



Asking people to give us more data: Data linkage and extra-survey tasks

Jonathan Burton, ISER, University of Essex
and many many others



Background: Part of the NIMLAS programme



- Network for Innovative Methods in Longitudinal Aging Studies
 - Sub-group on “Maximizing Rates of Consent to Record Linkage and Additional Data Collection”
 - <https://nimlas.isr.umich.edu/research-areas/consent-to-additional-data-collection/>
 - Aim of *this* presentation is to complement the forthcoming white paper on best practice in this area
-

Why are we asking people for more information?

- Powerful computers make it easier to collect, store, manage, and analyse large datasets
 - Widespread use of internet and smartphones has opened up opportunities to collect new types of data
 - We just have to persuade people to let us have these data!
-

Data collection technologies have changed...

- In the beginning...

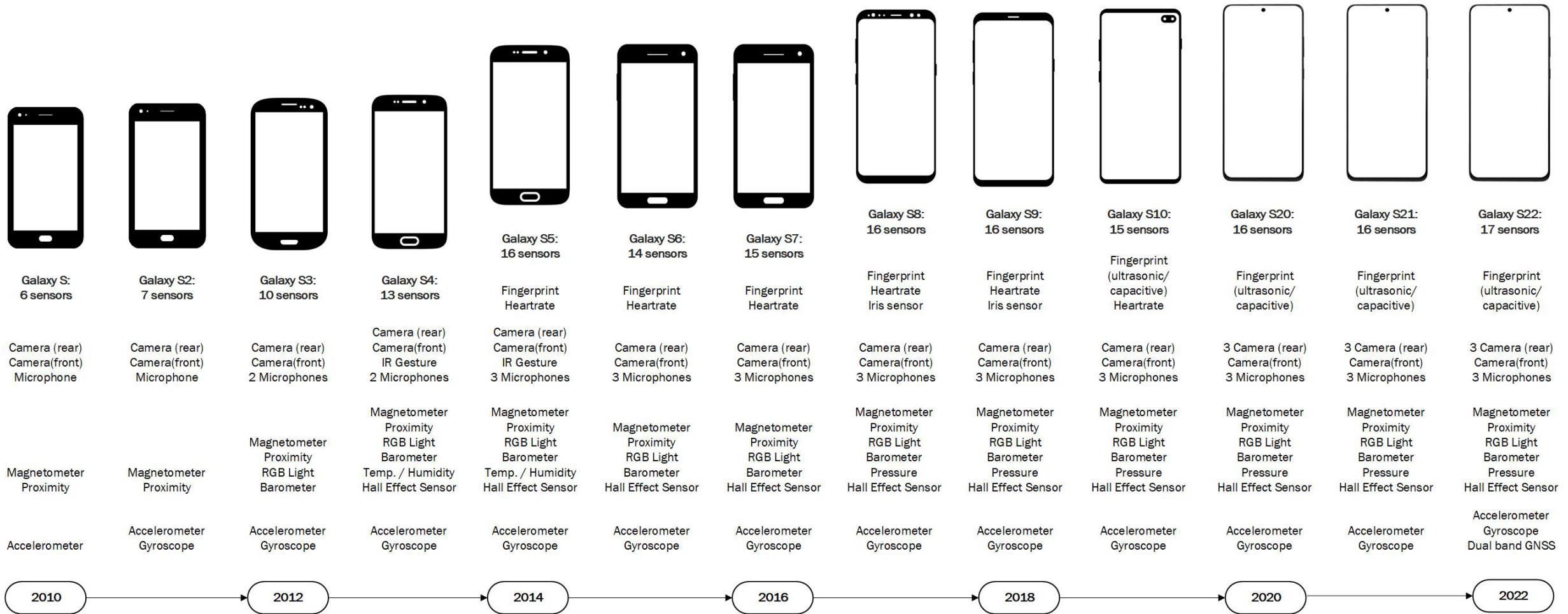
< mandatory picture of Literary Digest here >

Developments in technology lead to new opportunities for data collection

- Smartphones
 - Sensors, cameras, apps, trackers, microphones...



Smartphones are getting smarter



Developments in technology lead to new opportunities for data collection - 1

- Smartphones
 - Sensors, cameras, apps, trackers, microphones...
 - Wearable sensors
 - Accelerometry, air quality, GPS,
 - Digital process-generated data
 - Financial transactions, administrative records, social media, ...
 - Self-collection of biological samples
 - Blood, hair, saliva, stool, ...
 - AI...?
-

Researchers want more types of data



Admin data

Photos

Biological measures

Sensors

Event-triggered

Daily questions

Diaries

Gamified apps

Trackers



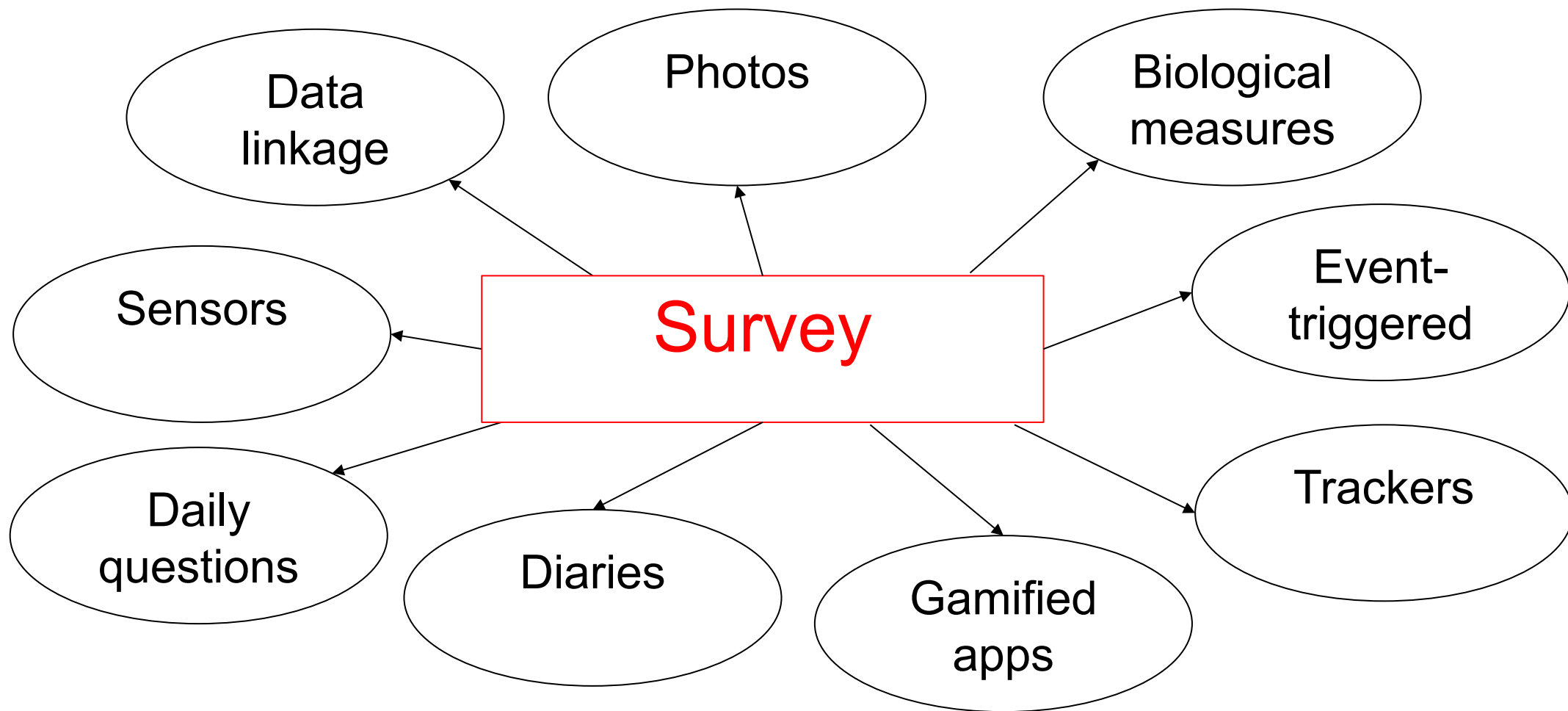
Challenge is getting a representative sample of your population

- Stand-alone studies using apps, sensors, trackers...
 - Often small-scale
 - Using volunteers to test
 - Self-selection (e.g., those who already own a specific type of device)
 - Lack of context information
 - Take place in ideal environment
 - Embed the request within a probability sample survey
-

Piggy-back on the strengths of surveys



- Probability samples improve the representativeness of the study
 - Relatively low barriers to participation
 - Can collect a wide range of information
 - ...On a large number of people
 - Able to incorporate requests for additional data collection
 - Allows researchers to look at selection and measurement effects of participating/consenting
-



Data collection systems



- Combining different methods of data collection
 - New concepts: information respondents do not know
 - Better quality: more accurate (?) information
 - More detail than can be asked in questionnaire
 - Continuous information over period of time
 - BUT...Are respondents willing to give consent or to take part in additional data collection tasks?
-

Before I go any further...



Understanding Society

THE UK HOUSEHOLD LONGITUDINAL STUDY

www.understandingsociety.ac.uk

Last century...

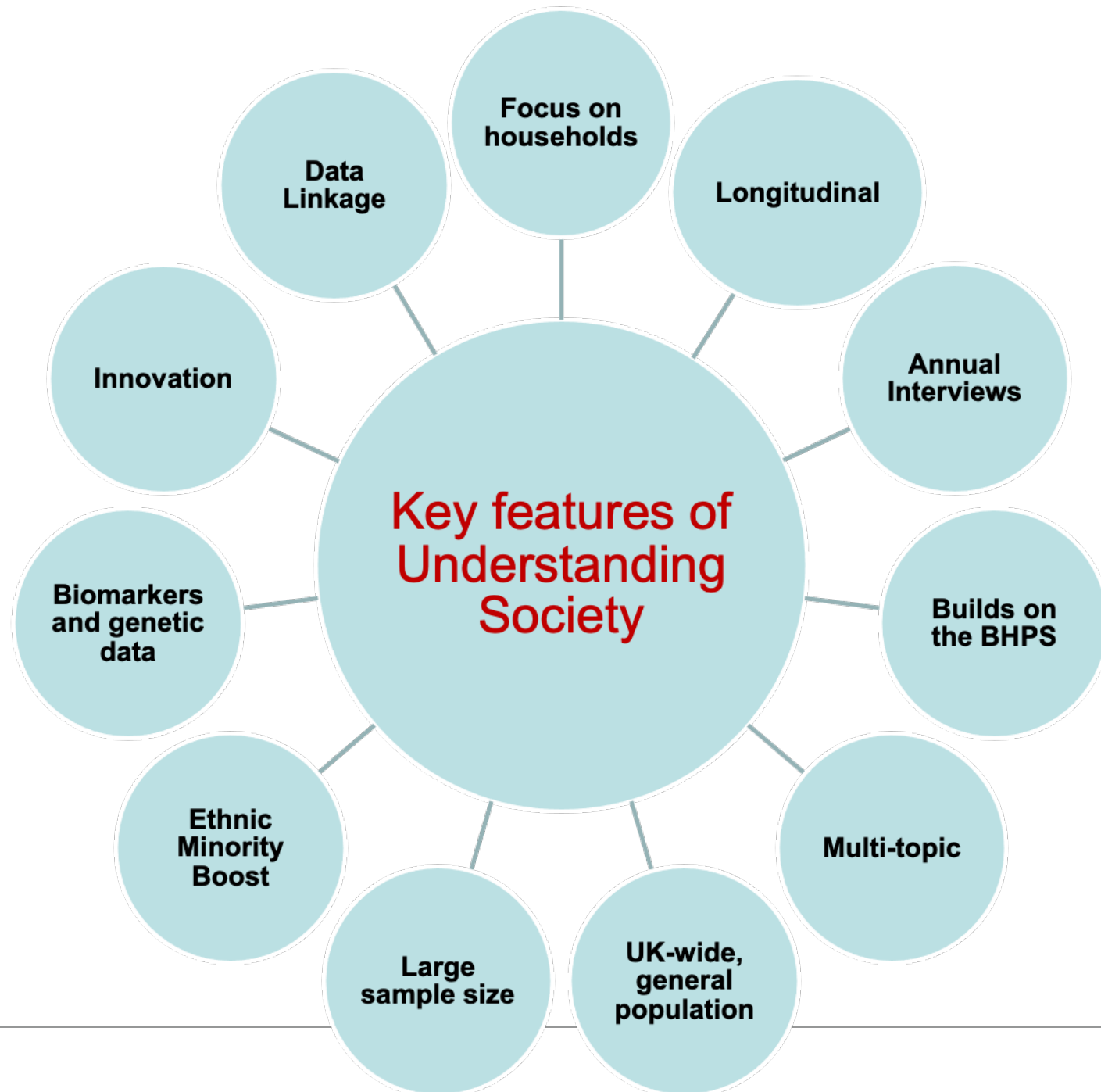
The British Household Panel Survey (BHPS)

- Started 1991, funded by the Economic and Social Research Council (ESRC)
 - Sample of residential households – 5,500 at Wave 1
 - All residents are ‘original sample members’
 - Aim: To follow and interview OSMs annually (and anyone they live with)
 - Design similar to other household panel surveys
-



Understanding Society: The UK Household Longitudinal Study (UKHLS)

- Increased interest in key groups → demand for larger sample sizes
 - 2009 – Understanding Society was launched
 - Funded by the ESRC
 - Same design as the BHPS...
 - ... but bigger and better!
-

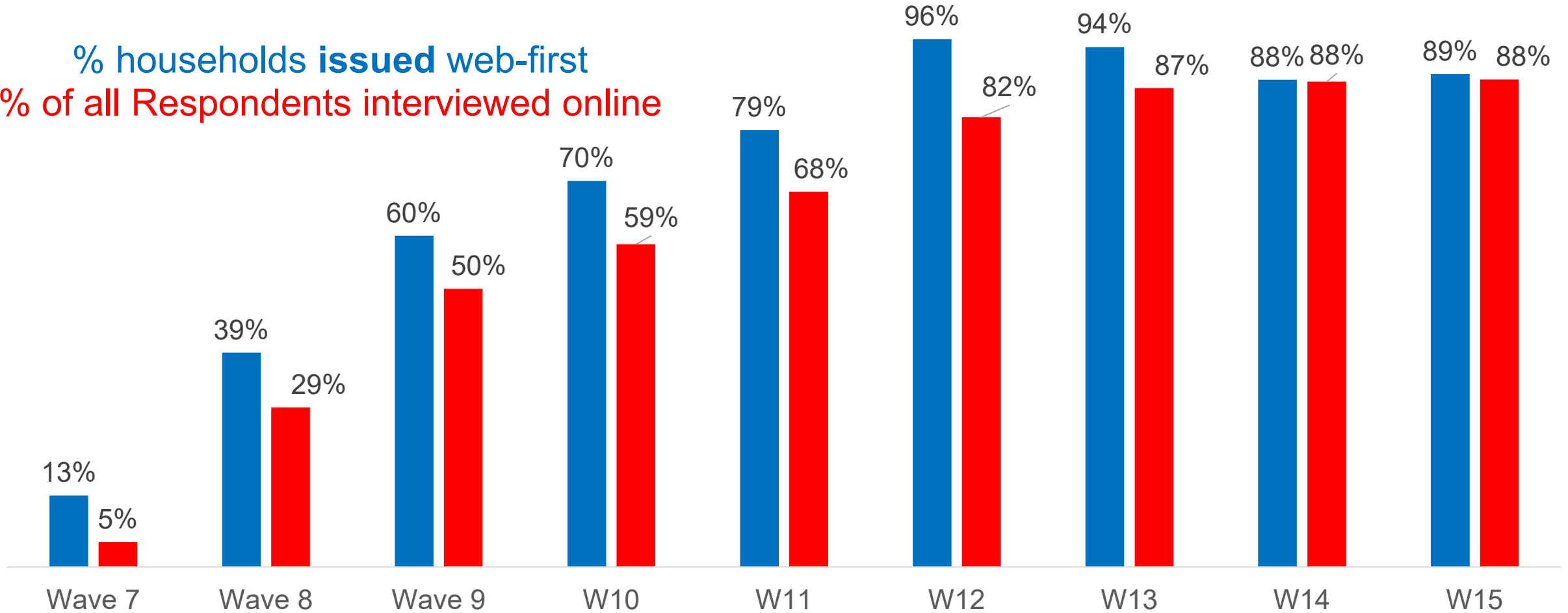


Innovation has always been a central part of Understanding Society

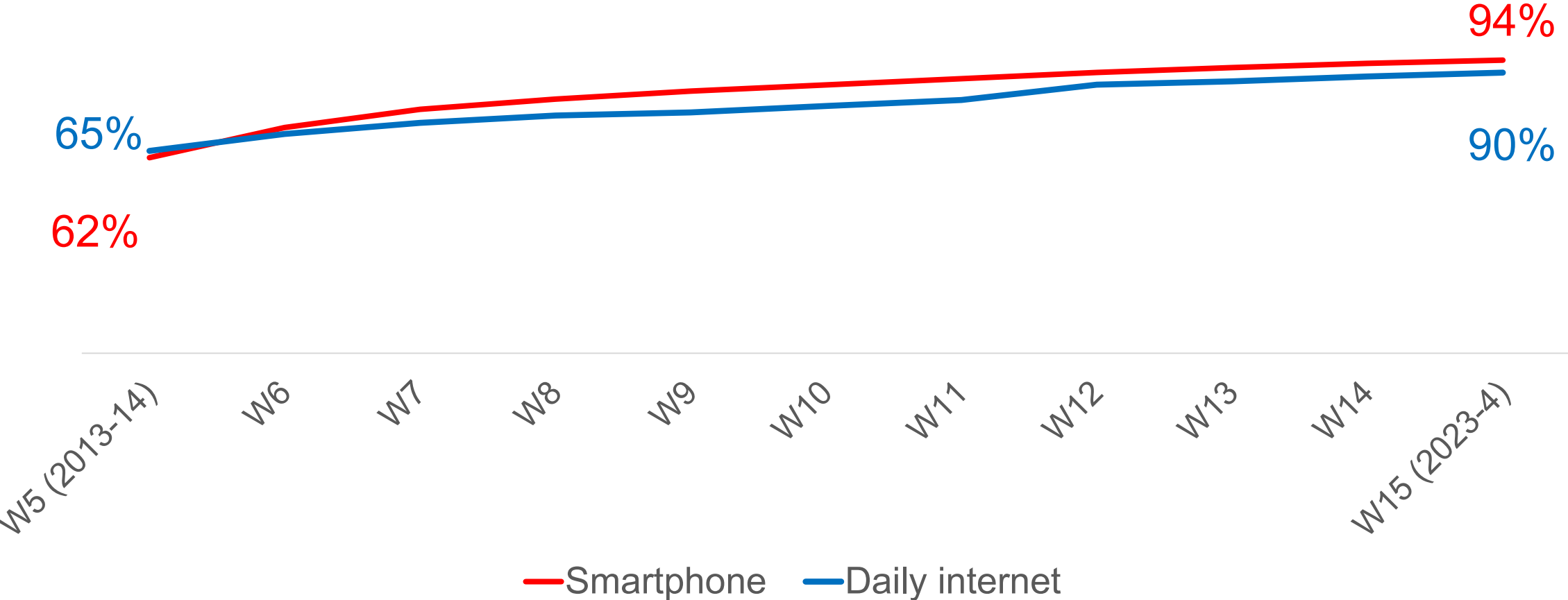
- Initial funding from Large Science Facilities Fund included “*innovative methods and types of data collection*”, and
 - “*Innovation Panel’ to trial new methods of data collection*”
 - The Innovation Panel (IP) started one year earlier than the main Understanding Society study
 - “Enhancements” budget in the current contract to try new things
-

Understanding Society has developed into a mixed-mode study

% households issued web-first
% of all Respondents interviewed online



Over the last ten years: Growth in smartphone ownership and daily internet use



What have we tried on Understanding Society?

- Consent questions
 - Link to admin data
 - Link to social media data



What have we tried on Understanding Society?

Apps

- Spending
 - Apps with upload pictures of receipts
 - Spending diary app
 - Well-being
 - Daily app-based questionnaire
 - Bio-measures
 - Body Volume scanning app
 - Cognition measures
 - “Sea Hero Quest” navigation game (Alzheimer’s Research)
-

What have we tried on Understanding Society? 2

- Event-triggered data collection
 - Monthly “life events”
- Bio-marker collection
 - Microbiome sample
 - Blood sample
 - Hair samples
- In progress
 - Data donation
 - Life History calendar



Experiments on the Innovation Panel to carry out additional tasks

- Annual call for proposals for new content and/or experiments
 - <https://www.understandingsociety.ac.uk/innovation-panel-competition/>
 - Either proposed and designed by us
 - Or proposed by external researchers
 - And we negotiate to add methodological experiments
-

With new content, we need to think...

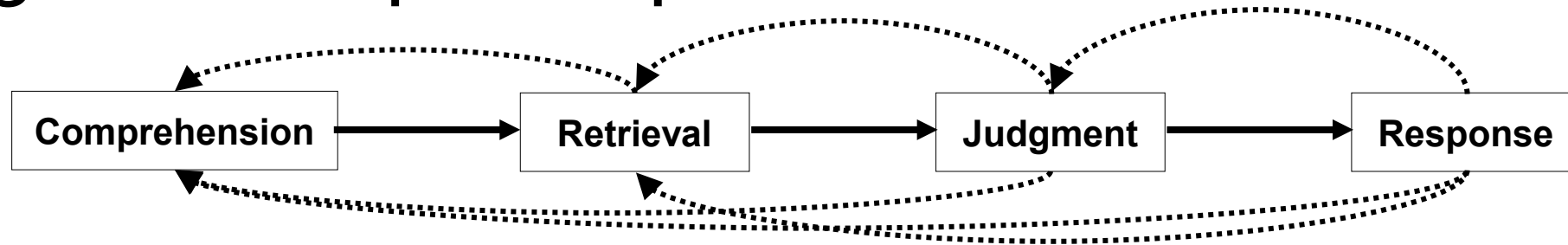
- What are the research questions?
 - How do we define the concepts of interest?
 - What do we want the data to “look like”?
 - Unit of analysis
 - Granularity
 - Regularity
-

What can we ask respondents to do?

- What can respondents do?
 - What are they willing to do?
 - How can we make the task as easy as possible?
 - Balance against burden
-

Can respondents answer question?

- Cognitive response process



- Do respondents know information?
- Easy to recall?
- Are they willing to report?

Innovation research process (ideally...)



