Mashfiqui Rabbi | Research statement

In recent times, much of mobile health (mHealth) research has focused on predicting health risks and 'activities of daily living' (ADL), with the primary objective of enhancing the self-management of health behavior. Given that 60% of adults experience at least one chronic disease [1] and 40% of premature deaths result from modifiable behavior [2], this aim of self-managing behavior has never been more imperative. Researchers, however, frequently focus solely on the predictions of risks and ADL, and fall short of investigating how such data can improve the delivery of self-management interventions. Few studies currently demonstrate how to use such data to effectively deliver interventions.

My research focuses on enhancing self-management by personalizing the content and timing of interventions and I have already made many contributions in this field. My Ph.D. thesis, MyBehavior (2015) [3], is the first ML application to deeply analyze phone ADL data and personalize suggestions that are easy to follow. My sub-goals app (2021) [4] moves beyond the traditional "one daily goal" paradigm in health apps (e.g., activity rings in Apple Watch) and divides the daily goal into smaller personalized time-bound subgoal plans. These subgoals prompt users to be active when they are usually active, and these subgoals sum to the daily goal. Another of my contributions is SARA (2018) [5,6], an app addressing the low-engagement problem in mHealth. SARA is the first app to deliver the right incentive at the right time to increase daily self-reports and medication adherence.

My research in personalized intervention spans over 11 years, and I am a pioneering researcher in this field. My research has produced several noteworthy contributions: (1) the first ML-based personalized mHealth intervention, (2) a 10-year impact award at the Tier 1 conference, (3) several research grants and a patent, and (4) a consumer app for Apple devices, currently used by millions [4]. I use an interdisciplinary approach in my research. When data is limited in the beginning, I innovate by steering the ML optimization to mimic the behavior change theories. I then validate and iteratively fine-tune this innovation by creating and deploying user-facing apps in real-world studies. My interdisciplinary approach both fosters collaboration and contributes uniquely to my field. My work has been cited more than 3220 times [7]. I also garnered broad working experience in academia and industry. Throughout my journey, I also identified gaps that require basic science and methodological development (e.g., developing the science and algorithms for context-specific personalized interventions). Since academia better accommodates such science and method development, I am applying to academia.

MyBehavior: Automatically turning sensor data into easy-to-follow suggestions

MyBehavior (2015) [3] is the first ML application that deeply analyzes fine-grained sensor data and issues personalized suggestions that are easy to follow. Although phones and wearables can frequently collect data, current feedback technologies either turn all the data into a single number (e.g., progress towards a 30-minute activity goal) or visualize almost the entire data without processing. These technologies increase user burden because users must manually analyze their data and figure out what to do. MyBehavior, on the contrary, remove user burden by automatically analyzing data using machine learning and goes one step further by creating suggestions that are less burdensome to follow. These suggestions ask for small changes to routine behavior or reinforce past behavior (e.g., take a 10-minute walk near your office. The same walk you did 112 times before).

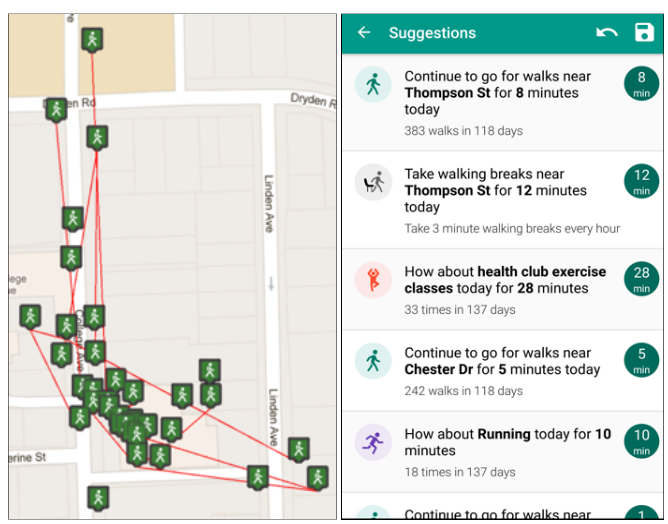

Figure 1: MyBehavior app screenshots (left) location traces of frequent walk, (right) suggestions generated by MyBehavior

