

An Offer of Control Over Data Shared with Researchers to Reduce Nonparticipation in a Passive Data Collection

Alexandra M. Brown Breslin

Joint Program in Survey Methodology

MASS Workshop

March 2024

Background

Social science research is shifting towards digital research methods, allowing for direct observation of participant behaviors and environment (through passive data collection) in place of self-reported information.

Benefits and Opportunities Offered through Passive Data Collection

- Data may be impossible to collect on a self-report survey
- Less measurement error compared to self-reports
 - E.g., internet usage in Revilla et al. (2017)
- Less burden on respondent compared to self-reporting
- Passive data can be linked to self-report data

BUT concerns over privacy is a barrier to participation (Keusch et al. 2019; Revilla 2019; Chin et al. 2012)

- Participants do not feel they have much control over what data is ultimately shared

Nonparticipation = refusal to consent or take required actions to share data after a sample member is invited to the study

Control = participants decide what data to share with researchers

Timing of control:

- Pre-emptive: control over data is offered before data is collected by temporarily turning off data collection
 - (+) Keusch et al. (2019): offering sample members the ability to temporarily turn off data collection increased WTP in a study that sample members knew to be hypothetical.
 - NOTE: Participants already have this control, so why would this be attractive?
- Reactive: control over data is offered after data is collected by allowing participants to edit data they do not want to share
 - (-) Struminskaya et al. (2020b): offering sample members control to edit after the data collection did not increase WTP in an app based study that sample members knew to be hypothetical.
 - (+) Struminskaya et al. (2021): offering control to edit after the data collection did increase the actual sharing of geolocation.
- Both?

Burden = number of interactions with the data collection tool + comfort with data collection tool:

- Pre-emptive: participant is asked to consider their actions in real time and then interact with the tool to turn off the data collection.
 - But, what happens if the respondent forgets to turn off the data collection for a certain activity?
- Reactive: participant goes about their normal activities and only has a single interaction point- to review and edit the collected data.
 - But, may be concern that edited information is still visible to the researcher?

Measure of Burden: Ability to complete tasks, following Jackle et al. (2017) and Read (2019)
How comfortable are you with installing new apps (e.g. from iTunes, Google Play Store) on a smartphone? (a) very comfortable (b) comfortable (c) not comfortable

	% (SE)
Not comfortable	8.1% (1.1)
Comfortable	27.9% (1.9)
Very comfortable	64.0% (2.0)

Measure of Privacy: Concern about sharing personal information on the Internet (IPSOS Household Survey)

How concerned are you about providing personal information over the internet? (a) not at all concerned (b) slightly concerned (c) somewhat concerned (d) very concerned

Very few of the respondents feel “not at all concerned” while over half (62%) have higher levels of concern, indicating that they are “somewhat” or “very” concerned with sharing their personal information over the internet.

	% (SE)
Not at all concerned	4.5 (0.9)
Slightly concerned	33.9 (2.1)
Somewhat concerned	36.9 (2.1)
Very concerned	24.7 (1.9)

Research Questions

- Does an offer of control (of any form) over the data collected during a passive data collection increase participation in a mobile app based passive data collection?
- Does offering sample members both pre-emptive and reactive control over their collected data increase participation in a mobile app based passive data collection compared to offering sample members only pre-emptive control, reactive control, and no control?
- Do privacy and burden concerns impact which offer of control is preferred?

Methods

Web Survey with Invitation to Passive Data Collection

- Fully crossed factorial design experiment with 4 levels of control
 - Control: no control offered, preemptive control, reactive control, both
 - 570 respondents: ~140 respondents assigned to each invitation
- Recruitment method: GfK KnowledgePanel
 - Deception study with debrief for all respondents
 - 79 respondents (12%) withdrew their data, evenly across experiments
- Invitation to download a mobile app geotracking study
 - Designed to match the study by Struminskaya et al. (2021)
 - Last question in short survey (survey programmed in Qualtrics)
 - Each respondent received one survey – varying control

Invitation

[PAGE ONE]

We would like to invite you to participate in an additional app based study that focuses on **where Americans use their smartphones**.

To participate in the study, please download the app provided on the next page and keep it on your phone for one week. You will receive \$20 for your participation, with \$10 provided for downloading the app and \$10 provided at the end of the week long data collection.

