Addressing Public Health Emergencies via Facebook Surveys: Advantages, Challenges, and Practical Considerations

André Grow¹(✉), Daniela Perrotta¹, Emanuele Del Fava¹, Jorge Cimentada¹, Francesco Rampazzo², Sofia Gil-Clavel¹, Emilio Zagheni¹

¹Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research, Germany

²Saïd Business School and Leverhulme Centre for Demographic Science, University of Oxford, United Kingdom

Abstract

Surveys of the general population can provide crucial information for designing effective non-pharmaceutical interventions to tackle public health emergencies, such as the COVID-19 pandemic. Yet, conducting such surveys can be difficult, especially when timely data collection is required. In this paper, we discuss our experiences with using targeted Facebook advertising campaigns to address these difficulties in relation to the COVID-19 pandemic. We describe central advantages, challenges, and practical considerations. This includes a discussion of potential sources of bias and how they can be addressed.

Keywords

Facebook, Online Surveys, Public Health Emergency, COVID-19

Introduction

As of September 9, 2020, the COVID-19 pandemic had caused over 27.4 million cases and over 894,000 deaths around the world [1]. To control its spread, national and local governments have implemented non-pharmaceutical interventions, including school closures, bans on large gatherings, mobility restrictions, and physical isolation, as well as unprecedented measures like local and nationwide lockdowns. Such measures have likely been critical for delaying and containing the COVID-19 pandemic [2]. However, individual behaviors, rather than governmental actions, may be crucial to control the spread of COVID-19 in the long run [3]. Timely and accurate data on human behaviors are thus of paramount importance to closely monitor the adoption of preventive measures, the emergence of symptoms, and changes in mobility and person-to-person contacts in the population.

In this context, surveys of the general population can provide central information needed to assess people's acceptance of, and compliance with, behavioral guidelines, but also to capture spontaneous bottom-up behavioral changes. Yet, researchers who want to conduct surveys that directly address ongoing epidemics are faced with unique methodological challenges: (1) these surveys need to be designed, implemented, and conducted quickly, as epidemics spread rapidly and are difficult to predict, especially when they involve new emerging diseases (*timeliness*); (2) they need to cover the entire population and, in the event of large-scale epidemics or pandemics, they need to be conducted simultaneously in multiple countries or regions, as regional differences could be relevant for designing effective interventions (*coverage*); (3) they

✉ Max Planck Institute for Demographic Research, Konrad-Zuse-Str.1, 18057 Rostock, Germany
e-mail: grow@demogr.mpg.de
phone: +49 381 2081 142

should be cost-effective, as obtaining large research funds quickly for an *ad hoc* survey can be difficult (*cost-effectiveness*). Furthermore, in the case of COVID-19, the nature of the recommended social distancing measures may limit some traditional modes of data collection, such as face-to-face interviews and even phone interviews, to the extent that they rely on large call centers.

In this contribution, we discuss the use of Facebook as a recruitment tool to address these challenges. Our assessment is based on the COVID-19 Health Behavior Survey (CHBS), that we conducted between March 13 and August 12, 2020, in eight countries (Belgium, France, Germany, Italy, the Netherlands, Spain, the United Kingdom, and the United States). Participant recruitment took place on a daily basis via targeted advertisements on Facebook, resulting in a total of 144,034 completed questionnaires. In what follows, we first provide an overview of the most important design aspects of the survey, and then discuss some of the central advantages, challenges, and practical considerations related to using Facebook advertisements in surveys that address public health emergencies, such as the COVID-19 pandemic. This includes a discussion and empirical assessment of potential sources of bias and how they can be addressed. Along the way, we make recommendations for those who want to implement similar surveys in the near future, and we hope that this will facilitate timely data collection to address the current—and possible future—public health and societal crises.

Methodological Approach

The CHBS is an online survey that focuses on people's reactions to the COVID-19 pandemic, targeting the population aged 18 and over. The questionnaire has four sections, encompassing socio-demographic characteristics (e.g., age, sex and education), health indicators (e.g., underlying medical conditions), behaviors and attitudes related to COVID-19 (e.g., perceived threat, preventive measures taken), and social contacts (i.e., the number of interactions with other people).

