Methodological pragmatism in doing data: A rejoinder to ‘Collecting qualitative data during a pandemic’ by David Silverman

SRIKANT SARANGI
Aalborg University, Denmark and Cardiff University, UK

Silverman’s discussion paper is focused on gathering qualitative data in contexts where physical presence of the researcher at the site of research had been the norm until COVID-19 struck. As he points out, conducting qualitative inquiry during the COVID-19 era has been particularly challenging with regard to accessing in-person, in-place ethnographic data, thus necessitating certain adjustments. He divides his discussion into problems and remedies, and overall it comes across as a statement from someone who has been a pioneer in ‘methodological pragmatism’ (Miller and Dingwall 1997: preface).

The COVID-19 situation certainly calls for methodological pragmatism in qualitative inquiry. In broad terms, pragmatism concerns situational awareness of affordances and constraints, with conditions of adaptability at its core. William James captures Charles Peirce’s position on pragmatism as follows: ‘to develope [sic] a thought’s meaning, we need only determine what conduct it is fitted to produce: that conduct is for us its sole significance’ (James 1975 [1907]: 29). In this sense pragmatism converges with praxis, including ‘a possible difference of practice’ (James 1975 [1907]: 29).

The very foundation of qualitative research – ‘doing data’ in person by ‘being there’ to gain a ‘thick’ sense of the context where the activity happens – has been threatened as a direct consequence of the ensuing pandemic. Certain modes of being (e.g. being a participant observer) to collect in-person, in-place data have become impractical. A brake – literal and metaphorical – has to be applied to ongoing projects in their initial stages of data collection, and any planned projects (for funding or otherwise) have to be shelved because of the lockdown measures restricting physical access. And even in cases where access is possible, social distancing and face mask requirements are likely to interfere with carrying out observational procedures. There is no doubt that such restrictions must limit the nature and extent of data collected, and this may amount to qualitative researchers losing out on their embodied experience as data gatherers.

My latest research concerns emergency medical calls, which are mediated via phones and recorded routinely for auditing and training purposes. There is no visible impact of COVID-19 on access to raw data for research projects of this kind. Nonetheless, the pandemic has affected the research process, which relies on collaborative interpretation involving nurse professionals as call takers, and in-house workshops which occasion opportunities for two-way ‘hot feedback’. The impact of the pandemic is felt beyond the data-gathering phase of qualitative inquiry.

In terms of pragmatism, some of the methodological remedies surrounding data gathering that Silverman discusses include alternative modes of ‘physical presence’ as one of the digital consequences of the pandemic. Among the
alternatives, he lists the following: shift from talk/interaction-based research to text-based research; adoption of digital media for data collection; transfer of responsibility of data collection to research participants; reuse of existing datasets etc. To single out the prospect of reuse of data – whether one’s own or others’ – may not be such a straightforward option, since it comes with ethical baggage concerning general data protection guidelines. Also, it is worth noting that reuse of existing data sources may not be compatible with certain research paradigms and analytical frameworks (e.g. ethnography).

Reuse of others’ data was more of an opportunity and less problematic decades ago. In the late 1980s my own PhD study at Lancaster University was based on job interview data which was obtained second-hand. When I decided to work on job interviews for my PhD project I knocked my head against many a brick wall seeking access to raw, first-hand recordings. Then, as now, such gatekeeping encounters were largely inaccessible because of their sensitive nature. In the absence of any easy downloading of digital data such as YouTube videos at the time, I relied on reusing data that existed as part of another funded project, with consent from the project coordinator. I was allowed access to selected videos of job interviews, which I transcribed to fit my purposes. I argued at the time – when PhD researchers were routinely expected to collect their own data – that recycling was a desirable and sustainable venture given how much collected data was generally wasted or remained under-analysed (Sarangi 1987).

The decision to use pre-existing data did impact my choice of analytical framework, while ensuring minimal compromises. Such reuse of others’ data would be almost impossible under the current General Data Protection Regulations (GDPR). More specifically, video and audio recordings are now required to be destroyed after a specific time-span; if the data is still alive, the participants – and not just the owner of the pre-recordings – must give consent again for any reinterpretation and this should go through relevant ethics committees, as it would constitute a new study design for a new purpose.

