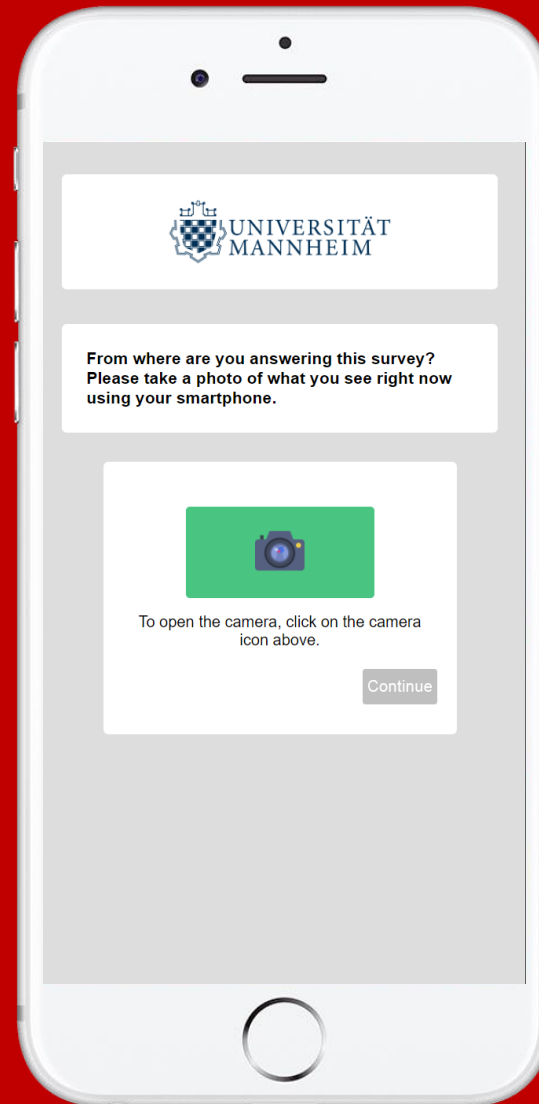
Visual data

Voice data

Metered data

In-the-moment surveys

VISUAL DATA

What kind of visual data can we ask for?
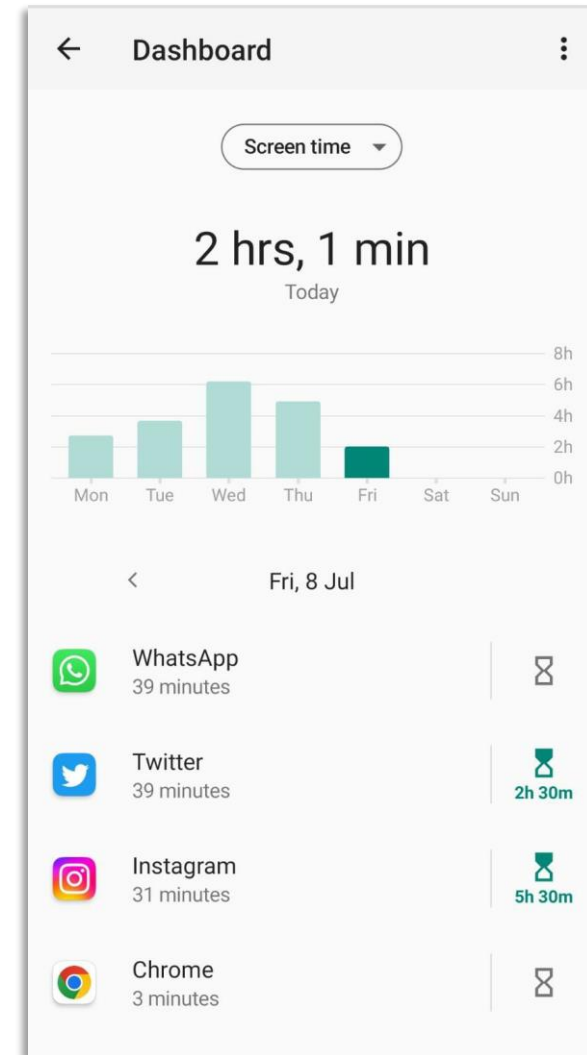
- Photos take in the moment

What kind of visual data can we ask for?

- Photos take in the moment
- Images already saved on the device

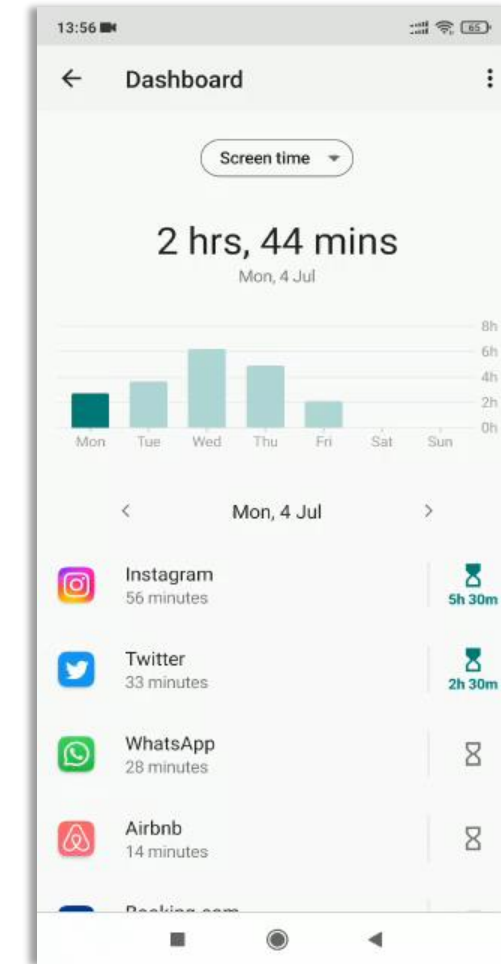# What kind of visual data can we ask for?

- Photos take in the moment
- Images already saved on the device
- Screenshots

What kind of visual data can we ask for?

- Photos take in the moment
- Images already saved on the device
- Screenshots

- Videos taken in the moment / already saved on the device
- Screen recordings

Plus videos!

Why would we want to ask for visual data?

- To obtain **more accurate** information

Why would we want to ask for visual data?

- To obtain **more accurate** information

Instead of relying on self-reports, we can observe the phenomenon of interest: e.g. **expenditure** or **online behavioural data**

Why would we want to ask for visual data?

- To obtain **more accurate** information

Instead of relying on self-reports, we can observe the phenomenon of interest: e.g. **expenditure** or **online behavioural data**

- For **hard / impossible** to collect data with survey questions

Why would we want to ask for visual data?

- To obtain **more accurate** information

Instead of relying on self-reports, we can observe the phenomenon of interest: e.g. **expenditure** or **online behavioural data**

- For **hard / impossible** to collect data with survey questions

Do people have dangerous moles?

# Why would we want to ask for visual data?

- To obtain **more accurate** information

- For **hard / impossible** to collect data with survey questions

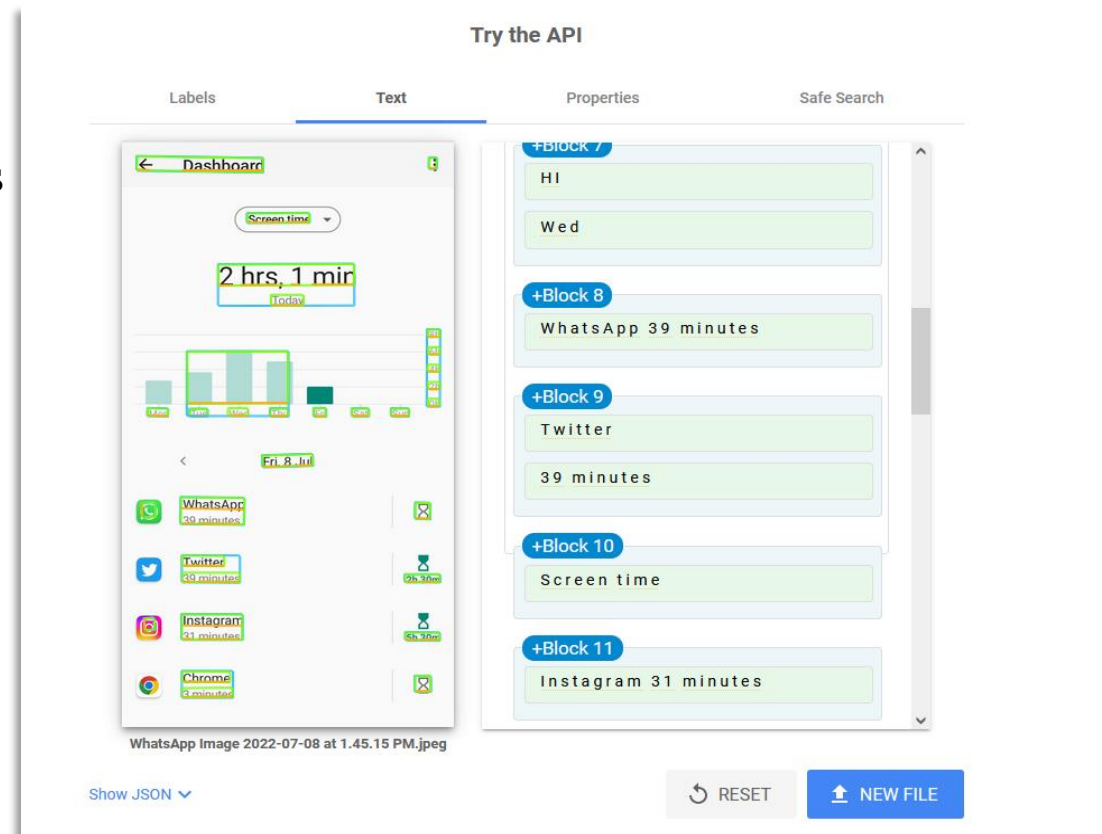- To make the survey experience **more natural and enjoyable**
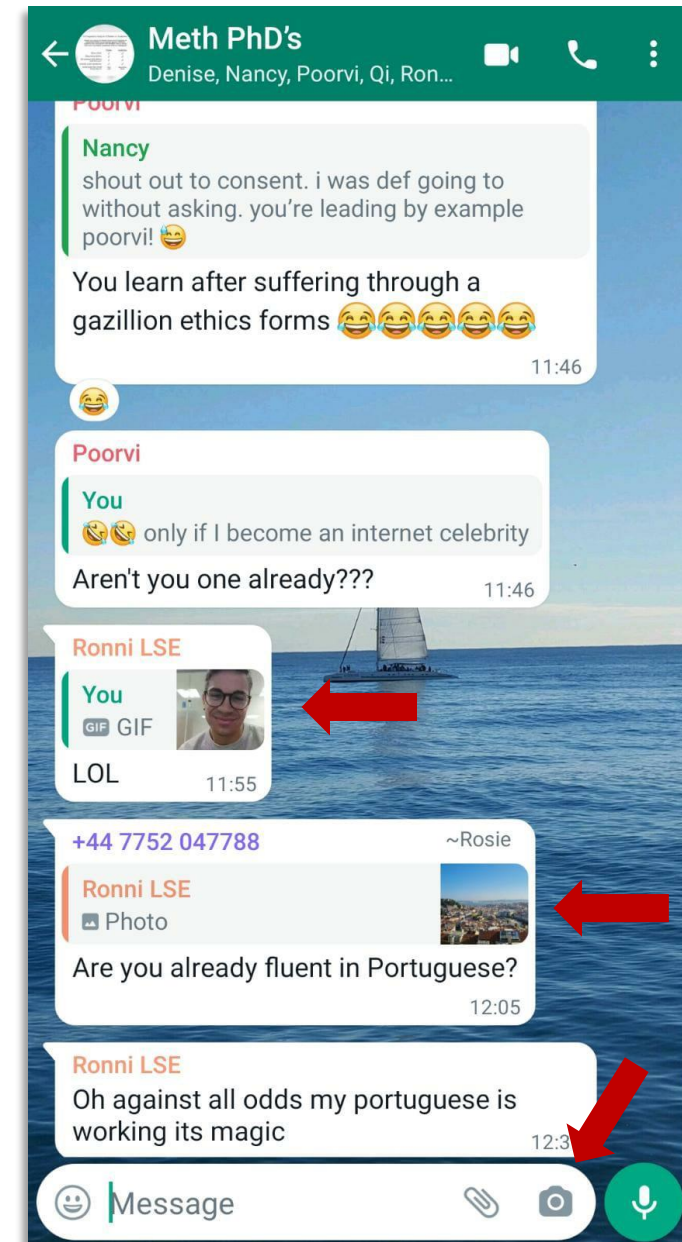
What do you need to ask for visual data?