Recruitment took place via Facebook, by means of advertisement campaigns that we created with the Facebook Ads Manager (FAM). Facebook's business model centers on targeted advertisements and the FAM enables advertisers to create campaigns that can be directed at specific user groups. Targeting can be based on both users' demographic characteristics and a set of characteristics that Facebook infers from their behavior on the social network. Advertising campaigns consist of three levels. The highest level is the campaign level, at which the goals of the campaign are defined (e.g., generating awareness or generating traffic). The second level is the ad set level, at which the target audience, budget, and schedule are defined. The third level includes the advertisements themselves, which can consist of multiple advertising materials (e.g., images, videos), an advertising text, and the link to the page where Facebook users should be directed when they click on the ad (see Pötzschke and Braun [4] for more details).

We created one campaign per country and stratified each campaign at the ad set level by users' sex (male and female), age group (18-24, 25-44, 45-64, and 65+ years), and region of residence (largely following the NUTS-1 classification in Europe, aggregated into larger macro regions, and the census regions in the United States; see Perrotta et al. [5] for details), resulting in 24 to 56 strata per country. Each ad set contained six different images, leading to a total of 1,776 different ads. Figure 1 provides an example of the ads shown to Facebook users in the United States. We launched the campaigns between March 13, 2020 (Italy, the United Kingdom, and the United States) and April 4 (Belgium). This difference in the timing of the inclusion of countries is owed to a trade-off between the time needed to translate and technically

Figure 1. Example of advertisement in the Facebook advertising campaign in the United States.

implement the country-specific surveys and starting data collection in a timely manner. We concluded the survey on August 12, 2020.

As indicated above, Facebook advertising campaigns can have different goals and the overall costs that are incurred will partly depend on the chosen goal. We chose the goal of generating traffic. This means that Facebook’s algorithms optimize ad delivery, to maximize the likelihood that people who are shown an ad will click on it. Advertisers can define different ways in which Facebook should use their budget to meet these goals. For example, they can set a budget that is evenly spread over a fixed period, or they can define an average daily budget that Facebook seeks to meet weekly over an unspecified period. We opted for the latter, as the duration of the COVID-19 pandemic was uncertain. Based on these parameters, ad delivery is determined through an automated bidding process, in which a given ad competes for delivery with ads from other advertisers who are targeting the same user groups. In this process, Facebook considers the budget that could be afforded to deliver a given ad, but also how likely this ad will be of interest to users compared to competing advertisements. Before an ad campaign is launched, the FAM provides an estimate of, e.g., the size of the target audience and daily reach, which makes it possible to gauge the likely performance of the advertising campaign.

It is important to note that our study is not the first to use targeted Facebook advertisements for participant recruitment in health research. Earlier research has used this approach to address topics such as smoking behavior [6], cannabis use [7], and mental health [8]. For systematic reviews of the literature see Whitaker et al. [9] and Thornton et al. [10]. However, compared to this earlier work, our survey stands out because of its cross-national character, duration, and population coverage, given that we continuously collected data in eight countries for five months, and recruited more than 140,000 participants from virtually all parts of society and subnational regions.

Advantages

Our use of Facebook for participant recruitment enabled us to address the above introduced challenges of *timeliness*, *coverage*, and *cost-effectiveness*.

First, we were able to design, implement and launch the survey in a timely manner. Preparing a Facebook advertising campaign involves, in a nutshell, creating an advertising account, a Facebook page that is associated with the advertisements/the survey, and the ads themselves (which can happen in bulk, e.g., by means of a CSV file upload). Once the ads have been created, they need to be submitted for review, during which their compliance with Facebook's advertising policies is assessed. This review can take between a couple of hours and a day or longer. In our case it was usually completed within 24 hours (note that Facebook warned advertisers in April/May 2020 that reviews could be delayed due to the COVID-19 pandemic [11]). Once reviewed, the ads can be delivered and only need to be reviewed again when major changes are made (e.g., changes to the advertising materials or ad text).

Second, our use of Facebook enabled us to draw multi-national samples of diverse parts of the respective national populations. Facebook is the largest social media platform, with 2.45 billion monthly active users worldwide, as of Fall 2019 [12]. In the United States, about 69% of adults used Facebook in 2018, whereas this rate is 56% in Germany, 75% in Italy and Spain, 76% in France, 79% in the Netherlands, 85% in the United Kingdom, 89% in Belgium [13]. In general, older adults tend to be under-represented on Facebook, compared to younger and middle-aged adults [14]. Nevertheless, its user population provides a cross section of the overall-population with access to the Internet, and a comparison of different social media platforms (Facebook, LinkedIn, Twitter, Tumblr and Reddit) in the United States suggests that Facebook is most representative, in terms of user's educational attainment and internet skills [15].

