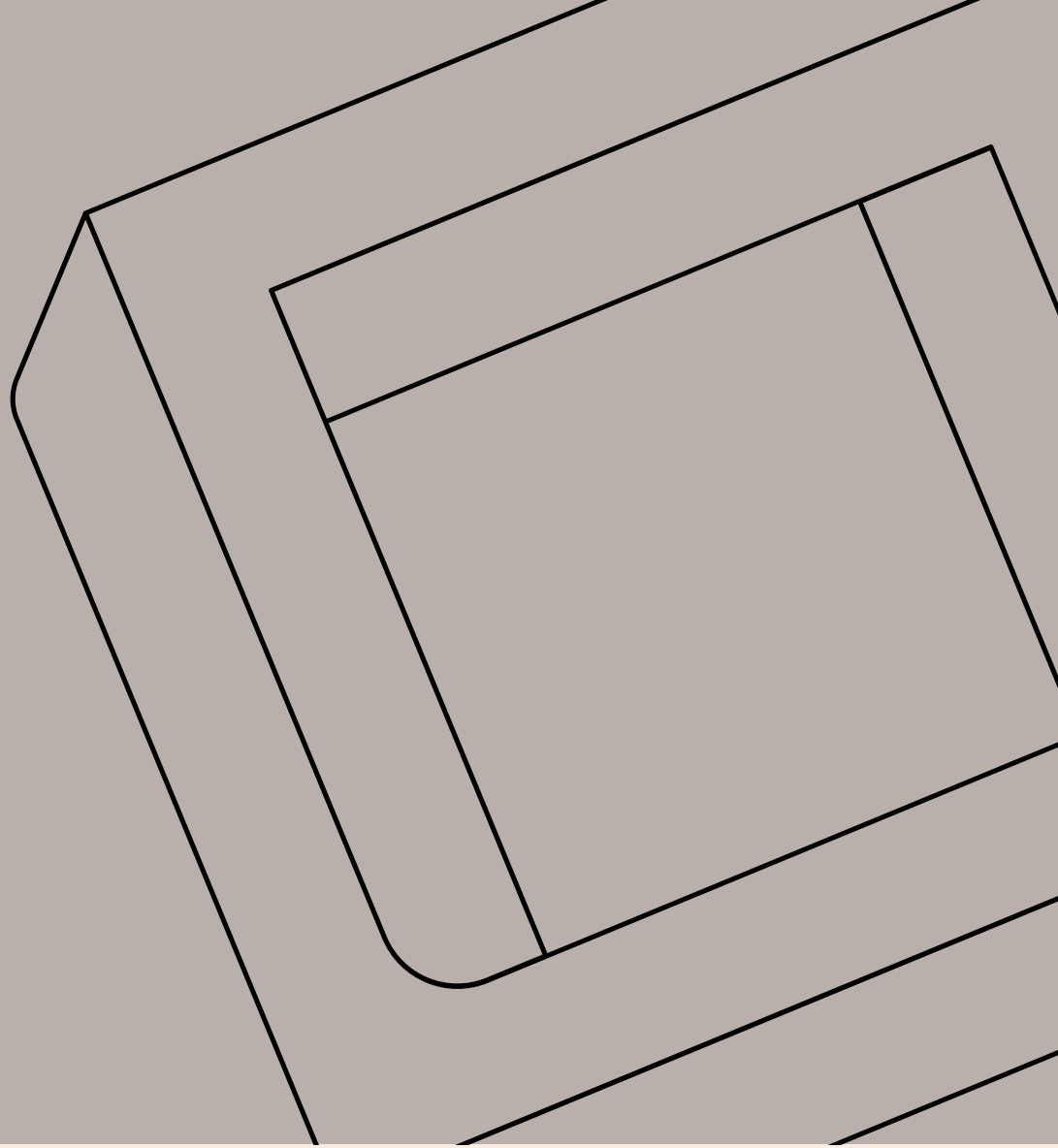
Differences in response characteristics between telephone- and web-survey modes in a large-scale, multi-country public attitudes survey

Tanja Kimova, Verian, presenter, author
James Hagan, Verian, author
Alexandrina Buruian, Verian, co-author
Patrick Moynihan, Pew Research Centre, co-author

18 March 2024

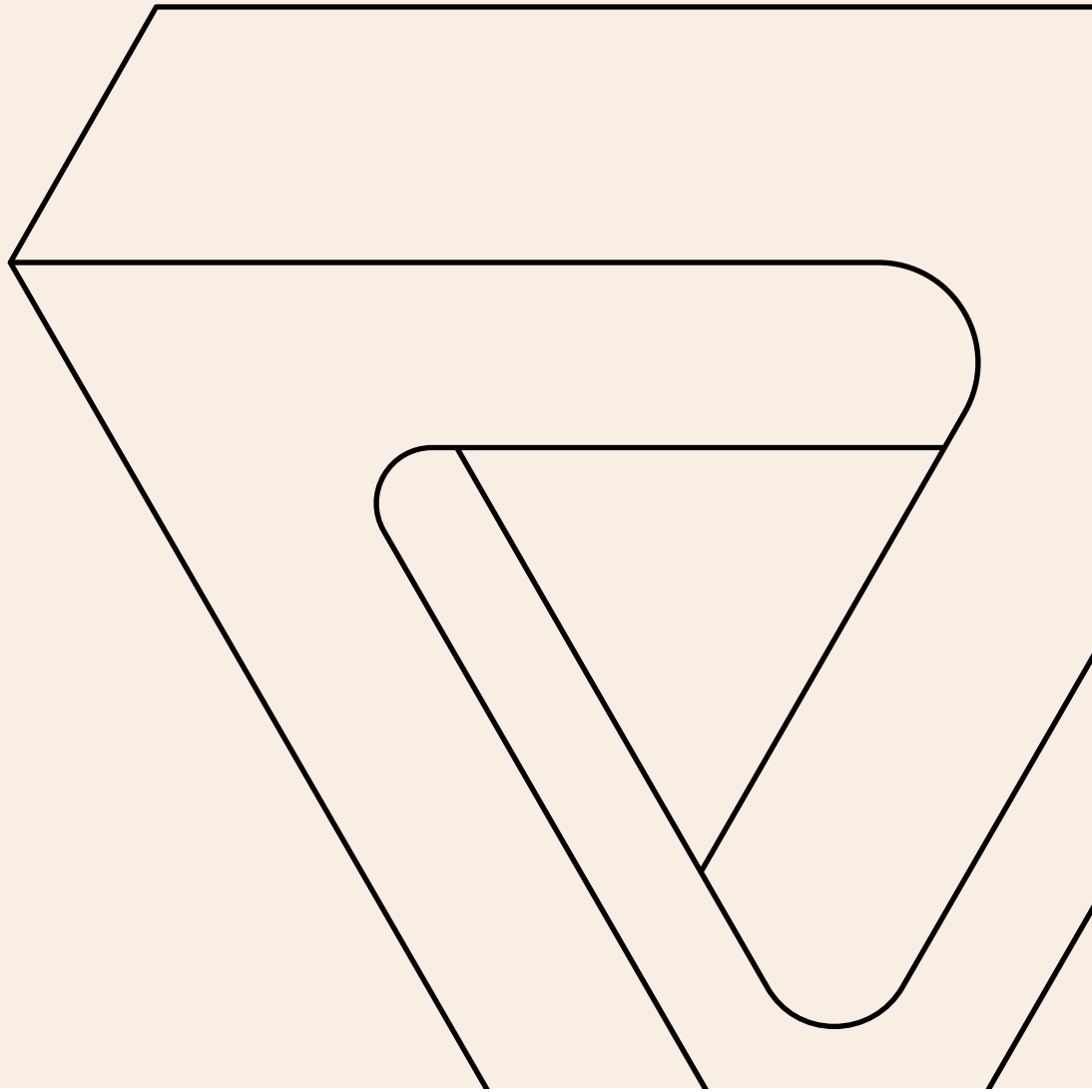


verian

Agenda

1. Motivation and research
2. Method and analysis plan
3. Results
4. Conclusions (and next steps)

Motivation and research design



verian

Motivations for the study

Mixed-mode research: the new norm

Chiefly in response to the challenges probability-based unimode approaches face (tighter project budgets, falling response rates and the higher costs required to maintain similar quality levels)

The literature gap

1. There is a lack of cross-national mixed-mode research aiming to understand differences in measurement quality between modes
2. Studies use less recent data sources, while respondents adapt to new technologies

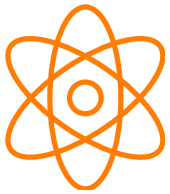
Web interviewing is on the rise: push & pull factors being at play

1. Almost always one of the data collection modes in current mixed-mode designs.
2. Compelling reasons to expect it will become the dominant survey mode in the social research over the next decade.

What is the impact of a mixed-mode design on the comparability of time-series data?

There is a need to evaluate and quantify the impact of mode (and specifically of measurement) differences on data comparability by disentangling sample and measurement biases on survey estimates.

Study design considerations

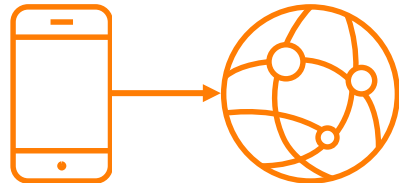


We compared two **high-quality designs** in 3 countries (the UK, France and Germany) in Q2 2023:

- Pew's Global Attitudes Project (GAP) run as a telephone survey (*classical approach*)
- Pew's GAP run via the PUBLIC Voice online probability-based panel (*exploratory survey*)

We adjusted the telephone questionnaire to make it fit for purpose for web interviewing and reduce the effect of mode on measurement (relied on literature review to inform design choices and our quality appraisal framework).

We take the assumption both samples are comparable after correcting for bias through the (elaborate) weighting procedures.



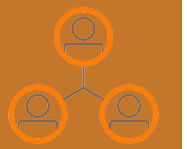
Hence, the main hypothesis is that the observed differences between the telephone and online survey estimates relate to the effects of mode on measurement.

Considerations

Understanding how respondent profiles differ between telephone-administered surveys and online surveys provides information about the **representativeness and generalisability of data collected through different modes.**

Despite the recognized need for flexibility in survey design**, the implications of sample and measurement bias resulting from changes in methodology remain poorly understood.**

Unclear influence of respondents' profile on response characteristics across modes, resulting in an incomplete understanding of the factors contributing to variability in response accuracy and robustness.



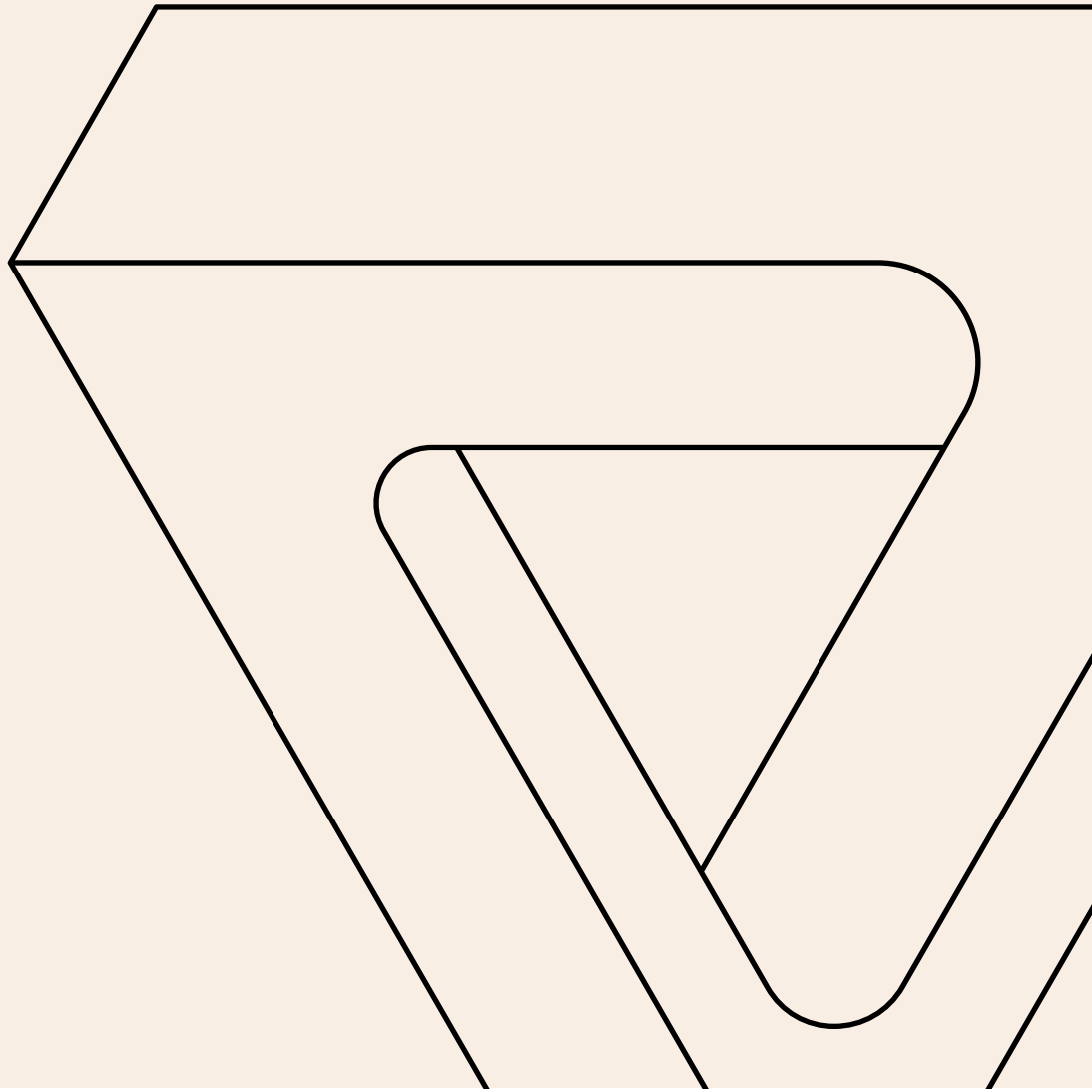
Research questions

0 Are there **differences in the demographic profiles** of the respondents who took part in the telephone-administered and online surveys (after weighting)?

1 Do we observe **differences in terms of response characteristics across telephone and web** for the same questions?

2 **Can we explain these differences by looking at individual-level differences** (gender, age, education, country)?

Method and analysis plan

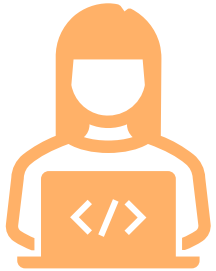


verian

Variable selection

Extracted a set of 50 comparable questions from the two different survey-mode datasets

PEW GAP telephone interviews



PEW GAP Public Voice web interviews

1 Verbatim answer question

41 Questions grouped into 9 batteries

8 Single-choice questions

Age

Sex

Education (ISCED)

Country

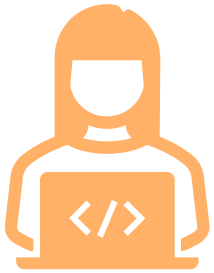
Examples of comparable questions

1 Verbatim answer question

PEW GAP telephone interviews

Question

We'd like to know more about your views of democracy in **[COUNTRY]**. What do you think would help improve the way democracy in **[COUNTRY]** is working?



PEW GAP Public Voice web interviews

Question

We'd like to know more about your views of democracy in **[COUNTRY]**. What do you think would help improve the way democracy in **[COUNTRY]** is working?

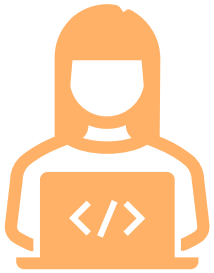
Examples of comparable questions

1 Verbatim answer question

PEW GAP telephone interviews

Question

We'd like to know more about your views of democracy in **[COUNTRY]**. What do you think would help improve the way democracy in **[COUNTRY]** is working?



