Journal of Media Psychology

Temporal Considerations for Self-Report Research Using Short Message Service

Erin Walsh and Jay K. Brinker

Temporal Considerations for Self-Report Research Using Short Message Service

Erin Walsh and Jay K. Brinker

Research School of Psychology, The Australian National University, Canberra, Australia

Abstract. When using Short Message Service (SMS) as a tool for data collection in psychological research, participants can be contacted at any time. This study examined how sampling frequency and time of day of contact impacted on response rates, response completeness, and response delay in repeated measures data collected via SMS. Eighty-five undergraduate students completed a six-item self-report questionnaire via SMS, in response to 20 SMS prompts sent on a random schedule. One group responded across 2 days, the other on a compressed schedule of 1 day. Overall, there was a high response rate. There was no significant difference in response rate, completeness, and delay of those responding across 1 or 2 days. Timing between prompts did not impact on response behavior. Responses were more likely to be complete if prompts were sent during the working day. The shortest time between prompts was 15 min, however, and use of an undergraduate sample limits generalizability. When conducting repeated measures sampling using SMS, researchers should be aware that more frequent sampling can be associated with poorer data quality, and should aim to collect data during the working day rather than mornings or evenings.

Keywords: methodology, data quality, Short Message Service, mobile telephone

There is almost total saturation of mobile telephone ownership in Australian adults (Australian Communications and Media Authority, 2013) and high levels of mobile ownership globally (Anhoj & Moldrup, 2009). This offers access to research participants regardless of their location (Haller, Sanci, Sawyer, Coffey, & Patton, 2006). Short Message Service (SMS) is a text-only communication method available on almost all mobile telephones. Although the rise of smartphones has led to considerable uptake of other mobile text-based communication technologies, SMS is used by many people daily and is likely to remain relevant in the future (Australian Communications and Media Authority, 2013; Deloitte Touche Tohmatsu Limited, 2014; Mackay & Weidlich, 2009). SMS is particularly suited for repeated measures research due to its preexisting integration into participants' everyday lives and low cost. Its bidirectional nature lends itself to signal-contingent sampling, where participants provide answers when prompted, rather than when an event occurs or on a response schedule without reminders to respond. Though there is a growing body of psychological research using SMS in repeated measures data collection, there is still a dearth of systematic methodological examination of its properties as a research mode (Cocco & Tuzzi, 2012; Tomlinson et al., 2009). One area to be investigated is the temporal properties of SMS as a tool for data collection – specifically, how the timing of sampling (both frequency and time of day) may impact on response behavior.

Sampling frequency can be thought of as the number of responses required over a fixed period of time (such as a day), or the amount of time that has passed since the last response was required, defined here as prompt spacing. Thinking of sampling frequency in terms of prompt spacing is particularly helpful if responses are sought on a random schedule, which is often done to avoid behaviors being altered by the expectation that a response will soon be required (Wheeler & Reis, 1991). The use of a particular sampling frequency is broadly driven by the expected change, variability, or frequency of what is being measured. For example, a mood-tracking questionnaire may be appropriate for shorter intervals (Cranford et al, 2006), but a voting habits questionnaire would be more appropriate over longer periods (e.g., yearly, as in Wright, 1993). When judging an appropriate sampling frequency, particularly when change can be expected to be rapid (e.g., mood), researchers face some decisions: More frequent sampling minimizes recall bias and increases study power (Raphael, 1987; Raudenbush & Xiao-Feng, 2001); however, increased frequency places a greater burden on participants (Bolger, Davis, & Rafaeli, 2003), which may reduce compliance and produce poorer quality data (Ebner-Priemer & Kubiak, 2007).