The app will record your location on a map once every hour, as displayed here. (Click "next page" to read more.)



[PAGE TWO]

Important additional study information:

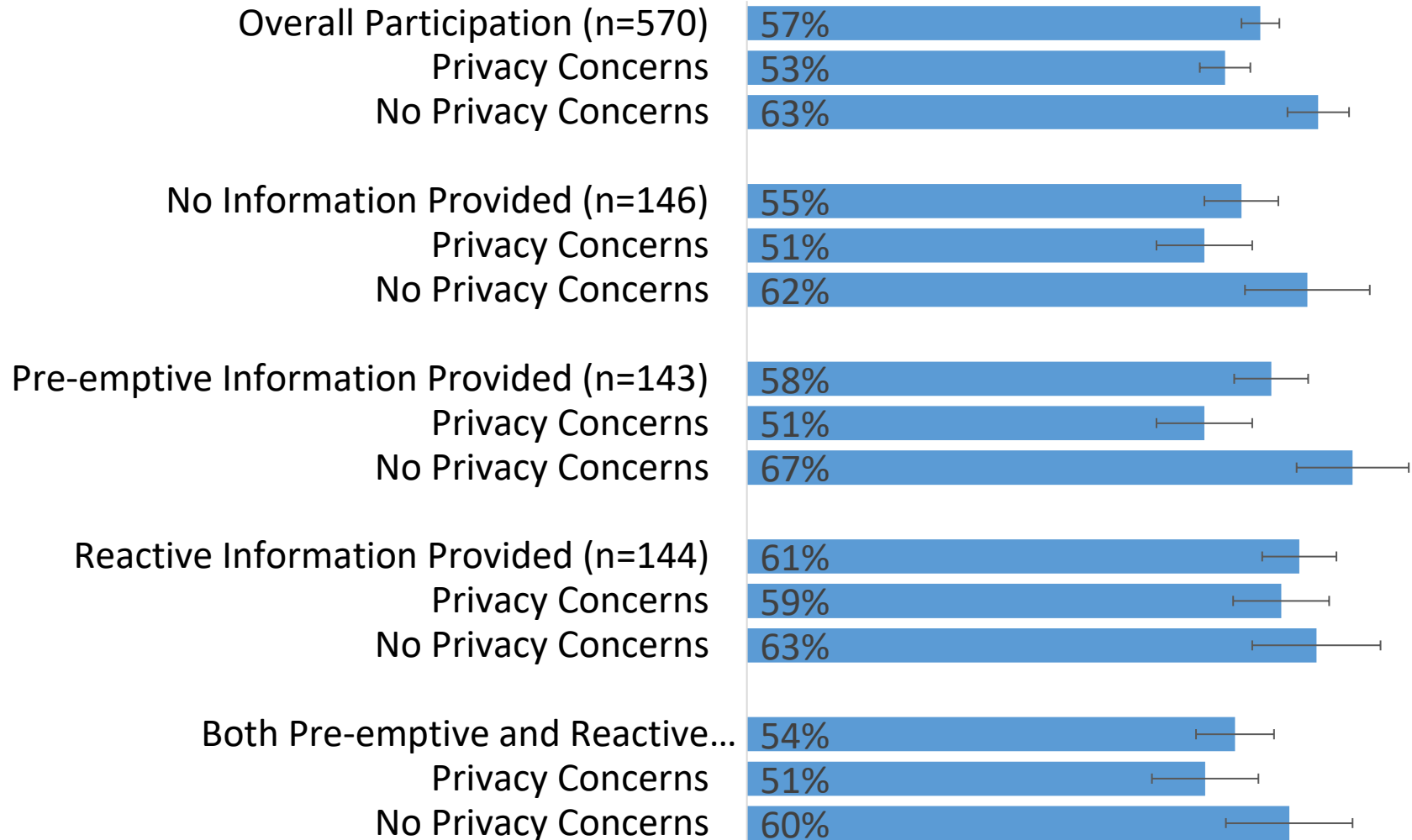
- Your data will remain anonymous and will never be used in a way that would personally identify you.
- *[Pre-emptive Control Offer]* The image collection can be turned off if you feel uncomfortable with sharing your location, as shown here.