PEW GAP Public Voice web interviews

Question

We'd like to know more about your views of democracy in **[COUNTRY]**. What do you think would help improve the way democracy in **[COUNTRY]** is working?

Please answer as thoughtfully as you can.

Examples of comparable questions

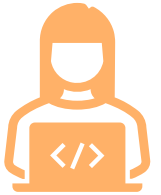
41 Questions grouped into 9 batteries

Question

I'm going to describe various types of political systems and ask what you think about each as a way of governing our country. For each one, would it be a very good, somewhat good, somewhat bad or very bad way of governing this country?

Question

Would each of the following political systems be a good or bad way of governing this country?



Examples of comparable questions

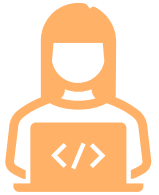
41 Questions grouped into 9 batteries

Question

I'm going to describe various types of political systems and ask what you think about each as a way of governing our country. For each one, would it be a very good, somewhat good, somewhat bad or very bad way of governing this country?

Question

Would each of the following political systems be a good or bad way of governing this country?



Very good ○ Somewhat good ○ Somewhat bad ○ Very bad ○

Examples of comparable questions

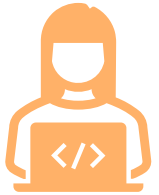
8 Single-choice questions

Question

Thinking about our economic situation, how would you describe the current economic situation in **[COUNTRY]** – is it very good, somewhat good, somewhat bad or very bad?

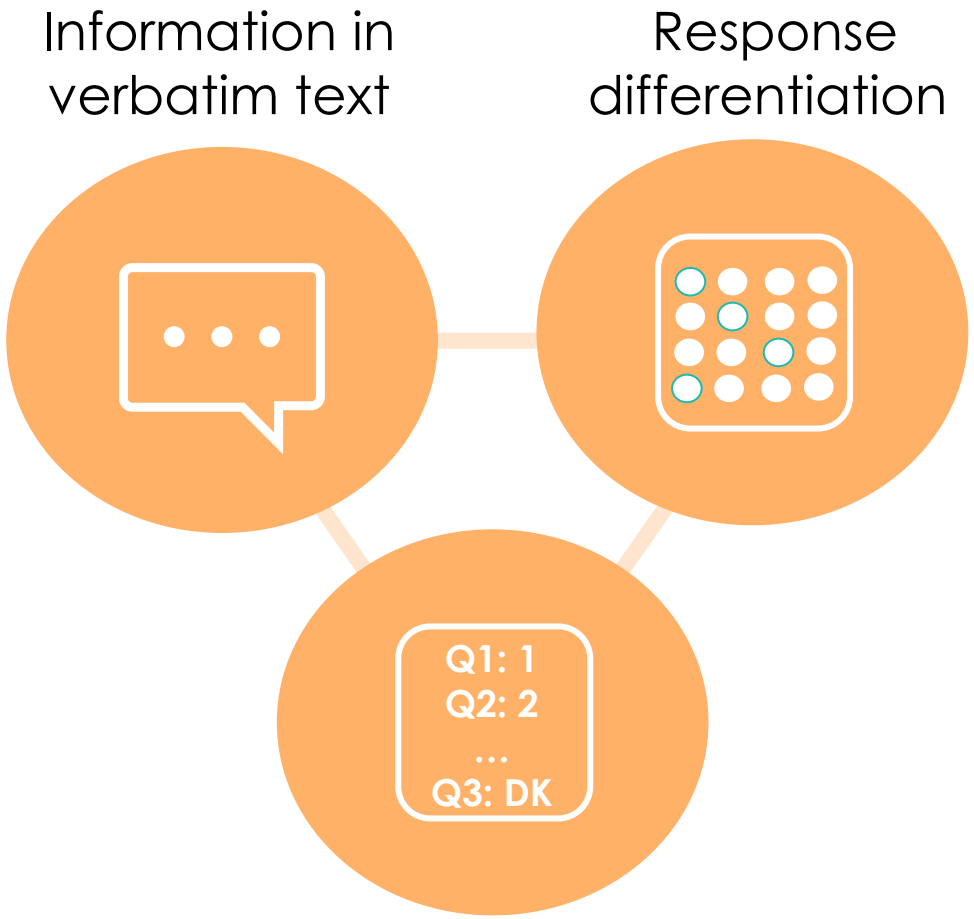
Question

Thinking about our economic situation, how would you describe the current economic situation in **[COUNTRY]**?

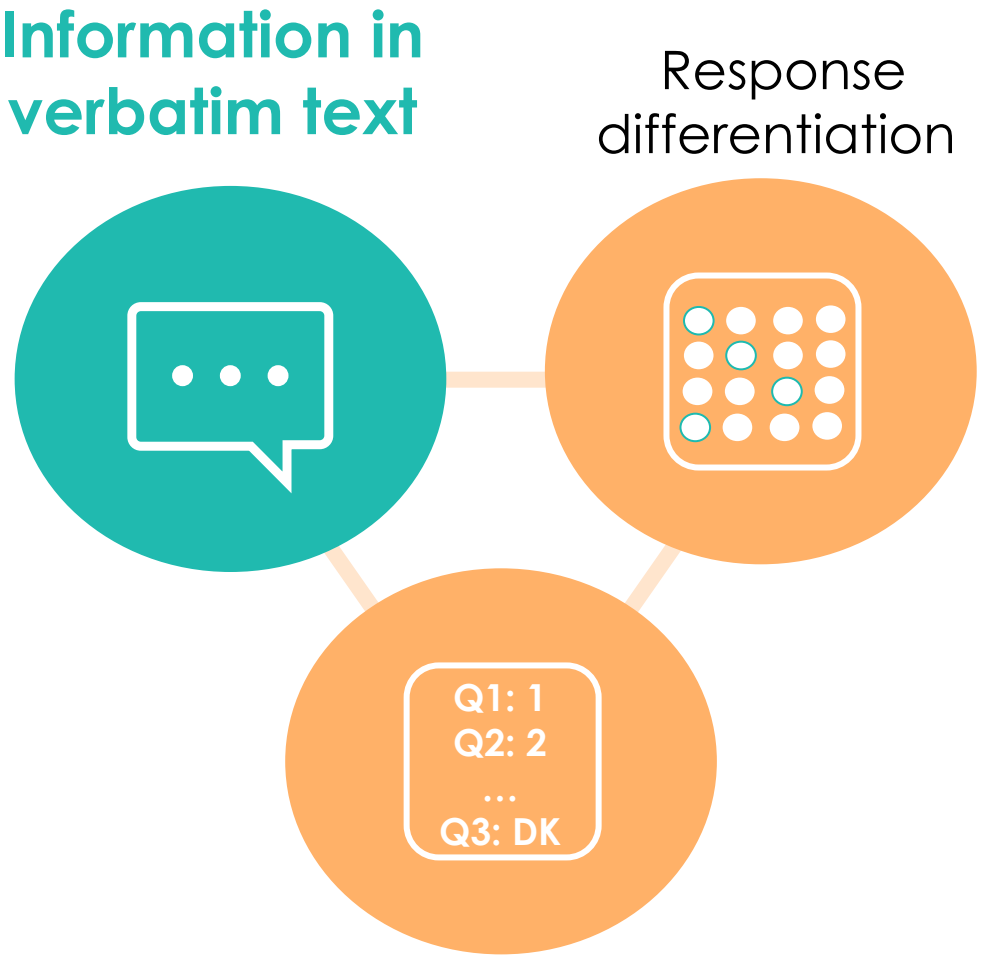


Very good	Somewhat good	Somewhat bad	Very bad
○	○	○	○

Measuring response characteristics



Measuring response characteristics



Response characteristics

Information in verbatim text

We'd like to know more about your views of democracy in **[COUNTRY]**. What do you think would help improve the way democracy in **[COUNTRY]** is working?

Answer

"Leaders just need to listen to what we want instead of them always second guessing and not just thinking about their wallet and how they choose affects the minority of people"

Answer

"Voting for trustful candidates and freedom of speech and thoughts"

Information in verbatim text

Used Xiao et al.'s (2020) **informativeness (I)** index to estimate the amount of information in a piece of text.

"Leaders just need to listen to what we want instead of them always second guessing and not just thinking about their wallet and how they choose affects the minority of people"

British National Corpus

word	$F(word_n)$
leaders	0.01
want	0.05
of	0.1
guessing	0.005
I	0.2
...	
minority	0.0025

$$= \sum \log_2 \frac{1}{F(word_n)}$$

I = 312 bits

Xiao et al.'s (2020): Informativeness

Information in verbatim text

Used Xiao et al.'s (2020) **informativeness (I)** index to estimate the amount of information in a piece of text.

"Voting for trustful candidates and freedom of speech and thoughts"

British National Corpus

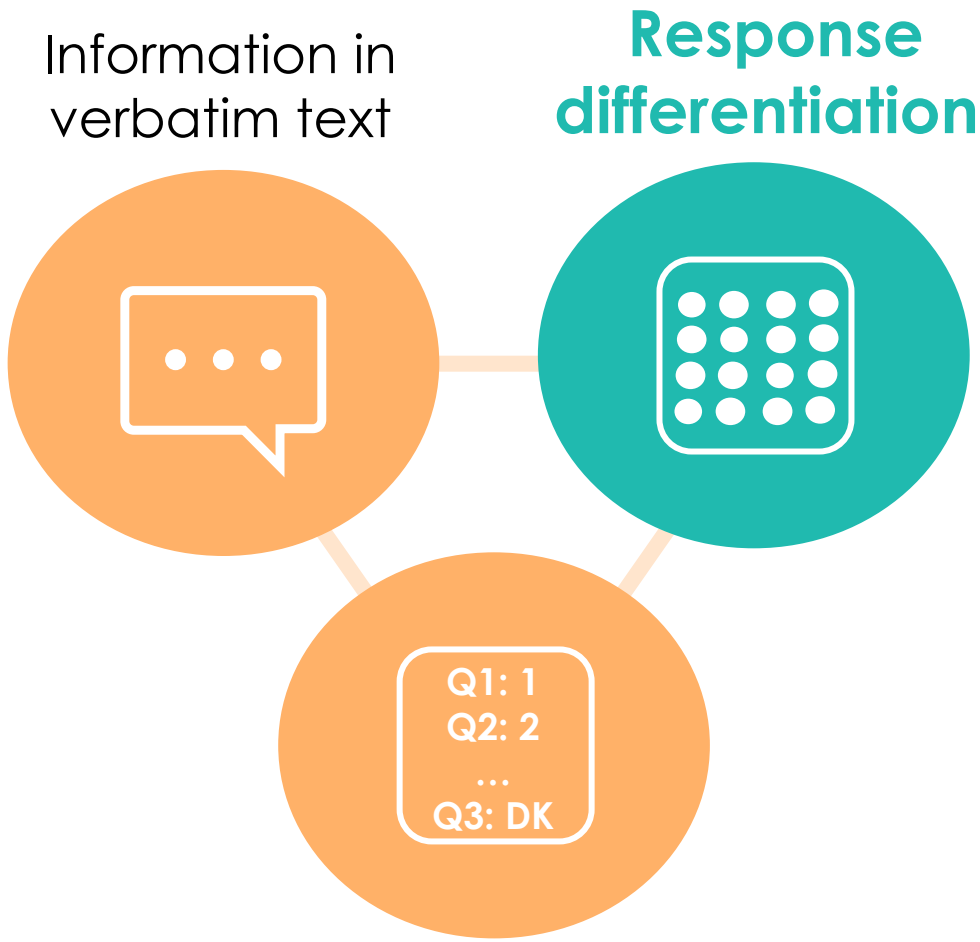
word	$F(word_n)$
voting	0.01
for	0.05
of	0.1
thoughts	0.005
and	0.2
...	
speech	0.0025

$$= \sum \log_2 \frac{1}{F(word_n)}$$

I = 119 bits

Xiao et al.'s (2020): Informativeness

Measuring response characteristics



Response characteristics

Response differentiation

Question

Do you think policies in [COUNTRY] would improve **(1)**, would get worse **(2)** or would mostly stay the same **(3)** if more elected officials were…

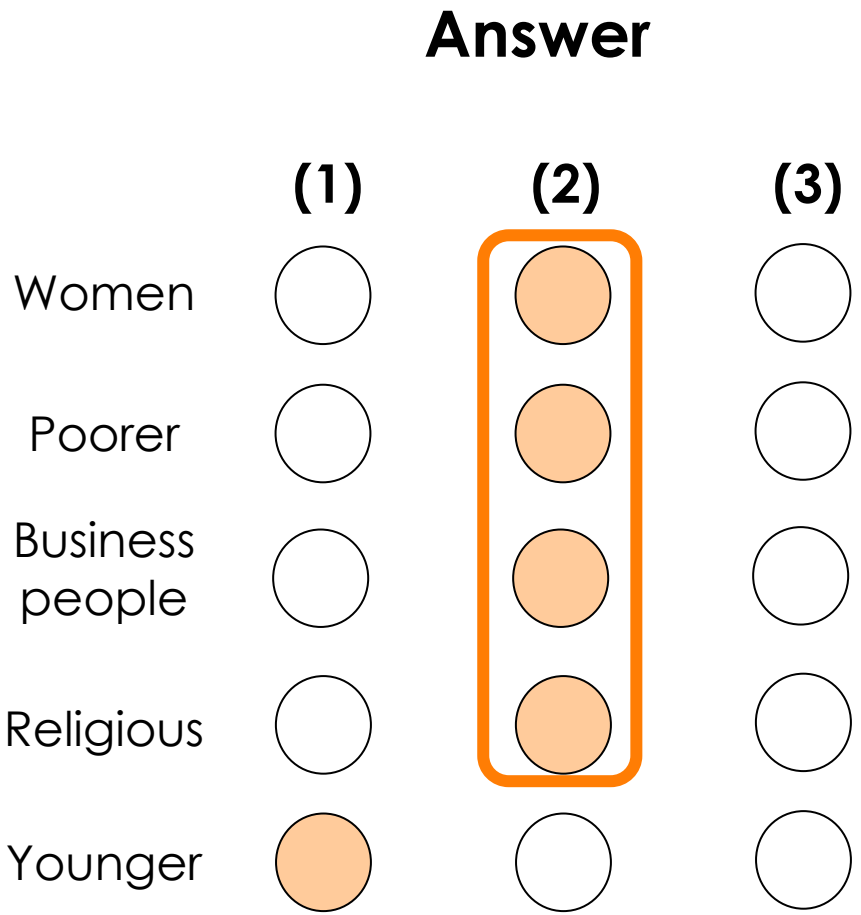
Answer

	(1)	(2)	(3)
Women	○	○	○
Poorer	○	○	○
Business people	○	○	○
Religious	○	○	○
Younger	○	○	○

Response differentiation

Question

Do you think policies in [COUNTRY] would improve **(1)**, would get worse **(2)** or would mostly stay the same **(3)** if more elected officials were…



Response differentiation: Mulligan's score

Distance-based measure of differentiation of a battery of responses

Question

Do you think policies in [COUNTRY] would improve **(1)**, would get worse **(2)** or would mostly stay the same **(3)** if more elected officials were…

