Methodological Reflections on Online Data Collection during the Covid-19 Pandemic

Sumedi P. Nugraha¹ and Dewi H. Susilastuti²,³

¹Department of Psychology, Faculty of Psychology and Socio-Cultural Sciences, Universitas Islam Indonesia, Yogyakarta, ²Master and Doctoral Program in Leadership and Policy Innovation, The Graduate School, Universitas Gadjah Mada, Yogyakarta, ³Center for Population and Policy Studies, Universitas Gadjah Mada, Yogyakarta

Corresponding Author: Sumedi P. Nugraha (email: sumedi.nugraha@uii.ac.id)

Abstract

The pandemic closed the door for the use of conventional, face-to-face data collection methods. At the same time, it built a momentum for the exploration and utilization of online data collection methods. However, the belief about superiority of the offline data collection persists. The literature review and the authors' research experience reveal that offline and online data collection methods yield similar result in terms of data completion and quality. All data collection methods contain weaknesses and strengths. Nonetheless, the online data collection methods are very versatile. They allow the researchers to choose the tools that best align with their research objectives.

Keywords: research amidst the pandemic; strengths and weaknesses of online data collection methods; research ethics

Introduction

Curiosity and questions about why people think the way they do, or why people behave the way they do, among other issues, are the triggers of research. The purpose of all research is to answer research questions or prove a hypothesis by collecting and analyzing data through the use of a combination of methods. There are a variety of true and tested ways to complete the research steps. Qualitative researchers are in a unique position as they are the main instrument in data collection. Meanwhile, the relationship between quantitative researchers and their respondents is mediated by some validated and ideally objective questionnaires.

The days of “armchair researchers” are long gone. There is an expectation for quantitative and qualitative researchers alike to go to the field to understand the geographical, social, cultural, and political context of the research. “Being in the field” enables researchers to get to know the people they study, engage with them, and in the case of qualitative researchers, immerse in their culture and get an in-depth understanding of the phenomena under study. The development of computer and information technology opens the door for researchers to apply computer-mediated research. Researchers in social sciences have interviewed people by using phone, skype, and instant messaging software.
for quite some time (Cater, 2011; Deakin and Wakefield, 2014; Jenner and Myers, 2019; Johnson et al., 2019; Sullivan, 2012). The computer-mediated research allows researchers to include informants from different parts of the country to participate in their research. The above-mentioned “non-conventional” data collection methods enable researchers to carry out research without actually going to the field. Before the Covid-19 pandemic, researchers chose computer-mediated research as one of the available methodological options. Teti et al. (2020) state that the pandemic is a social event that is disrupting our social order. The pandemic has thrust researchers back into the armchair, practically and metaphorically. Travel restriction, quarantine requirements imposed by local government for people who come from out of town, and physical distancing make it difficult for researchers to go to the actual geographical field and collect data based on face-to-face interaction with research participants. During the pandemic, conventional data collection methods such as participation observation and focus group discussion can turn into super spreader events. The pandemic closed the door to widely known and practiced conventional data collection. At the same time, it opened the possibility to use, and even maximize, computer-mediated data collection out of necessity.

When online research is compared with conventional research or offline research, some researchers discern the latter as superior due to the face-to-face interaction between researcher and research participants. As a result, online research is seen as “the second-best thing” compared to offline research (Holt, 2010). Offline research does allow researchers to build rapport with research participants more easily. Definition of rapport include “getting along with each other, a harmony with, a conformity to, and affinity for one another” (Seidman 2013: 98) and “conveying empathy and understanding without judgment” (Patton 2015: 458). Research participants can also ask researchers to clarify sentences in the questionnaire or questions posed by researchers. In addition, face-to-face research allows researchers to see nonverbal language in the form of body language and facial expressions of research participants. Due to the importance of intense face-to-face interaction with research participants in qualitative research, online research is deemed unsuitable for qualitative research (Holt, 2010). Despite the perceived weaknesses of online qualitative data collection, several researchers (for example, UNDP, 2018; Zhang & Watts, 2008) demonstrate that data quality from online research is on par with those collected offline.

As of now, nobody knows how long the pandemic is going to last. Meanwhile, researchers have to carry on with their work, with deadlines looming large in their timeline. Many research activities have to be conducted online. Even though discussions on the merits of online data collection are divergent online, it is imperative that researchers consider the strengths and weaknesses of online data collection to provide them with knowledge before they choose which data collection methods they are going to utilize. The purpose of this paper is to outline the methodological reflection of online data collection based on literature review and personal research experience. This paper will cover discussions pertinent to the research field, various online data collection methods, and the ethical aspects of online data collection methods.

