45 min
MRT Case Studies
SARA
JOOL
BariFit

MRT Design characteristics and considerations (for each MRT)
Treatment components
Decision points
Treatment/intervention factors, Availability
Data collection scheme (active, passive)
Outcomes: Proximal and Distal
Primary aim (primary scientific question)
Randomization scheme
Sample size calculation

Secondary Aim: Potential time-varying moderators
Technical issues encountered
Sensors, connectivity, etc.
HCI design issues
Working with app programmers

What is unique about the four MRTs discussed so far?

HeartSteps
Factors and proximal outcomes operating at
different time scales
NIH funded

JOOL
Funded by an app start-up company
Part of their app quality improvement process

SARA
The app is a data collection effort (as opposed to
being
therapeutic, as its primary goal)

BariFit
The primary focus of the MRT was engagement
(Kaiser Permanente)
Funded by a health management organization
Part of their effort to optimize the app prior to
rolling it out
Goals and Outline

Goal
• To illustrate the possibilities when designing a micro-randomized trial by describing three MRT case studies

Four MRT Case Studies

<table>
<thead>
<tr>
<th>HeartSteps</th>
<th>SARA</th>
<th>JOOL</th>
<th>BariFit</th>
</tr>
</thead>
</table>

• For each MRT, we describe: (1) decision points, (2) intervention components (availability), (3) passive/active data collection, (4) proximal and distal outcomes, (5) primary aim/question, (6) randomizations, (7) secondary aims
Target population: Young adult patients between ages of 14 and 24 recruited from the emergency room (ER), who have an Android or iPhone smartphone, and screen positive for past-month binge drinking or past-month cannabis use without a medical marijuana card.

Recruited at Umich emergency room.

Background: Mobile health apps are a promising data collection and intervention delivery tool for substance-using youth, as most young adults own a smartphone. However, engagement with data collection for most mobile health applications is low and often large fraction of users stop providing data after a week of use.

30 day study
**Financial incentives** SARA includes relatively minimal financial incentives and human support to supplement its core engagement strategies: for every three-day streak of self-reporting, i.e., completing the survey and the active tasks each day, participants can earn one dollar. For completing the longer weekly survey on Sunday, an extra 50 cents can be earned. For a 90% adherence rate 95% of participants can earn up to $12 over the 30 day study ($13 maximum). Note that this is a fraction of what daily substance use studies normally pay to collect data (e.g., $1-4 dollars per day).
The base engagement strategy is a **growing virtual aquarium**: In SARA, a virtual aquarium environment grows richer as more data is collected (see Figure 2). Every time a participant finishes a survey or the two active tasks, they earn 30 points towards the aquarium. For the longer survey each Sunday, 50 extra points are rewarded. New fish are unlocked as specific numbers of points are accumulated. Multimedia appendix 2 lists the fish in the SARA aquarium and the corresponding numbers of points that unlock them. SARA is set up so that one fish can be unlocked almost every day if both survey and active tasks are completed. Every time a fish is unlocked, a fun fact about the fish is also given; e.g., when a goldfish is unlocked participants see a message “Do you know goldfish can recognize faces?” An exception is made to the one fish a day rule for the first two days of the study, when SARA provides two fish per day. Initially, these extra fish are given to quickly condition the participant to the fact that interesting fish are unlocked if they self-report. SARA makes the aquarium environment more game-like by introducing levels: after 15 days of data collection, participants graduate from a fishbowl environment to a sea environment. Levels help prevent cluttering as more fish are unlocked, while increasing participant interest. In addition, for streaks of self-reporting, participants can earn treasures such as pearls and gemstones.
Four Push Intervention Components
1) Social-psychological perspectives suggest that reciprocity, i.e., returning a favor, is an innate human tendency. Drawing on these perspectives, SARA sometimes provides incentives before (i.e., unconditional on) self-reporting to facilitate participant reciprocation via subsequent self-reporting. SARA may issue a youth-focused inspirational message as a push notification at 4PM, two hours before the data collection period starts. 4PM was chosen because adolescents or emerging adults are likely to be out of school at that time and hence are likely to notice the notification, and 4PM is close enough to data collection time (6PM) so that providing an unconditional incentive may facilitate participant reciprocation via survey/active tasks completion. To facilitate participant reciprocation, we provide inspirational messages in the form of song lyrics/quotes that might be appealing to youth. Once again, this repository of messages was assembled by undergraduate Research Assistants who are the same age as our target population.

