



香港城市大學
City University of Hong Kong

專業 創新 胸懷全球
Professional · Creative
For The World

CityU Scholars

Surveying Spontaneous Mass Protests Mixed-mode Sampling and Field Methods

Yuen, Samson; Tang, Gary; Lee, Francis L. F.; Cheng, Edmund W.

Published in:
Sociological Methodology

Published: 01/02/2022

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY-NC

Publication record in CityU Scholars:
[Go to record](#)

Published version (DOI):
[10.1177/00811750211071130](https://doi.org/10.1177/00811750211071130)

Publication details:
Yuen, S., Tang, G., Lee, F. L. F., & Cheng, E. W. (2022). Surveying Spontaneous Mass Protests: Mixed-mode Sampling and Field Methods. *Sociological Methodology*, 52(1), 75-102.
<https://doi.org/10.1177/00811750211071130>

Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy

Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Surveying Spontaneous Mass Protests: Mixed-mode Sampling and Field Methods

Sociological Methodology

2022, Vol. 52(1) 75–102

© The Author(s) 2022



DOI: 10.1177/00811750211071130

<http://sm.sagepub.com>

Samson Yuen¹ , Gary Tang², Francis L. F. Lee³,
and Edmund W. Cheng⁴

Abstract

Protest survey is a standard tool for scholars to understand protests. However, although protest survey methods are well established, the occurrence of spontaneous and leaderless protests has created new challenges for researchers. Not only do their unpredictable occurrences hinder planning, their fluidity also creates problems in obtaining representative samples. This article addresses these challenges based on our research during Hong Kong's Anti-Extradition Law Amendment Bill Movement. We propose a mixed-mode sampling method combining face-to-face survey and smartphone-based online survey (onsite and *post hoc*), which can maximize sample sizes while ensuring representativeness in a cost-effective manner. Test results indicate that key variables from the survey modes are not statistically different in a consistent manner, except for age. Our findings show mixed-mode sampling can better capture protesters' characteristics in contemporary protests and is replicable in other contexts.

Keywords

protest survey, leaderless movements, mass protests, Hong Kong, Anti-Extradition Law Amendment Bill Movement

Over the past decade, mass protest movements have been a distinctive feature of political life across the globe. From the Arab Spring revolts to the Black Lives Matter movement, people all around the world are calling for social and political changes, including democratic reforms, climate justice, and racial equality. However, unlike previous waves of mass mobilizations that were often led by an organized leadership with a more or less scripted playbook, many contemporary protest movements are characterized by their leaderless, fluid, and spontaneous dynamics (Steinert-Threlkeld 2017; Tufekci 2017). They are often sparked by unpredictable and dramatic events, relying heavily on social media platforms to mobilize, voice their agenda, and garner support.

¹Hong Kong Baptist University, Kowloon, Hong Kong

²The Hang Seng University of Hong Kong, Shatin, Hong Kong

³The Chinese University of Hong Kong, Shatin, Hong Kong

⁴City University of Hong Kong, Kowloon, Hong Kong

Corresponding Author:

Edmund W. Cheng, Department of Public Policy and Centre for Public Affairs and Law, City University of Hong Kong, 83 Main Road, Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong SAR.

Email: ew.cheng@cityu.edu.hk

Once initiated, these protests are then guided on the fly by people whose preferences, motivations, and ideas may not always align, and their course of action often cannot be anticipated (Cheng and Chan 2017; Snow and Moss 2014).

Spontaneous protests are not new, but their frequent occurrence in recent years has created new methodological challenges for researchers who wish to study social movements on the spot, particularly in terms of surveying protest participants. For a long time, many researchers relied on onsite protest surveys as a tool to capture the composition, behavior, and attitudes of protesters (Onuch 2014; Onuch and Sasse 2016; Tufekci and Wilson 2012; Verhulst and Walgrave 2009). However, as popular protests are becoming more unpredictable, it has become increasingly difficult and costly for researchers to plan onsite surveys ahead of an event. Moreover, given that mass protests tend to be fluid and formless because of their decentralized nature, researchers may have more difficulty coming up with consistent sampling frames to obtain relatively representative samples. Furthermore, protestors' often-volatile interactions with the police may create safety issues for survey researchers in the field.

The emergence of the Internet and advanced mobile technologies—and the availability of online or Web-based surveys—presents new opportunities to overcome these methodological challenges. Across social science disciplines, online surveys are used extensively as substitutes or complements of traditional survey methods, such as telephone and face-to-face surveys (Evans and Mathur 2005, 2018). Online surveys certainly have their disadvantages, but the fact that they can be self-administered online, and often through mobile devices, has made them cheaper, faster, more flexible, and more capable of reaching a large population in a short period of time. A growing amount of research shows that online surveys can be as good as traditional surveys or laboratory experiments in terms of collecting high-quality data (Heerwegh 2009; Manago, Mize, and Doan 2021; Nandi and Platt 2017; Rada and Dominguez-Álvarez 2014; Simmons and Bobo 2015). Some research also highlights the unique advantages of online surveys, such as penetrating disadvantaged and hard-to-reach population groups (Keusch et al. 2019; Sugie 2018) and collecting real-time, relational, or sensitive information (Gaggioli et al. 2013; Raento, Oulasvirta, and Eagle 2009).

Can online surveys be used to survey protests, particularly the spontaneous protests described above? How similar are data collected from online surveys compared with data collected from face-to-face surveys? In this article, we evaluate the use of a mixed-mode sampling method, which combines face-to-face and online surveys, in conducting onsite surveys in protests. We draw on our fieldwork during Hong Kong's Anti-Extradition Law Amendment Bill (Anti-ELAB) Movement of 2019, a distinctively leaderless movement that developed spontaneously along its course, in which we conducted 26 onsite protest surveys and collected a total sample size of 16,386. Based on our fieldwork experience and survey data, we argue that a mixed-mode sampling method that combines face-to-face survey and smartphone-based online survey (both onsite and *post hoc*) enables researchers to collect large-enough samples while ensuring representativeness. Test statistics indicate that major variables from the three survey modes are not statistically different in a consistent manner, except for respondents' age—which can be tackled through the mixture of different survey modes. By

balancing the strengths and weaknesses of different modes, mixed-mode sampling can better capture the profiles and characteristics of actors participating in contemporary protest movements.

PROTEST SURVEYS: METHODS AND METHODOLOGICAL REFLECTIONS

Who were the protesters? Why did they protest? And how did they mobilize? To address these core questions about protesting, social scientists often rely on onsite surveys as a key source of information. In contrast to population surveys, which ask a random sample of citizens about their protest participation (along with other information) in general and can only cover a small percentage of participants in particular protests, onsite protest surveys directly target participants in contentious events while they are engaging in the act of protesting. Surveying protesters allows researchers to contextualize the results and design specific questions about individual participation, bridging macro, meso, and micro levels of analysis (Andretta and Della Porta 2014). First adopted by scholars in the 1960s (Boserup and Iversen 1966; Jenkins 1967; Parkin 1968), the method has been used more often since the 1990s as popular protests have increased around the world, but it is still a niche genre in today's social movement research (for a summary, see Walgrave and Verhulst 2011). More sophisticated protest survey designs have been introduced, and more studies have used the method to compare and understand protests across issues, time, and space.¹ Methodological reflections, although still inadequate, do suggest some improvements to how protest surveys should be designed and conducted (Fisher et al. 2019; Walgrave and Verhulst 2011; Walgrave, Wouters, and Ketelaars 2016).

Protest surveys are adapted from traditional exit polls, in which a systematic random sample of voters are interviewed about their voting decisions after they exit the polling stations (Levy 1983). Protest surveys apply a similar approach to protest events. Such surveys typically comprise face-to-face interviews conducted on the spot, mail-in questionnaires distributed at the protest venue, or a combination of both approaches.² Because there is no preexisting sampling frame based on the population (as there is no list or census of participants in a protest event), protest surveys often rely on a spatial sampling frame to ensure the selection of protest participants for interview is systematic and as random as possible, such that everyone has a equal chance of being surveyed. This sampling frame varies from protest to protest and depends on factors such as its form (e.g., moving or stationary) and the geography of its venue, but the most common approach is to arrange interviewers at various locations fairly distributed across the protesting crowd and instruct them to approach every n th protester for interview. This sampling procedure, if strictly enforced, can prevent interviewers from selecting only their approachable peers or making arbitrary decisions. Protests only allow a short time window for face-to-face interviews, so researchers often seek to enlarge sample sizes by distributing mail-in questionnaires at the venue, which some participants will fill in at home and send back via mail.

This general method has been adopted and refined by different teams of researchers over the years, such as Stefaan Walgrave and Joris Verhulst (2011), Stefaan Walgrave and colleagues (2016), Michael T. Heaney and Fabio Rojas (2015), Dana R. Fisher (2020), and Francis L. F. Lee and Joseph M. Chan (2010). Innovations and variations have been added to the basic method depending on the researchers using it and the nature of the protest being researched. Francis L. F. Lee and Joseph M. Chan (2010), for instance, partitioned a protest's gathering venue into numerous areas and instructed interviewers to move along a round trip within each of them, sampling one cluster of six respondents every 10 minutes. After the protest began, interviewers were redeployed along the marching route to capture participants who joined en route, interviewing the person who came closest to them one minute after a completed interview. Michael T. Heaney and Fabio Rojas (2007, 2015:12) propose an anchor-sampling method in which interviewers are instructed to first identify an arbitrarily selected anchor person from the crowd (who is not invited to take the survey) and then count five persons in a line from the anchor and invite that person for interview. The counting process is repeated until three surveys are accepted, after which the interviewer selects a new anchor and restarts the counting procedure. Stefaan Walgrave and Joris Verhulst (2011), in their rigorous methodological note, propose a sampling procedure to systematically survey a moving protest crowd that involves splitting the task of selecting respondents and interviewing them. The former task is handled by field supervisors, who stand at the start and end of a moving possession and select every n th row and every n th person in that row. Upon selecting a potential respondent, interviewers are tasked to either hand out a questionnaire to that individual or orally interview the person on the spot. To ensure a fair dispersion of questionnaires among the protest crowd, the selection is conducted alternately to cover different positions within a row. This helps to structurally prevent interviewers from making intuitive or arbitrary choices of respondents, which effectively reduces selection bias.

A systematic and random selection of respondents is crucial, but a high response rate is also necessary to ensure representativeness of the sample. If the response rate is low, it means the sample may have systematically excluded participants with some specific characteristics, thus resulting in the problem of response bias. In many protests, the response rate of face-to-face interviews tends to be quite high (given that people who attend protests are by default more willing to express their thoughts), but it is affected by a broad range of factors. For instance, longer questionnaires tend to result in lower response rates (Rüdig 2010), and so do chaotic protests (Walgrave et al. 2016). Black respondents consistently have lower response rates than do other racial groups, and women generally yield higher response rates than men (Heaney and Rojas 2015). Ideology of the protest also matters, with participants in conservative protests usually being less responsive than those in liberal ones (Heaney 2018). Even the identity of the interviewer can affect responses: Research shows participants are more responsive to interviewers who are female and older in age (Walgrave et al. 2016).

