45 min
What is a micro-randomized trial?
   MRT designs principles
      Typical Primary Aim
      Typical Secondary Aims
Using HeartSteps as the motivating example and to show results
A change in daily habits can significantly reduce a person’s risk for heart disease, or, for patients who are already diagnosed, the chance of having an acute coronary event [2,10,84]. Cardiac rehabilitation phase II (CR)—12 to 18-week outpatient programs of supervised exercise, dietary counseling, stress management, and health education—is a central component of secondary prevention of heart disease. Reviews and meta-analyses of cardiac rehabilitation trials have shown that CR can substantially impact physiological and behavioral risk factors—blood pressure, weight, physical activity, smoking, and diet—as well as cardiac mortality and morbidity [21,56,124]. Secondary prevention remains a challenge, however. Although cardiac rehabilitation can jump-start risk reduction, the rates of maintenance of lifestyle changes after cardiac rehabilitation ends have been very low. Fewer than half of patients who complete CR still exercise or eat a low-fat diet six months later [16,105,122]. Yet, after physical activity and heart-healthy diet are discontinued, most of their risk-lowering benefits are lost within a year and a half [77]. **Behavior-change maintenance is, therefore, a key challenge of cardiac risk reduction.**

---

**HeartSteps (PI Klasnja)**

**Goal:** Develop an mobile activity coach for individuals who have coronary artery disease

**Three iterative studies:**

- 42 day micro-randomized pilot study with sedentary individuals,
- 90 day micro-randomized study,
- 365 day personalized study
Questions to Optimize HeartSteps
Suggestions

- Do tailored activity suggestions have an effect at all?
- Do active and sedentary suggestions work equally well?
- Does the effect of suggestions change over time? (e.g., do people get tired of them after a while?)
- When should we send suggestions for optimal effect?
  - Do they work better during certain parts of the day?
  - Do they work better when weather is good vs. bad?
- Is the suggestion effectiveness, including contexts in which they work, different for different types of people?

Traditional evaluation methods don’t help with these questions.
Outline

Micro-Randomized Trials (MRTs)

- Elements of a MRT
- Inferential Target
- Sample size considerations

Using data to inform the development of JITAIIs

45 min
What is a micro-randomized trial?
MRT designs principles
    Typical Primary Aim
    Typical Secondary Aims
Using HeartSteps as the motivating example and to show results
Data from wearable devices that sense and provide treatments

- On each individual: $O_1, A_1, Y_2, ..., O_t, A_t, Y_{t+1}, ...$

- $t$: Decision point
- $O_t$: Observations at $t^{th}$ decision point
- $A_t$: Intervention option at $t^{th}$ decision point
- $Y_{t+1}$: Proximal outcome (e.g., reward, utility, cost)
The momentary times were selected because these times are the times at which most people
are able to be active
Pre-morning commute, mid-day, mid-afternoon, evening commute, after dinner.

Another example: The phone software monitors a risk measure at regular time intervals and if the
risk measures hits a criterion then a treatment is provided.
2) Observations $O_i$
   1) Passively collected (via sensors)
   2) Actively collected (via self-report)

HeartSteps: classifications of activity, location, step count, busyness of calendar, usefulness ratings, adherence........

Can include time of day or day of week and present weather.
Micro-Randomized Trial Elements

3) Intervention options $A_t$
   1) Types of treatments/engagement strategies that can be provided at a decision point, $t$
   2) Whether to provide an intervention

HeartSteps: tailored activity suggestion (yes/no)
Frequently the actions are primarily designed to have a near-term effect on the individual. E.g. Help then manage current craving/stress, help them manage or be aware of the impact of their social setting on their craving/stress.

**Micro-Randomized Trial Elements**

4) Proximal Outcome $Y_{t+1}$

**HeartSteps**: Activity (step count) over next 30 minutes.
Availability

- Interventions can only be delivered at a decision point if an individual is available
- Availability is known prior to delivering an intervention option
- Set $I_t=1$ if the individual is available at decision point $t$, otherwise, $I_t=0$

Availability is not the same as adherence, nor is it the same as interruptibility, receptivity

Available if
1. she is not currently potentially operating a car, (unethical to deliver)
2. she is not currently walking, and (not scientific to deliver) another example is available only if currently classified as at risk.
3. Participant has turned off intervention (unethical to deliver)
3. her phone is connected to the internet. (technical concerns) we added this when we realized there was a bug in the software code that prevented intervention delivery when phone was not connected.