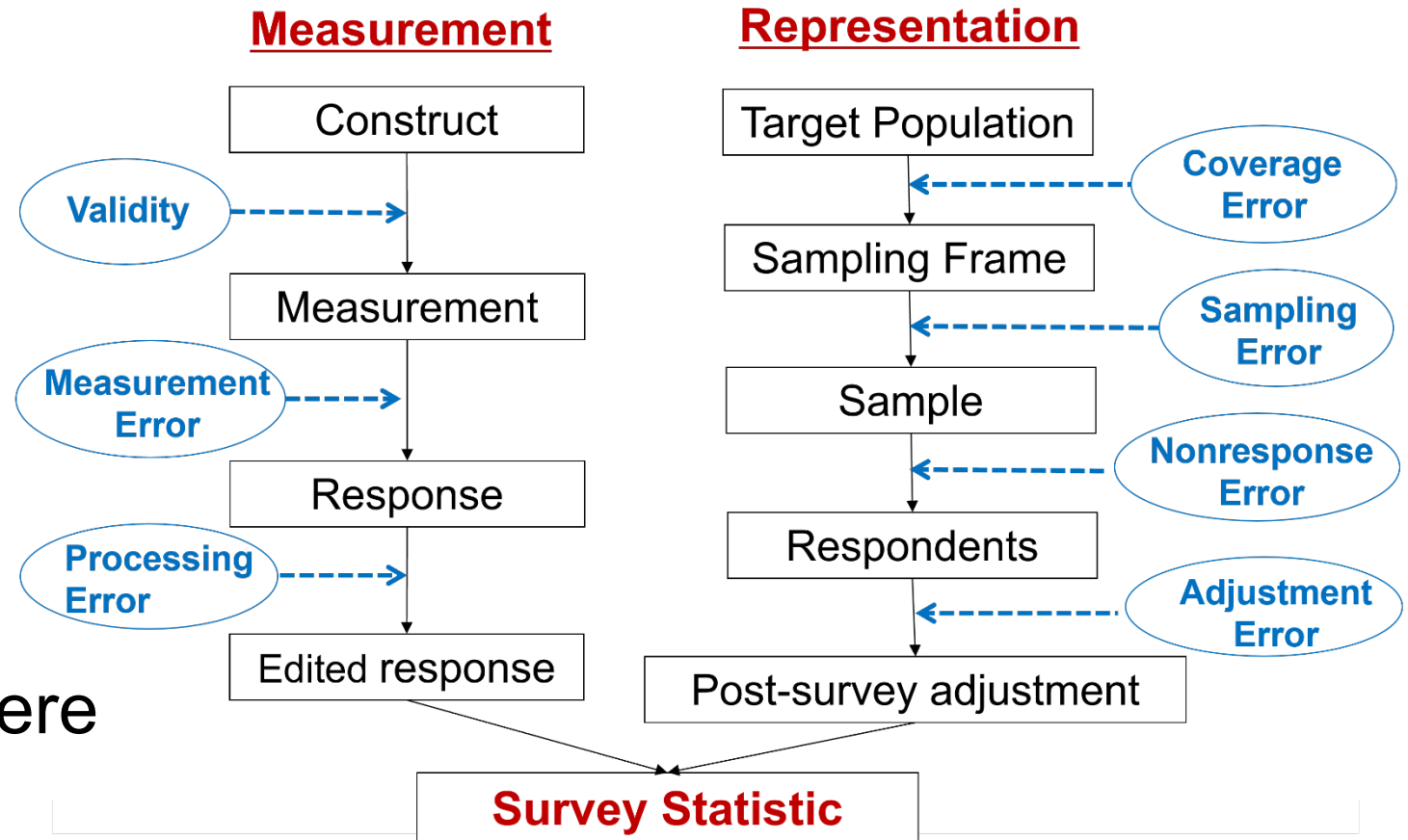
- Qualitative research
 - What do respondents know about the subject
 - What would they be willing to do
 - What obstacles might they face
 - What would engage them with the task
 - Invitation to the task
 - Experimental groups to test effects on participation rates/bias
-

Innovation research process (ideally...) - 2


- Context questions – effect on selection bias
 - Socio-demographics of respondents
 - Task-related behaviours (e.g., smartphone use)
 - Task-related outcomes (e.g., spending, health)
 - Feedback after invite
 - Reasons for/against taking part
 - Any problems or issues
 - Feedback after task
 - Ease of participation, obstacles/problems, willingness for future
-

Where can things go wrong?

- Total Survey Error framework
- Possible errors in selection and measurement
- Identify key points where design decisions intervene



Source: Groves et al. (2009) *Survey Methodology*. Hoboken: John Wiley & Sons.



Asking people to let us link
administrative data to their
survey answers

Why link data?

- Valuable, rich resource for researchers to link survey data with administrative (or other) data
 - (Potential to) reduce respondent burden
 - Better quality data (no recall or measurement error)
 - But admin data lacks context and is collected for other specific purposes, and for specific populations
 - Linkage combines depth of administrative data with breadth of survey data
-

Challenges for linking data

- Individual-level linking may require agreement – **consent** – for legal or ethical reasons
 - Non-consent affects the linked dataset
 - Fewer cases
 - Potential bias
 - We want to maximise informed consent
-

Legal issues in data linkage

- What is the legal basis for linkage?
 - UK/Europe – often Public Task/Interest, **not** “consent”
 - Is linkage compatible with original purpose?
 - How are identifiability and minimisation handled?
 - Do subject have sufficient information on which to made a decision?
 - Governance and safeguards in place (e.g., DPIAs, ethics, data sharing agreements, etc).
-

Linkage not always require consent

- No consent:
 - National Statistics Offices often rely on statutory mandate, and public task legal basis
 - UK: Office for National Statistics has a legal right to access and link data held by other public bodies – subjects are informed that the data are used for research and statistics
- Or implied consent – respondents provided with information before the survey and can withdraw consent to linkage:
 - Avon Longitudinal Study of Parents and Children (ALSPAC)
 - Family Resources Survey (FRS)

But ethically...

- We believe that it is ethical to ask respondents for their permission to link admin data
 - Which requires us to provide information at the point of consent
 - We also provide a Privacy Notice and FAQs online, and carry out a DPIA
 - We periodically remind people of consents they have given
 - ...But they can withdraw their consent at any time
 - Heightened awareness of data security and rights with GDPR – if we did not ask consent, risk of backlash against study
-

Asking consent for data linkage is not quite like other survey questions

- With most survey questions, the respondent answers
 - All the responses are anonymized and analysed together
 - Saying Yes to consent to linkage has further implications for the respondent
 - Their identifiable information is sent somewhere
 - Their records are retrieved and linked to their survey responses
 - This adds extra effort to the response process
-

“Understanding and improving data linkage consent in surveys”

- Initially a 2-year project funded by the Nuffield Foundation (2018-20)
- Continuing research programme as part of Understanding Society
- Project page:

<https://www.iser.essex.ac.uk/research/projects/understanding-and-improving-data-linkage-consent-in-surveys>



What we know from previous research



- Correlates of consent inconsistent
 - Between studies / within studies over time (e.g., Peycheva et al 2021)
 - Respondent / interviewer characteristics (e.g., Sala et al 2012, Jäckle et al 2021)
 - Large variation in consent rates between
 - Data to be linked / topics (e.g., Sakshaug et al 2012, Thornby et al., 2018, Edwards and Biddle 2021, Beuthner et al. 2024,
 - Interviewers (e.g., Korbmacher & Schröder 2013)
 - Modes of interview (Thornby et al 2018)
-

What we know from previous research 1



- Little/limited effect of question wording experiments, e.g.,
 - Length of consent question text (e.g., Edwards & Biddle 2021, Singer & Frankel 1982)
 - Emphasising benefits to respondent / survey (e.g., Pascale 2011)
 - Loss framing – inconsistent results (e.g., Kreuter et al 2016, Sakshaug et al 2019, Beuthner et al. 2024)
-

What we know from previous research - 2

- Many respondents do not understand request (Das & Couper 2014)
 - Asking earlier in interview (in context) increases consent (e.g. Sakshaug et al 2013, Sala et al 2014, Sakshaug et al 2019, Jäckle et al. 2023)
 - Half of non-consenters say 'yes' if asked again (Weir et al 2014, Mostafa & Wiggins 2018, Jäckle et al. 2018)
 - Multiple consents in one interview: latent willingness to consent (Jenkins et al 2006; Mostafa 2016)
 - But only weak latent willingness over time (Mostafa & Wiggins 2018)
-

What is the problem we wanted to tackle?

Increased interest in
data linkage

Increased shift to
web-first design

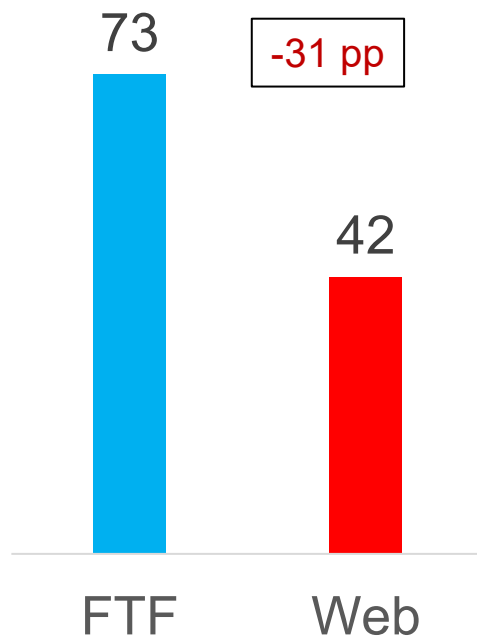
Lower
consent
rates
online!

The problem

- % Consent to link to tax records

IP11

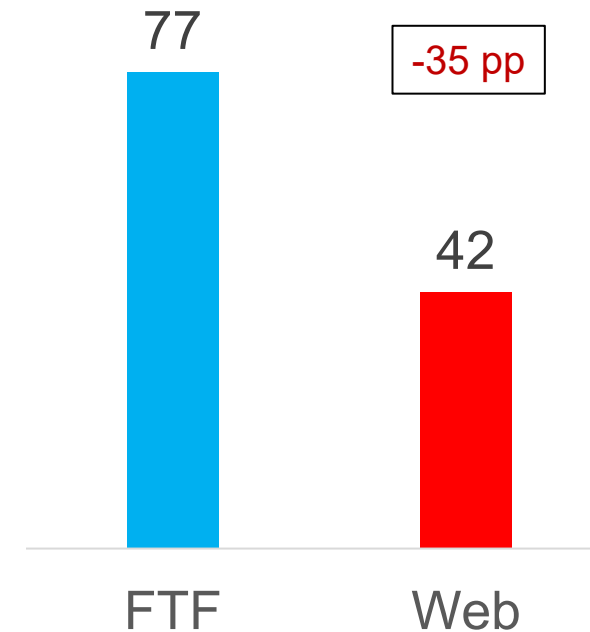
Mode of interview



- Reduces to -22 pp using Intention to Treat
- But back up to -30pp when using instrumental variables (local average treatment effect)

IP15

Mode of interview



The same in other studies

Next Steps (Age 25)

89%



69%



FTF

Web

Thornby et al., 2018, Peycheva et al, 2023

British Election Study, 2019

81%



59%



FTF

Web

Ormston et al, 2024

More evidence on this in a forthcoming Survey Futures publication

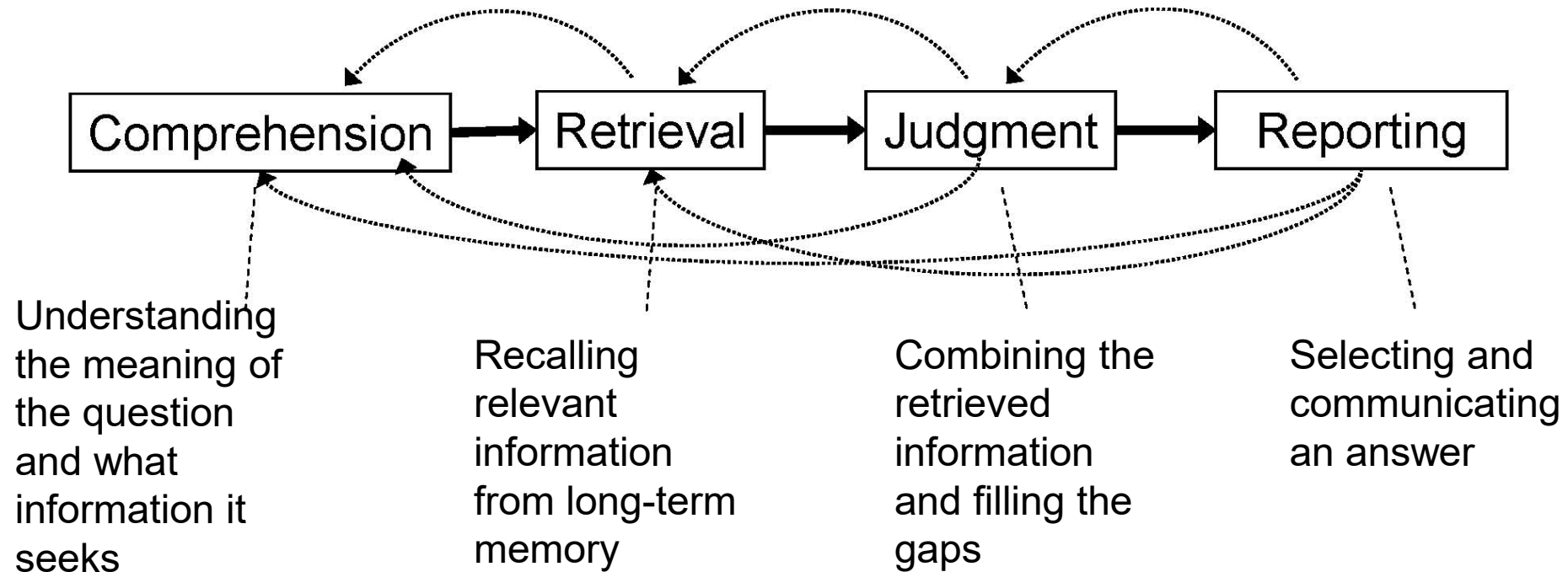


Question: Why are people less likely to consent online?

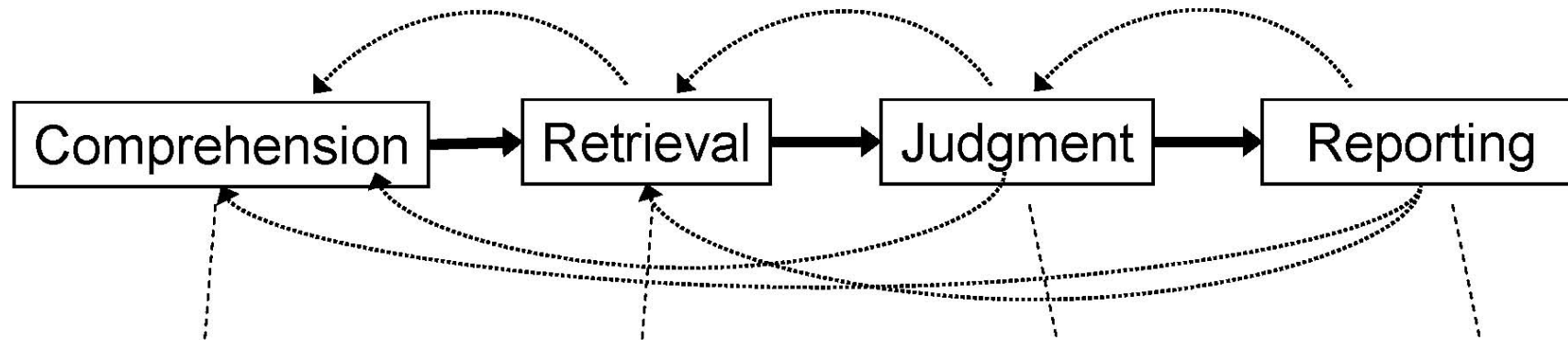
- First - *How* do people decide whether to consent?
-

How do respondents decide how to answer a survey question?

- Cognitive Aspects of Survey Methodology
- Survey response process (Tourangeau et al 2000)



How do respondents decide how to answer a survey question? Consent



- “Optimising” respondents make required effort at each stage
 - **Costs vs. benefits / consequences of consent**
- “Satisficing” respondents skip or shortcut stages (Krosnick 1991)
 - **E.g., Qualitative in-depth interviews on how respondents decide**
 - Nothing to hide
 - Trust in the survey
 - What friends/family would say
 - Habit, never provide personal information

Conceptual framework of consent decision



Based on:

- Qualitative interviews with IP respondents
 - Factors that influence consent decision (Beninger et al 2017)
 - Cognitive model of survey response process
 - How people answer survey questions (Cannell et al 1981, Tourangeau et al 2000)
 - Survey methods literature
 - Consent to data linkage, experiments
-

Conceptual framework of consent decision - 2



Based on:

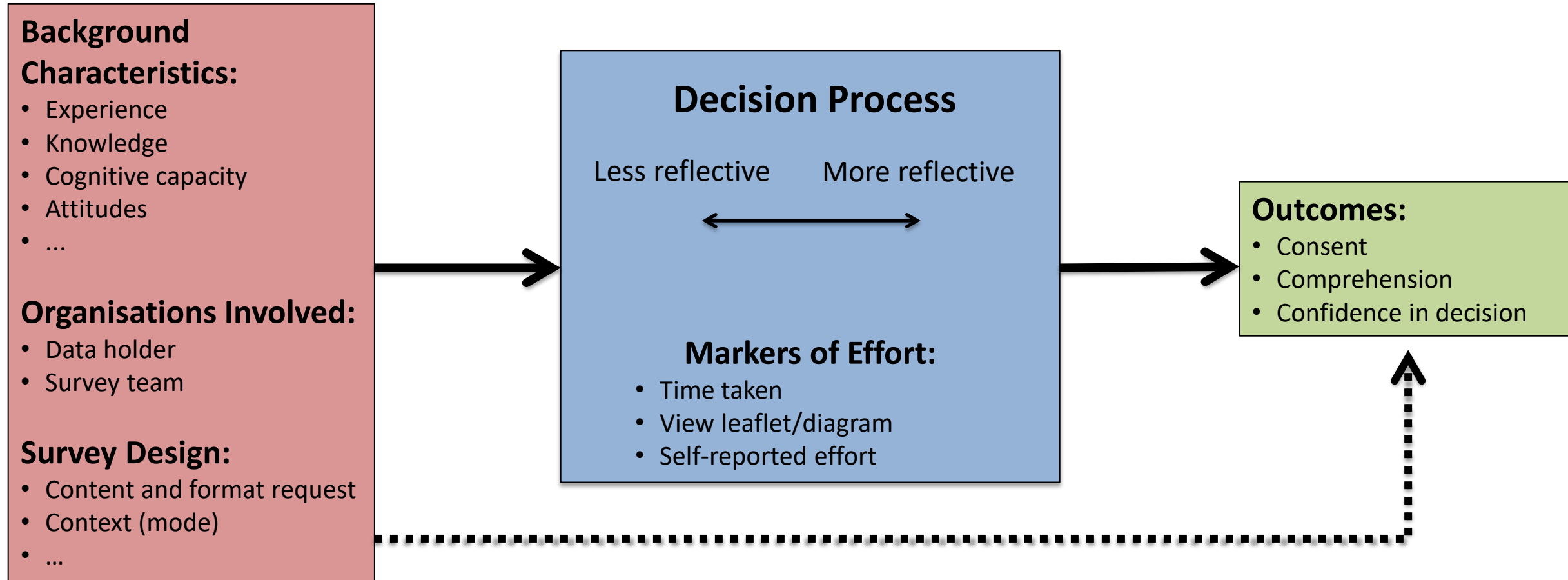
- Rational vs heuristic decision making
 - System 1 vs system 2 processing (Petty and Cacioppo 1986, Kahneman 2011)
 - Real-life decision making
 - People reduce amount of information considered (Galotti 2007)
-

Conceptual framework

Hypotheses

- Respondents necessarily employ heuristics decision processes...
 - Limited time and information when responding to a survey
- ...but heuristic employed varies across person and context
 - We label more information-intensive heuristics “**reflective**”
- More reflective decision processes associated with:
 - Higher consent rates
 - Greater comprehension
 - Greater confidence in the decision

Conceptual Framework



To test we need to collect more information

- Background questions, e.g.
 - Socio-demographics; attitudes to privacy, data security; data sharing
- Consent request
- Follow-up questions

<ul style="list-style-type: none">• Self-reported decision process• Subjective understanding of request• Objective understanding (test Qs)• Confidence in consent decision	<ul style="list-style-type: none">• Self-reported effort made• Factors taken into account• Reasons for consenting / or not
---	--

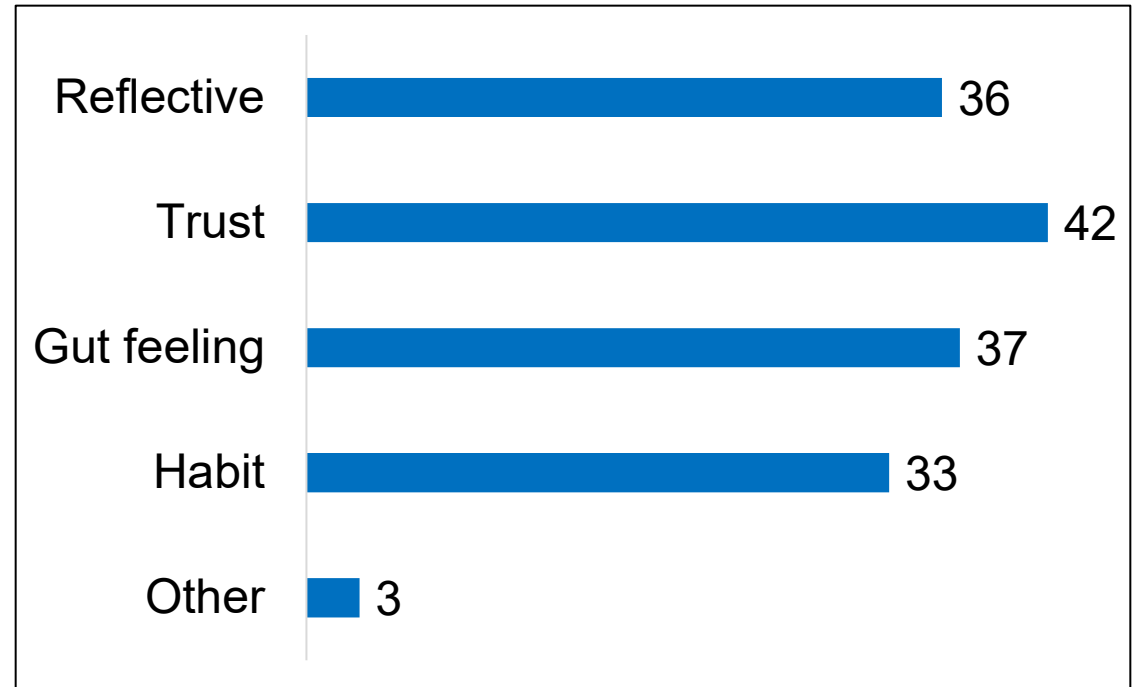
- Paradata: response times, clicks on links, interviewer observations
 - Audio recordings (CAPI): behaviour coding
-

Most respondents aren't giving the question much thought

How did you decide whether to say “yes” or “no” in response to the question about data linkage?