- An **interface** & **infrastructure** to allow for the collection and storage of visual data: **WebdataVisual**

More information available at: https://www.upf.edu/web/webdataopp/tools
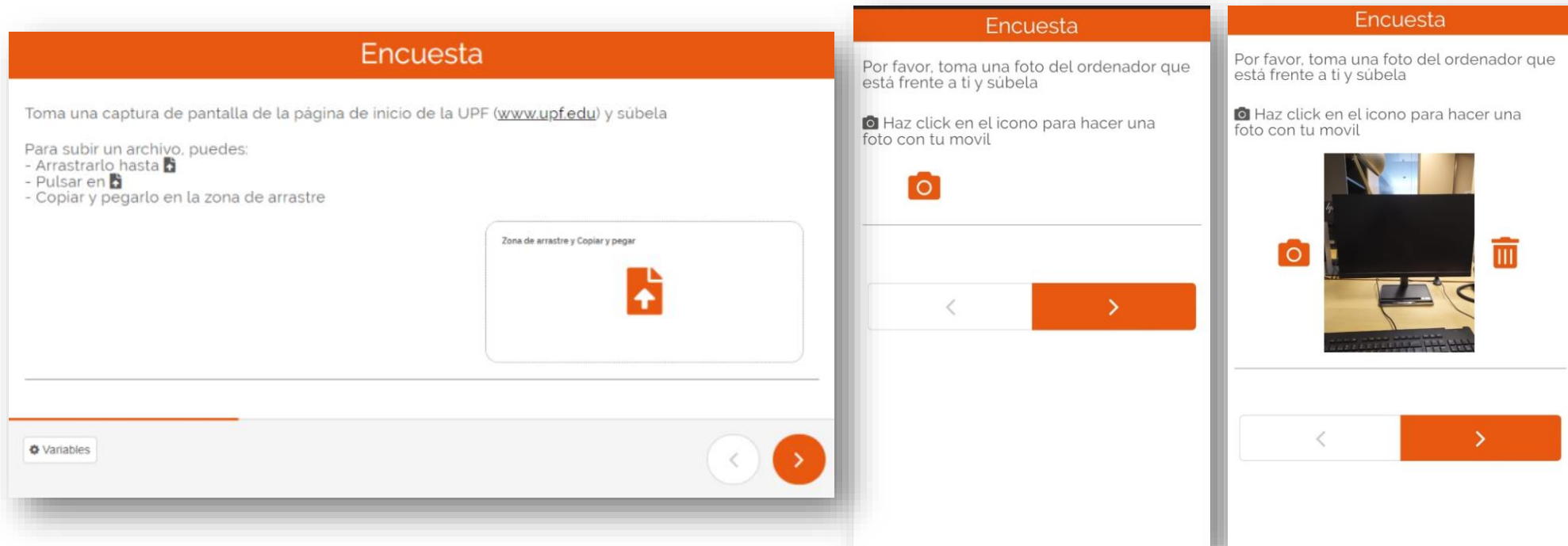
What do you need to ask for visual data?

- An **interface** & **infrastructure** to allow for the collection and storage of visual data: **WebdataVisual**

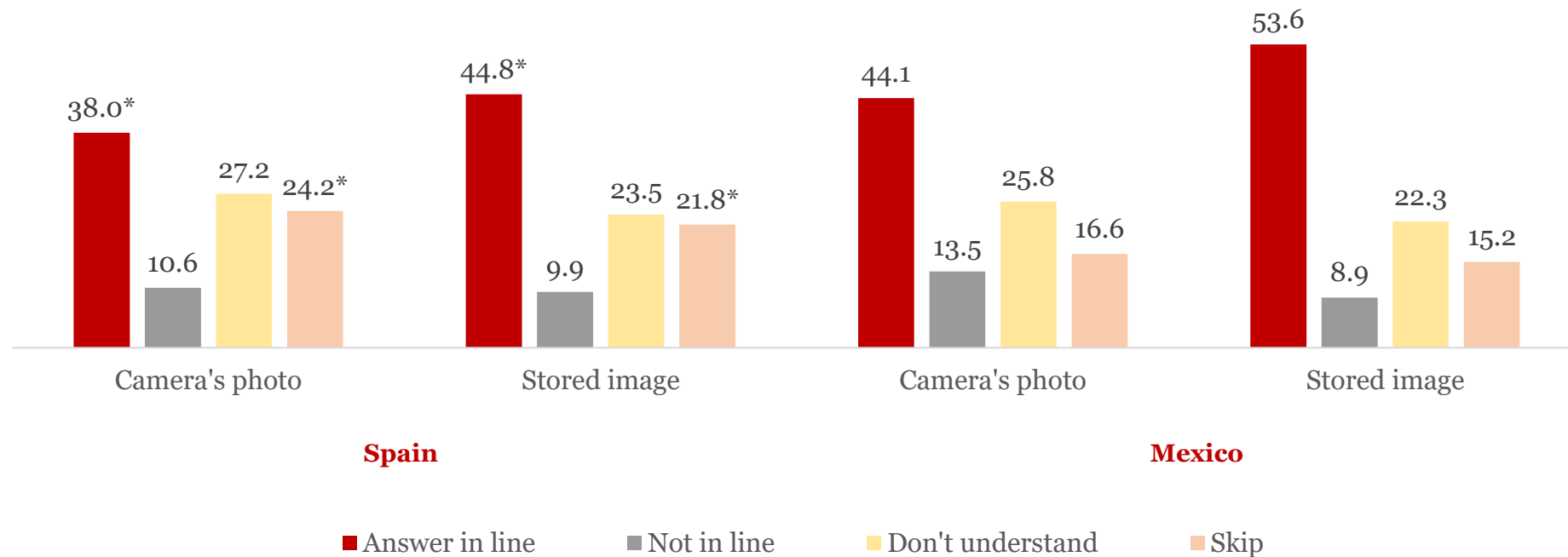- An approach to **extract information** from the images/videos

Google Vision: https://cloud.google.com/vision?hl=en

What have we found in our research?

What have we found in our research?

What is the real willingness of participants?

Percentage of Respondents Answering in Line, Answering not in Line, Not Understanding How to Do It, and Skipping
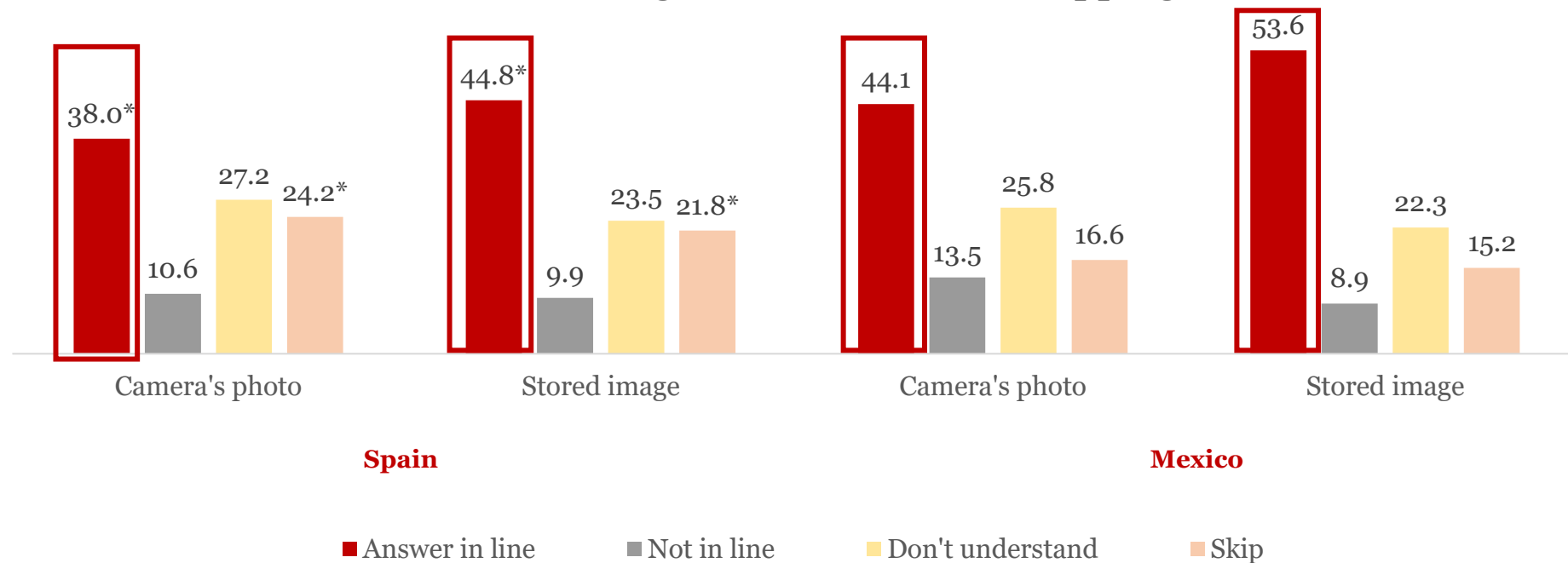

Spain

Camera's photo: 38.0* | 10.6 | 27.2 | 24.2*
Stored image: 44.8* | 9.9 | 23.5 | 21.8*

Mexico

Camera's photo: 44.1 | 13.5 | 25.8 | 16.6
Stored image: 53.6 | 8.9 | 22.3 | 15.2

■ Answer in line　■ Not in line　■ Don't understand　■ Skip

Note: * indicates significant differences between tasks at p<.05

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review*, *37*(5), 669-683.

What have we found in our research?

