



Understanding Society

Working Paper Series

No. 2019 – 16

December 2019

Event-triggered data collection

Annette Jäckle (University of Essex)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Michigan)



Non-Technical Summary

Every year in *Understanding Society*: the UK Household Longitudinal Study, we ask participants about key events that might have happened since the last interview. However, even with just a year in between interviews, there is a chance that we may miss some events, either because participants place them earlier in time than the last interview date or because they forget them altogether. In addition to this, there is interest in the research community in collecting additional information about these key events which may be difficult for the participant to recall at a later date.

Given this, we are testing the feasibility of collecting information about key events closer to the time that they occur. To accomplish this, we commissioned qualitative research with *Understanding Society* sample members to provide insights into the attitudes of participants to more frequent contacts. We also used a probability-based online access panel to run experiments to test the implementation of a monthly “life events” question.

Event-triggered data collection

Annette Jäckle (University of Essex)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Michigan)

Abstract: Each year, in *Understanding Society*: the UK Household Longitudinal Study, we ask respondents about key events that might have happened since they last took part in the study (e.g. changes in employment, family, partnership, health, etc). For this methodological briefing we have explored how the scope and/or quality of data about key transitions in people's lives could be improved by identifying events closer to their occurrence. For this purpose we have: 1) reviewed the annual *Understanding Society* questionnaire to identify candidate domains for event-triggered data collection, 2) carried out qualitative research to inform the design of event-triggered data collection, 3) developed prototype fieldwork protocols, and 4) experimentally tested aspects of these protocols.

Keywords: micro survey, life events, experiment, mode of survey invitation

JEL classification: C83

Acknowledgements: *Understanding Society* is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. This methodological briefing was funded by the Economic and Social Research Council UK Population Lab Innovation Development Grant (ES/S016651/1).

Corresponding author: Annette Jäckle, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, aejack@essex.ac.uk.

1. Review stage

We have reviewed the domains about which we collect data in the annual interviews and identified five candidate domains for event-triggered data collection: health, pregnancy, labour market activities, residential moves, and partnership changes. The rationale for capturing these events between the annual interviews varies.

Identifying residential moves close in time to their occurrence could help with tracing movers and reduce non-response in the annual interviews due to non-contact. Similarly capturing partnership changes could help with keeping in touch with sample members who split off from a sample household. This would in particular help retain separated parents of sample children, for whom contact rates tend to be low (in any survey) due to unknown addresses. For health, pregnancy and labour market activities the goal is to collect information about events that cannot be collected retrospectively, such as expectations.

We have discussed aspects of event-triggered data collection with individual *Understanding Society* Topic Champions and planned the next stage of consultation to coincide with the next Scientific Group meeting in the autumn of 2019 (date tbc). Once we have specified the information to be collected about events, a follow up stage will be to adapt the annual questionnaires to integrate it with this new between wave measurement of events.

2. Qualitative research about the expectations of participants

The aim of the qualitative research phase was 1) to investigate how best to capture information about life events that do (or do not) happen between the annual panel interviews, and 2) to explore under which conditions sample members would be willing to report on life events every month. For this purpose we contracted Kantar Public UK to carry out qualitative research with members of the *Understanding Society* Innovation Panel. The methodology and findings from this research are reported in the attached report by Horsley et al. (2019). The following is based on extracts of the Executive Summary from that report.

The study involved 42 participants across an online forum and focus group. Participants were recruited from the *Understanding Society Innovation Panel*. The sample was weighted towards participants most likely to experience, or who have recently experienced, one or more of five key life events: health, changes in labour market activity, pregnancy, moving to a new house, changes in relationship status. Researchers used a semi-structured discussion guide and a range of stimulus materials during the online and in-person discussions.

Acceptability of event-triggered data collection

The idea of event-triggered data collection was broadly accepted. Acceptability was supported by participants' existing relationship with *Understanding Society* and the prospect of financial incentives, and the emotional benefits participation provides to some participants. Except for highly sensitive topics like bereavement and financial circumstances, and health and relationships for some, events were not spontaneously identified as 'off limits.' Participants in our sample were more willing to respond to questions about certain life events, like relationships, health and work. Though, assumptions about the types and detail of questions likely to be asked about within each life event introduced some reservations. The topics of pregnancy and relationships were more likely to be viewed as sensitive and therefore less likely for participants to say they would respond to many follow-up questions.

