

# Providing extra incentives for open voice answers in web surveys

**Höhne<sup>1,2</sup>, Revilla<sup>3</sup>, Couper<sup>4</sup>, Claassen<sup>1,2</sup>**

<sup>1</sup> Leibniz University Hannover

<sup>2</sup> German Center for Higher Education Research and Science Studies (DZHW)

<sup>3</sup> RECSM at Universidad Pompeu Fabra

<sup>4</sup> University of Michigan

**7th MASS Workshop & WEB DATA OPP Project**

Barcelona (Spain) – June 4, 2026

The WEB DATA OPP project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 849165)

Melanie Revilla is Principal Investigator (PI)



This study was preregistered through Open Science Framework (OSF)



# Introduction

- Linking established methods with technological innovations
  - *Voice communication channel in the form of open narrative answers*
- However, voice answers struggle with missing data (e.g., item nonresponse)
- Revilla and Couper (2021) varied voice instructions but found no differences
- Revilla and Couper (2023) showed that voice answer provision is higher for ...
  - *... respondents using voice input in daily life*
  - *... respondents trusting that their answers are treated confidentially*
- We go beyond prior studies and provide extra incentives
  - *We focus on respondent varying in their likelihood of giving voice answers: low, medium, & high*
  - *In particular, we expect extra incentives to tip the scale in the medium likelihood group*

# Defining Likelihood Groups

| Likelihood groups | Descriptions                                                                                                                                                                                                    |
|-------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Low               | Respondents who report <u>not being aware</u> of the existence of voice recording or <u>never using</u> it in their daily life and <u>not completely trusting</u> that their answers are treated confidentially |
| Medium            | Respondents who report <u>not being aware</u> of the existence of voice recording or <u>never using</u> it in their daily life, but <u>completely trusting</u> that their answers are treated confidentially    |
|                   | Respondents who report <u>using at least sometimes</u> voice recording in their daily life, but <u>not completely trusting</u> that their answers are treated confidentially                                    |
| High              | Respondents who report <u>using voice recording at least sometimes</u> in their daily life and <u>completely trusting</u> that their answers are treated confidentially                                         |

# Study and Push-to-Voice Design

The figure displays three screenshots of the 'nicequest' web survey interface, illustrating different question types and voice recording options.

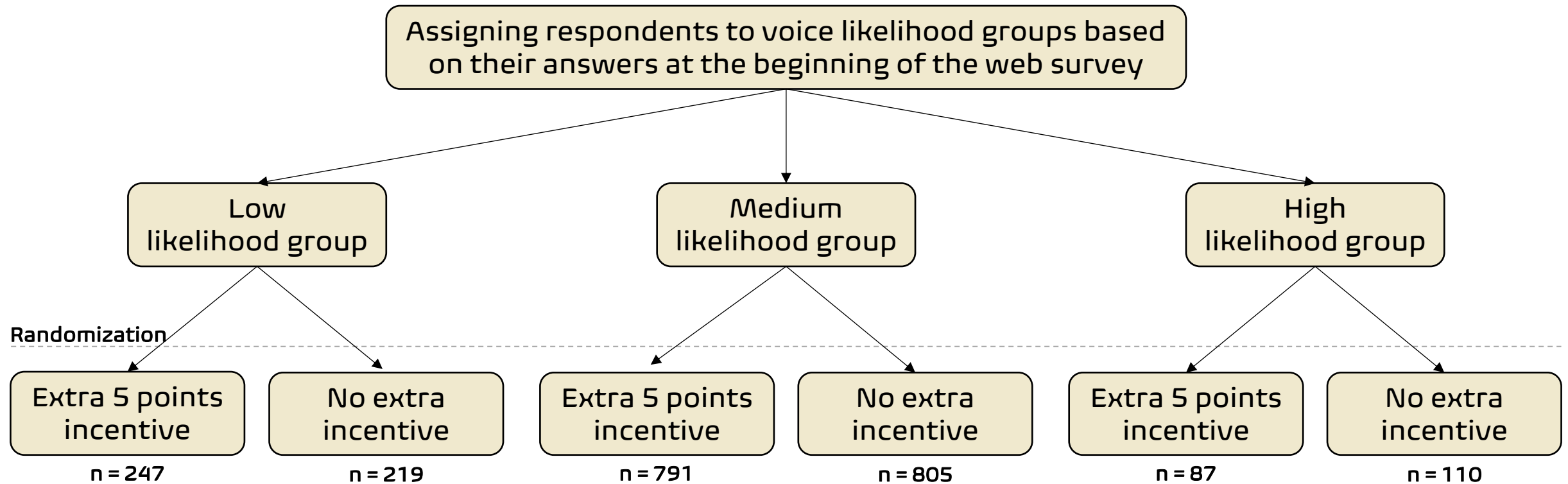
**1st closed question:** The first screenshot shows a question about the amount of information provided by services for the elderly in Spain. The response options are radio buttons: 'Ninguna información', 'Poca información', 'Alguna información', 'Mucha información' (selected), and 'Muchísima información'. A 'Variables' button is visible at the bottom left.