Answer

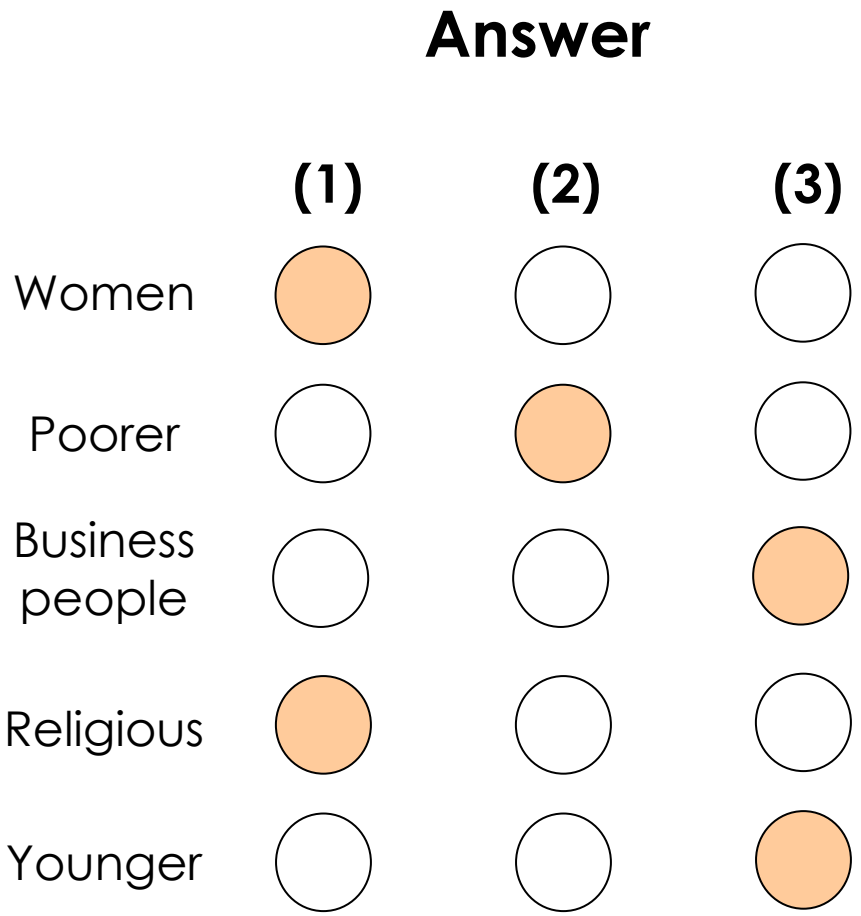
	(1)	(2)	(3)
Women	○	○	○
Poorer	○	○	○
Business people	○	○	○
Religious	○	○	○
Younger	○	○	○

Response differentiation: Mulligan's score

Distance-based measure of differentiation of a battery of responses

Question

Do you think policies in [COUNTRY] would improve **(1)**, would get worse **(2)** or would mostly stay the same **(3)** if more elected officials were…



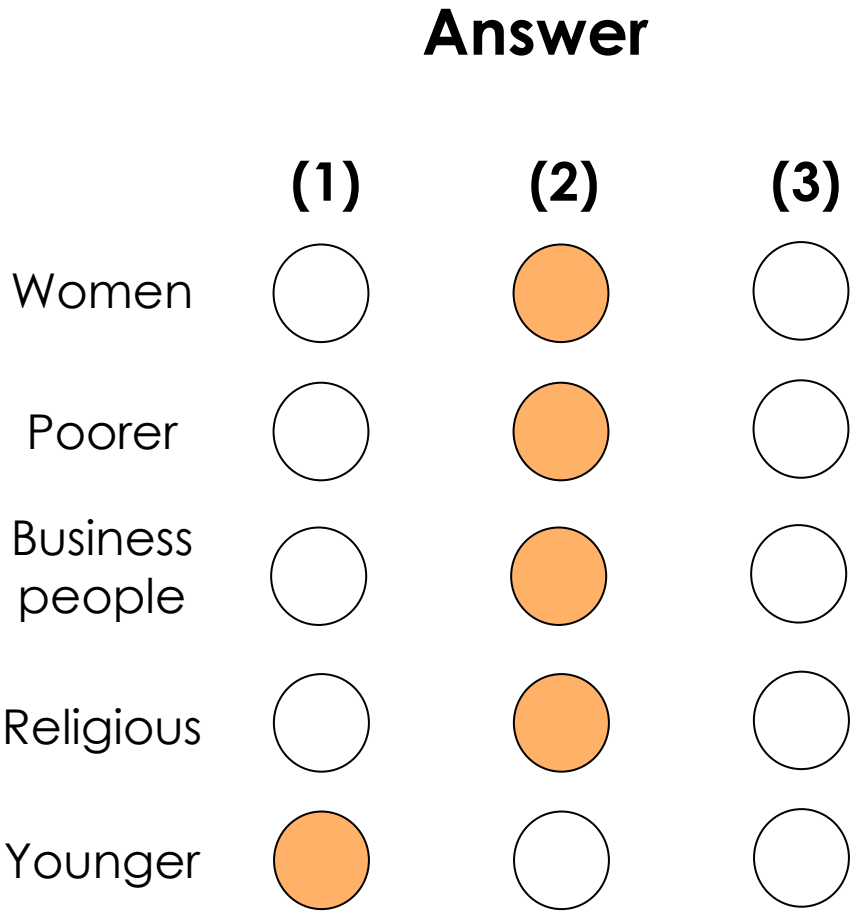
Mulligan's score = 1.93

Response differentiation: Mulligan's score

Distance-based measure of differentiation of a battery of responses

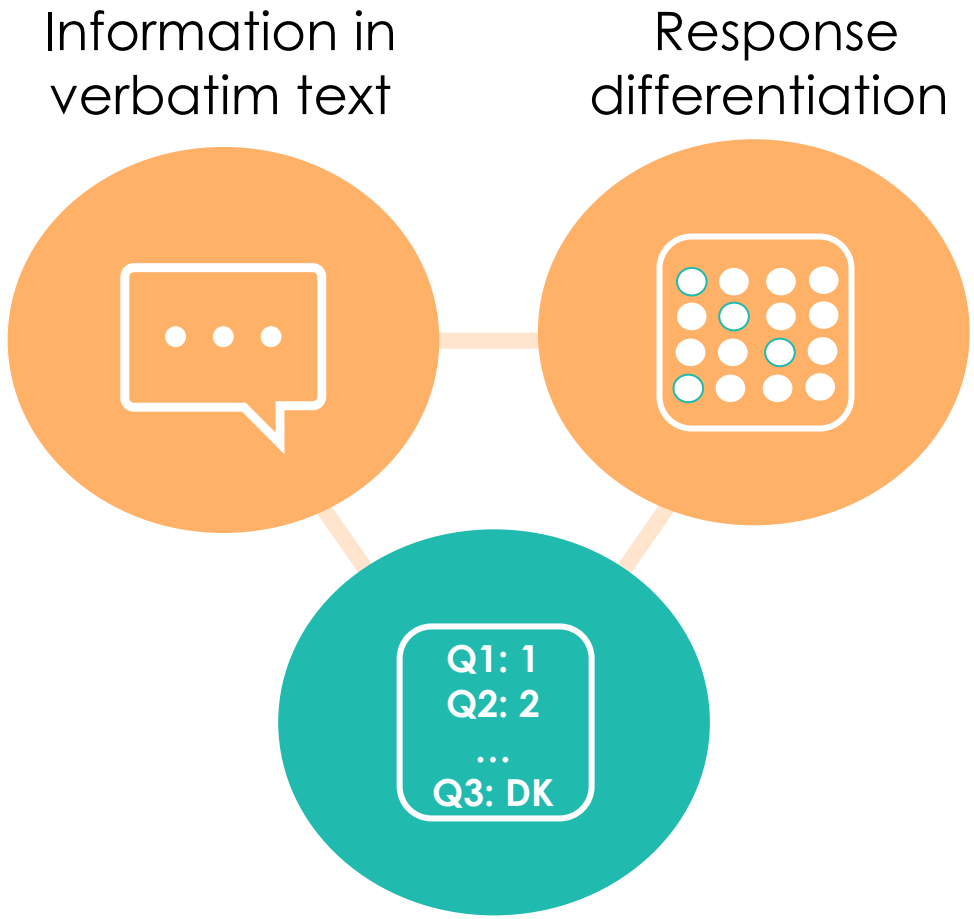
Question

Do you think policies in [COUNTRY] would improve **(1)**, would get worse **(2)** or would mostly stay the same **(3)** if more elected officials were…



Mulligan's score, with evidence of straightlinning = 0.8

Measuring response characteristics



Response characteristics

Item non-response

Calculated the number of **don't know** or **prefer not to say** answers across 49 comparable questions (excluding the verbatim).

Question	Answer
1	**DK**
2	6
3	2
4	AL
5	**DK**
6	3
7	**DK**
…	
49	5

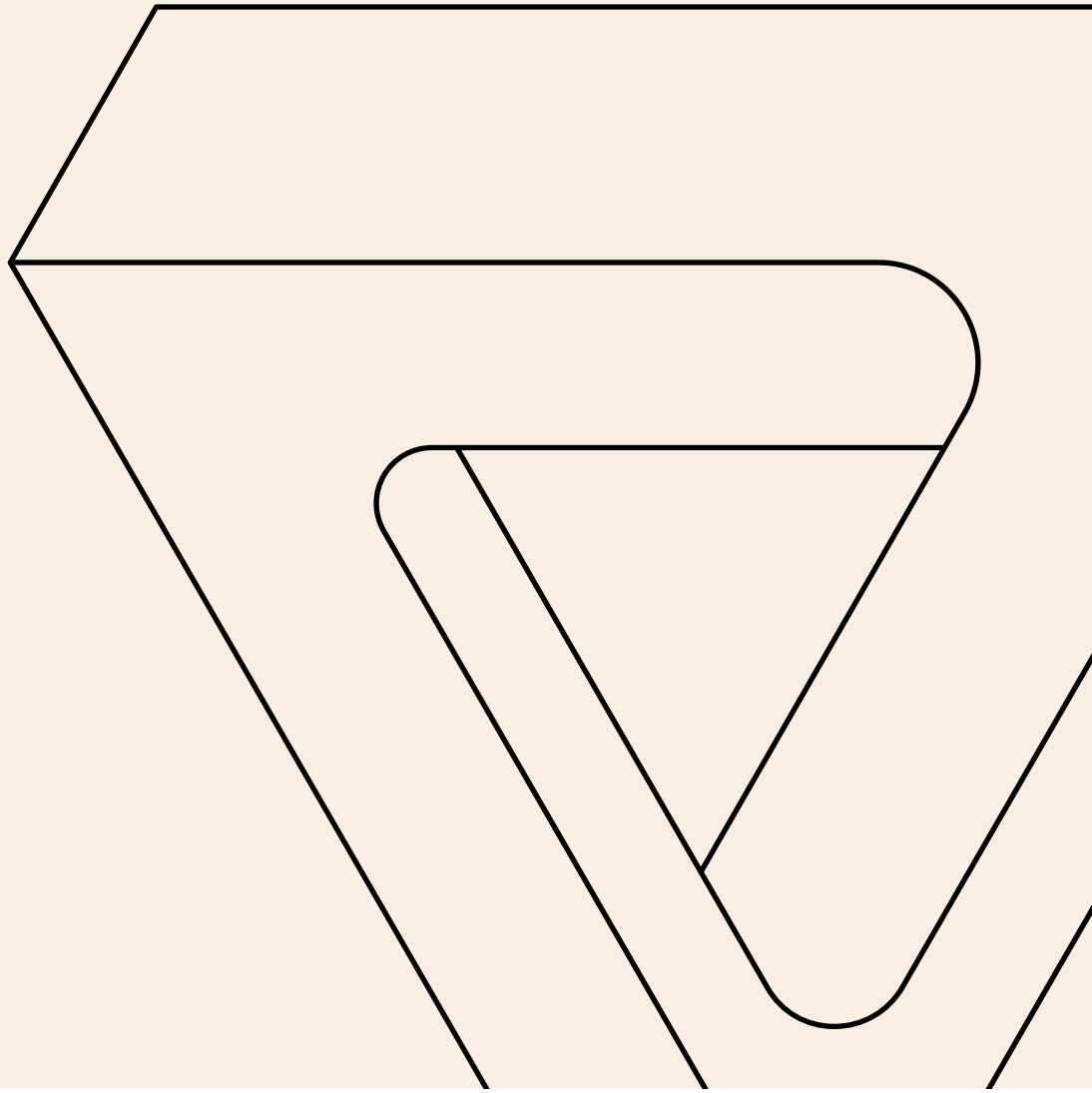
Statistical models

For each index, we fit a general linear model with method (telephone versus web) as a 0-1 dummy variable along with sex, age and education as covariates

B_1 – parameter is the effect of changing from telephone to web survey mode

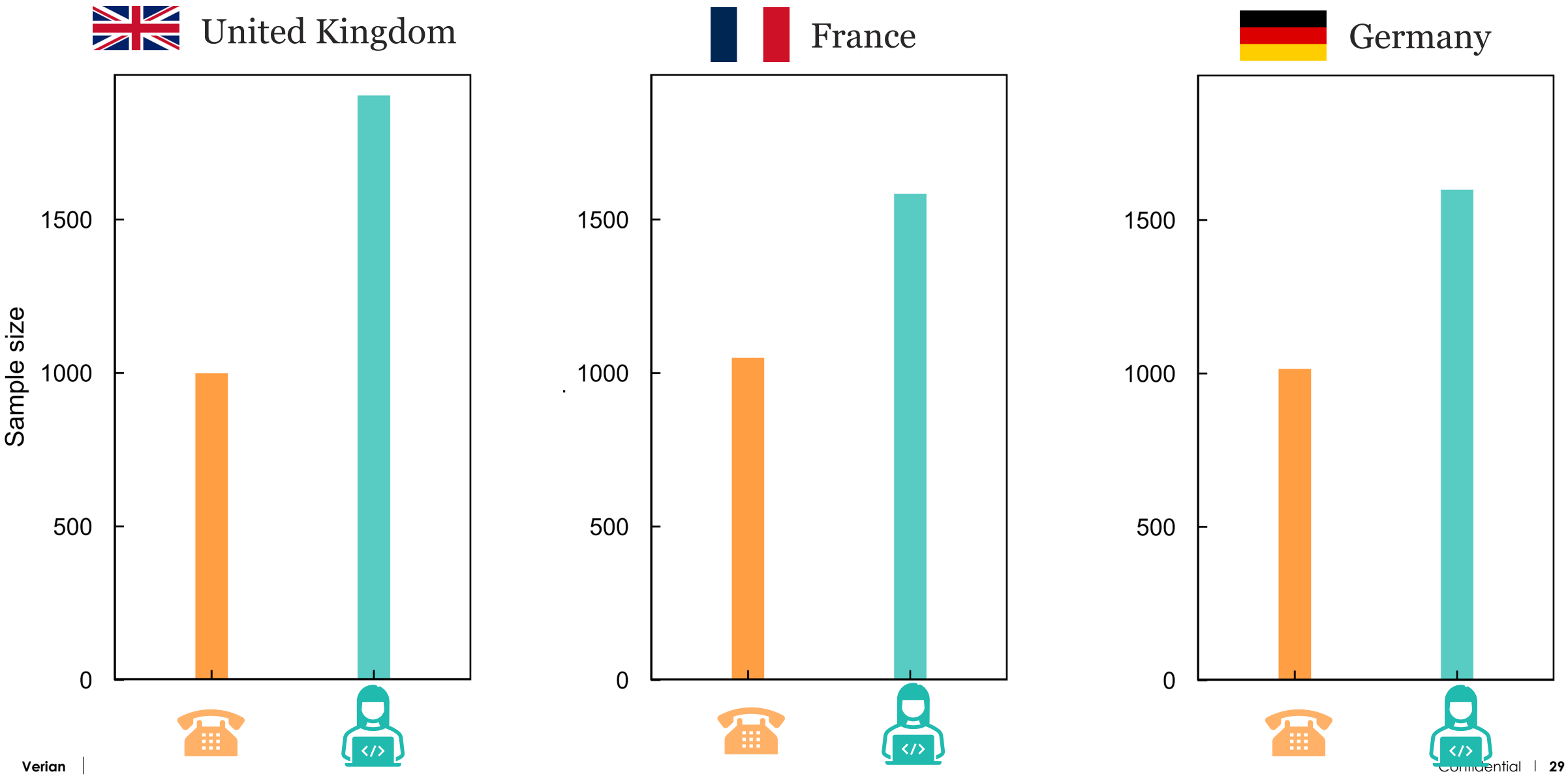
Informativeness (I)	Mulligan's score (M)	Item non-response (DK)
$\ln(I_i) \sim \text{Normal}(\mu_i, sd)$ $\mu_i = a + \mathbf{\beta_1 * method_i} + \beta_2 * sex_i + \beta_3 * age_i + \beta_4 * education_i$	$M_i^2 \sim \text{Normal}(\mu_i, sd)$ $\mu_i = a + \mathbf{\beta_1 * method_i} + \beta_2 * sex_i + \beta_3 * age_i + \beta_4 * education_i$	$N_{DK,i} \sim \text{Binomial}(50, p_i)$ $\text{logit}(p_i) = a + \mathbf{\beta_1 * method_i} + \beta_2 * sex_i + \beta_3 * age_i + \beta_4 * education_i$