Barriers to willingness to respond to event-triggered data collection

The more participants reflected on the idea of event-triggered data collection, the more reservations to participation emerged. Based on their experiences of the annual survey, participant expectations of time and effort were a key concern, and so it would be important to clearly state the time required to participate in the monthly data collection. Potential sensitivities to the data collection process were also identified and included asking questions too soon after a difficult event had taken place and asking about topics that some participants view as too intrusive.

Participants identified two risks to data collection if these sensitivities are not carefully handled. Variable data quality from capturing potentially inaccurate or “overly emotive” responses too soon after they were experienced. While this is always a risk with data collection, the frequency with which participants are expected to engage with event-triggered data collection means the risk may be higher. The second risk was low response rates or high dropout rates from overburdened or distressed participants. Regularly repeated questions about the same events that feel demanding or distressing may prompt participants to opt out of data collection. This highlights the need to fully inform sample members about the scope of the data collection, and the request that they only have to report on key changes in their life that have happened in the previous month.

Views on the frequency and length of data requests

The amount of time participants were prepared to spend on a monthly and quarterly basis varied, from a couple minutes to over an hour at each data collection point. Generally, participants were less willing to take part monthly and more willing to spend about 15 to 20 minutes in total every three months or so. Again, this reinforces the need to ensure that sample members are fully informed about the length of time they would be committing each month, from less than a minute for those who have not experienced any changes in the last month, to a few minutes for each major life change.

Views on data collection mode

There was little appetite across the sample for responding by text message beyond the initial question or a reminder to take part in the data collection. Alternative suggestions for data collection included an online forum or an app where participants can log on at times which suit them and contribute as much or as little as they wished.

Views on compensation for participation

Participants unanimously expected a financial incentive of some form for contributing to event-triggered data collection, with some high expectations of value

for time spent that may not be financially viable for ISER to fund. An incentive is necessary but unlikely to be enough on its own for regular responses to data collection requests.

Key recommendations

- Highlight the purpose and benefits of event-triggered data collection, especially to encourage people to respond when they do not have developments to report
- Manage expectations for event-triggered data collection to distinguish it from experiences of the length, depth and mode of participating in the annual survey. Be clear about how much time is likely required, the types of questions asked and reassure about anonymity and confidentiality
- If possible, consider communicating to participants that involvement in event-triggered data collection may reduce the length of the annual survey
- Include supportive messages around the request e.g. *We recognise this topic might be difficult for some people to reflect on depending on their circumstances*
- Consider the feasibility of the incentive plan because incentives will be essential for initial engagement, and maintaining responses over time
- Consider how to practically administer data collection at most every three months with a time requirement of around 20 minutes, which means necessarily prioritising some life events over others
- Tailor as much as possible to personalise the initial monthly request and avoid these questions being repetitive and irrelevant. For example, identify proxy measures in the annual survey to target event-triggered data collection or withhold specific questions to some respondents
- Allow some flexibility in data collection approaches, if possible give a choice of [modes] and the choice to defer responses to a later date
- Build on this research and conduct feasibility testing for the prototype event-triggered data collection approach

3. Prototype protocols for event-triggered data collection

Implementing event triggered data collection requires 1) identifying the occurrence of relevant events, and 2) implementing follow-up questions about any events that have occurred. There are several constraints on how this could be done:

- Budget constraints require minimising the costs of such additional data collection. This rules out using face-to-face interviewers as a mode of administration.
- The monthly question to identify events has to have a quick turn-around time if follow up questions are to be administered before the event question is sent out again the following month. This rules out paper questionnaires as the main mode.
- Respondents to the monthly events question should be representative of the full sample. This rules out using web as the only mode, given that internet usage is not universal.

Text messaging could be an ideal mode to administer a single question once a month to check for events. Text messaging is relatively cheap, widely used, and has a quick turnaround time. Text messaging does not require smartphones or internet access and coverage of mobile phones is high. Finally, text messaging will remain a viable option even once all mobile phones are smartphones.