The problem of survey nonresponse is more acute for mail-in questionnaires because respondents are required to mail back their responses. Failure to do so results

in “delayed refusal bias.” Indeed, not everyone given a paper survey makes such an effort. Stefaan Walgrave and Joris Verhulst (2011) find that older and female respondents are more likely to return their questionnaires; apart from that, they find small differences between the face-to-face and mail-in samples, which suggests mail-in surveys are not statistically biased toward specific respondents. Using follow-up surveys rather than mail-in questionnaires, Dana R. Fisher (2020) arrives at a similar conclusion that there are no statistically significant differences between those who are surveyed onsite and those who reply through follow-ups. However, a further study by Stefaan Walgrave and colleagues (2016) detects some delayed refusal biases in survey responses: protesters who are highly educated, who have higher political efficacy, and who are mobilized through mass media are more likely to mail back surveys. Despite the systematic refusal biases caused by individual characteristics, the silver lining is that responses are not affected by protest features, issues, or countries. This finding provides a credible basis for cross-national comparison of survey evidence on protests.

In summary, by calling attention to various biases, these methodological reflections have helped researchers improve the validity and representativeness of the data collected from protest surveys. They also provide a strong rationale to apply survey weights in mitigating such biases (Heaney and Rojas 2015). However, the evolving forms of mass protesting have also generated new challenges. As mass protests have become increasingly spontaneous and fluid, in part due to their leaderless and networked nature, researchers have more difficulty following the systematic sampling procedures discussed above. First, when protest crowds do not follow a fixed or preannounced route, it becomes more difficult to use a systematic spatial frame and ensure an equal dispersion of questionnaires among participants. And given the potential volatility of the protest and physical conditions of the protest site, it may be technically impossible for researchers to find anchors and follow counting procedures, or for field supervisors to point to potential respondents whom interviewers will separately approach. In the case of Hong Kong, for instance, streets were often packed with protesters during large rallies or demonstrations, making it difficult to identify anchors and deploy survey helpers. Second, the often sudden eruption of protests means it is difficult and costly for researchers to quickly mobilize a large and well-trained survey team. Also, face-to-face interviews are time-consuming. Interviewers need a lot of time to introduce themselves, explain what the survey is for, and read the questions. This may result in small sample sizes, which undermines researchers’ ability to make causal inferences. Third, given the declining importance of postal communication, it may be impractical to expect protest participants to send back their survey responses via postal mail. This, too, may reduce the sample size and may result in delayed refusal bias.

Some of these challenges are, to a large extent, unavoidable. For instance, the fact that mass protests have become more fluid and irregular inevitably undermines researchers’ ability to sample protest participants in a systematic fashion. Others, however, may be more remediable. Thanks to technological advancements, researchers have introduced new designs to adapt to these challenges. Dana R. Fisher (2020), for example, used electronic tablets instead of paper-based questionnaires to administer

onsite surveys during anti-Trump protests. Barbara L. Ley and Paul R. Brewer (2018) similarly used tablet-based questionnaires but combined them with paper-based ones in their research on the March for Science protest in the United States. More recently, Joost de Moor and colleagues (2020) used online surveys to study participants of the Fridays for Future climate protests by distributing flyers with a QR code. In contexts where onsite fieldwork is not possible, researchers have used *post hoc* online surveys posted on popular social media platforms to survey protesters after their participation (Grewal, Kilavuz, and Kubinec 2019; Hassanpour 2016; Ketchley 2017). These *post hoc* online surveys are less ideal for researchers to obtain a valid and representative sample of protest participants, but onsite tablet-based or online surveys can expedite data collection and analysis, particularly in spontaneous and fluid protests. These surveys are much faster than face-to-face ones and can be cheaply administered, thus providing a reliable and effective way to increase sample size. However, despite the increasing popularity of these approaches, there has been no serious methodological discussion regarding the use of online means to conduct protest surveys and whether it may generate biases. Researchers have also not explored the potential of different online-based approaches to expand the sample size of protest surveys. Nor have they devised ways to combine online surveys with pen-and-paper ones to balance their respective strengths and weaknesses.

HONG KONG'S ANTI-EXTRADITION BILL MOVEMENT

Hong Kong's Anti-ELAB Movement provided a valuable opportunity to apply online surveys alongside traditional face-to-face ones and test their reliability. Sparked by a government proposal to amend the semi-autonomous city's extradition laws, which would have allowed suspects to be extradited to Mainland China, the movement became the largest and most intense protest mobilization the city has ever seen. Starting from around March 2019 and continuing into early 2020, it consisted of hundreds of protest actions across the territory, some of which included tens of thousands of protesters. The protests on June 9 and June 16 reportedly drew one million and two million protesters, respectively, out of a population of seven million. Besides their occasionally enormous turnout, the protests also radicalized along their course. From August onward, many protests, despite their peaceful beginnings, would escalate into violent confrontations between protesters and the police. The police would quell the protests with heavy-handed tactics using tear gas, bean bag rounds, rubber bullets, and water cannons; protesters would defend themselves with umbrellas and gas masks, sometimes fighting back with sticks, bricks, and Molotov cocktails (Cheng et al. 2022).

Surveying the Anti-ELAB Movement presented many fieldwork challenges similar to those described above. First, except the mass rally on June 9, which was planned more than two weeks ahead, many of the subsequent protests were announced only a few days before they were held (partly because the authorities were reluctant to issue permits for protests to take place legally). Some protests were unplanned and spontaneous, initiated among informal networks and scaled up rapidly because they caught attention on social media. Some semi-spontaneous protests were relatively organized

and informed by collective identity at the beginning, but later became guided by autonomous and contingent elements (similar to Northern Ireland's "Bloody Sunday" protest, as described in Snow and Moss 2014:1131 and Spain's Indignados Movement, as described in Cristina Flesher Fominaya (2015:145). All these factors made survey planning difficult, leaving researchers with limited time to mobilize a large-enough team of well-trained survey helpers.

Second, given their massive turnouts, these protests often featured jam-packed crowds on streets and in parks, which left little space for interviewers to move freely and use established sampling techniques, such as anchor-sampling. Some of the protests did not have a preannounced route. Such fluidity rendered some existing survey methods (e.g., anchor-sampling and pointer-sampling) impractical, thus undermining systematic collection of survey data. Third, starting from late July, the Hong Kong police began to ban large protests by issuing Letters of Objection.³ This increased the difficulty for protest organizers to plan events, which compounded the fluidity of the protests, and it also meant some of the protests we surveyed became technically illegal. This required the research team to take more precautionary measures. Finally, the violent escalation that occurred in the latter half of the movement created further challenges. Not only did it raise safety concerns for the research team, it also meant the time frame for data collection in the protests would be cut short, thus significantly limiting the sample sizes.

In summary, these challenges made it difficult for researchers to adhere to traditional face-to-face surveys.⁴ The responsiveness of mail-in surveys was also questionable, given the limited reliance on postal services in the local context and the contentious social atmosphere resulting from the frequent occurrence of protests. Innovations were thus needed to overcome these practical issues and to ensure the survey samples were both sizable enough and representative of the protest crowds. Online surveys present an opportunity to tackle these challenges. From a methodological standpoint, the use of online surveys also provides an opportunity to test their validity and comparability to standard protest surveys.

METHOD AND DATA

Our research team conducted onsite surveys during 26 protests throughout the Anti-ELAB Movement. Due to resource constraints, we could not cover every protest (e.g., the Sheung Shui protest on July 13 and the Mongkok protest on August 3 are among the notable "missing" cases). We had to be selective based on our judgment of the scale, relative importance, and riskiness of the protests. Our aim was to cover the major protests that were announced at least two days before their occurrence and that were expected to feature large turnouts and provide a relatively safe environment for the research team to conduct fieldwork. The protests surveyed can be categorized into three types. Type I were mass rallies that began with protesters assembling at a specified point before following a planned route to a designated destination. Type II were static demonstrations that involved the gathering of protesters at a specific location. Type III were fluid demonstrations where protesters first gathered at a specific location but then became more mobile and unpredictable, spontaneously moving to various locations like wildcat strikes. Table 1 summarizes the basic features about each survey.

Table 1. Summary of Protest Surveys

Date	Venue	Form	Sample Size	Paper	Smartphone		Response (Paper, %) ^a	Response (Post Hoc, %) ^b
					Onsite	Post Hoc		
June 9 (2019)	Victoria Park—Central Government Office (CGO)	Mass rally (I)	285	285	N/A	N/A	74.0	N/A
June 12	CGO	Fluid demonstration (III)	175	90	85	N/A	N/A	N/A
June 16	Victoria Park—CGO	Mass rally (I)	875	610	265	N/A	89.0	N/A
June 17	CGO, Police HQ, and Revenue Tower	Fluid demonstration (III)	721	101	620	N/A	91.5	N/A
June 21	CGO	Fluid demonstration (III)	316	117	199	N/A	87.8	N/A
June 26	Edinburgh Place	Static demonstration (II)	418	254	164	N/A	90.7	N/A
July 1	Victoria Park—CGO	Mass rally (I)	1169	434	735	N/A	83.1	N/A
July 14	Shatin	Mass rally (I)	546	264	282	N/A	87.8	N/A
July 21	Victoria Park—Southorn Playground	Mass rally (I)	680	385	295	N/A	90.8	N/A
July 27	Yuen Long ^c	Mass rally (I)	235	N/A	N/A	235	N/A	13.1
August 4	Tseung Kwan O	Mass rally (I)	713	131	98	484	85.6	17.0
August 04	Sai Wan	Static demonstration (II)	559	100	146	313	92.7	16.9
August 10	Airport ^e	Static demonstration (II)	2309	73	N/A	2236	95.0	N/A
August 11	Shum Shui Po ^c	Mass rally (I)	636	82	112	442	90.1	14.3
August 13	Airport ^e	Static demonstration (II)	485	N/A	N/A	485	N/A	N/A
August 16	Chater Garden	Static demonstration (II)	632	N/A	N/A	632	15.8	N/A
August 18	Victoria Park—CGO ^c	Static demonstration (II)	806	126	67	613	82.8	18.3
August 25	Kwai Fong—Tsuen Wan	Mass rally (I)	372	68	7	297	91.5	9.3
August 31	Central and Sheung Wan ^c	Fluid demonstration (III)	527	N/A	N/A	527	13.2	N/A
September 8	Chater Garden—U.S. Consulate	Mass rally (I)	337	74	32	231	94.6	14.4
September 15	Causeway Bay, Wan Chai, and Central ^f	Fluid demonstration (III)	911	178	209	524	83.8	13.1
September 28	CGO	Static demonstration (II)	557	N/A	N/A	405	N/A	10.1
October 1	Causeway Bay, Wan Chai, and Central ^f	Fluid demonstration (III)	640	N/A	N/A	640	N/A	13.3
October 14	Chater Garden	Static demonstration (II)	662	N/A	N/A	662	N/A	16.6
October 20	Tsim Sha Tsui ^c	Fluid demonstration (III)	921	N/A	N/A	921	N/A	19.2
December 8	Victoria Park—Central	Mass rally (I)	902	142	145	615	88.0	15.1
January 1 (2020)	Victoria Park—Central	Mass rally (I)	1306	266	297	743	84.4	18.6

^aN/A means either paper-based surveys were not used or the response rate was not recorded (as on June 12).

