Willingness to participate in geolocation-based research.

March 2022

CARLOS OCHOA | RECSM - UPF

Introduction: the problem

Memory error

Surveys, a fundamental tool of empirical research in social sciences...

[5]

... but suffer **measurement and representation errors**[1] → wrong conclusions + implementation of non-optimal policies.

Memory error → major source of error in social science data[2].

Definition: difficulties to recall data related to events of interest[3] for researchers, also motivations and feelings [4][5].

About how we remember

Major classes of memory problems[2]

1. Non-encoding

We may never form a representation of an event in our memory

2. Post-encoding errors

Errors introduced after the original encoding.

3. Retrieval failures

We cannot remember the information that is there.

4. Reconstruction errors

*We fill in missing details based on our general knowledge.*

Factors increasing the chances of suffering memory errors:

+ Many events of the same category (e.g., supermarket visits)

+ Low distinctiveness

+ Low emotional impact

+ Short duration

+ Non-rehearsal (time spent thinking or talking about the event).

+TIME!

In-the-moment surveys

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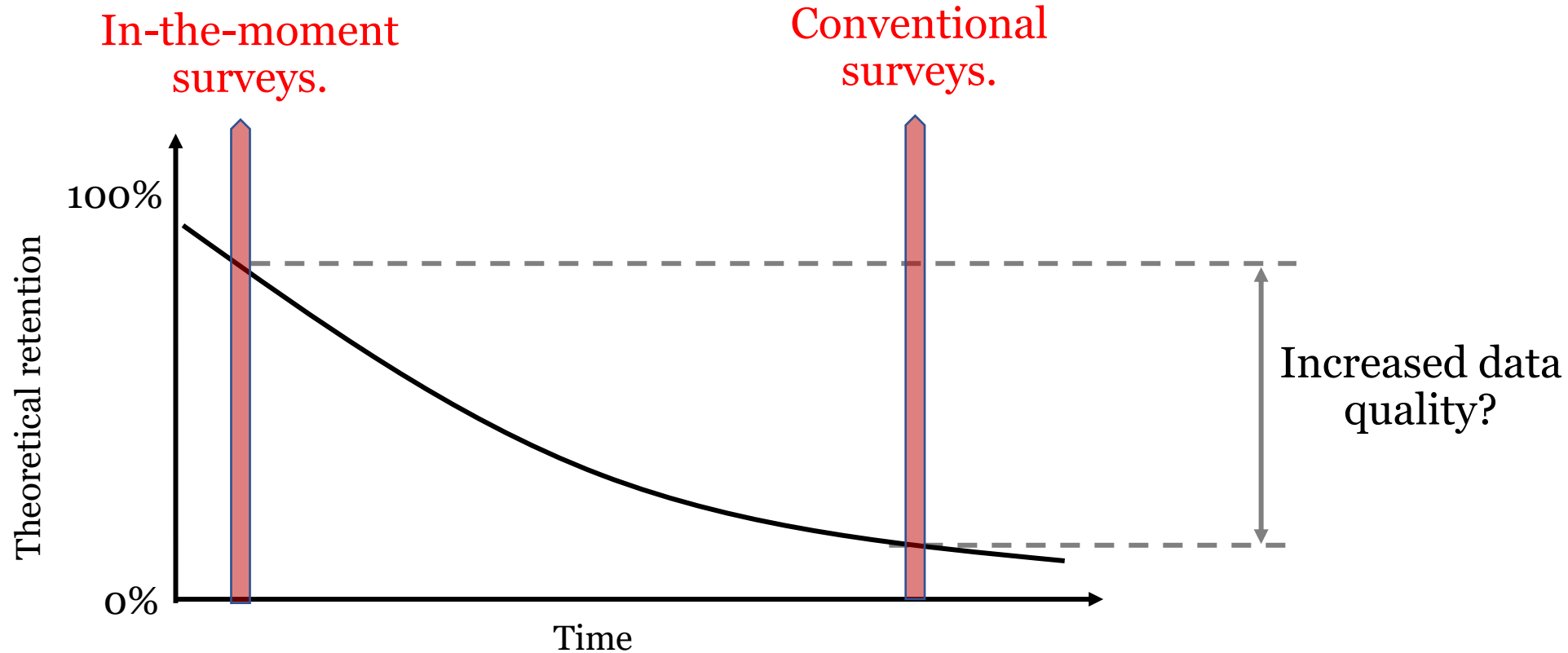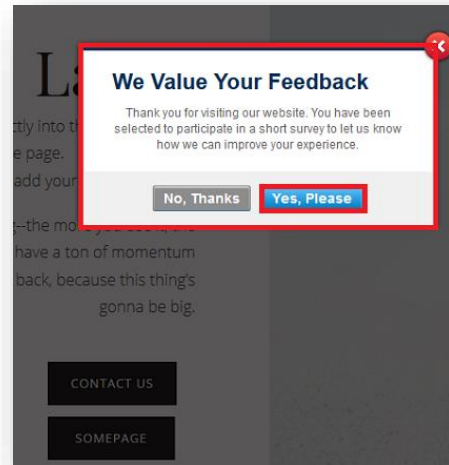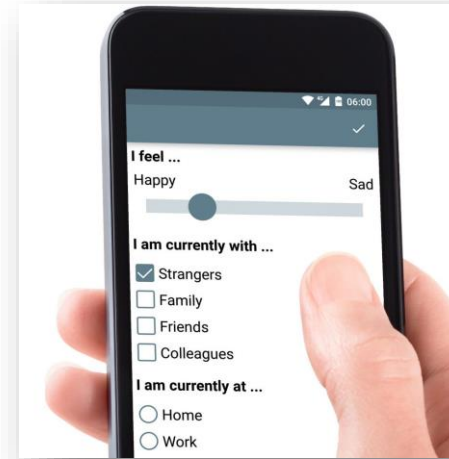
Existing in-the-moment surveys