Please select all that apply

1. I thought about what would happen if I said “yes” or “no” (**reflective**)
2. Instinct or gut feeling (**gut**)
3. I said what I usually say when I’m asked for information that is very personal (**habit**)
4. I thought about how much I trust the organisations involved (**trust**)
5. Something else



Only around 1/3 respondents make a reflective decision

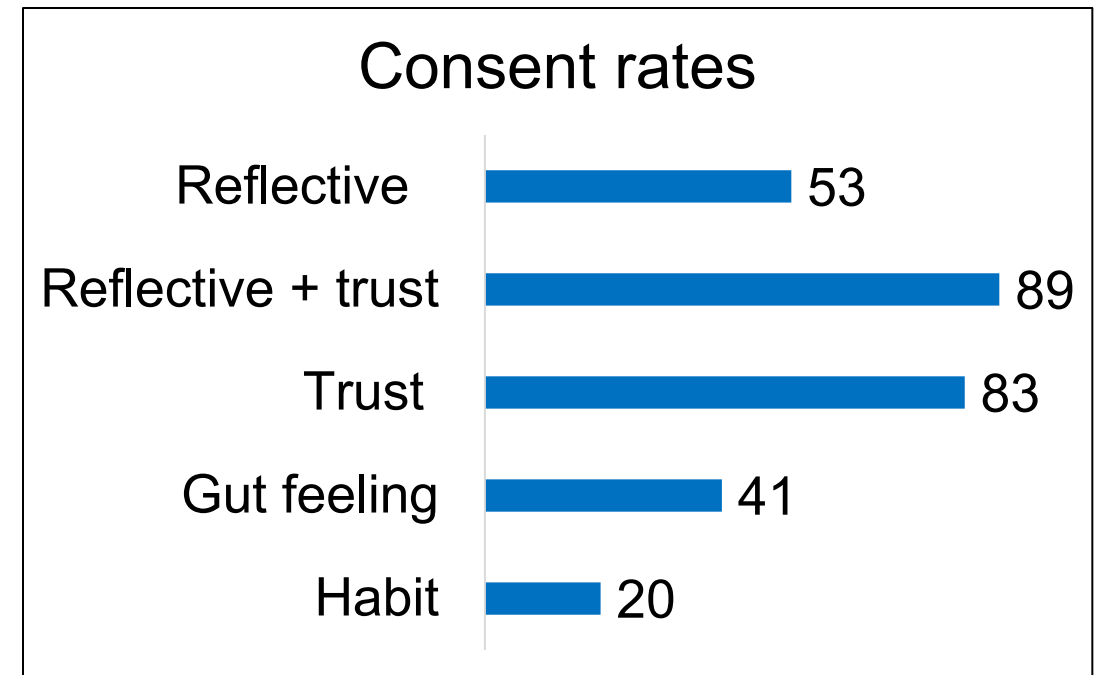
Self-reported decision process is corroborated by other indicators

- Decisions based on ‘habit’ or ‘gut feeling’ vs. ‘reflective’
 - Consent question answered more quickly
 - Less likely to click on links to leaflet or diagram explaining linkage
 - Lower self-reported effort to answer consent question
 - Less likely to say they considered ‘information’ in making their decision
 - Less likely to say they considered risks
 - Less likely to say they considered benefits

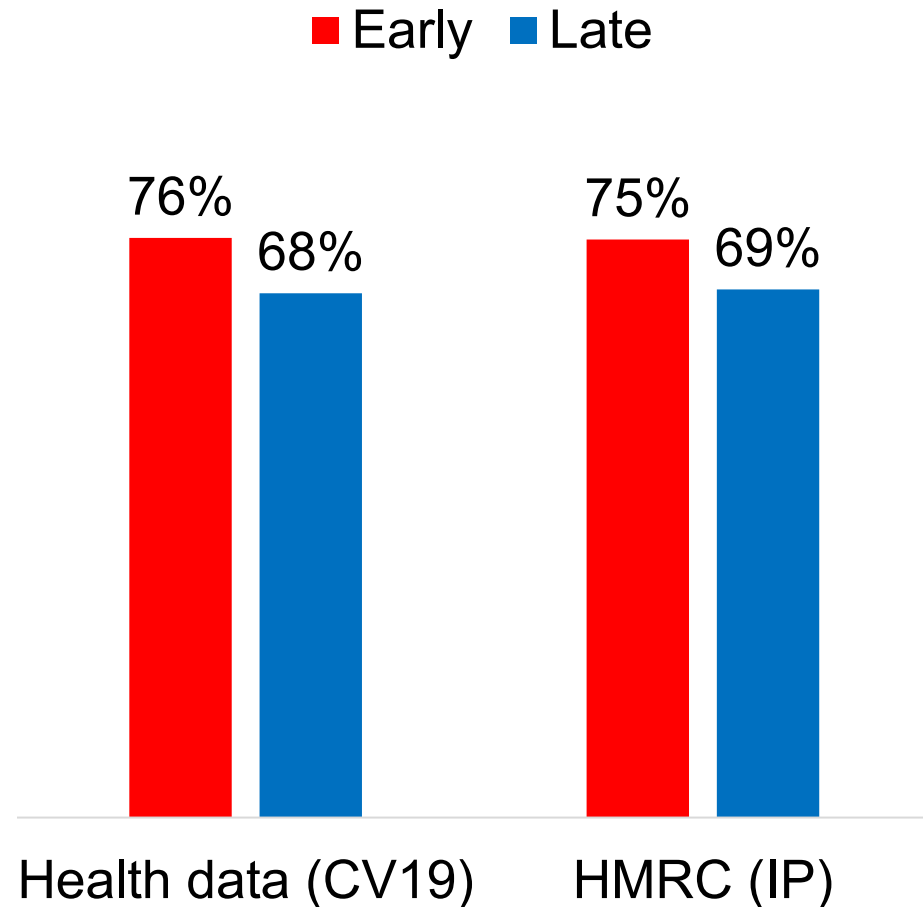
Self-reported decision process measures genuine differences between respondents

How people make the decision relates to consent outcomes

- Respondents making habit-based decisions
 - Least likely to consent
 - Over time: most likely to report same decision process and same consent decision



Asking early in the interview seems to help

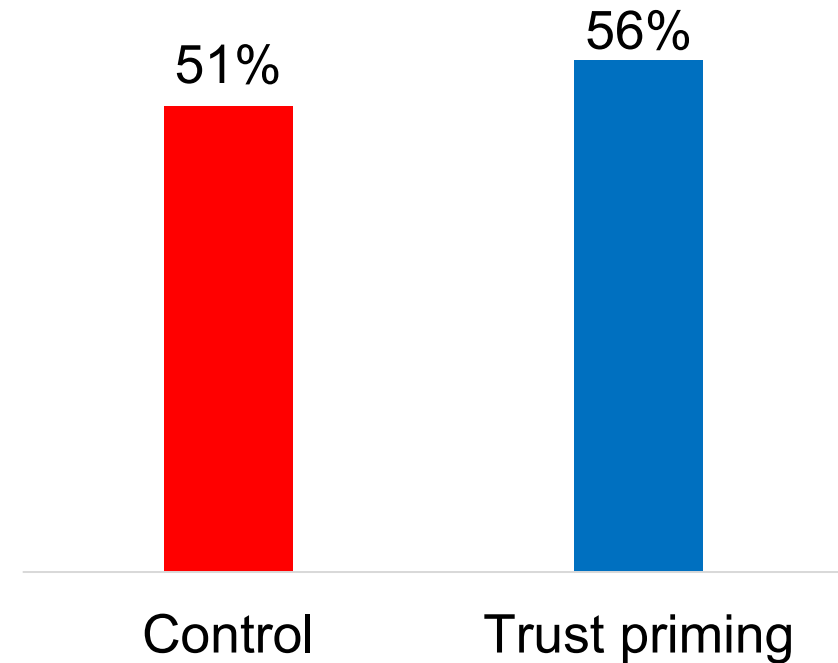
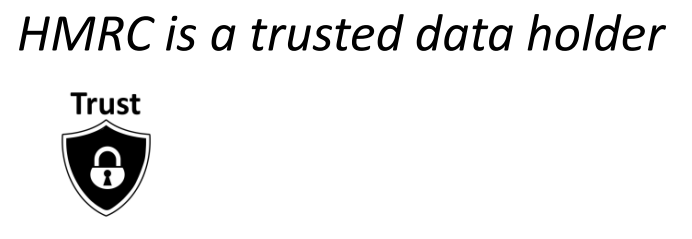


“Can you tell us why you did not give us permission...?”

- When consent asked late – significantly more likely to say:
 - Too tired to decide
 - Unclear about risks involved
 - Unclear why, what the data are for
 - Too personal, shared enough

Emphasising trust improves consent online

The next question is about linking the information you provide in this survey, to data that HM Revenue and Customs, or HMRC, hold about you.



Results from experimental studies



Ask for consent early in the interview, rather than at the end

- 'Giving fatigue' more likely if ask at the end

Emphasize trust in the organisations involved in the data linkage

Include all key information about the linkage in the consent question

- Respondents unlikely to read information leaflets/diagrams

When asking multiple consent questions, start with the one that gets highest consent rate

Simplify readability of consent question wording

- e.g. using Flesh-Kinkaid readability score implemented in MS Word

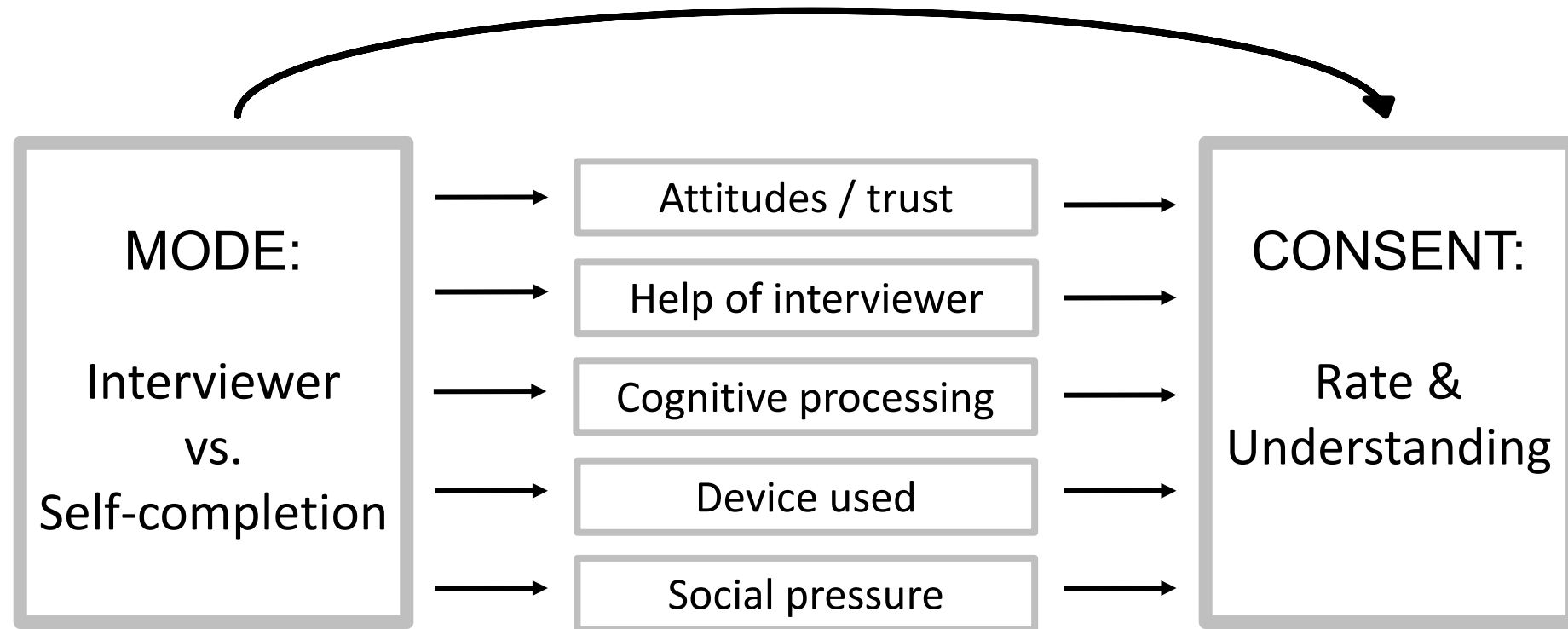
Multiple consent questions can be shown separately or on a single page

So what are online respondents doing?



Conceptual framework

How mode affects the consent decision process

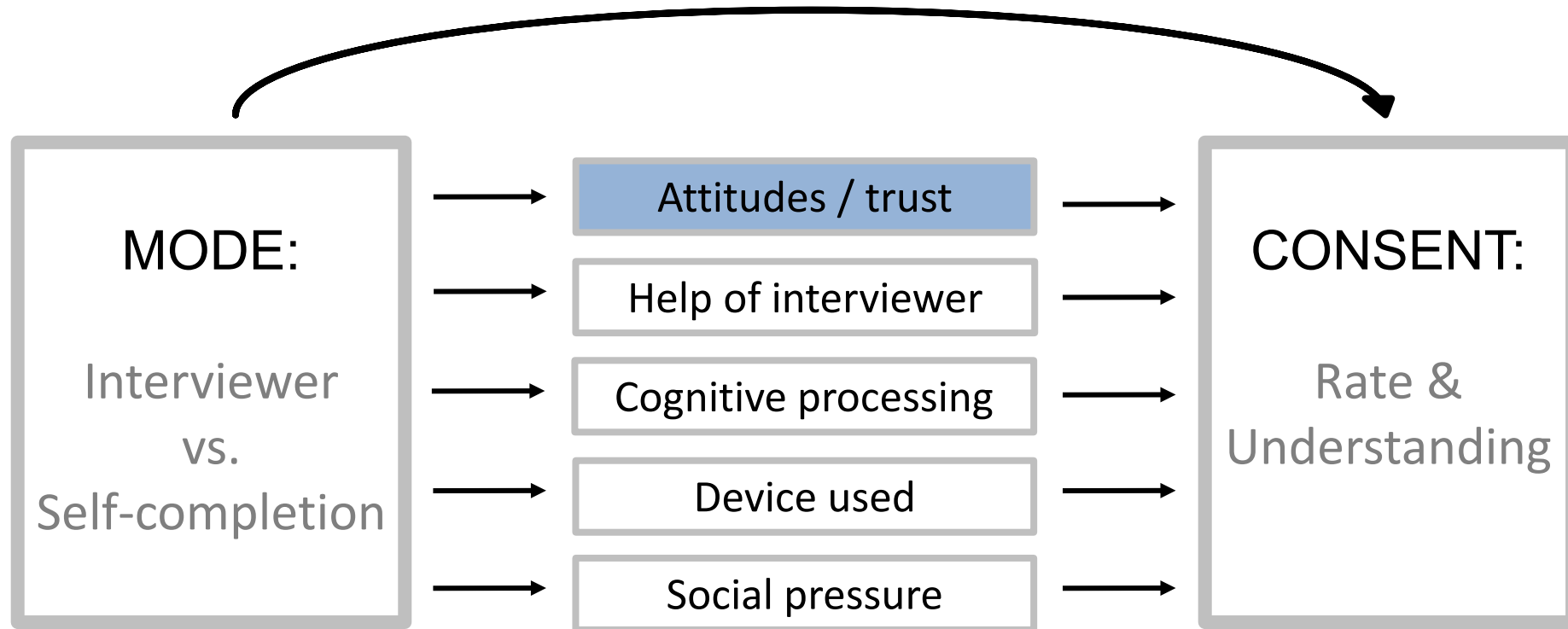


Consent and modes study 1

- Understanding Society Innovation Panel
 - Wave 11 (2018)
 - N=2,608
 - Consent to link tax record data
 - Mixed mode data collection
-

Conceptual framework

How mode affects consent





1. Potential mechanisms:

Respondent attitudes / trust

- Trust is positively associated with consent (Al Baghal 2016; Sakshaug et al 2017; Sala et al 2012)
 - Trust-based consent decision process: higher consent rates (Burton et al 2021)
-



1. Potential mechanisms:

Respondent attitudes / trust - 2

- Qualitative interviews (Beninger et al 2017):
 - *"My default is not to... I'd definitely, outside of the Understanding Society, say no but I **trust** them."*
 - *"I guess I am more likely to consent in person because I **don't know who is asking [questions] online.**"*
 - *"They may think **online it could go anywhere** because we all know once it's online it is for everyone to see. I think face to face they would feel more comfortable."*
-

Attitudes about privacy/data security

Asked early in survey, before consent

Privacy concerns

In general, how worried are you about your personal privacy?

1. Very worried
2. Somewhat worried
3. Not very worried
4. Not worried at all

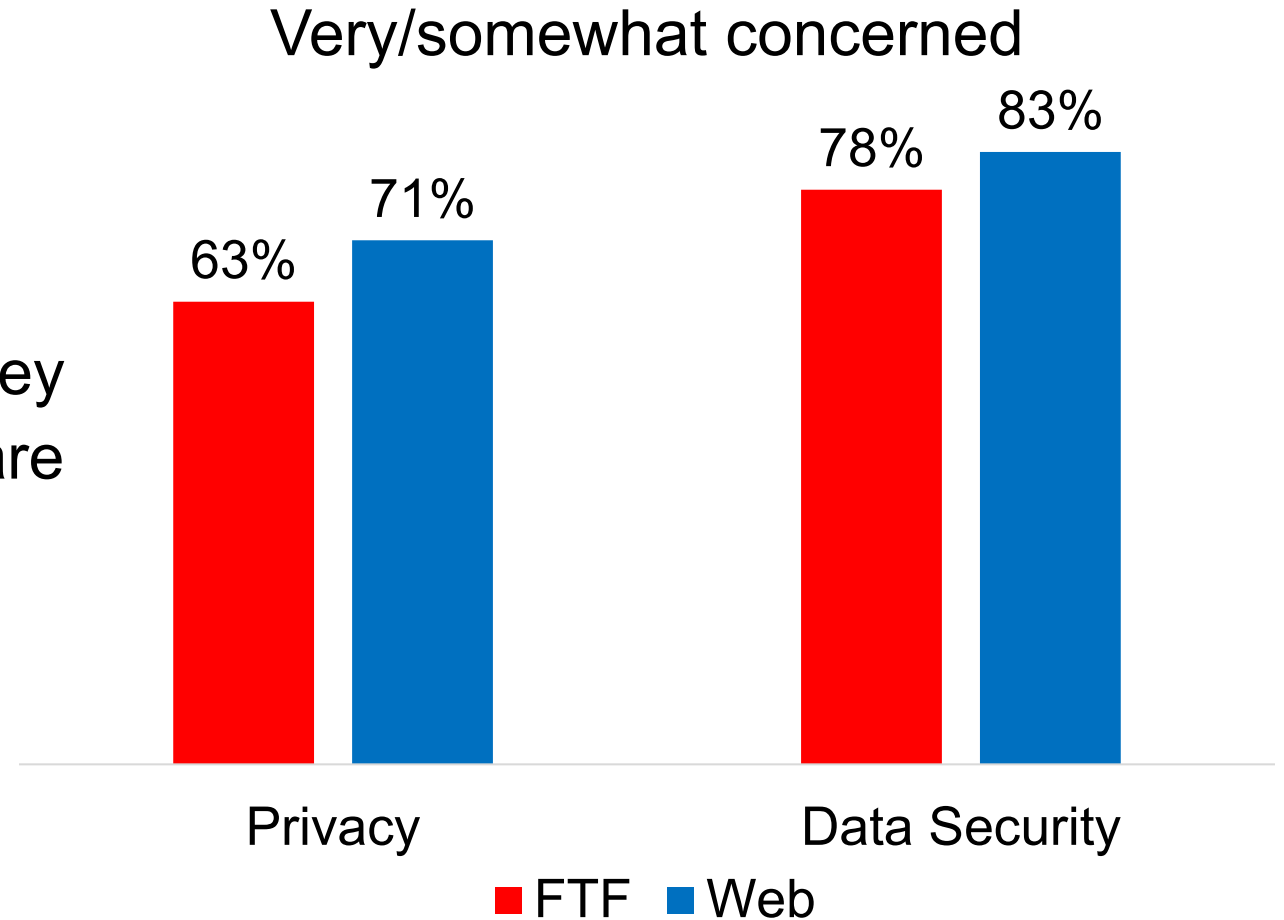
Data security concerns

Different private and public organizations have personal information about us. How concerned are you about whether or not they keep this information confidential?

1. Very concerned
2. Somewhat concerned
3. Not very concerned
4. Not concerned at all

Respondents more concerned about privacy when answering online

- But... no difference in how much they trust HMRC or belief that the data are sensitive



Try to “prime” the respondent to think about trust

- PopulusLive online access panel
- Great Britain
- Quota to match Innovation Panel
- Same questions
- ‘Easy’ consent question
- N=1,921

The next question is about linking the information you provide in this survey, to data that HM Revenue and Customs, or HMRC, hold about you.

HMRC is a trusted data holder



Experiment

Mentioning trust increased consent

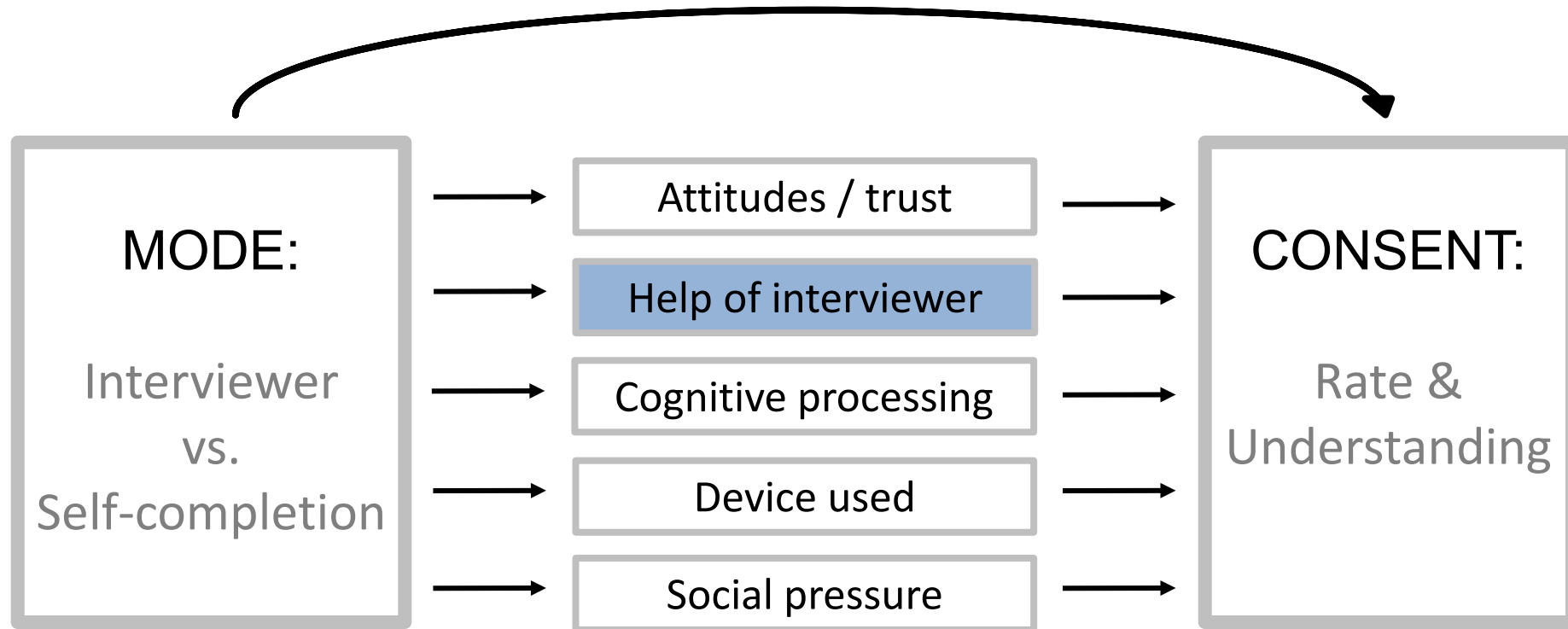
- Consent rate:
 - +5 pp
- No effects on
 - Understanding
 - Objective understanding
 - Confidence in decision
 - Decision process

Relatively large effect
Equivalent to 1/6th of mode gap

Trust is important
Survey design / setting can affect trust
Sub-conscious?

