

Mobile device tracking: Explaining discrepancies between survey self-reports and digitally tracked behavior

Claassen¹, Bosch², & Höhne¹

¹ DZHW, Leibniz University Hannover

² Verian Group

6th Mobile Apps and Sensors in Surveys (MASS) Workshop

London (UK) – June 4 to 5, 2025

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

Melanie Revilla is Principal Investigator (PI)



We also thank Netquest (www.netquest.com) for the data collection in their Spanish panel



Introduction

- App use and mobile browsing data are key for social and behavioral researchers
 - *Researchers increasingly utilize passive measurements through tracking applications* (Bach & Wenz, 2020; Müller & Bach, 2023; Wenz et al., 2024)
- Severe discrepancy between survey self-reports and digitally tracked behavior (Parry et al., 2021)
 - *Frequently attributed to errors in self-reports, such as recall error and social desirability bias* (Baghal et al., 2014; Kreuter et al., 2008)
- Errors in passive measurements are often overlooked (Bosch & Revilla, 2022)
 - *Tracking undercoverage of mobile devices* (Bosch et al., 2024)
 - *Sharing of mobile devices* (Revilla et al., 2017)

Research Questions (RQs)

- How prevalent is tracking undercoverage and sharing of mobile devices among participants in a web tracking panel? (RQ1)
- Can we explain the discrepancy between survey self-reports and digitally tracked behavior by investigating ...
 - ... tracking undercoverage of mobile devices? (RQ2)
 - ... sharing of mobile devices? (RQ3)

Method: Study Design

- Data collection (N = 1,191) in the Netquest Panel in Spain in June 2023
 - *Cross quotas on age and gender; additional quota on education*
- Web tracking data for a period of 30 days
 - *Collected on mobile devices and computers (as a benchmark)*
 - *Three variables (or digital behaviors): (1) daily internet usage, (2) social media usage, and (3) video-on-demand (VoD) streaming in hours and minutes*
- Web survey data on digital behavior
 - *Self-reports on (1) daily internet usage, (2) social media usage, and (3) VoD streaming in hours and minutes*
 - *Self-reports on number of devices owned and devices shared with third parties*

Method: Analytical Strategy

- Descriptive statistics on tracking undercoverage and device sharing (RQ1)
- Comparing the discrepancy between survey self-reports and digitally tracked daily internet usage between participants with different types of ...
 - ... *tracking undercoverage (RQ2)*
 - ... *device sharing (RQ3)*
- Two-level linear mixed effects regression with random intercepts (RQ2/3)
 - *Three types of digital behavior nested within participants*
 - *Dependent variable: Discrepancy in minutes*
 - *Independent variables: Tracking undercoverage, device sharing, age, gender, and education*

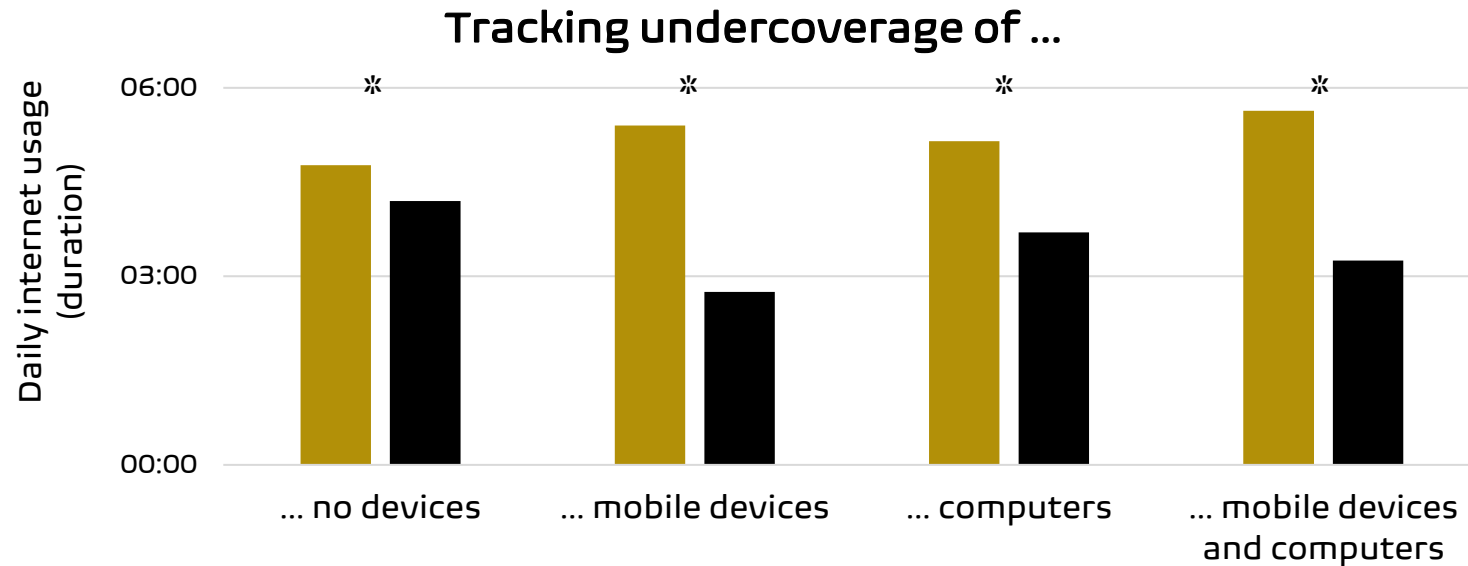
Results: Descriptive Statistics (RQ1)