2) Fogg argues that a notification may be necessary to increase persuasiveness. Past research has demonstrated that reminders can increase engagement. SARA provides a message at 6PM that reminds participants to report data. The reminder message is sometimes appended with additional content, such as “you are close to unlocking a new fish”, “you are close to finishing a streak and earning some money” or “it only takes a minute to collect data in SARA”. The additional content tries to spark interest by reminding about the engagement strategies or to inform participants that the self-report is low-effort.
Once participants complete the survey parts of the data collection, they may receive a meme or an animated gif. The topic of meme or gif is either funny or inspirational. Memes and gifs are chosen because they can generate positive emotions and provide encouragement. Further, the non-judgmental nature of included memes or gifs is consistent with other substance use interventions [66]. The memes and gifs in SARA were generated using Amazon’s Mechanical Turk, and reviewed by undergraduate Research Assistants (RAs) who are the same age as the target population.
participants may receive a life-insight after they complete the active tasks portion of the data collection. Life insights are visualizations of data collected in the past. SARA’s life-insights are trends of the various data collected using daily survey and active tasks over the past seven days. SARA contains a life insight for each of the following data types: (i) daily stress, (ii) amount of free time in the day, (iii) degree of loneliness in the day (iv) level of fun on the day (v) how new and exciting were the days (vi) tapping speed (vii) seconds took to finish spatial task. Note that (i)-(v) are gathered from the daily survey and (vi)-(vii) are gather from daily active tasks. Life insights are based on the notion that individuals strive to understand themselves and gain self-relevant knowledge. People are frequently unclear about their personal abilities and they learn about themselves by attending to and seeking self-relevant information. Consistent with this notion, previous work has demonstrated that participants are interested in receiving feedback about their past self-reported experiences; in fact, most health apps and wearables (e.g., fitbit) use visualizations of past data to provide feedback to their users.
Data

Passive data

- Interaction with Application
- Number of surveys completed

Active (self-report) data collection

- Daily free time, fun time, stress, loneliness
- Weekly report of substance use
Proximal and Distal Outcomes

**Proximal Outcomes**

1) Whether or not the survey/active tasks are completed between 6PM and midnight on the same day. The **4PM inspiration message** targets this outcome.

2) Whether or not the survey/active tasks are completed between 6PM and midnight on the same day. The **6PM reminder** targets this outcome.

3) Whether or not the survey/active tasks are completed the next day. The **reward for completing the survey** targets this outcome.

4) Whether or not the survey/active tasks completed between 6PM and midnight on the next day. The **reward for completing the active tasks** targets this outcome.

**Distal Outcome**

- Total number of completed survey/active tasks during the 30 day study.
Primary Hypothesis Tests

1) **4pm notification hypothesis**: Offering (vs. not offering) an inspirational quote increases the likelihood to fully complete current day’s survey and/or active tasks.

2) **After self-report completion reinforcement hypothesis**: Offering positive reinforcement in the form of a meme/gif (vs. not offering meme/gif) following survey completion increases the likelihood to fully complete next day’s survey and/or active tasks.
Randomization

1) 4 PM randomization for all participants:
   – .5 probability of receiving Youth-targeted inspirational quote/lyric
   – .5 probability of receiving no message.

2) 6 PM randomization for all participants:
   – .5 probability of receiving reminder with additional content
   – .5 probability of receiving reminder.
Randomization

3) After survey completion:
   – .5 probability of receiving Reward gif/meme for completing survey.
   – .5 probability of receiving nothing.