Using SMS for repeated measures self-report data collection has occurred across a wide range of sampling frequencies, from eight times daily for 21 days (Berkman, Dickenson, Falk, & Lieberman, 2011) to monthly over

the course of a year (Shrewsbury et al., 2010). Ebner-Priemer and Kubiak (2007) note the lack of structured investigation of methodological aspects of repeated sampling designs and limited justification for sampling frequency in the wider repeated measures literature, a criticism that applies to the mode of SMS. There is one example where sampling frequency of SMS was systematically manipulated (Conner & Reid, 2012); however, this was focused on the implications of oversampling on the construct being measured, rather than the impact on response behavior itself. Such research addresses the inevitable interactions between research topics, question types, sampling frequencies, and resultant response behavior, but more methodologically focused investigations are required to form basic foundations for expected response behavior.

One aspect of data quality is response behavior. This is how participants engage with research in terms of response rates, response completeness, and response delay. Response rates can be impacted by many factors, such as poor study design or a difficult-to-reach sample (Dillman, Smyth, & Christian, 2009), but high response rates maximize the sample size (Fox, Crask, & Kim, 1988) and minimize nonresponse bias, which can threaten the validity of a study's conclusions (Flick, 1988; Groves & Peytcheva, 2008; Shih, 2008). Meta-analyses spanning different topics and methodologies estimate response rates in self-report research to be less than 50% (Baruch, 1999; Baruch & Holtom, 2008; Yu & Cooper, 1983). Repeated measures research is particularly vulnerable to low response rates due to increasing attrition as studies progress (Little, 1995). Response rates in research using SMS for repeated measures also vary widely, from just 23% (Mutua et al., 2012) to 100% (Donaldson, Fallows, & Morris, 2014), but no investigation has examined what may be influencing these different rates.

Response completeness is reduced by skipping items or sending an unfinished response (Sax, Gilmartin, & Bryant, 2003). This may be due to accident or oversight, with participants unsystematically missing questions. It can also be intentional, if participants choose to not answer a particular question. Whether unsystematic or systematic, incompleteness can be problematic where total scores need to be calculated, and when the analysis of choice is not robust to missing data (Mogensen, 1963; Van Buuren, 2010). Perhaps due to the small size of the mobile telephone screen necessitating scrolling between reading and answering questions, SMS responses tend to be less complete than their paper counterparts (Gold et al., 2011). Another contributing factor may be participant burden. Thus overly burdensome sampling schedules could lead to less complete responses.

The delay between when a response is requested and when it is given is also important. In the case of data collection using SMS for research, response delays may be due to participants being away from their phone and thus unaware an SMS has arrived, or participants choosing not to respond immediately. The shorter the delay between when a response is prompted and when it is received, the less scope there is for recall bias to distort results. Self-report research using SMS has encountered a range of response delays, from an average delay of 2 min (Conner & Reid, 2012) up to 60 min (Lepper, Eijkemans, Beijma, Loggers, & Tuijn, 2013). While long delays may not be problematic for all research, it is desirable to minimize delay if the researcher is seeking to measure momentary experience (e.g., current mood).

The association between time of day and response behavior has generally focused on telephone interviews. Researchers are generally urged to contact people when they are likely to be at home and available – outside of business hours (D'Arrigo, Durrant, & Steele, 2009). These specific recommendations may not apply to SMS, because a voice call is typically taken or ignored as it arrives, while an SMS message may be left in the receiver's inbox to be dealt with at their discretion. This asynchronous nature of SMS is often used to provide thinking time in everyday SMS usage (Rettie, 2009), and may allow participants more flexibility in responding. However, this flexibility may cause response delays and missing responses due to participants forgetting that the SMS has been received. Therefore, the logic that time of day may be associated with availability and thus response behavior, may still hold for collecting data via SMS.

This experiment explored how the temporal factors of sampling frequency, prompt spacing, and time of day impacted on response rate, completeness, and delay in signal-contingent self-report repeated measures sampling using SMS. The study paradigm was designed with these methodological questions in mind, and used a short six-item questionnaire with a mixture of question types (binary, Likert, and open-ended). The topic of questions was mental time travel (e.g., remembering a previous meal or anticipating an upcoming social event). It was hypothesized that more frequent sampling and closer prompt spacing would be associated with lower response rates and less complete responses, and shorter response delays (as longer delays are more likely to result in responses being missed altogether). While there is some indication that availability may be related to response, it is not known how time of day will influence SMS response.