In-the-moment surveys are used nowadays (and were used in the past), but only in very specific environments (proprietary databases, no control on the sample., one-shot...)

Satisfaction surveys in public transportation.

Online satisfaction surveys.

Experience Sampling Method

Coincidental surveys: "are you listening to the radio?" instead of "did you listen to radio last week?"

New type of surveys: ppt-in online panel + passive + in-the-moment

web data opp

To overcome existing limitations of conventional surveys, I propose **a new type of in-the-moment surveys**.
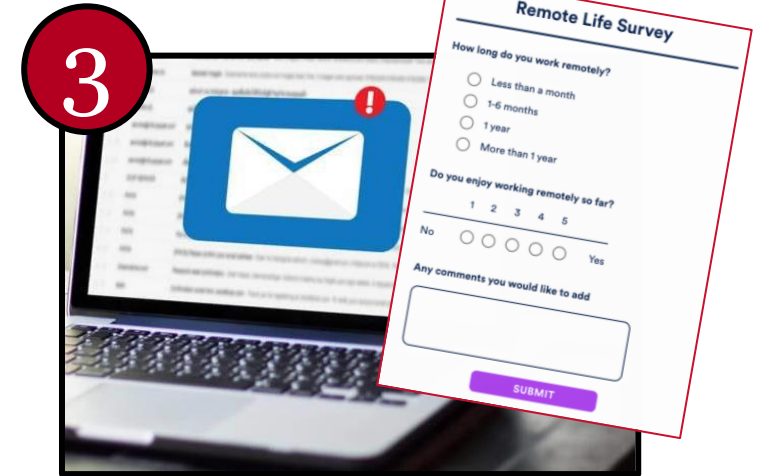
1
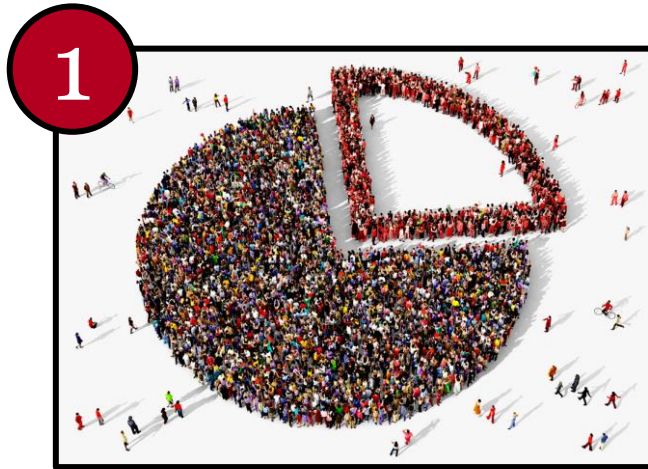
Opt-in online panels
Communities of people that voluntarily participate in research activities in exchange of reward.