Conceptual framework – 2

How mode affects consent



2. Potential mechanisms: Interviewer behaviours

- Qualitative interviews (Beninger et al 2017) :
 - "...if someone is talking to you about it they can ... **alleviate the concerns** that just can't be answered by reading on the screen. You never feel you are quite getting enough information to make an informed decision when you are doing it through a screen".
 - "If (regular interviewer) had done it face to face she usually **gives extra detail** and sometimes she can tell by the blank look on my face. She does tend to go into more detail sometimes."
-

2. Potential mechanisms: Interviewer behaviours - 2

- Audio-recordings
 - FTF interviews
- Coded what interviewers & respondents did
 - Consent question

Interviewers – rarely

- Emphasized confidentiality (4% cases)
- Offered additional information (14%)

Respondents – rarely

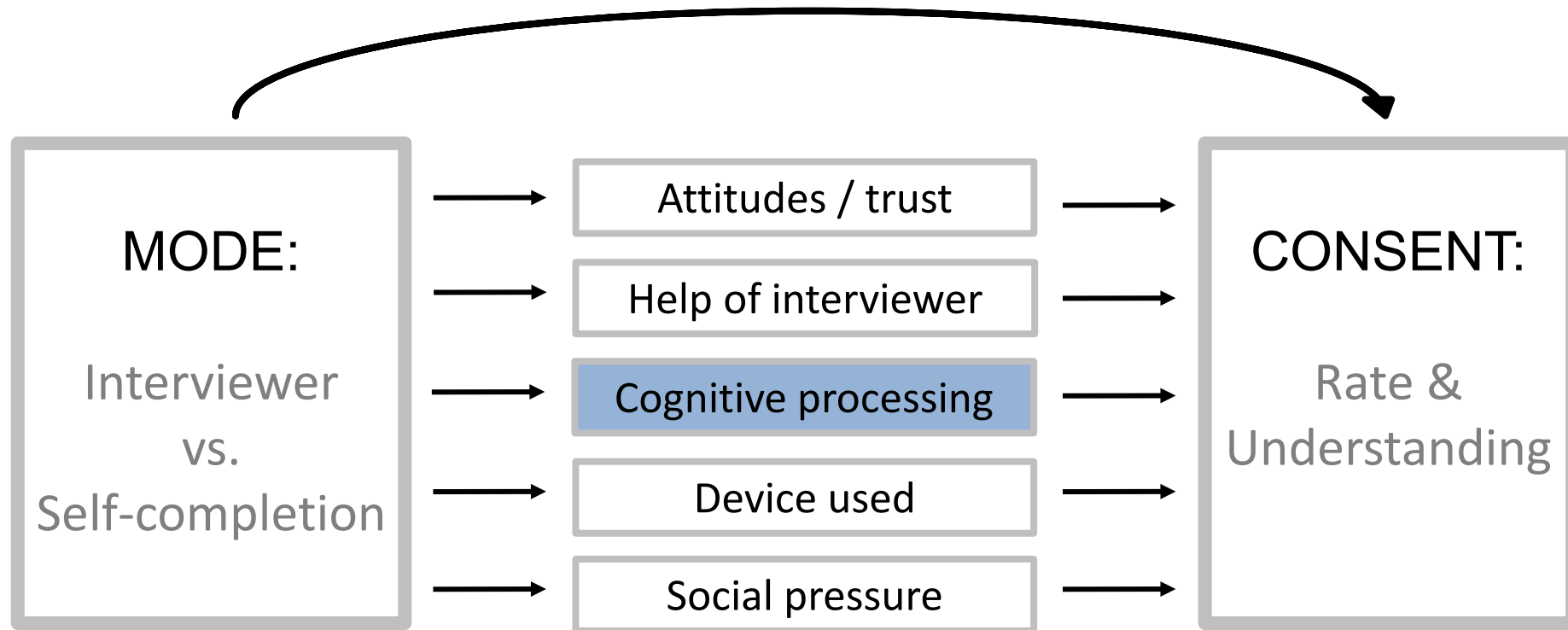
- Expressed concern/uncertainty (5%)
- Asked questions (16%)

Similar conclusions for CATI/web (Sakshaug et al 2021)

Verbal Interviewer behaviours not driving mode differences

Conceptual framework - 3

How mode affects consent



3. Potential mechanisms:

Cognitive processing

- Qualitative interviews (Beninger et al 2017) :
 - "Because the impact isn't as great is it, if you read something online. It's like, you know, **it's like reading an email at work - you might not take that information in.** But when you're doing it face to face, and you're sitting there, you know, I'm going to read this, aren't I, because you just asked me to."

Simplifying the question leads to increased understanding... but not higher consent

- Standard wording
 - UKHLS consent question (tax records)
 - Reading difficulty: Flesch-Kincaid Grade level = 14.5
- Easy wording
 - Lower reading difficulty: 8.2
 - Shorter sentences, shorter words, no passive sentences
 - Results from qualitative interviews
 - Text broken up into bullet points

Easy wording

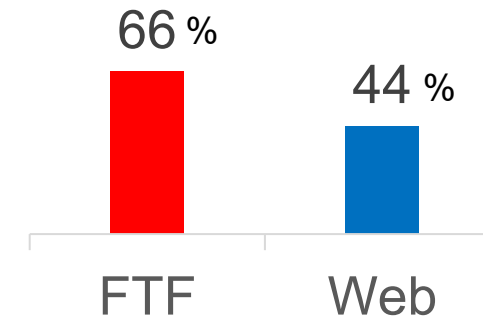
- Increased understanding
- No effect on consent
- No interaction with allocated mode

**Increasing understanding did not increase consent
Easier wording did not reduce mode effect on consent**

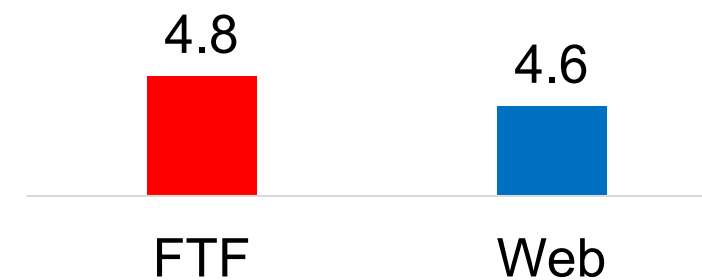
Respondents understand the request less when answering online

- Online vs CASI
- Subjective understanding: “How well do you think you understand what would happen with your data...”
- Objective understanding (8 true/false statements)
- But... no difference in confidence in decision

Completely/mostly understood:



Correct scores



3. Potential mechanisms: Cognitive processing

Consent decision process

- Decision process

“How did you decide whether to say “yes” or “no” in response to the question about data linkage?”
(*Select all*)

1. I thought about what would happen if I said “yes” or “no” (**reflective**)
2. Instinct or gut feeling (**gut**)
3. I said what I usually say when I’m asked for information that is very personal (**habit**)
4. Something else

Web treatment effect (IV)

- Reflective decision: -9 pp
- Habit-based decision: +12 pp
- Response time: 2.2 x faster
- Read/clicked leaflet: -32 pp
- Information presented was “too much”: +6 pp

Respondents answer consent Qs less thoroughly when answering online

Further attempts to encourage reflective decision making

Work in progress!!

- IP17
- Control + Treatments:
 - CAPI & Web: question text focus on benefits/risks

Focus on benefits and risks

- “Linking the information will give researchers detailed and accurate information, without having to ask you more survey questions. This will not affect the way you deal with the DWP now or in future and the DWP will not be given any information about answers you have given in this study.”
-

Further attempts to encourage reflective decision making - 2

- IP17
- Control + Treatments:
 - CAPI & Web: question text focus on benefits/risks
 - CAPI & Web: focus on importance for research and policy

Focus on benefits for research and policy



- “Linking the information will give researchers a fuller picture about who does and who does not claim pensions or benefits, or seek help from the DWP. This will tell us whether there are people who should be getting support but are missing out, and how much of a difference pensions and benefits make for people nowadays.”
-

Further attempts to encourage reflective decision making - 3

- IP17
 - Control + Treatments:
 - CAPI & Web: question text focus on benefits/risks
 - CAPI & Web: focus on importance for research and policy
 - Web: Ask respondents to think about the request
-

Asking to think more about the request



- We would like you to think about the request for permission to link information held by the DWP to your survey responses.
 - What do you think are reasons why you might want to give permission?
 - And what do you think are reasons why you might not want to give permission?
-

Further attempts to encourage reflective decision making - 4

- IP17
 - Control + Treatments:
 - CAPI & Web: question text focus on benefits/risks
 - CAPI & Web: focus on importance for research and policy
 - Web: Ask respondents to think about the request
 - Web: Ask the 4 knowledge questions before response
-

Asking the knowledge questions



- Before we ask you whether you wish to give consent to linkage with the DWP data or not, we would like to understand whether the explanation we gave you about DWP data linkage was clear or unclear.
 - Here are a few statements about how the linkage is done. Please specify whether you think each of the statements is true or false.
 - ...
-

And...?

Work in progress!

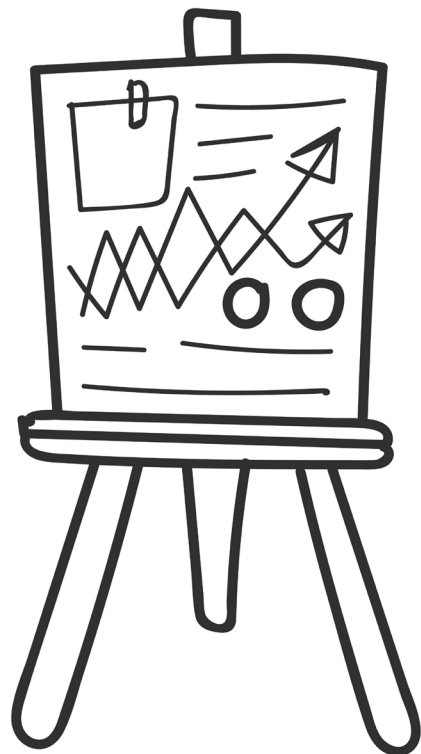
CAPI

- No significant differences
 - Those in treatment groups **not** more likely to:
 - Consent
 - Use reflective decision process
 - Have confidence in their decision
 - Put in more effort in answering
 - Mostly/completely understand process
-

But online...?

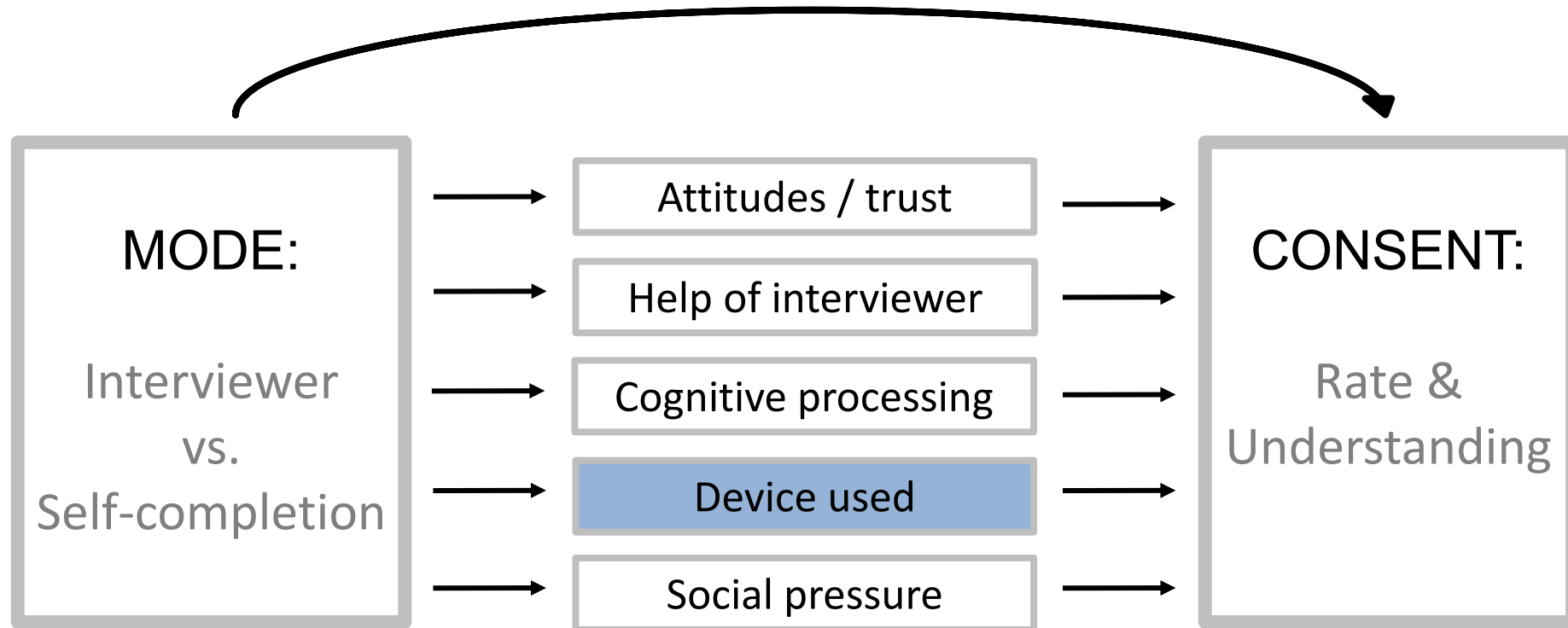
Work in progress!!

- Almost no significant differences
 - None for benefits/risks; importance for research & policy; asking knowledge questions
 - But – asking people to list reasons for/against...
 - Significantly **less** likely to consent (-8.4pp)
 - **Less** likely to understand process
 - **No** effect on reflective decision making
-



Conceptual framework - 4

How mode affects consent





4. Potential mechanisms:

Device used to complete web survey

- Privacy / data security concerns
 - Might be influenced by the devices respondents use to complete the survey
 - Especially if using public wifi
 - Smaller screen sizes
 - Might make it harder to read and process the consent question
-

4. Potential mechanisms:

Device used to complete web survey - 2

- Devices used by web respondents:
 - 57% PC, laptop, notebook
 - 29% tablet
 - 14% smartphone
- Logit models to control for respondent characteristics

No differences in

- Probability of consent
- Objective understanding
- Privacy / data security concerns
- Consent decision process
- Whether clicked leaflet/diagrams

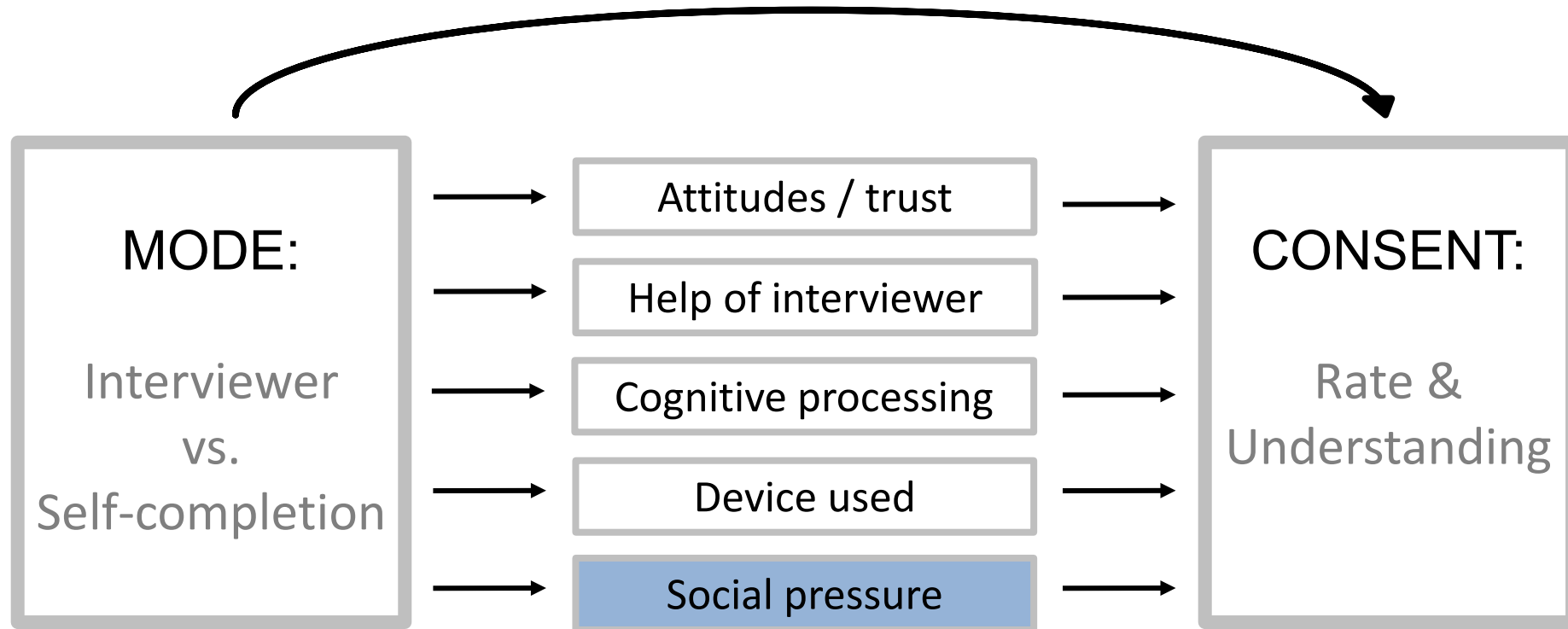
Smartphone users

- Information presented: “Too much”
- Took longer to answer consent Q

Devices not driving mode effects

Conceptual framework - 5

How mode affects consent



5. Potential mechanisms:

Social pressure

- Qualitative interviews (Beninger et al 2017):
 - "Doing it **online is much easier to say no** because usually it is just click a box whereas if someone is talking to you about it they can either be persuasive depending on whether they are supposed to be doing that ...".
 - "**Subtle pressures you get in a face to face...you don't want to let that person down.**"
-

Consent and modes study 2

- Understanding Society Innovation Panel 15 (2022)
- Consent to link tax record data: F2F respondents:
 - **CAPI**: interviewer reads Q and records answer [nudge to deeper cognitive processing, higher social pressure]
 - **Partial-CASI**: interviewers read out Q, respondent records answer in self-completion module [no social pressure]
 - **CASI**: consent Q in self-completion module [no nudge towards deeper processing, no social pressure]

5. Potential mechanisms:

Social pressure - CAPI

- Consent rates:

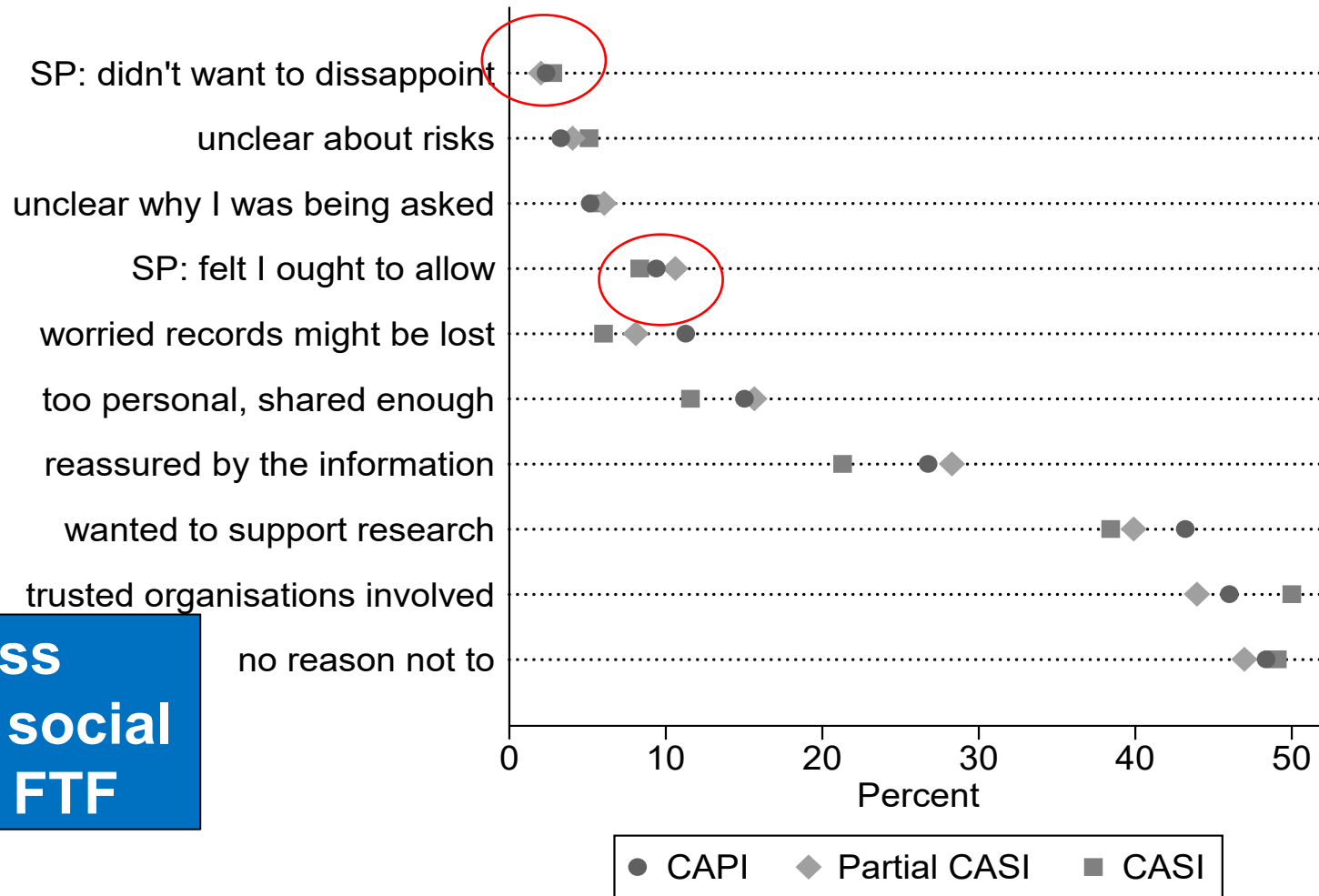
	%	n
CAPI	77	213
Partial-CASI	77	198
CASI	85	216

Expected:
Highest rates in
CAPI

- Consent decision process:
 - CAPI more likely to report gut feeling based decision
 - (Expected more reflective decisions with CAPI)
 - Objective knowledge test: n.s.
 - Reasons for consenting: n.s. (incl. indicators of social pressure)
-

5. Potential mechanisms: Social pressure

“When you were deciding... which of the following applied to you?”