Results

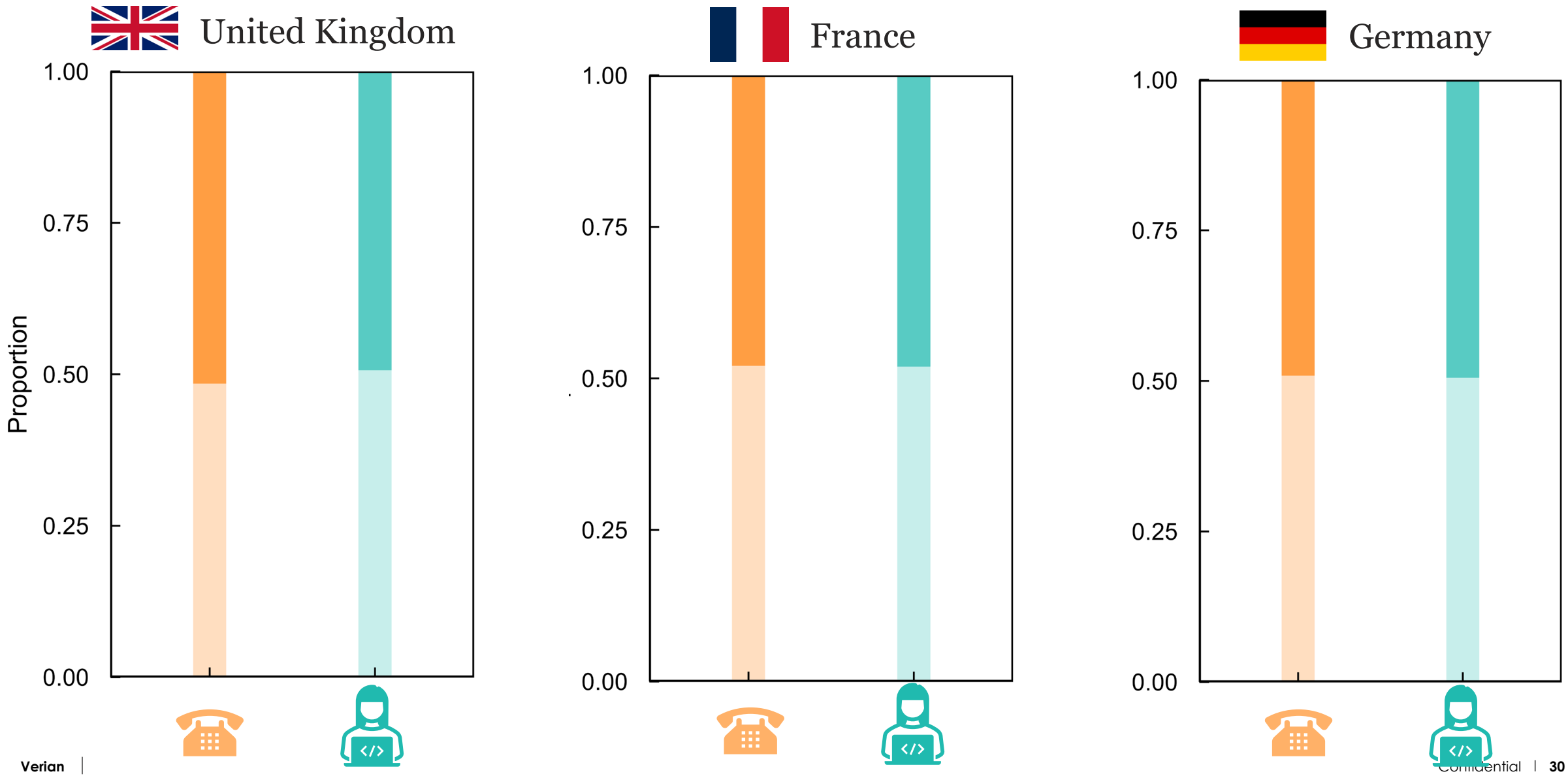


verian

Sample characteristics (n = 8151)

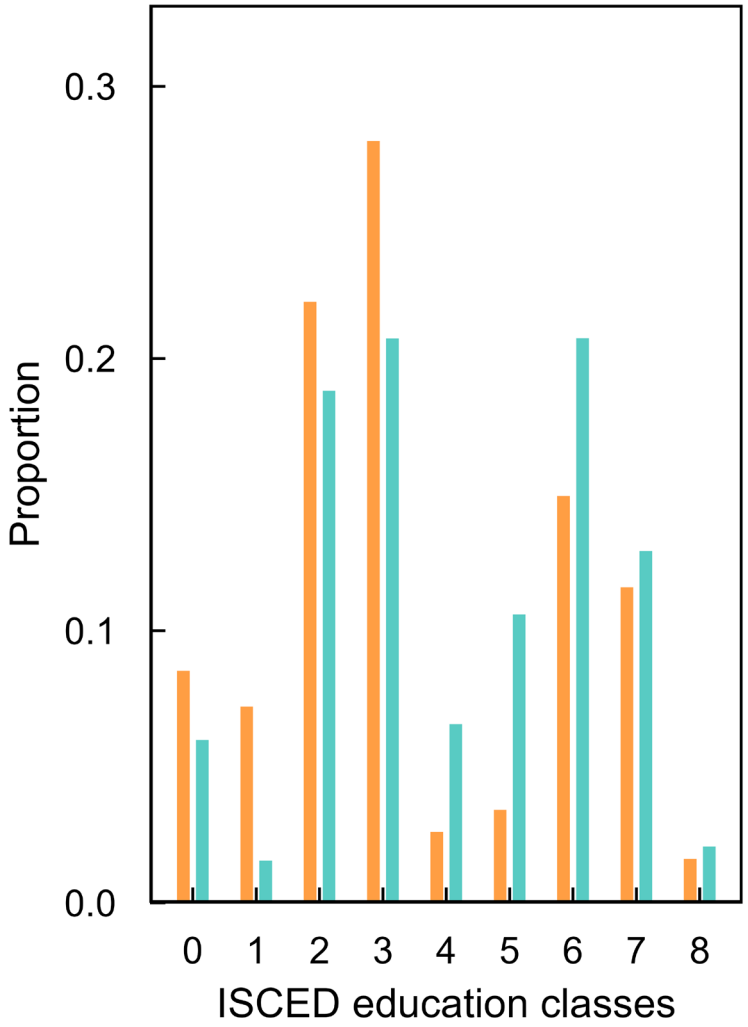


Gender/sex balance

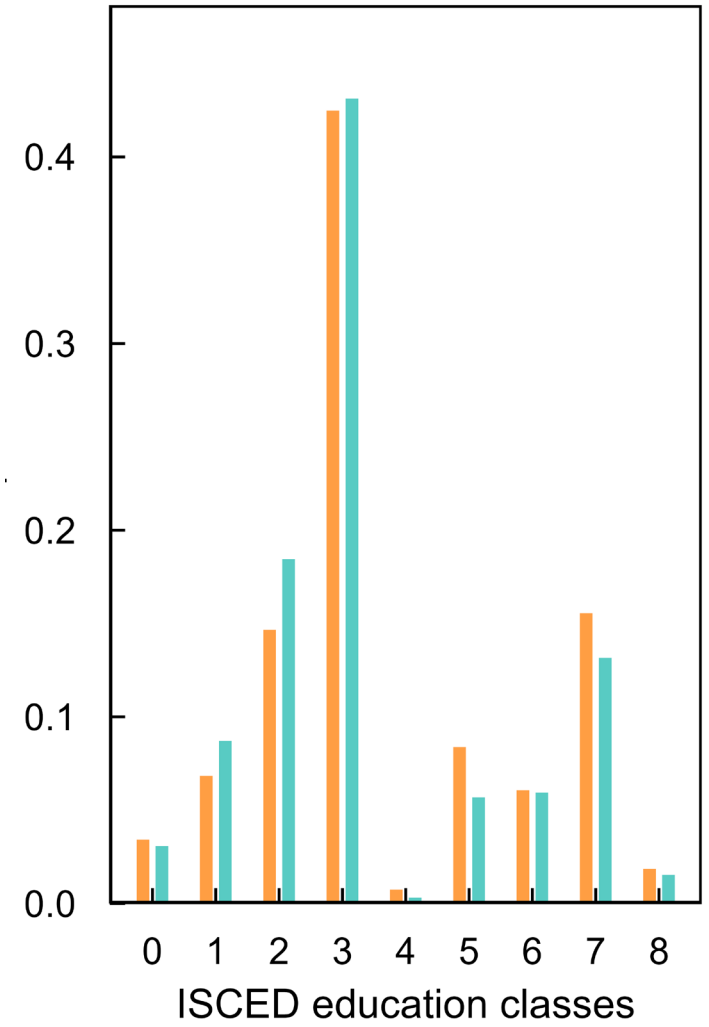


Education (ISCED scale)

