Varun Mishra Research Statement

My overarching research goal is to develop data-driven technologies to enable effective digital interventions for mental- and behavioral-health conditions. The pervasiveness of sensor-rich mobile, wearable, and IoT devices has enabled researchers to passively sense various user traits and characteristics, which in turn have the potential to detect and predict different mental- and behavioral-health outcomes. Upon detecting or anticipating a negative outcome, the same devices can be used to deliver in-the-moment interventions and support to help users. mHealth sensing and intervention is a growing area of research, and while significant work has been done in this domain, several challenges need to be overcome before effective interventions can be provided.

One important factor that determines the effectiveness of digital health interventions is delivering them at the *right time*: (1) when a person needs support, i.e., at or before the onset of a negative outcome, or a psychological or contextual state that might lead to that outcome (*state-of-vulnerability*); and (2) when a person is able and willing to receive, process, and use the support provided (*state-of-receptivity*). As part of my Ph.D. work, **I explore, advance, and contribute to the factors that inform the delivery of an intervention**, specifically by **(1) accurate sensing and detection of different states-of-vulnerability**, like stress and relapse in opioid use disorder (OUD) treatment, and **(2) exploring and identifying the states-of-receptivity** for behavior change interventions.

My research methodology involves building tools and systems to collect human-subjects data, employing various statistical and machine-learning methods for a thorough and data-driven approach to gain insights and build models for a sensing or intervention problem, and deploying those solutions or models in real-world studies to evaluate their efficacy. I am passionate about reproducibility and repeatability of my methods and models across different studies, and it is an underlying principle of my research. *My work thus far has made significant contributions to the respective sub-fields, by addressing fundamental challenges and advancing the current state-of-the-art* [1, 2], *and by contributing new knowledge that can guide the design, implementation, and delivery of future mHealth interventions* [3, 4, 5].

My research is in the broad field of Ubiquitous Computing and lies at the intersection of *mobile/wearable sensing, data science, human-centered computing,* and *behavioral science* . Given the interdisciplinary nature of my work, I have built a strong network of collaborators from several disciplines and institutions. I regularly collaborate with clinicians and psychologists to help lay a theoretical foundation of the mental- and behavioral-health condition we hope to detect, determine the appropriate intervention or support based on the sensed/detected state, and monitor the effectiveness of the intervention. I also collaborate with other computer scientists and engineers to design, build, and deploy the tools and systems needed to conduct studies to answer our research goals.

1. Sensing and detecting states-of-vulnerability

My work has thus far focused on sensing and detecting two different mental- and behavioral-health outcomes: stress and relapse in OUD treatment. I thus have expertise in a spectrum of sensing methodologies and modeling approaches, from fine-grained (*minute-by-minute*) time-series physiological signals for measuring stress to a long-term passive smartphone and wearable sensing and behavioral modeling for relapse during OUD treatment.

1.1. Detecting stress

Timely detection of an individual's stress level has the potential to improve stress management, thereby reducing the risk of adverse outcomes, e.g., smoking, anxiety, depression, or drug use. Recent advances in wearable sensor technologies have led to a variety of approaches for detecting physiological stress. Even with over a decade of research in the domain, however, there still exist many significant challenges, e.g., reliance on custom-made or clinical-grade sensors; lack of deployments with different user groups; and variance in devices, sensors, and methodologies used across studies, making it difficult to compare two studies. These challenges have resulted in a near-total lack of repeatability and reproducibility across studies. *My research makes significant contributions in trying to address these challenges by proposing reproducible methods for stress detection and providing concrete recommendations regarding the practical deployment of these methods and models.*

With collaborators from several disciplines, I designed and conducted multiple controlled and free-living studies for stress detection. My evaluations showed that with a careful data-processing pipeline, commodity devices like