MyBehavior creates personalized easy-to-do suggestions in the following way. We track walking, running, driving, and sitting from the phone along with their location. Other activities are tracked manually with a few clicks. We then group similar activities. For sedentary behavior, we cluster similar GPS coordinates from the sitting locations. For walking and running, each episode creates a trajectory of GPS points. We cluster similar trajectories using unsupervised clustering where we use Fréchet distance as the distance measure. We additionally group other activities by their tags. The subsequent challenge is to transform these behaviors into suggestions. We do not have prior data on the successes or failures of interventions to learn from. I overcame this challenge by incorporating psychological theories that tell us what kind of suggestions humans may follow. MyBehavior operationalizes three recommendations from Social Cognitive Theory from psychology: (1) *Small changes of past behavior are easy to do*. For each hour sitting in a place (excluding sleeping hours), MyBehavior suggests 3 minutes of walking for each sitting hour. (2) *Experience in doing some activity makes them easy*. MyBehavior asks users to continue past behavior of walking, running, etc. (3) *If people follow a suggestion, then it will feel easier to follow later*. MyBehavior tracks which suggestions are followed using a multi-armed bandit. At midnight each day, MyBehavior takes 80% suggestions that were most followed and picks 20% suggestions at random. Then at the end of the day, MyBehavior tracks how much of each of the suggestions were followed and updates the following day's suggestions. These recommendations are personalized as the suggestions are generated from one user's data. The computation is done on the phone to preserve the user's privacy.

In two studies, MyBehavior's personalized approach was more effective than generic suggestions from subject matter experts. In a 14-week study (N=16), for a population interested in weight loss, physical activity increased by 8 minutes per day (p<0.05) compared to control [3]. In a second 5-week study with patients with chronic back pain (N=10), walking increased by 5 minutes per day compared to control (p<0.05) [8].

MyBehavior made several contributions. As mentioned, MyBehavior is the first ML application to deeply analyze mobile data and issue personalized interventions. MyBehavior's "easy-to-do" suggestions are also unique to this day, as other efforts could not create less burdensome suggestions with an automated approach. Additionally, the second study on the chronic back pain population received funding from the National Institute of Aging, after subject matter experts found that MyBehavior's low-burden suggestions applied well to the chronic pain population, who are often hesitant and less motivated to be active. 110 million US adults are suffering from chronic pain and physical activity has been shown to reduce pain perception. But patients often fear getting hurt and 40% suffer from depression. MyBehavior's methodology has shown initial success and should be further investigated [8]. Finally, the papers related to MyBehavior have been cited more than 550 times, demonstrating its impact in the research community [7].

Sub-goals: Automated personalized plan to achieve a daily health goal

Goal tracking is a common form of health feedback mechanism on phones and wearables. The activity rings in the Apple Watch is an example of goal-tracking technology that provides real-time feedback on daily goals for moderate activity, stand hours, etc. This technology has not changed in the last two decades and simple goal tracking can fall short in two important ways: (1) Creating a plan to achieve a goal is burdensome and error prone. Real-life contexts are high-dimensional and evolving. Accurately planning around these evolving high-dimensional contexts can cause a significant burden. Many may not make the effort and thus not complete their goal, (2) Earlier in the day people may procrastinate because people discount rewards that are further in the future [9]. As one can complete daily goals at the end of the day, daily goals are less motivating earlier in the day. However, other life needs may show up or people may get tired later in the day, leading to incomplete goals.

Figure 2: Screenshots of the sub-goals app (left), traditional goal tracking, at 12pm, says 20 minutes is completed towards the 30-minute daily goals (right) sub-goals app, at 12pm is providing a more specific 5-minute sub-goal to be completed before 3pm.

The sub-goals app divides a daily goal into several sub-goals where each subgoal is assigned to a specific time segment of the day. When all sub-goals are added together, they sum to the daily goal. These sub-goals also reduce the burden as they ask users to be active when they are usually active. Take the daily 30-minute physical activity goal. Let an individual is usually active for 5 and 10 minutes respectively in the morning and afternoon. Then the suggested subgoals will focus on morning and afternoon, and suggest 10 and 20-minute subgoals respectively for morning and afternoon (note, 10 and 20 sum to the 30 minutes daily goal). The time horizon to complete each subgoal is sooner than the daily goal, thus users will likely discount those rewards less and be more motivated. Subgoals are also generated automatically thus removing the burden of manual planning.