Revisiting one’s own data archives has a distinct practical advantage, and this is not specific to the pandemic situation. In my professional career some years ago I fell victim to a technological mishap as my laptop suddenly died and I lost hundreds of documents, including data archives. I did not have the insight at the time to have backups of everything. This predates the automatic saving process via the institutionally available OneDrive. From a disused floppy disk I could recover some interview and classroom encounter data about a project that I had not analysed, so it was a second chance to engage with the existing data (Sarangi 2015). Reusing one’s own data archive remains an available route during the pandemic without circumventing GDPR and compromising the quality of qualitative research.

The alternative of switching from talk/interaction to text/document as the object of research is not an optimal answer to the problem, but is no doubt an option that is being taken up by many researchers. More specifically, it is being actively considered by doctoral researchers who have not been able to collect their data, thus forcing them to change the direction of their research topics. The disproportionality of this is felt distinctively by those who rely on fieldwork and interviews in physical settings, as opposed to those who mine data from the internet, including social media. In the latter case, there is no change of practice needed and there is no threat to research productivity.

Digitalisation of the data-collection instrument is both a topic and a resource for further consideration. It is useful to distinguish here between different kinds of research-data sites where digital practice is an integral part of routinised communicative activities, such as digital clinic consultations, university lectures, chatrooms. These are new social realities that present opportunities to be studied in their own right. Digitally mediated data collection is bound to increase manifold, as has been the case with online shopping during COVID-19: in the UK, online shopping increased by 36% within the COVID year, which would otherwise have taken more than a decade to reach (Islam 2021). This is comparable to online data mining, which resembles online shopping. In terms of digitalisation as a resource, qualitative researchers will need to embrace pragmatism regarding how to recruit participants, how to gain consent and how to conduct research interviews remotely when the researcher will have little control over intervening contextual variables.
Keeping with the testing regimen that we have all been subjected to during the ongoing pandemic, in what follows my comments will be mainly in the form of qualitative researchers undertaking three self-tests, with the distinct possibility of false-positive and false-negative outcomes! The tests are as follows: passing the essentialness test; passing the authenticity test; and passing the integrity test. The ‘essentialness’ dimension relates to the conduct of COVID- and non-COVID-related qualitative inquiry during the pandemic as well as the kind of data required to sustain one’s study; the ‘authenticity’ dimension concerns the qualitative nature of in-person, in-place vis-à-vis digitally mediated data; and the ‘integrity’ dimension spans both research integrity and researcher integrity during the ongoing pandemic. I elaborate on each of these dimensions below.

**Passing the essentialness test**

During COVID-19, what constitutes ‘essential’ in different spheres of social life has come to the fore. This particularly relates to what travel is considered essential travel. Neither international nor local travel for data collection qualifies as essential travel, even if we persist that in-person, in-place data is integral to our research project. Besides travel, another area of what is essential concerns our shopping practices/preferences and by extension which shops or sections of shops should remain physically accessible. A video went viral of a man in his boxer shorts and a face mask entering the Tesco supermarket in Wales (Cole 2020) as a protest against the government’s decision to declare clothing retail as non-essential. Intriguingly, the sale of alcohol did not fall under non-essential shopping. The dichotomous essential/non-essential categorisation has been very salient in the context of urgent and non-urgent healthcare delivery. A glaring example is how COVID-19 patients have taken precedence over other kinds of patients needing cancer treatment, surgery, management of chronic conditions – the latter have been labelled as non-urgent interventions, affecting many individuals and families and leading to a monstrous backlog.

In extrapolating the above scenarios, the broader questions are as follows. Is doing qualitatively oriented social science research ‘essential’ when we find ourselves in the midst of a pandemic? And by extension, if social scientific expertise does not contribute towards combating the pandemic here and now, should such qualitative research be shelved?

We need to distinguish here between two different strands of research: (1) qualitative studies directly addressing COVID-related issues in examining the impact/experience of the pandemic; and (2) qualitative studies that concern other non-COVID-related topics being conducted during the pandemic. As regards the former, this readily meets the ‘essential’ criterion, although perhaps not at the level at which research dealing with innovative treatment methods or vaccine development fronted by the biomedical sciences does. Certainly, sociological and psychological research depicting perceptions about and the impact of COVID-19 in terms of behaviour change and attitudes towards risk taking or risk aversion is regarded as essential for informed policy decisions.