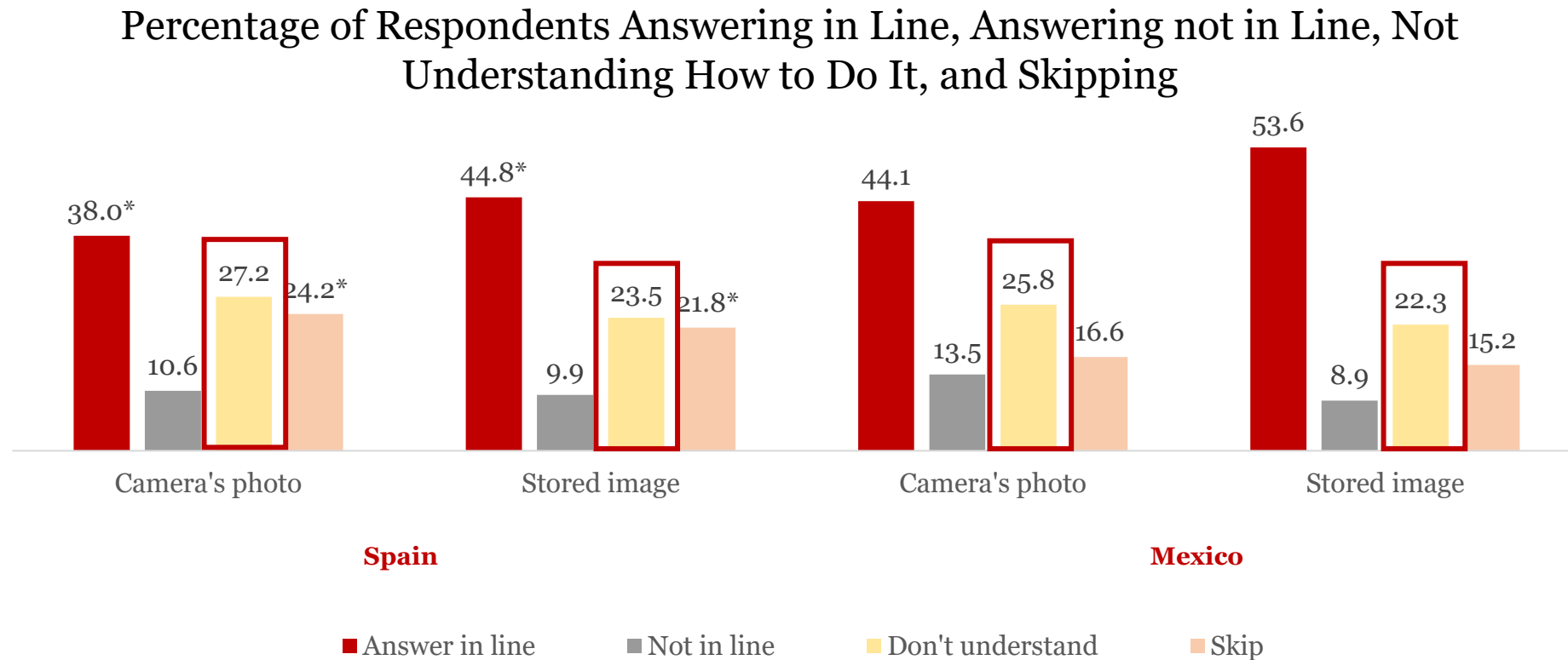## What is the real willingness of participants?

Percentage of Respondents Answering in Line, Answering not in Line, Not Understanding How to Do It, and Skipping

*Note: * indicates significant differences between tasks at p<.05*

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review, 37*(5), 669-683.

What have we found in our research?

What is the real willingness of participants?

Percentage of Respondents Answering in Line, Answering not in Line, Not Understanding How to Do It, and Skipping

Note: indicates significant differences between tasks at p<.05

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review, 37*(5), 669-683.

VISUAL DATA

What have we found in our research?

What is the impact on data quality compared with text?

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).
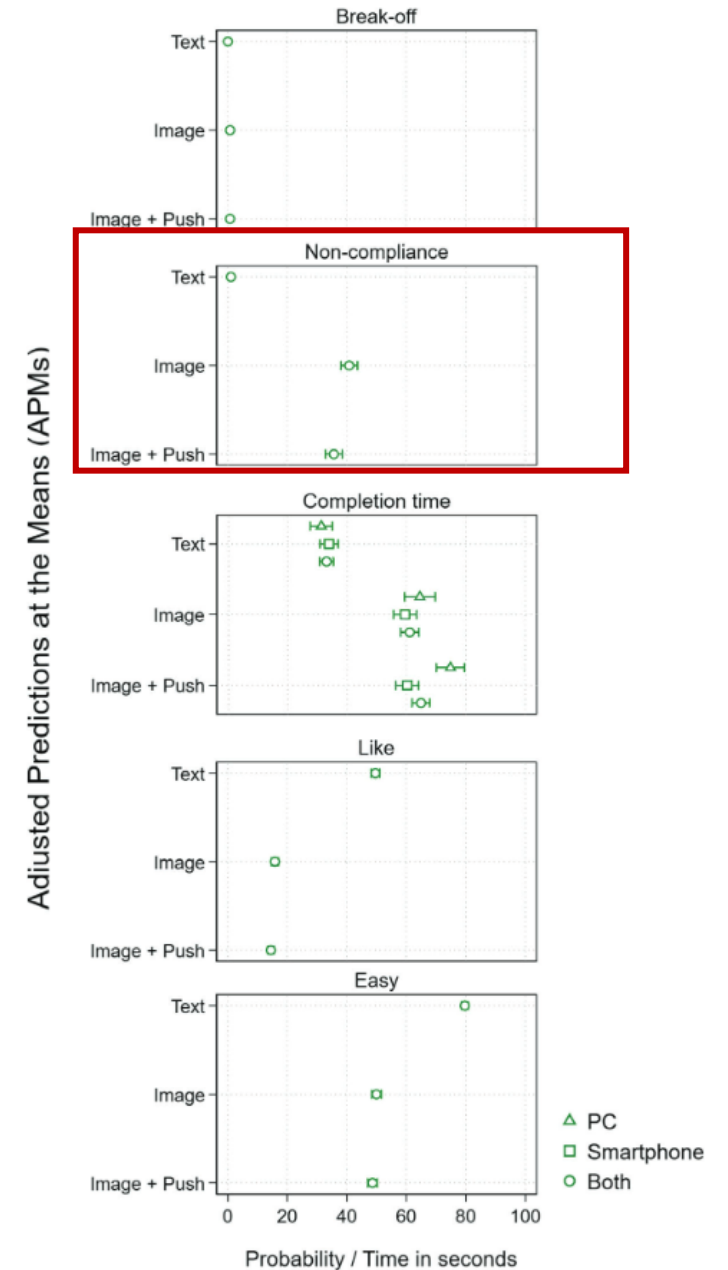
What have we found in our research?

What is the impact on data quality compared with text?

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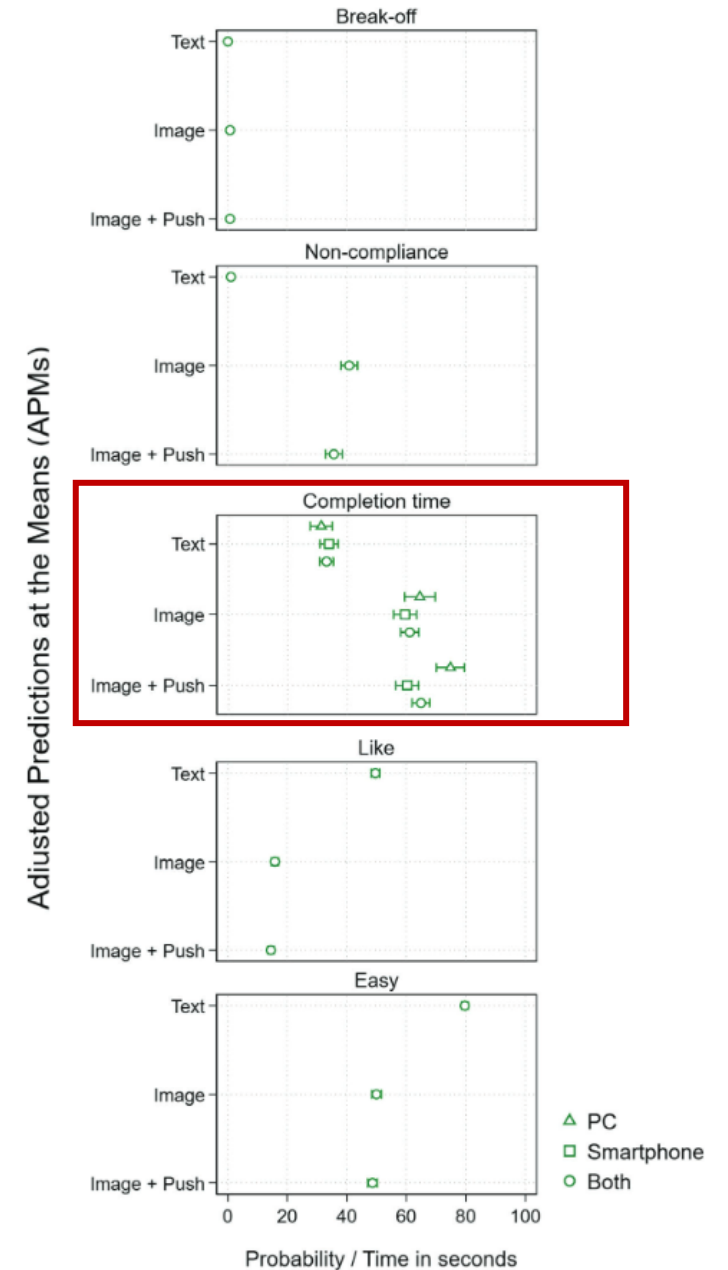- **Increases the probability of item nonresponse** (34-39 p.points higher probability)

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).

What have we found in our research?

What is the impact on data quality compared with text?

Asking for images:

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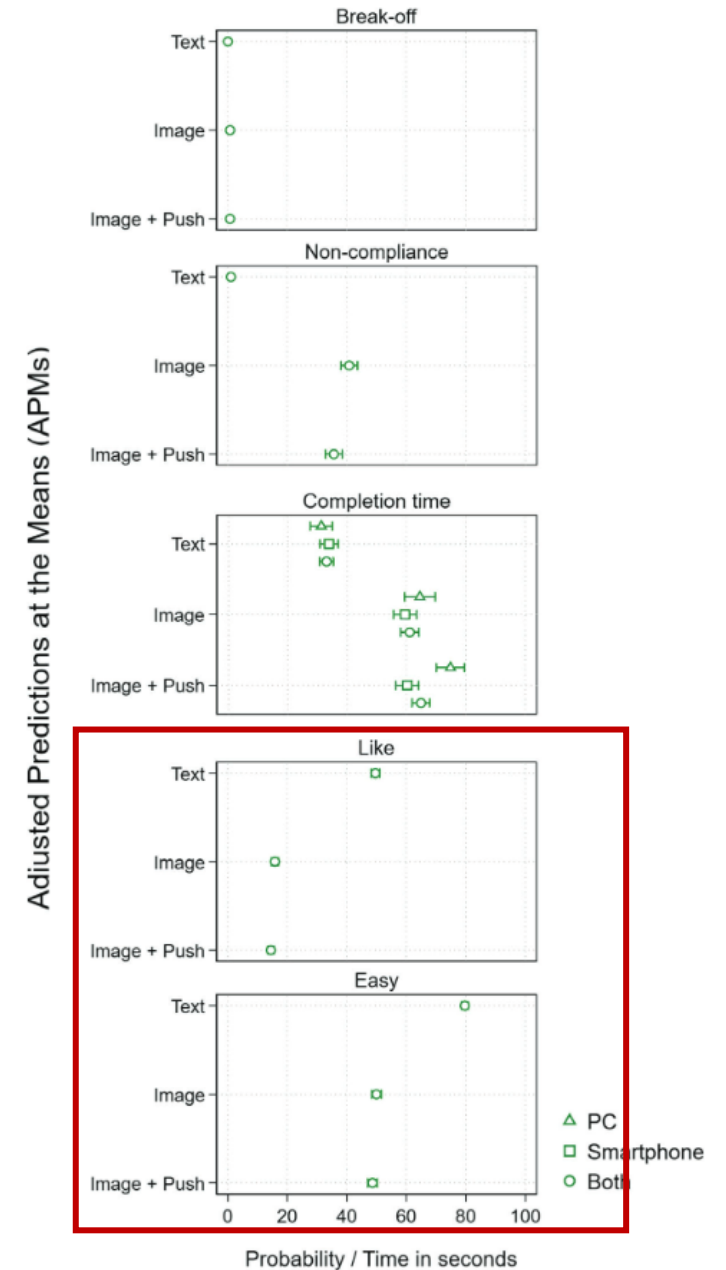- **Increases completion times** (25.6 to 43.52 seconds more)

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).

What have we found in our research?

What is the impact on data quality compared with text?