Sub-goal generation works in three steps. We use daily 30-minute daily moderate physical activity as our running example. In step 1, we divide the day into six segments: 6-9am, 9am-12pm, 12-3pm, 3-

6pm, 6-9pm, 9pm-12am. For these segments, we estimate how many minutes of physical activity a user usually does. We use linear regression for estimation with 'time of day' and 'day of the week' as one hot encoded feature. In step 2, we form subgoals by adding a small change to the estimated usual minutes in the first step. We also ensure the sum of the subgoals equals the daily goal. We use a constrained quadratic program to accomplish this task. We minimize the squared distance between the subgoal and the estimated usual minute to ensure a small change to usual behavior. This minimization is done under the constraint that subgoals sum to the daily goal. In step 3, we coalesce small subgoals to give users something meaningful to do. This step is done because we noticed steps 1 and 2 sometimes generate subgoals that are small (0 to 2 minutes). In pilot testing, users told us they wanted subgoals in the 5-minute range. We follow this feedback by coalescing small sub-goals into 5 or more minutes.

We ran an 8-week study on 200 participants where the daily goal was 30 minutes of moderate activity. Compared to the control with one daily goal, participants' daily moderate activity increased by 6.6% per day when subgoals were provided (p=0.09). However, for participants between the age of 35-45, moderate activity increased by 10.4% (p=0.05) per day for subgoals compared to one daily goal.

This sub-goal generation idea had several contributions in the field. In iOS 17, Apple's Fitness+ shipped a similar feature where the 30-minute daily moderate activity goal is divided into subgoal plans [4]. We worked on a follow-up where we replaced steps 1 and step 2 with a data-intensive generative variational encoder neural network and a finite-horizon dynamic decision-making problem. The latter resulted in a patent. A publication is currently in the works.

SARA: Just-in-time intervention for improving engagement

Lack of engagement is a thorny problem in mobile health. Market research shows that nearly 50% of users stop using health apps within their second use [6]. Engagement is crucial for interventions for at least two reasons: (1) interventions that depend on self-reported measures need users to stay engaged and self-report regularly (2) behavior change is difficult, and people may have difficulty acting on therapies [10]. An engaging platform may keep people interested when people struggle to change. How can then one increase engagement? One consideration is the sequential nature of engagement. Engagement strategies may have short-term positive effects due to novelty but have long-term negative effects due to habituation [13]. Engagement strategies also vary in their value and how long their effects last. e.g., money is often more valuable than virtual badges [11]. An altruistic reward can have more lasting effects than a funny meme [12]. Finally, the efficacy of the engagement strategy depends on person-specific factors (e.g., age, culture). Thus, a systemic approach is needed to understand the right engagement strategy for the right time for the right person.

SARA is a platform to experiment and learn the optimal strategy to improve engagement of self-reports. SARA includes various engagement strategies that can either be sequentially randomized or delivered under a fixed schedule without randomization. Such sequential randomization is also known as Micro-randomized trials or MRTs. Currently, four different engagement strategies can be randomized at different probabilities based on an investigator's choosing. One of these strategies is a reciprocity message in the form of an inspirational quote before self-reporting. The other three strategies are post self-report incentives in the form of memes, altruistic messages, and life-insight (life-insights are visualizations of self-tracked data). Two types of fixed incentives are also available: (1) money incentives which are provided for each 3-day streak of self-report (2) a virtual environment, with 60 animated animals and fish, one of which is rewarded after each self-report.