**1st probe with voice answer option:** The second screenshot shows a question about the reasons for providing services. It includes a text input field and a 'Grabar' (Record) button. A 'Variables' button is visible at the bottom left.

**1st probe with voice + text answer option:** The third screenshot shows a question about the reasons for providing services. It includes a text input field and a 'Grabar' (Record) button. A 'Variables' button is visible at the bottom left.

- Web survey lasted 10 min (N = 2,259)
  - Respondents got an incentive of 12 points – 2,000 points equal 100€
  - Cross-quotas on gender and age plus education
- 2 follow-up probes on 2 closed questions
  - We used the WebdataVoice tool (Revilla et al. 2022)
- <-- Example screenshots (PC screen)
  - If respondents refuse to answer the probe through voice, they receive a text option
- Sample characteristics
  - Mean age (48 years), female (52%), medium (25%), and high education (34%), and smartphone (72%)

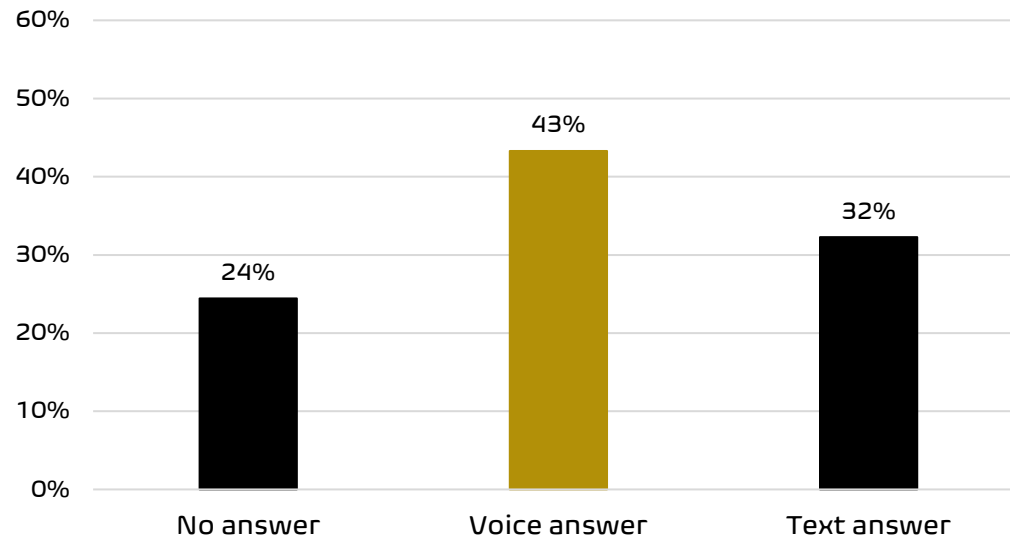