Method

Participants

Eighty-five undergraduate students in Australia, aged 17–46 ($M = 21$, $SD = 5.38$), 74% female, participated in return for course credit.[1] To standardize the response

---

[1] In accordance with the ethical principle of provision for withdrawal without penalty, this incentive was not contingent on full completion of all prompts: Course credit was offered regardless of response rate.

experience, a condition of participation was ownership of an iPhone.

Materials

Participants completed two computer-administered questionnaires: an initial demographic information instrument and an exit instrument reflecting on the participation experience. The ongoing participation consisted of a six-item self-report instrument with a mix of question types. The questionnaire was designed to explore methodological issues but was part of a larger study on mental time travel.

The repeated measures questionnaire consisted of (1) a Likert-style rating of current mood, on a scale of 0 = *poor*, to 5 = *good*; (2) a categorical nomination of current temporal orientation of thoughts (remembering/knowing/present/imagining/future thought/other); (3) a binary yes/no response to whether participants intended to do anything based on current thoughts; (4) an open-ended question asking for more information about plans; (5) a categorical nomination of current location (home/work/university/transport/other); and (6) a categorical nomination of the presence of other individuals (alone/with others and not engaging with them/with others and engaging with them). The full questionnaire was sent via SMS to participants once. Prompts to complete it consisted of the text "Please complete the MTT questionnaire. Questions? Email [researcher's email]." As all participants used iPhones, which organize incoming SMS according to number, upon receiving each prompt, participants scrolled to the top of the conversation and replied to the original SMS.

Procedure

Data collection took place in 2011 and 2012. Participants met with the researcher to complete the initial computer-administered instrument. At this time, they confirmed they owned an iPhone and had sufficient credit to send the SMS required by the study, and provided their mobile telephone numbers. The researcher sent all participants an SMS containing the six-item questionnaire, and informed participants they would each receive 20 SMS prompts sent between 8:00 a.m. and 10:00 p.m., on weekdays only, on a randomized schedule. Participants were informed of whether they would be participating in the 1- or 2-day condition. They were asked to respond as soon as possible by replying with their answers. Research has demonstrated successful sampling at 15 min (Ebner-Priemer & Kubiak, 2007), so this was chosen as the shortest interval; however, because we did not want to influence participant expectations, this minimum interval was not disclosed.

Data collection began with all participants responding across 2 days, receiving 10 prompts per day (a total of 20 prompts across 2 days). When preliminary analyses indicated unexpectedly high response rates, the researchers chose to add a condition of higher sampling frequency,

where participants would receive 20 prompts over a single day. While random assignment to responding across 1 or 2 days would have been ideal, the exploratory nature of this research and discovery of an unexpected capacity for a compressed sampling frequency made this unfeasible. This resulted in all participants recruited first being assigned to the 2-day condition. Those recruited after the decision to change sampling frequency were assigned to the 1-day condition. The larger number of participants in the 2-day condition reflected the fact that the 1-day condition was added relatively late in the data collection process. Participants were unaware of the two different sampling schedules when signing up for the study. Aside from sampling across 1 or 2 days, all other aspects of the method (recruitment, instructions, and questions) were identical across the two groups. Once repeated measures data collection was complete, participants met once again with the researcher to complete the computer-administered exit questionnaire. At this point, those not on unlimited texting plans were reimbursed for the cost of sending reply SMS in the course of participation.

Results

A series of general linear multilevel (also known as hierarchical) models, with response behavior nested by participant, was used. Where any variable of interest was missing, that observation (e.g., particular time point for a particular participant) was removed from analysis. Significance of predictors was established by comparison of the model fit with and without the predictor (χ^2), and bootstrapping to create 95% confidence intervals (95% CI) around the slopes (at 10,000 replicates). Response time variables (response delay, and time since last response) were log transformed for analysis due to skewness and bounding.