Given these constraints on data collection mode, we developed an optimal protocol for frequent, short, and timely updates of critical events between waves of data collection. In order to further constrain the design parameters for initial testing, we assumed a single question (either a multi code question asking respondents to check all that apply, or a short series of yes/no questions) survey administered monthly to identify the occurrence of a small number of key events. Those events that respondents reported had occurred would then be followed up by short surveys identifying key elements of those events, potentially for feeding forward to the next wave of data collection for more detailed questions on the topic.

This protocol is sketched out in Figure 1. The protocol would start with an advance letter sent by post. This would introduce the event-triggered data collection exercise, provide a rationale for why we are introducing this, describe how the data would be used, and provide information on the likely time commitment each month and incentives involved. This letter would include instruction on how to opt out of the between-wave event-triggered data collection.

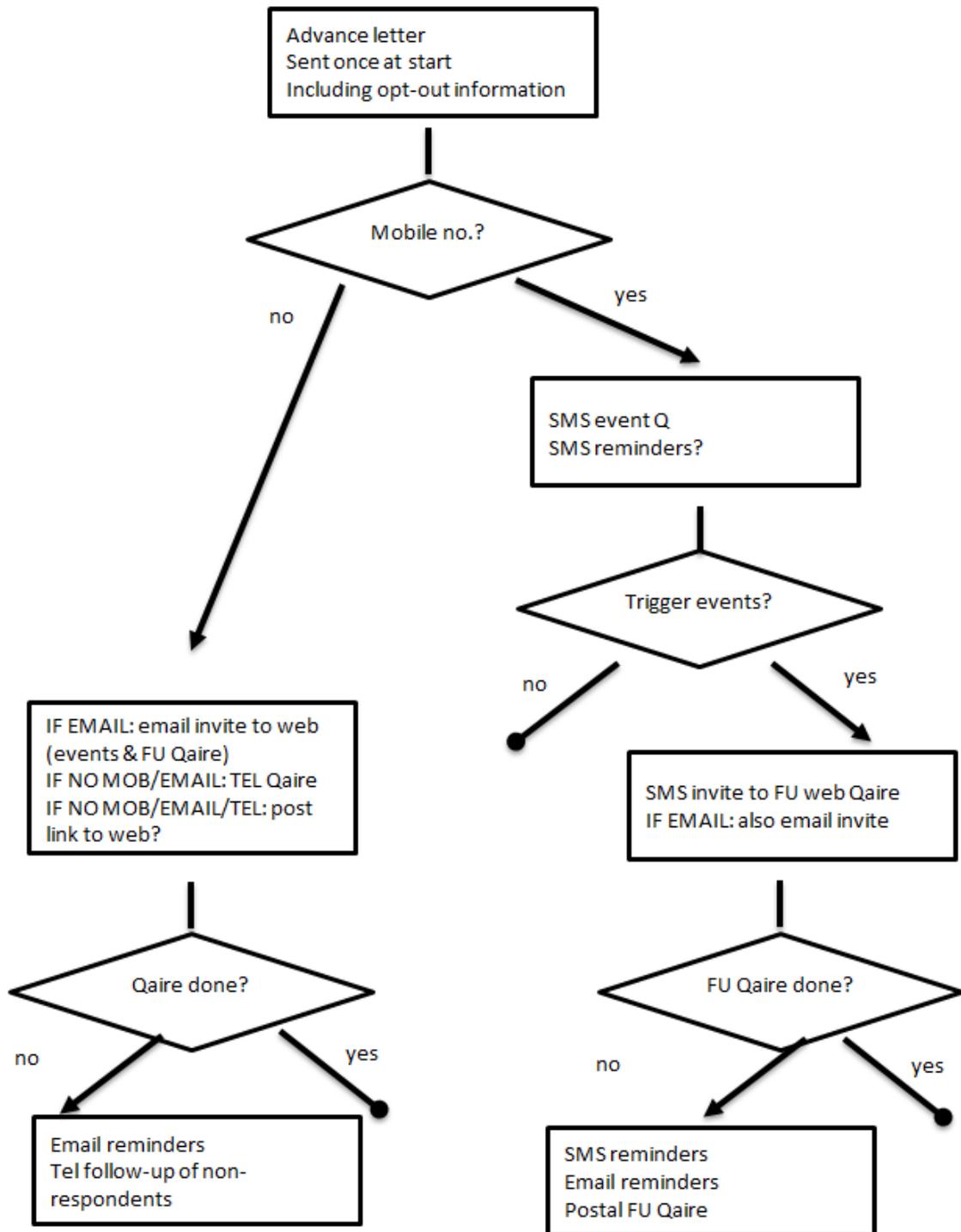
The protocol for the subsequent monthly surveys would then depend on contact information obtained during the prior wave of data collection. That is, sample persons would receive a different protocol if they provided a mobile phone number and/or an email address. The primary means of communication would be by SMS (text messages). If a mobile phone number was provided and the sample person did not opt out, they would be sent an SMS. Ideally they would respond by simply texting their answers. Alternatively, the text message would contain a link to a web survey where they could answer this question. If respondents reported any of the events to have occurred in the past month, they would be sent a link via SMS to complete the relevant follow-up questions online.