Table 1. Prevalence of tracking undercoverage and device sharing

	Mobile devices	Computers	Either device
Tracking undercoverage			
Participants who are not fully tracked	42%	48%	55%
Average number of non-tracked devices	2.28	1.66	2.75
Sharing of tracked devices			
Participants who share at least one device	11%	19%	24%
Average number of shared devices	1.03	1.05	1.18

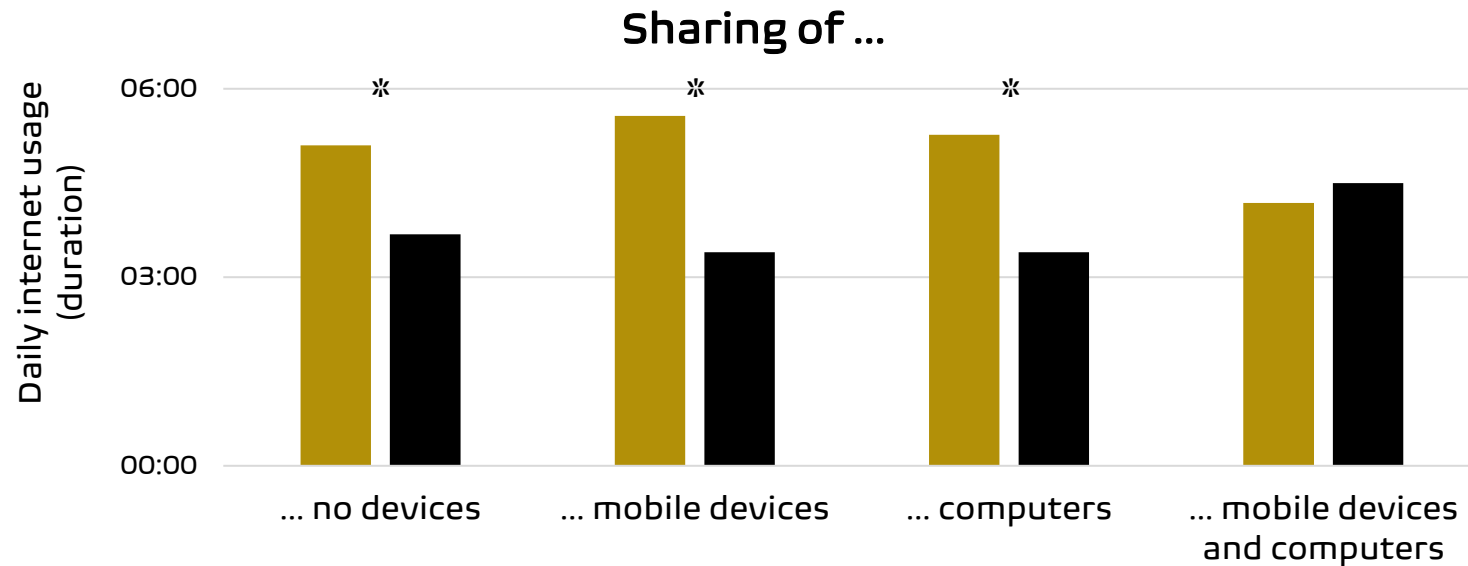
Note. When calculating the average number of non-tracked devices and shared devices, we only considered participants who own are not fully tracked and share at least one device, respectively.

Results: Tracking Undercoverage (RQ2)



Note. * $p < 0.05$ (paired t-Tests). Self-reported (gold) vs. tracked (black) daily internet usage by type of tracking undercoverage.

Results: Device Sharing (RQ3)



Note. * $p < 0.05$ (paired t-Tests). Self-reported (**gold**) vs. tracked (**black**) daily internet usage by type of device sharing.