# Experiment & Likelihood Groups



**Important: Respondents only receive the extra 5 points if they answer both probes through voice**

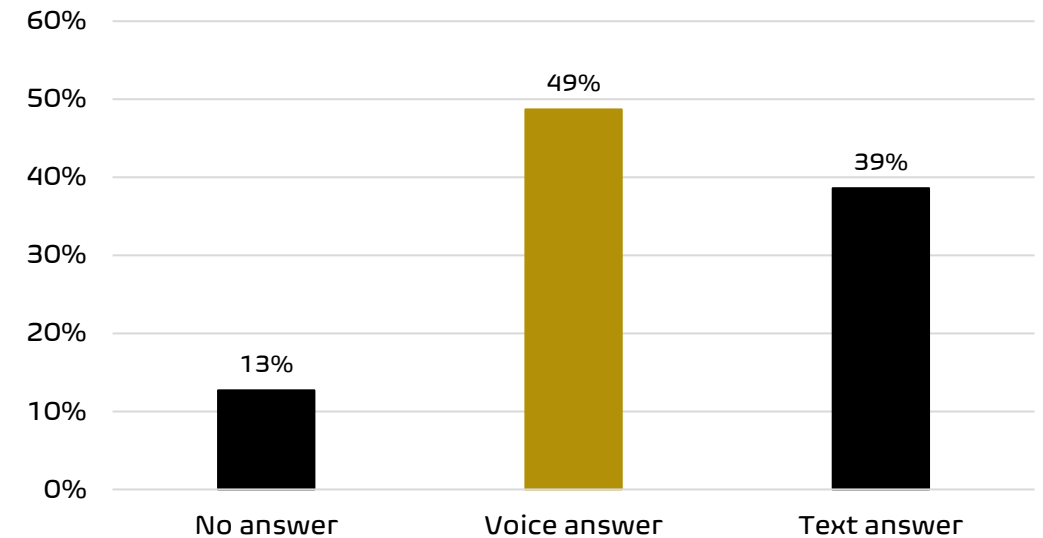
# Results I

Answer behavior (%)  
Probe question 1



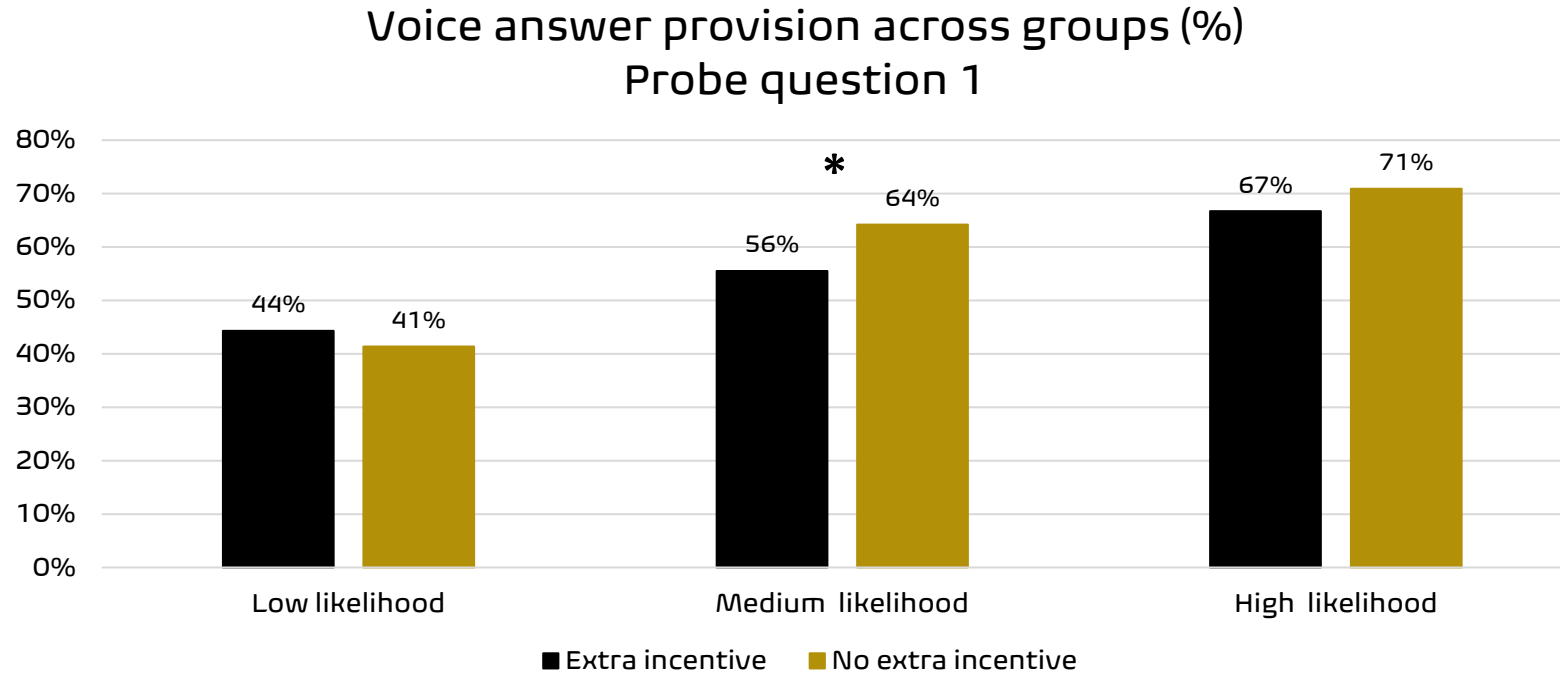
Note. We only consider respondents reaching the first probe question in the analyses.

Answer behavior (%)  
Probe question 2



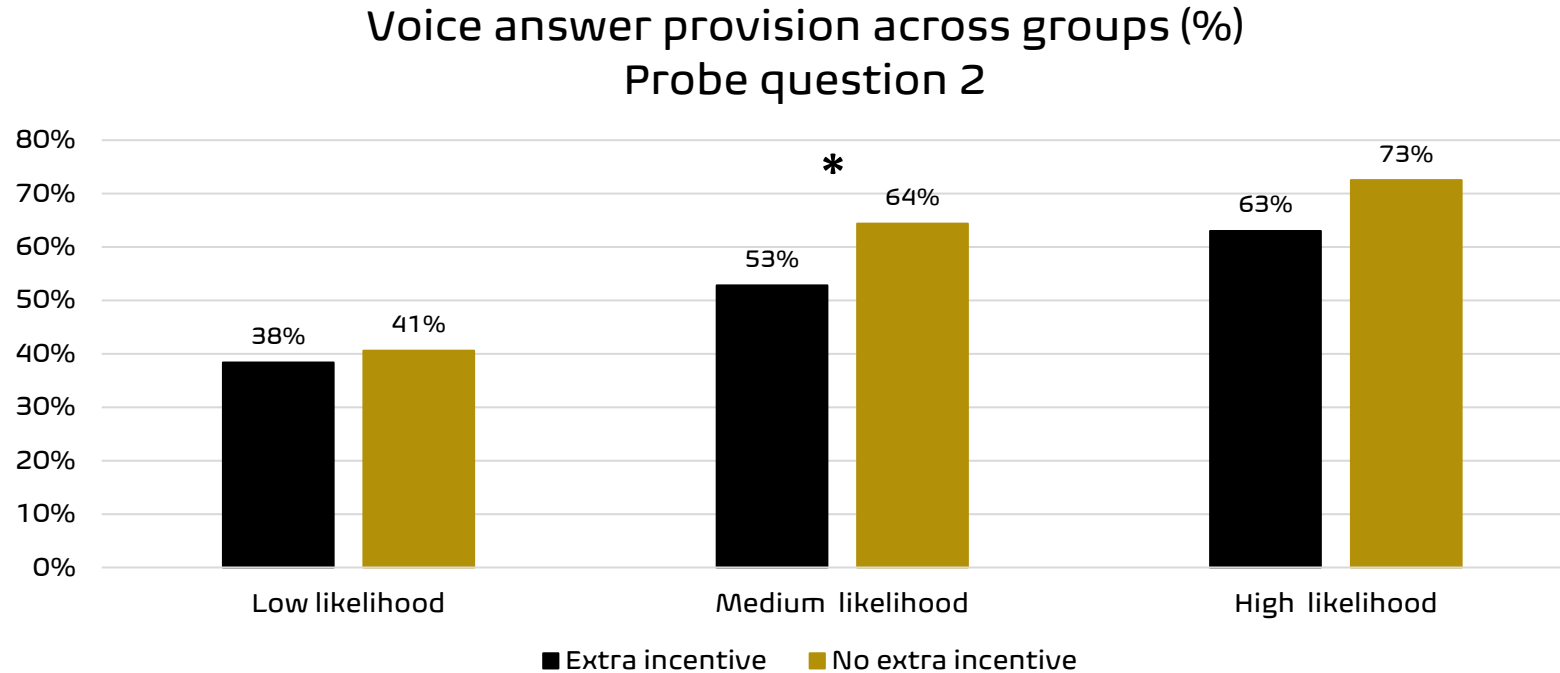
Note. We only consider respondents reaching the second probe question in the analyses.