2
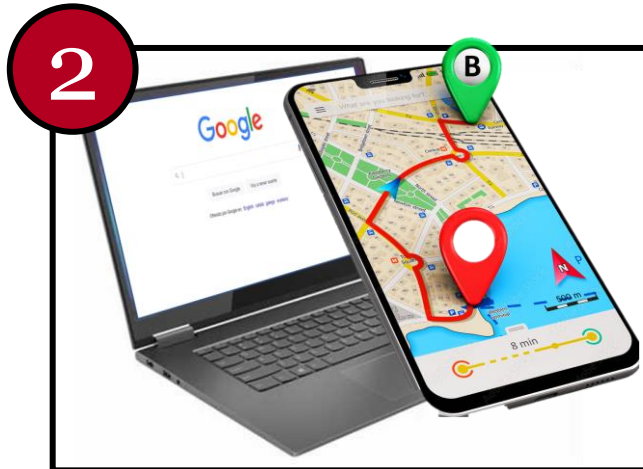
Passive measurement
- Metered data -> online events
- Geolocation data -> offline events

3

In-the-moment survey
When an event of interest is detected (e.g., visiting a political party Facebook page) a survey is sent.

Potential use cases

web data opp

Examples of potential uses of these new in-the-moment surveys triggered by metered data:

FAKE NEWS

Nyhan and Reifler (2018)[6] used **meter data to research consumption of fake news**: do Trump's supporters read more fake news? Surveys used only to profile participants.

In-the-moment surveys answer:

- "Do you give credibility to this news?"
- "Read this fact-checking information, do you still give credibility to...?
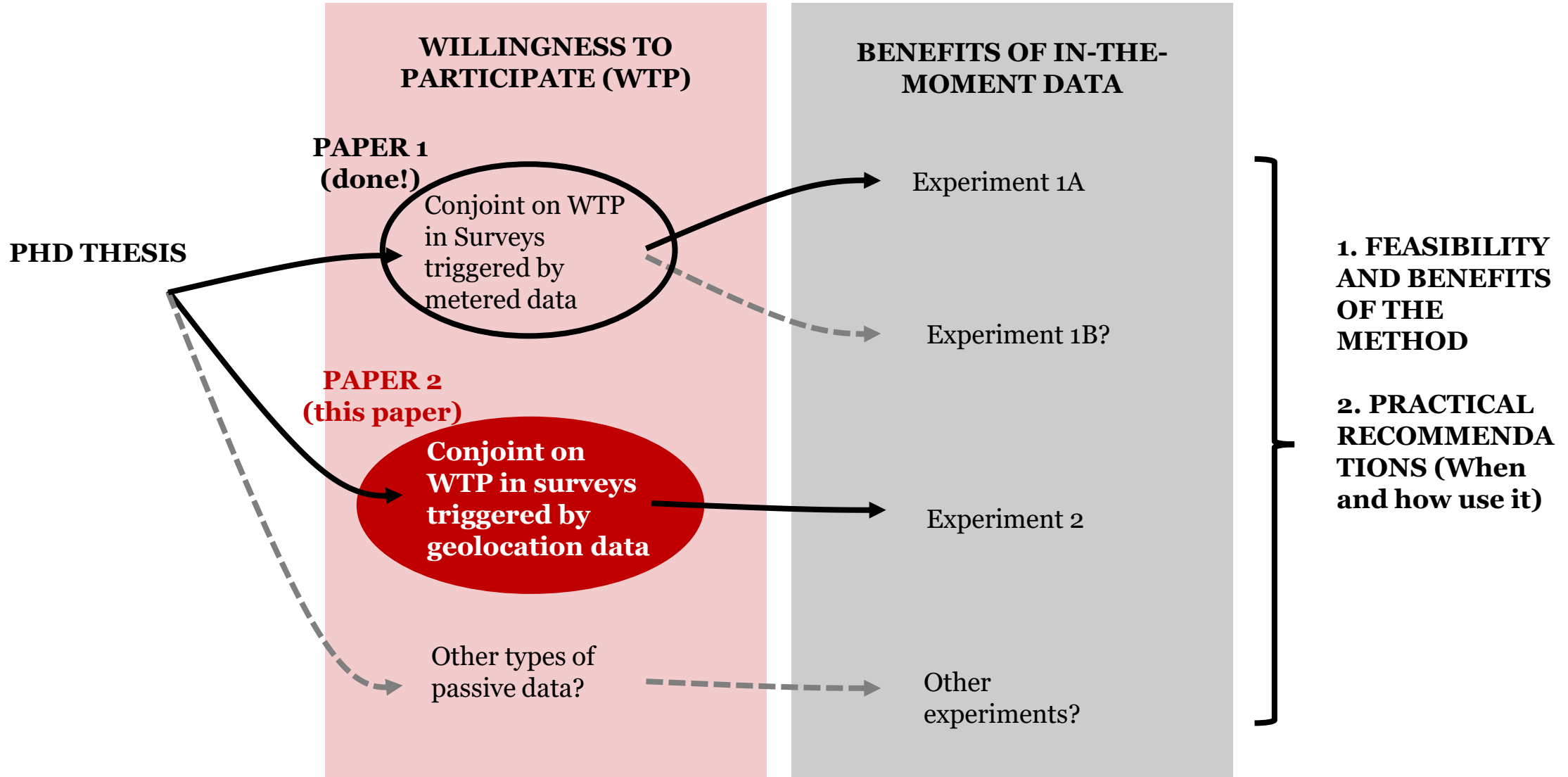
TRAVEL RESEARCH