The effect of treatment at a decision time is the difference in proximal response between available individuals assigned an activity recommendation and available individuals who are not assigned an activity recommendation.

Adherence (i.e. compliance) is very different from availability. Suppose a person is available at a decision point. However the phone is in their purse across the room. So they don’t hear whether the phone pings/ see the lockscreen light up. This person is non-adherent at this decision point. Primary analyses will be intention-to-treat and thus will average over non-compliance.
1.1.2. N-of-1 trials. At first glance, the micro-randomized trial design appears similar to the N-of-1 trial design frequently used in the behavioral sciences. However the estimand is quite different. We will, as is typical in statistical causal inference, consider average causal effects, possibly conditional on covariates. In the behavioral field N-of-1 trials are used most often to ascertain individual level causal effects [McDonald et al., 2017]. A variety of nuanced assumptions about individual behavior using behavioral science theory is brought to bear as scientists attempt to triangulate on individual level effects; see the section on “Measuring behavior over time” in McDonald et al. [2017] for a discussion.

In the clinical field, N-of-1 trials were developed for settings in which scientists wish to compare the effect of one treatment versus another (treatment A versus treatment B) on an outcome but it is very expensive to recruit many participants. In both settings a common assumption underlying the analysis of N-of-1 trials is that there are no carry-over effects. Additionally one often assumes that the treatment effect is constant over time.
An excellent overview of N-of-1 designs and their use for evaluating technology based interventions is Dallery et al. [2013]. See Kravitz et al. [2014] for a review of this design in pharmacotherapy trials.
Micro-Randomized Trial Elements

1. **Record** outcomes
   - Distal (scientific/clinical goal) & Proximal Outcome
2. **Record** context (sensor & self-report data)
3. Randomize among intervention options at decision points
4. **Use data after study ends to assess treatment effects**, develop warm-start JITAI
Why Micro-Randomization?

- Randomization (+ representative sample) is a gold standard in providing data to assess causal effects.

- Sequential randomizations (+ representative sample) will enhance replicability of data analyses (moderation, decision rule development).
Outline

Micro-Randomized Trials (MRTs)

- Elements of a MRT
- Inferential Target
- Sample size considerations

Using data to inform the development of JITAI's

45 min
What is a micro-randomized trial?
MRT designs principles
  Typical Primary Aim
  Typical Secondary Aims
Using HeartSteps as the motivating example and to show results
Micro-Randomized Trial

How to justify the experimental trial costs?
• Address a question that can be stated clearly across disciplinary boundaries and be able to provide guarantees.
• Design trial so that a variety of further interesting questions can be addressed.

First Question to Address: Do the intervention options differentially impact the proximal outcome? (aka, is there a main effect?)
Micro-Randomized Trial for HeartSteps

• 42 day trial
• Whether to provide a tailored activity recommendation? $A_t \in \{0, 1\}$
• Test for main effects on proximal outcome
• Randomization in HeartSteps

\[ P[A_t = 1] = .4 \quad t = 1, \ldots, T = 210 \]

A_t = 1 if pushed a message
planned study 2160 decision times.
J=42*5=210 momentary randomizations
Time-varying Main Effects

Time varying potentially intensive/intrusive intervention options $\rightarrow$ potential for accumulating habituation and burden

$\rightarrow$

In the test statistic allow the main effect of the intervention options on proximal outcome to vary with time
Availability & the Treatment Effect

- Interventions can not be delivered at a decision point if an individual is unavailable.

- The effect of an intervention at a decision point is the difference in proximal outcome between available individuals assigned an activity suggestion and available individuals who are not assigned an activity suggestion.
Potential Outcomes

- Define
  \[ \tilde{A}_t = \{A_1, A_2, \ldots, A_t\}, \tilde{a}_t = \{a_1, a_2, \ldots, a_t\} \]

- Define \( Y_{t+1}(\tilde{a}_t) \) to be the observed response,
  \( Y_{t+1} \) if \( \tilde{A}_t = \tilde{a}_t \), e.g., \( Y_{t+1} = Y_{t+1}(\tilde{A}_t) \)

- Define \( I_t(\tilde{a}_{t-1}) \) to be the observed “available for treatment” indicator if \( \tilde{A}_{t-1} = \tilde{a}_{t-1} \)
Main Effect

- Define the main effect at time $t$ as

$$E[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0)|I_t(\bar{A}_{t-1}) = 1]$$

- What does this main effect mean?