If sample persons did not provide a mobile phone number, they would be sent an email with a link to the single multi code question. For those events that are reported to have occurred, the follow-up questions would be administered online immediately following response to the single question. For those who did not provide a mobile phone number or email address, follow-up will occur either by sending a postal invite to complete the survey online or by administering the short survey by phone.

Depending on the mode of initial contact, subsequent reminders will be sent by SMS, email, or post. Given the expected ease of SMS for the initial task, the plan would be to prioritise SMS over email, if this approach proves. Failing this, email will be used as much as possible, for both speed and cost reasons. In other words, maximising the proportion who complete the event-triggered surveys by electronic means will

enable ISER to deliver appropriately customised surveys to participants on a monthly cycle.

Figure 1: Sketch of protocols for event triggered data collection



Notes: FU – follow-up, Q – question, Qaire – questionnaire

4. Experimental testing of data collection protocols

With the protocol outlined in the previous section in mind, we conducted an exercise to identify key gaps in knowledge and features of the protocol that were important to test before launching a more formal test on the Innovation Panel. This preliminary testing was also driven by time constraints and existing technical capabilities of potential vendors. This exercise included a review of literature (both published and unpublished) and consultation with knowledgeable researchers and practitioners about the feasibility of implementing surveys via text messaging or SMS (short message service).

Some key lessons from the literature review (see, e.g., Conrad et al., 2017; Hoe and Grunwald, 2015; Marlar and Hoover, 2019; Schober, 2018; Van der Heijden, 2017) and consultations are as follows:

- Text messaging (SMS) is increasingly used for interventions and ecological momentary assessment (EMA) or similar rapid repeated measurements in the health field, often on a small scale (see, e.g., Cárdenas and Stormshak, 2019; Lawson et al., 2019; Turner et al., 2019).
- While text messaging is increasingly used for invitations and reminders to surveys, its use as mode of survey data collection is still relatively rare.
- Obtaining permission before sending text messages is important.
- Messages longer than 160 characters will be chunked and concatenated, but may not arrive in sequence.
- We found no-one who has successfully field a multi-code question in SMS; most argued for using a series of single-code questions.

In summary, this review identified text messaging (SMS) as a potentially useful mode of data collection for short frequent surveys such as we envision. However, careful development and testing is needed for effective implementation at scale.

This preliminary work led to the decision to use an access panel to trial the use of text messaging to administer or invite participants to complete a single (check-all) question survey to eligible respondents asking about 5 types of events.

During the procurement phase we found that none of the five fieldwork organisations we had approached for quotes had the technical capacity to use text messaging to administer survey questions. We were therefore not able to compare the SMS implementation of the events question with a web-based implementation. We did identify one market research organisation who said they could administer text message questions on our (*Understanding Society*) sample, but they did not have an online access panel that we could have used for this initial stage of experimental testing. This led to a change in focus to testing text messages for survey invitations.

We contracted NatCen Social Research to collect experimental survey data, to test two aspects of the protocols outlined in the previous section: 1) how best to word the monthly events question, and 2) how best to invite sample members to this events question.

4.1 Data

All sample members invited to this survey had been part of the British Social Attitudes survey in 2015 and 2016, before being recruited into the NatCen panel. Only sample members for whom both an email address and a mobile phone number were known were eligible for this study, which was fielded in June 2019. The issued sample size was 2,224.