Third, our use of Facebook made participant recruitment comparatively cost-effective, even though it can be difficult to determine the exact costs in advance. This is partly due to the nature of the bidding process that determines ad delivery and variation in the competition for advertising space. A central performance measure of Facebook advertising campaigns is the cost per click (CPC) and in their review of studies that have used Facebook for recruiting participants for health research, Whitaker et al. [9] reported CPC values between .17€ (\$.20) and 1.46€ (\$1.74). This variation is likely due to differences in the definition of the targeted user groups, the competition from other advertisers, and the likelihood of users clicking on the respective ad. Our costs are similar to those reported in earlier research. Between March 13 and August 12, 2020, we collected 144,043 questionnaires, at an overall CPC of about .14€ (\$.17) and an overall cost per completed questionnaire (CPQ) of about 1.05€ (\$1.25), excluding VAT.¹ The difference between CPC and CPQ is due to Facebook users who clicked on one of our ads but did not complete our questionnaire.

Thanks to these advantages, we were able to collect data that provide key insights into attitudes and behaviors that shape—and are shaped by—the COVID-19 pandemic. Figure 2 illustrates this by plotting for the entire observation period the average number of face-to-face social contacts that respondents reported for the day before participating in the survey. Face-to-face social contacts are the main vehicle for virus spread, given that the virus is mainly transmitted by infected secretions or respiratory droplets [16]. Figure 2 shows that there was great variation in face-to-face contacts over time, especially those that occurred outside the

¹ Note that as a relief measure, in Germany the VAT was reduced temporarily from 19% to 16% on July 1, 2020.

Figure 2. Average number of contacts at home and outside of home per week between March 13–August 12, 2020, across the eight countries. Lines show averages and shaded areas show 95% confidence intervals. Post-stratification weighting has been applied.

home. When paired with external information on, e.g., lock-down measures and infection rates, these data provide valuable insights into the effectiveness of different policies. When the data is additionally broken down by respondents’ demographic characteristics, it becomes also possible to assess whether different demographic groups responded differently to different policies. The data can also be used to calculate central epidemiological metrics (e.g., the effective reproduction number R_t) and to design more realistic epidemiological models. For more details on the contact patterns that we found in the CHBS, see Del Fava et al. [17]. For insights into other behaviors and attitudes in response to COVID-19 see Perrotta et al. [5].

To better illustrate the comparative strengths of our approach, it is helpful to contrast our work with similar *ad hoc* digital survey efforts that also focus on COVID-19. One prominent example is the work of Fetzer et al [18], who used an open survey to recruit 108,075 participants from 58 countries between March 20 and April 7, 2020. In this study, recruitment took place via link sharing on social media and similar channels. A second prominent example is the work of De Coninck et al.[19], who used an existing opt-in online panel maintained by a commercial polling company to recruit 1,000 participants in the Flemish region of Belgium between March 17 and March 22, 2020. Both studies used online surveys to collect information about people’s attitudes and behaviors in relation to COVID-19. Compared to Fetzer et al [18], our use of Facebook offered more control over the recruitment process, given that when the survey link is shared via social media, it is not possible to control who is invited to participate in the survey.

Hence, our targeted advertising arguably made it easier to ensure that our samples were demographically balanced. In comparison to De Coninck et al. [19], our use of Facebook arguably offered less control, given that existing online panels typically offer more detailed information about prospective participants than is available from the FAM. This makes it easier to collect demographically balanced samples. Yet, our use of Facebook offered a larger reach in terms of the number of countries that could be included and the time frame that was covered. In terms of costs, the CPQ that our paid advertisements incurred was higher than the CPQ incurred by Fetzer et al. [18], as their approach to link-sharing did not incur any costs, but it was similar to the CPQ incurred by De Coninck et al. [19],² who paid a commercial polling company for data collection.

Challenges

Eliciting information via self-administered (online) surveys faces some general challenges. For example, issues with recall inaccuracy often occur when some time has elapsed between a specific event of interest and participation in the survey [20,21]. This puts limits on the information that can be collected and should therefore be considered in light of the goals of the respective study. If, for example, detailed and accurate medical information is essential (e.g., exact blood pressure measurements), an online survey may not be the best choice and an in-person assessment with medically trained personnel may be preferable. Discussing the methodological challenges of online surveys is out of the scope of this paper (see Eysenbach and Wyatt [22] for such a discussion). Instead, here we focus on the challenges that are specific to using Facebook for participant recruitment.