Marginal over randomization treatment policy (and effects thereof), conditional on those who have intervention on.

The group who have the intervention turned on is a selected group of people likely depending on the intervention dose they experienced up to time $j$. This intervention dose $\bar{A}_{j-1}$ may have caused burden, may have caused learning.
Main Effect

- The randomization implies that

\[
E[Y_{t+1}(A_{t-1}, 1) - Y_{t+1}(A_{t-1}, 0)]_{I_t(A_{t-1}) = 1} = \\
E[Y_{t+1}|I_t = 1, A_t = 1] - E[Y_{t+1}|I_t = 1, A_t = 0]
\]

- Put

\[
\beta(t) = E[Y_{t+1}|I_t = 1, A_t = 1] - E[Y_{t+1}|I_t = 1, A_t = 0]
\]

Assuming “consistency” THESE ARE CAUSAL EFFECTS!!!!!!
These equations only make sense when the randomization is constant, or time varying but only depending on an exogeneous variable. If randomization is stratified based on an endogeneous time varying variable then the formula changes.
Outline

Micro-Randomized Trials (MRTs)

- Elements of a MRT
- Inferential Target
- Sample size considerations

Using data to inform the development of JITAI

45 min
What is a micro-randomized trial?
MRT designs principles
Typical Primary Aim
Typical Secondary Aims
Using HeartSteps as the motivating example and to show results
Design of MRT

Determine the number of participants so that micro-randomized trial can detect a main effect on proximal outcome

The main effect is a time-varying main effect $\beta(t), \ t=1, \ldots, T$

The main effect is a causal effect.

In heartsteps beta(t) is the effect of tailored activity suggestion on next 30min step count.

Delayed effects which are akin to higher order interactions would be investigated in secondary analyses.
Instead of a sparsity bet, we place a smoothness bet. We are not assuming that the main effect has a quadratic form.

Since the test statistic for the main effect does not depend on time of day, we are averaging any variation in main effect across the occasions during the day (recall we are sizing the study; a primary analysis might be a little more complex and in secondary data analyses one would likely estimate and test if the effect varies by time of day and/or varies by day in study).
Sample Size Calculation

Alternative hypothesis is low dimensional → assessment of the effect of the activity suggestion uses contrasts of between subject outcomes + contrasts of within subject outcomes.

--The required number of subjects will be small.

The contrasts become within person contrasts due to smoothness over time in the targeted quadratic alternative. If the main effect at each time point were to be estimated separately then it would be like a two arm study at each time j.
Centering of $A_j$ protects estimation of $\beta(j)$ from misspecification of $\gamma(j)$. This works for consistency of $\hat{\beta}(j)$ and variance estimation in the data analysis but the sample size formula uses a simplified variance—the simplification assumes the working assumption that $\gamma$ is correctly specified.

Essentially we are estimating the projection of this conditional mean on this space.
Emphasize here that the goal is to construct a test statistic as opposed to estimating beta(t).

Since the model for the proximal effect of Aj on Yj does not depend on time of day, we are averaging any variation in proximal effect across the occasions during the day (recall we are sizing the study; a primary analysis might be a little more complex and in secondary data analyses one would likely estimate and test if the proximal effect varies by time of day and/or varies by j, since j denotes duration in study).

\gamma(t) can include baseline variables: gender, baseline activity level
Alternative for Sample Size Calculation

- One calculates a sample size to detect a given alternative with a given power.

- Alternative:
  \[ H_1 : \beta_i = d_i \bar{\sigma}, i = 0, 1, 2 \]

where \( \bar{\sigma}^2 \) is the average conditional variance.
Alternative for Sample Size Calculation

- Average conditional variance is

\[ \bar{\sigma}^2 = (1/T) \sum_{t=1}^{T} E[VAR(Y_{t+1} | I_t = 1)] \]
Specify Alternative for Sample Size Calculation

