

Case study

Data quality in citizen social science

SOURCE: Heiss, R., & Matthes, J. (2017). Citizen science in the social sciences: A call for more evidence. *GAIA-Ecological Perspectives for Science and Society*, 26(1), 22-26.

"[...] in the citizen science project Young Adults' Political Experience Sampling (YAPES), school students are engaged in independent data collection using their smartphones (comments and pictures sent via Email or WhatsApp). Based on their own daily experience, they followed the predefined research question, how young people get in touch with political issues in their everyday lives. The key objects of interest were the places (e.g., school, on the street, on social media, etc.), the channels (advertisement, media, political talk, etc.) and the content (what issues) of political youth communication in Austria. Whereas researchers hardly have access to this information and usually have to rely on either memory-based survey data or a small number of qualitative interviews, the citizen science approach allows to collect large-scale real experience data. Even though YAPES was designed as just a small pilot project, the participating 254 students collected a total of 1,768 observations. In this participatory research context, the young participants have evolved from a mere subject role to an active scientific role, as they actively screened and reported information about their political environment. Selected results were published on the project blog in order to provide the participating students with an outcome of their effort. YAPES, as other similar projects currently evolving in the social sciences, show that there are spaces where citizens can add value to social sciences research (SSR) by engaging in tasks which have traditionally been implemented by scientists."

The authors identify three main problems for implementation of citizen science in social sciences: "(1) difficulties in mobilizing target groups, (2) the relation between human observation and data quality, and (3) ethical questions.

First of all, there is reason to believe that it might be a considerable challenge for social scientists to mobilize target groups. [...] Second, data quality is one of the core challenges in citizen science in general. Yet there is reason to believe that it is an even bigger problem for SSR [social science research]. Rather than objectively measuring observations, for instance through technical devices (like in many natural science projects), SSR does rely much more on human observation. In fact, human observation such as coding, counting, documenting, photographing, recording, etc. is potentially biased because it is built upon human perception. [...] Furthermore, the measurement process needs to be independent from the individual characteristics of the volunteers, such as political leanings or personal opinions about social scientific research objects. [...] Third, ethical issues arise from the fact that the public is co-producing the data and hence has an "active stake" in the data production (Riesch and Potter 2014). Resnik et al. (2015) identify four main areas of ethical concern: data quality and integrity, data sharing and intellectual property, conflicts of interest, and exploitation of volunteers. As compared to research in the natural sciences, ethical issues regarding data sharing maybe of particular

concern in SSR, which mostly carries out research on humans. This is especially the case when citizens use their own smartphones and may provide sensitive data about themselves or other humans they observe. For example, sharing data about a person's political ideology or interpersonal behavior may raise more ethical issues than sharing data about local soil."

Questions for discussion:

1. Data quality is one of the main challenges in citizen science in general, but arguably it might be an even bigger problem for social sciences research. What are the challenges for data quality that researchers might face in the case above? What are the reasons of those challenges? What can be done to mitigate these problems?
2. Develop a list of recommendations that the researchers should implement in this case and in citizen social sciences in general in open social sciences to ensure data quality (use the table below for this task)!

| Challenges for ensuring data quality (in this case and in open social sciences in general) | Recommendations |
|--|-----------------|
| | |
| | |

| | |
|--|--|
| | |
| | |
| | |
| | |