**No success
manipulating social
pressure in FTF**

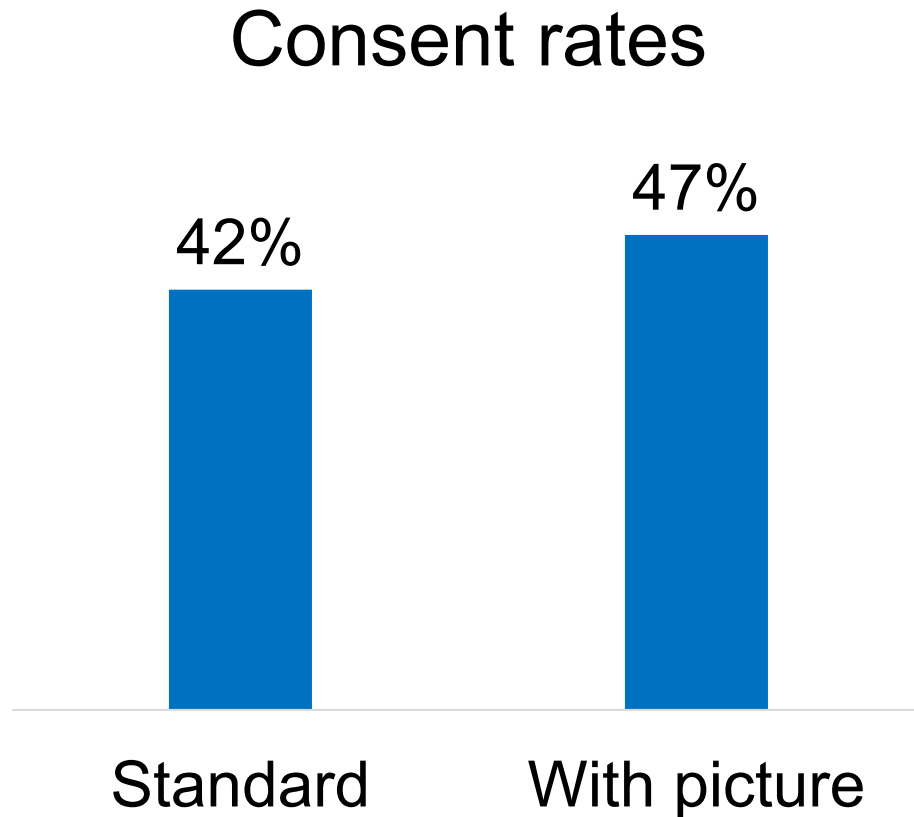
Consent and modes study 2 - Web

- Understanding Society Innovation Panel 15 (2022)
- Consent to link tax record data: Web respondents:
 - Standard question
 - Question with message from Study Director



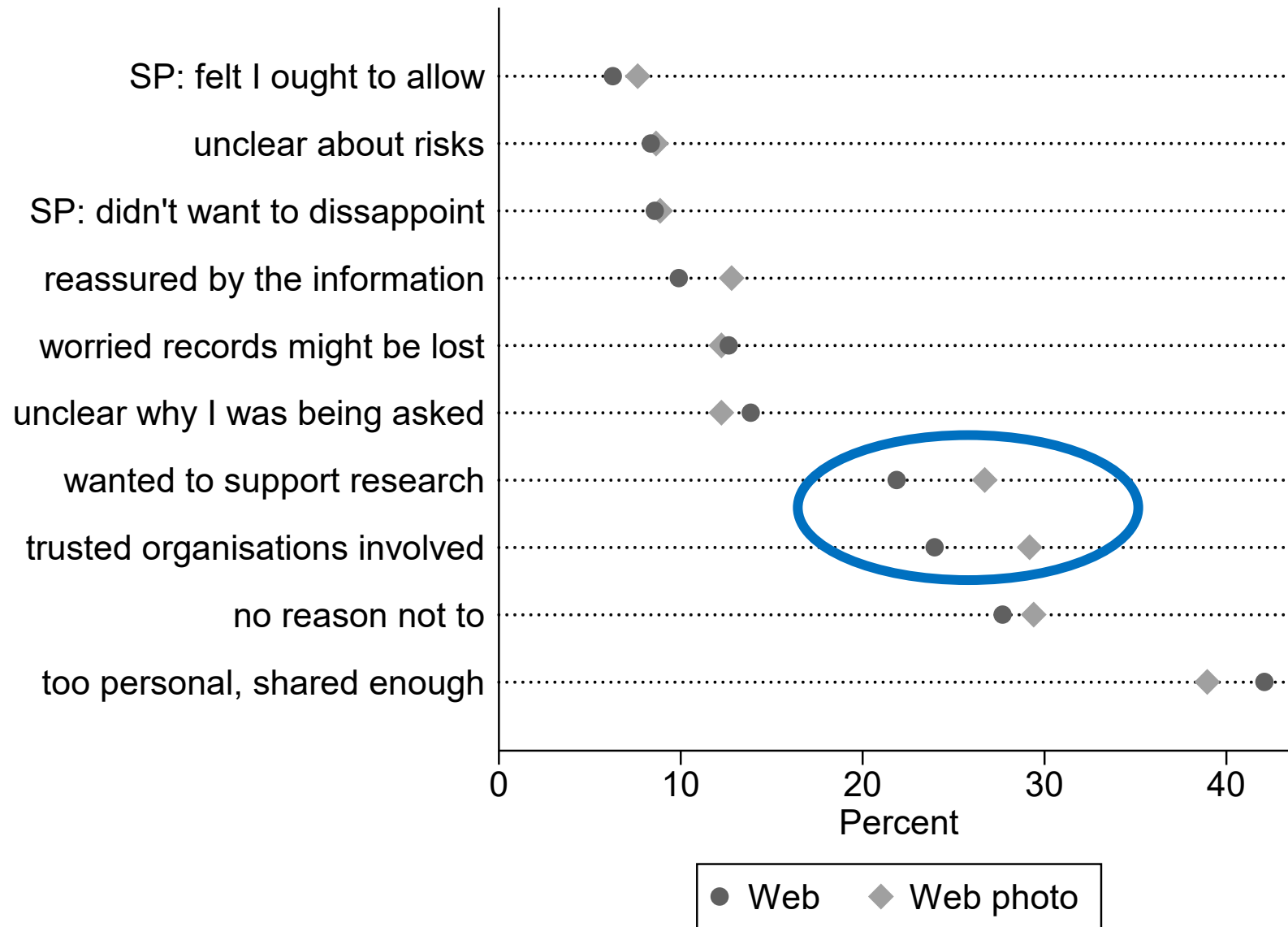
Hello, I am Professor Michaela Benzeval, Director of Understanding Society. Linking your survey answers with administrative records while preserving the confidentiality of your data increases the value of the information you provide. Please read the following carefully.

5. Potential mechanisms: Social pressure - Web



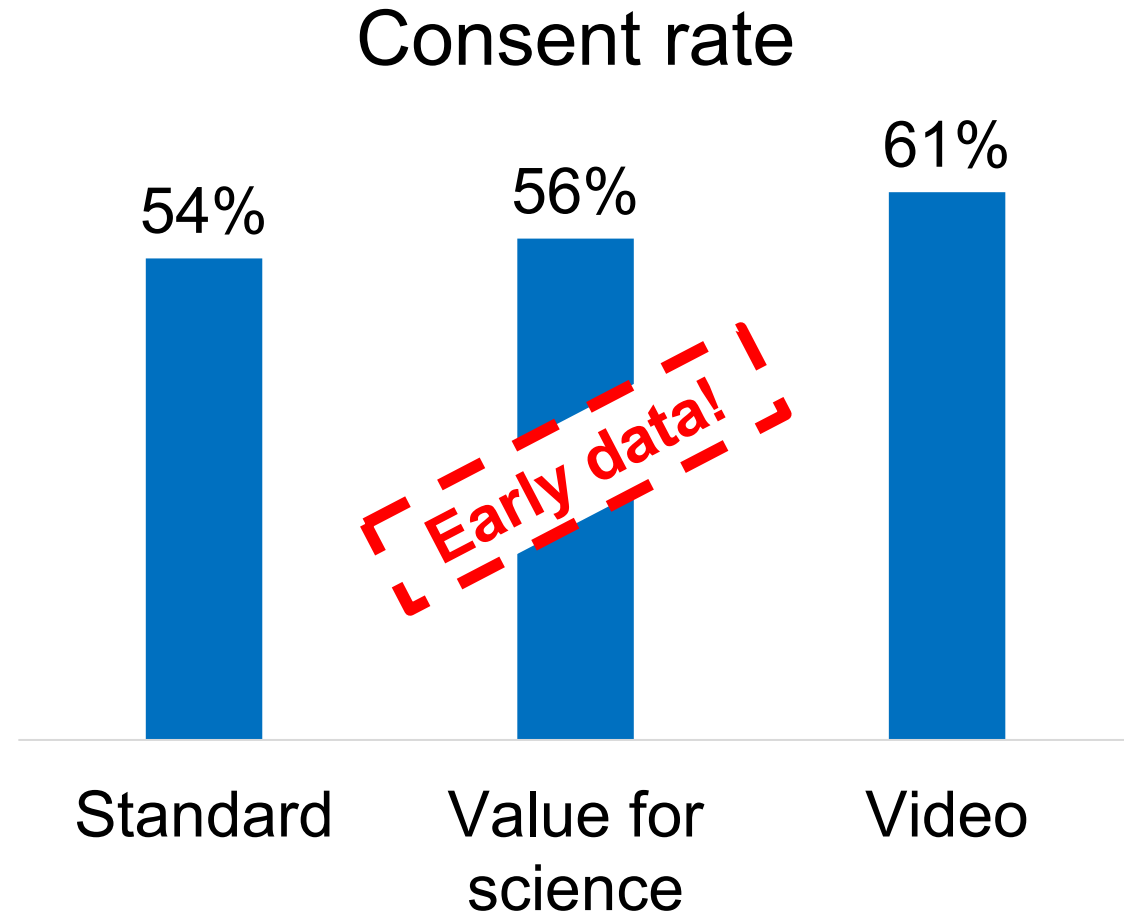
- Significantly higher consent rate
- More likely to consider trust when making decision
- No effect on objective knowledge

Those who saw the picture more likely to say they consented because of trust or support research



5. Potential mechanisms: Social pressure II

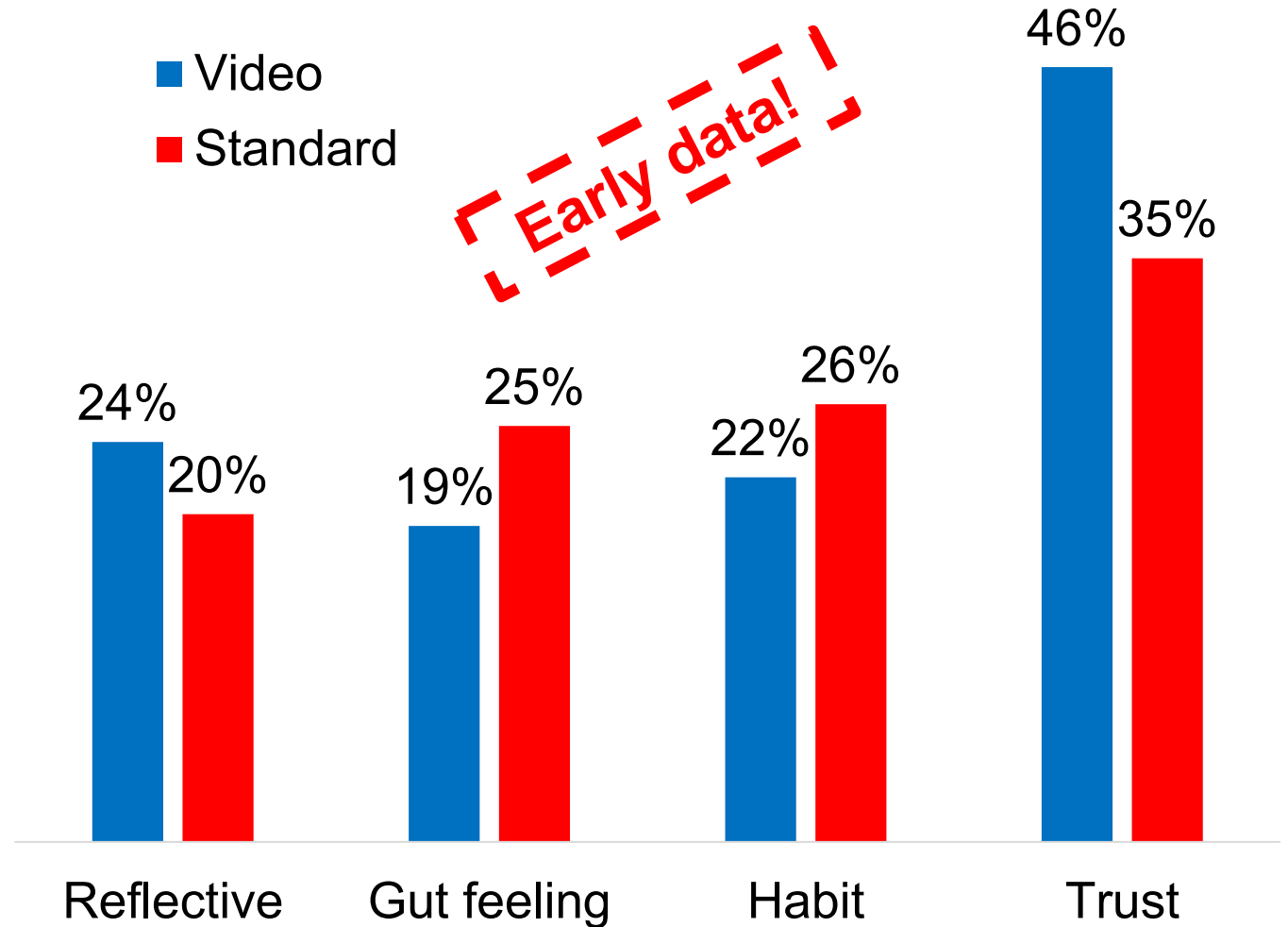
- **IP18**
- Experiment with video of interviewer reading question
- vs standard
- vs “value of science”
- **Final data due in March**



5. Potential mechanisms: Social pressure II - 2

Video treatment

- More likely to prompt reflective decision process
- More likely to think about trust in the organisation
- Also higher reported cognitive effort



Summary of how mode may affect consent



- Respondents use different decision processes
 - Only a minority use a reflective decision process
 - Reflective decision process associated with consent and comprehension
 - Limitation: causality not established
-

Respondents answer less thoroughly when answering online

- Less likely to make a reflective decision (-9pp)
- More likely to make a habit-based decision (+12pp)
- Respond faster (2.2x)
- Less likely to read information leaflet (-32pp)
- More likely to feel information given was “too much” (+6pp)

It doesn't appear to be...

- Devices
 - No differences in consent, understanding, privacy/security concerns, decision processes, accessing information
- (audible) Interviewer-interaction behaviours
 - Rarely emphasise confidentiality (4% cases) or offer additional information (14%)
 - Respondents rarely express concern (5%) or ask questions (16%)

Will making the question easier help understanding?

- Yes
 - “Easy” wording, increases objective understanding (4.5 → 5 correct)
 - Higher objective understanding associated with higher consent (not causal)
- But...
 - Easy wording did not reduce the mode effect
 - Interview mode affected consent, regardless of wording

What helps?

- ✓ Asking early in the interview
 - ✓ Priming with trust
 - ✓ Introducing 'social pressure' online
 - ? Making the question wording easier
 - X Providing additional information
-



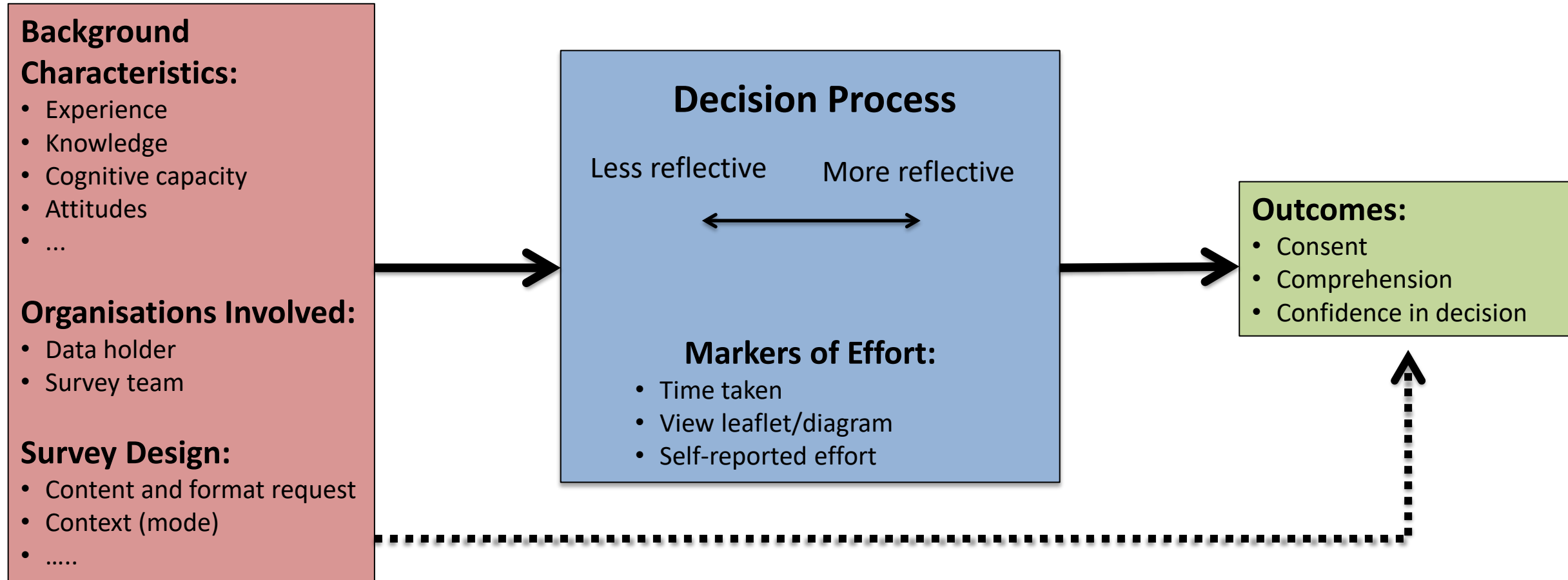
Moving to more active participation



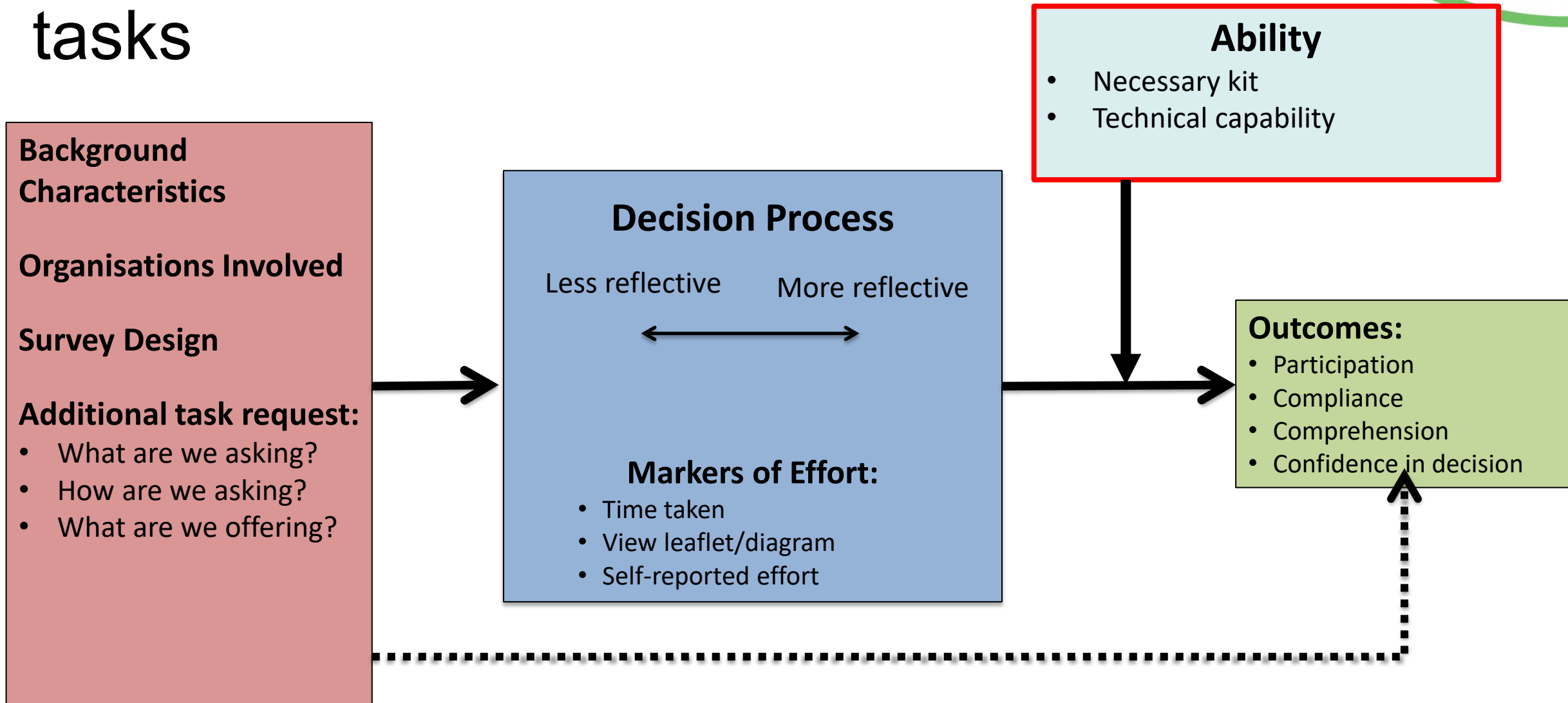
Moving to more 'active' decisions...