Sample members were randomly allocated to four treatment groups of equal sizes:

- 1 = Email invitation to the survey & single code event question
- 2 = Email invitation to the survey & multi code event question
- 3 = SMS invitation to the survey & single code event question
- 4 = SMS invitation to the survey & multi code event question

For all four groups the questionnaire was implemented as an online survey. Groups 1 and 2 received an email invitation to this survey, while groups 3 and 4 received the invitation in a text message. In both groups, non-respondents received up to two reminders in the same mode.

The 'multi code' event question was worded as follows, with randomised response options:

Which of the following, if any, have you experienced since 1st January 2019?

Please select all that apply

1. *Diagnosed with a new health condition or entered hospital as an in-
/outpatient*
2. *Became pregnant / partner became pregnant*
3. *Changed jobs, started or stopped working*
4. *Moved house*
5. *Stopped or started living with a partner*
6. *None of these*

The 'single code' event question was wording as follows, with a randomised list of events:

Have you experienced any of the following since 1st January 2019?

[list of events, as in previous question]

- 1 *Yes*
- 2 *No*

Respondents who answered 'yes' were asked the 'multi code' question as a follow-up (with events randomised in the same order as in the single code question), to ascertain which events the respondent had experienced.

The motivation for testing the 'single code' version of the question was that we found no examples of existing surveys that have asked multi code questions by text messaging. The concern is that respondents can answer such a question in any way, since it is not possible to restrict answer formats in text messages. This could require

a lot of processing to identify which events have been reported, before triggering any follow-up questionnaires. With the 'single code' question, the follow up questionnaire could be triggered for everyone answering 'yes', and the question to ascertain which events were experienced could be asked as the first question in the follow-up questionnaire and used to determine which modules to route the respondent into.

Note that we asked about events experienced in the previous six months (since January 2019), in order to maximise the number of respondents who would report events.

Following the events question, all respondents were asked a series of follow-up questions. For each event reported they were asked when this had occurred. In addition, everyone was asked a series of health questions and a question about how frequently they use text messaging. Socio-demographic questions were not asked as part of this study, instead characteristics of sample members were fed forward from previous surveys. These included gender, age, education, labour market activity, relationship status, household type, number of own children (0-18) in the household, housing tenure, and Government Office Region.

4.2 Results

RQ1: What is the best way of asking about events in the last month?

The wording of the events question had no effect on the percentage of respondents who reported at least one event, on the number of events respondents reported, or on the frequency of reporting the individual event types (Table 1). These results therefore suggest that the monthly event question could potentially be asked as a monthly yes/no question by text messaging, without affecting the rate of reporting that events have occurred.

Table 1: Reporting of events by event question wording

	Single code events Q		Multi code events Q		P-value
	N	%	N	%	
Reported 1+ events	105	25.4	110	26.4	0.738
No. of events reported: 0	309	74.6	307	73.6	-
No. of events reported: 1	93	22.5	96	23.0	-
No. of events reported: 2-3	12	2.9	14	3.4	0.906
Reported health event	48	11.6	58	13.9	0.317
Reported pregnancy	2	0.5	7	1.7	0.096
Reported labour market activity change	40	9.7	35	8.4	0.523
Reported residential move	18	4.3	20	4.8	0.757
Reported partnership change	12	2.9	4	1.0	-
N	414		417		

Notes: P-values from χ^2 tests.

Although we asked about events reported over a six month period (January to June 2019), the rates of events reported were relatively low: only a quarter of the sample reported any events, and most reported only a single event. Note that we do not attempt to draw any conclusions about the accuracy of the reporting of events, given the nature of the sample which was selected to ensure comparability of experimental treatment groups (in terms of availability of email and mobile phone contact details) rather than population representativeness. However we flag up this low rate of events, since the qualitative research suggested that respondents could tire of reporting 'no events' and become annoyed about repeated questions about events that they are unlikely to experience. On the other hand, the burden of answering follow-up questions on events that occurred is likely to be low, given the relative infrequency of these events occurring. This suggests that we could ask about more frequently occurring events monthly and rarer events less frequently. It also suggests that the single question (did any of these events occur) may be less annoying since respondent do not have to identify which events did not occur.

The numbers of respondents reporting events were too small to test for differences in the correlates of reported events.

RQ2: What is the effect of the invitation mode on response rates to the event question?