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).
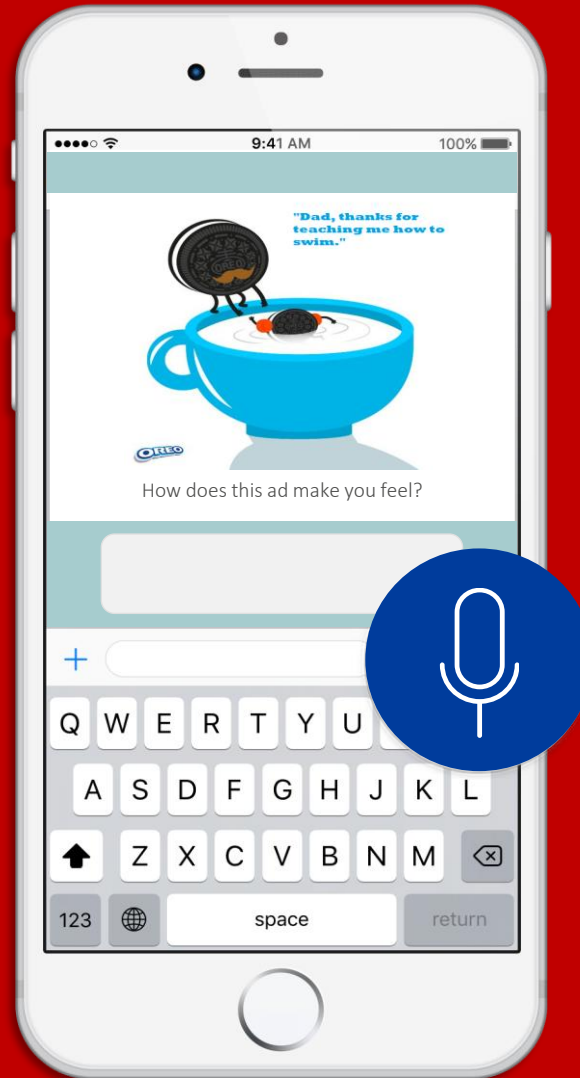
What have we found in our research?

How suitable is a Computer Vision API compared with human coders?

How to code all these images to extract workable information? ➡ Computer vision

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review, 37*(5), 669-683.

What have we found in our research?

How suitable is a Computer Vision API compared with human coders?

How to code all these images to extract workable information? ➡️ Computer vision

We assessed the adequacy of using a computer vision API (Google Vision)

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review, 37*(5), 669-683.

What have we found in our research?

How suitable is a Computer Vision API compared with human coders?

How to code all these images to extract workable information? ➡ Computer vision

We assessed the adequacy of using a computer vision API (Google Vision)

- Cost x 1,818 images = Human: **553€** / Google Vision: **2.50€**
- Time x 1,818 = Human: **35 hours** / Google Vision: **<5 minutes**
- Between **52%** and **65%** of the images similarly coded by the human coder and Google Vision

Bosch, O. J., Revilla, M., & Paura, E. (2019). **Answering mobile surveys with images: an exploration using a computer vision API.** *Social Science Computer Review, 37*(5), 669-683.

VOICE DATA

What kind of voice data can we ask for?

- **Voice memos**: capturing the actual audio
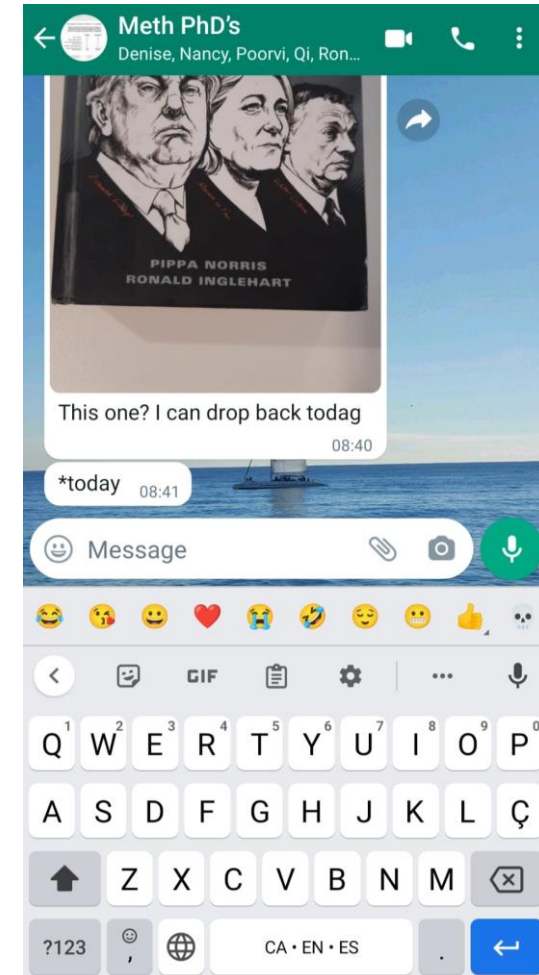
What kind of voice data can we ask for?

- **Voice memos**: capturing the actual audio

- **Voice dictation**: directly transforming the voice to data without having access to the audio
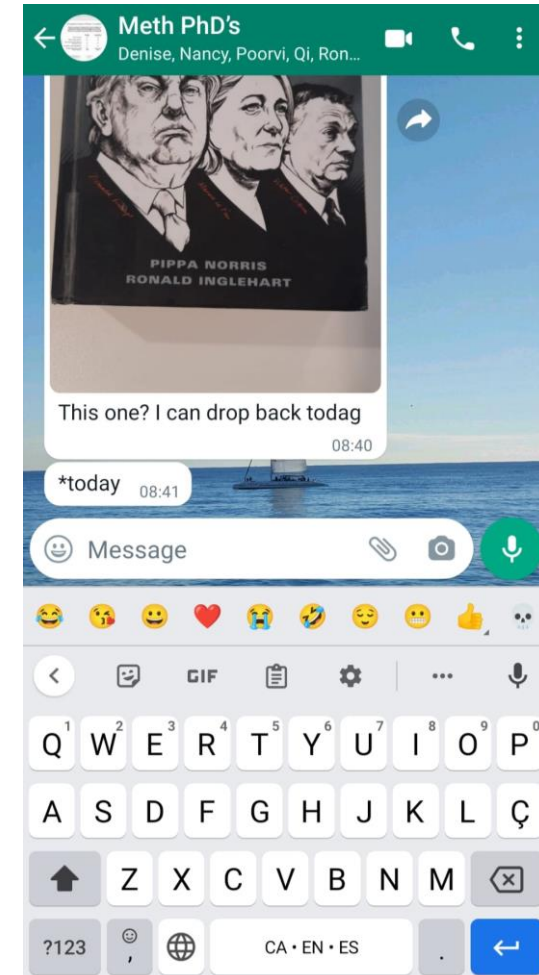
Why would we want to ask for voice data?

- To obtain **longer** and **richer** answers to **open-ended questions**

Why would we want to ask for voice data?

- To obtain **longer** and **richer** answers to **open-ended questions**

- To **extract information** from the **voice** (e.g., sentiment or intensity of opinion)

Why would we want to ask for voice data?

- To obtain **longer** and **richer** answers to **open-ended questions**

- To **extract information** from the **voice** (e.g., sentiment or intensity of opinion)

- To measure **contextual data** (e.g., noise pollution, social networks, TV/streaming consumption)

Why would we want to ask for voice data?

- To obtain **longer** and **richer** answers to **open-ended questions**

- To **extract information** from the **voice** (e.g., sentiment or intensity of opinion)

- To measure **contextual data** (e.g., noise pollution, social networks, TV/streaming consumption)

- To make the **survey experience** more **natural** and **enjoyable**

web
data
opp

What do you need to ask for voice data?

- Again, an **interface** & **infrastructure** to allow for the collection and storage of voice data: **WebdataVoice**

More information available at: https://www.upf.edu/web/webdataopp/tools
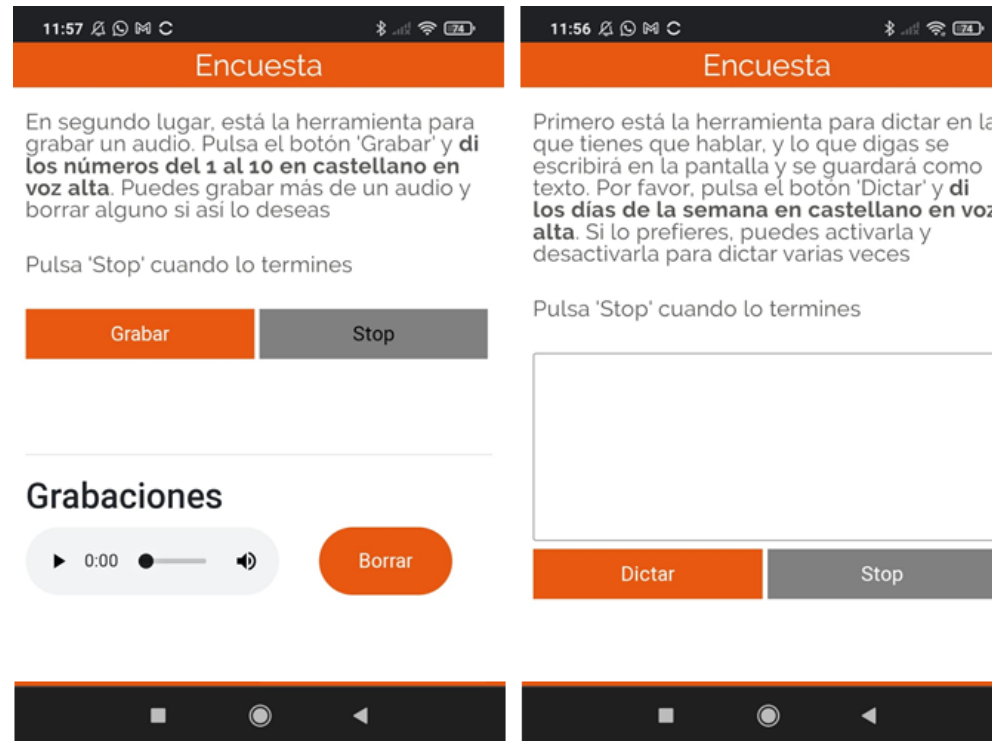
What do you need to ask for voice data?

- Again, an **interface** & **infrastructure** to allow for the collection and storage of voice data: **WebdataVoice**

- An approach to extract information from the voice
  - This can mean **simple transcription**
  - Or more **complex algorithms** to extract e.g., sentiment

Google Speech-to-text: https://cloud.google.com/vision?hl=en

What have we found in our research?

How well does it work to ask respondents to use voice input options **to answer open questions**?

Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). **Testing the use of voice input in a smartphone web survey**. *Social Science Computer Review*, *38*(2), 207-224.

What have we found in our research?

How well does it work to ask respondents to use voice input options **to answer open questions**?

- **Item nonresponse (% did not answer all 6 experimental questions)**
  - Android-Voice: **63.3%**
  - iOS-Dictation: 3.3%
  - Control groups: 1.5% in Android; 3.0% in iOS

Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). **Testing the use of voice input in a smartphone web survey**. *Social Science Computer Review, 38*(2), 207-224.

What have we found in our research?

How well does it work to ask respondents to use voice input options **to answer open questions**?

- **Item nonresponse (% did not answer all 6 experimental questions)**
  - Android-Voice: **63.3%**
  - iOS-Dictation: 3.3%
  - Control groups: 1.5% in Android; 3.0% in iOS

- **Problems reported**
  - Android-Voice: **26.3%**
  - iOS-Dictation: 8.7%
  - Control groups: 5.4% in Android and 6.0% in iOS

Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). **Testing the use of voice input in a smartphone web survey**. *Social Science Computer Review*, *38*(2), 207-224.

What have we found in our research?

For those answering, effect of using voice input on:

- **Data quality**
  - **Longer and more elaborated answers for Android-Voice**
  - Fewer valid answers and a smaller amount of information for iOS-Dictation

Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). **Testing the use of voice input in a smartphone web survey**. *Social Science Computer Review*, *38*(2), 207-224.

What have we found in our research?