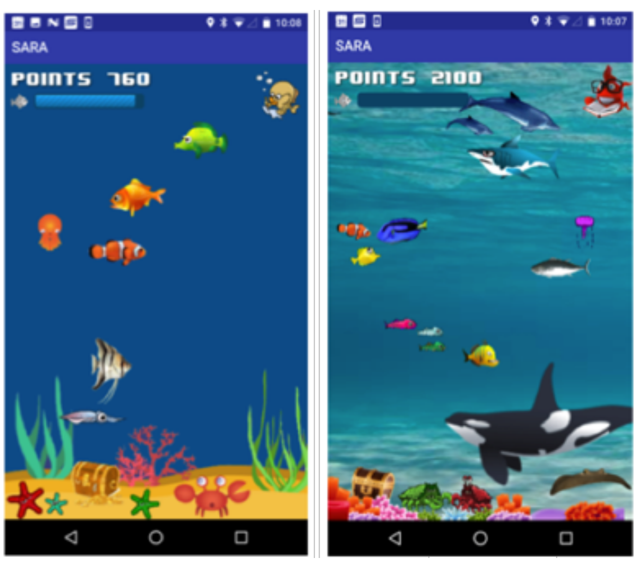

Figure 3: Screenshots of the SARA application

Two MRTs were conducted with SARA. In the first MRT, the target population was younger adults at high risk of substance abuse. In a 30-day trial with 70 participants, we asked participants to complete a survey and two active tasks each day. Participants were randomized to receive inspirational quotes, memes, and life insights at 0.5 probability. We found life-insight increased next-day self-report by 23% if participants did not self-report the day prior. Inspirational messages increased same-day self-report by 6% [5]. All the incentives together had equivalent self-report adherence to a similar study that used 7 times more money [11]. Regular self-report was correlated with decreased substance use (alcohol r=0.4, cannabis r=0.1). In the second MRT (N=30, 6 months), the target population was younger adults who were in remission from leukemia. The goal was to capture daily self-reports of socio-psychological factors that affect medication adherence. Participants were randomized to receive inspirational quotes, memes, and altruistic messages at 0.5 probability. We found memes increased next-day self-report by 51% on weekends and inspirational quotes increased the same day self-report by 8% when self-report was completed the day prior.

NIH funded several grants that are currently using SARA or its technology to improve engagement (K08CA241335, P50DA054039, R01NR020781-01A1). These grants target youth with substance abuse disorder, sickle cell disease, and cancer. We are also developing a reinforcement learning (RL) algorithm to optimize the right strategy for the right time. For sample efficient learning, we are using reward engineering that operationalizes various psychological theories.

Other contribution in sensing and sequential experimentation

Building data-driven personalized interventions requires both sensing and experimentation, and I contributed to sensing and experimental research individually. I began my career working on sensing. I later familiarized myself with sequential experimentation as I started working on sensing-driven interventions. My notable contributions in sensing: In 2011, I lead-authored the first paper that diagnose mental health symptoms from sensor data (cited 289 times) [14]. Later I co-authored an audio-based stress detection on a phone that received a 10-year impact award at a tier 1 conference (Ubicomp, cited 608 times) [17]. Later my system was used by more than a thousand participants in mental health studies at CMU, UW, Dartmouth, and Georgia Tech (e.g., StudentLife [15], CampusLife [16]). My notable contributions in sequential experimentation: I ran a two-week MRT (N=56) that showed, when people are presented with context from earlier in the day, people can more accurately recall stress, mood, etc. from earlier in the day. Such evening recall increased self-report adherence by 26% compared to asking surveys in the moment [18]. In another MRT (N=20), I showed that when notifications are presented with personalized content about physical activity, then people increase their activity and this effect increases when notifications notify about daily goal progress (e.g., at 3 PM, a notification may say "Do 5 minute of activity next hour. Then your usual activity for rest of day should complete your goal"). I did other work in both areas, and I will continue such work to build and validate sensing-driven interventions.

Future work

Building the science and algorithms for just-in-time interventions

Behavior change is hard. Interventions that target one psychological construct (e.g., lower burden or planning) typically cannot change behaviors to a healthy level [24]. Thus, complex intervention, which is a combination of many psychological constructs, is typically used. Experts traditionally handcraft complex interventions by setting rules on which psychological construct to use in which context. However, in mobile health, context is high-dimensional. Phones also capture new measures that are absent in prior psychology literature (e.g., weather, and calendar events). Additionally, behavioral intervention often has longer-term effects (e.g., initial positive effect due to novelty but negative long-term effect due to habituation). Such long-term effect is studied as “Temporal credit assignment” problems which add many more dimensions because we must consider many future possibilities. Given limited human memory, an expert cannot optimally handcraft rules under such high dimensions.

