

SARA: Availability and Missing Data Statistics

Jamie Yap, Mashfiqui Rabbi, Meredith Kotov

September 05, 2019

This document provides summary statistics on availability and missing data from the SARA Micro-Randomized Trial (MRT) and is divided into three sections.

Section 1: summary statistics on availability for each aim.

Section 2: summary statistics on missing intervention assignment for each aim.

Section 3: summary statistics on control covariates for each aim.

The *csv* files under the column “File Name” in the tables below each contain a list of participants and dates that were impacted by various issues/events leading to unavailability or missing intervention assignment. A brief description of these issues/events is given below. Other material relevant to curation and analysis of the SARA MRT data are provided at <https://github.com/jamieyap/SARA>

Table 1: Summary of known issues affecting the outcome, intervention assignment, or control variables by aim

Issue No.	File Name	Aim 1 Impacted?	Aim 2 Impacted?	Aim 4 Impacted?	Issue/Event
0	issue_00.csv	Yes	Yes	Yes	Participant lost access to aquarium app but contacted study staff to regain access.
1	issue_01.csv	No	Yes	Yes	A bug affected the delivery of memes.
2	issue_02.csv	Yes	No	No	Participant did not receive any 4pm notification due to a scheduling error.
3	issue_03.csv	Yes	No	No	A bug caused 4PM notifications to fail to be sent within a set 2 second time limit.
4	issue_04.csv	Yes	No	No	No 4PM notification was sent on a given day because server was full.
5	issue_05.csv	Yes	No	No	App's initial registration with the server failed: after installing the app, the app is registered with external servers before data can be transferred between the app and the server. If the registration fails, it is re-attempted again until registration is successful.
6	issue_06.csv	Yes	No	No	Participant did not provide read access to aquarium app to mobile phone's SD card.
7	issue_07.csv	Yes	No	No	Participants did not have data collection app installed on their phone on certain days of the study.
8	issue_08.csv	Yes	No	No	Participant is assumed to not have charged their phone.
9	issue_09.csv	No	Yes	No	Participant closed main SARA app too soon after completing a survey, preventing randomization from occurring.
10	issue_10.csv	No	Yes	No	Randomization to Aim 2 intervention assignment very likely happened but did not get recorded for these participant days due to issues uploading randomization assignment from aquarium app to the cloud. Details: Randomization assignment from aquarium app can fail to be uploaded to the cloud when (1) internet connection is slow or (2) javascript is a single thread and has no background operations.
11	issue_11.csv	No	No	Yes	Data on randomization assignment on the phone did not sync to the cloud.
12	issue_12.csv	No	No	Yes	Unknown cause unrecorded missing randomization assignment.
13	noncompletion_aim2.csv	No	Yes	No	Survey was not completed on a given participant day.
14	noncompletion_aim4.csv	No	No	Yes	Active task was not completed on a given participant day.
15	-	Yes	Yes	Yes	On last day in study, survey completion or active task completion data in the aquarium app may not have been uploaded to the cloud.

Table 2: Issue 01: Total Participant days impacted by bug in days 1-28 (Aim 2)

No. Participants	No. Participant Days
18	314

Table 3: Issue 01: Total Participant days impacted by bug in days 3-28 (Aim 4)

No. Participants	No. Participant Days
16	279

In Sections 1-3, summary statistics on availability and missing data were calculated after excluding data corresponding to all participants who experienced the bug impacting the delivery of memes (issue 1).

1 Availability

Table 4: Total participant days not available

	Aim 1	Aim 2	Aim 4
No. participant days available	1888	879	740
No. participant days not available	66	521	560
Total participant days in data for analysis	1954	1400	1300

Table 5: Aim 1: Total participants and days not available for 4PM notification per issue/event within days 1-29

File Name	No. of Participants	No. of Participant Days
issue_00.csv	5	10
issue_02.csv	2	16
issue_03.csv	9	19
issue_04.csv	19	19
issue_05.csv	2	2
Total	37	66

Note:

0 participant days are impacted by more than one issue/event.
Hence, $66 - 0 = 66$

Table 6: Aim 2: Total participants and days not available for meme/GIF per issue/event within days 1-28

File Name	No. of Participants	No. of Participant Days
issue_00.csv	4	9
noncompletion_aim2.csv	48	512
Total	52	521

Note:

0 participant days are impacted by more than one issue/event. Hence, $521 - 0 = 521$

Table 7: Aim 4: Total participants and days not available for life insights per issue/event within days 3-28

File Name	No. of Participants	No. of Participant Days
issue_00.csv	4	9
noncompletion_aim4.csv	48	551
Total	52	560

Note:

0 participant days are impacted by more than one issue/event. Hence, 560
- 0 = 560

2 Missing Intervention Assignment

Table 8: Total participant days having missing intervention assignment among available participant days

	Aim 1	Aim 2	Aim 4
No. available participant days having an intervention assignment	1780	858	703
No. available participant days having missing intervention assignment	108	21	37
Total available participant days	1888	879	740
Total participants with any missing intervention assignment on available days	6	13	22

Table 9: Aim 1: Total participants and days having missing intervention assignment among available participant days within days 1-29

File Name	No. of Participants	No. of Participant Days
issue_06.csv	2	55
issue_07.csv	3	52
issue_08.csv	1	1
Total	6	108

Note:

0 participant days are impacted by more than one issue/event.
Hence, 108 - 0 = 108

Table 10: Aim 2: Total participants and days having missing intervention assignment among available participant days within days 1-28

File Name	No. of Participants	No. of Participant Days
issue_09.csv	3	3
issue_10.csv	11	18
Total	14	21

Note:

0 participant days are impacted by more than one issue/event.
Hence, 21 - 0 = 21

Table 11: Aim 4: Total participants and days having missing intervention assignment among available participant days within days 3-28

File Name	No. of Participants	No. of Participant Days
issue_11.csv	22	32
issue_12.csv	1	5
Total	23	37

Note:

0 participant days are impacted by more than one issue/event.

Hence, $37 - 0 = 37$

3 Control Covariates

Table 12: Total participant days with “YES”, “NO”, or “UNKNOWN”among available

	Aim 1	Aim 2	Aim 4
	No. Participant Days	No. Participant Days	No. Participant Days
isCompleted_yesterday_yes=1	1156	705	591
isCompleted_yesterday_no=1	723	170	149
isCompleted_yesterday_unknown=1	9	4	0
Total available participant days	1888	879	740

Table 13: Total participant days with “YES”, “NO”, or “UNKNOWN”among available

	Aim 1	Aim 2	Aim 4
	No. Participant Days	No. Participant Days	No. Participant Days
contact_yes=1	993	370	262
contact_no=1	877	497	478
contact_unknown=1	18	12	0
Total available participant days	1888	879	740

Table 14: Total participant days with “YES”, “NO”, or “UNKNOWN”among available

	Aim 1	Aim 2	Aim 4
	No. Participant Days	No. Participant Days	No. Participant Days
appusage_yes=1	1319	781	635
appusage_no=1	508	79	105
appusage_unknown=1	61	19	0
Total available participant days	1888	879	740