The most important set of challenges relates to the issue of *self-selection bias*. Facebook's user base is a rough cross-section of the overall population with Internet access, but not all demographic groups are equally well represented [23]. Additionally, among those who use Facebook, there may be variation in their interest in the survey topic. Hence, there is no guarantee that the resulting samples will be representative in central demographic characteristics (e.g., age and sex) and important unobservable characteristics. This issue is potentially exacerbated by the algorithmic optimization that Facebook uses for ad delivery. If certain demographic groups are more likely to click on an ad, Facebook might increasingly deliver the ads to these groups, thereby reinforcing existing self-selection bias. This is particularly difficult to correct if participation and ad delivery are affected by user characteristics that cannot be easily considered in defining the relevant sampling strata.

Furthermore, if a survey is conducted over a long period of time, there may be changes in Facebook user activity. This may be due to seasonal variation in people's engagement with social media [cf. 24], but it seems possible that the development of the COVID-19 pandemic may also have led to changes in the composition of our samples over time. Especially in the early days of the pandemic, the virus dominated the news and lockdown measures were put in place to curb its spread. This may have increased participation in our survey in two ways. First, a lack of alternative activities due to lockdown measures may have led some population members to spend more time on Facebook than they would normally do, and this may have increased the likelihood that they see our ad. Second, the salience of the pandemic may have increased the chance that those who see our ads click on them. Over time, as the number of infections decreased and lockdown measures were eased, participation in the survey may have decreased, and the resulting samples may have become more selective.

² Source: Personal communication.

Figure 3. Number of daily (DAU) and monthly (MAU) active users between March 13–August 12, 2020, across eight countries based on all ad sets. Numbers are standardized per country by dividing the value for a given day by the average DAU/MAU over the entire observation period. Lines show 7-day moving averages.

Note: We collected DAU and MAU values every six hours, as these estimates can change within one day. We averaged these multiple observations to obtain one number per day. Data collection started between March 18 and April 4; no data was collected on April 11, April 12, and between August 3–6, due to technical issues.

Figures 3 and 4 explore changes in user behavior and survey participation over time, by plotting, for the entire study period, the average numbers of daily (DAU) and monthly (MAU) active Facebook users, as well as the click-through rate (CTR) for all countries. All three measures are based on estimates from Facebook, which we obtained via its Application Programming Interface (API). The DAU is the number of unique active users on a given day, whereas the MAU provides the number of unique users who have been active on Facebook within the last 30 days [25]. Both estimates are commonly used to assess the potential reach of advertising campaigns and we systematically collected DAU and MAU estimates for all our strata over the entire study period. The CTR is defined as the share of people who had clicked on an ad after seeing it and becomes available after a campaign has started and has been delivered to users [26]. Hence, changes in the DAU and MAU provide insights into changes in Facebook usage, whereas changes in the CTR provide insights into topic salience/interest among Facebook users. Note that between March 21 – March 26, we experienced technical

Figure 4. Click-through rate (CTR) between March 13–August 12, 2020, across the eight countries. The line shows the 7-day moving average. The shaded area shows the period during which technical problems with ad delivery occurred.

problems with ad delivery across countries, leading to a substantially lower number of participants than in the other weeks of our study. The CRT values for this period are thus less reliable than for the rest of the survey period.

As can be seen from Figure 3, in most countries the DAU was largest in the early weeks of the observation period, but then gradually decreased, usually by three to five percentage points. This means that the number of unique users who could have seen our ads on a given day decreased over time. The only exceptions from this are the United Kingdom and the United States, where the DAU was more erratic. The trends in the MAU values somewhat deviate from this. Typically, there was an initial increase in the number of monthly active users, whereas this number later decreased again. Hence, while the number of individuals who may have seen our ads on a daily basis decreased over time, towards the middle of the observation period we may have reached users who we would not have reached in other months. As Figure 4 illustrates, also the CTR values show a clear trend over time. Across countries, the click-through rate was initially high, then decreased, and ultimately increased again. This means that in the early phases of the survey, Facebook users were more likely to click on our ads than in later phases. Hence, our data suggest that over time the processes by which users selected themselves into the survey may have changed, but they do not allow us to assess precisely why these changes occurred.

How can the problem of self-selection be addressed? We suggest four methodological steps that can help alleviate this problem. First, in line with common approaches in traditional survey research, we suggest *stratifying ad campaigns*, based on characteristics that are known to relate to survey participation and the outcome of interest (cf. also Pötzschke and Braun [4]). Evidently, in the case of new emerging diseases, the relevant individual characteristics are difficult to know in advance. Even more, some relevant characteristics will not be available for creating strata in the FAM (e.g., pre-existing medical conditions). However, those characteristics that *are* available (e.g., age and sex) should be considered for stratifying the advertising campaigns. We suggest to also stratify ads by region within countries, as people's responses might vary locally. In this way, the bias that Facebook's ad delivery algorithms may generate is counteracted, leading to more balanced samples.