- Consent to data linkage is passive
 - The respondent says Yes/No
 - They don't have to do anything else
 - If they agree, additional tasks happen elsewhere by other people
 - Other examples of additional tasks require the respondent to actively do something
-

Conceptual Framework for consents



Conceptual Framework for additional tasks



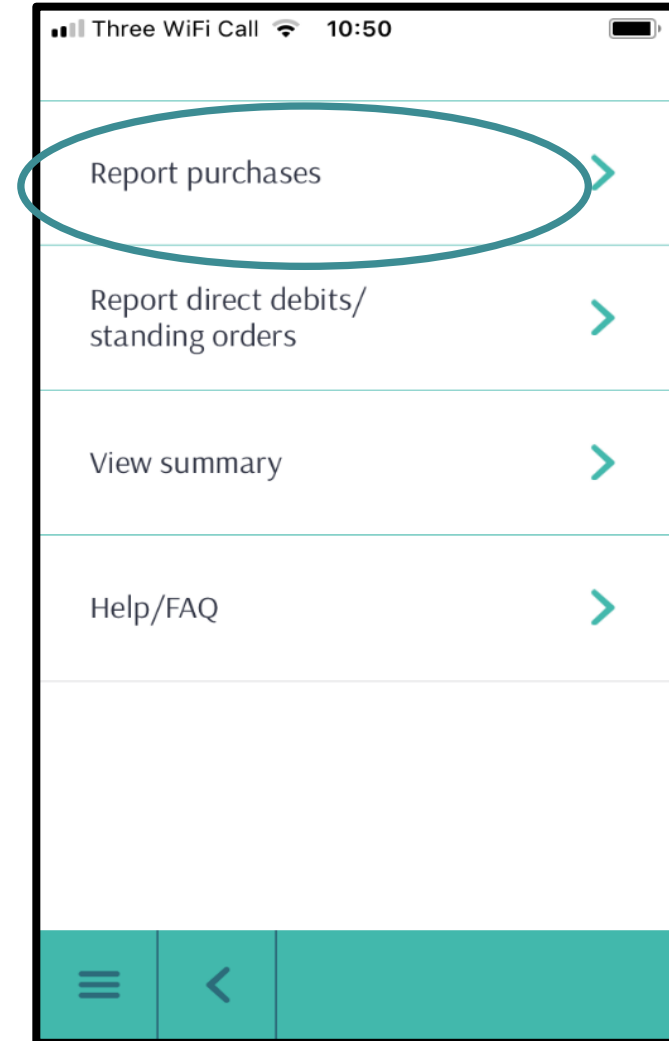
Additional Task Requests...



- What are we asking the respondent to do?
 - Time
 - Effort
 - Discomfort
 - Risk of disclosure
 - Intrusion of privacy
 - What are we offering?
 - Incentive
 - Feedback
 - “benefits to science”, altruism
 - How are we asking?
 - Wording
 - Placement
 - Additional information
 - Mode
 - Timing
 - ...
-

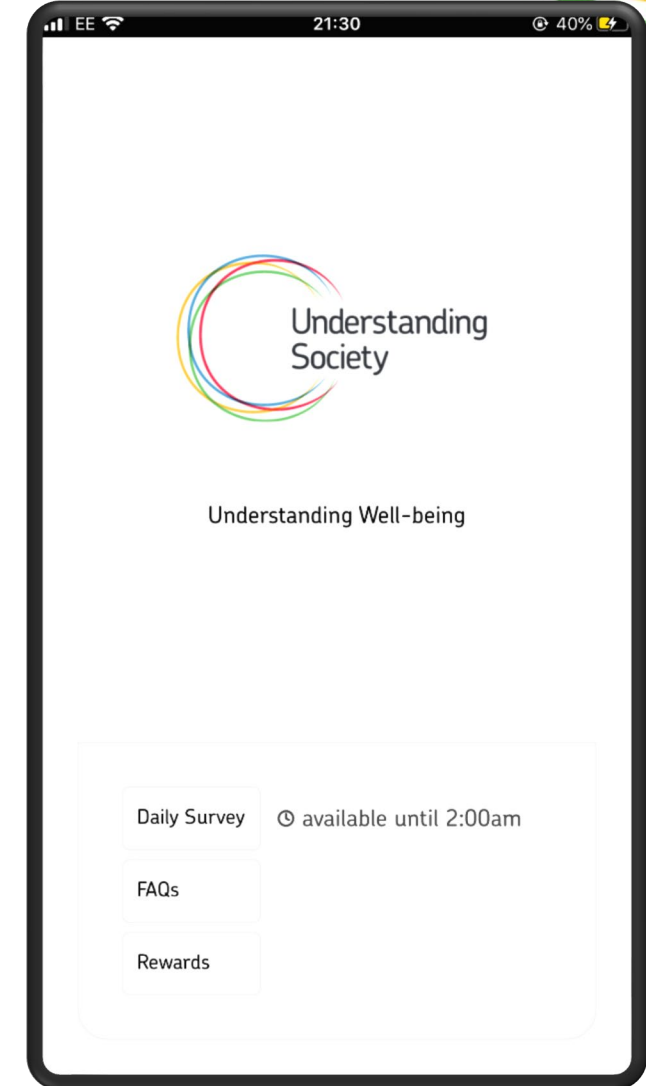
Understanding Society: App studies... 2

- 2016 – Spending Study 1
 - Ask people to upload photos of receipts for 1 month
- 2018 – Spending Study 2
 - Daily spending diary for 1 month
- 2020 – Well-being study
 - Daily (evening) app-based survey for 14 days
- 2022 – Body Volume app
 - Ask people to use an app to take two scans of their body
- 2023 – “Sea Hero Quest” app
 - Navigation game to measure spatial cognition



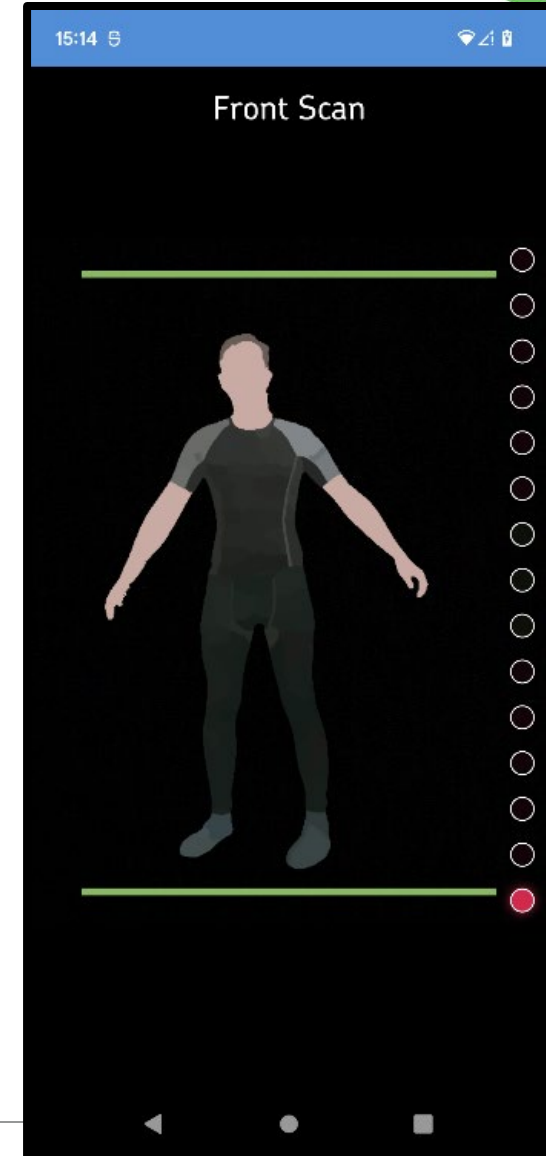
Understanding Society: App studies... 3

- 2016 – Spending Study 1
 - Ask people to upload photos of receipts for 1 month
- 2018 – Spending Study 2
 - Daily spending diary for 1 month
- 2020 – Well-being study
 - Daily (evening) app-based survey for 14 days
- 2022 – Body Volume app
 - Ask people to use an app to take two scans of their body
- 2023 – “Sea Hero Quest” app
 - Navigation game to measure spatial cognition



Understanding Society: App studies... 4

- 2016 – Spending Study 1
 - Ask people to upload photos of receipts for 1 month
- 2018 – Spending Study 2
 - Daily spending diary for 1 month
- 2020 – Well-being study
 - Daily (evening) app-based survey for 14 days
- 2022 – Body Volume app
 - Ask people to use an app to take two scans of their body
- 2023 – “Sea Hero Quest” app
 - Navigation game to measure spatial cognition

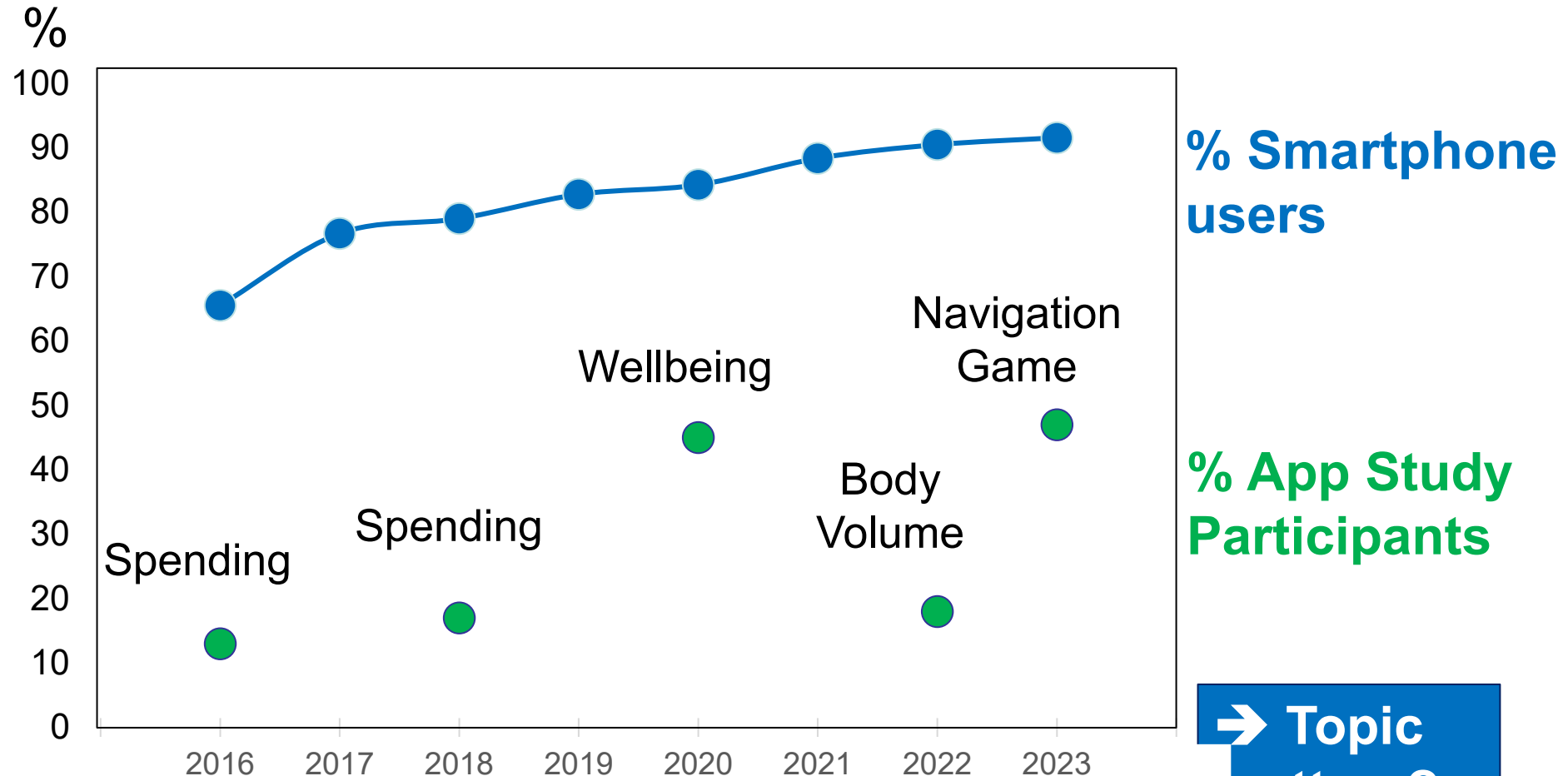


Understanding Society: App studies... - 5

- 2016 – Spending Study 1
 - Ask people to upload photos of receipts for 1 month
- 2018 – Spending Study 2
 - Daily spending diary for 1 month
- 2020 – Well-being study
 - Daily (evening) app-based survey for 14 days
- 2022 – Body Volume app
 - Ask people to use an app to take two scans of their body
- 2023 – “Sea Hero Quest” app
 - Navigation game to measure spatial cognition




Innovation Panel respondents



% Smartphone users

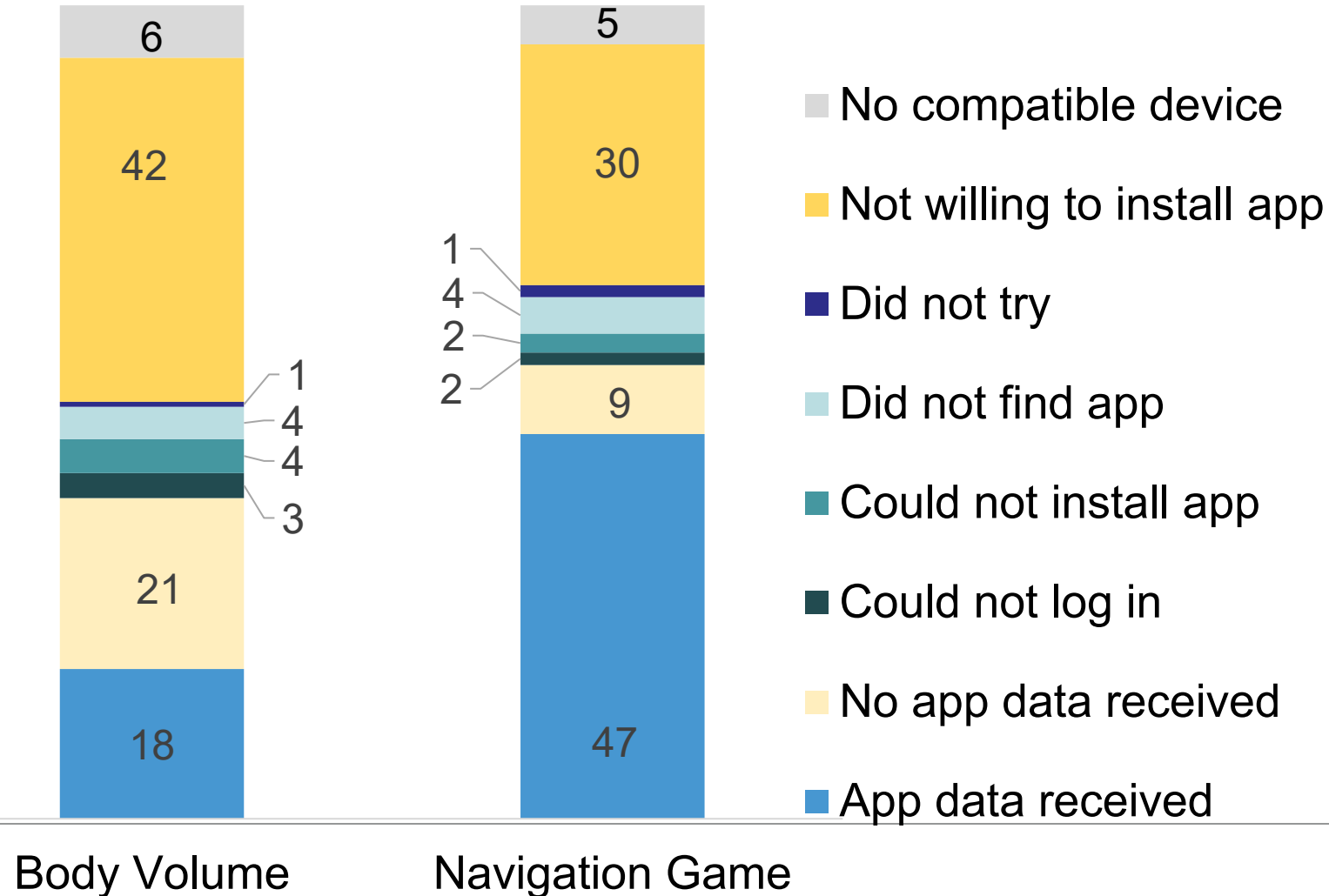
% App Study Participants

→ Topic matters?



Where do we lose
respondents?

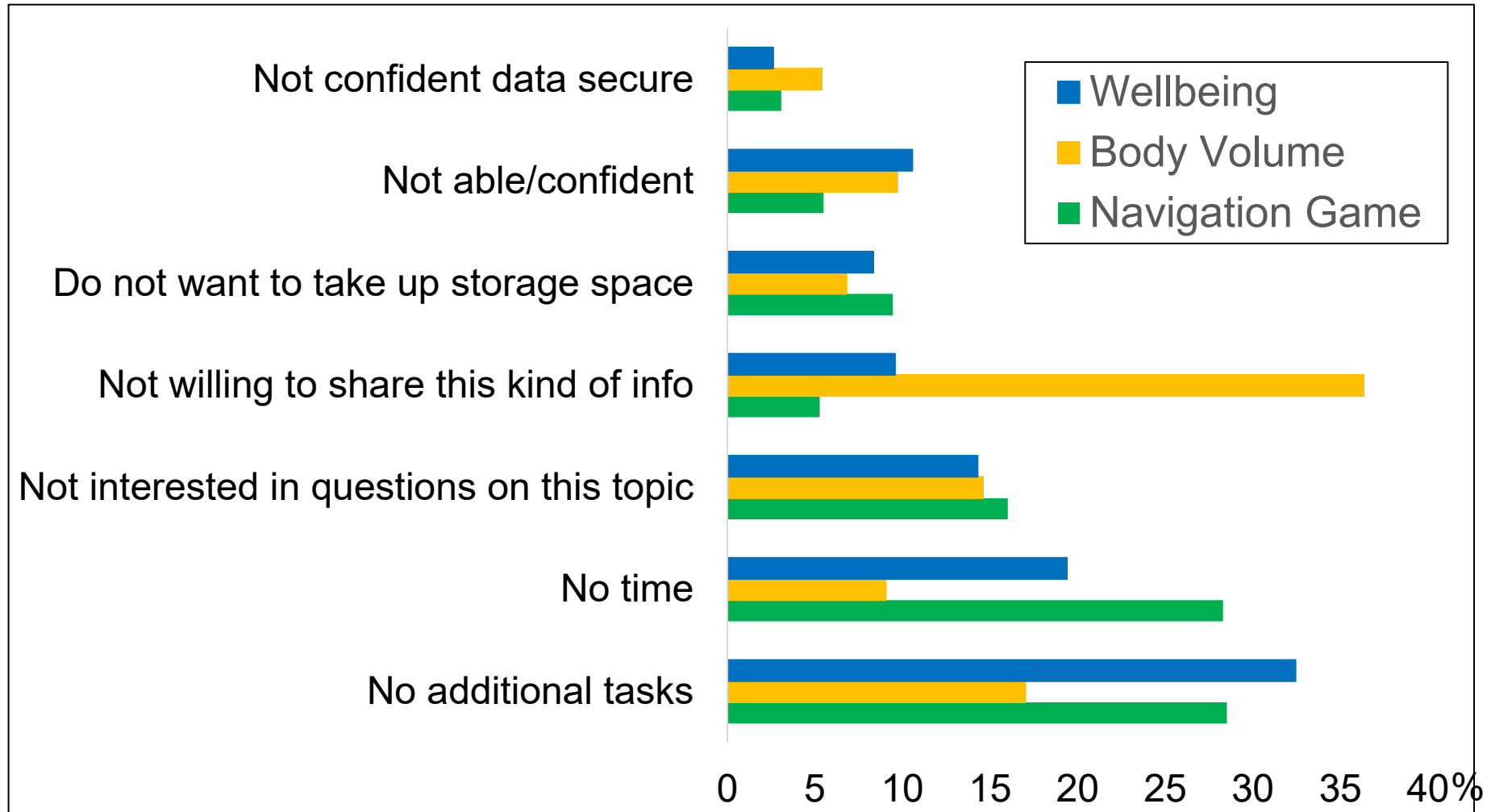
The main obstacle to app study participation is willingness, rather than technical reasons





Why do respondents not want
to install our apps?