In the public media sphere, at least in the UK, the discipline of psychology (health psychology in particular) is more visible and takes pride of place as individual/public behaviour patterns including belief systems and motivations underpinning action have emerged as a central issue. Under this rubric fall language/communication-based studies concerned with discursive representations of perceptions and actions about COVID-19, where data is collected from public media and/or social media and does not require participants’ involvement. Rather than being inhibitive, the COVID-19 pandemic has presented new opportunities – thematically and methodologically – for specific types of language/communication research trajectories, which have had a huge surge. Textual data mining (as distinct from talk/interaction-oriented data gathering) has increased exponentially during the COVID-19 era; whether or not they constitute essential or non-essential research somehow bypasses deliberation.

With regard to the second strand described above, if the topic of research is not about COVID-19, could this be shelved until better times, in the same way as non-COVID-related medical interventions have been deferred by hospitals? In this sense, such research may struggle to pass the ‘essential’ test, although there is no
denying that early career researchers, especially doctoral students, will find themselves in harm’s way when such arbitrary categorisations of ‘essential’ and ‘non-essential’ take root and become consequential.

Then, the next question surfaces: is in-person, in-place data essential for pursuing qualitative research, whether COVID-19 related or not? At a very material level, we need to consider what might be lost if data is collected remotely/digitally rather than physically, which overlaps with the ‘authenticity’ dimension discussed in the next section.

**Passing the authenticity test**

In qualitative research, the authenticity of the data being analysed constitutes the Holy Grail. In-person, in-place data in any ethnographic inquiry serves as the benchmark, and so a suspension of physically embodied data collection in favour of adopting a remote, digital format may be seen as generating inauthentic data. But is such a presumption valid?

Qualitative research rooted in researcher-generated data is invariably contaminated (cf. the observer’s paradox, also referred to as the Hawthorne effect), and is a matter of degree rather than either/or. The online context can act as a variable, but in pre-pandemic days, when social proxemics was a given, whether the interview took place at work, at home or at a mutually neutral place always altered the nature of the data collected. All kinds of data are mediated, but data collected digitally is likely to be mediated in a distinct way. As Malinowski (1935) very aptly observed a long time ago, even ethnographic field notes in the case of researchers being physically co-present are selective and made up of ‘constructive drafting’; i.e., ‘constructed by the observer out of a multitude of manifestations of varying significance and relevance [...] by inductive computation’ (Malinowski 1935: 317). The imperative, then, is to account for the context of the data collection in our analytical commentaries — something that is routinely overlooked in interview- or focus-group-based studies adopting grounded theory or thematic/content analysis procedures.

The context in which a digital interview unfolds acts as an intervening variable which is not always apparent to the researcher. This happens in journalistic interviews. It is useful to recall the famous 2017 BBC World News interview involving Professor Robert Kelly, an expert on South Korea. While the interview was in progress, two toddlers entered the camera space. There was a brief disruption, but the interviewee kept a straight face; a woman, Korean in appearance, entered the room and took away both the children. In the follow-up public comments, the woman was referred to as the nanny, when in fact she was the mother of the toddlers and the wife of the white interviewee. This was both racial and gender stereotyping of an Asian woman in such a context. Even more strikingly, the way she dragged the children out of the room by one hand was seen as uncaring and inhuman. As more background context emerged, it became known that at the time of the invasion of the children into the camera space, the wife/mother was in the bathroom, and she only had one hand free to deal with the children, with her other hand keeping her skirt in place. Visibly, she was in a state of panic. The accusing finger was also pointed at the husband, as to why he did not get up to steward the children away from the camera lens. Some interpreters even went to the extent of suggesting that he might not have been wearing proper trousers, as these would be redundant in the context of online performances where only the upper part of the body is visible. Indeed, there was a follow-up interview involving the same BBC interviewer and the couple to seek their accounts of the situation and the ensuing public comments, which provides us with secondary data for triangulation purposes. This scenario underscores how interpretation of remotely accessible data is likely to be ‘thin’ on context.

In relation to qualitative inquiry, one nagging question that often raises its ugly head concerns the overall quality of remotely generated data and its authenticity. To what extent are online recordings naturalistic? Does any data, for that matter, qualify as naturalistic? Even in face-to-face settings, Cicourel (1964) refers to the paradox of interaction in the research interview setting — where the framing of the interview as research instrument may clash with the framing of the interview as social interaction — for both researchers and research participants. As he puts it: ‘A double responsibility is imposed on the interviewer; he
must simulate spontaneous participation while evaluating the subject's views toward the interview, the observer, and their relationship' (Cicourel 1964: 75). The digitally mediated interview will be no exception in interactional terms and will introduce similar or different intervening variables. If data is always constructed by the researcher's own positioning, then what emerges as data in an in-person, 'being there' situation will be different from data collected remotely, via Zoom etc. Viewed from another perspective, as people in their everyday life get used to Zoom meetings with family members and friends, having the research interview in a remote format might appear more naturalistic. In contrast, as already indicated, although there is a surge in research outputs based on media data, it is worth pointing out that access to raw data in the public domain may still be suppressed by authorities and the researchers will have no control over this, in which case the authenticity of the data in hand can be called into question.