Polar H7, a consumer chest-worn fitness band, performed at par with clinical-grade devices for stress detection. I also proposed a novel two-layered approach using Bayesian Network Models that accounted for the temporal dynamics of stress, that performed significantly better than the traditional methods for detecting stress [1]. I evaluated the reproducibility of my methods with data from 90 participants from four studies, with varying study protocols and research goals. I found that my methods consistently led to improved stress-detection performance across all studies, irrespective of the device type, sensor type, or the type of stressor [2]. In addition, to enable easy deployment of these models to new participants, I proposed a clustering-based approach to determine the stressed/not-stressed threshold, which consistently performed better than choosing a pre-determined threshold based on the training data. *My methods, findings, and recommendations have important implications and could be utilized not just for stress detection but also for other health outcomes that leverage continuous physiological signals*.

Although I built robust models for stress detection, deploying them to free-living conditions remains a significant challenge. The major reason being that free-living conditions have various confounding factors that can cause similar physiological arousal as stress, e.g., physical activity, drinking caffeine, or even listening to music. My work at IBM Research showed that the *context* of a person plays a significant role in their perception of stress, and just using physiological signals was not enough. Combining physiological signals and the contextual information led to the most accurate stress detection results in a study with 23 participants in free-living conditions [7]. I plan to continue working in this direction by exploring ways to incorporate high-level context and fine-grained physiological signals to enable just-in-time interventions for stress.

1.2. Detecting relapse in Opioid Use Disorder (OUD) treatment

Across the U.S., the prevalence of Opioid Use Disorder (OUD) and the rates of opioid overdoses have risen precipitously in recent years. While several effective medication-assisted treatments for OUDs exist, treatment retention is a challenge. Many individuals do not consistently take their medication or remain engaged in treatment -- typically resulting in continued opioid use.

I am currently collaborating with researchers from the Center for Technology and Behavioral Health (CTBH) at Dartmouth, IBM Research, and Kaiser Permanente Medical Center to conduct a study to provide a deeper understanding of patient adherence to OUD treatment. In the study, which is currently underway, we are collecting smartphone and wearable sensor data, Ecological Momentary Assessments (EMA), social media data, and Electronic Health Records (EHR) to track the treatment trajectory of at least 50 patients for three months. I am particularly interested in leveraging this data to detect *relapse*, i.e., use of unprescribed drug, in treatment. As a first step, I am currently using the data collected to model at-risk indicators like mood, pain, stress, and craving. I subsequently plan to use these indicators to estimate the likelihood of relapse at various time-frames in the future. *This work is the first to evaluate the feasibility of estimating relapse from passively-collected data and could unlock a wide variety of research in this domain, not just in detection but also in developing effective interventions based on at-risk indicators*.

2. Exploring and detecting states-of-receptivity

Along with accurately detecting states-of-vulnerability, it is crucial to determine moments when a person is able and willing to receive the intervention. However, there is limited research on states-of-receptivity to actual mHealth interventions. *My work is one of the first to take concrete steps in this direction*. In particular, I worked on understanding and detecting user receptivity to interventions aimed at (a) improving physical activity behavior and (b) improving affective well-being while driving.

2.1. Improving physical activity behavior

I collaborated with colleagues from the Center for Digital Health Interventions (CDHI) at ETH Zürich to develop a chat-bot-based digital coach, *Ally*, to deliver activity interventions aimed at improving physical activity behavior by targeting daily step goals. We conducted a study with 189 participants, representative of the German-speaking part of Switzerland. I defined various metrics for gauging participants' receptivity to interventions. I found that various intrinsic factors (i.e., participant-specific characteristics like age, device type, and personality) and contextual factors (e.g., time, location, activity, device usage) played an important role in determining participant receptivity.