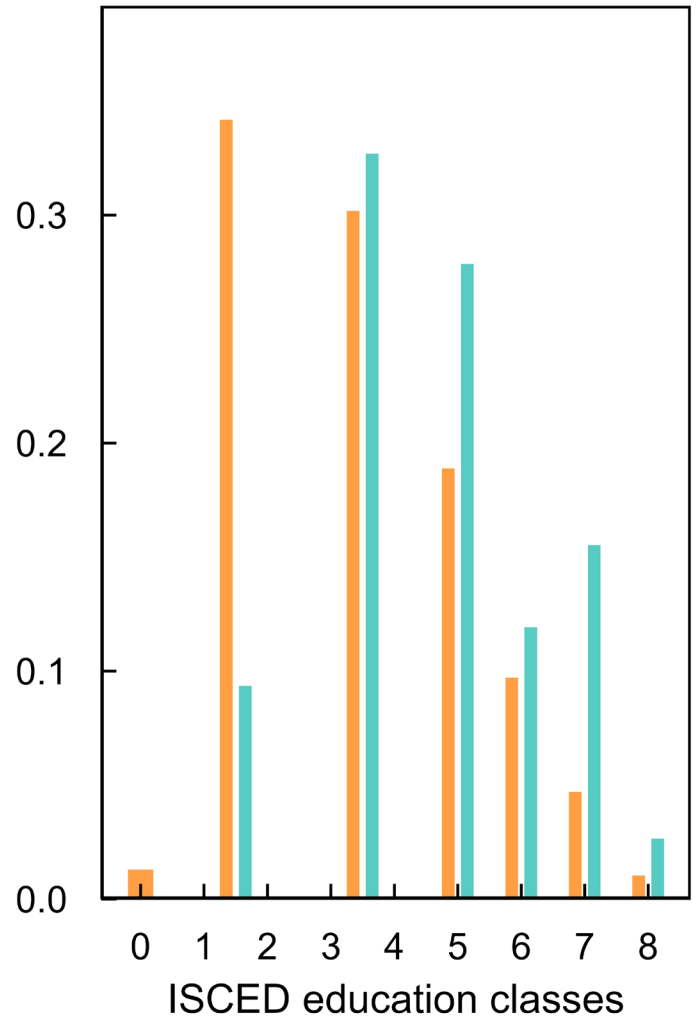
United Kingdom



France

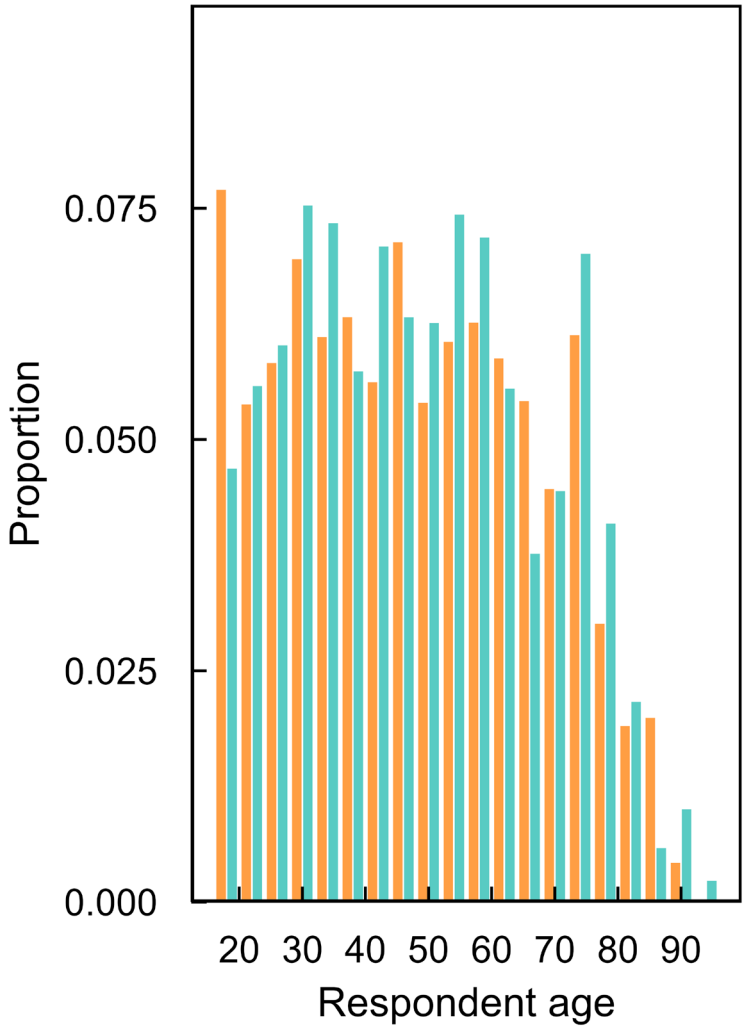


Germany

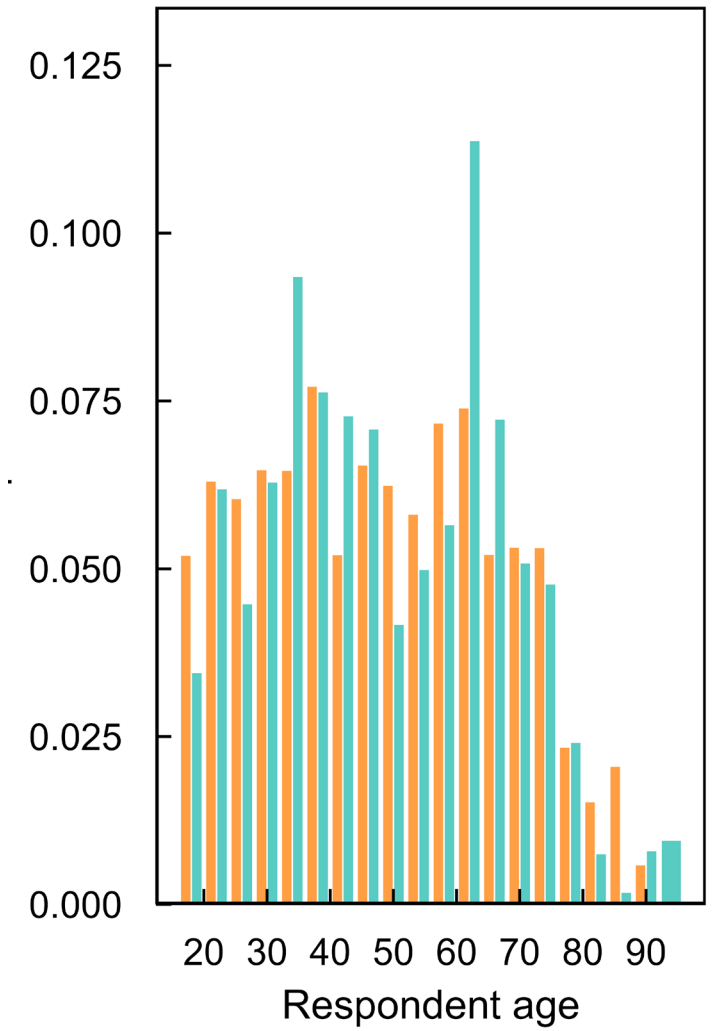


Age

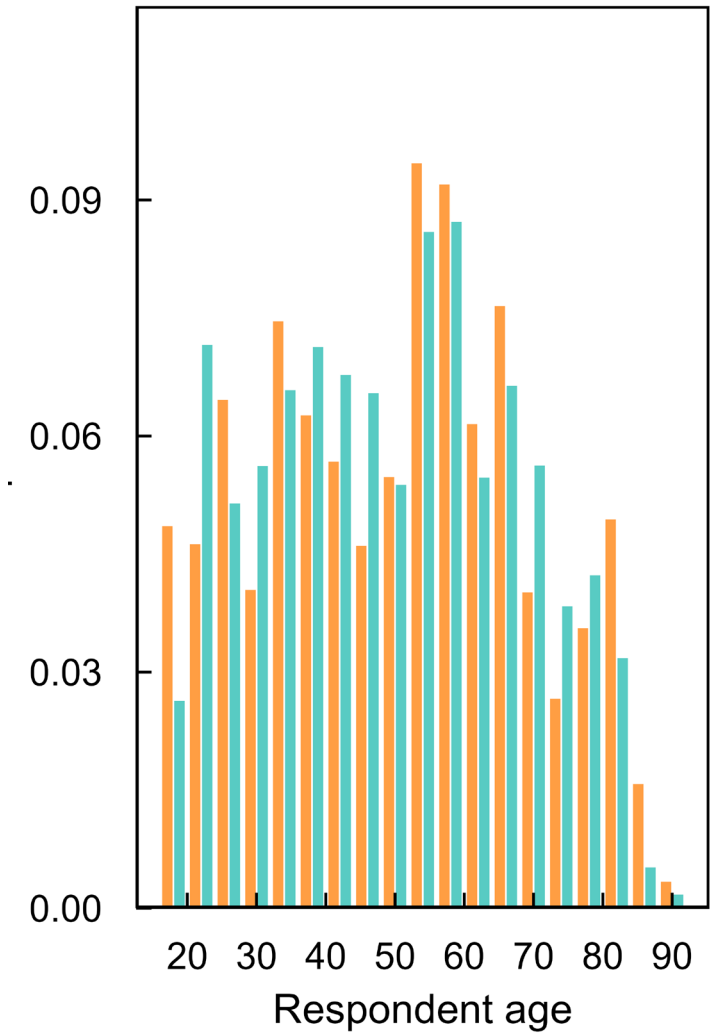
United Kingdom



France

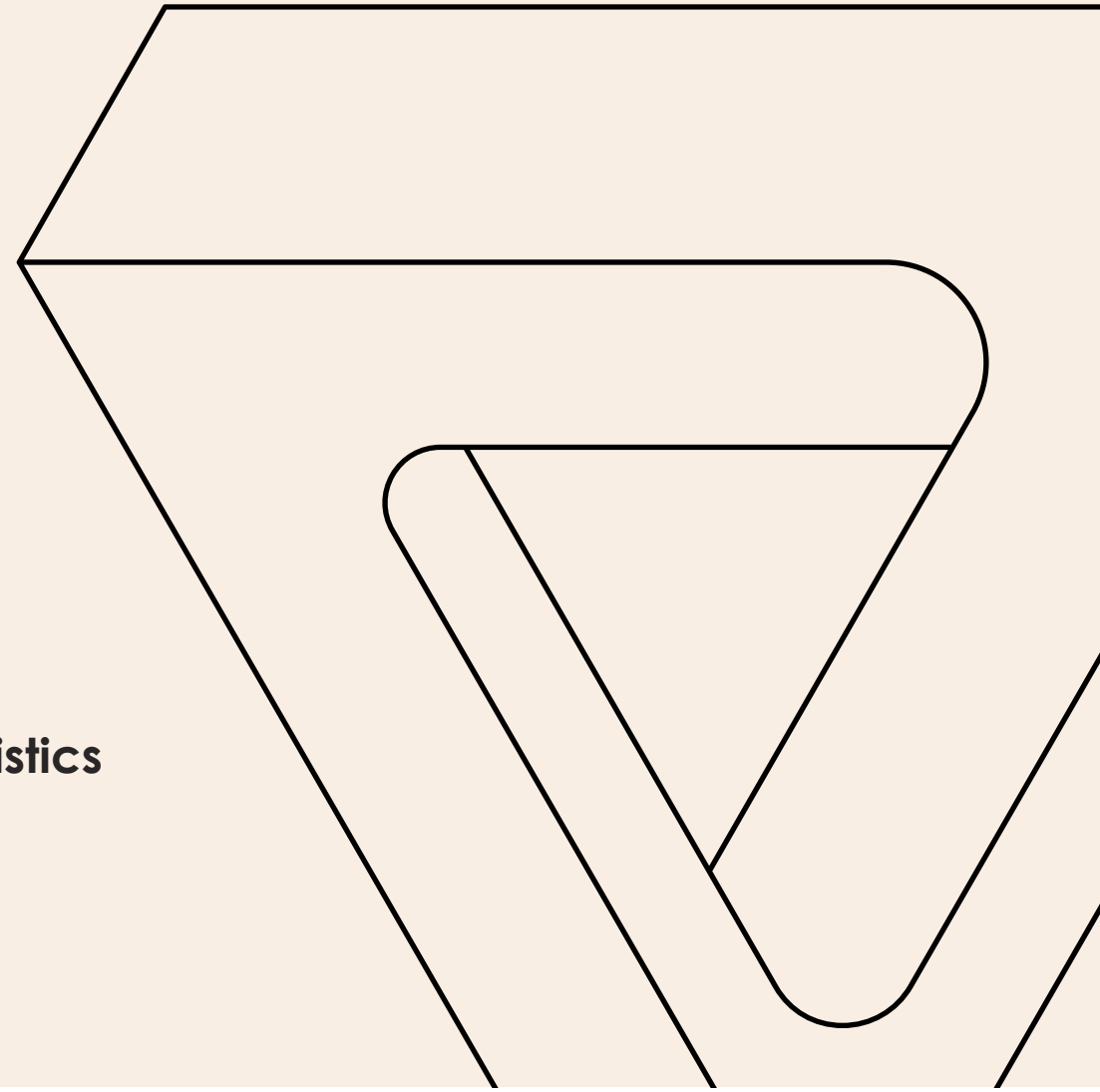


Germany



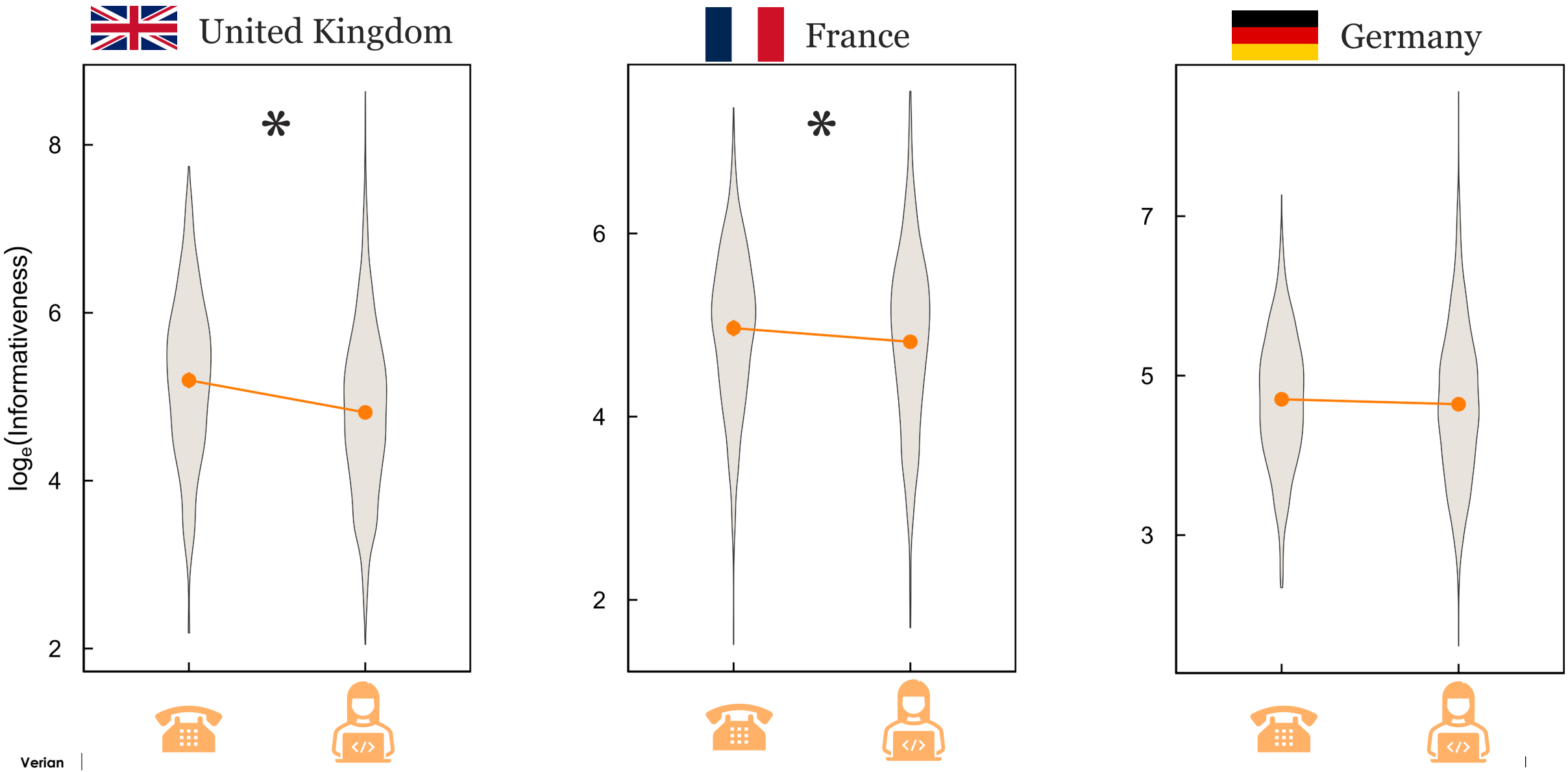
RQ1:

Do we observe differences in terms of response characteristics across telephone and web for the same questions?

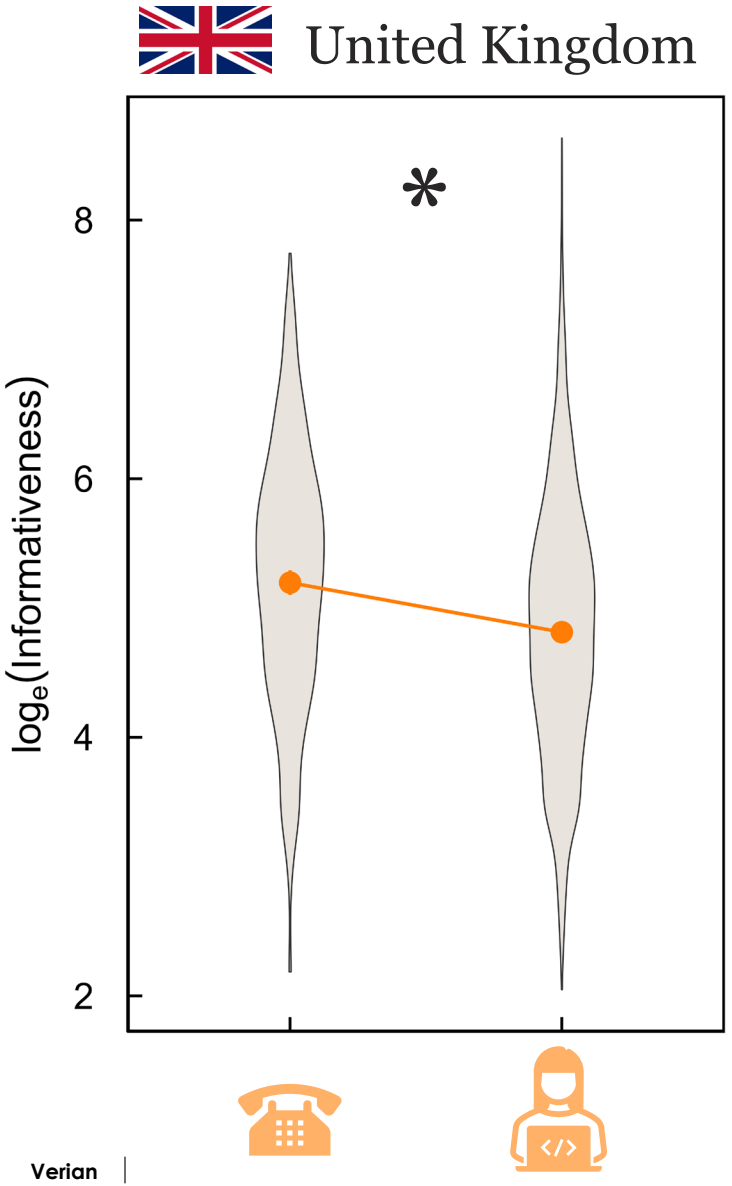


verian

Information in verbatim text



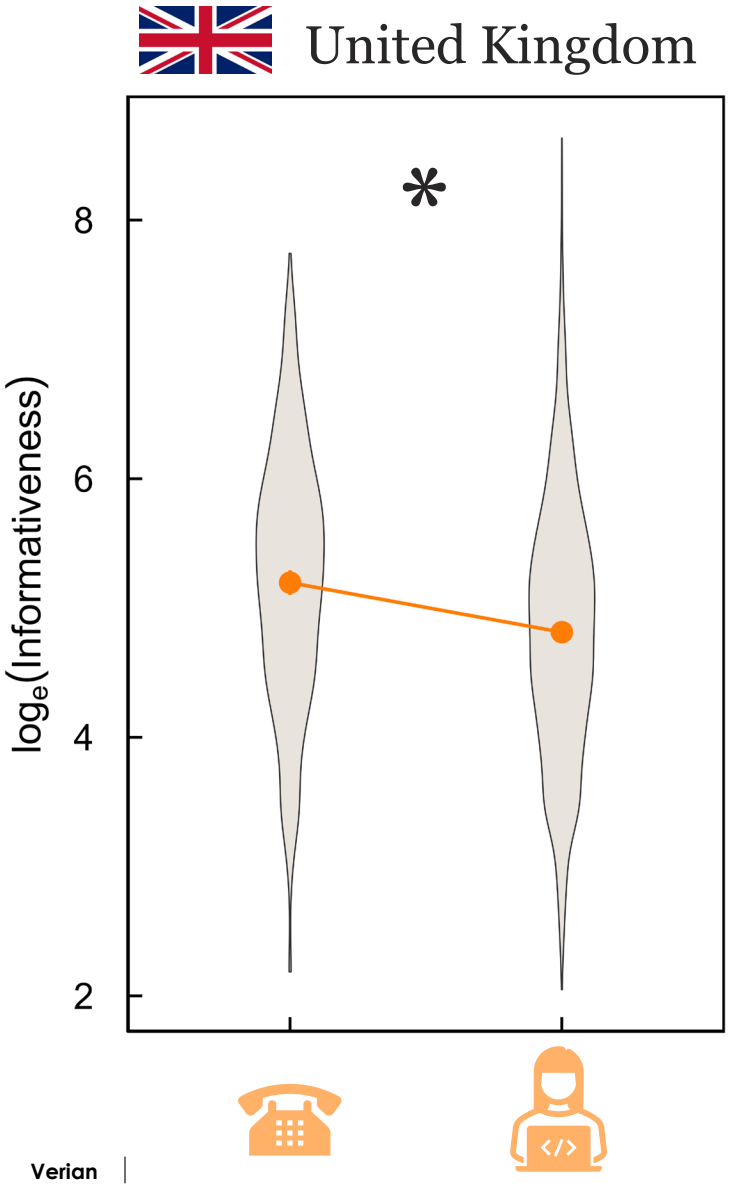
Information in verbatim text



$B \pm SE = -0.38 \pm 0.04$

In the UK, the effect was the strongest with an average change in log-informativeness of 0.38 between telephone and web surveys.