- Scientist indirectly specifies standardized $d_i$’s
  - initial main effect: $d_0$,
  - average main effect over trial duration:
    \[
    \frac{1}{T} \sum_{i=1}^{T} \left( d_0 + d_1 \left\lfloor \frac{t-1}{5} \right\rfloor + d_2 \left\lfloor \frac{t-1}{5} \right\rfloor^2 \right),
    \]
  - and day of maximal main effect: $-\frac{d_1}{2d_2}$
- We solve for $d_0$, $d_1$, $d_2$

Specify alternative so that scientist can provide $d_i$’s
Test Statistic for Sample Size Calculation

- Put $Y_i = (Y_{i2}, \ldots, Y_{iT+1})^T$ for $i^{th}$ subject

$q+3$ is the total number of parameters;

$X_i$ is the associated design matrix ($T$ by $q+3$)

$N$ is sample size

Last 3 columns of $X_i$ contain row entries:

\[
I_{it}(A_{it} - q_t), I_{it}(A_{it} - q_t)\left\lfloor \frac{t-1}{5} \right\rfloor, \quad I_{it}(A_{it} - q_t)\left\lfloor \frac{t-1}{5} \right\rfloor^2
\]
Test Statistic for Sample Size Calculation

- “GEE” test statistic is

\[
N \hat{\beta}^T (K \hat{\Sigma} K^T)^{-1} \hat{\beta} = N \hat{\beta}^T (\hat{\Sigma}_\beta)^{-1} \hat{\beta}
\]

where \( \hat{\Sigma} \) is the usual sandwich estimator of the variance-covariance and \( K \) is a \( 3 \times (3+q) \) matrix picking out columns associated with coefficients \( \beta \)

\( \hat{\Sigma}_\beta \) accommodates the within person correlation across time.
Working Assumptions for Sample Size Calculation

1) $E(\epsilon_{it}\epsilon_{it'} \mid I_{it} = 1, I_{it'} = 1, A_{it}, A_{it'})$ is constant.
2) $E(\epsilon_{it} \mid I_{it} = 1, A_{it}) = 0$
3) $Var(\epsilon_{i,t} \mid I_{it} = 1, A_{it})$ is constant.

$$\epsilon_{it} = Y_{i,t+1} - \left( \gamma(t) + \beta(t)(A_{it} - q_t) \right)$$

These working assumptions are false!
Sample Size Calculation

- Under the working assumptions, $\Sigma_\beta$ only depends on polynomials in $\left[\frac{t-1}{5}\right]$, the marginal distribution of $I_t$ and on the randomization probabilities.

- $\Sigma_\beta$ does not depend on the form of $\gamma(t)$

$\Sigma_\beta = (Z^T Z)^{-1}$ divided by $q(1-q)\tau$  
$Z$ is number of time points by 3. The 3 columns correspond to the columns $[A_{j-6}, (A_{j-6}) \lceil \frac{j-1}{5} \rceil, (A_{j-6}) \lceil \frac{j-1}{5} \rceil^2]$ in the design matrix.
Sample Size Calculation

- Under standard moment assumptions, the asymptotic distribution of the “GEE test statistic” is a Chi-Squared on 3 degrees of freedom with non-centrality parameter:
  \[ Nd^T (\Sigma_\beta)^{-1} d \]

- Instead of a Chi-Squared on 3 degrees we use \( \frac{3(N-q-1)}{N-q-3} F_{3, N-q-3} \) with the same non-centrality parameter.

\( N \) is no. of subjects
\( q \) is no of control variables including intercept.
\( q_t \) is randomization prob. at time t
\( p \) is the number of parameters in \( \gamma(t) \). These are the degrees of freedom recommended by Lloyd A. Mancl and Timothy A. DeRouen A Covariance Estimator for GEE with Improved Small-Sample Properties BIOMETRICs 57, 126-134 March 2001

\( \Sigma_\beta = (Z^T Z)^{-1} \) divided by \( q(1-q) \cdot \tau \)

\( Z \) is number of time points by 3. The 3 columns correspond to the columns \([ A_{j-q-t}, (A_{j-q-t}) \cdot \frac{j-1}{5}, (A_{j-q-t}) \cdot \frac{(j-1)^2}{5} \] in the design matrix.
HeartSteps Example

- Standardized $d_i$’s
  - initial effect: $d_0 = 0$
  - output average main effect
  - day of maximal main effect: $-\frac{d_1}{2d_2} = 28$