^bThe *post hoc* response rate is calculated by the number of responses received from smartphone-based *post hoc* surveys divided by the number of leaflets distributed. N/A means smartphone-based *post hoc* surveys were not used.

^cA Letter of Objection was issued by the Hong Kong police for the protest.

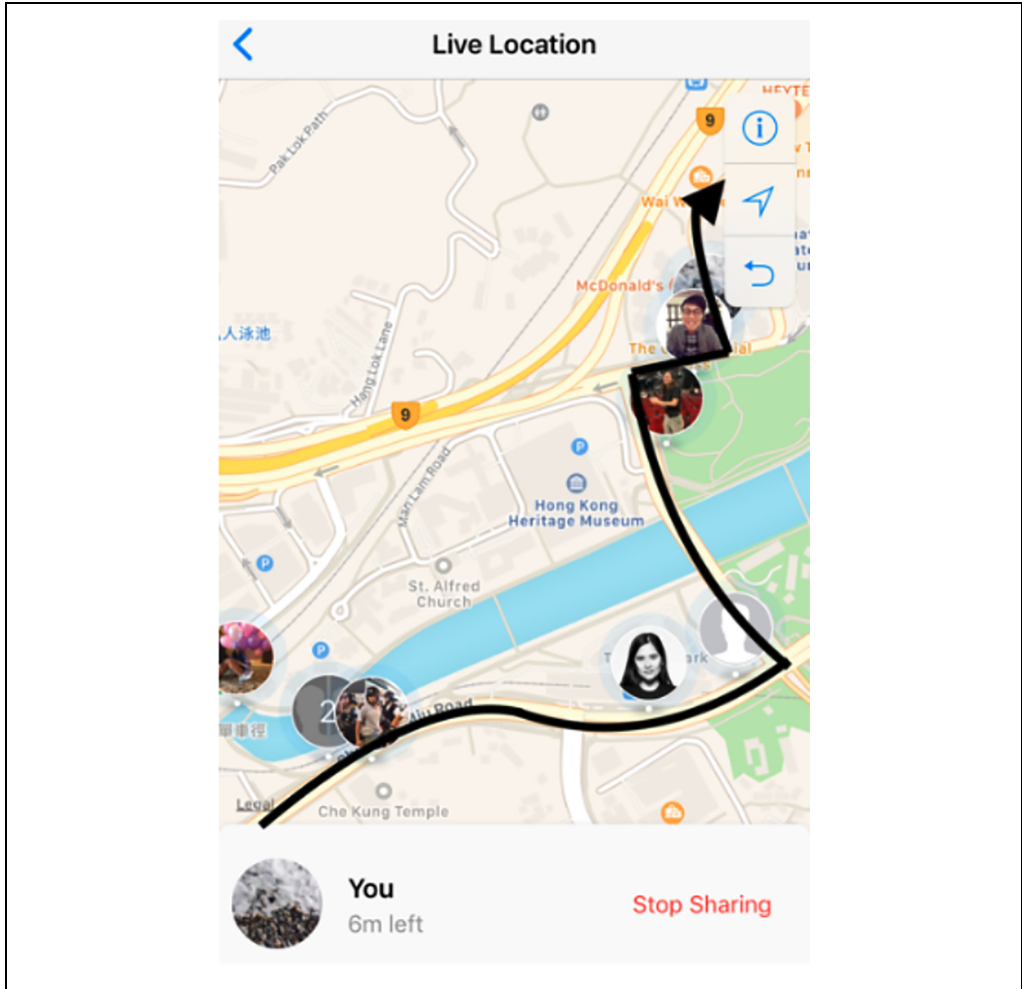


Figure 1. WhatsApp live location in the July 14 protest.

Note: The arrow path indicates the protest route on that day. The profile pictures indicate the live locations of the survey helpers during the protest. This information allowed survey supervisors to redeploy interviewers along the protest route so they could be evenly spread out.

Generally, interviewers were asked to invite every 10th person they saw within a designated area or route to complete the survey.⁵ If they were turned down, they would then ask the next 10th person, and so on. However, because each protest had unique spatial dynamics, it was necessary to make slight adjustments in the sampling method. In each protest, the fieldwork team was led by two to three supervisors. The number of interviewers ranged from 10 to 25, which varied according to the expected turnout.⁶ Interviewers were asked to switch on their live location function on WhatsApp so their actual location could be tracked (see Figure 1).⁷ This was to avoid clustering of interviewers and facilitate their redeployment to new locations if needed. To avoid interviewers being identified as protesters, we gave them hanging badges that showed the

institutional affiliation of the research team, and we requested they not dress in the same color as the protesters, who were usually in black. Whenever protests turned confrontational with the police, we would stop the survey and ask interviewers to go to safe locations.

For the protests in June and July, except for June 9, the survey team interviewed protest participants through two modes: (1) face-to-face survey using a paper-based questionnaire; and (2) smartphone-based, online survey using a self-administered, Qualtrics-based questionnaire linked to a QR code.⁸ The second mode was added after the June 9 protest to obtain larger samples, given the limited size of the survey team. The survey team was normally divided into two groups: around two thirds of the team would approach potential respondents with the first mode, and around one third would approach with the second mode. The logic behind this allocation was to ensure a large-enough group of face-to-face respondents so we could compare them with the smartphone-based respondents because face-to-face surveys were much more labor intensive. Interviewers responsible for the face-to-face survey would read out each question and then fill in the questionnaire together with the interviewee. Face-to-face interviews typically lasted around 10 minutes (on average, this was longer than surveys filled out by respondents on the smartphone). For both modes, interviewers recorded when a person declined to complete a survey.

The smartphone-based surveys were effective in enlarging sample sizes, largely because they were self-administered and could be completed by respondents on their own capacity without the help of interviewers. From the June 12 to July 21 protests, responses from this mode account for 51 percent of the survey sample size, on average. However, despite their cost-effectiveness, smartphone-based questionnaires have their limitations. First, although Hong Kong has one of the highest smartphone penetration rates in the world,⁹ not everyone in the protests had a smartphone or was good at using a smartphone to complete surveys. Such protesters, who were typically older in age, were more likely to decline survey requests. Even if they started the online survey with the help of interviewers, they were more likely to drop out along the process. Second, because protesting has become an increasingly risky activity under Hong Kong's repressive political climate, some participants worried that completing an online questionnaire through their own phones might leave behind digital footprints and bring legal consequences. These two limitations may systematically exclude a portion of protest participants, thus resulting in response biases. A final limitation is that smartphone-based questionnaires depend on good Internet connection. In large and crowded protests, Internet connection can be slow and unstable. Signals may jam or may be disrupted by the authorities.

Face-to-face, paper-based surveys allowed us to counterbalance against these limitations. Although they require substantially more time to be completed, the upside is that they will not exclude people who are not used to technology or who are worried about the risk of leaving digital footprints. Moreover, they are not affected by the quality of the Internet connection. Use of on-the-spot, paper-based surveys can thus ensure that at least a portion of the sample is representative of protest participants, compared with

the sample collected through online surveys. As such, they offer a representative snapshot of the protester population in each protest.

However, as Stefaan Walgrave and Joris Verhulst (2011) warn, selection bias can be a problem in protest surveys. This occurs when interviewers systematically (with or without intent) exclude people with certain characteristics from the survey. For instance, young interviewers may be inclined to invite young protesters to complete the survey, whom they assume to be less likely to turn them down. Interviewers are also generally more inclined to talk to friendly-looking people than unfriendly ones. To mitigate selection bias, Stefaan Walgrave and Joris Verhulst (2011) suggest having fieldwork supervisors serve as pointers and select interviewees. We see the merits of this method, but it was not practical in Hong Kong's Anti-ELAB protests. The method requires a large team of fieldwork supervisors, which is difficult to put together, given the insufficient planning time. More fundamentally, protests in Hong Kong are often very large in scale, involving hundreds of thousands of people in jam-packed streets where it is difficult to move freely. It is virtually impossible for pointers to systematically select interviewees from the protest crowd. For these reasons, we decided to delegate selection to the interviewers. To ensure interviewers would faithfully follow the sampling procedures, we provided a training session before each protest to brief them about the questions and sampling procedures. We also preferred to recruit experienced interviewers. In addition, we imposed a quota limiting the maximum number of samples they could collect within an hour (around 6 respondents for face-to-face surveys and around 15 for online surveys), so they would space out interviews according to the procedures rather than doing them on groups of protesters.

We added a third survey mode—*post hoc* online surveys self-administered through smartphone—using the same questionnaire starting from the Yuen Long protest on July 27. The Yuen Long protest was the first protest since June that received a Letter of Objection from the police, which made the protest “illegal.” The protest also took place a week after pro-government thugs launched an indiscriminate attack against citizens in Yuen Long who returned from a protest. Due to safety concerns for the survey team, we changed our sampling approach from onsite face-to-face and online surveys to simply *post hoc* online surveys. We did so by dispatching four experienced interviewees to distribute flyers printed with a QR code through which protesters could access the online questionnaire with their smartphones. The survey had to end prematurely due to the violent escalation of the protest, but by noon the next day, we had collected 235 responses from the 1,800 flyers distributed (i.e., 13-percent response rate). In light of this promising result and the violent escalation of the protests, we decided to combine this survey mode with the other two modes in subsequent surveys. Whenever protests were banned by the police or were expected to become violent, the research team would solely rely on *post hoc* online surveys for data collection, similar to the July 27 protest.