Given that some attrition may be expected (Little, 1995), responses were examined for systematic degradation of response rates as the study progressed, to establish whether sampling occasion (1 through 20) would need to be controlled for in subsequent analyses. There was no evidence of attrition over the course of the study (Figure 1), and in all cases, adding prompt spacing or time of day to models containing sampling occasion as a predictor led to significantly better fit (at $p < .001$). This indicated that systematic attrition did not need to be controlled for by adding sampling occasion as a covariate into subsequent analyses.

Response behavior was examined in terms of response rates (the number of prompts receiving a response), response completeness (the percentage of questions answered among responses received), and response delay (the number of minutes between sending a prompt and receiving a response). Those responding across 2 days ($n = 62$, 74% response rate, with mean completion 63% and 14-min delay) and 1 day ($n = 23$, 78% response rate, with mean completion rate 70% and 11-min delay) did
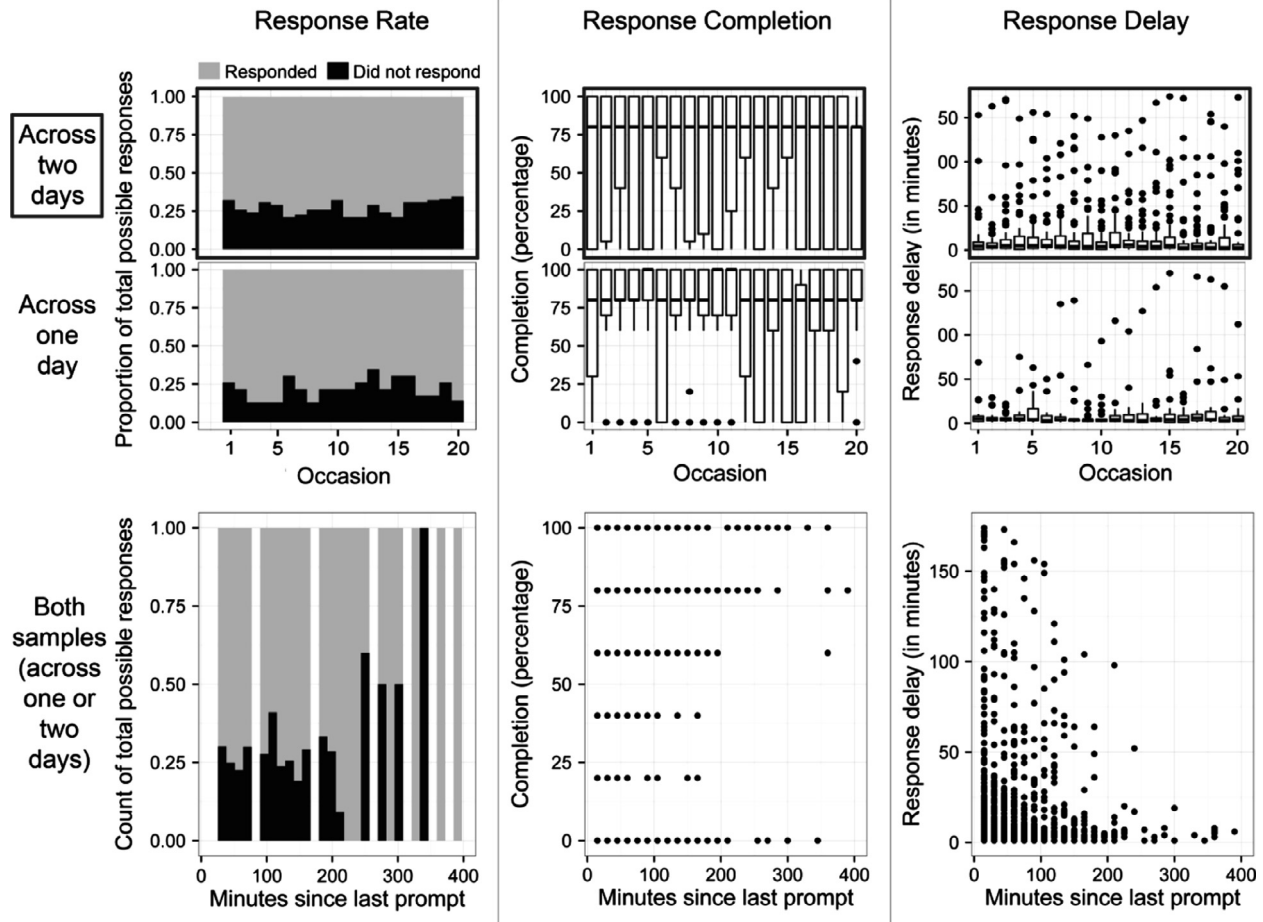
Figure 1. Response behavior, by sampling occasion (top row) and by prompt spacing (bottom row). Separation by sampling across 1 or 2 days in the top row demonstrates similarity in response behavior across different sampling frequencies. The y-axes in the second row have been truncated at 400 min in the interests of readability; outliers of up to 800 min were present.