---

1 In this paper, research participants refer to respondents in quantitative research and informants in qualitative research.
Redefining the Field

During the course of the research, researchers will perform fielding activities. “Going to the field” and “leaving the field” are inherent activities in field research. Researchers have to leave one’s “home institution to acquire data, information, or insights that significantly inform one’s research” (Kapiszewski et al., 2015: 1). At the completion of the study, they will then leave the field. In pre-pandemic time, field refers to the actual geographical space. In the world of travel restriction and physical distancing, the field is cyberspace (Christia, 2021). For qualitative researchers, the migration to the virtual space is problematic. Research is “based on personal interaction with research [participants] in their own setting” for an extended period (Wood, 2007: 123). The setting provides context – cultural, social, political, economic - for the phenomena under study. The protracted timeline affords an opportunity for researchers to immerse themselves in the daily activities of the study participants as well as their culture. When researchers pivot towards online research, they are inclined to redefine the field. What is, or where is, the field when researchers stay at home or at the office while talking to research participants over the phone or staring at them from the researchers’ computer screen. Beaulieu (2010) states that the concept of location in cyberspace is vague, as researchers have to switch from offline “co-location” to online “co-presence”. Both researchers and participants of research are aware that they are not in the same location, although they can hear or see each other. The realization of the non-existence of “co-location” is emphasized by questions such as: “It’s raining here. Is it raining where you live?” or “Today is extremely hot here. Is it also hot over there?”. In spite of the misgiving of the lack of the “actual field” in online data collection methods, Howlett (2021) found that she was able to be grounded on the research site even though she was in London and her research participants were in Ukraine. Similar earlier research (Beaulieu, 2004; Pink et al., 2016) back Howlett (2021)’s research experience.

Online Quantitative and Qualitative Data Collection Tools

This part of the paper outlines different data collection tools, starting from the very simple one, such as Short Message Service (SMS) which does not require a smartphone to a more sophisticated tool such as Computer Assisted Web Interviewing (CAWI). Since no data collection tool is perfect, this section constitutes the advantages and disadvantages of each tool. Before delving into details of various online data collections tools, it is important to understand the technological context in Indonesia. Table 1 shows the digital divide between rural and urban areas in Indonesia. Access to smartphones, computers, and stable internet connection is greater in urban areas. This means that it is easier for researchers to conduct online data collection in and from urban areas. However, with online data collection methods, researchers are able to cover much wider research area, which otherwise will be too expensive for offline data collection. Online data collection is more economical than offline data collection since it does not constitute funding for travel. Generally speaking, the biggest allocation in any research falls under the category of travel cost.

2 For a more complete description of various online data collection tools see Susilastuti, DH, Abritaningrum, YT & H Murti, SWUH (2020) Penelitian di tengah Pandemi Covid-19: Petunjuk Praktis. Pusat Studi Kependudukan UGM. Unpublished manuscript
Table 1. Information Communication Technology Indicators

<table>
<thead>
<tr>
<th></th>
<th>Rural Area (%)</th>
<th></th>
<th>Urban Area (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Household Access</td>
<td>Individual Usage</td>
<td>Household Access</td>
<td>Individual Usage</td>
</tr>
<tr>
<td>Internet</td>
<td>36.0</td>
<td>26.3</td>
<td>32.5</td>
<td>48.5</td>
</tr>
<tr>
<td>Types of internet access</td>
<td>36.0</td>
<td>-</td>
<td>-</td>
<td>93.3</td>
</tr>
<tr>
<td>a. Mobile broadband</td>
<td>93.3</td>
<td>95.4</td>
<td>-</td>
<td>93.3</td>
</tr>
<tr>
<td>b. Fixed broadband</td>
<td>7.8</td>
<td>7.4</td>
<td>-</td>
<td>14.3</td>
</tr>
<tr>
<td>Handphone</td>
<td>84.4</td>
<td>79.5</td>
<td>70.1</td>
<td>90.7</td>
</tr>
<tr>
<td>a. Smartphone</td>
<td>-</td>
<td>-</td>
<td>59.2</td>
<td>-</td>
</tr>
<tr>
<td>b. Non-smartphone</td>
<td>-</td>
<td>-</td>
<td>61.5</td>
<td>-</td>
</tr>
<tr>
<td>c. Both</td>
<td>-</td>
<td>-</td>
<td>20.7</td>
<td>-</td>
</tr>
<tr>
<td>Computer</td>
<td>31.4</td>
<td>22.1</td>
<td>20.4</td>
<td>43.4</td>
</tr>
<tr>
<td>Fixed phone</td>
<td>4.5</td>
<td>1.4</td>
<td>-</td>
<td>8.5</td>
</tr>
<tr>
<td>TV</td>
<td>87.7</td>
<td>82.6</td>
<td>67</td>
<td>94.2</td>
</tr>
<tr>
<td>Radio</td>
<td>40.0</td>
<td>26.3</td>
<td>20.5</td>
<td>48.5</td>
</tr>
</tbody>
</table>