# Results II



Note. Z-test. \* $p < 0.05$ . We only consider respondents reaching the first probe question in the analyses.

# Results III



Note. Z-test. \* $p < 0.05$ . We only consider respondents reaching the second probe question in the analyses.

# Discussion and Conclusion

- We face an imbalance across experimental/likelihood groups
  - *The low and high likelihood groups are under-represented*
- Although not reported, breakoff and item-nonresponse are about 20%
- We did not find support for our pre-registered hypotheses
  - *Extra incentives do not appear to increase voice answer provision – Medium likelihood group*
- Data quality beyond item-nonresponse
  - *Testing our pre-registered hypotheses on text-based quality metrics (e.g., length and topics)*
  - *Testing transcription quality of Automatic Speech Recognition (ASR) systems for voice answers*
- Take home message: It seems that extra incentives do not increase voice answer provision and missing data remains a concern

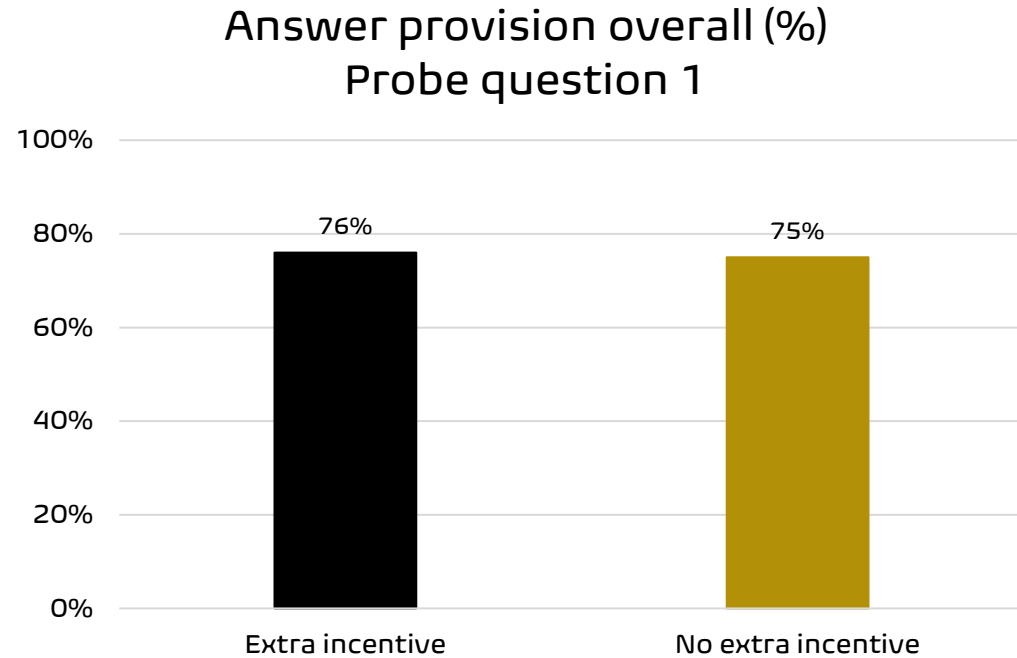
# Many Thanks for your Attention!