- Projection used to form test statistic:
  \[
  \gamma(t) + \beta(t)(A_{it} - .4), \quad t = 1, \ldots, 210
  \]
  where
  \[
  \gamma(t) = \gamma_0 + \gamma_1 \left[\frac{t-1}{5}\right] + \gamma_2 \left[\frac{t-1}{5}\right]^2
  \]

Specify alternative so that scientist can provide $d_i$’s

42 days in our study
Average main effect is standardized.
#parameters=6
Meaningful increase in stepcount is 1000/day
Usual std is 2000/day
Roughly a standardized treatment effect of 200/666= .3

<table>
<thead>
<tr>
<th>Standardized Average Main Effect over 42 Days</th>
<th>Sample Size For 70% availability or 50% availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.06 standard deviations</td>
<td>81 or 112</td>
</tr>
<tr>
<td>0.08 standard deviations</td>
<td>48 or 65</td>
</tr>
<tr>
<td>0.10 standard deviations</td>
<td>33 or 43</td>
</tr>
</tbody>
</table>
Same Test Statistic for Analysis

- “GEE” test statistic is
  \[ N \hat{\beta}^T (K \hat{\Sigma} K^T)^{-1} \hat{\beta} \]
  
  where \( K \) is a 3 by 3\(+p \) matrix picking out columns associated with \( \beta \) coefficients

- No working assumptions
The resulting estimator of the variance of \( \hat{\beta} \) is consistent. We are not assuming that \( \gamma(j) \) is correct.
Simulation Results
Type 2 Error Rate (2000 data sets)

<table>
<thead>
<tr>
<th>Average Main Effect (Sample Size)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05(115)</td>
<td>0.790</td>
</tr>
<tr>
<td>0.06(81)</td>
<td>0.794</td>
</tr>
<tr>
<td>0.07(61)</td>
<td>0.800</td>
</tr>
<tr>
<td>0.08(48)</td>
<td>0.801</td>
</tr>
<tr>
<td>0.09(39)</td>
<td>0.798</td>
</tr>
<tr>
<td>0.10(33)</td>
<td>0.803</td>
</tr>
</tbody>
</table>

E[I]=0.7, no initial effect, maximal effect on day 29, randomization prob.=0.4 (originally we thought we would have randomization prob. =.4 but then changed this to .6 after a run in of several subjects prior to real study)

Simulations indicate:
Method is sensitive to
Guess of average amount of time user is available: $\frac{1}{J} \sum_{j=1}^{J} E[I_j]$. Choose on the low side to be safe
Guess of average main effect. Choose on the low side to be safe.
Heteroscedasticity of errors variance of Y when A=1 is larger than variance of Y when A=0 is problematic

Simulations indicate robustness to
$I_{j+1}$ a function of past A_j’s
Guess at day of maximal proximal effect (we use different function from quadratic when this day is less than ½ of the way through the study—this is not presented here)
Non-symmetry or skewness to residual error distribution.
Positive correlated across time residual errors

Mixture of people, some of whom have the intervention turned off x % of time and some who have their intervention turned off y% of the time where overall % time turned off is .7 or .5
Discussion

The micro-randomized trial is a sequential factorial trial with multiple factors, e.g. in HeartSteps:

Factor 1: Activity recommendation is randomized 5 times per day (yes/no)
Factor 2: Daily activity planning is randomized each evening (yes/no)

• https://cran.r-project.org/web/packages/MRTSampleSize/index.html
• https://pengliao.shinyapps.io/mrt-calculator/
This project tests the feasibility and effectiveness of providing, via a smartphone, just-in-time tailored physical activity suggestions as well as evening prompts to plan the following day's physical activity so as to help sedentary individuals increase their activity. The resulting data will be used to inform the development of a JITAI for increasing physical activity.

**PI:** Predrag Klasnja

**Location:** University of Michigan

**Funding:** NHLBI/NIA R01HL125440

heartsteps MRT

https://www.clinicaltrials.gov/ct2/show/NCT03225521?titles=HeartSteps&r=1
Planning a Micro-Randomized Trial?

Be conservative in planning the trial:

1) Under-estimate the amount of time participants are available for the intervention.
2) Under-estimate the average standardized effect
Micro-Randomized Trials: When are they (not) useful?