For those answering, effect of using voice input on:

- **Data quality**
  - **Longer and more elaborated answers for Android-Voice**
  - Fewer valid answers and a smaller amount of information for iOS-Dictation

- **Survey evaluation**
  - **More negative ratings for Android-Voice**
  - No difference for iOS-Dictation

Revilla, M., Couper, M. P., Bosch, O. J., & Asensio, M. (2020). **Testing the use of voice input in a smartphone web survey**. *Social Science Computer Review*, *38*(2), 207-224.
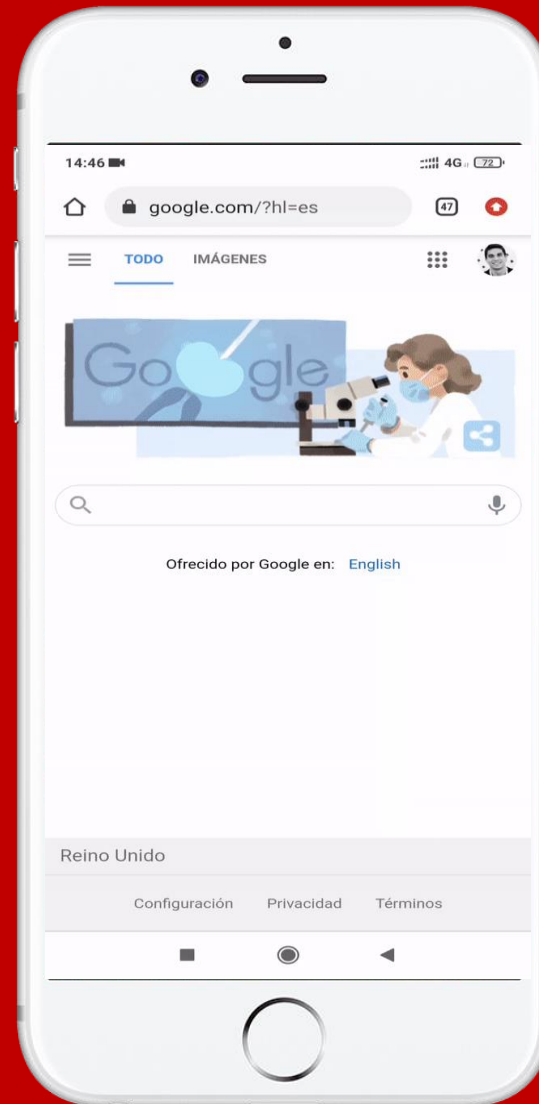
METERED DATA

What is metered data?

Approach to 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
ALLOWS MASSIVE DATA COLLECTION OF THE ONLINE ACTIVITY OF USERS.

| URLs of the web-pages visited | The use of apps | Time and length of the visit | Ad exposure |
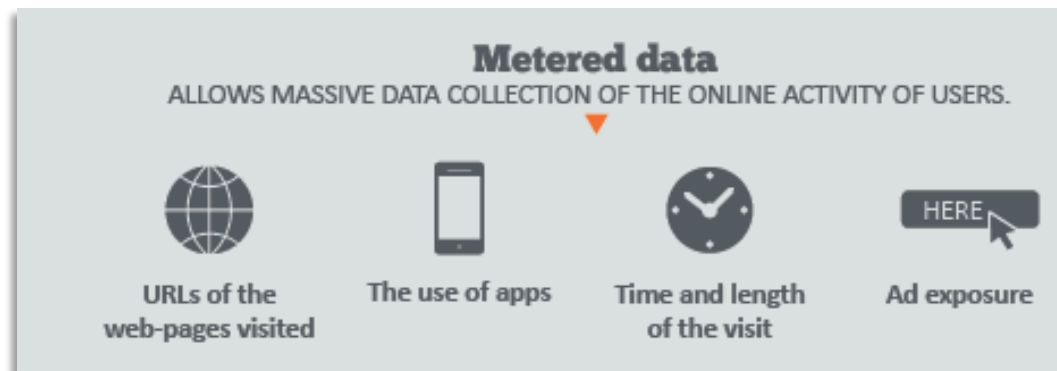
Why using metered data?

- **Unbiased observations** of their online behaviour

Why using metered data?

- **Unbiased observations** of their online behaviour

- More **granular data** (imagine asking if they have visited 700 different outlets)

Why using metered data?

- **Unbiased observations** of their online behaviour

- More **granular data** (imagine asking if they have visited 700 different outlets)

- To obtain **new types of data** (e.g., HTML information)

Why using metered data?

- **Unbiased observations** of their online behaviour

- More **granular data** (imagine asking if they have visited 700 different outlets)

- To obtain **new types of data** (e.g., HTML information)

- To **reduce** the number of **questions** and **burden** of participants

What do you need to collect metered data?

Tracking technologies which participants can **install or configure** in their devices:

What do you need to collect metered data?

Tracking technologies which participants can **install or configure** in their devices:

- Desktop apps (i.e., VPNs)
- Web browser plug-ins
- Smartphone apps
- Manually configured proxies
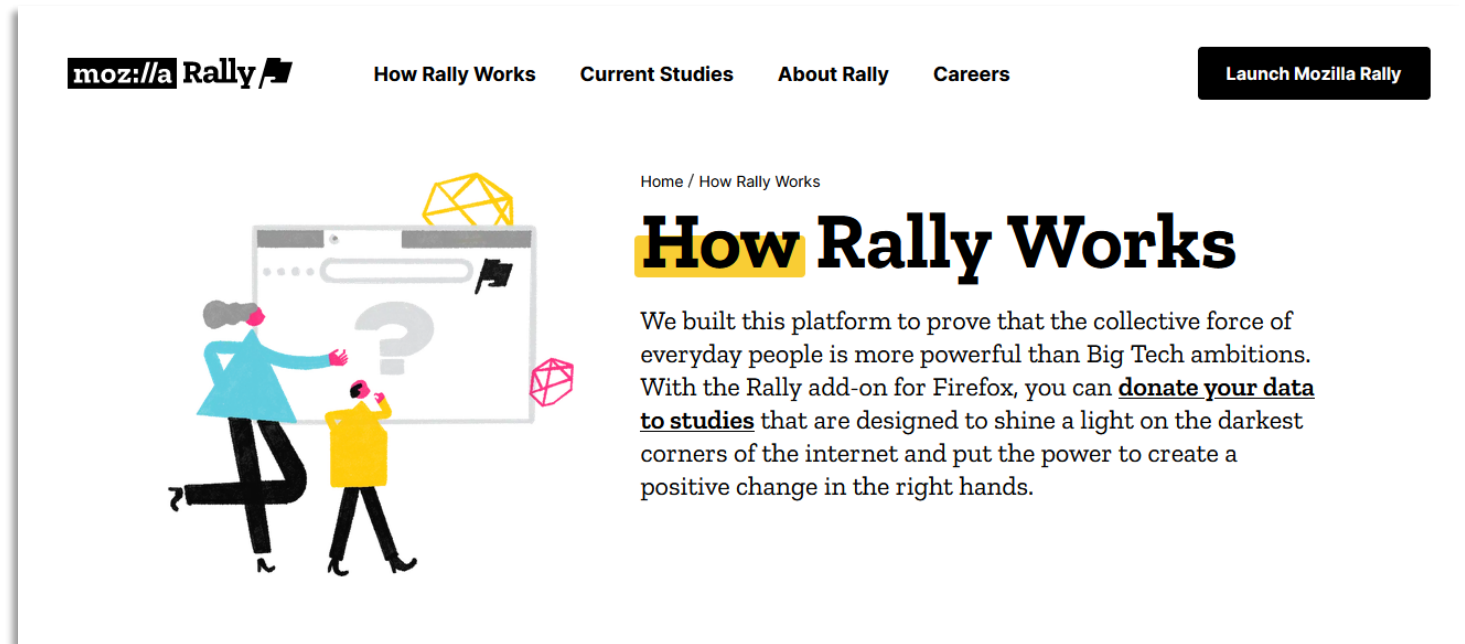
What have we found in our research?

What have we found in our research?

Our main interest has been to test whether metered data is actually **unbiased**

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
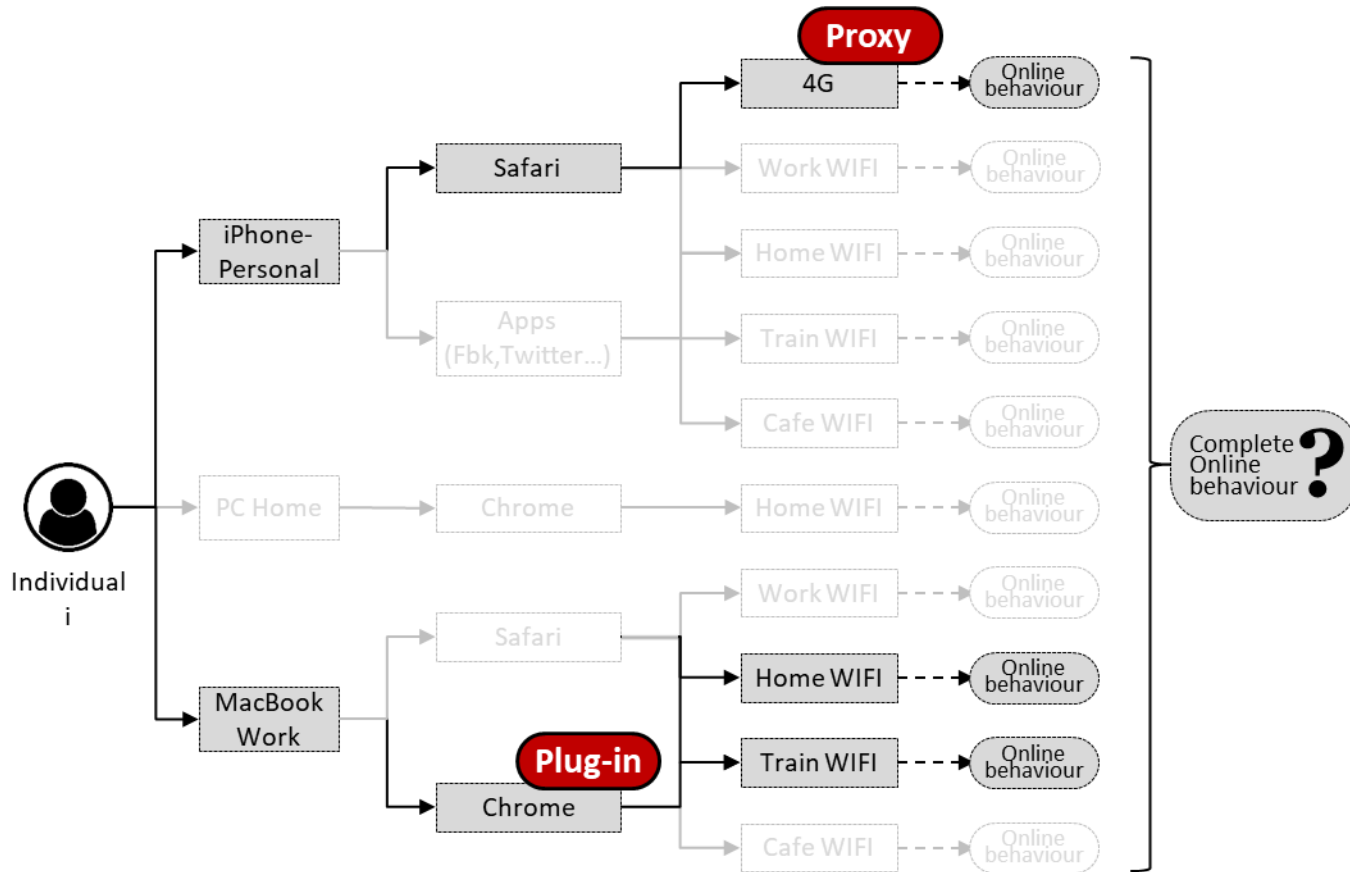
Biases of metered data: tracking undercoverage
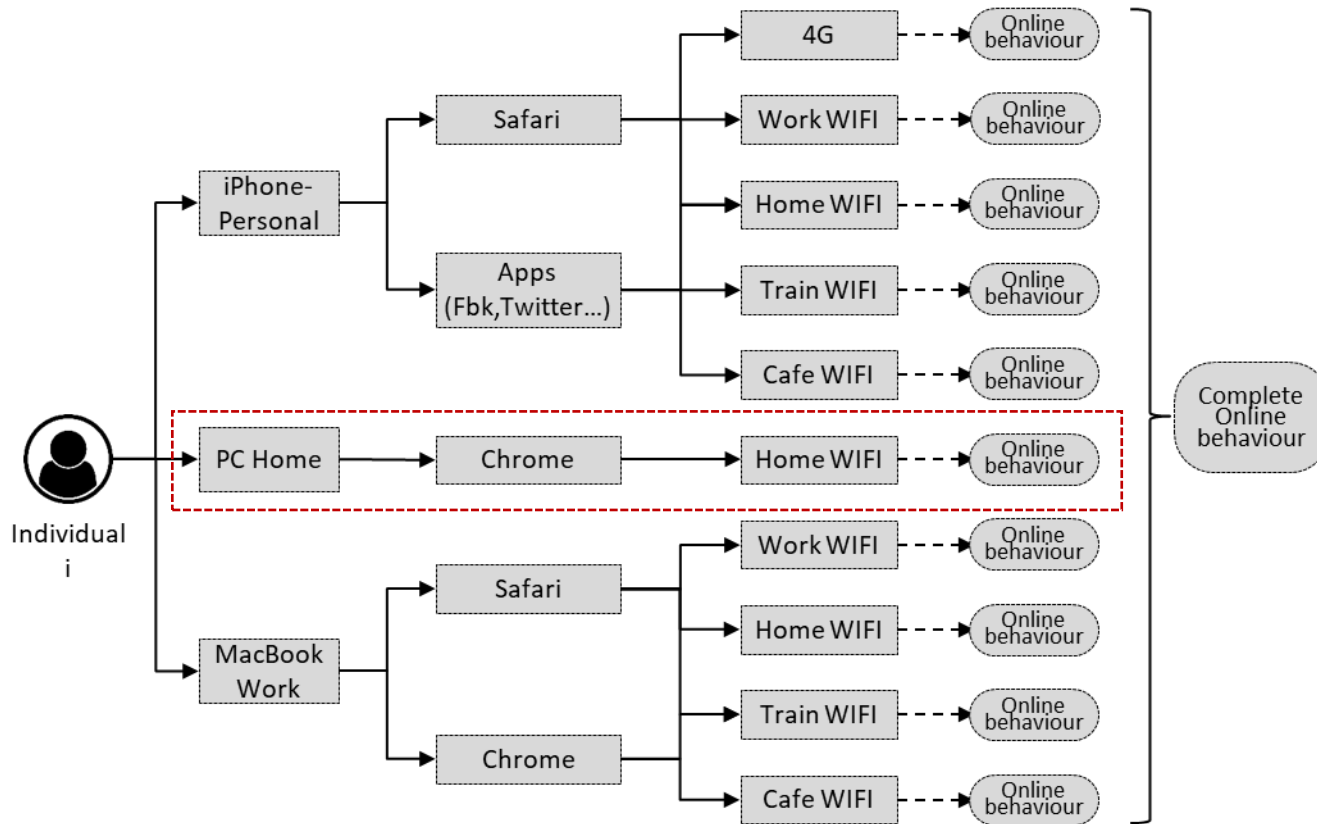
Objective: measuring individuals' behaviours