not significantly differ in terms of response rates, percentage completion, and response delays (Figure 1), $\chi^2(1) = 1.03$, $p = .31$; $\chi^2(1) = 1.34$, $p = .25$; and $\chi^2(1) = 1.03$, $p = .31$, respectively. The intraclass correlation coefficient (ICC) for a logistic nested model, calculated as in Zeger, Liang, and Albert (1988), revealed a vanishingly small advantage gained by nesting data by whether participants samples across 1 or 2 days (ICC < 0.01). Accordingly, it was not included as a covariate or grouping variable.

Prompt spacing did not have a significant effect on response rates, completeness, or delay. As specific ranges in time may be associated with different response behavior patterns (e.g., workday vs. evening), regression trees (with a liberal .0015 complexity cutoff) were used to find split points within the data, to bin time into meaningful groupings. In the following models, time of day was added as a predictor to models already containing prompt spacing. There was no relationship between the time of day a prompt was sent and response rates, or with response delay; the regression tree indicated no split points, and adding time

as a continuous predictor did not significantly improve model fit, $\chi^2(1) = 3.4$, $p = .07$. Neither was there a relationship between time of day and response delay: Adding time binned into seven categories (as indicated by the regression tree) as a predictor did not significantly improve model fit, $\chi^2(7) = 7.2$, $p = .409$.

There was a significant relationship between time of day and response completeness. Regression trees of this relationship revealed four split points, resulting in five bins which we arbitrarily named according to time or most likely activity: morning (8:00 a.m. – 8:30 a.m.), day (8:30 a.m. – 4:30 p.m.), early afternoon (4:30 p.m. – 6:00 p.m.), late afternoon (6:00 p.m. – 6:45 p.m.) and evening (6:45 p.m. – 10:00 p.m.). The mean percentage completion for morning responses was 40%, for day 66%, for early afternoon 60%, for late afternoon 71%, and for evening 60%. Adding these bins as a predictor of percentage completion significantly improved model fit, $\chi^2(4) = 16.379$, $p = .002$. Commute and evening responses were significantly less complete than day responses ($b = -7.49$,

95% CI [−12.57, −2.46]; and $b = −5.69$, 95% CI [−10.00, −1.40], respectively), but morning and afternoon responses did not significantly differ from day responses.

Discussion

This study explored the impact of sampling frequency, prompt spacing, and time of day on SMS response behavior in a signal-contingent self-report repeated measures paradigm. Response rates were higher than the average 50% response rate estimated in the wider psychological literature (Baruch, 1999; Baruch & Holtom, 2008; Yu & Cooper, 1983), and surprisingly did not exhibit the expected attrition given the repeated measures nature of sampling (Little, 1995). Though the lack of attrition across sampling occasions may be due to the relatively short time frame of sampling and thus would not hold for research undertaken over a longer time period, the high retention rate is promising for using SMS for intensive sampling. This suggests that SMS is a viable data collection tool for research where frequent self-report data is required.

