



# Consent to linking social media, webtracking, and survey data

Johannes Breuer, Sebastian Stier,  
Pascal Siegers, Tobias Gummer,  
Arnim Bleier

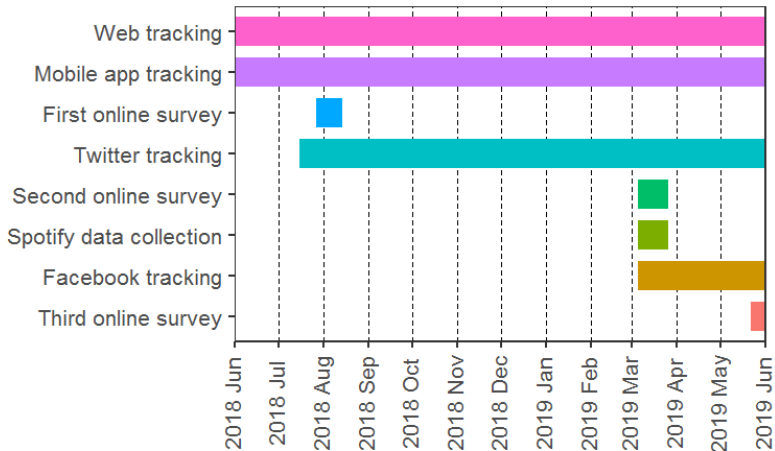
# Background & motivation

- surveys and social media data both have limitations
- combining them holds great promise for research on online behavior (see [Stier et al., 2019](#))
- linking at the individual level requires informed consent

## Research interest

- **Research question:** *Which factors influence consent to the tracking and linking of social media data?*
- replication & extension of Al Baghal, T., Sloan, L., Jessop, C., Williams, M. L., & Burnap, P. (2019). Linking Twitter and survey data: The impact of survey mode and demographics on consent rates across three UK studies. *Social Science Computer Review*. Advance online publication.  
<https://doi.org/10.1177/0894439319828011>

# Project timeline



# Methods

- *respondi* web tracking panel ( $N \sim 2000$  for each month)
  - website visits on domain level
- 2 online surveys among panelists
  - survey 1: 2018-07-27 - 2018-08-14,  $N = 1347$ , 48.11% female, age: 16 to 70 ( $M = 44.66$ ,  $SD = 13.2$ )
  - survey 2: 2019-03-05 - 2019-03-26,  $N = 1240$ , 48.95% female, age: 16 to 68 ( $M = 45.4$ ,  $SD = 12.81$ )
  - topical focus in both surveys: media use & politics

# Social media data

- Twitter
  - continuous collection of Tweets through (public) streaming API
- Spotify
  - web app developed at KU Leuven
    - collects last 50 plays, playlists, preferences
- Facebook
  - browser plugin developed at LMU Munich (see [Haim & Nienierza, 2019](#))
    - for Firefox & Chrome
    - collects public posts (+ some metadata) from users' feed

# Incentives

- additional incentives for social media data
  - Twitter: 5 Euros, prepaid (regardless of consent) vs. postpaid (only if consent = yes)
  - Facebook: 5 Euros
  - Spotify: 2.5 Euros

## Informed consent

- short informed consent in questionnaire
- extended data privacy information on GESIS website (link in questionnaire)
- translated & adapted with permission from Al Baghal et al. (2019)

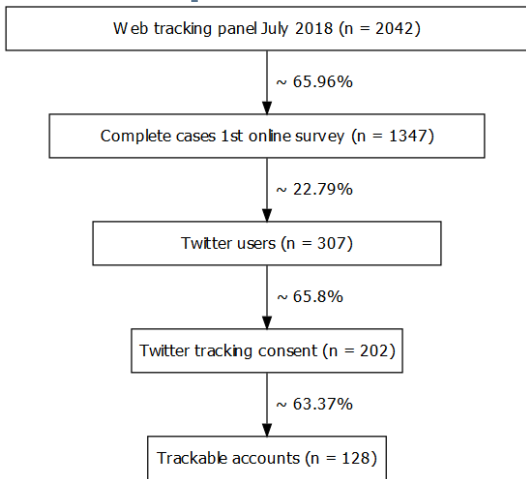
Platform	Visits	Visitors	Avg_Duration
Twitter	11	6	57.09
Facebook	9	7	34.11
Spotify	7	4	19.29

*Note:*

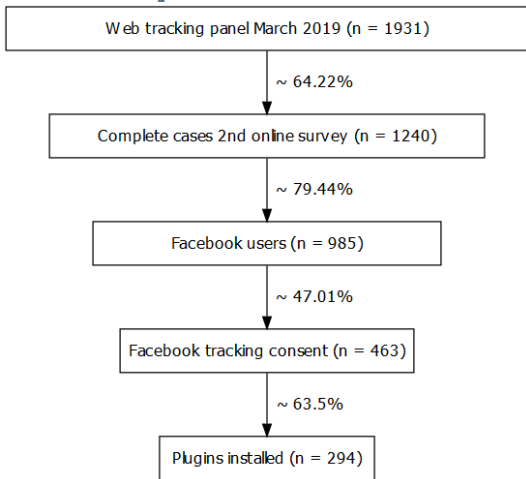
3 participants who visited gesis.org consented to both the Facebook and the Spotify data collection.