Second, in line with Zagheni and Weber [27], we suggest *applying post-stratification techniques* to the samples obtained from Facebook, to the extent that they deviate from the overall population in important characteristics. In this regard, using Facebook offers distinct advantages over other, less controlled ways of recruiting online samples. As indicated above, prior to launching a campaign, the FAM provides an estimate of the audience size with the characteristics of interest. Arguably, this feature is similar to a sampling plan and has been used in earlier research to conduct a 'virtual census' of the overall population (e.g., Zagheni et al. [28]). Furthermore, the ad performance estimates that FAM provides after a campaign has been launched (e.g., the number of users to whom a given campaign, ad set, or ad has been delivered, and how many of them had clicked on the ad) can be paired with information about survey completion rates by stratum. This makes it possible to calculate performance measures, such as approximate participation rates. However, it is important to keep in mind that many of the measures that Facebook reports are only estimates. The resulting indicators should thus be viewed as informed proxies.

As a complement to this approach, Zhang et al. [29] recently illustrated that by selectively activating and deactivating ad sets over the course of the survey period, it is possible to obtain representative samples from Facebook that do not require post-stratification. This approach is feasible if the specific timing of participation over the study period does not matter. Yet, in our case this was not feasible, as our goal was to obtain balanced samples daily. Selectively closing and opening ads over the course of a day or week would have implied that responses from certain subgroups may have been concentrated during a certain time of the day or days of the week. As an alternative, researchers may opt for dynamically adjusting the budget, so that more money is spent on strata that are under-represented in the survey. With this approach, it is important to keep in mind that large changes in the budget allocated to an ad set trigger the ad review process again, which can lead to a gap in data collection. As this would have undermined the goals of our study, we decided against this approach and instead continuously recruited members of all strata with a stable budget.

Third, we suggest that the issue of self-selection based on participant characteristics that are difficult to observe before people take part in the survey can be partly addressed by *considering possible sources of bias in the design of the survey and the advertising campaign*. In the case of our study, we expected that individuals who are particularly concerned about COVID-19 might be more likely to participate, and such concerns might also be reflected in reported behaviors and attitudes. As it is not possible to stratify Facebook advertising campaigns based on such concerns, we considered this issue in two ways. First, in the survey, we directly assessed participants' concerns about COVID-19 and other factors that may raise such concerns. Second, when selecting the images for our ads, we aimed to create variation in

1 – Male athlete
©Adobe Stock/grki

2 – Group of athletes
©Adobe Stock/nd3000

3 – Woman blowing nose
©iStockphoto/Goodboy
Picture Company

4 – Couple blowing noses
©iStockphoto/Goodboy
Picture Company

5 – Woman wearing mask
©Adobe Stock/shintartanya

6 – Man wearing mask
©iStockphoto/Михаил
Руденко

Figure 5. Images used in the Facebook advertising campaigns

how closely they are linked to the topic of COVID-19. To the extent that all these factors affect participation and answers to other questions, we can control for them in our analyses.

For illustration, Figure 5 shows the different images that we used in the ads. We considered images 1 and 2 to be least strongly linked to COVID-19, and images 5 and 6 most strongly. About 74% of respondents arrived at our survey via Image 5, about 16% via Image 6, about 7% via Image 3, and the rest via the remaining images. To assess whether the image through which participants arrived at the survey was related to their concerns about COVID-19, we conducted a Kruskal-Wallis test by rank. In this test, we assessed whether respondents' perception of how large a threat COVID-19 was for themselves personally (5-point Likert-type scale; 1 = very low threat/5 = very high threat) was associated with the picture via which they arrived at the survey. Table 1 shows the share of respondents who selected scale points 4 or 5, thereby indicating a high or very high threat perception. The share of respondents with a high to very high threat perception was largest among those who arrived via Image 3 (77%), and lowest among those who arrived via Image 2 (63%). The observed variation in threat perceptions across images was significant at the 1% level ($\chi^2(5) = 801, p < .01$). A broader analysis of how the different images related to respondents' self-reported attitudes and behaviors would be important, but our assessment suggests that the inclusion of different images helped recruiting more diverse samples in terms of concerns for COVID-19.