The mode of invitation had a large effect on response rates: 45.1% of those invited to the survey by email completed the questionnaire, compared to only 29.7% of those invited by text message ($P < 0.001$).

Comparing the composition of the two respondent samples, there were however no differences between those invited by email and those invited by text messaging, in terms of gender, age, education, Government Office Region, relationship status, household size, number of own children (0-18) in the household, household type, labour market activity, and self-reported health status. There was however a slightly larger proportion of owner-occupiers in the respondent sample invited by email compared to respondents invited by text messaging (79.4% versus 73.6%, $P = 0.035$).

These results suggest that there is no advantage of inviting sample members to the survey using text messaging: this mode of invitation does not seem to bring in different types of respondents than the email invitation does.

5. Conclusion

Our initial development work (including conceptual activities, literature review and qualitative and quantitative testing) has been very effective at identifying how best to implement event-triggered data collection on *Understanding Society*. Among other things, we have learnt:

- At the point of inviting sample members to this new data collection task, we should clearly set out what we would like sample members to do, and why we are doing it.
- We should also clearly set out the expected time commitments, to make it a clear and distinctly different task to the annual survey.
- We should carefully consider how frequently to ask about certain events (e.g., asking every month about changes in stable situations like partnerships, jobs or housing) may be annoying. Similarly, asking about pregnancy status every month

(especially for those who have no pregnancy intentions) may similarly be too intrusive.

- We should consider matching the frequency of the event trigger questions to respondents' situations (e.g., ask monthly about pregnancy status for those who state an intention to become pregnant; ask less frequently for those with no intentions). The single code event question may also help alleviate this concern.
- Given the difficulty of asking a multi code question via text message, and the relative effectiveness of email, we should consider using text messages to supplement (rather than replace) email invitations as a mode of invitation to complete a short web survey (rather than attempting to administer the question via SMS).

In summary, we have learned a lot from the qualitative work about what *Understanding Society* participants would not want to do and how we would best communicate this task to them. From the quantitative work we have also learned about alternative ways to invite sample persons to participate and to ask the event trigger question. This gives us a solid foundation to continue developing and testing the event triggered data collection protocol on the *Understanding Society* Innovation Panel.

References

- Cárdenas, L.E., and Stormshak, E.A. (2019), "Measuring Daily Activity of Emerging Adults: Text Messaging for Assessing Risk Behavior." *Journal of Child and Family Studies*, 28 (2): 315-324.
- Conrad, F.G., Schober, M.F., Antoun, C., Hupp, A.L., and Yan, H.Y. (2017), "Text Interviews on Mobile Devices." In P. Biemer, S., Eckman, B. Edwards, E. de Leeuw, F. Kreuter, L. Lyberg, C. Tucker, and B. West (eds.), *Total Survey Error in Practice*. New York: Wiley, pp. 299-318.
- Hoe, N., and Grunwald, H. (2015), "The Role of Automated SMS Text Messaging in Survey Research." *Survey Practice*, 8 (5).
- Horsley, A., Beninger, K., Day, N. Dhillon, G., Jäckle, A., Burton, J., and Couper, M.P. (2019), "The acceptability and feasibility of asking monthly "life-event" questions in between waves of a panel study." *Understanding Society Working Paper 2019-06*. Colchester: University of Essex.
- Lawson, A., Dalfen, A., Murphy, K.E., Milligan, N., and Lancee, W. (2019), "Use of Text Messaging for Postpartum Depression Screening and Information Provision." *Psychiatric Services*, 70 (5): 389-395.
- Marlar, J., and Hoover, M. (2019), "Leveraging SMS for Survey Research." AAPOR webinar, June 20th.
- Schober, M.F. (2018), "Text Message Interviewing: Data Quality, Efficiency, and Design Considerations." MAPOR Webinar, October 17th.
- Turner, C.M., Arayasirikul, S., Trujillo, D., Le, V., and Wilson, E.C. (2019), "Social Inequity and Structural Barriers to Completion of Ecological Momentary Assessments

for Young Men Who Have Sex With Men and Trans Women Living With HIV in San Francisco." *JMIR mHealth and uHealth*, 7 (5), e13241.

Van der Heijden, P. (2017), "The Practicalities of SMS Research." *International Journal of Market Research*, 59 (2): 157-172.