The rationale of inviting protesters to complete a *post hoc* online survey through distributing flyers was the same as the mail-in survey. Similar to mail-in surveys, *post hoc* surveys rely on whether protesters who are given a flyer will take the initiative to complete the survey on their own. The advantage with the online survey tool is that

interviewees incur almost no cost in submitting completed questionnaires, in contrast to mail-in ones. Thus, it provides an effective channel to maximize the sample without exposing interviewers to excessive risks. From the August 4 to January 1 protests, we combined all three survey modes during five protests. On average, responses from *post hoc* online surveys account for 65 percent of the overall sample. Besides enlarging the sample size in hostile protest conditions, *post hoc* surveys may also help reduce the selection bias caused by interviewers inviting more approachable respondents to complete the survey. By simply distributing flyers instead of needing to explain the survey to potential respondents, interviewers are less likely to pick and choose their recipients.

Responses from the paper-based surveys were entered into an online database by interviewers in 24 hours; the cut-off time for the onsite and *post hoc* online surveys was set at noon the next day. After the data were collected in each survey, the field supervisor cleaned the data according to the following procedure. For the paper-based survey data, the supervisor conducted a basic cleaning exercise by randomly checking if some of the values matched the answers from the paper surveys. For smartphone-based onsite surveys, the cleaning procedure comprised three steps. The first step was to check the start time of the surveys. Interviewers were instructed to invite interviewees by every 10 protesters passing by, so we assumed every interviewer needed to spend at least one minute after finishing an interview to invite another protester to conduct the survey. Based on this assumption, for each interviewer, if more than one respondent started the survey within the same minute, only the first would be counted. The second step was to check the IP addresses of the devices respondents used to conduct the survey. IP addresses are a unique identification of computing devices. If there were repeated IP addresses, it meant more than one survey was conducted by the same device. When this occurred, we only retained the first response for every duplicated IP address.¹⁰ The third step was to check the time respondents took to complete the survey. According to our pilot survey conducted to test our questionnaire, a reasonable interviewee would need at least five minutes to complete the survey. Responses completed in less than five minutes were thus considered to have questionable quality and were excluded from the sample.

We followed a similar cleaning procedure for the *post hoc* online survey. We checked the data for their completion time and whether there were duplicate IP addresses, and we removed unqualified entries. Field supervisors verified whether respondents actually participated in the relevant protest because the flyers and QR code might be circulated to people who did not join the protests. To do so, we relied on an additional screening question to the *post hoc* survey questionnaire, which asked protesters about the location at which they started the protest (there was an option of “Did not attend the protest”). Respondents who clicked this option were removed from the sample. Table 2 shows a summary of the cleaning procedure of the three modes of data collection. A total of 6,949 responses were excluded after cleaning the sample. On average, 25 percent of responses were removed from each survey.

We continued to conduct these onsite surveys for six months, so we had the opportunity to improve our survey methods over time and identify some best practices. For instance, after adopting QR code online surveys, we could gradually identify the

Table 2. Cleaning Procedures for Different Survey Modes

	Paper-based Onsite Survey	Smartphone-based Onsite Survey	Smartphone-based <i>Post Hoc</i> Survey
Correctness of data input	✓		
Check IP address			✓
Check start time		✓	
Check response duration		✓	✓
Check protest participation			✓

optimal number of interviewees to be recruited to maintain a large-enough sample size because we could estimate the approximate number of respondents obtained from the online surveys for each interviewer. We also learned about the response rate for online surveys, so we could estimate the approximate amount of leaflets we should print out and distribute to protest participants. Over time, we were able to build a large contact list with more than 30 experienced interviewers, which allowed us to readily recruit for planned or spontaneous protests. In a few instances, we removed interviewers from the list because they did not abide by our sampling procedures (e.g., conducting too many interviews within an hour, which runs the risk of undermining sampling quality).

MODE EFFECTS AND RESPONSE BIASES

An important task is to investigate whether online surveys draw respondents who are systematically different from those who take paper-based surveys. To test for “mode effects,” we follow Walgrave and colleagues’ (Walgrave and Verhulst 2011; Walgrave et al. 2016) suggestion to adopt the Mann-Whitney U test and compare the distribution of various variables across different survey modes.¹¹ For our surveys, we identified 12 variables that were most commonly asked across the questionnaires and could serve as anchors to check whether and how respondents differ across survey modes. These variables include basic demographics: age, gender (0 = female; 1 = male), education level (1 = primary education or below; 2 = secondary education; 3 = university or diploma; 4 = graduate school), and perceived socioeconomic status (1 = upper class; 2 = middle class; 3 = lower class); political behavior and attitude: frequency of participating in past protests (0–7), political inefficacy (0 = government will offer concessions after protests; 1 = government will not offer any concessions after protest), and whether respondents self-identified as a localist¹² (0 = nonlocalist; 1 = localist); frequency of using particular media as sources of information about the protests (1 = not at all; 5 = very frequent): traditional media, online media, Facebook, Telegram, WhatsApp, and online forums; and finally their views on four statements about radicalization (1 = very much oppose; 5 = very much agree). Table 3 presents the mean values for each variable for each survey mode.

Table 4 presents the Mann-Whitney U test results for surveys conducted between June 16 and July 21 to compare the paper-based and smartphone-based survey mode. Statistical significance is tested on three levels: 5 percent, 1 percent, and 0.1 percent.

Table 3. Comparison of Mean Values across Protest Surveys

Survey mode ^a	6/16		6/17		6/21		6/26		7/1		7/14		7/21	
	0	1	0	1	0	1	0	1	0	1	0	1	0	1
Age	33.17	26.49	30.53	24.23	24.94	21.74	27.03	26.74	36.15	28.53	32.87	27.27	32.03	28.38
Gender (male)	0.51	0.53	0.50	0.44	0.56	0.45	0.43	0.43	0.53	0.52	0.57	0.54	0.52	0.48
Education	2.67	2.69	2.95	2.83	2.81	2.77	2.96	2.93	2.76	2.85	2.73	2.88	2.74	2.88
Socioeconomic status	2.42	2.49	2.47	2.54	2.49	2.53	2.54	2.46	2.40	2.44	2.23	2.32	2.26	2.29
Past protest frequency	NA	NA	2.31	1.88	1.70	1.58	2.12	2.10	1.79	1.81	2.41	2.10	1.93	1.88
Localist	0.18	0.21	0.26	0.25	0.26	0.30	0.42	0.37	0.26	0.28	0.28	0.35	0.29	0.29
Political inefficacy	NA	NA	0.44	0.40	0.41	0.52	0.54	0.50	0.58	0.56	NA	NA	0.61	0.50
Traditional media	NA	NA	NA	NA	NA	NA	3.32	3.48	3.68	3.75	3.44	3.91	3.52	3.81
Online media	NA	NA	NA	NA	NA	NA	4.44	4.51	4.16	4.46	4.31	4.67	4.30	4.68
Facebook	NA	NA	NA	NA	NA	NA	4.29	4.35	4.00	4.37	4.01	4.51	4.07	4.53
Instagram	NA	NA	NA	NA	NA	NA	3.73	3.71	3.05	3.68	3.02	3.68	2.99	3.72
WhatsApp	NA	NA	NA	NA	NA	NA	3.70	3.62	3.26	3.73	3.38	3.96	3.32	3.72
Telegram	NA	NA	NA	NA	NA	NA	3.60	3.57	2.37	2.61	2.94	3.24	2.77	3.23
LIHKG forum ^b	NA	NA	NA	NA	NA	NA	4.42	4.02	3.25	3.77	3.78	4.13	3.54	4.22
Radical actions are justifiable if the government fails to listen	NA	NA	3.44	3.46	3.57	3.68	3.55	3.58	3.20	3.25	NA	NA	3.55	3.62
Radical actions will alienate ordinary citizens	NA	NA	3.62	3.61	3.46	3.63	3.37	3.39	3.46	3.54	NA	NA	3.07	3.17
Radical actions are best when combined with peaceful ones	NA	NA	4.15	4.27	4.39	4.39	4.33	4.39	4.01	4.19	NA	NA	4.22	4.53
Radical actions are understandable	3.89	3.87	NA	NA	NA	NA	NA	NA	4.32	4.41	NA	NA	4.59	4.74
	7/27		8/4		8/11			8/18			8/25		8/31	
Survey mode ^a	2	0	1	2	0	1	2	0	1	2	0	1	2	2
Age	31.45	31.94	29.04	29.91	31.83	26.83	29.45	30.65	27.75	32.44	26.74	23.57	32.25	28.99
Gender (male)	0.57	0.58	0.50	0.51	0.59	0.50	0.54	0.46	0.40	0.45	0.57	0.57	0.50	0.51
Education	2.99	2.83	2.85	2.93	2.71	2.97	2.82	2.85	2.85	2.96	2.85	2.86	2.90	2.98
Socioeconomic status	2.43	2.15	2.21	2.18	2.17	2.29	2.27	2.45	2.44	2.41	2.29	2.43	2.17	2.27
Past protest frequency	NA	1.98	1.92	1.87	2.02	2.03	1.92	1.79	1.60	2.04	1.96	1.29	2.03	1.98
Localist	0.37	0.39	0.36	0.37	0.23	0.48	0.37	0.26	0.30	0.28	0.46	0.29	0.35	0.38
Political inefficacy	0.36	0.17	0.18	0.22	0.17	0.13	0.10	0.13	0.18	0.11	0.10	0.00	0.14	NA
Traditional media	3.72	3.58	3.70	3.83	3.40	3.69	3.70	3.75	3.75	3.83	3.68	3.43	3.77	3.81
Online media	4.76	4.65	4.84	4.80	4.52	4.81	4.77	4.54	4.64	4.74	4.65	4.43	4.70	4.77
Facebook	4.61	4.35	4.59	4.55	4.11	4.70	4.50	4.17	4.57	4.51	4.53	4.14	4.44	4.43
Instagram	3.43	3.57	3.72	3.54	3.18	3.96	3.37	3.44	3.99	3.31	3.57	4.29	3.47	3.61
WhatsApp	4.12	3.86	3.90	3.93	3.70	4.06	3.87	3.79	4.21	4.01	3.60	3.71	3.99	3.79
Telegram	3.42	3.43	3.39	3.56	3.48	3.98	3.79	3.14	3.18	3.41	3.94	3.57	3.62	3.79
LIHKG	4.18	4.06	4.28	4.23	4.07	4.54	4.40	3.98	4.27	4.04	4.24	4.29	4.11	4.36
Radical actions are justifiable if the government fails to listen	3.90	3.43	3.52	3.48	3.36	3.51	3.59	3.44	3.59	3.48	3.60	4.14	3.64	NA
Radical actions will alienate ordinary citizens	3.14	3.13	3.22	3.17	3.17	3.45	3.26	3.17	3.24	3.25	3.00	3.14	2.97	NA
Radical actions are best when combined with peaceful ones	4.55	4.44	4.53	4.53	4.32	4.50	4.56	4.35	4.42	4.44	4.37	5.00	4.55	NA
Radical actions are understandable	4.79	4.70	4.73	4.75	4.67	4.70	4.76	4.62	4.58	4.68	4.63	4.86	4.75	NA