Table 1. Share of respondents who assessed COVID-19 as a high (4) or very high threat (5) for themselves personally on a 5-point Likert-type scale by ad image

Image	Percentage	Total <i>n</i>
1 - Male athlete	72%	586
2 - Group of athletes	63%	1,461
3 - Woman blowing nose	77%	9,280
4 - Couple blowing noses	72%	2,519
5 - Woman wearing mask	68%	102,061
6 - Man wearing mask	75%	22,451

Note: No weighting has been applied

Fourth, in recent years, the *multilevel regression and post-stratification* (MRP) approach to making inferences from highly selected survey data [30] has proved effective in producing unbiased population estimates [31,32]. In the first stage of MRP, the sample is partitioned into a large number of demographic strata (e.g., each combination of age group, sex, and region), and a multilevel regression model is used to estimate the outcome of interest, e.g., the average number of contacts or the percentage of people wearing a face mask, in each stratum. In the second stage, the stratum-level estimates are used to produce a final population-level estimate, using post-stratification weights to account for the proportion of each stratum in the population. This approach, used in combination with the previously mentioned steps, and given a fine partition of the sample into demographic strata, enables researchers to make proper inference at the population level even in presence of strong selection bias.

Another challenge relates to *trust in online surveys*. Online surveys may face suspicion, as they could be used to elicit personal information for non-research purposes (e.g., marketing, identity theft, etc.) [cf. 33]. In addition, over the last years, there have been several incidents that may have negatively affected the trust that the public has in the data protection measures put in place by Facebook. A prominent example is the Cambridge Analytica scandal, in which personal data of Facebook users was harvested without consent, with the goal to influence the 2016 US presidential election through micro targeting [34]. While fielding our survey, we encountered such suspicions in the commenting sections of our ads, and one notable concern was that Facebook would transfer personal user information to us. We addressed this issue by highlighting that the survey is anonymous and that no personal information was exchanged with Facebook, providing additional information about our research institute, the research team, and the goals of our survey, providing a link to our data protection policy, and providing information about preliminary results and reports as they became available. Despite these measures, it seems likely that Facebook users who are more concerned about data privacy were less inclined to participate in our survey. Additionally, while anonymous online surveys have the potential to reduce the likelihood that respondents provide socially desirable answers compared to personal interviews [35], privacy concerns may have rendered them reluctant to answer questions that they perceive as sensitive. We addressed this issue by offering the possibility of not answering questions that they feel uncomfortable with, to avoid forcing answers on sensitive topics.

Practical considerations

Finally, there are some practical aspects that need to be considered when using Facebook ads for survey research. First, the possibility to target certain user groups makes it easier to recruit

members of certain sub-populations, even when they are under-represented on Facebook. Yet, it is important to keep in mind that as the number of strata in the campaign increases, so will their selectivity and the costs, in particular if the members of certain strata are less likely to participate than members of other strata. For example, if the goal is to stratify a campaign based on 5-year age groups, the number of strata will increase considerably compared to strata based on 10- or 20-year age groups. This means that overall, more responses will need to be collected to have enough observations per stratum to apply post-stratification weighting. Furthermore, when members of some of these groups engage little with Facebook, a larger share of the budget will need to be devoted to recruiting them.

Second, both the advertisements and the study page to which they are linked need to be actively managed. Facebook advertisements are similar to user posts, meaning that users can react to them (e.g., liking), comment on them, and share them with friends. At the same time, they can leave posts on the study page and can review it. In our experience, it is important to engage with user comments and to provide additional information if needed, to maintain trust for current and prospective participants. The time investment that this requires should not be underestimated, especially for studies that run for a long period of time and are conducted in multiple languages. On average, our ads received about 19,300 impressions and about 135 comments per day, while it took us about one hour to manage between 50-100 comments. The exact time it takes to manage depended on the length of the comment and the complexity of the answer that was required. Over time, we became more experienced and efficient in managing comments. Based on these numbers, researchers who want to conduct a study on a similar topic and with a similar number of impressions per day should expect to spend about 1-2.5 hours per day on managing comments.

Third, the ad review process involves an automatic assessment of the links that are provided in the ads. In this process, the webpage to which the link leads is accessed. Hence, submitting ads for review can generate a large amount of traffic to the online survey, and it is important to keep in mind that the number of times that the survey page was accessed is not equivalent to the number of potential respondents who have accessed it, as some of the traffic may have been generated by the review process. It would therefore not be valid to approximate the survey completion rate by dividing the number of completed questionnaires by the number of page accesses. Furthermore, it is important to schedule sufficient time between submitting ads for review and launching the data collection. In our case, the review process was usually completed in a timely manner, but there might be delays (as those predicted by Facebook due to COVID-19) and problems (rejections due to violations of Facebook's advertising policies) that increase the time between review completion and delivery.