Information in verbatim text

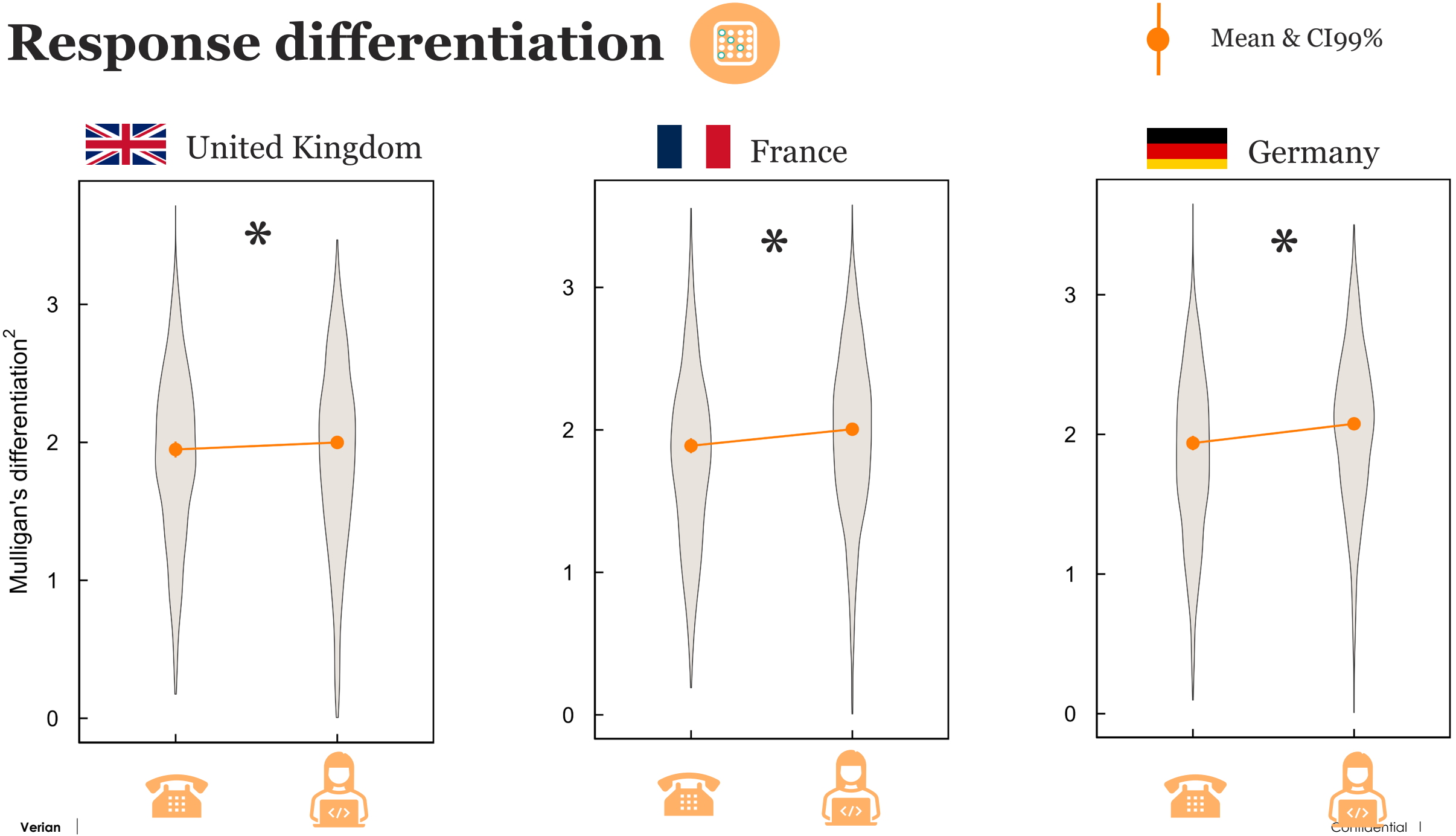


$B \pm SE = -0.38 \pm 0.04$

"Different ways of voting. I think people should have working experience, knowledgeable of the office that they represent"

"People are too divided these days, we need to meet in the middle"

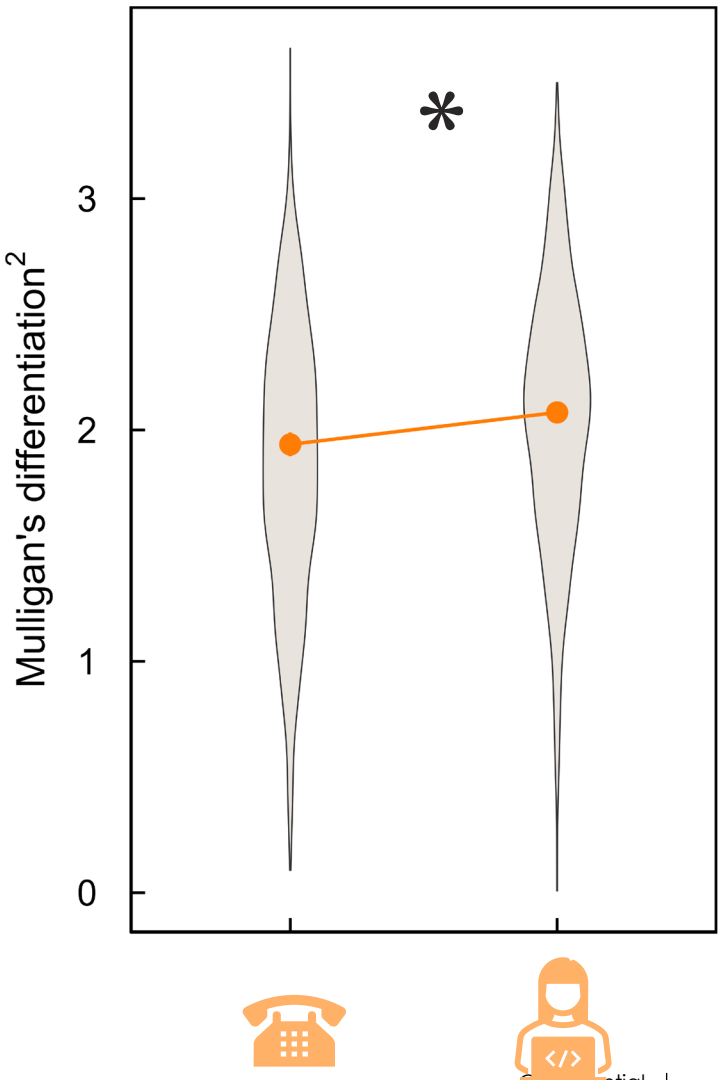
Response differentiation



Response differentiation

Mean & CI99%

Germany



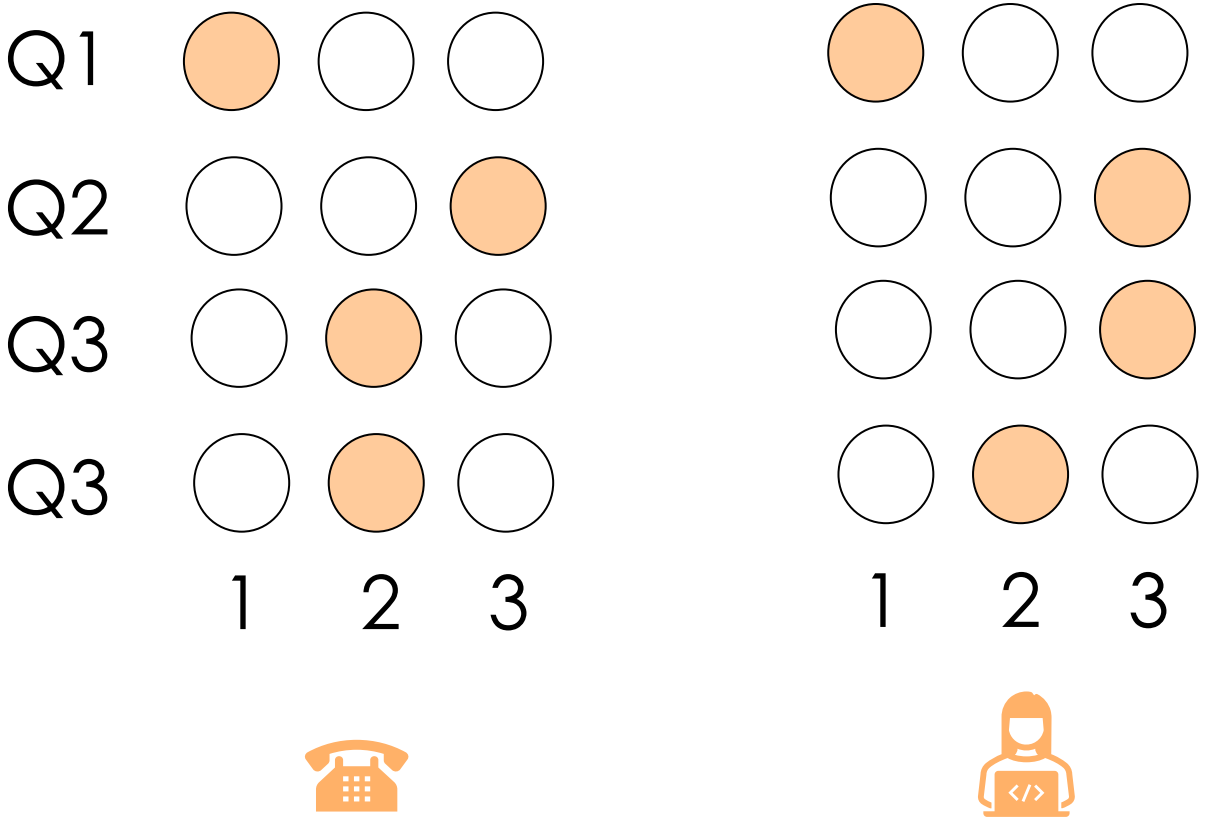
However, although there was a statistically detectable effect of survey mode on response differentiation, the effect was minimal.

Even in Germany, where the effect was strongest, the difference is minor.