- As the paper is under review in PlosOne, we have some questions for discussion:
  - *Why did the extra incentive backfire?*
  - *What are theoretical approaches that can we learn from?*
  - *How can we encourage respondents to use voice (or can't we)?*
  - *Do you have any ideas on further variables for defining likelihood groups?*

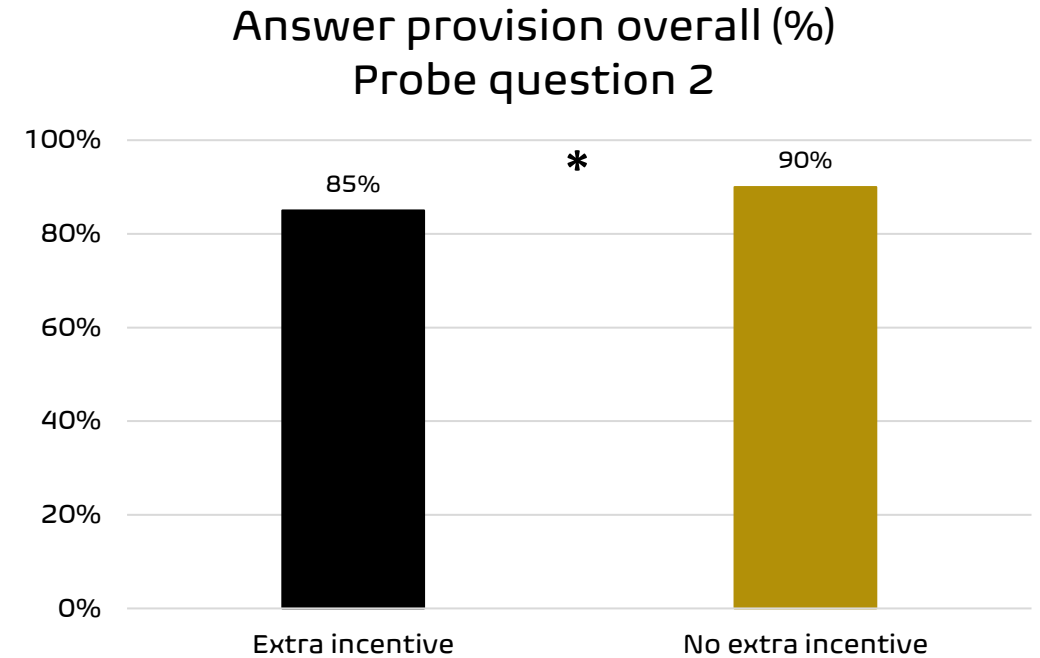
# Literature

- Revilla, M., & Couper, M. P. (2021). Improving the use of voice recording in a smartphone survey. *Social Science Computer Review*, 39, 1159-1178.
- Revilla, M., Couper, M. P. (2023). Combining dictation and/or voice recordings with text to answer narrative open-ended survey questions. Paper presented at the Conference of the European Survey Research Association (ESRA) in Milan (Italy): <https://www.upf.edu/documents/244683118/246905697/ESRA-Revilla-Couper-Voice.pdf/9f24fc51-b23b-a996-c7fb-54af1f872688?t=1702657915024>
- Revilla, M., Iglesias, P., Ochoa, C., & Antón, D. (2022). WebdataVoice: a tool for dictation or recording of voice answers in the frame of web surveys. Open Science Framework (OSF): <http://doi.org/10.17605/OSF.IO/B2WYZ>

# Appendix: Results I



Note. Z-test. We only consider respondents reaching the first probe question in the analyses.