Detecting when someone has travelled using geolocation data + in-the-moment survey to ask:

- The purpose of the travel.

- Satisfaction with the mode of transport.

- Confirm whether a particular location was visited.

This paper
Willingness to participate in geolocation-based research

Research plan

LITERATURE ABOUT GEOLOCATION

- Willingness to share geolocation data: **20%-40%.**
- Studied factors:
  - Offline samples vs. online panels
  - One-time capture vs. continuous sharing.
  - Willingness to participate vs. actual participation
  - Others: country, specific online panel…

HOWEVER

- Little research about the **effect of the conditions offered to individuals to participate** (incentive, duration of the project).

LITERATURE ABOUT <u>IN-THE-MOMENT</u> SURVEYS

- 1ˢᵗ paper of this PhD: "willingness to participate in in-the-moment surveys triggered by online behaviors of <u>metered</u> panelists".

- Four survey attributes studied:
  1. Survey length
  2. Invitation lifetime (maximum time allowed to participate)
  3. Incentivization level (compared to a conventional survey)
  4. Triggering activity (that causes to be invited to take a survey)

MAIN FINDINGS

- High willingness to participate (**68.5%** to **94.7%).**
- Preference for longer surveys and longer times to participate.
- The tracked activity that triggers the survey plays a minor role.
- Survey length + incentive level = **75.9% of the importance**.
- Few differences among panelists.

RESEARCH QUESTIONS

RQ1 – What are the levels of **willingness to participate** in geolocation-based research:

    (a) share geolocation data

    (b) in-the-moment surveys triggered by geolocation data.

RQ2 – How the **attributes** of geolocation-based research influence the willingness to participate in such surveys?

RQ3 – Are there **significant differences** among panelists?

RQ4 –Main reasons for deciding whether or not to participate stated by the panelists?

ABOUT THE ATTRIBUTES

web
data
opp

We study the effect of 6 attributes, 2-6 levels per attribute.

Research activity:	**Project duration:**	**Invitation lifetime*:**	**Geolocation incentive:**	**Survey incentive level:**	**Length of the interview:**
Sharing	1 week	15 min	1 point/week	X 1 (normal)	1 min
geolocation	1 month	30 min	2 points/week	X 1.5	5 min
vs.	3 month	1 h	3 points/week	X 2	10 min
In-the-moment	6 month	2 h	4 points/week	X 3	15 min
surveys triggered	1 year	3 h	6 points/week	x 4	20 min
by geolocation	Indefinite	6 h	8 points/week		
		12 h			

* =maximum time to participate

METHODS

Choice Based Conjoint analysis:

- A method to assess the influence of each attribute by the analysis of choices.