Results: Mixed Effects Regression (RQ2/3)

Table 2. Two-level linear mixed effects regression with random intercepts

DV: Discrepancy (in minutes)	Model 1		Model 2	
	COEF	SE	COEF	SE
Intercept	<u>84.91</u>	4.52	<u>59.76</u>	7.28

Tracking undercoverage of (reference: no devices) ...				
... mobile devices			<u>54.45</u>	12.02
... computers			14.67	12.69
... mobile devices and computers			<u>55.24</u>	12.24
Sharing of (reference: no devices) ...				
... mobile devices			12.15	17.37
... computers			6.23	13.37
... mobile devices and computers			-40.78	26.53
ICC	0.22		0.21	
N	3,243		3,243	

Note. Coefficients with $p < 0.05$ in bold and underlined. COEF = Coefficient. SE = Standard error. Participants with missing values for any of the independent variables were excluded from the analyses. The results do not change when controlling for age, gender, and education (see Model 3 in the Appendix).

Discussion and Conclusion

- High discrepancy between survey self-reports and digitally tracked behavior
- Tracking undercoverage of mobile devices is common (about 40%)
 - *Discrepancy increases with tracking undercoverage of mobile devices*
 - *But: Discrepancy is not associated with tracking undercoverage of computers*
- Sharing of mobile devices is less common (about 10%)
 - *Computers are shared nearly twice as often*
 - *Sharing of mobile devices is not associated with the discrepancy*
- Take home message: The discrepancy between survey self-reports and digitally tracked behavior can partly be explained by investigating tracking undercoverage of mobile devices

Many thanks for your attention!

claassen@dzhw.eu
@jclaass.bsky.social

Literature

- Bach, R. L., & Wenz, A. (2020). Studying health-related internet and mobile device use using web logs and smartphone records. PLoS ONE, 15(6), Article e0234663. <https://doi.org/10.1371/journal.pone.0234663>
- Baghal, T. A., Belli, R. F., Phillips, A. L., & Ruther, N. (2014). What are you doing now? Activity-level responses and recall failures in the American Time Use Survey. Journal of Survey Statistics and Methodology, 2, 519-537. <https://doi.org/10.1093/jssam/smu020>
- Bosch, O. J., & Revilla, M. (2022). When survey science met web tracking: Presenting an error framework for metered data. Journal of the Royal Statistical Society, 185(S2), S408-S436. <https://doi.org/10.1111/rssa.12956>
- Bosch, O. J., Sturgis, P., Kuha, J., & Revilla, M. (2024). Uncovering digital trace data biases: Tracking undercoverage in web tracking data. Communication Methods and Measures. <https://doi.org/10.1080/19312458.2024.2393165>
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys. Public Opinion Quarterly, 72(5), 847-865. <https://doi.org/10.1093/poq/nfn063>
- Müller, P., & Bach, R. L. (2023). Populist alternative news use and its role for elections: Web-tracking and survey evidence from two campaign periods. News Media & Society, 25(10), 2663-2683. <https://doi.org/10.1177/14614448211032969>
- Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. Nature Human Behavior, 5, 1535-1547. <https://doi.org/10.1038/s41562-021-01117-5>
- 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. <https://doi.org/10.1177/0894439316638457>
- Wenz, A., Keusch, F., & Bach, R. L. (2024). Measuring smartphone use: Survey versus digital behavioral data. Social Science Computer Review. <https://doi.org/10.1177/08944393231224540>

Appendix: Full Regression Models

Table A1. Two-level linear mixed effects regression with random intercepts

	Model 1		Model 2		Model 3	
DV: Discrepancy (in minutes)	COEF	SE	COEF	SE	COEF	SE
Intercept	<u>84.91</u>	4.52	<u>59.76</u>	7.28	<u>89.42</u>	21.57
Tracking undercoverage of (reference: no devices) ...						
... mobile devices			<u>54.45</u>	12.02	<u>54.51</u>	12.04
... computers			14.67	12.69	16.37	12.72
... mobile devices and computers			<u>55.24</u>	12.24	<u>55.69</u>	12.29
Sharing of (reference: no devices) ...						
... mobile devices			12.15	17.37	9.86	17.35
... computers			6.23	13.37	9.18	13.32
... mobile devices and computers			-40.78	26.53	-44.12	26.50
Age (in years)					<u>-0.80</u>	0.34
Female (reference: male)					<u>22.75</u>	9.04
Education (reference: low education)						
Medium education					6.05	11.54
High education					-17.63	11.52
ICC	0.22		0.21		0.21	
N	3,243		3,243		3,243	

Note. Coefficients with $p < 0.05$ in bold and underlined. COEF = Coefficient. SE = Standard error. Participants with missing values for any of the independent variables were excluded from the analyses.