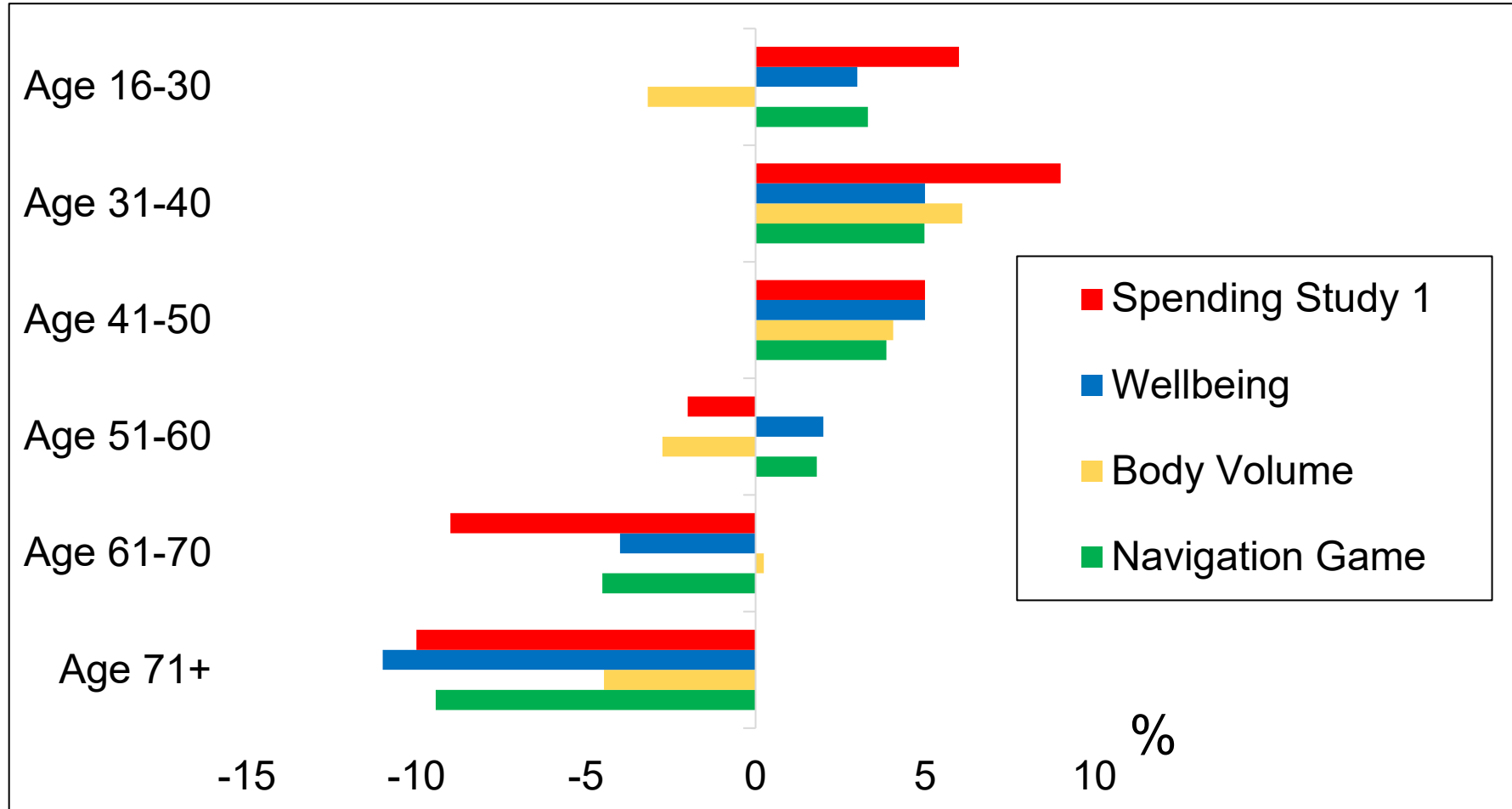
Why not willing to install app?





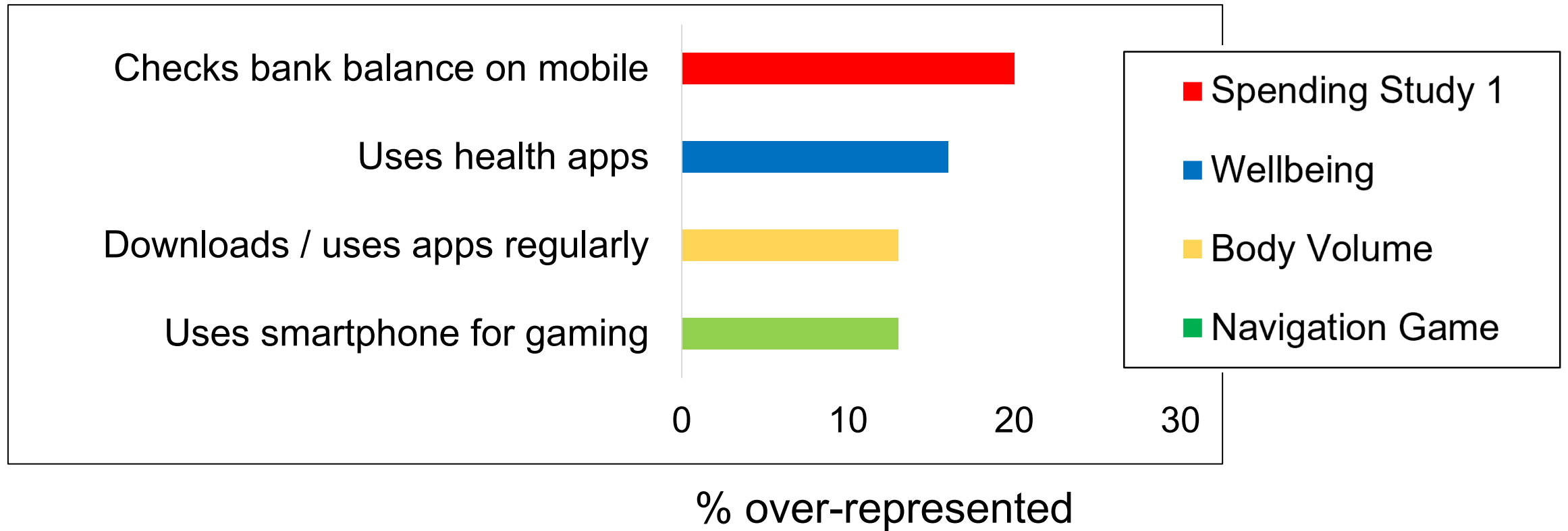
What types of respondents
do participate?

App participants are different to the full sample



Compared to the full sample, younger people are over-represented, older people under-represented

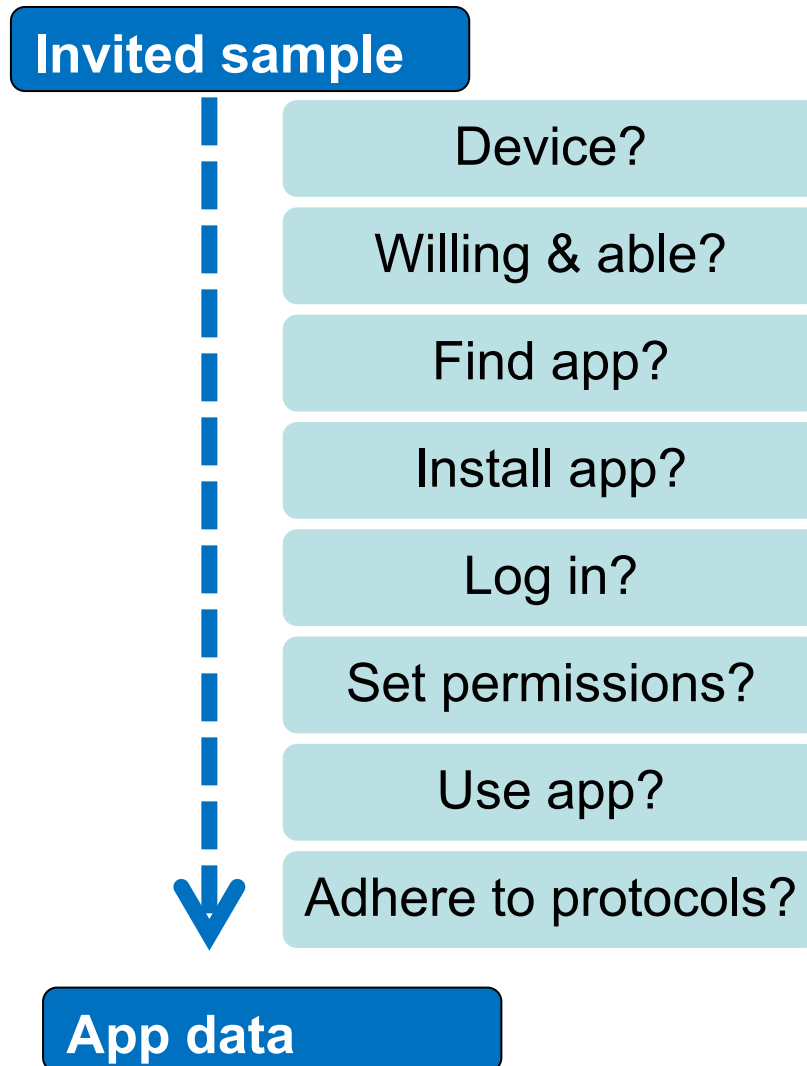
But task-related behaviour is more predictive.





So, what can we do?

Participation in data collection using apps



What are the barriers?

Compatible Device

No device, wrong OS, outdated OS



What are the barriers? - 2

Compatible Device

No device, wrong OS, outdated OS

Willing and able to participate

Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much



What are the barriers? - 3

Compatible Device

No device, wrong OS, outdated OS

Willing and able to participate

Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much

Find app

Mis-type app name, does not recognise app, link broken

What are the barriers? - 4



Compatible Device

No device, wrong OS, outdated OS

Willing and able to participate

Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much

Find app

Mis-type app name, does not recognise app, link broken

Install app

Lack of storage space, limited data plan

What are the barriers? - 5

Compatible Device

No device, wrong OS, outdated OS

Willing and able to participate

Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much

Find app

Mis-type app name, does not recognise app, link broken


Install app

Lack of storage space, limited data plan

Log in to app


Typo, log-in not working

What are the barriers? - 6




Compatible Device	No device, wrong OS, outdated OS
Willing and able to participate	Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much
Find app	Mis-type app name, does not recognise app, link broken
Install app	Lack of storage space, limited data plan
Log in to app	Typo, log-in not working
Set relevant permissions	Not willing, misunderstanding why required

What are the barriers? - 7



Compatible Device	No device, wrong OS, outdated OS
Willing and able to participate	Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much
Find app	Mis-type app name, does not recognise app, link broken
Install app	Lack of storage space, limited data plan
Log in to app	Typo, log-in not working
Set relevant permissions	Not willing, misunderstanding why required
Use app once	Not willing, delay before using, forgets

What are the barriers? - 8



Compatible Device	No device, wrong OS, outdated OS
Willing and able to participate	Doesn't want to, not interested, too busy, privacy concerns, sensitive information, data security concerns, survey asking too much
Find app	Mis-type app name, does not recognise app, link broken
Install app	Lack of storage space, limited data plan
Log in to app	Typo, log-in not working
Set relevant permissions	Not willing, misunderstanding why required
Use app once	Not willing, delay before using, forgets
- Adhere to study protocol	Not willing, forgets, weak internet signal

What can we do to reduce barriers?

Device

- Loan device
- Compatibility: OS / OS version / tablet & phone

- Loan devices?
 - Expensive
 - Require training materials
 - Technical helpline
- Check OS in questionnaire
 - Requires respondent to know it
 - Or instructions on how to find it
- Design for lower/older OS
 - Loss of functionality for app

What can we do to reduce barriers? - 2

Device

- Loan device
- Compatibility: OS / OS version / tablet & phone

Willing & able

How can we
increase
willingness?

How do respondents decide?

Topic
Interest

Value for
science?

Burden?

Invitation
wording?

Personal
benefit?

Intrusive?

Technical
requirements?

Risks?

Emotional
state?



How can we increase willingness?

Levers?

- Incentives
 - Feedback
 - Length of task
 - Question wording
 - Mode of invite
 - Placement within interview
-

Incentives may work if high enough or targeted

- Spending study 1
 - £2 vs £6 to install app **X**
- Well-being
 - £10 bonus for all 14 days vs £2.50 if complete on 4 random days vs no bonus **X**
- Body volume
 - **£5 conditional on using the app** vs extra £5 unconditional for the survey ✓ +7pp
- Sea Hero Quest spatial navigation
 - £10 vs **£30** for using the app ✓ +9pp

Incentives may work if high enough or targeted - 2

Other authors:

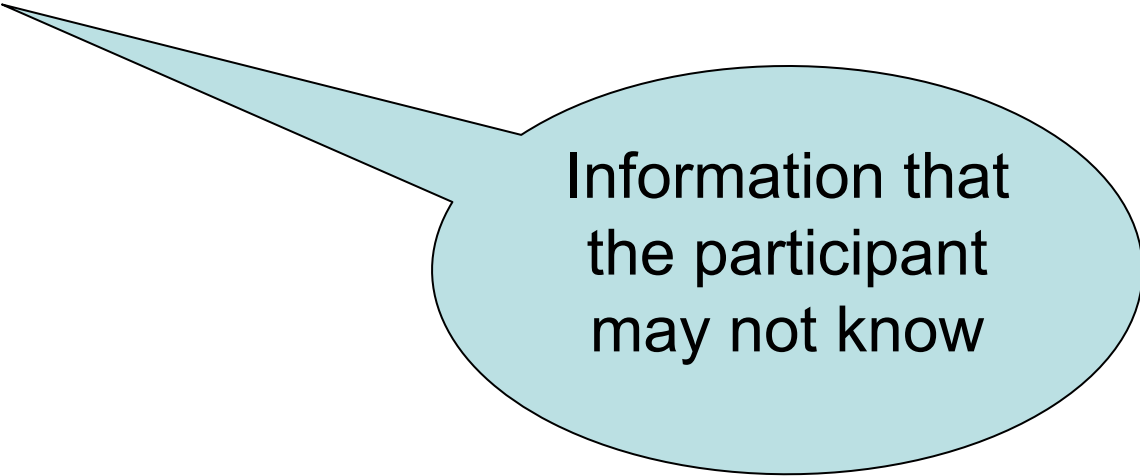
- Haas et al (2020)
 - €20 vs. €10 (IAB-SMART) +3pp
- Keusch et al (2021)
 - €30 vs. €0 (refugees) n.s.
- McCool et al (2021)
 - Bonus if app active 7 days (travel) +10pp
 - €5 vs. €10 vs. €20 +6pp

Non-financial incentives? Offering personalised feedback?

- Spending Study 2
 - Feedback on spending in app vs none? **X**
 - Body Volume
 - Body fat vs visceral fat vs none? **X**
 - Rodenburg et al (2023): feedback on Household Budget Survey app – no effect on participation or drop-out
-

But feedback may work in certain circumstances

- Feedback on blood analytes (IP12)
 - Offering blood test results increased take-up in CAPI and online



Information that the participant may not know

Ask people to spend less time?

- Well-being app
 - Daily questions on app, open in the evening
 - 2 minutes or 10 minutes of questions **X**

When and how we ask people may have an effect

- Spending Study 2
 - Invite in interview ✓
 - Invite by letter post-interview

Web: **X**
CAPI: +20 pp if invited during interview

When in the interview?

- Well-being app
 - Early ✓ +9pp
 - Late

See also Rodenburg et al (2023); household budget survey.
Interviewer +13pp vs letter

But mode of interview may not have an effect

- CAPI vs Web interviews (controlling for selection)
 - Body Volume **X**
 - Sea Hero Quest navigation **X**

What can we do to reduce barriers? - 3

Device	<ul style="list-style-type: none">• Loan device• Compatibility: OS / OS version / tablet & phone
Willing & able	<ul style="list-style-type: none">• Incentives• Feedback• Length of task (# days / daily time)• Wording: purpose / privacy / data security...• Mode of invitation: letter / in interview• Placement of invitation in interview
Find app	<ul style="list-style-type: none">• Search / link / QR code• Name and logo of app
Install app	<ul style="list-style-type: none">• App size
Log in	<ul style="list-style-type: none">• <u>User name</u> and/or password• Password usability
Set permissions	<ul style="list-style-type: none">• Wording: justification• Incentives
Use app	<ul style="list-style-type: none">• App design & instructions
Adhere to protocols	<ul style="list-style-type: none">• Reminders• Bonus incentives

Summary

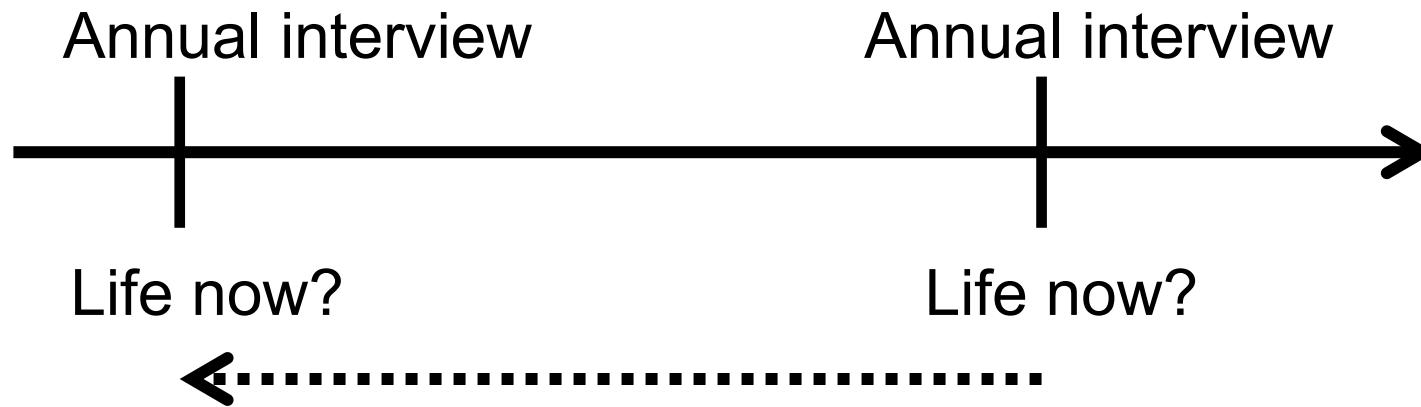


- Mobile device coverage now high
 - Few respondents have technical problems (finding, installing, logging in)
 - Willingness is main problem
 - Topic of app study matters
 - Whether respondents already do similar tasks matters
 - Recruitment in interview / with interviewer more successful
 - Time required not main barrier
 - Incentives / feedback do not compensate for other barriers
-

Additional (non-app) tasks

- Event-triggered data collection
 - Biological measures
-

Event-triggered data – what we do now



Recall questions – Annual Events:

- Health
- Pregnancy
- Partnerships
- Work
- Moves

Difficult to collect retrospectively:

- Subjective: expectations, plans, wellbeing
- Time-sensitive factual: e.g. during pregnancy

→ Develop methods to collect data when events occur

Event-triggered data collection

Developing protocols

- Qualitative research with Innovation Panel sample members
 - What willing to do?
 - Under what conditions?
- Experimental test of Life Event questions
 - *“Any of these events in last calendar month?”*
- Monthly **Life Events survey**
 - Every month for 12 months (2020)
- Respondent debrief questions

Life Events Study

Monthly filter question

Did you experience **any** of the following during <last month> 2020?

- Diagnosed with a new health condition or entered hospital/clinic as an in- or outpatient
- Had a pregnancy confirmed / partner had a pregnancy confirmed
- Changed jobs, started or stopped working
- Moved home
- Stopped or started living with a partner

Yes / No

IF No:

End of survey

IF Yes:

Follow-up questions about event(s)

Life Events Study

Key findings

- No effect of invitation to monthly study on annual interview
 - Attrition or reporting of events
- Monthly response rates stable
 - Increasing from 52% - 57%
- Reasons for non-participation


Next steps:

Questionnaire consultation
Phase 2 testing (IP) – NOW!
Roll out on main sample?

Outcome	%	Problem
No internet users in HH	10	→ Non-coverage
Email and mobile number unknown	15	→ Non-contact
Non-respondent in all waves	25	→ Non-response
Completed 1-10 waves	16	→ Wave non-response / attrition
Completed 11-12 waves	34	



Biological measures



Early in the study (Waves 2 & 3), nurses collected a range of measures

- Blood pressure
- Height
- Weight
- Waist measurement
- Body fat
- Grip strength
- Lung function

Including blood samples

- Measures of fat
 - Indicator of diabetes
 - Measures of inflammation
 - Measures of anaemia
 - Liver & kidney function
 - Hormones
- DNA
 - Genotyped data
 - Epigenetics
 - Proteomics
 - Polygenic scores



Our challenge now is how to collect
biomeasures without nurses

- Or the use of clinics
 - And limited use of interviewers (mostly web-first survey)
-

We experimented using the Innovation Panel (IP12) and the Wave 16 pilot

- IP12: Interviewer-led vs participant-led (with nurse visit as comparison)
 - Dried blood spots
 - Hair
- W16 pilot:
 - Microbiome
 - Blood sample

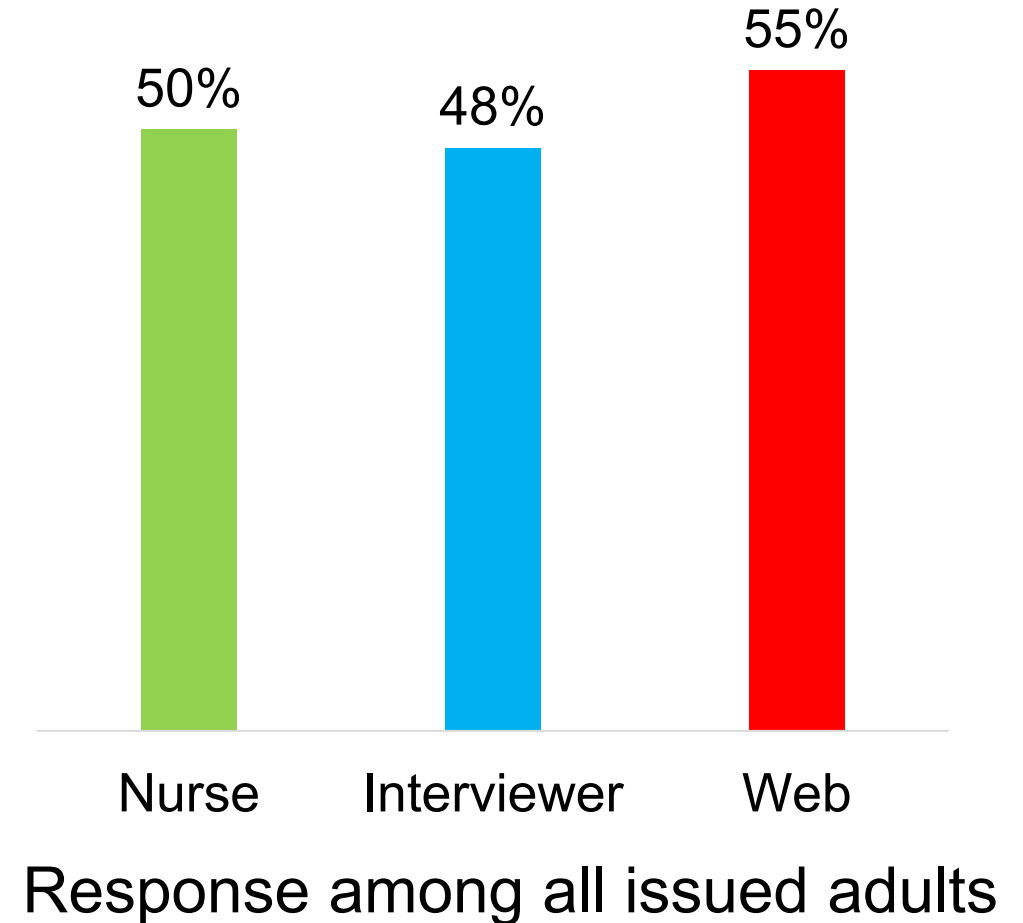


At IP12 (2019) we experimented with interviewer- and participant-led collection

- Hair sample and Dried Blood spots
 - (Taken by nurse)
 - Kit left by interviewer / sent to online consenters
-