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

Biases of metered data: tracking undercoverage

Undercoverage can prevent tracking a participant's complete online behaviour.

Different reasons:

- **Non-trackable targets**
- **Meter not installed**
- **Meter uninstalled**
- **New non-tracked target**

Proportion undercovered in the Netquest opt-in panel

	Spain	Italy	Portugal
Overall	80.5	83.1	85.7
Device*	69.7	76.1	77.5
Browser	35.1	26.8	39.3

Very high prevalence, with differences between device and browser
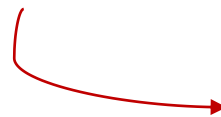
* 68% in the Pew Research Centre report, in the USA, using a probability-based panel and a different tracking provider

Is undercoverage evenly distributed across devices?

Proportion of users who use a specific type of device and not all of them are tracked

	Spain	Italy	Portugal
Windows PC	50.5	54.0	49.2
MAC	69.3	78.2	67.2
Android	44.7	47.8	53.1
iOS	93.4	80.9	95.4

Apple devices present a substantially higher prevalence

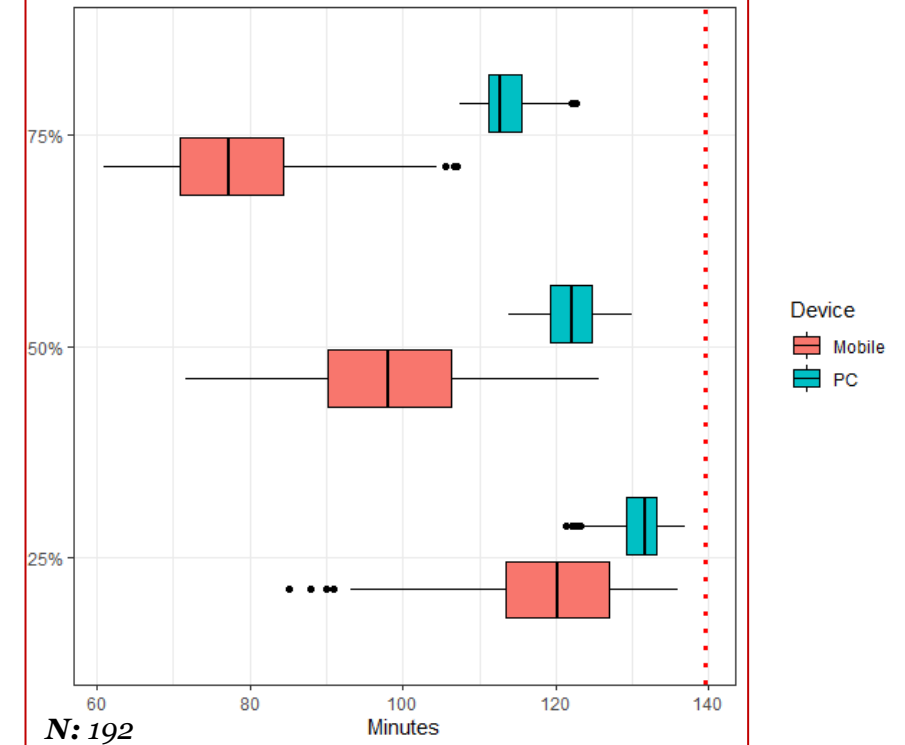# Undercoverage biases univariate estimates...

Average time spent on the internet
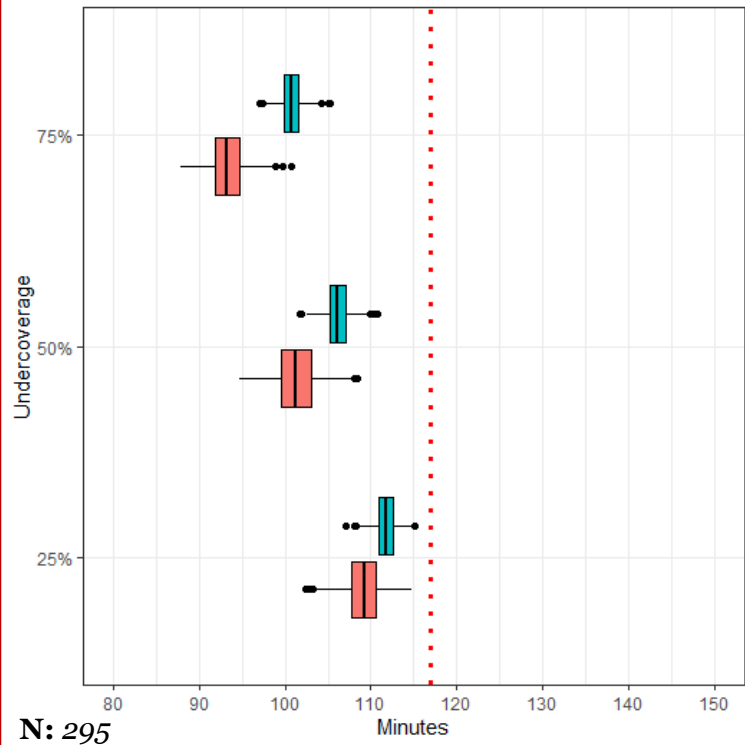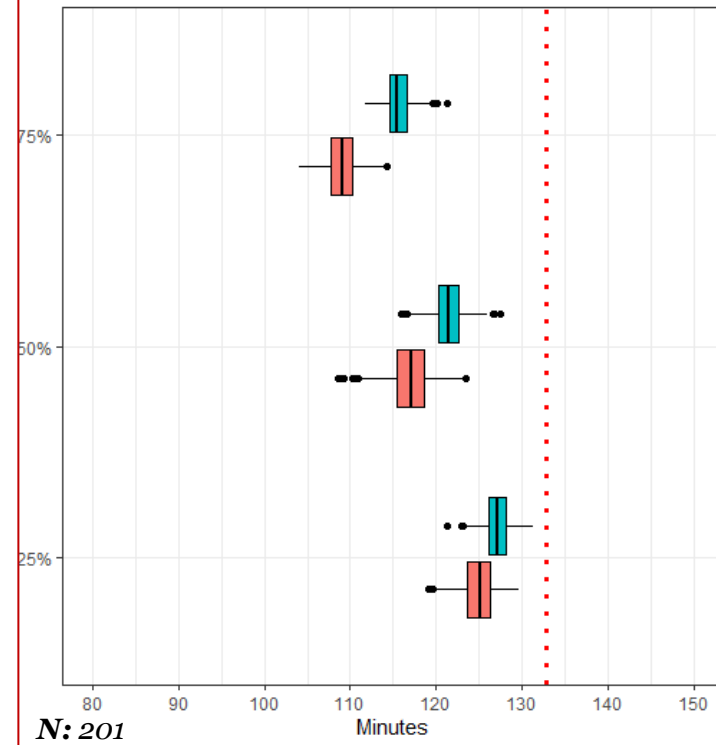
Avg. bias: *5 – 38 minutes*   *5 – 23 minutes*   *5 – 24 minutes*