Overall, responses were prompt, if not quite as quick as the 2-min turnaround of Conner and Reid (2012). This was likely due to the longer instrument used in the current study. It should be noted that the current study could not comment on whether delays were due to participants not seeing the prompt or because they chose to wait to respond. As accidental or active choice to delay responding may have different impacts on response quality, this issue should be investigated further. For example, evening responses were significantly less complete than day responses and this may have been due to participants being asleep and thus missing incoming SMS. Regardless of the cause of response delay, this indicates that researchers considering SMS as a tool for self-report data may expect low response delays. This is of particular utility in areas where retrospective recall bias or intervening events may diminish the validity of self-report data, as may be the case in recording of passing thoughts or of emotions.

In line with the literature, there was a potentially problematic number of incomplete responses (Gold et al., 2011). This problem could be addressed in a number of ways. Firstly, one could simply avoid collecting data at times known to be problematic. The relationship between completeness and time of day is informative, as it suggests that researchers working with an undergraduate sample are likely to maximize the completeness of responses by avoiding times when participants may be traveling and evenings, where possible. Future research could build on the association between response completeness and time of day by examining nonundergraduate samples, as, for example, it is quite likely that working hours would be unsuitable for a professional adult population. Another possible solution to this problem is alternative data collection methods that are similar to SMS but can prevent submission of incomplete responses, such as mobile applications or app-like website environments. Forced response completeness raises ethical concerns by removing the capacity to refuse to answer particular questions. Reminder messages which highlight unanswered questions, but allow the participant to move on, are an alternative which may improve response rates without raising ethical concerns.

The hypothesis that more frequent sampling and closer prompt spacing would be associated with lower response rates and less complete responses, but shorter response delays, was not supported. The indication that higher sampling frequency (and associated closely spaced prompts) does not adversely impact on response behavior is promising. Firstly, the current results suggest that SMS can be a very useful tool for research regarding constructs known to fluctuate on the order of minutes – for example, mood (Cranford et al, 2006). Secondly, it leaves considerable scope to see how far a researcher can push the sampling frequency and prompt spacing before response behavior breaks down. Such investigation should be mindful of ethical considerations of response burden and intrusiveness of repeated questioning, and physical possibility (as prompt spacing must exceed the time it takes to type a response). This relates to both the sampling frequency and the length of the instrument to be completed. It would also be useful to establish how the length of the instrument may impact on the relationship between sampling frequency and response behavior, keeping in mind that surprisingly lengthy instruments (10 or more items) can be administered via SMS (Walsh & Brinker, 2014).

Though this study was explicitly designed to address broad methodological issues, the generalizability of results is still somewhat bound to the particular questions asked. For example, open-ended questions take physically longer to type than multiple-choice questions, and questionnaires with more open-ended questions may expect longer delays. The particular topic may impact on many aspects of response behavior, including response rates (Cook et al., 2000), where difficult or sensitive questions may translate into greater participant burden and thus lower response rates (Bolger, Davis, & Rafaeli, 2003; Ebner-Priemer & Kubiak, 2007). Generalizability is further limited by the choice to use an undergraduate sample, and limiting respondents to those using iPhones. Despite these limitations, this study sheds some light on what prospective research using SMS may expect, and is a useful beginning for further investigation of sampling frequency when using SMS for self-report data collection.

This experiment explored how sampling frequency, prompt spacing, and time of day impacted on response rate, completeness, and delay. The remarkably low attrition and lack impact of sampling frequency on response behavior suggest that SMS can be a useful method when the researcher wishes to track a self-reported phenomenon across frequent sampling occasions. However, response incompleteness is a potential drawback which is difficult to address without turning to other modes, such as mobile apps. SMS, through lengthier instruments, more frequent sampling, and longer sampling durations. This study is the first step in understanding the temporal issues related

to signal-contingent self-report data collection using SMS. By systematically varying elements of the current study design, such as sample, topic, question format, and sampling frequency, future research can build a valuable framework to guide the application of SMS as a tool for data collection.