- 10 questions per participant: 2 proposals + "I would not participate".

- Orthogonal design (minimum correlation between attribute-levels)

- Multinomial model + Bayesian analysis using simulation (MCMC*).

- "Utilities" (coefficients) used to estimate **importance** of attributes and **willingness to participate** in each scenario, for each participant.

*_=Markov Chain Monte Carlo_

DATA

- Data collection: 21st of February – 7th of March 2022.

- Netquest opt-in online panel in Spain.

- **1,016 valid surveys** (2,306 invited, 1,847 started the survey, 461 discarded due to quotas and filters)

- Survey length: mean = 8.8 min.

- Quotas on age(3)+gender(2) and education(3), representative of the Spanish online population.

- **<u>27% of the participants have installed a meter</u>** (already sharing online behaviors).

Preliminary results
(work in progress)

INFLUENCE OF EACH ATTRIBUTE-LEVEL

**web
data
*opp***

Average utilities of the multinomial model
(larger utilities = higher preference)
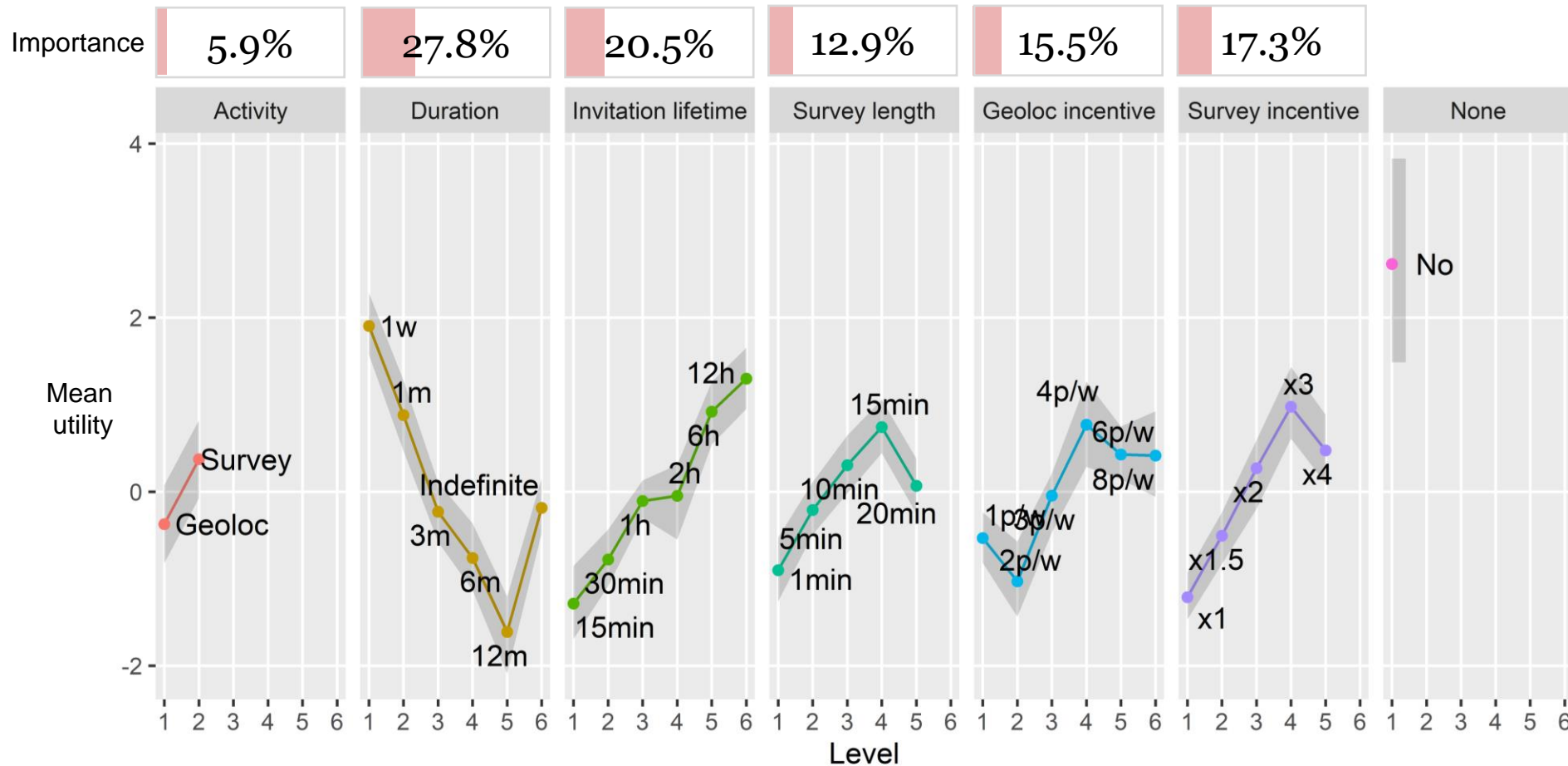
Preference for:

1. In-the-moment surveys over just sharing geoloc

2. Shorter project durations.

3. Larger invitation lifetimes.

4. Larger survey lengths up to 15 min.

5. Larger incentives, (with some inconsistencies).

WILLINGNESS TO PARTICIPATE

Moderate willingness to participate

Scenario	Mean	90% credible interval
Best	79.3%	75.1% ⇔ 83.0%
Average	45.4%	44.5% ⇔ 46.3%
Worst	24.8%	19.7% ⇔ 28.9%

Best scenario:

Survey / duration: 1 week / invitation lifetime: 12h / survey length: 15 min / 4 points per week / x3 survey incentive

Worst scenario:

Geoloc / duration: 1 year / invitation lifetime: 15min / survey length: 1 min / 2 points per week / x1 survey incentive
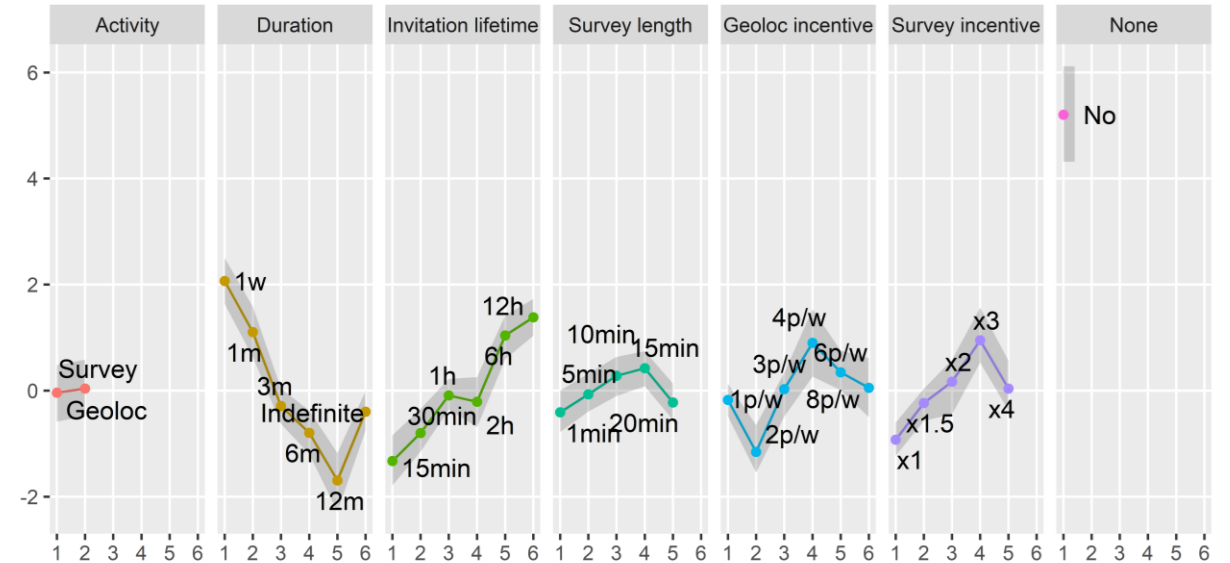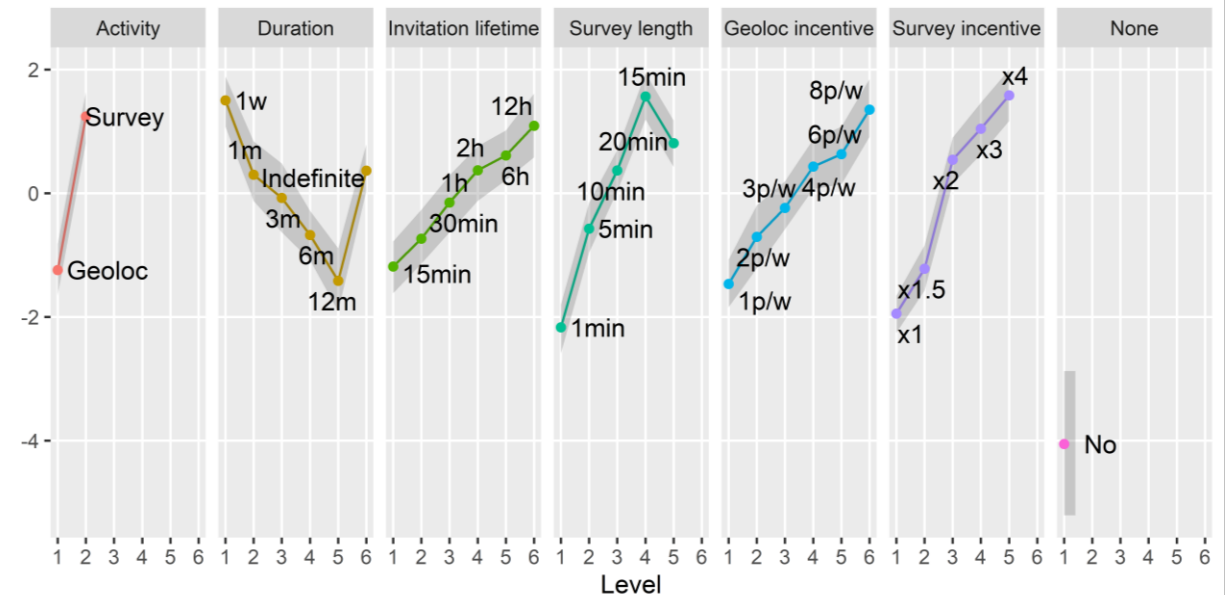
PRELIMINARY CONCLUSIONS