*Source: MCIT, 2016, p. 5-43 in Hadi (2018)*

Despite the regional restriction due to the digital divide, our research experience show that online data completion and data quality are on par with offline data collection. During the Covid-19 pandemic, many people have been conducting some of their activities online, acquiring technological competence and ease in the process. Many research participants are adept not only at using electronic gadgets but also at sharing their life and thoughts in virtually conducted research. Hence online data collection is proficient at generating good quality, complete and rich data.

Below is the outline and simple description of different data collection tools.

**Types of Online Qualitative and Quantitative Data Collection Tools**

1. **Short Message Service (SMS)**

   Advantage: cheap.
   Disadvantages: determining the sampling frame is quite challenging, it cannot be used for research with complex questions, there is no room to look for nonverbal cues, rapport development is not easy.

2. **WhatsApp messages**

   Advantages: cheap, it can be used to track respondents with high mobility.
   Disadvantages: difficulty in determining sampling frame, researchers cannot observe nonverbal language, not everyone has a smartphone, it depends on mobile phone credit and a strong and stable internet network, rapport development is quite challenging.

3. **Telephone**

   Advantages: there is direct contact with research participants through voice, the researcher can probe, meaning ask for clarification and more detailed answers from research participation, research participants have less inhibition to talk about sensitive matters.
Methodological Reflections on Online Data Collection during the Covid-19 Pandemic

Disadvantages: determining sampling frame is challenging, researchers cannot see nonverbal cues, mobile phone credit cost is relatively high, it is dependent on a strong and stable telephone connection.

4. Online meeting applications (Skype, Zoom, Webex, Google Meet, etc.)

Advantages: researchers and research participants can see nonverbal language through facial expressions, rapport can be built more easily, probing can be done relatively smoothly.
Disadvantages: these applications require a reliable internet network, high cost, the research participants must have the skills to use the selected application, it depends on a reliable internet network.

5. WhatsApp videos

Advantages: researchers and research participants can see each other’s faces so they can see nonverbal language through facial expressions, rapport can be built more easily, probing can be done relatively easily.
Disadvantages: not everyone has a smartphone, it requires a reliable internet network, research participants are charged with mobile phone credit.


Advantages: cheap, data is collected in a data center so time to process data can be reduced.
Disadvantages: not everyone has a computer or internet access, many people are reluctant to fill out questions online, it is dependent on a reliable internet network.


The interviewer reads the questions that have been inputted by a programmer into the application, then he/she enters the research participants’ answers into the program on the computer.
Advantages: there is direct contact through the voice of researcher and research participation, rapport can be built more easily, in the event that research participants do not understand particular questions, they can ask for explanations or clarifications from the researcher, data are more accurate because the interviewer directly enters the answers from research participants into the computer.
Disadvantages: expensive, research participants are charged with mobile phone credit fees if they use WA phone, sometimes research participants think that interviewers are telemarketers, so they are reluctant to participate in the study.

8. Computer-assisted Personal Interviewing (CAPI)

CAPI is very similar to CATI, but CAPI is usually used for offline research. A comparison of the advantages and disadvantages of CAPI, CATI, and CAWI can be seen in the following table.
Table 2. CATI vs CAWI vs CAPI