Algorithms can find the optimal interventions for high-dimensional contexts while considering short- and long-term effects. To this end, I want to conduct two types of research: (1) micro-randomized trials (MRTs): In MRTs, participants are sequentially randomized with different interventions over time [19]. From MRTs, one can find causal effects and build the science of which interventions work in which context. The science of “the right intervention at the right time” is underexplored and ripe for scientific discovery. MRTs, however, pose unique challenges in study design and post-study causal inference (e.g., time-varying confounding). I was a postdoc with the creators of the MRT methodology. I also ran five MRTs. In the future, I intend to run MRTs that focus on behavioral and mental health. I want to consider both engagement and therapy together because an engaging platform is necessary for therapy to take effect (2) reinforcement learning algorithms: MRT can be used to further personalize intervention timing using reinforcement learning (RL) [20]. MRT data is equivalent to pure exploration in RL. I’ll use off-policy learning on MRT data to find a better policy to deliver the right intervention at the right time. Some of the key methodological challenges for RL include large state space of contexts, the noisy nature of human behavior, longer-term effect of interventions. These challenges make ‘RL learning’ hard due to large sample size requirements. I will deal with these challenges with sample efficient exploration strategies (e.g., reward learning, hierarchical RL).

Personalized intervention content

Time and again in my research, I found personalized content in interventions is more effective (e.g., low-burden suggestions of MyBehavior [3], life-insights in SARA [5]). The likely reason for this is that personalized messages are relevant and can be highly persuasive if they are based on behavior change theories. In this regard, I want to do two types of work: (1) ML-based personalized messages grounded in behavior change theories: health interventions are often theory-based as theory-based interventions are generally more effective. If personalized interventions are theory-based, they can be even more effective. I intend to continue my work on ML-based personalized intervention content that also operationalizes behavior change theories. One such direction can be personalized "social proof", where two people with similar activity levels can follow each other's success. Furthermore, my past work on sub-goal and low-effort suggestions can be extended to other health behaviors. (2) Notification for personalized intervention content: data-driven personalized content is often related to context. Notifications in the right context can make the personalized suggestion more persuasive. e.g., for the subgoals app, a timely notification before a subgoal can increase the chances of following the subgoal. However, a challenge for such notifications is, that there are many such candidate moments to issue notifications (e.g., we could send a notification for each subgoal). Notifications at all these times can cause burden and app abandonment. I'll learn the best notification times by experimenting uniformly across potential opportune times within a budget (say 1 notification per day). Liao et al [21] described a similar method and I'll extend this method to personalized intervention content.

Actionable latent states for intervention

In behavioral and clinical science, treatments are defined on latent variables [22]. For example, mental health treatments are defined for hopelessness, stress, and sleep clock disturbances. Estimates of these latent variables can be used for automated (e.g., algorithmically) or manual (e.g., by a provider) delivery of interventions. One can manually capture these latent variables using self-report or automated sensing. However, self-reports are burdensome. Automated detection may have large confidence intervals for these latent variables. I intend to use a mix of automated prediction and low-frequency self-reporting to predict the latent states. One direction to explore is a planned missingness design [23], where it can be shown that, with a random sample of self-reports where the randomness is controlled by a researcher, one can obtain unbiased estimates of mean, variance, and covariance of latent states. However, the confidence interval can be large because of less self-reports. One can reduce the uncertainty by controlling how to induce randomness, e.g., using a prediction model to ask when uncertainty is high. This research can enable high-confidence reliable estimates of latent states that an algorithm or medical provider can use to intervene.

Aligning with provider and payer incentives

Healthcare is typically a three-party system of providers, payers, and patients. While improving self-management is patient-focused, aligning the incentives of providers and payers can ensure these systems have a higher impact. Providers (e.g., doctors, and nurses) are interested in lowering burnout and their payments are dependent on making patients healthy. Payers (e.g., insurers) are interested in behavioral health and incentivize programs that reduce costs in the long term. Behavior change interventions can help providers because they can make people healthy. Also, the prediction of latent states can make decision-making easier for providers. Regarding payers, I intend to do both efficacy evaluation as well as how much money the intervention saves in the long run. Mental health intervention also must be coupled with general behavioral care because payers typically sell behavioral health as a service. I intend to work in those directions and make my solutions compatible with payers and providers.

Low-cost solutions for less-privileged community

Social determinants of health (SDoH) are among the highest predictors of health outcomes, and less privileged communities are at higher risk. While diverse solutions are needed to address SDoH, I will focus on early prevention. AI-based interventions can be significant differentiators for early prevention as they are less costly due to automation and can be of higher quality as they are evidence-based. I have run studies in underserved communities and built interventions using AI and user-centered design. This research has already resulted in grants for sickle cell disease, substance abuse, and mental health issues, which disproportionately affect the less privileged. I will continue this research with personalized and low-burden interventions that the underprivileged can follow despite their low resources.