(continued)

Table 3. (continued)

Survey mode ^a	9/8		9/15		9/28		10/1		10/14		10/20		12/8		1/1			
	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
Age	34.20	30.94	31.61	26.42	30.46	32.92	37.30	38.91	32.77	34.65	30.90	26.12	31.97	32.18	29.03	35.12		
Gender (male)	0.45	0.38	0.50	0.53	0.50	0.51	0.49	0.55	0.47	0.48	0.51	0.54	0.51	0.49	0.49	0.49		
Education	2.99	2.97	2.94	2.85	2.85	3.01	2.98	3.02	2.98	2.97	2.72	2.85	2.89	2.71	2.87	2.94		
Socioeconomic status	2.33	2.61	2.48	2.57	2.45	2.43	2.34	2.32	2.30	2.36	2.46	2.46	2.37	2.43	2.50	2.36		
Past protest frequency	NA	NA	NA	NA	NA	NA	2.52	2.38	2.18	2.07	1.75	1.63	1.86	1.93	1.90	2.00		
Localist	0.43	0.44	0.37	0.43	0.38	0.35	0.35	0.31	0.39	0.32	0.35	0.48	0.37	0.33	0.37	0.32		
Political inefficacy	NA	NA	NA	0.40	0.29	0.41	NA	NA	NA	NA	NA	NA	NA	NA	0.41	0.47		
Traditional media	3.46	3.69	3.58	3.80	3.72	3.71	3.72	3.80	3.68	3.68	3.61	3.84	3.66	3.34	3.42	3.51		
Online media	4.59	4.75	4.74	4.85	4.60	4.76	4.78	4.73	4.77	4.76	4.65	4.77	4.75	4.24	4.34	4.35		
Facebook	4.27	4.63	4.54	4.67	4.39	4.40	4.44	4.47	4.55	4.41	4.39	4.29	4.31	4.41	4.58	4.62		
Instagram	3.47	3.72	3.44	4.20	3.80	3.34	2.97	3.00	3.44	3.28	3.97	4.15	3.61	4.11	4.22	4.16		
WhatsApp	3.58	3.88	3.93	4.14	3.88	3.75	3.80	4.04	4.12	3.90	3.90	3.95	3.81	3.46	3.67	3.16		
Telegram	3.72	3.72	3.86	4.21	4.00	3.90	3.70	3.56	4.10	3.74	4.15	4.45	3.95	3.73	3.71	3.62		
LIHKG	4.24	4.31	4.30	4.47	4.32	4.28	4.03	3.98	4.19	4.19	4.04	4.44	4.10	3.72	3.85	3.67		
Radical actions are justifiable if the government fails to listen	NA	NA	NA	3.89	3.59	3.75	NA	NA	3.87	3.79	3.77	4.07	3.79	NA	NA	NA		
Radical actions will alienate ordinary citizens	NA	NA	NA	2.98	2.85	3.03	NA	NA	3.22	3.04	2.80	2.97	2.93	NA	NA	NA		
Radical actions are best when combined with peaceful ones	NA	NA	NA	4.60	4.39	4.59	NA	NA	4.60	4.60	4.43	4.53	4.55	NA	NA	NA		
Radical actions are understandable	NA	NA	NA	4.81	4.68	4.76	NA	NA	4.85	4.83	4.72	4.85	4.86	NA	NA	NA		

^aFor survey mode, 0 refers to paper-based onsite survey; 1 refers to smartphone-based onsite survey; and 2 refers to smartphone-based *post hoc* survey. ^bLIHKG is a Reddit-like Internet forum that was widely used during the Hong Kong 2019 protests.

Table 4. Mann-Whitney *U* Tests for Surveys Conducted between June 16 and July 21: Paper-based versus Smartphone-based Surveys

	6/16	6/17	6/21	6/26	7/1	7/14	7/21
Total sample size	875	721	316	418	1,169	546	680
Paper-based	590	101	117	254	434	264	385
Smartphone-based	285	620	199	164	735	282	295
Age	—***	—***	—***	n.s.	—***	—***	—**
Gender (male)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Education	n.s.	—*	n.s.	n.s.	+**	+**	+
Socioeconomic status	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Past protest frequency	n.s.	—**	n.s.	n.s.	n.s.	n.s.	n.s.
Localist	NA	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Political inefficacy	NA	n.s.	n.s.	n.s.	n.s.	NA	—**
Traditional media	NA	NA	NA	n.s.	n.s.	+***	+**
Online media	NA	NA	NA	n.s.	+**	+***	+***
Facebook	NA	NA	NA	n.s.	+***	+***	+***
Instagram	NA	NA	NA	n.s.	+***	+***	+***
WhatsApp	NA	NA	NA	n.s.	+***	+***	+***
Telegram	NA	NA	NA	n.s.	+	+	+
LIHKG	NA	NA	NA	—**	+***	+**	+***
Radical actions are justifiable if the government fails to listen	NA	n.s.	n.s.	n.s.	n.s.	NA	n.s.
Radical actions will alienate ordinary citizens	NA	n.s.	n.s.	n.s.	n.s.	NA	n.s.
Radical actions are best when combined with peaceful ones	NA	n.s.	n.s.	n.s.	+**	NA	+***
Radical actions are understandable	n.s.	NA	NA	NA	NA	NA	+***

Note: Negative sign means the mean value of the variable from smartphone-based onsite surveys is significantly lower than that from paper-based onsite surveys. *p* values are available in the online appendix: <https://sites.google.com/view/surveying-mass-protests/>.

p* < .05. *p* < .01. ****p* < .001.

A common finding here is that the mean age of respondents is consistently lower for smartphone-based questionnaires than for paper-based ones, as shown by the negative coefficients. This is not a surprise. As explained earlier, smartphone-based questionnaires were more likely to be completed by tech-savvy respondents, who also tended to be younger. Significant differences in the education level appear in some of the surveys, although the pattern is not as consistent as that for age. We find no statistically significant differences for gender and socioeconomic status. This shows the online survey method did not create response biases for demographic variables, except for age. Yet, we do see significant differences in the media and radicalization variables. Respondents in the smartphone-based samples were more reliant on online and social media as sources of information, and they were also more radical in how they saw protests. We suspect these differences may be due to the fact that the smartphone-based samples are younger in age.

To verify our instinct, we performed a series of ordinary least squares (OLS) regressions in which age is treated as the independent variable. Political participation, political attitudes, and use of media are designated as the dependent variables. Gender, education level, and perceived socioeconomic status are controlled in each model.

Table 5. OLS Regression with Age as Independent Variable of Main Interest and Listed Attributes as Dependent Variable

	6/16	6/17	6/21	6/26	7/1	7/14	7/21
Past protest frequency	n.s.	n.s.	n.s.	n.s.	-0.128***	n.s.	n.s.
Localist	-0.140***	-0.127**	n.s.	n.s.	-0.197***	-0.246***	-0.225***
Political inefficacy	N.A.	n.s.	n.s.	-0.122*	-0.079*	N.A.	-0.093*
Traditional media	N.A.	N.A.	N.A.	n.s.	0.110**	0.141**	0.127**
Online media	N.A.	N.A.	N.A.	n.s.	-0.307***	-0.235***	-0.208***
Facebook	N.A.	N.A.	N.A.	0.138**	-0.220***	-0.146***	-0.191***
Instagram	N.A.	N.A.	N.A.	-0.238***	-0.579***	-0.409***	-0.547***
WhatsApp	N.A.	N.A.	N.A.	0.119*	-0.110***	-0.093*	n.s.
Telegram	N.A.	N.A.	N.A.	-0.169***	-0.392***	-0.428***	-0.438***
LIHKG	N.A.	N.A.	N.A.	-0.192***	-0.504***	-0.401***	-0.533***
Radical actions are justifiable if the government fails to listen	N.A.	-0.142***	n.s.	n.s.	-0.219***	N.A.	-0.119**
Radical actions will alienate ordinary citizens	N.A.	n.s.	n.s.	n.s.	0.105**	N.A.	n.s.
Radical actions are best when combined with peaceful ones	N.A.	n.s.	n.s.	n.s.	-0.236***	N.A.	-0.216***
Radical actions are understandable	-0.161**	N.A.	N.A.	N.A.	N.A.	N.A.	-0.195***

Note: Numbers displayed are standardized coefficients for the OLS regression; statistically insignificant coefficients are not displayed. *p* values are available in the online appendix: <https://sites.google.com/view/surveying-mass-protests/>. OLS = ordinary least squares.
 p* < .05. *p* < .01. ****p* < .001.

Table 5 presents the regression results; the first column lists the dependent variables that are predicted by age. The results show age is indeed a determinant of political attitudes and use of media. Younger respondents are more likely to rely on online and social media and to show radical attitudes. The findings are consistent across most of the surveys, except on June 26, in which older respondents were more frequent users of Facebook and WhatsApp. This is probably because respondents on June 26 were overall younger than those in the July protests. Given a narrower age range, the positive relationship between age and use of Facebook and WhatsApp could be due to an intracohort difference in the use of social media (i.e., comparison between relatively young cohorts), instead of an intercohort difference. Facebook and WhatsApp are more popular among senior young people than among other young people; social media use is less common among most of the older citizens in general.

These findings support our conjecture that respondents who completed paper-based surveys were somewhat different from those who completed online surveys. The latter were younger and probably more tech-savvy, which leads to differences in media habits and political attitudes. These differences, given that the smartphone-based samples were larger than the paper-based samples in some surveys, necessitate sample weighting. After each survey, we first compared paper-based samples with the smartphone-based samples in terms of their age, gender, and education level by running a *t* test. If we found significant differences in these demographics between the two samples, we weighed the online sample based on the demographic distributions of the paper-based one. We assume the latter to be more representative of the protester

Table 6. Mann-Whitney *U* Tests for Smartphone-based Surveys Conducted between August 4 and January 1: Onsite versus *Post Hoc* Surveys

	8/4	8/11	8/18	8/25	9/8	9/15	12/8	1/1
Total sample size	1,272	636	806	372	337	911	902	1,306
Paper-based	231	112	126	68	74	178	142	266
Smartphone-based (onsite)	244	82	67	7	32	209	145	297
Smartphone-based (<i>post hoc</i>)	797	442	613	297	231	524	615	743
Age	n.s.	n.s.	***	+	n.s.	***	****	****
Gender (male)	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Education	+	-**	n.s.	n.s.	n.s.	****	n.s.	+
Socioeconomic status	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	-***
Past protest frequency	n.s.	-*	n.s.	n.s.	NA	NA	n.s.	n.s.
Localist	n.s.	n.s.	-*	n.s.	n.s.	n.s.	-*	n.s.
Political inefficacy	n.s.	n.s.	n.s.	n.s.	NA	***	NA	n.s.
Traditional media	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Online media	n.s.	n.s.	n.s.	n.s.	n.s.	****	n.s.	n.s.
Facebook	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
Instagram	n.s.	-***	-**	n.s.	n.s.	-***	n.s.	n.s.
WhatsApp	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	-***
Telegram	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	-***	n.s.
LIHKG	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	-***	n.s.
Radical actions are justifiable if the government fails to listen	n.s.	n.s.	n.s.	n.s.	NA	n.s.	-***	NA
Radical actions will alienate ordinary citizens	n.s.	n.s.	n.s.	n.s.	NA	+	n.s.	NA
Radical actions are best when combined with peaceful ones	n.s.	n.s.	n.s.	n.s.	NA	****	n.s.	NA
Radical actions are understandable	n.s.	n.s.	n.s.	n.s.	NA	***	n.s.	NA

Note: A negative sign means the mean value of the variable from *post hoc* smartphone-based surveys is significantly lower than that from smartphone-based onsite surveys. *p* values are available in the online appendix: <https://sites.google.com/view/surveying-mass-protests/>.

p* < .05. *p* < .01. ****p* < .001.

population because paper-based surveys were not affected by possible biases introduced by digital literacy. Therefore, paper-based samples are assumed to be the “true value” for weighing the entire sample of a specific survey.