Conclusion

To conclude, we suggest that targeted advertisements on Facebook can be a powerful tool to recruit participants for *ad hoc* surveys of the general population during a public health emergency, as long as certain methodological steps are taken to address the issue of self-selection. We hope that the experiences that we have described here, together with our recommendations, will make it easier for other researchers to implement similar surveys to tackle the current and future pandemics.

Acknowledgements

We thank the editor and an anonymous reviewer for helpful comments on an earlier version of this paper. Francesco Rampazzo conducted most of his research for this article as a PhD student

at the Department of Social Statistics and Demography of the University of Southampton with a scholarship from the Economic and Social Research Council (South Coast Doctoral Training Partnership) [Project: ES/P000673/1].

References

1. World Health Organization. Weekly operational update on COVID-19: 9 September 2020 [Internet]. Available from: https://www.who.int/docs/default-source/coronaviruse/weekly-updates/wou-9-september-2020-cleared.pdf?sfvrsn=d39784f7_2
2. Ferguson N, Laydon D, Nedjati Gilani G, Imai N, Ainslie K, Baguelin M, Bhatia S, Boonyasiri A, Cucunuba Perez Z, Cuomo-Dannenburg G, Dighe A, Dorigatti I, Fu H, Gaythorpe K, Green W, Hamlet A, Hinsley W, Okell L, Van Elsland S, Thompson H, Verity R, Volz E, Wang H, Wang Y, Walker P, Walters C, Winskill P, Whittaker C, Donnelly C, Riley S, Ghani A. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand [Internet]. 20. 2020 Mar. [doi: 10.25561/77482]
3. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet Elsevier*; 2020;395(10228):931–934.
4. Pöttschke S, Braun M. Migrant sampling using Facebook advertisements: A case study of Polish migrants in four European countries. *Social Science Computer Review* 2017;35(5):633–653.
5. Perrotta D, Grow A, Rampazzo F, Cimentada J, Del Fava E, Gil-Clavel S, Zagheni E. Behaviors and attitudes in response to the COVID-19 pandemic: Insights from a cross-national Facebook survey. *medRxiv* 2020; [doi: <https://doi.org/10.1101/2020.05.09.20096388>]
6. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *J Med Internet Res* 2012 Feb 23;14(1):e28. PMID:22360969
7. Borodovsky JT, Marsch LA, Budney AJ. Studying cannabis use behaviors with Facebook and web surveys: Methods and insights. *JMIR Public Health Surveill* 2018 May 2;4(2):e48. PMID:29720366
8. Choi I, Milne DN, Glozier N, Peters D, Harvey SB, Calvo RA. Using different Facebook advertisements to recruit men for an online mental health study: Engagement and selection bias. *Internet Interventions* 2017 Jun 1;8:27–34. [doi: 10.1016/j.invent.2017.02.002]
9. Whitaker C, Stevelink S, Fear N. The use of Facebook in recruiting participants for health research purposes: A systematic review. *J Med Internet Res* 2017 28;19(8):e290. PMID:28851679
10. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: Systematic review. *Internet Interventions* 2016 May 1;4:72–81. [doi: 10.1016/j.invent.2016.02.001]
11. Facebook Inc. Delays expected for requested ad reviews [Internet]. 2020 [cited 2020 Apr 15]. Available from: <https://www.facebook.com/business/help/285854632398488>
12. Facebook Inc. Facebook reports third quarter 2019 results [Internet]. 2019 [cited 2020 Apr 15]. Available from: <https://investor.fb.com/investor-news/press-release-details/2019/Facebook-Reports-Third-Quarter-2019-Results/default.aspx>
13. Pew Research Center. Social media use 2018: Demographics and statistics [Internet]. 2018 [cited 2020 Apr 15]. Available from: <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>

