Assessing Online Social Media Recruitment for Reaching and Recruiting Rare Survey Participants

Zaza Zindel & Simon Kühne

Faculty of Sociology
Quantitative Methods of Empirical Social Research
Bielefeld University

ESRA Conference 2021
July, 2nd
Rare Populations

- Often, there is no sampling frame available for rare populations (e.g., gay men)
- Classic sampling approaches for general population surveys are inefficient when applied to rare or hard-to-reach populations
  - Random walks
  - Random digit dialing
- Main problem: Large number of screenouts = High costs

**Alternatives:**

- Give up (survey-based) project
- Non-probability sampling, often online convenience sampling
Social Media Usage

- Facebook and Instagram: Two of the largest social networks worldwide
- Coverage in Germany (ARD & ZDF 2020):
  - 32% of 14+ use Facebook at least once per month
  - 24% of 14+ use Instagram at least once per month
  - 65% of 14-29 use Instagram at least once per week
- Recruiting via ads displayed on the networks (timeline, stories) based on socio-demographics, behavior, and interests for targeted advertising
Advantages & Disadvantages

Advantages

- Costs
- Reach
- Time-efficiency
- Targeting
- Ease of use

Great potential to reach potential participants efficiently!

Disadvantages

- Sampling Error / (Self-) Selection Bias
- (Potential) Coverage Bias

Often low data quality and high potential for biased estimates.

(see Burke-Garcia et al. 2020; Lehdornvirta et al. 2021; Pötzschke Braun 2017)
Project 1 - LGBielefeld

▶ Online-Survey, in which lesbian, gay, bisexual, trans*, queer people, and rainbow families in Germany were recruited

▶ Recruitment via ads on Facebook’s Timeline (Desktop and Smartphone) and Instagram’s Timeline and Stories (Smartphone only)

Recruiting Phase: July 3th - August 2nd, 2019
Examples of the Ads on Facebook & Instagram
Campaign Results - Clicks, Respondents, & Costs

- Costs for ads: 2,710 €
- 784,820 impressions and 335,461 individuals reached
- 43,063 link-clicks (80% through Instagram)
- 12,263 started the questionnaire
- 7,129 complete interviews
  - Only 1.7% reporting a heterosexual orientation
  - 28.4% living with children in household
- Average net cost for a complete interview: 38 Cents
- 98 percent used a smartphone

(see Kühne & Zindel 2020)
## Data Quality - Sample Composition (unweighted)

<table>
<thead>
<tr>
<th>Variable</th>
<th>LGBielefeld</th>
<th>SOEP-Q</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td><strong>Sexual Orientation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Heterosexual</em></td>
<td>121</td>
<td>1.7</td>
</tr>
<tr>
<td><em>Homosexual</em></td>
<td>4,294</td>
<td>60.3</td>
</tr>
<tr>
<td><em>Bisexual</em></td>
<td>1,867</td>
<td>26.2</td>
</tr>
<tr>
<td><em>Other Orientation</em></td>
<td>811</td>
<td>11.4</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Male</em></td>
<td>2,748</td>
<td>39.4</td>
</tr>
<tr>
<td><em>Female</em></td>
<td>3,778</td>
<td>54.1</td>
</tr>
<tr>
<td><em>Trans</em></td>
<td>133</td>
<td>1.9</td>
</tr>
<tr>
<td><em>Other Gender</em></td>
<td>324</td>
<td>4.6</td>
</tr>
<tr>
<td><strong>Total n</strong></td>
<td>7,129</td>
<td></td>
</tr>
</tbody>
</table>
## Data Quality - Sample Composition (unweighted)

<table>
<thead>
<tr>
<th>Variable</th>
<th>LGBielefeld</th>
<th>SOEP-Q</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>4,159</td>
<td>59.9</td>
</tr>
<tr>
<td>25-34</td>
<td>1,550</td>
<td>22.3</td>
</tr>
<tr>
<td>35-49</td>
<td>748</td>
<td>10.8</td>
</tr>
<tr>
<td>50+</td>
<td>484</td>
<td>7.0</td>
</tr>
<tr>
<td><strong>School Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower</td>
<td>331</td>
<td>4.7</td>
</tr>
<tr>
<td>Medium</td>
<td>1,288</td>
<td>18.2</td>
</tr>
<tr>
<td>Higher</td>
<td>4,687</td>
<td>77.1</td>
</tr>
<tr>
<td><strong>Total n</strong></td>
<td>7,129</td>
<td></td>
</tr>
</tbody>
</table>
Project 2 - Showing Your Religion

- Online-Survey on experiences of discrimination in the employment history of Muslims
- Recruitment via ads on Facebook’s Timeline (Desktop and Smartphone)
- Recruiting Phase: January 15th - February 14th, 2021
- Targeted to muslim women with and without headscarves, muslim men and non-muslim population living in Germany
Examples of the Ads on Facebook & Instagram

Empirische Sozialforschung · Uni Bielefeld
Gespert · Finanziert von Universität Bielefeld

Du bist Muslim? Dann nimm jetzt an unserer Befragung zum Thema „Arbeitswelt und Jobsuche“ teil!

Empirische Sozialforschung · Uni Bielefeld
Gespert · Finanziert von Universität Bielefeld

Nimm jetzt an unserer Befragung zur Situation von Musliminnen auf dem deutschen Arbeitsmarkt teil!

WWW.UNI-BIELEFELD.DE
Hier geht es zur Umfrage.
Befragung der Universität Bielefeld

5 Kommentare 1 Mal geteilt

153 37 Kommentare 13 Mal geteilt
Campaign Results - Clicks, Respondents, & Costs

- Costs for ads: 4,259 €
- 1,068,760 impressions
- 21,486 link-clicks with 18,040 individual link-clicks
- 9,282 started the questionnaire
- 3,120 complete interviews (before first data cleaning)
  - Overall reached 37.6% Muslim participants
  - 14.3% female Muslims wearing a headscarf in everyday life
- Average net cost for a complete interview: 1,53 €
- 79 percent used a smartphone
Data Quality - Violation of Quality Indicators

- During data collection process: Evidence of systematic manipulation of study results
- Initial control based on eight known quality indicators, e.g., implausible and inconsistent data, duplicate enrollments, suspicious survey behavior.
- First results: Suspicious or fraudulent enrollments detected especially in the target group of Muslim men.

<table>
<thead>
<tr>
<th>No. of violated QI</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No violation</td>
<td>2,500</td>
<td>80.13</td>
</tr>
<tr>
<td>1 violation</td>
<td>489</td>
<td>15.67</td>
</tr>
<tr>
<td>2 violations</td>
<td>104</td>
<td>3.33</td>
</tr>
<tr>
<td>3 violations</td>
<td>24</td>
<td>0.77</td>
</tr>
<tr>
<td>4 violations</td>
<td>3</td>
<td>0.10</td>
</tr>
<tr>
<td><strong>3,120</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nearly **20%** of all interviews have at least one violated Quality Indicator!
Social media sampling = Cost-efficient way to recruit rare populations that are difficult to impossible to reach via other means

Methodological Issues
- Undercoverage
- Nonresponse & Selection bias
- Trolls & fake interviews

An expansion to other social networks (e.g., TikTok, Snapchat, WhatsApp) and different ad formats connected with the steadily increasing share of users further improves the applicability of the method in the future


Thank you.