References

Anhoj, J., & Moldrup, C. (2009). Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): Response rate analysis and focus group evaluation from a pilot study. *Journal of Medical Internet Research, 6*, e42. doi: 10.2196/jmir.6.4.e42

Australian Communications and Media Authority. (2013). *ACMA Communications report 2012–2013*. Retrieved from http://www.acma.gov.au/theACMA/Library/Corporate-library/Corporate-publications/communications-report-2012-13

Baruch, Y. (1999). Response rate in academic studies: A comparative analysis. *Human Relations, 52*, 421–438. doi: 10.1177/001872679905200401

Baruch, Y., & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human Relations, 61*, 1139–1160. doi: 10.1177/0018726708094863

Berkman, E. T., Dickenson, J., Falk, E. B., & Lieberman, M. D. (2011). Using SMS text messaging to assess moderators of smoking reduction: Validating a new tool for ecological measurement of health behaviors. *Health Psychology, 30*, 186–194. doi: 10.1037/a0022201

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. doi: 10.1146/annurev.psych.54.101601.145030

Cocco, M., & Tuzzi, A.. New data collection modes for surveys: A comparative analysis of the influence of survey mode on question-wording effects. *Quality and Quantity 47*, 3135–3152. doi: 10.1007/s11135-012-9708-1

Conner, T. S., & Reid, K. A. (2012). Effects of intensive mobile happiness reporting in daily life. *Social Psychological and Personality Science, 3*, 315–323.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. *Educational and Psychological Measurement, 60*, 821–836. doi: 10.1177/00131640021970934

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32*, 917–929. doi: 10.1177/0146167206287721

D'Arrigo, J., Durrant, G., & Steele, F. (2009). Using field process data to predict best times of contact conditioning on household and interviewer influences. Retrieved from http://eprints.soton.ac.uk/66318/1/s3ri-workingpaper-M09-12.pdf

Deloitte Touche Tohmatsu Limited. (2014). *Short messaging services versus instant messaging: Value versus volume*. Retrieved from https://www2.deloitte.com/content/dam/Deloitte/au/Documents/technology-media-telecommunications/deloitte-au-tmt-short-messaging-services-versus-instant-messaging-011014.pdf

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys* (3rd ed.). Hoboken, NJ: Wiley.

Donaldson, E. L., Fallows, S., & Morris, M. (2014). A text message based weight management intervention for overweight adults. *Journal of Human Nutrition and Dietetics, 27*, 90–97. doi: 10.1111/jhn.12096

Ebner-Priemer, U. W., & Kubiak, T. (2007). Psychological and psychophysiological ambulatory monitoring. *European Journal of Psychological Assessment, 23*, 214–226. doi: 10.1027/1015-5759.23.4.214

Flick, S. N. (1988). Managing attrition in clinical research. *Clinical Psychology Review, 8*, 499–515. doi: 10.1016/0272-7358(88)90076-1

Fox, R., Crask, M., & Kim, J. (1988). Mail survey response rate a meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly, 52*, 467–491. doi: 10.1086/269125

Gold, J., Aitken, C. K., Dixon, H. G., Lim, M. S. C., Gouillou, M., Spelman, T., Wakefield, M., & Hellard, M. E. (2011). A randomised controlled trial using mobile advertising to promote safer sex and sun safety to young people. *Health Education Research, 26*, 782–794. doi: 10.1093/her/cyr020

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly, 72*, 167–189. doi: 10.1093/poq/nfn011

Haller, D., Sanci, L., Sawyer, S., Coffey, C., & Patton, G. (2006). R U OK 2 TXT 4 RESEARCH? Feasibility of text message communication in primary care research. *Australian Family Physician, 35*, 175–176.

Lee, S. S. S., Xin, X., Lee, W. P., Sim, E. J., Tan, B., Bien, M. P. G., . . . Thumboo, J. (2013). The feasibility of using SMS as a health survey tool: An exploratory study in patients with rheumatoid arthritis. *International Journal of Medical Informatics, 82*, 427–434. doi: 10.1016/j.ijmedinf.2012.12.003

Lepper, A. M. De, Eijkemans, M. J. C., Beijma, H. Van, Loggers, J. W., & Tuijn, C. J. (2013). Response patterns to interactive SMS health education quizzes at two sites in Uganda: A cohort study. *Tropical Medicine and International Health, 18*, 1–6. doi: 10.1111/tmi.12059

Little, R. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association, 90*, 1112–1121. doi: 10.1080/01621459.1995.10476615

Mackay, M. M., & Weidlich, O. (2009). *Australian Mobile Phone lifestyle index*. Sydney, Australia: Australian Interactive Media Industry Association Mobile Industry Group.