4) After active task completion:
   – .5 probability of receiving life insights for completing survey.
   – .5 probability of receiving nothing.
The Substance Abuse Research Assistance (SARA) is an app for gathering data about substance use in high-risk populations. App developers are using an MRT to improve engagement with completion of the self-report data collection measures. At the time this summary was written, this MRT is unique in that it has an engagement component, but not a treatment one. 30 days

**PIs:** Maureen Walton, Susan Murphy, and Mashfiqui Rabbi Shuvo

**Location:** University of Michigan

**Funding:** Michigan Institute for Data Science (PI S. Murphy), University of Michigan Injury Center (PI M. Walton), NIDA P50 DA039838 (PI Linda Collins), NIAAA R01 AA023187 (PI S. Murphy), CDC R49 CE002099 (PI: M. Walton)

https://clinicaltrials.gov/ct2/show/NCT03255317

And

https://osf.io/whgfp/
Sample Size

- Study funding determined the sample size
- Had funding for recruitment until 2/1/2018
- 73 subjects were recruited from emergency room
  - 3 against protocol as had no phone with them
  - 18 suffered through software bugs

We recruited 73. 3 were mistakes because they only had an ipad/no phone. Furthermore 18 suffered from a bug in the code (this bug resulted from an update to the Software in late Dec/early Jan.)
Secondary Aims

1) 6pm reminder notification hypothesis: Offering a reminder with an additional persuasive message (vs. a simple reminder) increases the likelihood to fully complete current day’s survey and/or active tasks.

2) After active task completion reinforcement hypothesis: Offering positive reinforcement in the form of life insights (vs. not offering life-insights) following active task completion increases the likelihood of to fully complete next day’s survey and/or active tasks.
Exploratory Analyses

Examine potential moderators (gender, weekend/weekday, day of study, prior reward accumulation) of the effects of providing the 4PM inspirational message, 6PM reminder, the reward (post survey completion), or the life-insights (post task completion).

Challenges

Creating and iterating prototypes with designers to develop non-monetary rewards.

Running focus groups and pilot studies to fine-tune rewards that are relevant for the young adults.

Developing for both Android and iPhone.

We lost 18 subjects due to bugs in code.
JOOL is a purpose-driven solution for promoting healthy change. It could be categorized as a behavior health and well-being app. Excellent information about the app can be found here: https://www.joolhealth.com/ Below we list information that is about the first micro-randomized trial that JOOL mounted. The goal of the micro-randomized trial was to better when and in what context it is best to push-notifications to encourage users to engage with the app.

89 day study

The primary aim of this quality improvement study was to assess the effect of sending a push notification containing a contextually tailored health message versus not sending a push notification on proximal engagement with the app.
From JOOL’s website

JOOL is an ecosystem of change activation and support. JOOL connects with individuals’ lives and organizations’ cultures dynamically across a range of integrated, data-driven behavioral engagement points.

Remember why we all decided to work in healthcare in the first place? Caring is the great, big beating heart of our industry. Caring about people. Their health. Their lives. Time to get back to what’s always mattered most.
There are 8 different types of messages. The type of message depended on current context.

Decision Points 8:30am, 12:30pm, 5:30pm, 6:30pm, 7:30pm, 8:30pm
Availability
Individuals were only randomized to push or not push if “available=yes”.
Availability depended on various factors:

No individuals were available at 830AM on weekends.
Once a push is made, the individual is no longer available for any remaining time points that day. That is, two pushes can never be sent on the same day.
The remaining availability rules depended on time since last chart:

- Users who have not charted today or yesterday: not available less than 3 days after the last push
- Users who have not charted for between 2 to 9 days: not available less than 2 days after the last push
- Users who have not charted for between 10 to 29 days: not available less than 6 days after the last push
- Users who have not charted for more than 30 days: not available less than 15 days after the last push
Data

Passive data
- Time of day, day of week
- Frequency of charting

Active (self-report) data collection
- Baseline characteristics
- Charting self-reports on sleep, presence, activity, creativity, and eating
Proximal and Distal Outcomes

Proximal Outcomes

1) User engagement with the app, operationalized as whether or not the user charts in the app over the next 24 hours

Distal Outcome

1) Using the app (in particular, charting) is expected to improve SPACE outcomes—sleep, presence, activity, creativity, and eating. These, in turn, are expected to improve energy and willpower, which are expected to improve purpose.