In-the-moment surveys triggered by geolocation data...

1. ... are feasible in terms of willingness to participate.

2. ... are preferable to just sharing geolocation data for an unspecific purpose.

To ensure high levels of participation:

3. Short project durations with reasonable invitation lifetimes.

4. Up to 15 min survey duration

5. Incentive is also key

Developing geolocation-based research on "panelists already sharing online behaviors" may be <u>effective</u> and would allow to <u>combine offline and online research</u>.

References

1. Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., Tourangeau, R., (2009). Survey Methodology, 2nd Edition, Wiley series in survey methodology. Wiley

2. Tourangeau, R. (1999). Remembering What Happened: Memory Errors and Survey Reports. Chapter from the book "The Science of Self-report". Psychology Press.

3. Gray, P. G. (1955) The Memory Factor in Social Surveys, Journal of the American Statistical Association, 50:270, 344-363, DOI: 10.1080/01621459.1955.10501269

4. Walker, W. Richard; Skowronski, John J. (November 2009). The Fading affect bias: But what the hell is it for? (PDF). Applied Cognitive Psychology. 23 (8): 1122–1136. doi:10.1002/acp.1614.

5. Kanheman, D. (2011). Thinking, Fast and Slow. Farrar, Straus and Giroux.

6. Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. Nature human behaviour, 4(5), 472–480. https://doi.org/10.1038/s41562-020-0833-x

9. Sun, B., Morwitz, V.G. (2010). Stated intentions and purchase behavior: A unified model, International Journal of Research in Marketing, Volume 27, Issue 4, Pages 356-366, ISSN 0167-8116, https://doi.org/10.1016/j.ijresmar.2010.06.001

10. Klein, S. B., Babey, S. H., & Sherman, J. W. (1997). The functional independence of trait and behavioral self-knowledge: Methodological considerations and new empirical findings. Social Cognition, 15, 183–203.

Thanks!

Questions?

CARLOS OCHOA | RECSM - UPF

Carlos.ochoa@upf.edu

https://www.upf.edu/web/webdataopp