A second issue is whether the same ethical guidelines can be upheld when accessing research participants remotely. This also pertains to gaining informed consent digitally and in spoken format, as written consent forms might hinder the data-collection process. Much of the sociological and psychological research about/during COVID-19 has continued using remote data-collection techniques, although such mediated events are likely to intervene and influence the quality of data collected. Context is a key variable of interaction, and the researcher may not have control over who might be co-present during an online interview and how this might affect what an informant might disclose on record.

However, when remotely interviewing an informant, other family members may be listening in, which would dilute the authenticity of the accounts offered, especially if the topic is sensitive and private in nature. The variable of participation framework in the Goffmanian sense can be either inhibitive or facilitative. In being mindful that eavesdroppers do not assume the status of primary participant, we can remain positive that a conversation which includes family members can generate rich data.

Between the mid-1990s and early 2000s, when I was conducting research on genetic counselling, it used to be common practice for specialist genetic nurses to arrange home visits ahead of clients' clinic attendance. The openness of participation of family members in these home visits went hand-in-hand with the guidelines for genetic counselling consultations, which encouraged family members to attend the clinic – given the familial nature of genetic conditions and the potential at-risk status of family members – and even insisted on a companion being present when test results were given. This reminds me of Barrett's (2004) exploration of the interplay between ethnographic context and interview conversation in psychiatry – how in a Borneo longhouse setting the dyadic conversation is transformed into collective interlocution in the sense that it becomes a public affair, thus compromising confidentiality. Under the rubric of observational studies, the contingent, intervening variables must be accounted for in the analysis.

Of course, protocols exist whereby research participants rather than researchers can take over the responsibility of collecting data on their practices (as in the case of family communication research where the participants are given recording equipment for volitional use). The COVID-19 pandemic does not alter that scenario. In an ongoing ESRC project which investigates the deliberation practices of clients in their decision-making trajectories, the datasets include different types of data such as self-diaries, clinic recordings, interview accounts. With regard to self-diaries, our participants had the option of either digitally or manually recording their experiences and expectations about genetic counselling, including deliberations regarding decision making. It is noteworthy that the language of these distinct datasets is different, offering differential perspectives. Although data collection using self-diaries is least likely to be hampered during the COVID-19 pandemic, self-isolation may have had an influence on the content of the diary data. This data, unlike interview data, is very reflective in offering deeper insights. Besides the proven merits of multi-method approaches to data collection, such a scenario raises the question about which data, i.e. which account of the client, should be taken at face value and considered more 'authentic' when making analytical claims. The crux here is not so much whether a specific dataset is authentic or inauthentic in itself, but whether the interpretation can be regarded as being authentic in its attention to the context of data generation. Interpretation of
data when the researcher is in self-isolation might pose other specific difficulties, which I take up in the next section.

**Passing the integrity test**

Whether or not a pandemic becomes the social reality, research integrity has to be maintained, and this applies to qualitatively oriented social scientific research. As part of research integrity, all research must go through usual ethics committees and meet data protection regulations, informed consent and anonymity procedures etc. When research is interventionist and has immediate societal impact (e.g. clinical trial for vaccine development) research integrity is of utmost importance – the protocols for recruitment of participants, inclusion/exclusion criteria etc. need to be disclosed to render credibility to the research findings. There is urgency involved in the race of virus vs vaccine, so biomedical research has to be fast paced. In a climate of competition and the race for ‘being first’, research integrity may be compromised, which may particularly apply to clinical trials and related interventionist research projects (cf. increasing retraction of peer-reviewed published articles in biomedical research). Some researchers may be too proactive in short changing ethical issues in their eagerness to study the important topic of COVID-19.