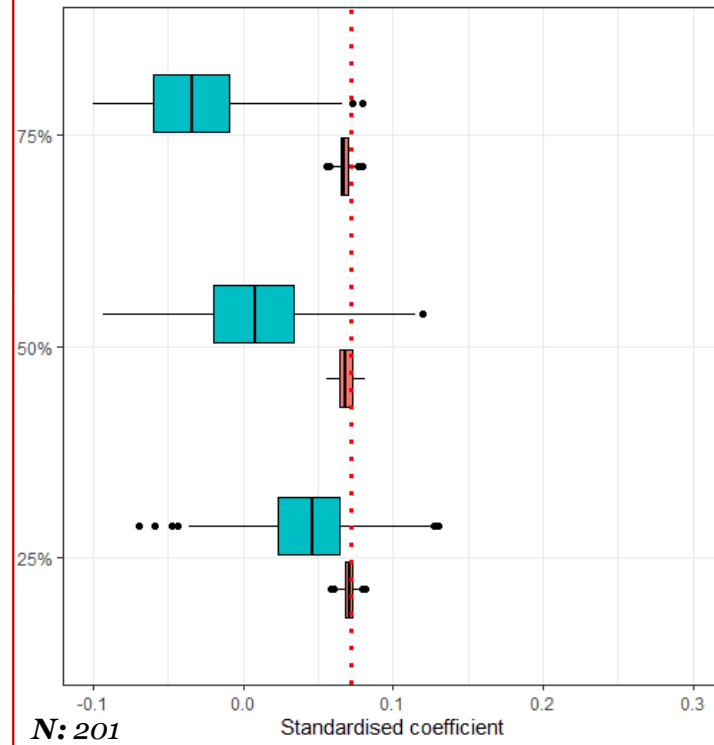# ...as well as multivariate estimates

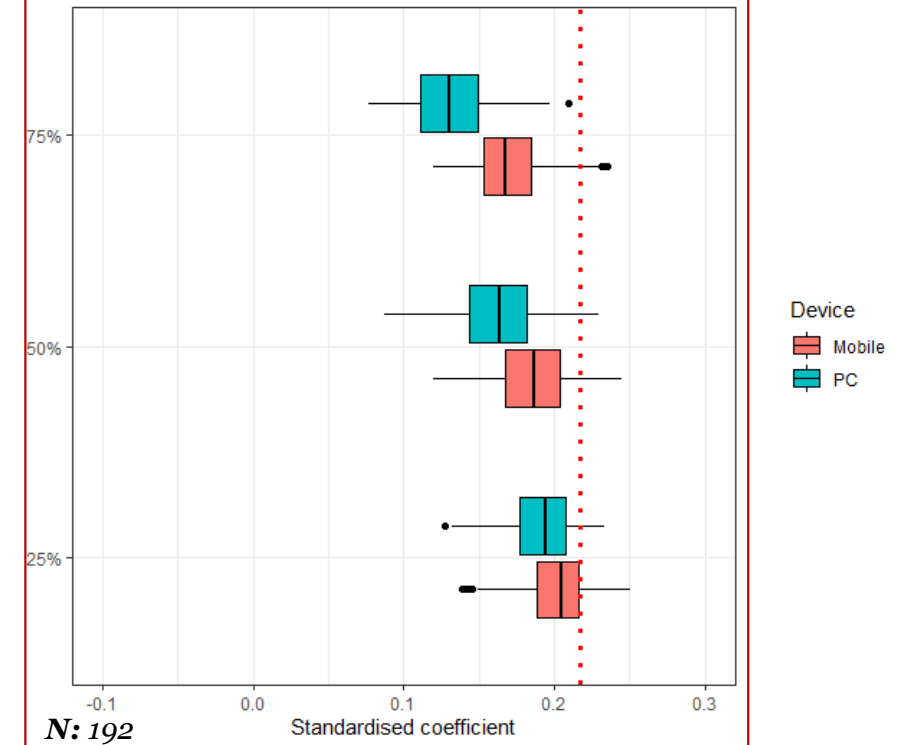**OLS coefficient: Political Knowledge ~ Nº visits to online news**
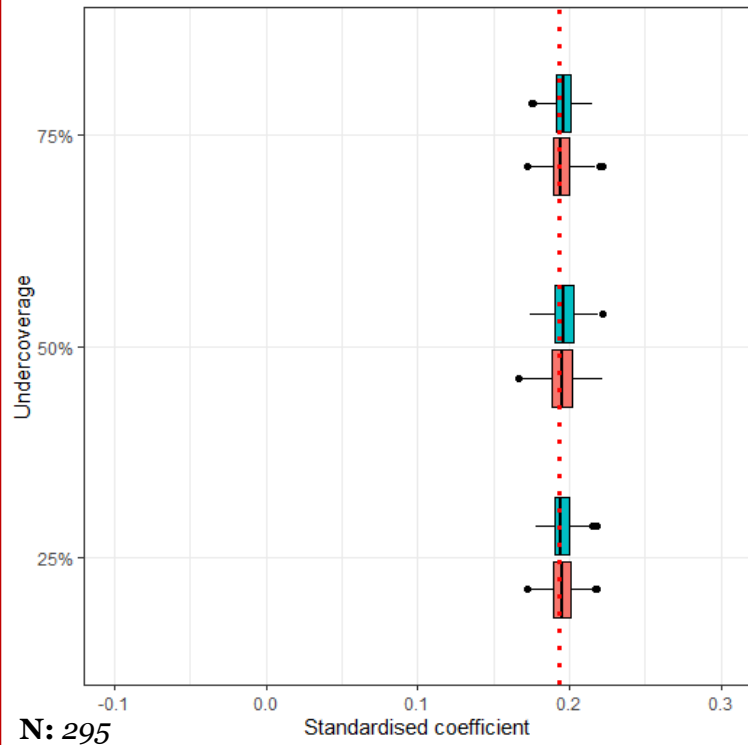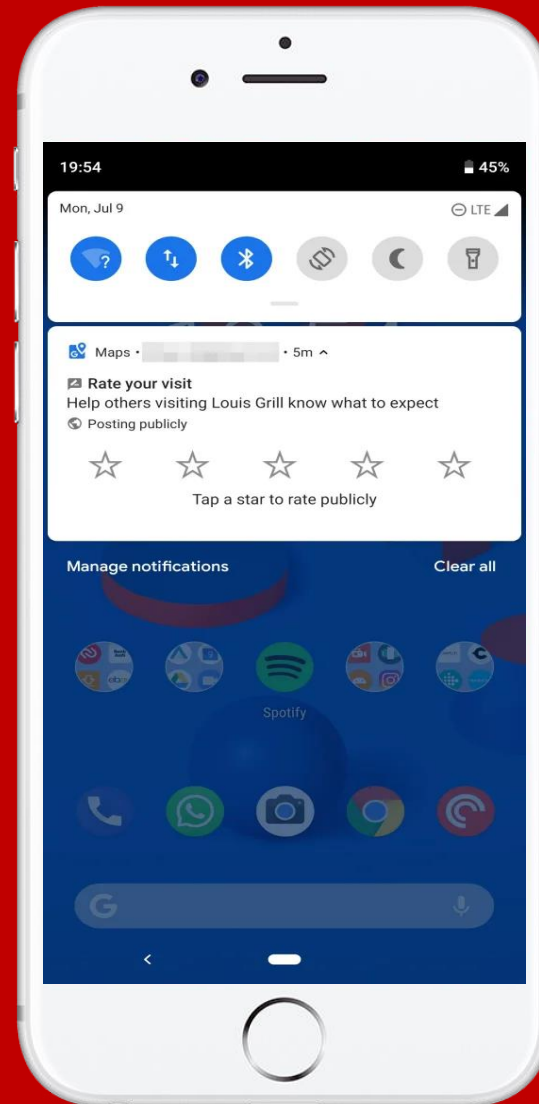

Avg. bias: *0.002 -0.003*   *0.00 – 0.11*   *0.01 – 0.09*
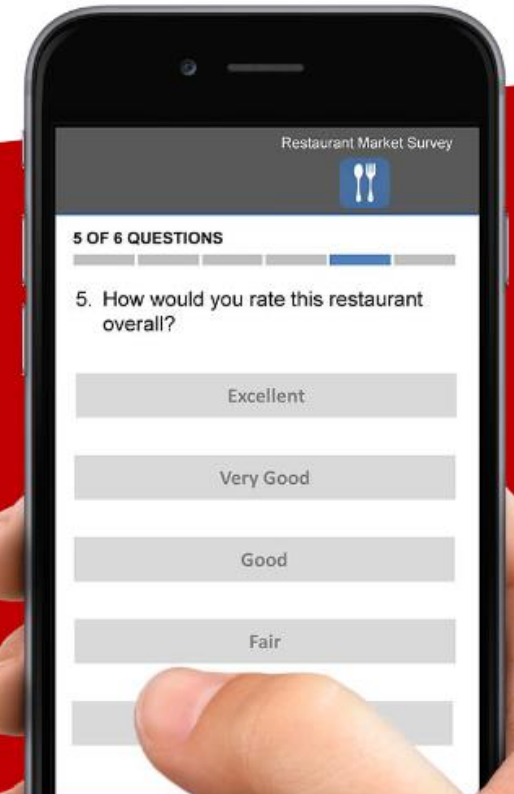
IN-THE-MOMENT SURVEYS

What are in-the-moment surveys

- Surveys which are **triggered by specific behaviours** of participants

What are in-the-moment surveys

- Surveys which are **triggered by specific behaviours** of participants

- These can be triggered by different kinds of data:
  - **GPS data**
  - **Metered data**
  - **Accelerometer data**
  - **Etc.**

web
data
opp

Why would we want to ask for in-the-moment surveys?

- Surveying a sample of individuals right in the moment – or short time after – an event of interest happens **may reduce memory errors.**
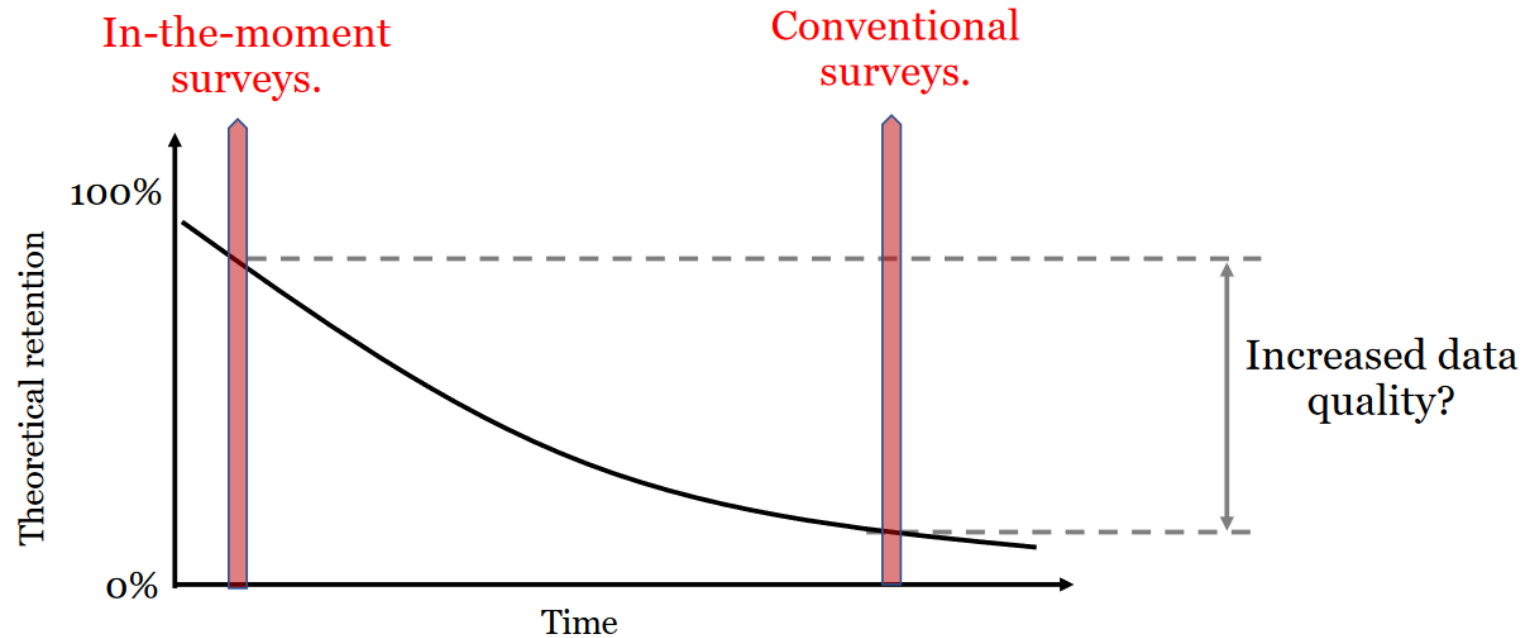
Why would we want to ask for in-the-moment surveys?

- Surveying a sample of individuals right in the moment – or short time after – an event of interest happens **may reduce memory errors.**

- Survey data can be used to **validate passive data**, and vice versa!