Proximal outcome(s)
To encourage engagement with the app. Specifically, the intended goal of the push notifications was to improve “charting” or engaging with the app over the next 24 hours. (as this is the first and most important interaction a user engages in after opening the app); this is a binary, longitudinal outcome

Distal outcome(s)
Using the app (in particular, charting) is expected to improve SPACE outcomes—sleep, presence, activity, creativity, and eating. These, in turn, are expected to improve energy and willpower, which are expected to improve purpose.
Primary Hypothesis

1) Test whether sending a push notification containing a tailored health message vs not sending push notification (in moments of availability) results in an increased likelihood of proximal engagement with the app (i.e., the user interacted with the app at some point in the next 24 hours).
There were two randomization protocols:

The first was a time-of-day randomization: Users were randomized (uniformly) to one of six times of the day (at which we could push or not push a notification): 8:30AM, 12:30PM, 5:30PM, 6:30PM, 7:30PM, 8:30PM

The second was a push vs no push randomization: Users were randomized with equal probability (50/50) to push or not push an engagement notification.
JOOL is a behavioral health and well-being app that is designed to help people monitor and improve their sleep, presence, activity, creativity, and eating, with the ultimate goal of helping people move closer to fulfilling their life’s purpose. This MRT aims to understand whether push notifications of tailored health messages are useful in promoting engagement with the JOOL app; and, if so, when and under what circumstances they are most effective. 1,255 users of a commercial workplace wellbeing intervention product over 89 days

**PI:** Victor Strecher, PhD, MPH, CEO of JOOL Health

**Location & Funding:** Ann Arbor, MI

**URL:** [https://www.joolhealth.com](https://www.joolhealth.com)
All individuals who had the app installed on their phone and had push notifications enabled between March 2017, when the randomization software was rolled out, until August 2017 were included in this study.

**Sample Size**

- All users who had enabled push notifications during a specific calendar period
- 1255 users
Secondary Aims

1) To test whether the effect of sending a push notification containing a tailored message (vs not sending a push notification) on the likelihood of proximal engagement with the app decreases with time in the study.

2) To test whether the effect of sending a push notification containing a tailored message (vs not sending a push notification) on the likelihood of proximal engagement with the app differs by weekday (Monday to Friday) vs weekend (Saturday or Sunday).

89 day study

Secondary Aim 1: Hypothesis: We hypothesize that the effect of sending a push notification containing a tailored message will differ by week in the study. Specifically, we hypothesize that the effect will be greater earlier in the study than later in the study.

Secondary Aim 2: Hypothesis: We hypothesize that the effect of sending a push notification containing a tailored message will differ by weekday vs weekend. Specifically, we hypothesize that the effect will be greater when the push notification is on weekends than on weekdays.
Exploratory Analyses

Examine potential moderators (gender, BMI, age, time since last chart, usage rate) of the effects of providing a tailored reminder on charting in the subsequent 24 hours.

Baseline features hypotheses include

- BMI: among people with higher BMI, there will be weaker effect of push vs no push
- Gender: among females, there will be a stronger effect of push vs no push
- Age: among younger individuals, there will be a stronger effect of push vs no push

  Implication is that we ought to push more for younger people

  Factor in age when adjusting push probabilities because for older people too much push will turn them off

  The others are the same

Combo feature: younger people are more likely to respond to a push (vs no push) at later parts of the day (after 7:30). Older people are more likely to respond uniformly across all time periods. Younger people will have more preferential times.

Dynamic features hypotheses

- When the time elapsed since the last chart is longer, there will be a weaker effect of push vs no push

  Implication is to not let too much time elapse since the last chart
When the usage rate (# of charts / # of program days) decreases, there will be a weaker effect of push vs no push.

Implication is to push when we see usage rate going up or steady.
BariFit
PI: Pedja Klasnja

Goal: To promote weight loss through increased activity and adherence to Bariatric diet among people who received bariatric surgery.

