Augmenting Survey and Experimental Designs with Digital Trace Data[1]

Jeffrey Boase

University of Toronto

Author Note

Jeffrey Boase, Institute of Communication, Culture, Information and Technology and

Faculty of Information, University of Toronto.

Correspondence concerning this article should be addressed to Jeffrey Boase, Institute of

Communication, Culture, Information and Technology , University of Toronto Mississauga,

3359 Mississauga Rd, Mississauga ON L5L 1C6, Canada. E-mail: j.boase@utoronto.ca

---

Augmenting Survey and Experimental Designs with Digital Trace Data

In many cases, digital traces of communication exposure alone do not provide enough information to test theoretically significant hypotheses. Testing these hypotheses can require the addition of self-report information or the manipulation of conditions such that causality can be inferred. Accordingly, a multimethod approach in which digital trace data augments survey or experimental designs may be warranted. How might researchers effectively integrate digital trace data into these types of classic research designs? Rather than focusing on the validity or reliability of digital trace data, this article explores this question through the discussion of two studies developed by the author, Dr. Jeffrey Boase, and his collaborator, Dr. Tetsuro Kobayashi.

The first study involved the development of a mobile app that drew on information found in mobile phone address books, along with calling, texting, and email logs, in order to deliver on-screen survey questions. The purpose of these survey questions was to collect information about individuals with whom participants were actively communicating, without collecting any identifying or potentially compromising information in the process. A company was hired to develop a mobile app that randomly selected an address book name of an individual with whom there had been at least one logged voice call, text message, or email message in the previous 24 hours. The app then displayed the name of this individual in a set of on-screen survey questions. For example, the selected address book name would be included in questions such as, "Is _____ a family member of yours?", or, "Do you talk to _____ in person during a typical day?", "Do you ever talk about politics or social issues with _____?". This process of selecting a name from the address book and using it in a series of on-screen survey questions occurred once per day, over a period of approximately 30 days.

These on-screen survey questions were derived from a "name generator" survey

instrument. The name generator instrument requires an interviewer to first ask the participant to recall the name of an individual based on a given criteria, and the interviewer then ask a series of follow-up questions about this individual. Unfortunately, asking respondents to recall the names of individuals fitting criteria can present a high amount of cognitive burden. This is particularly problematic when respondents are asked about individuals with whom they have recently communicated (Boase & Ling, 2013). We avoided this problem by having the app select a name for the respondent, and having this name selected using recently logged communication data. Moreover, our analysis was able to utilize both the responses to these on-screen survey questions, and non-identifying information gathered from the calling, texting, and email logs. For further information on this study see Boase and Kobayashi (2012) and Kobayashi and Boase (2012).

In a second study a mobile app was developed to leverage digital trace data in an experimental design. The purpose of this study was to examine the extent to which smartphones could help individuals maintain personal networks by countering the effects of relational atrophy. Given that relational atrophy and mobile app use both occur in everyday life, studying this topic in a laboratory would give rise to issues of external validity. To ensure a high degree of external validity a mobile app was developed to collect digital trace data from participants in both the experimental and control groups, and to further deliver a stimulus to participants in the experimental group during their daily lives.

Participants in both the experimental and control groups installed the mobile app which was designed to collect non-identifying logged communication behavior. The app further delivered regular on-screen stimuli to the experimental group. These stimuli asked participants to reconnect with individuals that were in their address book but for whom there had been no

logged calling, texting, or email interactions in the past 60 days. Our assumption was that a lack

of logged calling, texting, and email communication in the past 60 days would be indicative of a

lack of communication in-person and by other means as well. This assumption was supported

through a manipulation check in which the respondents answered several survey questions

regarding how well the app identified individuals with whom they had lost touch.

In this experimental group, digital trace data was also used to measure instances of

reconnection after stimuli had been delivered. For example, if person X's name appeared on a

participant's screen in day two of the experiment, and the phone log indicated that there had been

a phone call to person X on day four of the experiment, this was taken as evidence of

reconnection. This log data was complemented by self-report data in which the respondents

indicated if they had reconnected by means other than calling, texting, or emailing. A detailed

description of this study is available in Kobayashi and Boase (2015) and Boase, Kobayashi,

Schrock, Suzuki, & Suzuki (2015).

Both of these studies demonstrate how it is possible to enhance survey and experimental

research through the use of digital trace data. The first study shows how it is possible to

strengthen an existing survey instrument by drawing on digital trace data to avoid cognitive

burden. Rather than requiring respondents to remember the names of individuals with whom they

had recently interacted, the app drew on the address book and logged interactional data to

provide these names for the respondents. The second study shows how digital trace data can be

woven into an experimental design in order to ensure a high degree of external validity. In this

experiment a mobile app was designed to collect naturally occurring communication data from

both experimental and control groups, and further used to deliver a stimulus to the experimental

group. Although further work is necessary to explore how digital trace data can augment studies

focusing on other kinds of communication exposure, such work is likely to yield fruitful results.

References

Boase, J., & Kobayashi, T. (2012). Mobile communication networks in Japan and America. *China Media Research, 8*(4), 90-98.

Boase, J., Kobayashi, T., Schrock, A., Suzuki, T., & Suzuki. T. (2015). Reconnecting here and there: The reactivation of dormant ties in the US and Japan. *American Behavioral Scientist, 59*(8), 931-945. doi:10.1177/0002764215580588

Boase, J., & Ling, R. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication,18*(4), 508-519. doi:10.1111/jcc4.12021

Kobayashi, T., & Boase, J. (2012). No such effect? The implications of measurement error in self-report measures of mobile communication use. *Communication Methods and Measures, 6*(2), 126-143. doi:10.1080/19312458.2012.679243

Kobayashi, T., & Boase, J. (2015). Emerging from the cocoon? Revisiting the tele-cocooning hypothesis in the smartphone era. *Journal of Computer-Mediated Communication, 20*(3), 330-345. doi:10.1111/jcc4.12116