<table>
<thead>
<tr>
<th></th>
<th>CATI</th>
<th>CAWI</th>
<th>CAPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population coverage</td>
<td>★★★</td>
<td>★★</td>
<td>★★★★</td>
</tr>
<tr>
<td>Cost per interview</td>
<td>HIGH</td>
<td>LOW</td>
<td>MEDIUM-HIGH</td>
</tr>
<tr>
<td>Initial investment</td>
<td>HIGH</td>
<td>LOW</td>
<td>MEDIUM-HIGH</td>
</tr>
<tr>
<td>Reliability of collected data</td>
<td>★★★</td>
<td>★★</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Accuracy of answers</td>
<td>★★★★</td>
<td>★★</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Redemption</td>
<td>★★★</td>
<td>★★</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Fast survey completion</td>
<td>★★</td>
<td>★★★★</td>
<td>★★★</td>
</tr>
<tr>
<td>Rate of dropout</td>
<td>MEDIUM-HIGH</td>
<td>HIGH</td>
<td>LOW</td>
</tr>
<tr>
<td>Good for long questionnaire</td>
<td>★★★★</td>
<td>★★</td>
<td>★★★★★</td>
</tr>
<tr>
<td>Good for complex questionnaire</td>
<td>★★★</td>
<td>★</td>
<td>★★★★★</td>
</tr>
</tbody>
</table>


Some research constitutes a predetermined population, for example, a list of people who reside in specific regions, or who participate in particular government programs. In this case, sampling will be drawn by the researchers who work in the office or from home. The local research assistants who are recruited from the research areas will then contact the people who are selected to participate in the research.

When field assistants contact potential research participants, they will ask for their telephone numbers. At the same time, they will inform the prospective research participants that somebody from our office in Yogyakarta will contact them through the phone. The field assistant will emphasize that the person who will call from Yogyakarta is a research assistant and not a telemarketer.

Another challenge in computer-mediated research comprises of establishing cooperation from the research participants so that they will be willing to complete the interview process. In an attempt to secure the research participants’ cooperation, some researchers provide compensation for them. The subject of compensation is elaborated further in the section of ethical considerations below.

9. Netnography/cyber ethnography/virtual ethnography

Information in various platforms such as Facebook, Instagram, and Twitter, provides rich data for online ethnographic studies. Social scientists, particularly anthropologists, see social networking as a mirror of social life and belief that ethnography can be done online through what is eventually named netnography (Kozinets, 2015). Some researchers question the need for a new name for online ethnographic research. Is the conventional ethnographic study quite different from the online ethnographic research, so that it is deemed necessary to attach a new name to the latter? Lombardi (in Kozinets, 2015: 4) argues:

“The worlds of research and intellectual innovation are strewn with neologisms that might’ve sounded odd or wrong when brand-new: cybernetics, psycholinguistics, soft-ware. So yes, new mappings of reality sometimes call for new names, and sometimes the names take a while to settle in.”
Netnography highlights discussion on the fluidity of the concepts of culture and community. The availability of online archived communication that represents social life, combined with the opportunity to conduct interviews online, has changed the way people do ethnographic research online. Kozinets (2016) observes that ethnographic research constitutes the use of big data and discourse analysis.

Advantages: very diverse and rich data sources. This method can be applied by researchers from various scientific backgrounds.

Disadvantages: researchers must pay attention to relatively complex research ethics issues.

“While analyzing media content, policy documents, and other official public content is straightforward, ethically speaking, content generated online by the public (e.g. forums, blogs, vlogs, reader comments) can be more ethically controversial. The key consideration is what constitutes ‘public’ or ‘private’ online and how might such research be received by those individuals or communities whose content has been used. Researchers should also check if their professional bodies have any specific guidance regarding online data collection. For example, the British Psychological Society has ethical guidelines for internet-mediated research....” (Jowett, 2020)³

Ethical Considerations

Conducting online research during the Covid-19 pandemic is not only related to logistics but also to research ethics. Research ethics are applicable in all kinds of research, online research included, and they include respect for persons, anonymity-pseudonymity, risk or benefits for participants, nondisclosure, conflict of interest, justice, subject compensation (Ess & Ha’rd af Segerstad, 2019). Since online research is devoid of direct interaction with research participants, researchers have to be cognizant of the fact that there is always a “person” who may be affected by the research (Marcham and Buchanan, 2012).

Participation in research is voluntary. Consent from research participants to partake in research activity is very important. Ethically, researchers must explain the purpose of the research, the impacts of the research on the research participants, the non-compulsory nature of their participation, then ask for approval from the research participants to participate in the research. Technically researchers can send a written informed consent form through an e-mail and request a signature from research participants prior to data collection activity. However, this method of obtaining consent is often impractical because not all research participants are familiar with how e-mail works. The alternative to a written consent from research participants will be a verbal consent. Researchers can read from the written informed consent to research participants, ensure that they understand its content, and then ask the research participants: “Can I please start asking questions?”