Funding: Internal funding to improve care at Kaiser

Target population: Patients who have gone through Roux-En-Y Gastric Bypass surgery or The Laparoscopic Sleeve Gastrectomy

- Target population: Patients 18 or older, who have undergone Bariatric surgery at Kaiser Permanente Washington between Jan and May, 2017.
- Background: Sustained weight loss in patients who undergo Bariatric surgery requires them to monitor their weight, increase physical activity, and track their diet on an ongoing basis. After discharge from surgery they have limited support to help initiate these health behaviors as they recover and transition to a normal routine.

4 months following surgery
Reminders to weight themselves regularly

These are reminder messages sent once a week to weight themselves regularly. We will have a bucket of such messages, we just pick one randomly from a bucket and send at a predetermined time every week. There is no randomization. I am not sure if participant decided when he/she wanted to receive this message or the study decided.

Nutrition advice messages

These are messages about how patients should be eating after surgery. What patients are able to eat depends on how far from the surgery they are and what kind of surgery they had. There were buckets of these messages for different numbers of weeks since surgery.

Delivery

We will send one of these messages each afternoon at .6 probability, so people don’t stop paying attention.
Four Push Intervention Components

1) **Set Goals:** at baseline participants are assigned daily variable percentile step goals or always daily 60th percentile step goals

2) **Rest days:** at baseline participants are assigned to a rest day from the goals on average once every 7th day or no rest day from daily step goals

Goals are pushed/delivered via a daily text message in the morning during study.

---

*Factor 1: Percentile goal*

Level 1: Participant always gets 60th percentile step goal. You basically take their daily step count for the previous 10 days, rank order them, and then use the 6th highest step count as the goal for today. (since we will only have 7 days of baseline data when the account is created, for the first two days of the study we can use 5th highest step count, and then from 3rd day on, we can just use the 6th highest step goal).

Level 2: Each day we micro-randomize among the following percentiles: 50%, 70%, 90%, and 120%. Basically, we use 5th, 7th, and 9th highest step count from the previous 10 days, and the highest step count + 20% of that highest count. (e.g., if the highest count over the last ten days is 6530 steps, the 120% percent goal would be 6530 * 1.2 = 7836 steps).

*Factor 2: Rest days*

Level 1: no rest days. The participants in this condition always get a step goal.

Level 2: yes rest days. Every morning we micro-randomize with probability .15 (basically 1 day out of 7) whether to give the person a rest day. If there is a rest-day, we just send them a message “Good morning from Bar-Fit! There is no step goal today. Have a great day.”

For participants who are in this condition and also variable goal condition, the rest
day is the first randomization we do since there is no point in randomizing among 4 goal levels if they are not going to get a goal anywa
Reminder to track food
This is a message that we micro-randomize every morning about the importance of logging what they eat. We will pull a reminder message from a bucket of messages.

Delivery
When they are randomized to receive the reminder, we can just send it right after the goal message. Basically, if a person gets both the step goal and the reminder, they would just get two messages back-to-back from us in the morning.

BariFit contained two types of contextually-tailored activity suggestions: (1) suggestions that recommend that the participant takes a walk of roughly 10 minutes / 1000 steps (walking suggestions); and suggestions that the participant get up and move around a bit—go to a water cooler, stretch her legs, etc. (anti-sedentary suggestions).

Activity suggestions Tailor on weather, weekday/weekend, and time of day, but we won’t have location, so we can either use a heuristic (weekend = home; weekday lunch, mid-afternoon, and evening commute = work; else home) or just use messages that are applicable across locations.

---

**Four Push Intervention Components**

3) **In the morning:** participants are to receive a text message reminder to track their food intake, or no reminder.

4) **Five decision points per day:** Participants are to receive a text message with a tailored activity suggestion, or no suggestion.

*Availability:* Participants were always available for these components (no access to sensors on phone or wearable)
Contextually-tailored suggestions delivered up to five times per day at participant-defined times.
Data

**Passive data collection**
- Step count via Fitbit Charge 2
- Weight via Withings Body wireless scale
- Whether participant tracks food via Fitbit app
- Minutes of MVPA via Fitbit Charge 2

**Active (self-report) data collection**
- Food intake (self-report via Fitbit app)

Moderate to Vigorous Physical Activity (MVPA)
Proximal and Distal Outcomes

Proximal Outcomes

- Total number of steps taken in the following 30 minutes. The tailored activity suggestion targets this outcome.
- Whether or not they tracked their food intake that same day. The reminder to track food targets this outcome.