The June and July surveys indicate that the major problem of smartphone-based onsite surveys is age response bias, which skewed the sample to a younger age. The addition of a third sampling mode—*post hoc* smartphone-based surveys—created an opportunity to rectify this bias, even though we did not initially adopt it for this purpose (but to ensure a bigger sample given that protest conditions became more uncertain and chaotic after August). As Table 3 shows, the mean age of respondents is consistently higher in the *post hoc* surveys than in the onsite smartphone-based surveys. This is confirmed by the Mann-Whitney *U* test between the two online survey modes. Table 6 presents the results, which indicate the age difference between these two online survey modes is statistically significant in five out of the eight surveys; the test does not show statistically significant differences for most other variables, except occasionally for education, media use, and attitudes on radicalization.

The results from Tables 4 to 6 indicate age is the key factor contributing to the response bias among the three modes of survey, and it is age that shapes differences in other attributes, including use of media and attitudes on radicalization. But might other

Table 7. Logistic Regression for Paper-based versus Smartphone-based (Onsite) Surveys

Dependent Variable: Survey Mode (0 = Paper-based; 1 = Smartphone-based Onsite)	6/17	7/1	7/14	7/21
Age	-0.032**	-0.035***	-0.066***	n.s.
Gender (male)	n.s.	n.s.	n.s.	n.s.
Education level	n.s.	n.s.	0.493*	0.417*
Socioeconomic status	0.681**	n.s.	n.s.	n.s.
Past protest frequency	n.s.	n.s.	0.131*	n.s.
Localist	n.s.	n.s.	n.s.	n.s.
Political inefficacy	n.s.	n.s.	NA	n.s.
Traditional media	NA	n.s.	0.384***	n.s.
Online media	NA	n.s.	n.s.	n.s.
Facebook	NA	n.s.	0.434***	n.s.
Instagram	NA	n.s.	n.s.	n.s.
WhatsApp	NA	0.284***	n.s.	n.s.
Telegram	NA	-0.186**	n.s.	n.s.
LHK	NA	n.s.	n.s.	0.209*
Radical actions are justifiable if the government fails to listen	n.s.	n.s.	NA	n.s.
Radical actions will alienate ordinary citizens	n.s.	0.269**	NA	0.243*
Radical actions are best when combined with peaceful ones	n.s.	n.s.	NA	0.388*
Radical actions are understandable	NA	NA	NA	n.s.

Note: Numbers displayed are standardized coefficients for the OLS regression; statistically insignificant coefficients are not displayed. *p* values are available in the online appendix: <https://sites.google.com/view/surveying-mass-protests/>. OLS = ordinary least squares.

p* < .05. *p* < .01. ****p* < .001.

factors, such as education, also contribute to the response bias? For instance, are more highly educated people more likely to take online surveys irrespective of age? To ensure age is indeed the most important source of response bias across the survey modes, we conducted binary logistic regressions based on the findings from the Mann-Whitney *U* tests presented in Tables 4 and 6, following the test adopted by Wolfgang Rüdiger (2010). For the test, we chose surveys in which significant mean differences are simultaneously found in age and the other variables. These include surveys on June 17, July 1, July 14, and July 21 from Table 4, and surveys on August 18, September 15, December 8, and January 1 from Table 6. Here, survey modes are designated as the dependent variable, and age and other attributes are treated as independent variables.

Tables 7 and 8 present results of the logistic regressions and show how different variables predict the survey mode. Table 7 compares paper-based and onsite smartphone-based surveys; Table 8 compares onsite and *post hoc* smartphone-based survey. Both tables support our observation that age is the most important factor in differentiating how respondents are drawn to different survey modes. Age is statistically significant in three out of four protests. Other attributes—particularly education level—are occasionally found to be significant, but the effect of age is more consistent. When controlling for other factors, younger protesters are more likely to take smartphone-based surveys, compared with paper-based ones (as shown by the negative coefficients for age in Table 7), and older protesters are more likely to submit *post*

Table 8. Logistic Regression for Smartphone-based Surveys, Onsite versus *Post Hoc*

Dependent Variable: Survey Mode (0 = Smartphone-based Onsite; 1 = Smartphone-based <i>Post Hoc</i>)	8/18	9/15	12/8	1/1
Age	0.049*	n.s.	0.038**	0.037***
Gender (male)	n.s.	n.s.	n.s.	n.s.
Education level	0.604*	0.880***	n.s.	n.s.
Socioeconomic status	n.s.	n.s.	n.s.	n.s.
Past protest frequency	n.s.	n.s.	n.s.	n.s.
Localist	n.s.	0.597*	n.s.	n.s.
Political inefficacy	n.s.	-0.563*	NA	-0.443*
Traditional media	n.s.	n.s.	n.s.	n.s.
Online media	n.s.	1.013***	n.s.	n.s.
Facebook	n.s.	n.s.	0.233*	n.s.
Instagram	n.s.	-0.246*	n.s.	n.s.
WhatsApp	n.s.	n.s.	n.s.	n.s.
Telegram	0.292**	n.s.	n.s.	n.s.
LHKKG	n.s.	n.s.	n.s.	n.s.
Radical actions are justifiable if the government fails to listen	n.s.	n.s.	-0.404**	NA
Radical actions will alienate ordinary citizens	n.s.	n.s.	n.s.	NA
Radical actions are best when combined with peaceful ones	n.s.	n.s.	n.s.	NA
Radical actions are understandable	n.s.	n.s.	n.s.	NA

Note: Numbers displayed are standardized coefficients for the OLS regression; statistically insignificant coefficients are not displayed. *p* values are available in the online appendix: <https://sites.google.com/view/surveying-mass-protests/>. OLS = ordinary least squares.

p* < .05. *p* < .01. ****p* < .001.

hoc surveys, compared with smartphone-based onsite ones (as shown by the positive coefficients for age in Table 8).

Findings from these statistical tests provide strong evidence on the difference in age among different modes of survey, and also on how the difference in age is associated with differences in protesters' attributes. As these findings reveal, smartphone-based onsite surveys have a response bias toward younger respondents, and older citizens "prefer" paper-based surveys. As such, smartphone-based *post hoc* surveys might strike a balance between the respective bias of the other two survey modes. To illustrate this relationship, we conducted cross-tabulations between age groups of respondents and survey modes. As shown in Table 9, in the eight surveys where all three survey modes were adopted, we find significant relationships between age groups and survey modes in six of them. In five of these surveys, respondents from *post hoc* surveys tended to be more evenly distributed across age groups, as compared with onsite online surveys, which were often skewed to people in their 20s. The result on September 15 is an outlier that will be discussed in the conclusion.

The relatively older samples collected from the *post hoc* surveys may thus help balance the consistently younger samples from the onsite smartphone-based surveys. They could also help make up for selection bias that might result from relying on a team of university student interviewers, who were likely inclined to approach young peers for interviews. As explained earlier, this might be related to the method of flyer

Table 9. Comparison of Age Group Distribution across Survey Modes

	19 or Below (%)	20–24 (%)	25–29 (%)	30–39 (%)	40 or Above (%)
August 4 ($\chi^2 = 22.43^{**}$)					
Paper-based	5.6	25.9	19.9	14.7	33.9
Smartphone-based (onsite)	7.6	34.6	17.7	20.7	19.4
Smartphone-based (<i>post hoc</i>)	8.6	27.2	22.1	18.7	23.4
August 11 ($\chi^2 = 53.14^{***}$)					
Paper-based	10.8	14.4	18.9	9.0	46.8
Smartphone-based (onsite)	11.2	27.6	34.7	20.4	6.1
Smartphone-based (<i>post hoc</i>)	14.6	20.7	23.5	18.3	23.0
August 18 ($\chi^2 = 28.14^{***}$)					
Paper-based	7.4	31.1	23.6	12.8	25.0
Smartphone-based (onsite)	4.1	32.9	38.4	21.9	2.7
Smartphone-based (<i>post hoc</i>)	8.9	23.0	27.6	21.3	19.1
August 25 ($\chi^2 = 9.12$)					
Paper-based	13.2	32.4	16.2	20.6	17.6
Smartphone-based (onsite)	28.6	42.9	14.3	14.3	0.0
Smartphone-based (<i>post hoc</i>)	10.1	23.2	23.6	20.2	22.9
September 8 ($\chi^2 = 0.15$)					
Paper-based	1.4	27.0	14.9	18.9	37.8
Smartphone-based (onsite)	3.1	37.5	15.6	15.6	28.1
Smartphone-based (<i>post hoc</i>)	3.9	29.0	24.2	22.1	20.8
September 15 ($\chi^2 = 47.02^{***}$)					
Paper-based	12.9	38.8	20.2	20.2	7.9
Smartphone-based (onsite)	8.6	30.1	16.7	15.3	29.2
Smartphone-based (<i>post hoc</i>)	9.7	21.0	20.8	21.9	26.5
December 8 ($\chi^2 = 42.91^{***}$)					
Paper-based	17.3	19.9	22.4	9.6	30.8
Smartphone-based (onsite)	27.3	31.5	18.9	14.0	8.4
Smartphone-based (<i>post hoc</i>)	13.8	20.6	22.9	16.6	26.1
January 1 ($\chi^2 = 49.31^{***}$)					
Paper-based	18.7	13.5	23.5	13.5	30.8
Smartphone-based (onsite)	18.9	22.8	25.2	20.2	12.9
Smartphone-based (<i>post hoc</i>)	14.8	15.9	20.4	16.2	27.6

* $p < .05$. ** $p < .01$. *** $p < .001$.

distribution: When distributing flyers in a protest crowd, interviewers were more likely to do so randomly than to select only their peers. Even though responses depend on whether the potential respondent takes the initiative to complete the survey after the protest, test statistics show that such respondents do not significantly differ from those in either the paper-based or onsite electronic mode, except they tend to have higher education levels, which other scholars have also identified (Fisher 2020; Walgrave and Verhulst 2011). In summary, *post hoc* smartphone-based surveys can serve as an effective method alongside paper-based and onsite smartphone-based surveys to maximize overall sample sizes without creating unacceptable response biases.