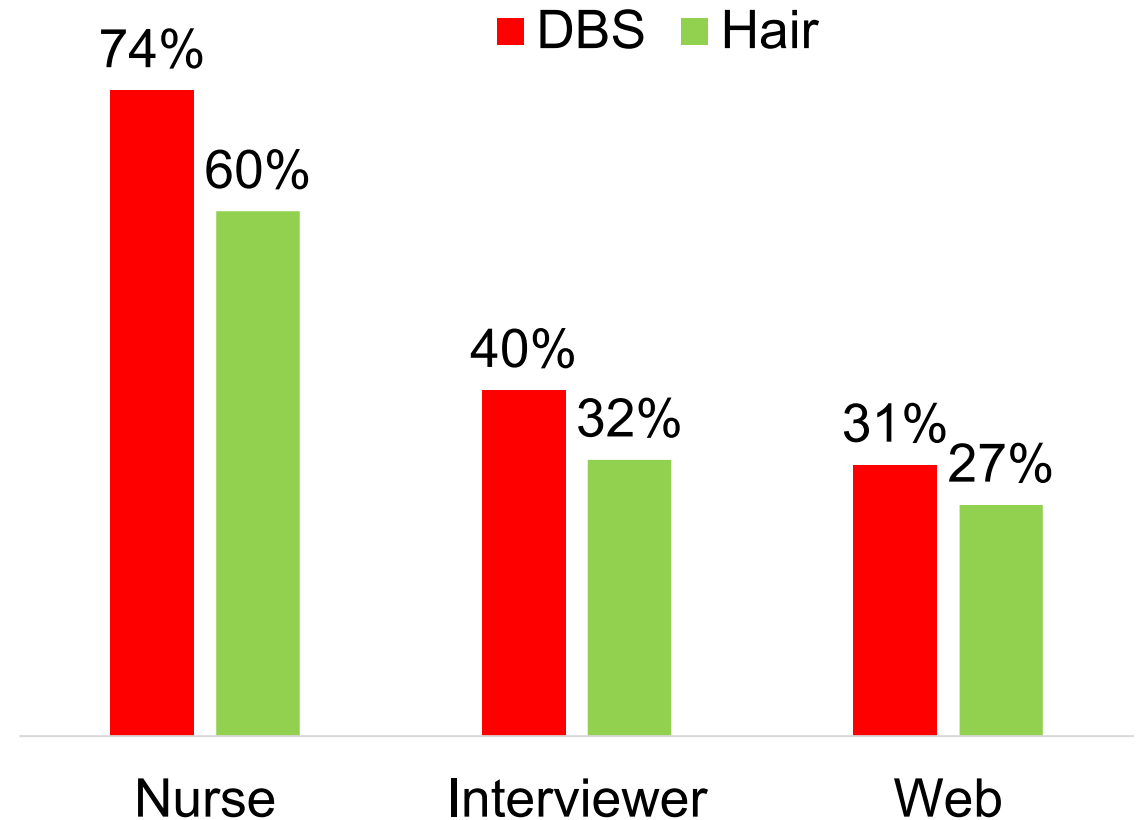
Higher response to the survey for those issued web-first

- Response to the survey
Highest in web-first
No difference between interviewer-first or nurse



But, as expected, higher return of samples from those with a nurse visit

- Return rate for samples (hair/DBS)
 - Highest for nurses
 - Interviewer-led higher than participant-led
- But offering **feedback** of results was most effective for web and interviewer modes

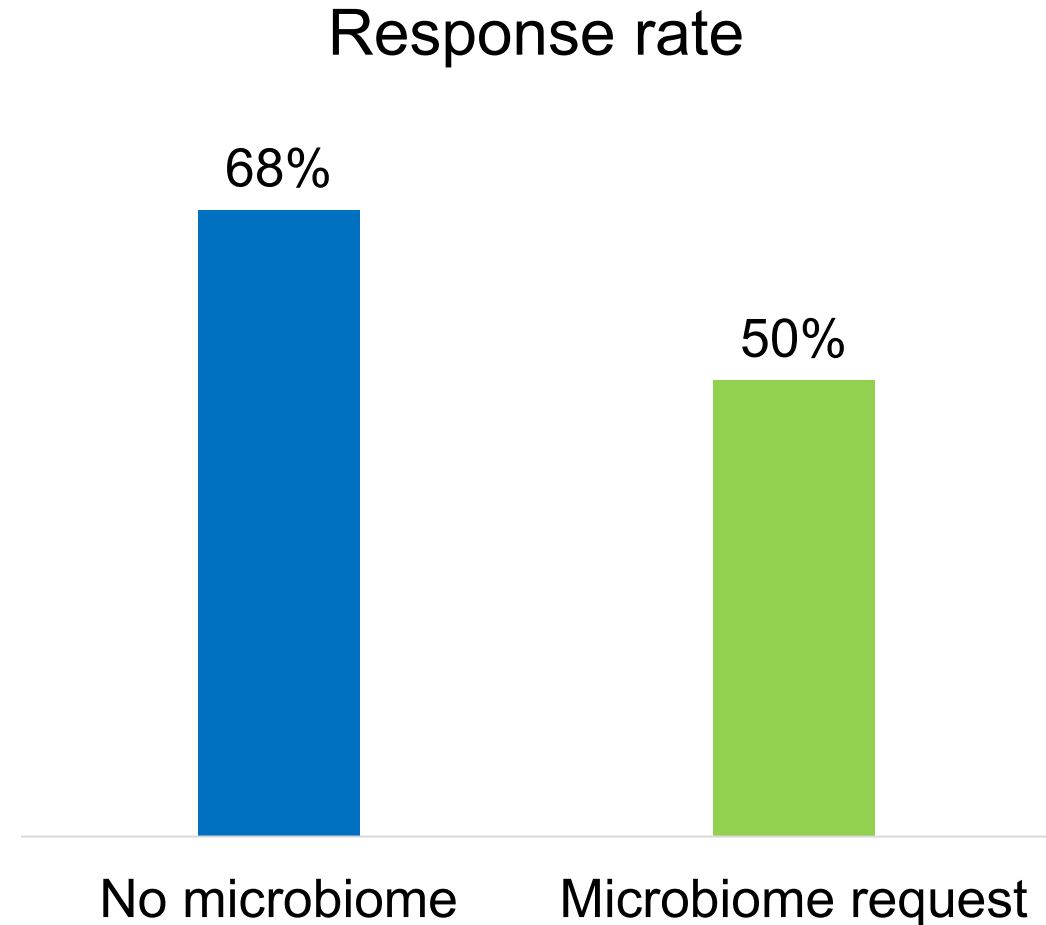


Microbiome – new for Understanding Society (Wave 16 pilot)

- Gut microbiome: implicated in depression, anxiety, dementia, obesity and metabolic syndrome.
- Evidence suggests that the environment plays a greater role in shaping gut microbiota than genetics
- Association of environmental and social factors with the gut microbiota requires further investigation
- Experiment in the Wave 16 pilot
 - **Half the sample** allocated to be asked for permission to send a collection kit

However, mentioning microbiome in the advance letter affected response to the survey

- Request to send a kit included in the questionnaire
- But information about this measure also included in advance materials
- Additional testing at IP18 (data due soon)



Wave 16 pilot: Those interviewed in-person more likely to agree to have a blood kit sent to them

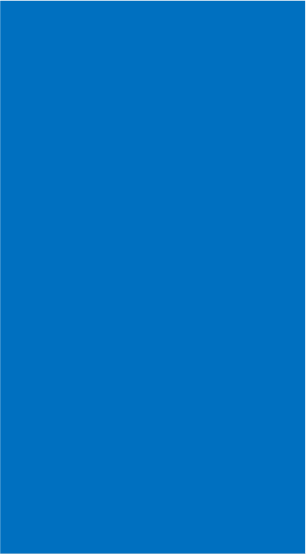


73%



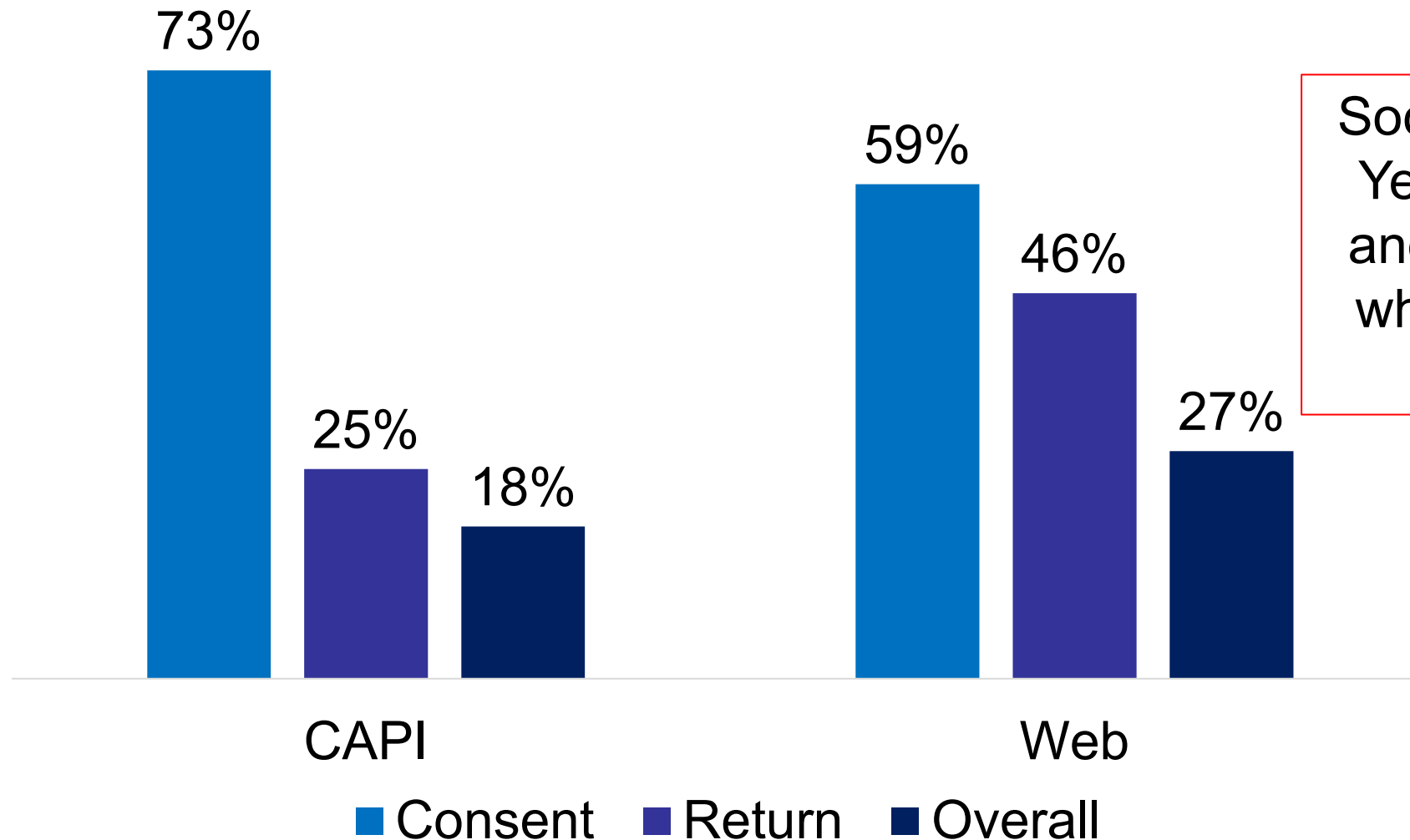
CAPI

59%



Web

But less likely to return the blood sample



Social desirability? Say Yes to the interviewer and then ignore the kit when sent? Delay the actual decision?

A couple more examples

- Less detail
- Areas that we're looking at now
- But other people in this (virtual) room know much more...





Data Donation

Asking for the data rather than collecting it



- Alternative to asking respondents to download an app to collect digital trace data
 - Ask respondent to ‘donate’ their data – often by uploading it to an online portal
 - Made easier with European GDPR (and similar) rules on data portability – the right for citizens to get a copy of their data
 - Digital platforms provide a Data Download Package (DDP)
 - Respondents can review this and upload all or part of the DDP; pre-processing to extract the required data
-

Asking for the data rather than collecting it - 2



- Examples with Instagram (van Driel et al 2022), Facebook (Kmetty and Németh 2022), smartphone health monitoring apps (Keusch et al 2025); but note differences between platforms (Wedel et al 2025).
 - Best Practice guide by Carrière et al (2025):
<https://doi.org/10.1007/s11135-024-01983-x>
-

Often a gap between willingness and ability to donate data

- E.g., Keusch et al (2025), LISS panel, over 50s with device: 29% willing, 12% successful
- Hase & Haim (2024), two studies, 63% & 52% willingness, 20% & 12% donated
- Requires respondents who want to give this detailed data
- *And* are technically comfortable in navigating the data donation process
 - But this is getting easier
- (to be tested in IP19 later this year)



Life History Calendars

Presenting a life history graphically



- Data collection tools used to reconstruct key events (e.g., residences, education, partnership, childbirth, employment)
 - Aim to improve quality and completeness of retrospective reporting
 - Use key events as an ‘anchor’ when thinking about other events
 - Strength is reducing recall bias and improve consistency (vs questionnaire-based) (Glasner and van der Vaart, 2009; Smith et al., 2021a; Morselli and Berchtold, 2023)
-

Limited research on impact of asking as additional task

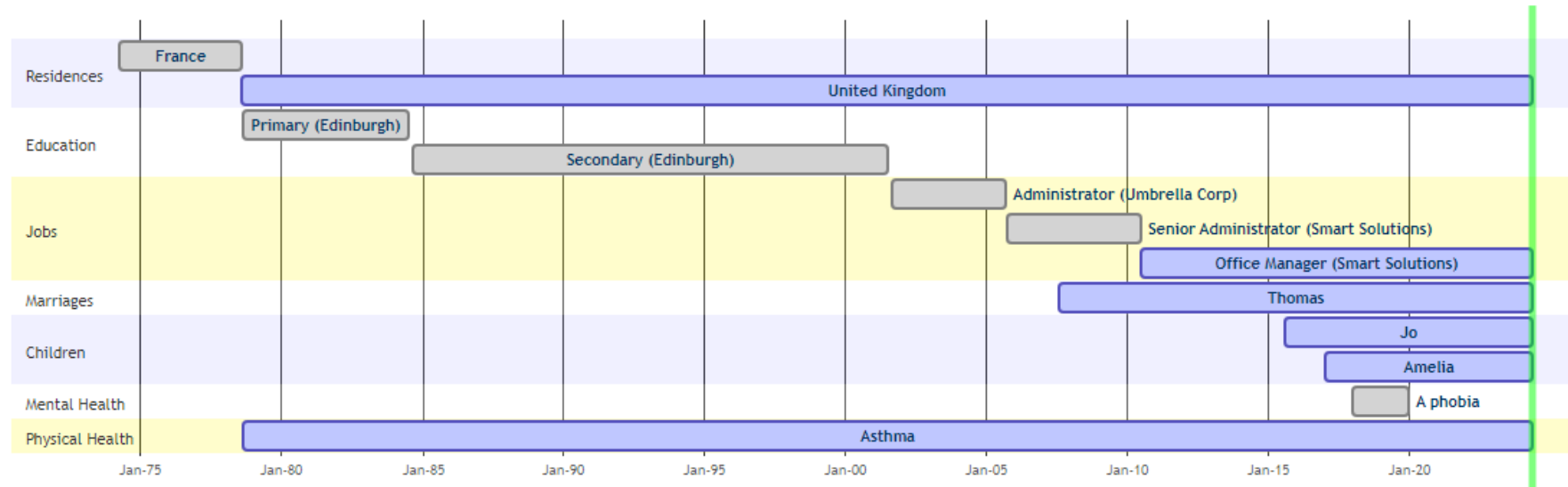
- But limited research on the effect on response/attrition
 - Little or no impact: Kominski, 1990; Belli et al, 2001
 - Negative impact (where calendar is foregrounded): Van der Vaart and Glasner, 2005; Yoshihama et al., 2005.
- And limited research on online calendars
 - Mostly referring to data quality/break-offs rather than response (e.g., Calderwood, 2012; West et al, 2022; Yu et al, 2026)

Testing on Understanding Society



- Internal piloting/testing
 - Using formr – open-source survey framework built in R
 - Iterative and extensive testing
 - Piloting – just under 500 pilot sample members invited to take part (outside of annual interviews)
 - Pilot sample not used often
 - 39% started; 37% completed at least one section; 33% completed all
-

Test version: Participants able to download their version of the calendar



Thank you for all of the information you have provided. Once you click "Submit" below, **no further changes** can be made to the timeline. Please check the timeline above and please go back if you need to correct anything.

When you are sure that everything is correct, click "Submit". The timeline will be submitted and complete. Thank you once again.

Submit

Back to

[Video](#)

[Instructions](#)

[Residences](#)

[Education](#)

[Jobs](#)

[Marriages](#)

[Children](#)

[Mental Health](#)

[Physical Health](#)

[Complete Calendar](#)

Testing on Understanding Society - 2

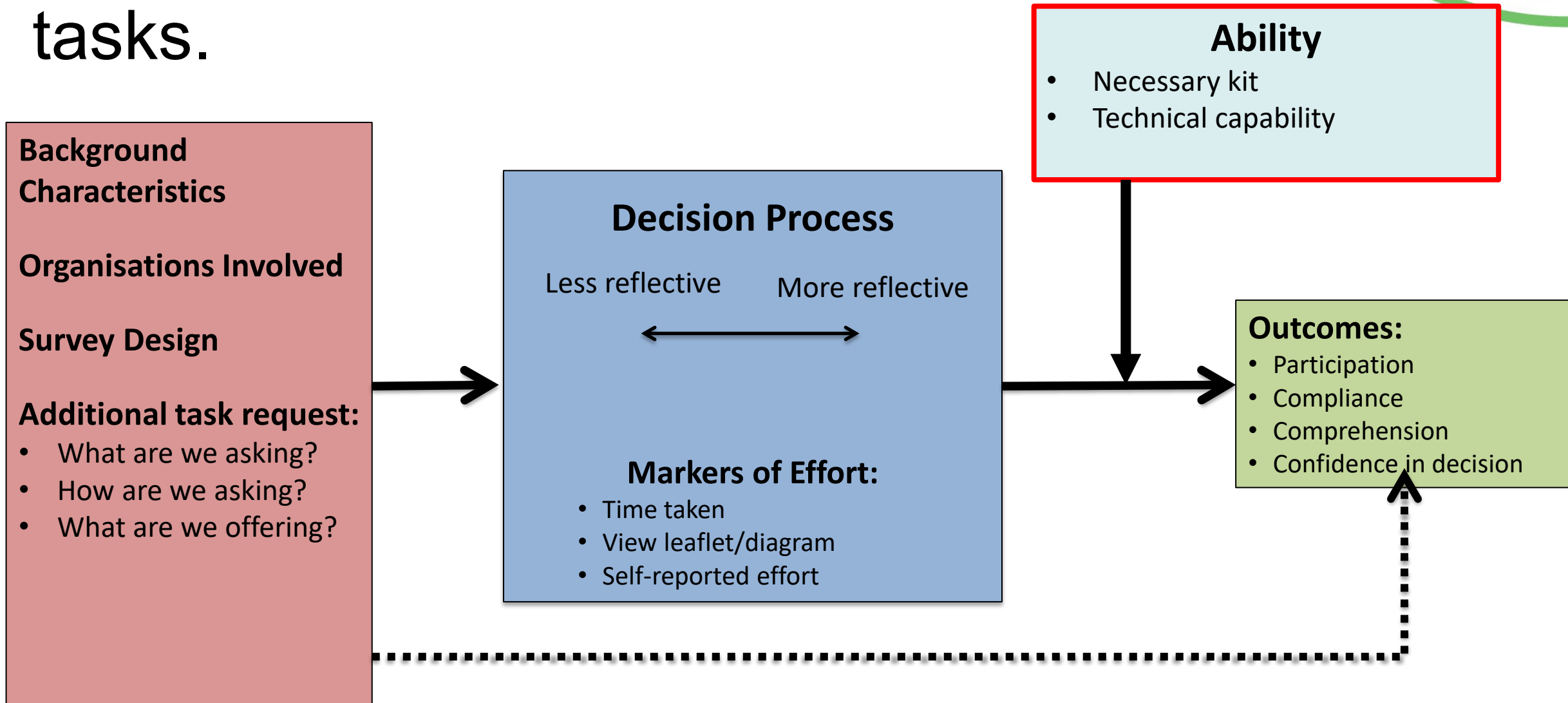


- Further usability testing Spring/Summer 2026
 - Testing on IP sample (after IP19)
 - <https://www.understandingsociety.ac.uk/participants/projects/life-history-calendar-faqs/>
-



How can we persuade people to do additional tasks?

Conceptual Framework for additional tasks.



Participation in additional tasks depends on...

- Design of baseline survey
 - Survey mode (interviewer / self-completion)
 - Respondent incentives (for baseline survey / for additional task)
 - How the task is presented (affecting willingness)
- Nature of task
 - Domain/topic of data (health, wellbeing > finance)
 - Duration of task (number of days / time per day)
 - Technical ease – lack of friction – to carry out the task

Participation in additional tasks depends on... 2

- Protocols for implementation of task
 - Incentives (bonuses for compliance)
 - Feedback (if information not known to respondents)
 - Mode of invitation to task (in interview > by post)
 - Placement of invitation to task in baseline survey (early > late)
 - Reminders
 - Reviewing the points at which people may drop out and reducing the 'friction'
-

On a final note – can we ask too much?

- In a recent working paper Mitchell and Jäckle (2026) use the multiple requests in the IP
 - Sample members participated in just under half of tasks they were invited to
 - Controlling for selection bias, each additional task:
 - Decreases the probability of participating in a task (by 4pp)
 - Decreases the probability of taking part at the next wave (by 0.5pp)
- We can ask people to do more, but we need to weigh up the benefits to the study with the potential risk
- This is one of the values of the Innovation Panel – to test the effects of these additional task and work out how best to implement them.

Thank you



Colleagues

- Annette Jäckle, Mick Couper, Tom Crossley, Sandra Walzenbach, Jim Vine, Michaela Benzeval, Brienna Perelli-Harris, Brendan Read, Paulo Serodio, Alexander Wenz, Daniel Horn, plus advice and help from many other colleagues

Fieldwork organisations

- Verian, NatCen Social Research, Ipsos MORI

App developers

- Kantar Worldpanel, Qmob, Connect Internet Solutions, Select Research, Glitchers

Funder

- Economic and Social Research Council UK

Data

- University of Essex, Institute for Social and Economic Research (2025). Understanding Society: Innovation Panel, Waves 1-17, 2008-2024. [data collection]. 14th Edition. UK Data Service. SN: 6849, [DOI: http://doi.org/10.5255/UKDA-SN-6849-17](https://doi.org/10.5255/UKDA-SN-6849-17)
-

What about you?

- Consent to data linkage
 - Are you required to ask for consent?
 - What other legal/ethical issues do you encounter?
 - Additional tasks
 - What additional tasks have you asked respondents to do?
 - Have you been able to test these experimentally?
 - What have you found?
 - What is next....?
-