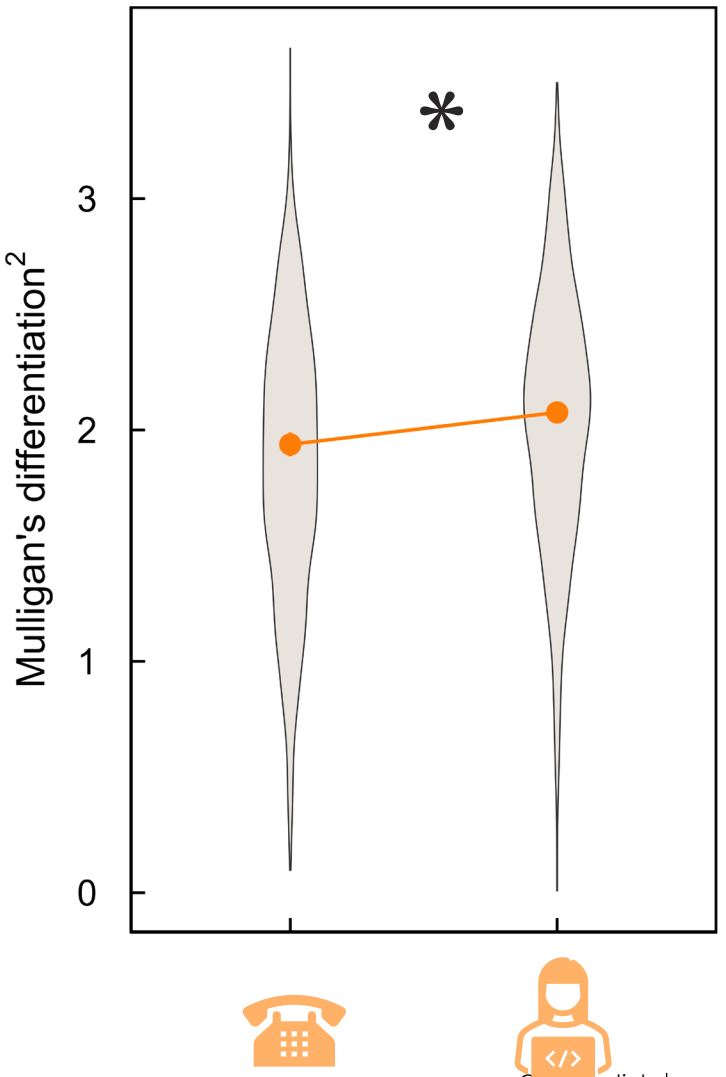
Response differentiation

Mean & CI99%

$B \pm SE = 0.14 \pm 0.02$



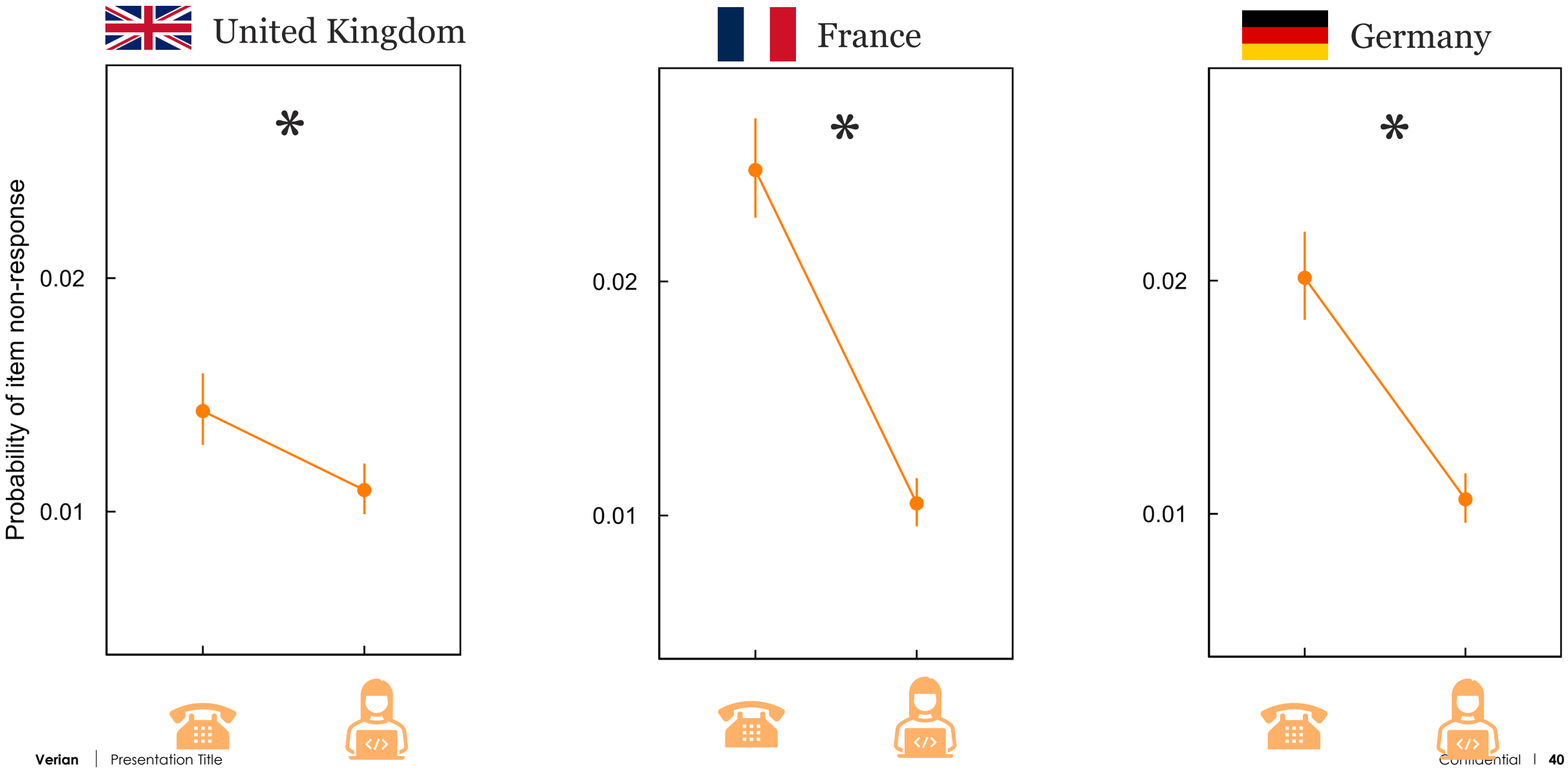
Germany



Item non-response

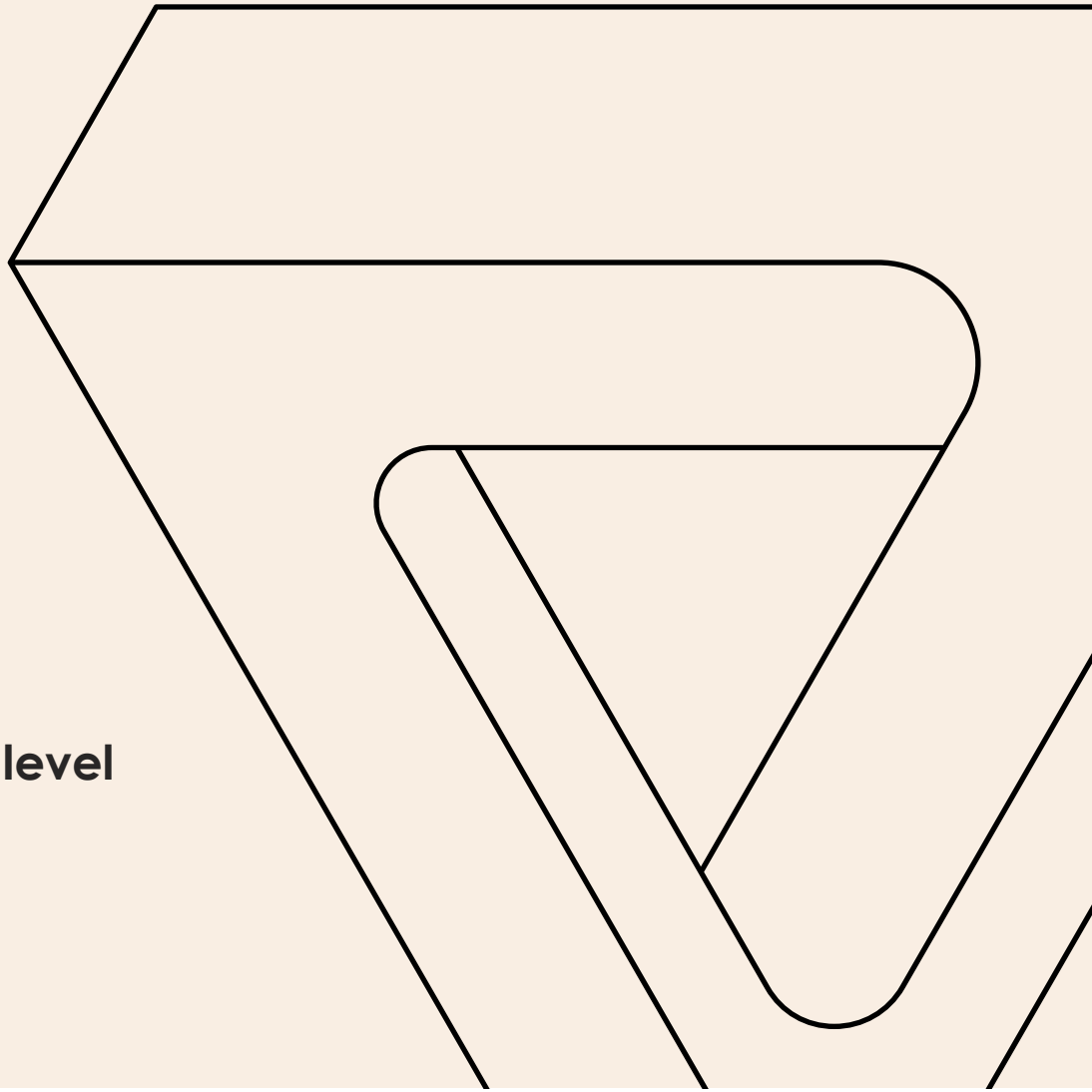
Q1: 1
...
Q3: DK

Mean & CI99%



RQ2:

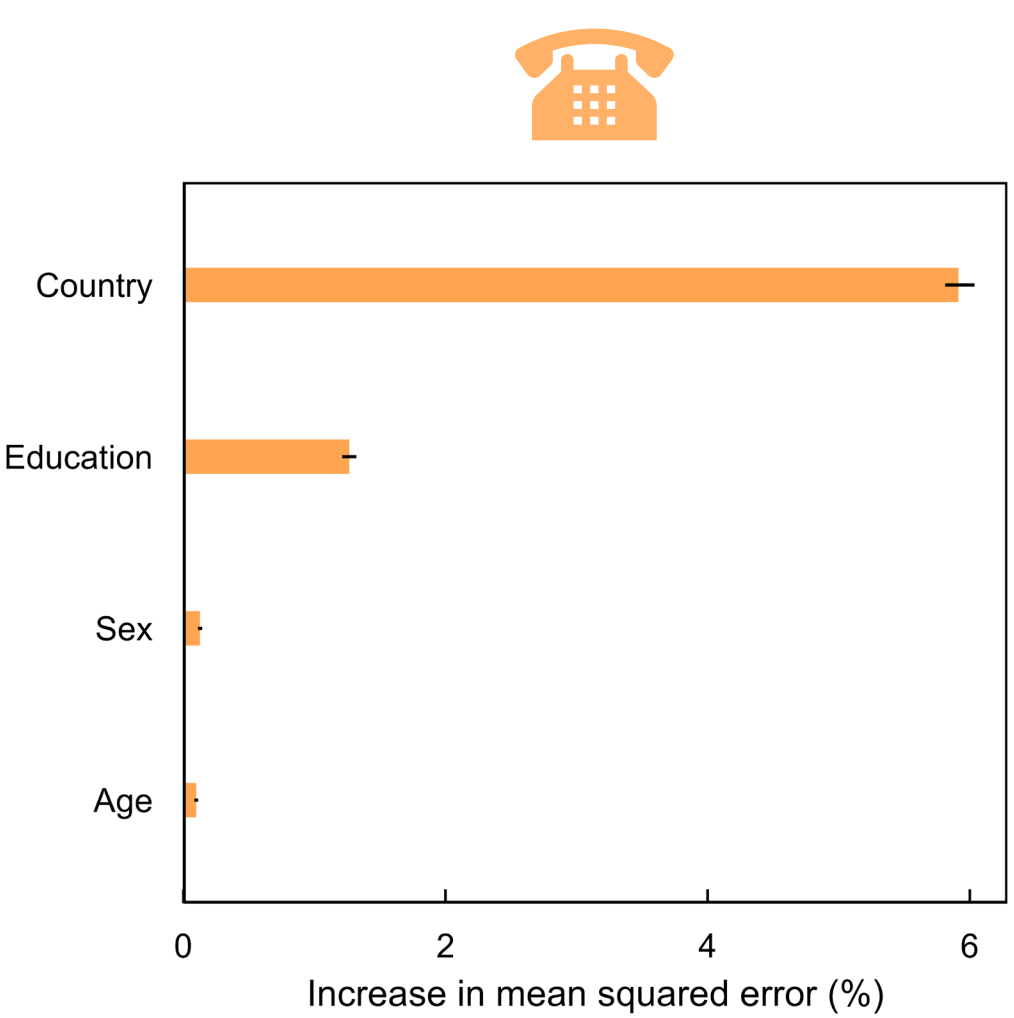
Can we explain these differences by looking at individual-level differences (gender, age, education, country)?

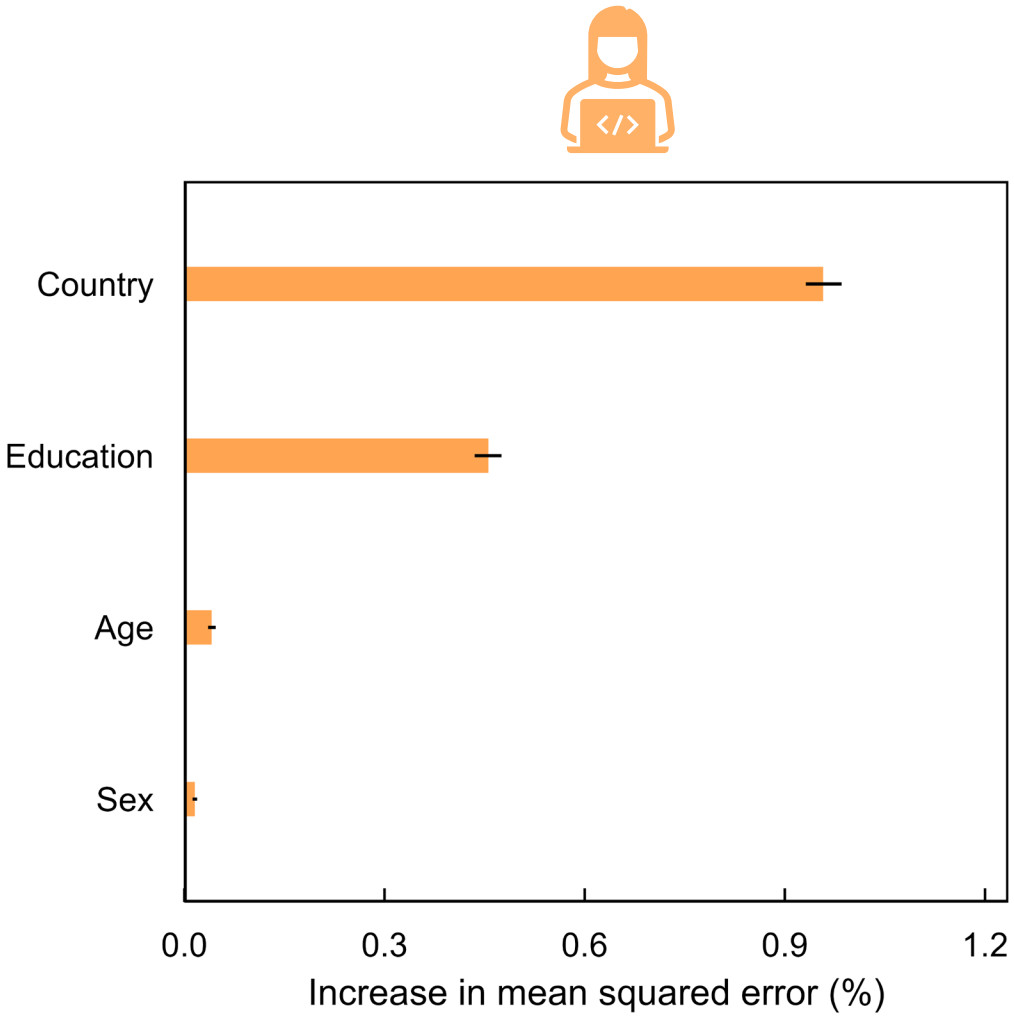


verian

Relative predictive importance

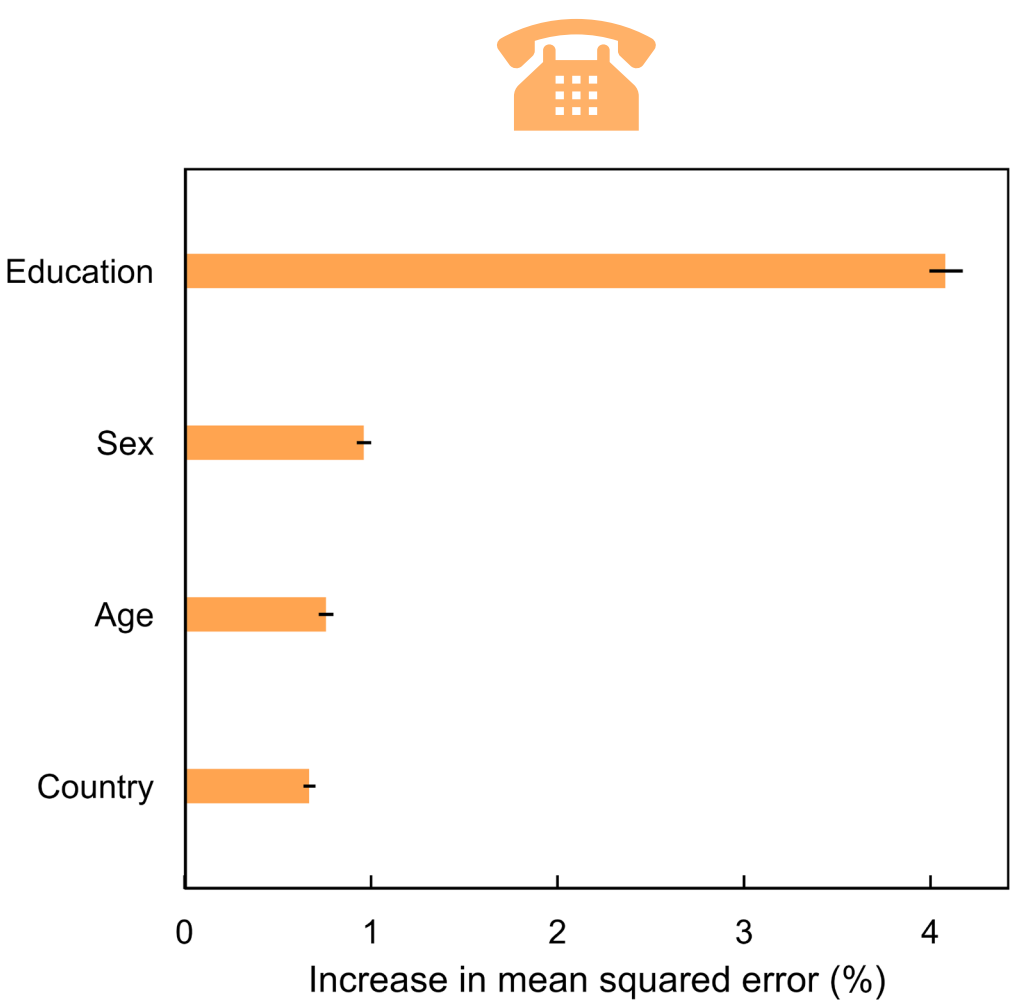
Information in verbatim text: Informativeness