- *[Reactive Control Offer]* At the end of the week, you will be provided with the images to review. You may edit the series of images before they are sent to our researchers, as shown here.



Results



Dependent variable: Participation	Equation 1		Equation 2	
	Coeff	p	Coeff	p
Level of Control (No Control is Comparison)				
Pre-emptive Control	0.160	0.501	1.009	0.295
Reactive Control	0.287	0.228	-0.318	0.750
Both Pre-emptive and Reactive Control	-0.004	0.987	-0.288	0.747
Respondent Characteristics				
Privacy			-0.304	0.383
Burden			0.951	0.138
Interactions (Both Pre-emptive and Reactive Control)				
No Information * Privacy			0.050	0.920
Pre-emptive Control * Privacy			-0.367	0.456
Reactive Control * Privacy			0.170	0.734
No Information * Burden			-0.302	0.724
Pre-emptive Control * Burden			-1.011	0.289
Reactive Control * Burden			0.235	0.810

Discussion

- Unexpectedly high willingness to participate – 57% overall
- Sample was recruited from an active online based probability panel – smartphone and survey savvy
- Study timing - two years after the last study
- Respondents awareness - still needed to take the additional step of downloading an app to their phone...

Why not participate?

<i>n=195</i>	n	%	Example
Study request is “too much”	83	43%	“I am nervous about being tracked in this way”
Burden	37	19%	“I don’t have room on my phone.” “I don’t have time for that.”
Miscellaneous	35	18%	“Just no.” “I am not interested.”
Privacy	16	8%	“Privacy concerns.”
Trust	15	8%	“Don’t trust this.”
Too low of incentive	4	2%	“Not enough monetary incentive. I would not accept anything below \$100.”
Negative previous experience	3	2%	“I downloaded an app once for vpn and it kept me from logging in to my job and caused issues”

Thank you!

Contact: abrown53@umd.edu

References

Chin, E., Felt, A. P., Sekar, V., & Wagner, D. (2012, July). Measuring user confidence in smartphone security and privacy. *In Proceedings of the eighth symposium on usable privacy and security* (pp. 1-16).

Jäckle, A, Burton, J, Couper, M. P, & Lessof, C. (2017). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases. Institute for Social and Economic Research, University of Essex: Understanding Society Working Paper Series, 9.

Keusch, F, Struminskaya, B, Antoun, C, Couper, M. P, & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public opinion quarterly*, 83(S1), 210-235.

Read, B. (2019, April). Respondent burden in a mobile app: Evidence from a shopping receipt scanning study. *In Survey Research Methods* (Vol. 13, No. 1, pp. 45-71). European Survey Research Association.

Revilla, M, Ochoa, C, & Loewe, G. (2017). Using passive data from a meter to complement survey data in order to study online behavior. *Social Science Computer Review*, 35(4), 521-536.

Revilla, M, Couper, M. P, & Ochoa, C. (2019). Willingness of online panelists to perform additional tasks. *methods, data, analyses*, 13(2), 29.

Struminskaya, B., Toepoel, V., Lugtig, P., Haan, M., Luiten, A., & Schouten, B. (2020). Understanding willingness to share smartphone-sensor data. *Public opinion quarterly*, 84(3), 725-759.

Struminskaya, B., Lugtig, P., Toepoel, V., Schouten, B., Giesen, D., & Dolmans, R. (2021). Sharing Data Collected with Smartphone Sensors: Willingness, Participation, and Nonparticipation Bias. *Public Opinion Quarterly*.

Why not participate?

	Open Response Category	n	%
No Info (n=55)	Study request is “too much”	25	45%
	Burden	10	18%
	Privacy	9	16%
	Trust	3	5%
Pre-emptive (n=50)	Study request is “too much”	18	36%
	Burden	11	22%
	Privacy	6	12%
	Trust	5	10%
Reactive (n=47)	Study request is “too much”	19	40%
	Burden	5	11%
	Privacy	7	15%
	Trust	8	17%
Pre-emptive & Reactive (n=42)	Study request is “too much”	21	50%
	Burden	2	5%
	Privacy	4	10%
	Trust	2	5%