When people agree to participate in research activity, they set aside some time to answer questions from a researcher. Compensation for their time is deemed appropriate. It should be noted that several studies on the relationship between incentives and increased collaboration with respondents did not show conclusive results (Ballivian, et.al., 2013; Demombynes, et.al., 2013; Hoogeveen et.al., 2014; Leo et.al., al., 2015).

³ Research ethics for internet-mediated research issued by the British Psychological Association can be viewed at this link https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Pages/ResearchEthicsGuidelinesInternetMediatedResearch%202017%20Eng.pdf
Even so, compensation is a nice way for the researchers to show appreciation to the research participants. Researchers should select the type of compensation carefully to prevent undue influence, especially during the pandemic where economic insecurity is on the steep rise (MacKenzie et al., 2007; Head, 2009). What constitutes appropriate compensation is debatable, but Head (2009) suggests it should be useful for the research participants. Based on our research experience, IDR 50,000,- worth of mobile phone credit can be considered as a suitable compensation. Perhaps due to Covid-19 induced economic precarity, some research participants asked the researchers to transfer the IDR 50,000,- to their Ovo or GoPay account. The researchers politely declined the request as sending the money to their OVO or GoPay account will amount to undue influence.

Another ethical consideration that researchers have to adhere to is the commitment to protect the well-being of research participants. Research must not injure or harm the people being studied (Hugman et.al., 2011; Kaplan et.al., 2020; Vankley, 2013). Kaplan et.al. (2020) argue that the no-harm principle should also apply to researchers, especially field assistants. Based on their research in various countries, Kaplan et.al. (2000) find that many research assistants work in environments that do not support their physical and emotional well-being. Bisoka (2020) argues that the pandemic shines a light on the gap of privilege among researchers from developed countries as opposed to those from developing countries. The first cannot, or choose not to, travel to the field research amidst the pandemics. Yet, some of them still expect “people who are already there”, namely researchers from developing countries to carry out offline research in spite of the risk of contracting Covid-19. Some of these “local researchers” or assistant researchers are contract workers with insufficient insurance coverage. The pandemic is an opportunity to reflect on the unequal position of researchers from developed and developing countries. The reflection can be a catalyst for the creation of collaborative agenda-setting that rectifies practical, structural, and labor inequalities (Marks and Zakayo, 2021)

Conclusion

The Covid-19 pandemic has been wreaking havoc in many aspects of our life, both personal and professional. It upended the widely held conventions regarding basic aspects in research, such as the field and data collection methods. When the government implemented various measures, such as travel restriction, reduced mobility and physical distancing to curb the spread of the virus, many researchers were not sure on how to conduct their research activities. Some reluctantly visit, or revisit, online data collection methods. Even though online data collection is not novel, researchers in some quarters still perceive it as “second best” compared to the conventional, face to face, data collection methods. At the end, the reluctance, or sometimes hesitation, to try the online data collection methods was replaced with the curiosity to implement them. After all, as Indonesians, are not we all very familiar with the legendary saying *tidak ada rotan, akarpun jadi* (when there is no rattan, we can use roots)?

Literature review and our personal research experience have provided us with pleasant surprises. Data quality from online research are comparable to those collected offline. Online data collection increases the awareness among research participants that they have to use their time wisely to anticipate dropped call due to poor mobile phone signal. They tend to spend less time thinking about normative answers, hence the complete data from the interview. Moreover, during the Covid-19 pandemic people move some of their offline activities online. Some of them
use a very simple device, such as a non-smartphone to trade news with people in their circle. Other people use more sophisticated gadgets, such as smartphones, tablets, and computers. These groups have more liberty to migrate to the cyberspace and carry out a larger part of their daily activities there. In doing so, they develop ease and mastery in using different gadgets and platforms. These people form a pool of potential research participants. Some of them actually participate in research during the course of the pandemic. Their technological savviness and their ease in sharing their thoughts and feeling explain the quality and completeness of data collected online.

Research activities amidst the pandemic are unique, and they often lead to methodological reflection. The encouraging field results and the positive methodological reflection make the Covid-19 induced uncertainty less daunting. If and when Omicron, the newest Covid-19, enters Indonesia, many researchers are more equipped to carry out their online data collection than they were a couple of years ago.

References


Demombynes, G., Gubbins, P., & Romeo, A. 2013. Challenges and Opportunities of Mobile Phone-Based Data Collection Evidence from South Sudan. The World Bank.


Head, E. 2009. The ethics and implications of paying participants in qualitative
UNDP. 2018. *Below the surface: Results of a WhatsApp Survey of Syrian refugees and host communities in Lebanon.*


**Other source:**