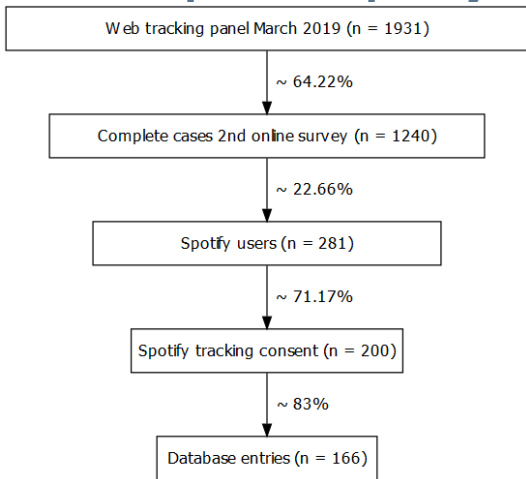
# Participation Twitter



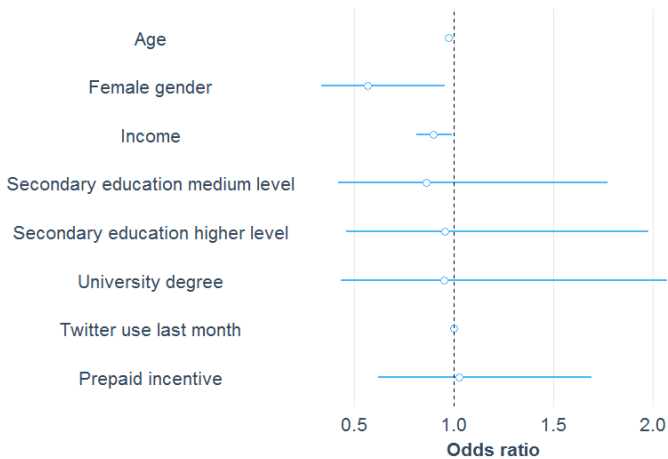
# Participation Facebook



# Participation Spotify



# Twitter tracking consent



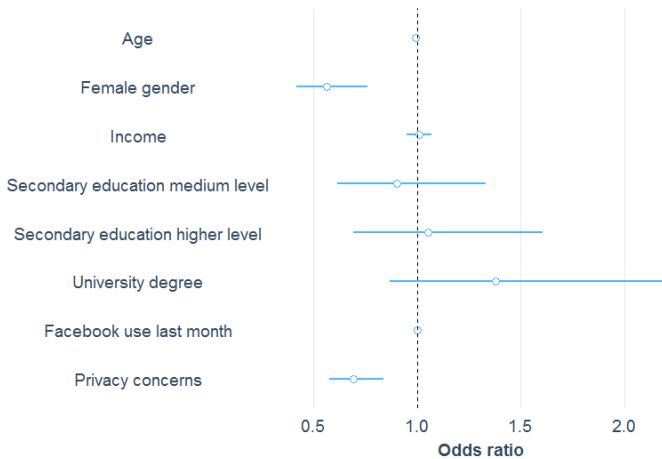
## Reasons for opting out - Twitter

Reason	Percentage
Tweets are private	45.63
Privacy concerns	38.83
Usefulness unclear	24.27
Risk of non-scientific use	16.50
Other	12.62

*Note:*

Multiple selection was possible.

# Facebook tracking consent



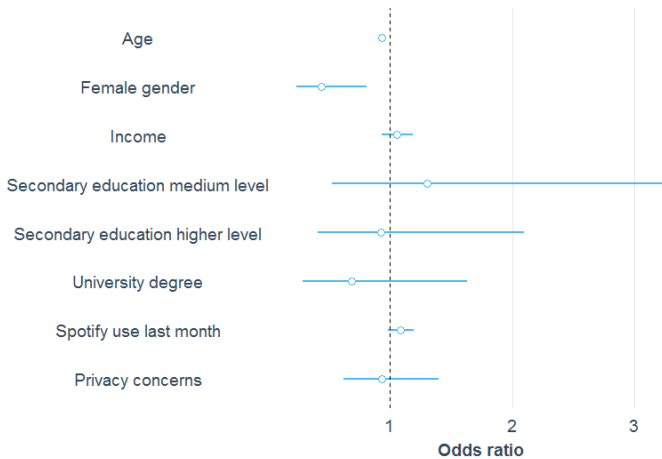
## Reasons for opting out - Facebook

Reason	Percentage
Privacy concerns	41.88
Risk of non-scientific use	29.06
Usefulness unclear	22.81
Other	19.06
Wrong browser	13.75

*Note:*

Multiple selection was possible.

# Spotify consent





## Reasons for opting out - Spotify

Reason	Percentage
Privacy concerns	42.50
Risk of non-scientific use	35.00
Other	31.25
Usefulness unclear	27.50

*Note:*

Multiple selection was possible.

# Conclusion

- results for NatCen Panel from [Al Baghal et al. \(2019\)](#) replicated for gender & age (for Twitter & Spotify)
- privacy concerns as additional (negative) predictor for Facebook data
- detailed privacy information ignored by most participants
  - still: detailed informed consent important to have
- discrepancies between consent and actual use of tracking/data collection tools
  - technical barriers & difficulties are important
- consent rate in our study possibly = upper bound due to nature of the sample & study design