14. Gil-Clavel S, Zagheni E. Demographic differentials in Facebook usage around the world. *Proceedings of the International AAAI Conference on Web and Social Media* 2019;13:647–650.
15. Hargittai E. Potential biases in big data: Omitted voices on social media. *Social Science Computer Review* SAGE Publications Sage CA: Los Angeles, CA; 2018;38(1):10–24.
16. World Health Organization. Transmission of SARS-CoV-2: Implications for infection prevention precautions [Internet]. [cited 2020 Sep 29]. Available from: <https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions>
17. Del Fava E, Cimentada J, Zagheni E, Perrotta D, Grow A, Rampazzo F, Gil-Clavel S. The differential impact of physical distancing strategies on social contacts relevant for the spread of COVID-19. *medRxiv* 2020; [doi: <https://doi.org/10.1101/2020.05.15.20102657>]
18. Fetzer T, Witte M, Hensel L, Jachimowicz J, Haushofer J, Ivchenko A, Caria S, Reutskaja E, Roth C, Fiorin S, Gómez M, Kraft-Todd G, Götz FM, Yoeli E. Perceptions of an insufficient government response at the onset of the COVID-19 pandemic are associated with lower mental well-being [Internet]. *PsyArXiv*; 2020 Apr. [doi: 10.31234/osf.io/3kfmh]
19. De Coninck D, d’Haenens L, Matthijs K. Perceived vulnerability to disease and attitudes towards public health measures: COVID-19 in Flanders, Belgium. *Personality and Individual Differences* 2020;166(Online first).
20. Ge S, VonKorff M. Recall of psychiatric history in cross-sectional surveys: Implications for epidemiologic research. *Epidemiol Rev* 1995 Jan 1;17(1):221–227. PMID:8521941
21. Newell S, Girgis A, Sanson-Fisher R, Ireland M. Accuracy of patients’ recall of Pap and cholesterol screening. *Am J Public Health* 2000 Sep;90(9):1431–1435. PMID:10983202
22. Eysenbach G, Wyatt J. Using the Internet for Surveys and Health Research. *J Med Internet Res* [Internet] 2002 Nov 22;4(2). PMID:12554560
23. Ribeiro FN, Benevenuto F, Zagheni E. How biased is the population of Facebook users? Comparing the demographics of Facebook users with census data to generate correction factors. *arXiv* [Internet] 2020 May 16; Available from: <http://arxiv.org/abs/2005.08065>
24. Facebook Inc. Capturing the spirit of summer [Internet]. Facebook IQ. [cited 2020 Sep 28]. Available from: <https://www.facebook.com/business/news/insights/capturing-the-spirit-of-summer>
25. Facebook Inc. Graph API reference v8.0: Ad campaign delivery estimate - documentation [Internet]. Facebook for Developers. [cited 2020 Oct 9]. Available from: <https://developers.facebook.com/docs/marketing-api/reference/ad-campaign-delivery-estimate/>
26. Parameters - Marketing API - Documentation [Internet]. Facebook for Developers. [cited 2020 Oct 9]. Available from: <https://developers.facebook.com/docs/marketing-api/insights/parameters/v8.0/>
27. Zagheni E, Weber I. Demographic research with non-representative internet data. *International Journal of Manpower* 2015;36(1):13–25.
28. Zagheni E, Weber I, Gummadi K. Leveraging Facebook’s advertising platform to monitor stocks of migrants. *Population and Development Review* 2017;43(3):721–734.
29. Zhang B, Mildemberger M, Howe PD, Marlon J, Rosenthal SA, Leiserowitz A. Quota sampling using Facebook advertisements. *Political Science Research and Methods* Cambridge University Press; 2020 Jul;8(3):558–564. [doi: 10.1017/psrm.2018.49]
30. Gelman A, Little TC. Poststratification into many categories using hierarchical logistic regression [Internet]. Citeseer; 1997. Available from: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.5270>
31. Downes M, Gurrin LC, English DR, Pirkis J, Currier D, Spittal MJ, Carlin JB. Multilevel regression and poststratification: A modeling approach to estimating population quantities

- from highly selected survey samples. *American Journal of Epidemiology* Oxford University Press; 2018;187(8):1780–1790.
32. Wang W, Rothschild D, Goel S, Gelman A. Forecasting elections with non-representative polls. *International Journal of Forecasting* 2015;31(3):980–991.
 33. Evans JR, Mathur A. The value of online surveys. *Internet Research* 2005 Apr;15(2):195–219. [doi: 10.1108/10662240510590360]
 34. Lapowsky I. How Cambridge Analytica sparked the great privacy awakening. *Wired* [Internet] [cited 2020 Oct 9]; Available from: <https://www.wired.com/story/cambridge-analytica-facebook-privacy-awakening/>
 35. Stern MJ, Bilgen I, Dillman DA. The state of survey methodology: Challenges, dilemmas, and new frontiers in the era of the tailored design. *Field Methods* SAGE Publications Inc; 2014 Aug 1;26(3):284–301. [doi: 10.1177/1525822X13519561]