Before we conclude, it is important to revisit the issue of sample weighting. As mentioned, the online surveys allowed us to significantly increase the sample size, but the fact that they require a certain level of digital literacy might bias the samples toward

Table 10. Sample Weighting for August 4 and August 11 Surveys

	August 4		August 11	
	Unweighted	Weighted	Unweighted	Weighted
Age (%)				
19 or below	10.0	7.8	14.1	13.4
20–29	49.8	49.2	46.6	45.1
30–39	21.5	18.3	23.7	17.1
40–49	10.2	10.4	8.3	7.4
50–59	6.6	10.0	5.2	9.7
60 or above	1.8	4.3	2.2	7.3
Gender (% female)	47.7	42.0	46.2	41.5
Education level (%)				
Primary level or below	0.2	0.4	0.5	1.8
Secondary level	17.1	22.2	20.9	31.1
Tertiary education or above	82.7	77.4	78.7	67.1
Socioeconomic status (%)				
Lower class	38.7	38.5	51.8	51.1
Middle class	60.7	60.9	47.5	48.1
Upper class	0.6	0.6	0.7	0.8
Localists (%)	39.6	39.0	40.5	39.7
Political inefficacy (%)	51.5	50.9	62.6	61.9
Past protest history (0–7)	1.93	1.96	1.95	1.94
Media use (1–5)				
Traditional media	3.77	3.78	3.67	3.65
Online media	4.79	4.76	4.75	4.70
Facebook	4.54	4.50	4.52	4.43
Instagram	3.60	3.49	3.51	3.35
WhatsApp	3.92	3.89	3.91	3.86
Telegram	3.55	3.45	3.87	3.69
LIHKG	4.24	4.16	4.42	4.27
Political attitudes (1–5)				
Radical actions are justifiable if the government fails to listen	3.48	3.48	3.54	3.56
Radical actions will alienate ordinary citizens	3.18	3.18	3.28	3.26
Radical actions are best when combined with peaceful ones	4.52	4.5	4.52	4.49
Radical actions are understandable	4.73	4.73	4.73	4.71

younger protesters. Thus, whenever large differences existed between the samples from different survey modes, it was necessary to weigh the online samples, which we did by assuming the demographic distribution of the paper-based ones as the “true values.” Table 10 illustrates the result of sample weighting and shows the unweighted and weighted findings in the August 4 and August 11 surveys, in which all three survey modes were adopted. The results show the percentage of elderly respondents expanded after weighting. The distributions of education level were also adjusted downward because the education level correlates with age in both surveys ($r = -.196$ and $-.190$, respectively, $p < .001$ for both). The values relating to use of social media and political attitudes were adjusted downward. This pattern is largely consistent with our expectation that age was a defining factor for other attributes of respondents. It thus lends support to our rationale for doing sample weighting, which is to minimize the potential bias caused by online surveys—especially onsite ones.

DISCUSSION AND CONCLUSION

The recent emergence of spontaneous and leaderless protests presents new methodological challenges to researchers who study protest events through surveys. This article discussed some of these challenges and proposed a new approach to address them. Based on our fieldwork in Hong Kong's Anti-ELAB Movement, we delineated a mixed-sampling survey method that combines traditional paper-based questionnaires with onsite and *post hoc* smartphone-based ones. Use of these online survey techniques is intended to ensure the collection of large-enough samples despite the volatile conditions of spontaneous and leaderless protests, substituting for mail-in questionnaires that may result in low responses due to the decline of postal communications. Combining onsite and *post hoc* online surveys is also purposive. Because the former technique is likely to be biased toward younger respondents, it is important to use the latter to reach older respondents, such that the age bias can be mitigated. Test statistics indicate that the differences in major variables from the three survey modes are not statistically significant in a consistent manner, except for respondents' age, which can be tackled through the combination of different survey modes.

Our research offers fresh evidence to support incorporating the use of online survey methods into protest surveys, contributing to the broader literature that evaluates the application of online surveys in the social sciences. Online and other electronic survey techniques have become increasingly popular among researchers who survey protest events, but there has been a lack of serious methodological reflection on whether they are reliable methods and how they should be implemented. Our study provides, to the best of our knowledge, the first empirical inquiry to understand the potential biases produced by online surveys in studying protests and to explore possible ways to overcome them. Although online surveys have their limitations during the implementation stage (e.g., cybersecurity issues), they offer a cheap and effective way to maximizing samples in difficult protest conditions, and by mixing onsite and *post hoc* survey modes, they can ensure the representativeness of samples. Given that it is easily replicable as long as one has enough interviewers and the necessary technological infrastructure (high smartphone penetration and good Internet connection), the proposed mixed-mode sampling method can be applied to many other popular protests, especially protests in urban locations that have massive turnouts and more fluid forms. Recent examples include the Thailand anti-monarchy student protests, the Belarus uprising, and the Black Lives Matter protests in the United States.

Our discussion mostly focused on the methodological challenges and possible solutions stemming from the spontaneous and fluid nature of the protests. It is important to note, however, that challenges also came from protesters during our fieldwork in Hong Kong. Response rates were generally very high across the protest events (likely because protesters were willing to express their views, just as in other countries), but on multiple occasions, our surveys were doubted for their authenticity. In June 2019, early in the movement, some protesters "reported" our survey on an open group in Telegram, a popular messaging app, suspecting we were working with the police to extract protesters' data. Fortunately, an anonymous netizen quickly posted on the same group and clarified that we were real university researchers. Later, during an August

protest, our survey was “reported” again on LIHKG, a Reddit-like online forum that gained wide popularity during the protests. The post gained traction immediately and people began to sabotage the online questionnaire, filling in nonsensical answers. Eventually we had no choice but to call off the survey. These experiences taught us that respondents’ suspicion is a major risk factor in implementing online surveys during live protests, and gaining trust is the only way to mitigate such a risk. The team decided to include more details about the research (e.g., its objectives), the profile of team members (e.g., our personal Web pages), and Weblinks to journal articles as well as media reports that team members had published, in both the online survey interface and the flyers we distributed to respondents. We also included an office phone number of our university research center, through which people could verify our identity and inquire more about the research. Such information served as our credentials and made surveys—and us—look more credible. The effort seemed to pay off, as similar incidents no longer occurred.

The distribution of mean values across the three survey modes on September 15 was an exception that deserves explanation. In that survey, paper-based survey respondents were significantly younger, which is the opposite of what we observed on other dates, and thus contradicts our argument that less tech-savvy (and thus older) respondents tend to be drawn to paper-based surveys. In retrospect, we believe the problem might have to do with the protest dynamics on that day and our interviewer allocation. In the September 15 protest, some interviewers in charge of paper-based surveys were assigned to locations that became bottlenecks where the protest crowd was stuck. As a result, these interviewers were jammed in the crowd and only managed to interview the people around them, who might have been a group of young protesters. Bottlenecks can be caused by numerous factors—the form of protests, the landscape of protest areas, or the police’s crowd management tactics—but on September 15, police tactics mattered most. Because that protest was not authorized by the police, many roadblocks were set up, and protesters became stuck at the starting point. This was, to a large extent, unavoidable, but it was a valuable experience for the research team.

To conclude, this article provided a methodological review on the combination of face-to-face and online modes of protest surveys. Limitations certainly abound, primarily because the surveys we conducted in the Hong Kong protests were not designed to conduct methodological reflections in the first place. We developed our methods basically by “crossing the river while feeling the stones,” that is, we adapted continuously to the changing protest environment and contingent events. Our method is certainly far from perfect, but it provides a useful foundation researchers can adopt when studying protests of a similar nature. More could be done to check for the representativeness of data collected from online surveys. For example, researchers can include questions that gauge fundamental social and political attitudes but are theoretically unaffected by protest dynamics—such as views about the role of the government in the economy or attitudes on lesbian, gay, bisexual, transgender (LGBT) rights—to segment protesters and allow for comparison with data collected from paper-based surveys. More could also be done to test for other types of bias. For instance, similar to

Stefaan Walgrave and colleagues (2016) who devised ways to identify delayed refusal biases from mail-in questionnaires, researchers could do the same for *post hoc* online surveys and compare the results. Researchers could also follow Stefaan Walgrave and colleagues (2016) to investigate how characteristics of interviewers can induce selection biases in responses. In short, this article builds on previous methodological reflections and presents a method of using online surveys to study contemporary protests. It joins a growing literature that attests to the validity of using online surveys to collect social science data. Further research is needed to find more sophisticated ways through which online surveys can be administered.

Acknowledgments

The authors would like to thank the editor and the three anonymous reviewers for their valuable feedback on earlier drafts. We also thank our survey helpers, without whom it would not have been possible to complete the research.


Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for research, authorship, and/or publication of this article: This work was partly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Ref. no. 11605820).

ORCID iD

Samson Yuen  <https://orcid.org/0000-0002-7816-7180>

Notes

1. One pertinent example is the collaborative research project “Caught in the act of protest: Contextualizing contestation,” which uses protest surveys to compare demonstrations in seven European countries (see <https://protestsurvey.eu/>).
2. For an overview of the protest survey from design to implementation, see Massimiliano Andretta and Donatella Della Porta (2014).
3. In Hong Kong, although the right to assembly is enshrined in the law, protest organizers must first notify the police before an event. A protest is only legal if the police either grant a Letter of No Objection or do not issue a Letter of Objection. Otherwise, protesting can be a punishable crime.
4. The research team did rely on this method in the June 9 protest, the first mass protest that formally kicked off the movement. However, due to crowded conditions, response was low. Only 285 surveys were collected on that day.
5. We follow the survey method adopted by Francis L. F. Lee and Joseph M. Chan (2010) during the annual July 1 rallies and Cheng W. Edmund and Wai-Yin Chan (2017) during the Umbrella Movement occupation. Instead of inviting every fifth person as other scholars have typically done, we opted for 10 because of the often-crowded conditions in Hong Kong’s mass protests.
6. Interviewers were mostly university students who had prior experience in conducting surveys. We paid them an hourly rate. Interviewers’ contact information was stored in a WhatsApp group, which allowed the research team to swiftly assemble a group of interviewers.