Researcher integrity is equally important in qualitative inquiry, especially relating to collection and interpretation of data as well as dissemination of findings. If the data is collected remotely, the researcher has to acknowledge the known and unknown contextual variables and account for them. On dissemination, one may feel encouraged to consider alternative avenues, but there are risks attached to dissemination via social media in bypassing the peer review process. It is also an issue if qualitative researchers in the humanities and social sciences become opportunistic in mining particular kinds of data from media and social media. Numerous studies report their newly kindled interest in COVID-19 in the context of public health messaging and social media, encompassing themes such as risk, trust and quality of life and focusing on linguistic phenomena such as metaphors. Corpus linguistic methodology is at the forefront of application – readily applicable to easily accessible COVID-19 data. While this research trend is very much visible, the impact of such studies is suspect.

Regarding the ESRC project involving diary data discussed above, the impact that such data (which included suicidal thoughts) has had on the researcher in the act of interpretation during the lockdown period is the topic of reflection:

A particular challenge whilst working in isolation during the first wave of the COVID-19 pandemic was analysing diary data. This data was unlike the interview and clinical consultation data for two reasons; firstly, the researcher usually participates in the generation of interview and clinical consultation data, and secondly, the diary data was more emotional, raw and contained unanticipated content. More specifically, the diary data contained upsetting accounts regarding the care of loved ones as well as plans of suicide. The emotional toll of analysing the diary data was further impacted by not having physical access to other researchers in the team. This experience raises the general question about how research networks need to come up with ways to safeguard researchers’ wellbeing, on par with participants’ wellbeing, in the research process. (Lisa Ballard, pers. comm.)

For me, this relates to ‘ethics of interpretation’ (Sarangi 2019); this is particularly salient when one is compelled to work in self-isolation because of the pandemic, and it has clear implications in relation to sense making, mental wellbeing and uptake of qualitative research.

With regard to researcher integrity, the early career PhD student is at the most risk. Alternative modes of data collection may also mean a change of focus, including study design. I am reminded of recently supervising a visiting international PhD student in the early stages of her PhD journey who had collected some pilot data for our supervisory consultations. Her topic is end-of-life training sessions for palliative care physicians, and the main dataset is intended to include recordings of training sessions and ethnographic field notes. The project, though, is currently at risk of not progressing, as the training sessions have been discontinued and/or the physical presence of any
researcher would be restricted. For early career researchers, there are delays, deferrals, alteration of topics and methodologies, and even abortion of the project itself. This means ethnography-oriented qualitative researchers begin to see themselves in the role of objects of study (accounting for how COVID-19 has impacted their research trajectories and by extension their professional careers) rather than being agents of research exploring others’ lived experiences.

Instead of a conclusion

While research in areas of public health, epidemiology, virology, clinical trials etc. has boomed during COVID-19, qualitatively oriented social science research seems to have dwindled. In healthcare research, the poor cousin status of qualitative research specifically relates to a lack of demonstrable impact on the researched and the society as a whole, mainly because by design qualitative research does not deal with numbers and findings are not quantifiable and generalisable. Against the backdrop of the powerful rhetoric of numbers in quantitative studies, there is overarching scepticism about COVID-19 related qualitative research, primarily linked to the nature of data sampling and the subjective interpretive frameworks. While it is unthinkable what impact qualitative research can have regarding COVID-19 at a societal level, it is intriguing to see the ongoing – and lasting – impact COVID-19 has on qualitative research. We are still in the pandemic and we can expect the experiences of qualitative researchers ‘before, during and after’ the pandemic to abound, requiring critical reflexivity and pragmatism. More specifically, data collection practices ‘before, during and after’ COVID-19 very likely will trigger salient methodological reflections, parallel to how alternative insights are gained by switching analytical lenses after being exposed to new tricks and tools of the trade.

Very recently I participated in a panel discussion as part of a doctoral summer school in Finland. The panel was titled ‘Research ethics in exceptional times,’ which primarily meant conducting research ethnographically. The basic parameters of conducting ethnography – going native, getting a sense of context, participant observation, field work – may become extinct as routine research practice. A dominant methodology like ethnography will then gradually be silenced by COVID-19. Many questions present themselves but will remain unanswered. Does a pandemic like COVID-19 present an existential threat to qualitative research? Does the quality of qualitative research suffer under lockdown and social distancing regulations? Can researcher-researched intimacy be fostered through zoom-mediated interviews? What would a ‘stay at home’ policy translate to in terms of conducting qualitative research, including interpreting data in self-isolation? Does it mean ‘research from home’ could dilute the quality of data gathered and its interpretation, failing the tests of authenticity and integrity? Some of the challenges to qualitative research practice – and by extension, research ethics – existed in pre-pandemic times and will remain with us in the post-pandemic era.

References