Finally, I have the interdisciplinary experience and partners to deliver on my future research agenda. I have a rich history of interdisciplinary collaboration with computer scientists, statisticians, clinicians, and behavioral scientists. My human-centered approach to AI solutions has already been adopted by individuals, providers, and big tech. I am currently working for the largest payer in the US health system, UnitedHealth, which is also one of the biggest networks of providers. This experience taught me how to build patient-centered interventions that payers and providers are incentivized to support. I will use my expertise and connections to realize my above-mentioned research vision and help people lead healthier lives.

References

[1] https://www.cdc.gov/chronicdisease/resources/infographic/chronic-diseases.htm
[2] Mokdad, Ali H., et al. "Actual causes of death in the United States, 2000." *Jama* 291.10 (2004): 1238-1245.
[3] Rabbi, Mashfiqui, et al. "MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones." *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 2015.
[4] https://support.apple.com/guide/watch/create-a-custom-plan-in-apple-fitness-apd07cef67fc/watchos
[5] Nahum-Shani, I., Rabbi, M., Yap, J., Philyaw-Kotov, M. L., Klasnja, P., Bonar, E. E., ... & Walton, M. A. (2021). Translating strategies for promoting engagement in mobile health: A proof-of-concept microrandomized trial. *Health Psychology*, *40*(12), 974.
[6] Rabbi, Mashfiqui, et al. "Toward increasing engagement in substance use data collection: development of the Substance Abuse Research Assistant app and protocol for a microrandomized trial using adolescents and emerging adults." *JMIR research protocols* 7.7 (2018): e9850.
[7] https://scholar.google.com/citations?hl=en&user=TjQI6gIAAAAJ
[8] Rabbi, Mashfiqui, et al. "Feasibility and acceptability of mobile phone–based auto-personalized physical activity recommendations for chronic pain self-management: pilot study on adults." *Journal of medical Internet research* 20.10 (2018): e10147.
[9] Ainslie, George, and Nick Haslam. "Hyperbolic discounting." (1992).
[10] Hagger, Martin S., et al., eds. The handbook of behavior change. Cambridge University Press, 2020.
[11] Rabbi, Mashfiqui, et al. "Translating behavioral theory into technological interventions: Case study of an mHealth app to increase self-reporting of substance-use related data." arXiv preprint arXiv:2003.13545 (2020).
[12] Seligman, Martin EP. Authentic happiness: Using the new positive psychology to realize your potential for lasting fulfillment. Simon and Schuster, 2002.
[13] Rankin, Catharine H., et al. "Habituation revisited: an updated and revised description of the behavioral characteristics of habituation." Neurobiology of learning and memory 92.2 (2009): 135-138.
[14] Rabbi, Mashfiqui, et al. "Passive and in-situ assessment of mental and physical well-being using mobile sensors." Proceedings of the 13th international conference on Ubiquitous computing. 2011.
[15] Wang, Rui, et al. "StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones." Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing. 2014.
[16] https://smartcities.gatech.edu/news/understanding-campus-life-through-digital-data
[17] Lu, Hong, et al. "Stresssense: Detecting stress in unconstrained acoustic environments using smartphones." Proceedings of the 2012 ACM conference on ubiquitous computing. 2012.
[18] Rabbi, Mashfiqui, et al. "ReVibe: a context-assisted evening recall approach to improve self-report adherence." Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 3.4 (2019): 1-27.
[19] Qian, Tianchen, et al. "The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations." Psychological methods (2022).
[20] Liao, Peng, et al. "Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 4.1 (2020): 1-22.
[21] Liao, Peng, et al. "Just-in-time but not too much: Determining treatment timing in mobile health." Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 2.4 (2018): 1-21.
[22] Kerr, J., Privatdozen, R. W., & Moretti, M. (2005). ABC of Behavior Change (pp. 58-59). Philadelphia: Elsevier Ltd. [23] Little, Todd D., and Mijke Rhemtulla. "Planned missing data designs for developmental researchers." Child Development Perspectives 7.4 (2013): 199-204.
[23] Little, Todd D., and Mijke Rhemtulla. "Planned missing data designs for developmental researchers." *Child Development Perspectives* 7.4 (2013): 199-204.
[24]Duckworth, Angela L., Katherine L. Milkman, and David Laibson. "Beyond willpower: Strategies for reducing failures of self-control." *Psychological Science in the Public Interest* 19.3 (2018): 102-129.