- NOT USEFUL: When malleable circumstances are rare: Want to learn the best type of alert to prevent suicide attempt.
- USEFUL: When malleable circumstances change rapidly: Stress, urges to smoke, adherence, physical activity, eating.
- NOT USEFUL: Proximal outcome cannot be feasibly assessed.
- USEFUL: Proximal outcome can be unobtrusively sensed or unobtrusively self-reported.
Micro-randomized Trials

A new type of factorial design
i. Time varying factors → time varying main effects, time-varying two-way interactions, different delayed effects
ii. Design studies specifically to detect interactions between factors.

iii. Calculator:
https://sites.google.com/a/umich.edu/pengliao/
Single-patient trials or individual-patient trials or single-case experiments

MRTs vs Other designs

• RCT
• N-of-1 Trials (& Crossover Trials)
• Factorial Designs
MRT vs. Randomized Control trial (RCT)

A randomized control trial (RCT) evaluating a JITAI compared to a suitable control.

- Assumes evidence exists to develop a high-quality JITAI including the
  - choice of tailoring variables & decision rules
- The primary aim of an RCT is to confirm the JITAI’s effectiveness compared to an alternative
  - Is not well suited to constructing or optimizing a JITAI
- RCT is optimal for evaluation
MRT vs. N-of-1 Trial in Clinical Setting

N-of-1 Trials are usually multiple cross-over trials in which the order of the treatments are randomized within a person.

- RCT is too expensive or not feasible
  - Test: Is one-time treatment A better than one-time treatment B?
  - Ideally the treatments should have minimal delayed effects (minimal carryover effects) or N-of-1 design should incorporate a suitable washout period


ABBA trials, single case designs, quasi-experimental

Design and Implementation of N-of-1 Trials: A users’ guide


In the clinical field, N-of-1 trials were developed for settings in which scientists wish to compare the effect of one treatment versus another (treatment A versus treatment B) on an outcome but it is very expensive to recruit many participants. In both settings a common assumption underlying the analysis of N-of-1 trials is that there are no carry-over effects. Additionally one often assumes that the treatment effect is constant over time.
An excellent overview of N-of-1 designs and their use for evaluating technology based interventions is Dallery et al. [2013]. See Kravitz et al. [2014] for a review of this design in pharmacotherapy trials.
N-of-1 Trials are usually multiple cross-over trials in which the order of the treatments are randomized within a person.

- Goal is to ascertain individual level causal effects
  - Test: Is one-time treatment A for Jane better than one-time treatment B for Jane
  - Nuanced assumptions about individual behavior based on theory are brought to bear in order to triangulate on individual level effects.

McDonald et al., 2017

ABBA trials, single case designs, quasi-experimental

1.1.2. N-of-1 trials. At first glance, the micro-randomized trial design appears similar to the N-of-1 trial design frequently used in the behavioral sciences. However the estimand is quite different. We will, as is typical in statistical causal inference, consider average causal effects, possibly conditional on covariates. In the behavioral field N-of-1 trials are used most often to ascertain individual level causal effects [McDonald et al., 2017]. A variety of nuanced assumptions about individual behavior using behavioral science theory is brought to bear as scientists attempt to triangulate on individual level effects; see the section on “Measuring behavior over time” in McDonald et al. [2017] for a discussion.

Suzanne McDonald, Francis Quinn, Rute Vieira, Nicola O’Brien, Martin White, Derek W. Johnston & Falko F. Sniehotta (2017): The state of the art and future opportunities for using longitudinal n-of-1 methods in health behaviour research: a systematic literature overview,

Health Psychology Review, DOI: 10.1080/17437199.2017.1316672
MRTs vs Factorial Experiments

A factorial design

- is an experimental design involving more than one components (e.g., factors); the levels of the components can be meaningfully crossed.

<table>
<thead>
<tr>
<th>Treatment A</th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO</td>
<td>Neither A nor B</td>
<td>A only</td>
</tr>
<tr>
<td>YES</td>
<td>B only</td>
<td>Both A and B</td>
</tr>
</tbody>
</table>

A MRT

- is a special form of a factorial; components are employed sequentially in time within a person.
- components can operate at different time scales
- randomization to subsequent components in a MRT may depend on outcomes of prior components

No notion of availability
MRTs vs Factorial Experiments

Components can be randomized at different time scales, e.g. in HeartSteps:

Factor 1: Tailored activity recommendation is randomized 5 times per day (yes/no)

Factor 2: Daily activity planning is randomized each evening (yes/no)
Collaborators!