What do you need to ask for in-the-moment surveys?

web
data
opp

- **Tracking technologies** installed on participants devices that allow passively collecting GPS, accelerometer or metered data
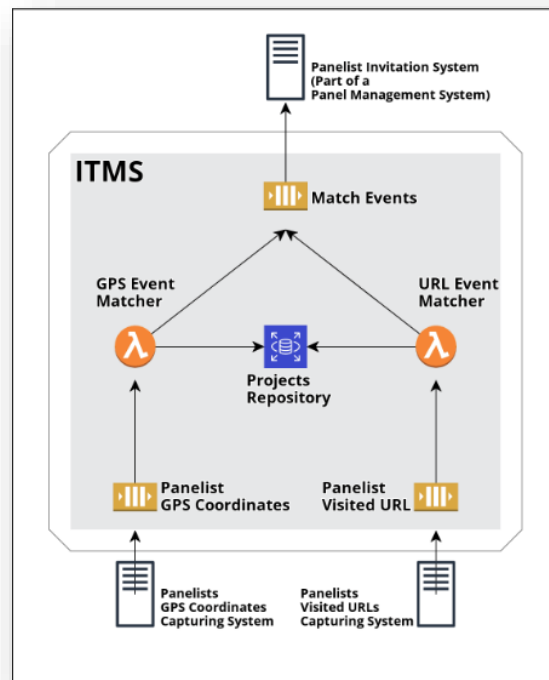
What do you need to ask for in-the-moment surveys?

- **Tracking technologies** installed on participants devices that allow passively collecting GPS, accelerometer or metered data

- A system that can **collect the passive data**, **detect events** of interest, and **send the surveys** through the desired channels: **WebdataNow**

More information available at: https://www.upf.edu/web/webdataopp/tools
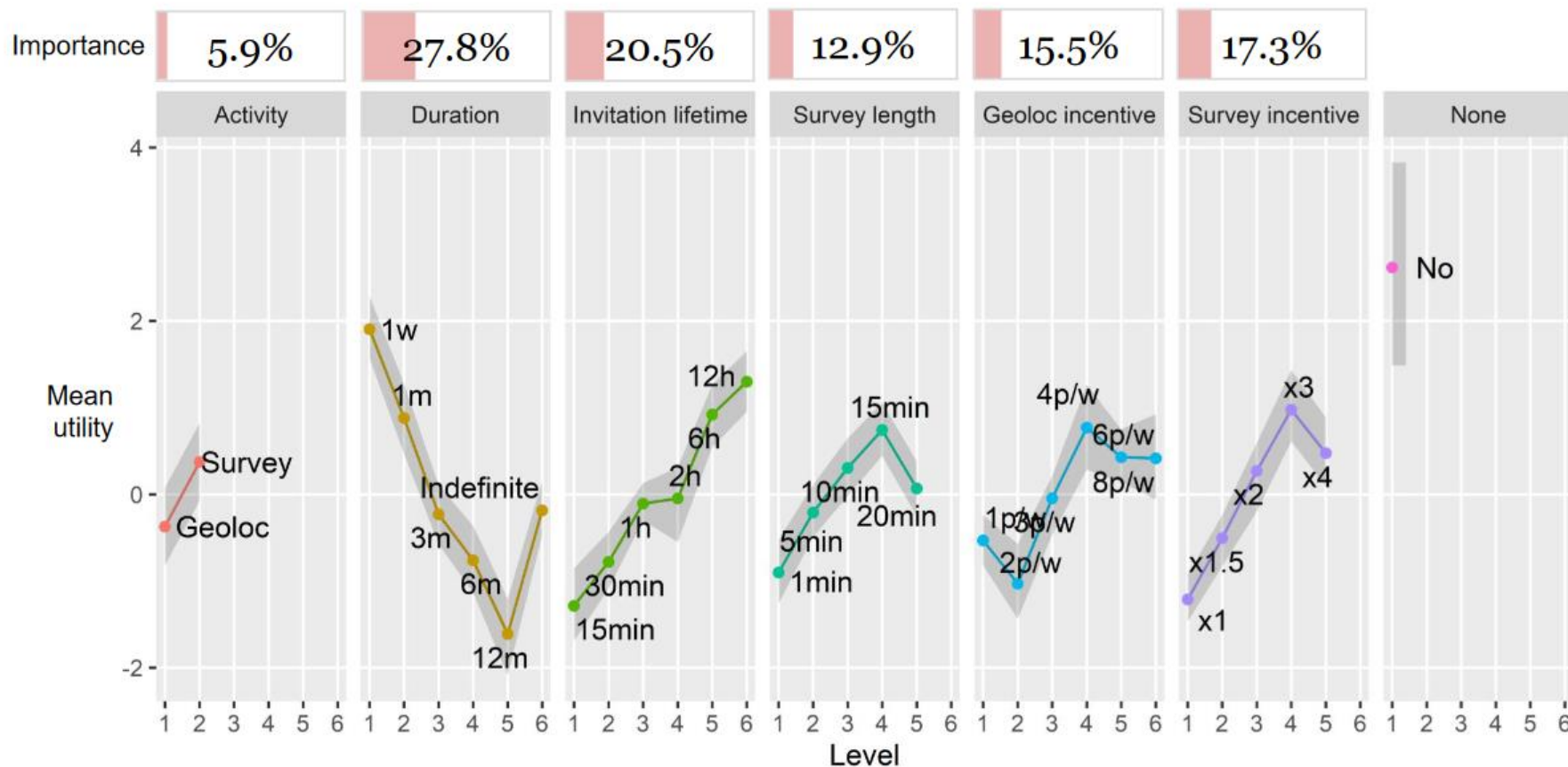
What have we found in our research?

Ochoa, C. and Revilla, M. (2022). **Willingness to participate in in-the-moment surveys triggered by online behaviors.** Behavior Research Methods (2022). https://doi.org/10.3758/s13428-022-01872-x.

What have we found in our research?

- The average willingness of panellists to participate in in-the-moment surveys triggered by:

    - **GPS location data**: **45.4%** (for people not being tracked by GPS already)
    - **Online behaviours**: **93.2%** (only for people already tracked with meters)

Ochoa, C. and Revilla, M. (2022). **Willingness to participate in in-the-moment surveys triggered by online behaviors.** Behavior Research Methods (2022). https://doi.org/10.3758/s13428-022-01872-x.

What have we found in our research?

The **factors influencing** panellists: triggered by **GPS**

Ochoa, C. and Revilla, M. (2022). **Willingness to participate in in-the-moment surveys triggered by online behaviors.** Behavior Research Methods (2022). https://doi.org/10.3758/s13428-022-01872-x.

What have we found in our research?

The **factors influencing** panellists: triggered by **online behaviours**
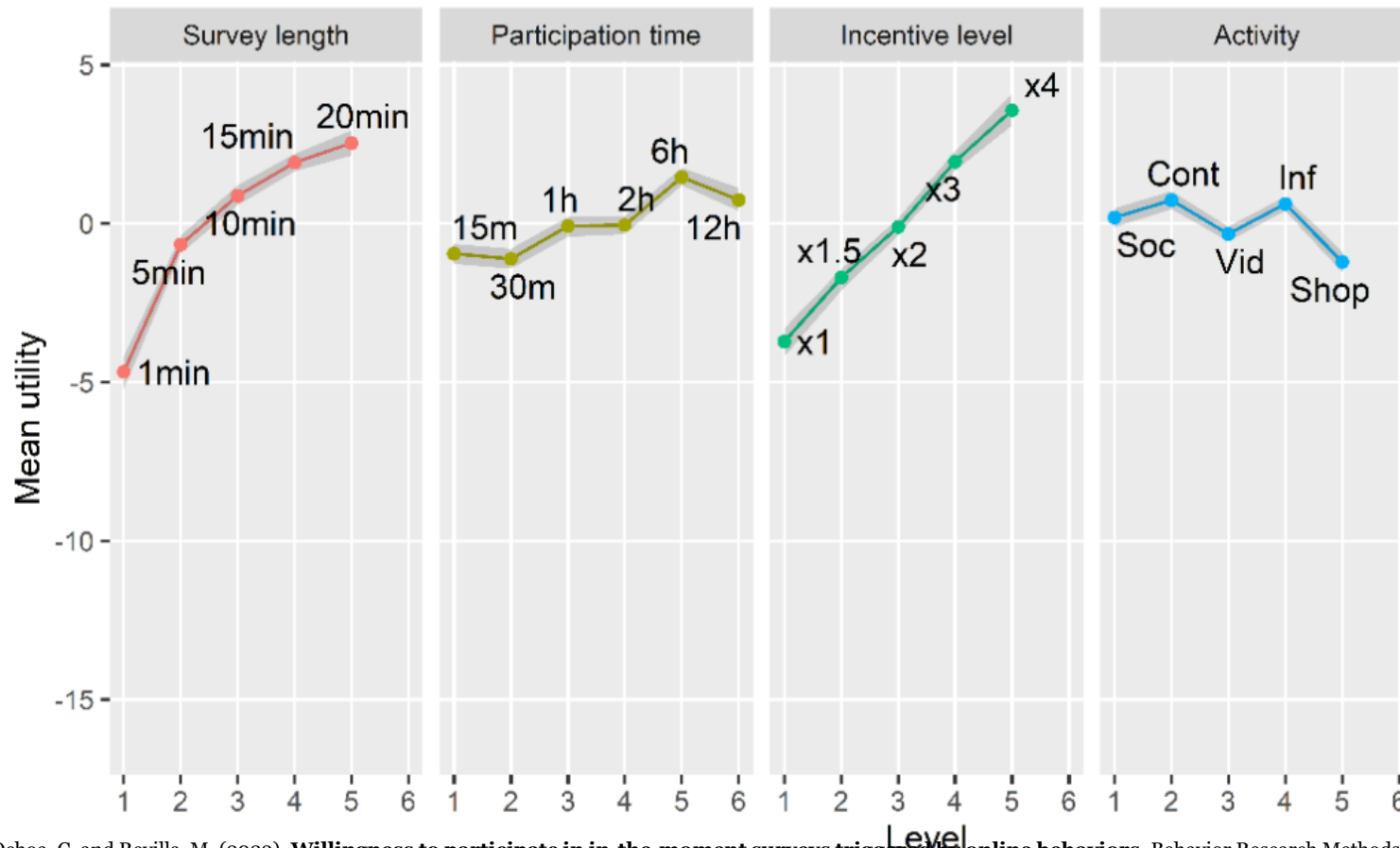
- Incentive level is the most important attribute.
- Survey length is also highly relevant, maybe because it affects the total incentive.
- Participation time and activity, much less relevant.

Ochoa, C. and Revilla, M. (2022). **Willingness to participate in in-the-moment surveys triggered by online behaviors.** Behavior Research Methods (2022). https://doi.org/10.3758/s13428-022-01872-x.

Closing remarks

Take-home messages

- Web surveys open the door to collecting many **new kinds of data**

Take-home messages

web
data
opp

- Web surveys open the door to collecting many **new kinds of data**

- However, they all bring challenges

  - We should expect **high nonresponse rates**
  - People do **not necessarily enjoy performing** these new tasks
  - And **data cannot be assumed to unbiased**

Take-home messages

- Web surveys open the door to collecting many **new kinds of data**

- However, they all bring challenges

  - We should expect **high nonresponse rates**
  - People do **not necessarily enjoy performing** these new tasks
  - And **data cannot be assumed to unbiased…but what can be assumed to be?**

Three ways to measure UK coronavirus deaths

Deaths with positive test result*

40,597

Death certificate mentions Covid-19**

50,107

Deaths over and above the usual number at this time of year**

63,708

*Figure to 7 Jun. Source: DHSC
**Figures to 29 May (31 May, in Scotland). Source: ONS, NRS, NISRA
Source: DHSC, ONS, NRS, NISRA BBC

Take-home messages

- Web surveys open the door to collecting many **new kinds of data**

- However, they all bring challenges

  - We should expect **high nonresponse rates**
  - People do **not necessarily enjoy performing** these new tasks
  - And **data cannot be assumed to unbiased…but what can be assumed to be?**

- Not realistic to aim for perfect data
  - What we need is to **be aware of the errors and their consequences**
  - Try to minimize them / correct for them / look from different perspectives

Thanks!

Questions?

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

✉ o.bosch-jover@lse.ac.uk

🐦 orioljbosch

💻 https://orioljbosch.com/