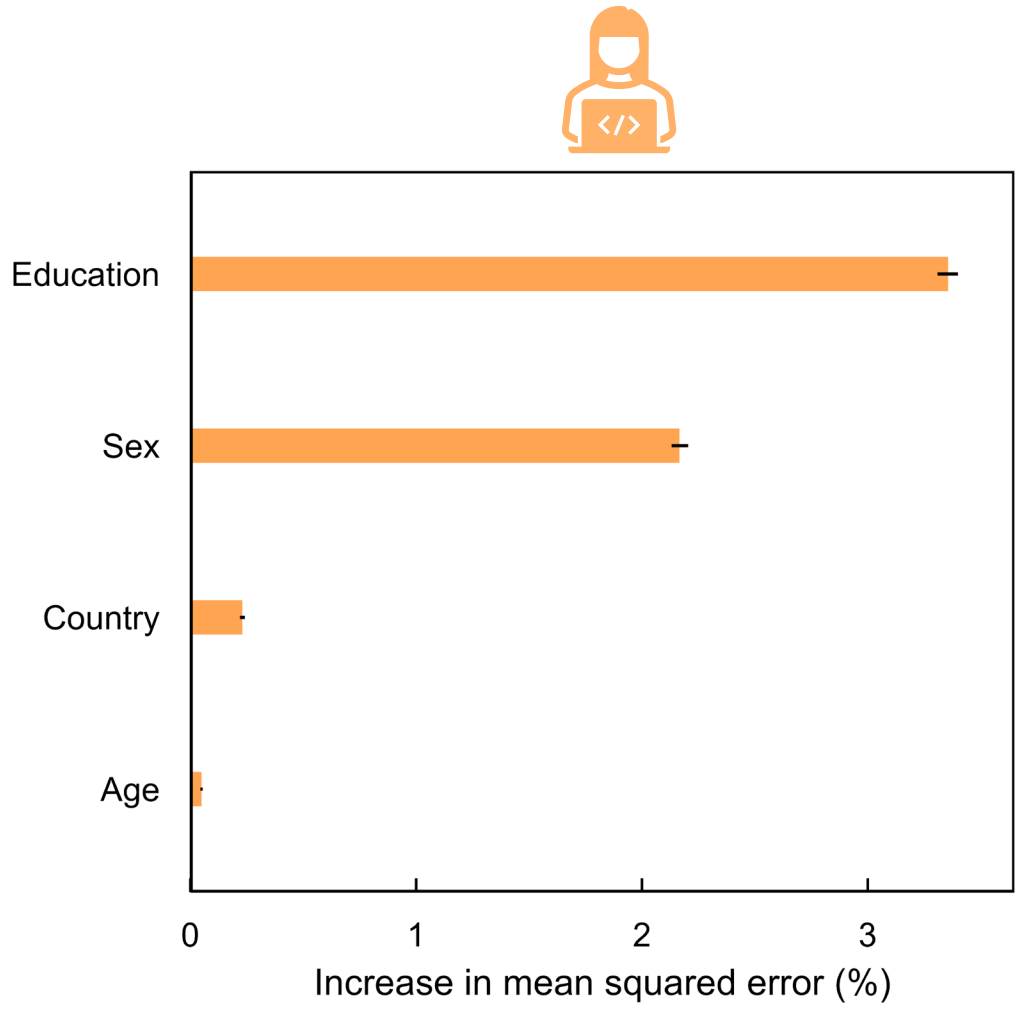




Relative predictive importance

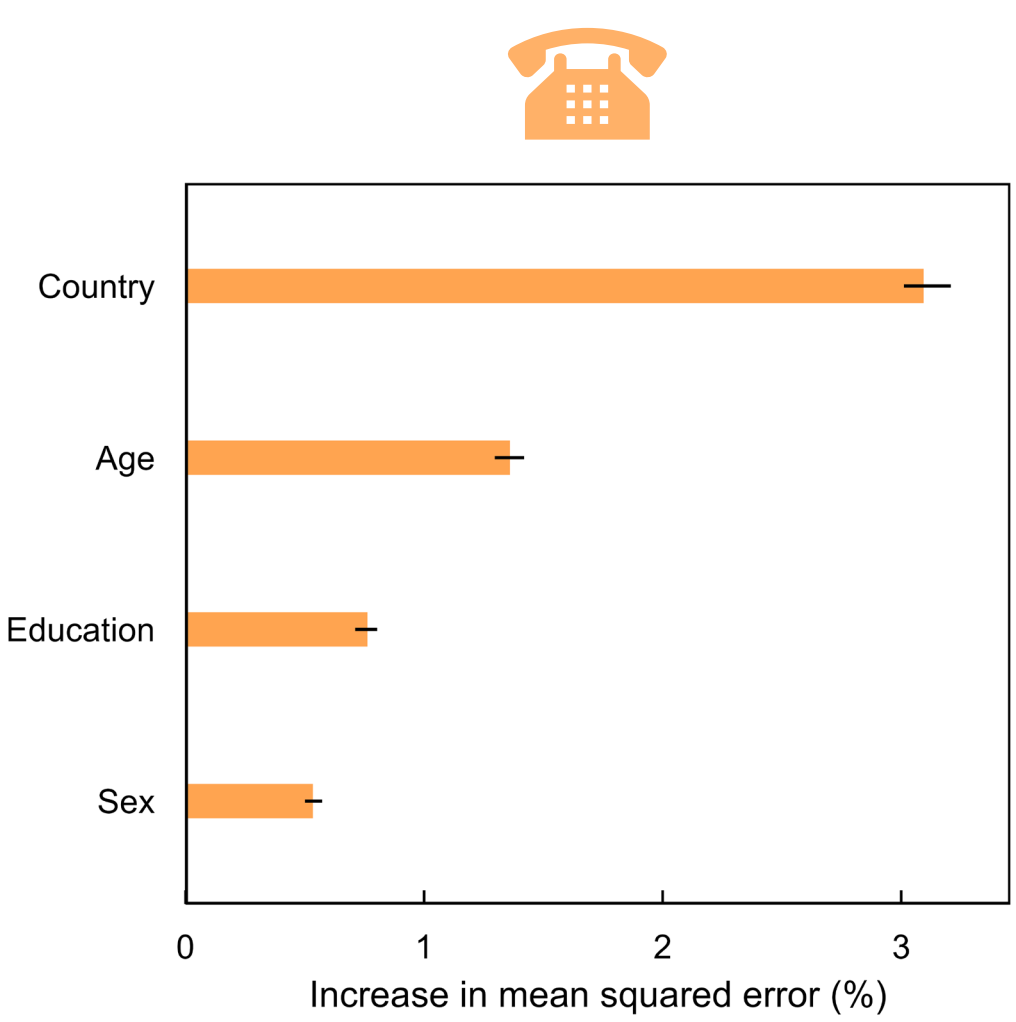
Response non-differentiation: Mulligan's score

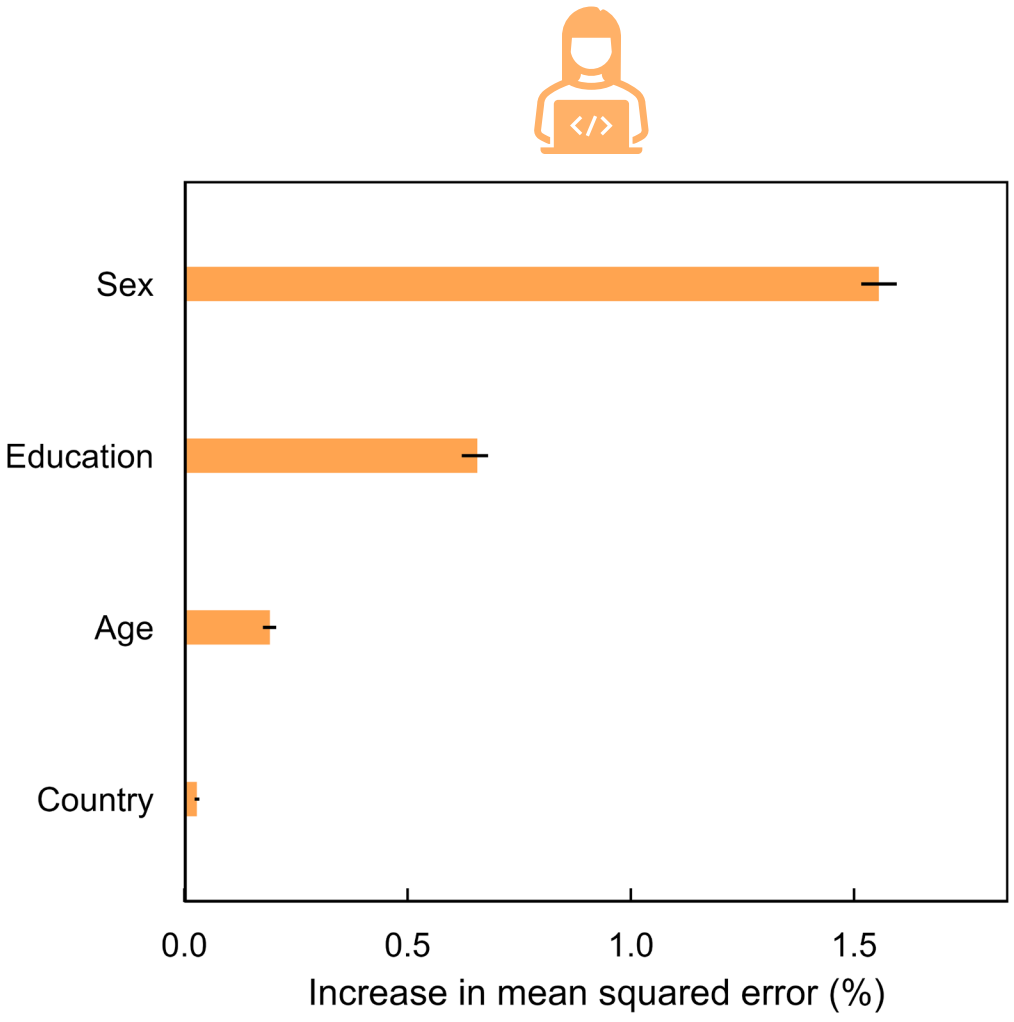




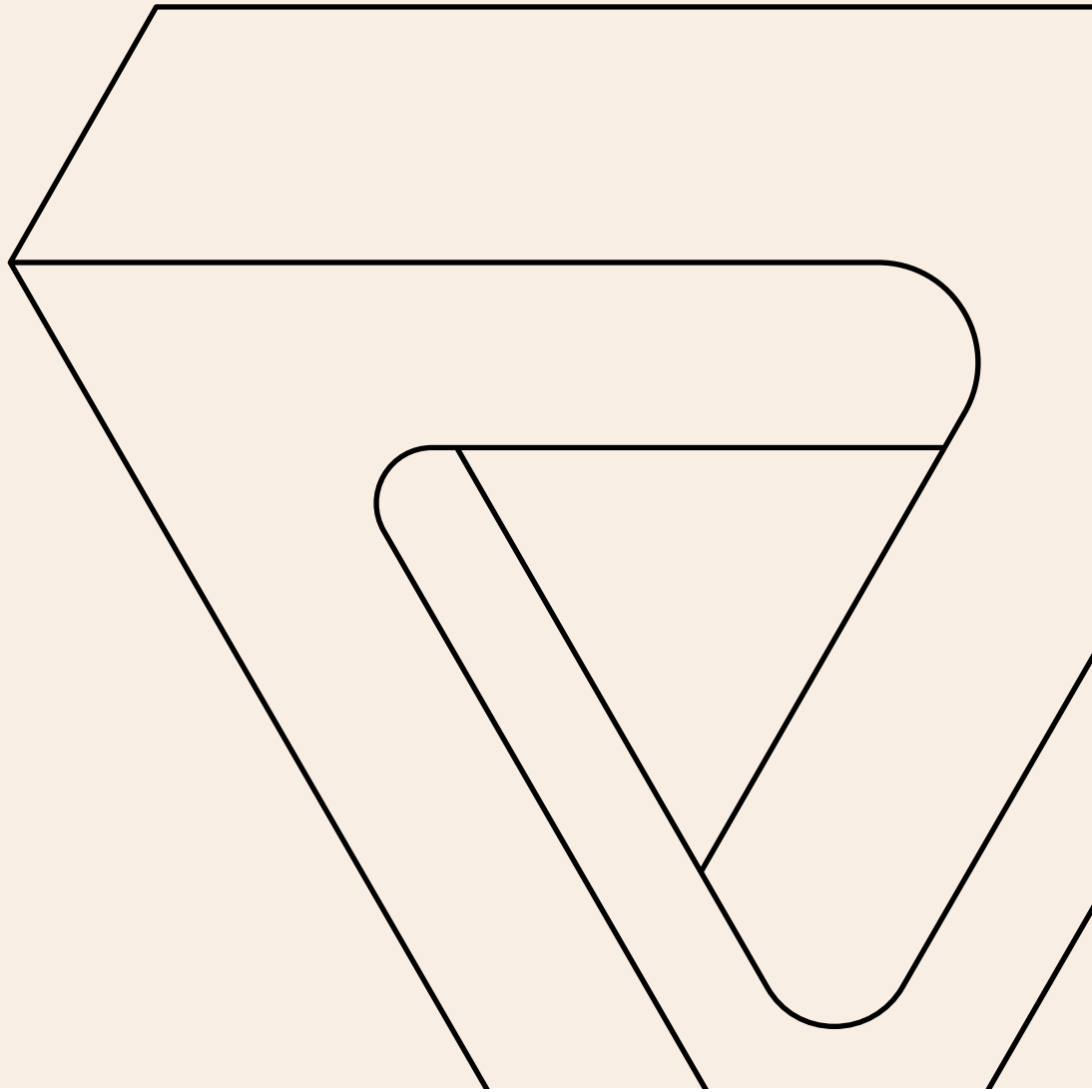
Relative predictive importance

Item non-response rates





Conclusions (and next steps)



verian

Conclusions and next steps

- No sizable differences between online and phone on the three metrics we've looked at
- The analysis had some caveats and was limited to some areas so would be interesting to look at the A/B verbatim experiment and whether the cohort encouraged to provide a more throughout responses did better?
- Additionally, we would like to take a closer look at is whether non-response changes as the survey goes on and does this differ across modes?
- Considering some large attitudinal differences by mode (we presented on at ESRA) we'd like a closer look at the substantive questions

Thank you

Tanja Kimova
tanja.kimova@veriangroup.com

James Hagan
James.hagan@veriangroup.com

Alexandrina Buruian
Alexandrina.buruian @veriangroup.com

Patrick Moynihan
PMoynihan@PewResearch.org

verian

Powering decisions
that shape the world.