Distal Outcome

- Growth in step count and minutes of moderate or vigorous activity across the study. The Set Goals and Rest Days components target this outcome.
Primary Aims

This factorial MRT aims to test each of the four main effects:

1) Test the main effect of providing the 60th percentile step goals versus the variable percentile step goals on the growth of stepcount across the study?

2) Test the main effect of providing rest days from step goals versus no rest days from step goals on the growth of stepcount across the study?

3) Test the main effect of providing a reminder to track food (versus not providing a reminder) on whether or not participants tracked their food intake that same day?

4) Test the main effect of providing an activity suggestion (versus not providing an activity suggestion) on the stepcount in the following 30 minutes?
Randomization

1) At baseline decision point:
   - .5 probability of receiving daily text messages with 60% percentile step goals for the duration of the study.
   - .5 probability of receiving daily text messages with variable percentile step goals for the duration of the study.

2) At baseline decision point:
   - .5 probability of receiving rest days from the daily step goals.
   - .5 probability of receiving no rest days from the daily step goals

1) **Set Goals:** at baseline participants are assigned daily variable percentile goals or always daily 60th percentile goals based on past 10 days

2) **Rest days:** at baseline participants are assigned to a rest day from the goals on average once every 7th day or no rest day from the goals
Randomization

3) At daily morning decision point:
   – .5 probability of receiving a reminder to track food.
   – .5 probability of receiving no reminder to track food.

4) At five decision points every day:
   – .3 probability of receiving a tailored activity suggestion tailored.
   – .7 probability of receiving no activity suggestion.
Researchers are conducting this quality-improvement MRT aiming to promote weight maintenance through increased activity and improved diet among people who received bariatric surgery. At the time it was developed, this project was novel in that it implemented separate randomizations at the start of the study, on a daily basis, and five times throughout the day. N=50; 4months

**PI:** Pedja Klasna

**Location & Funding:** Kaiser Permanente
Sample Size

- Study funding determined the sample size
- Had funding for N=50 participants
Technical issues

Since the intervention relied on SMS to increase accessibility, we could not capture context from phones or when (or whether) messages were read.

All procedures conducted remotely (via mail and phone). Participants had variable technological expertise, which led to 4 participants abandoning the study.

Unreliable APIs from Weather Underground server led to some data loss.

Secondary Analyses

1) Test if the main effect of providing a reminder to track food (versus not providing a reminder) on whether or not participants tracked their food intake that same day decreases with time in study.

2) Test if the main effect of providing an activity suggestion (versus not providing an activity suggestion) on the step count on the following 30 minutes decreases with time in study.
Collaborators!
This project tests the feasibility of conducting an MRT aiming to investigate whether real-time sensor-based assessments of stress are useful in optimizing the provision of just-in-time prompts to support stress-management in chronic smokers attempting to quit. The resulting data will be used to inform the development of a JITAI for smoking cessation. 10 days postquit, 4 days prequit.

**PI:** Santosh Kumar, Center of Excellence for Mobile Sensor Data-to-Knowledge (MD2K, [https://md2k.org](https://md2k.org))

**Location:** Northwestern University, Bonnie Spring, (site P.I.)

**Funding:** NIBIB through funds provided by the trans-NIH Big Data to Knowledge initiative ([www.bd2k.nih.gov](http://www.bd2k.nih.gov)). U54EB020404

MD2K smoking cessation study
[https://www.clinicaltrials.gov/ct2/show/study/NCT03184389?recrs=a&lead=Northwestern+University&cntry1=NA%3AUS&state1=NA%3AUS%3AIL&draw=1](https://www.clinicaltrials.gov/ct2/show/study/NCT03184389?recrs=a&lead=Northwestern+University&cntry1=NA%3AUS&state1=NA%3AUS%3AIL&draw=1)