7. According to WhatsApp, the live location function is protected by end-to-end encryption, similar to its private messaging. This means no one can access a user's location except the people the user shares it with (see <https://faq.whatsapp.com/android/chats/how-to-use-live-location/?lang=en>).
8. Sample survey questionnaires can be accessed in this repository: https://drive.google.com/drive/folders/1-WKkH5tN5CQeJ_dZcIPjCBxr3-8U7Rt9?usp=sharing.
9. According to the Census and Statistics Department of the Hong Kong Government, Hong Kong had a smartphone penetration rate of 91.5 percent in 2019, which puts it at the top of the global chart (see Thematic Household Survey Report No. 69, The Government of Hong Kong Special Administrative Region, March 26, 2020, <https://www.info.gov.hk/gia/general/202003/26/P2020032600444.htm>).
10. Once we downloaded the data, we removed all IP addresses to ensure the anonymity of respondents as much as possible.
11. The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed.
12. In Hong Kong's context, a localist is someone who supports self-determination or political independence and who tends to support a more confrontational approach to politics. They are distinguished from pan-democrats, who generally support fighting for democracy within the existing constitutional framework of "One Country, Two Systems" (for details, see Veg 2017).

References

- Andretta, Massimiliano, and Donatella Della Porta. 2014. "Surveying Protestors: Why and How." Pp. 308–34 in *Methodological Practices in Social Movement Research*, edited by D. Della Porta. Oxford, England: Oxford University Press.
- Boserup, Anders, and Claus Iversen. 1966. "Demonstrations as a Source of Change: A Study of British and Danish Easter Marchers." *Journal of Peace Research* 3(4):328–48.
- Cheng, Edmund W., and Wai-Yin Chan. 2017. "Explaining Spontaneous Occupation: Antecedents, Contingencies and Spaces in the Umbrella Movement." *Social Movement Studies* 16(2):222–39.
- Cheng, Edmund W., Francis L. F. Lee, Samson Yuen, and Gary Tang. 2022. "Total Mobilization from Below: Abeyance Networks, Threats and Emotions in Hong Kong's Freedom Summer." *China Quarterly*.
- Evans, Joel R., and Anil Mathur. 2005. "The Value of Online Surveys." *Internet Research* 15(2):195–219.
- Evans, Joel R., and Anil Mathur. 2018. "The Value of Online Surveys: A Look Back and a Look Ahead." *Internet Research* 28(4):854–87.
- Fisher, Dana R. 2020. *American Resistance: From the Women's March to the Blue Wave*. New York: Columbia University Press.
- Fisher, Dana R., Kenneth T. Andrews, Neal Caren, Erica Chenoweth, Michael T. Heaney, Tommy Leung, L. Nathan Perkins, and Jeremy Pressman. 2019. "The Science of Contemporary Street Protest: New Efforts in the United States." *Science Advances* 5(10):eaaw5461.
- Flesher Fominaya, C. 2015. "Debunking Spontaneity: Spain's 15-M/Indignados as Autonomous Movement." *Social Movement Studies* 14(2):142–163.
- Gaggioli, Andrea, Giovanni Pioggia, Gennaro Tartarisco, Giovanni Baldus, Daniele Corda, Pietro Cipresso, and Giuseppe Riva. 2013. "A Mobile Data Collection Platform for Mental Health Research." *Personal and Ubiquitous Computing* 17(2):241–51.
- Grewal, Sharan, M. Tahir Kilavuz, and Robert Kubinec. 2019. *Algeria's Uprising: A Survey of Protesters and the Military*. Washington, DC: Brookings Institution.
- Hassanpour, Navid. 2016. *Leading from the Periphery and Network Collective Action*. Vol. 42. Cambridge, England: Cambridge University Press.
- Heaney, Michael T. 2018. "Making Protest Great Again." *Contexts* 17(1):42–47.
- Heaney, Michael T., and Fabio Rojas. 2007. "Partisans, Nonpartisans, and the Antiwar Movement in the United States." *American Politics Research* 35(4):431–64.
- Heaney, Michael T., and Fabio Rojas. 2015. *Party in the Street: The Antiwar Movement and the Democratic Party after 9/11*. New York: Cambridge University Press.

- Heerwegh, Dirk. 2009. "Mode Differences between Face-to-face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects." *International Journal of Public Opinion Research* 21(1):111–21.
- Jenkins, Robin. 1967. "Who Are These Marchers?" *Journal of Peace Research* 4(1):46–60.
- Ketchley, Neil. 2017. *Egypt in a Time of Revolution*. Cambridge, England: Cambridge University Press.
- Keusch, Florian, Mariel M. Leonard, Christoph Sajons, and Susan Steiner. 2019. "Using Smartphone Technology for Research on Refugees: Evidence from Germany." *Sociological Methods & Research* 50(4):1863–94.
- Lee, Francis L. F., and Joseph M. Chan. 2010. *Media, Social Mobilisation and Mass Protests in Post-colonial Hong Kong: The Power of a Critical Event*. New York: Routledge.
- Levy, Mark R. 1983. "The Methodology and Performance of Election Day Polls." *Public Opinion Quarterly* 47(1):54–67.
- Ley, Barbara L., and Paul R. Brewer. 2018. "Social Media, Networked Protest, and the March for Science." *Social Media + Society* 4(3):1–12. doi:10.1177/2056305118793407.
- Manago, Bianca, Trenton D. Mize, and Long Doan. 2021. "Can You Really Study an Army on the Internet? Comparing How Status Tasks Perform in the Laboratory and Online Settings." *Sociological Methodology* 51(2):319–47.
- Moor, Joost de, Katrin Uba, Mattias Wahlström, Magnus Wennerhag, and Michiel De Vydt. 2020. "Protest for a Future II: Composition, Mobilization and Motives of the Participants in Fridays for Future Climate Protests on 20–27 September, 2019, in 19 Cities around the World." Retrieved September 1, 2021. (<https://osf.io/asruw/>).
- Nandi, Alita, and Lucinda Platt. 2017. "Are There Differences in Responses to Social Identity Questions in Face-to-face versus Telephone Interviews? Results of an Experiment on a Longitudinal Survey." *International Journal of Social Research Methodology* 20(2):151–66.
- Onuch, Olga. 2014. "Who Were the Protesters?" *Journal of Democracy* 25(3):44–51.
- Onuch, Olga, and Gwendolyn Sasse. 2016. "The Maidan in Movement: Diversity and the Cycles of Protest." *Europe-Asia Studies* 68(4):556–87.
- Parkin, Frank. 1968. *Middle Class Radicalism: The Social Bases of the British Campaign for Nuclear Disarmament*. Manchester, England: Manchester University Press.
- Rada, Vidal Diaz de, and Juan Antonio Dominguez-Álvarez. 2014. "Response Quality of Self-administered Questionnaires: A Comparison between Paper and Web Questionnaires." *Social Science Computer Review* 32(2):256–69.
- Raento, Mika, Antti Oulasvirta, and Nathan Eagle. 2009. "Smartphones: An Emerging Tool for Social Scientists." *Sociological Methods & Research* 37(3):426–54.
- Rüdiger, Wolfgang. 2010. "Assessing Nonresponse Bias in Activist Surveys." *Quality & Quantity* 44(1): 173–80.
- Simmons, Alicia D., and Lawrence D. Bobo. 2015. "Can Non-full-probability Internet Surveys Yield Useful Data? A Comparison with Full-probability Face-to-face Surveys in the Domain of Race and Social Inequality Attitudes." *Sociological Methodology* 45(1):357–87.
- Snow, David A., and Dana M. Moss. 2014. "Protest on the Fly: Toward a Theory of Spontaneity in the Dynamics of Protest and Social Movements." *American Sociological Review* 79(6):1122–43.
- Steinert-Threlkeld, Zachary C. 2017. "Spontaneous Collective Action: Peripheral Mobilization during the Arab Spring." *American Political Science Review* 111(2):379–403.
- Sugie, Naomi F. 2018. "Utilizing Smartphones to Study Disadvantaged and Hard-to-reach Groups." *Sociological Methods & Research* 47(3):458–91.
- Tufekci, Zeynep. 2017. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. New Haven, CT: Yale University Press.
- Tufekci, Zeynep, and Christopher Wilson. 2012. "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square." *Journal of Communication* 62(2):363–79.
- Veg, Sebastian. 2017. "The Rise of 'Localism' and Civic Identity in Post-handover Hong Kong: Questioning the Chinese Nation-state." *The China Quarterly* 230:323–47.

- Verhulst, Joris, and Stefaan Walgrave. 2009. "The First Time Is the Hardest? A Cross-national and Cross-issue Comparison of First-time Protest Participants." *Political Behavior* 31(3):455–84.
- Walgrave, Stefaan, and Joris Verhulst. 2011. "Selection and Response Bias in Protest Surveys." *Mobilization: An International Journal* 16(2):203–22.
- Walgrave, Stefaan, Ruud Wouters, and Pauline Ketelaars. 2016. "Response Problems in the Protest Survey Design: Evidence from Fifty-one Protest Events in Seven Countries." *Mobilization: An International Quarterly* 21(1):83–104.

Author Biographies

Samson Yuen is an assistant professor in the Department of Government and International Studies at Hong Kong Baptist University. He researches contentious politics, civil conflicts, public opinion, and health and food politics, focusing particularly on East Asia. His articles have been published in *Political Studies*, *Government and Opposition*, *Mobilization*, *Social Movement Studies*, *China Quarterly*, and *Journal of Contemporary China*. He holds a DPhil in politics from Oxford University.

Gary Tang is an assistant professor in the Department of Social Science and an associate director of the Centre for Public Policy Research at the Hang Seng University of Hong Kong. His research interests include political communication, social media, and social movements. His articles have been published in *Social Science Computer Review*, *China Quarterly*, and *Social Indicators Research*.

Francis L. F. Lee is the Director and Professor in the School of Journalism and Communication at the Chinese University of Hong Kong. He is the lead author of *Memories of Tiananmen: Politics and Processes of Collective Remembering in Hong Kong, 1989–2019* (Amsterdam University Press 2021) and *Media and Protest Logics in the Digital Era* (Oxford University Press 2018). He is the Chief Editor of the *Chinese Journal of Communication* and an elected Fellow of the International Communication Association.

Edmund W. Cheng is an associate professor in the Department of Public Policy and an associate director of the Centre for Public Affairs and Law at the City University of Hong Kong. His research interests include comparative politics, contentious politics, political communication, and research methods. He is the coeditor of the special issues "Revisiting the Public Sphere in 20th and 21st Century China" in *China Quarterly* and "Organizing without Organizations: Networked Movements in Asia" in *Social Movement Studies*.