Note. Z-test. \* $p < 0.05$ . We only consider respondents reaching the second probe question in the analyses.

# Appendix: Results II

| Likelihood                  | Low  |      |      | Medium      |             |                | High |      |      | Total |      |         |
|-----------------------------|------|------|------|-------------|-------------|----------------|------|------|------|-------|------|---------|
| Extra incentive             | Yes  | No   | Diff | <i>Yes</i>  | <i>No</i>   | <i>Diff</i>    | Yes  | No   | Diff | Yes   | No   | Diff    |
| OQ1                         |      |      |      |             |             |                |      |      |      |       |      |         |
| ... Voice answer rate (%)   | 44.3 | 41.4 | 2.9  | <i>55.5</i> | <i>64.2</i> | <i>-8.8**</i>  | 66.7 | 70.9 | -4.3 | 53.9  | 60.7 | -6.8**  |
| ... Breakoff (%)            | 15.0 | 20.6 | -5.6 | <i>11.5</i> | <i>18.8</i> | <i>-7.3***</i> | 11.5 | 11.8 | -0.3 | 12.3  | 18.4 | -6.2*** |
| ... Item nonresponse (%)    | 10.9 | 7.8  | 3.2  | <i>11.9</i> | <i>5.6</i>  | <i>6.3***</i>  | 12.6 | 10.0 | 2.6  | 11.7  | 6.4  | 5.3***  |
| ... Overall answer rate (%) | 74.1 | 71.7 | 2.4  | <i>76.6</i> | <i>75.7</i> | <i>1.0</i>     | 75.9 | 78.2 | -2.3 | 76.0  | 75.1 | 0.9     |

# Appendix: Results III

| Likelihood                  | Low  |      |      | Medium      |             |                  | High |      |        | Total |      |                  |
|-----------------------------|------|------|------|-------------|-------------|------------------|------|------|--------|-------|------|------------------|
|                             | Yes  | No   | Diff | <i>Yes</i>  | <i>No</i>   | <i>Diff</i>      | Yes  | No   | Diff   | Yes   | No   | Diff             |
| Extra incentive             |      |      |      |             |             |                  |      |      |        |       |      |                  |
| OQ2                         |      |      |      |             |             |                  |      |      |        |       |      |                  |
| ... Voice answer rate (%)   | 38.4 | 40.6 | -2.2 | <i>52.8</i> | <i>64.4</i> | <i>- 11.6***</i> | 63.0 | 72.5 | -9.5   | 50.6  | 61.0 | <i>- 10.5***</i> |
| ... n                       | 159  | 138  |      | <i>566</i>  | <i>562</i>  |                  | 54   | 80   |        | 779   | 780  |                  |
| ... Breakoff (%)            | 2.7  | 3.1  | -0.4 | <i>2.0</i>  | <i>2.3</i>  | <i>-0.3</i>      | 0.0  | 1.1  | -1.1   | 2.0   | 2.3  | -0.3             |
| ... Item nonresponse (%)    | 13.2 | 11.7 | 1.5  | <i>12.0</i> | <i>7.0</i>  | <i>5.1**</i>     | 21.7 | 9.0  | 12.8*  | 13.0  | 8.1  | 5.0**            |
| ... Overall answer rate (%) | 84.1 | 85.2 | -1.1 | <i>86.0</i> | <i>90.8</i> | <i>-4.8**</i>    | 78.3 | 89.9 | -11.6* | 85.0  | 89.7 | -4.6**           |