Mogensen, A. (1963). Item-skipping and right and wrong solutions in a preliminary version of a multiple-choice vocabulary test. *Acta Psychologica, 21*, 49–54.

Mutua, G., Sanders, E., Mugo, P., Anzala, O., Haberer, J. E., Bangsberg, D., . . . Priddy, F. H. (2012). Safety and adherence to intermittent pre-exposure prophylaxis (PrEP) for HIV-1 in African men who have sex with men and female sex workers. *PLoS One, 7*, e33103. doi: 10.1371/journal.pone.0033103

Raphael, K. (1987). Recall bias: A proposal for assessment and control. *International Journal of Epidemiology, 16*, 167–170.

Rettie, R. (2009). SMS: Exploiting the interactional characteristics of near-synchrony. *Information, Communication & Society, 12*, 1131–1148.

Raudenbush, S., & Xiao-Feng, L. (2001). Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. *Psychological Methods, 6*, 387–401.

Sax, L., Gilmartin, S., & Bryant, A. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*, 409–432. Retrieved from http://link.springer.com/article/10.1023/A:1024232915870

Shih, T. H. (2008). Comparing response rates from Web and mail surveys: A meta-analysis. *Field Methods, 20*, 249–271. doi: 10.1177/1525822X08317085

Shrewsbury, V., Chou, A., Steinbeck, K., Nguyen, B., Baur, L., Lee, A., . . . Thumboo, J. (2010). Adolescent engagement in additional therapeutic contact for overweight management via short message service and electronic mail: The loozit study. *Journal of Adolescent Health, 46*, 15–16.

Tomlinson, M., Solomon, W., Singh, Y., Doherty, T., Chopra, M., Ijumba, P., . . . Jackson, D. (2009). The use of mobile phones as a data collection tool: A report from a household survey in South Africa. *BMC Medical Informatics and Decision Making, 8*, 1–8. doi: 10.1186/1472-6947-9-51

Van Buuren, S. (2010). Item imputation without specifying scale structure. *Methodology, 6*, 31–36. doi: 10.1027/1614-2241/a000004

Walsh, E., & Brinker, J. (2014). *As you Likert: Cross-mode equivalence of administering lengthy self-report instruments via text message*. Paper presented at the ACSPRI Social Science Methodology Conference, Sydney, Australia.

Wheeler, L., & Reis, H. (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of personality, 59*, 339–354. doi: 10.1111/j.1467-6494.1991.tb00252.x/full

Wright, G. C. (1993). Errors in measuring vote choice in the National Election Studies, 1952–88. *American Journal of Political Science, 37*, 291–316. doi: 10.3886/ICPSR08475.v14

Yu, J., & Cooper, H. (1983). A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research, 1*, 36–44.

Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, 1049–1060.

Erin Walsh is a postdoctoral research fellow at the Centre for Research on Ageing, Health and Wellbeing. Her PhD focused on the strengths and weaknesses of text messaging (SMS) as a tool for self-report psychological research. Alongside research, she works as a scientific illustrator, offering assistance with stimuli preparation, data visualization, and figure preparation for scholarly publication or in engaging the general public with science.

Erin Walsh
Research School of Psychology
The Australian National University
Canberra 2615
Australia
Tel. +61 2 6125-5585
E-mail erin.walsh@anu.edu.au

Jay K. Brinker (PhD) is a visiting fellow at the Australian National University College of Medicine, Biology and Environment. Her research interests include the neurocognitive factors in depression in both younger and older adults. She is also interested in the links between rumination, executive functioning, and depressed mood, and neurocognitive and interpersonal factors in depression among younger and older adults.