Our results have major implications not just on determining receptivity to interventions but also on how researchers could design interventions for specific sub-populations. For example, I found that participants with higher neuroticism scores were faster in replying to the chat-bot but that did not translate to higher conversation engagement. Hence, intervention options for a study with neurotic participants could be designed to include shorter (single-tap) messages spread throughout the day instead of requiring them to engage in a conversation. We also found evidence that receptivity to interventions led to higher goal completion that day, and completing goals on a day motivated the participants to be more receptive the next day, suggesting that receptivity and goal completion have a virtuous effect on each other. Researchers could potentially get a sense of the effectiveness of an intervention based on how participants engage with that intervention and hence could improve, modify, or change the intervention being administered [3].

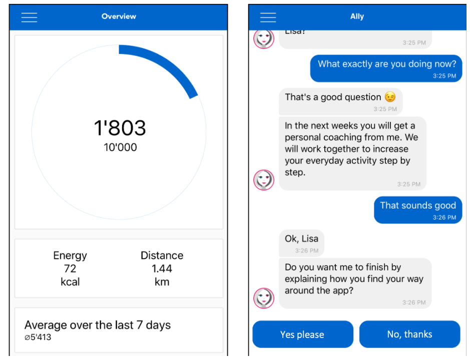

Figure 1: The Ally app: The screenshot on the left shows the dashboard with a step count, calories burnt, and distance walked. The second screenshot shows the chat interface.'

I also built machine-learning models to infer receptivity. Even with fairly simple classifiers, I observed significant improvements in F1-score of up to 77% over a naive classifier baseline. We were awarded a pilot grant from CTBH to deploy these models in a real-world study and get a true sense of their effectiveness. I leveraged the data from the initial study to implement two machine-learning models in the Ally app: a *static* pre-trained model and an *adaptive* model that continuously learned the receptivity of individual participants and updated itself as the study progressed. We conducted a within-subject study with 83 participants spread across the east coast of the United States to compare these ML models' effectiveness to deliver interventions at moments they determined as receptive to delivering interventions at random periods. Our results showed that receptivity to messages delivered by the static model was significantly higher than those delivered randomly. We also observed that although the adaptive model did not show significant improvements for the entire study duration, it showed an increasing trend as the study progressed (Figure 2), suggesting that as the model got more data about a participant, it got better at detecting receptive moments [4].

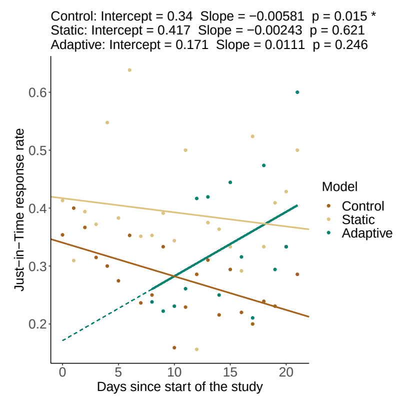

Figure 2: Effect of the different intervention delivery models on the just-in-time response rate over time.

Ours is the first work to demonstrate the feasibility of deploying receptivity-detection models in real-world studies. The findings and methods used in my work would not only be valuable for future work in detecting state-of-receptivity but also in the related field of interruptibility to more general smartphone notifications. I am currently working to make the machine-learning models available to other researchers to use in their studies. In parallel, I am working to improve the adaptive model by using reinforcement learning and evaluating the methods with data from over 1500 participants collected from multiple studies.

2.2. Improving affective well-being while driving

In addition to receptivity to interventions in daily living conditions, I explored how participants interacted with interventions in a constrained environment like in a vehicle. Leveraging the significant amount of time people spend diving in a dedicated space, several in-vehicle interventions aimed at improving drivers' well-being have been developed. It is important, however, to ensure that these interventions do not distract the user from their primary objective of driving. Hence, understanding interaction and receptivity to interventions while driving is of great importance.

To this end, I collaborated with colleagues from the Bosch IoT Lab at ETH Zürich to analyze longitudinal data from 10 participants who were given a retrofitted study car for their daily commute for two months. During each trip, the participants received audio-based mindfulness or music interventions to improve affective well-being while simultaneously recording driving behavior. We found that several high-level trip factors (traffic flow, trip length, and vehicle occupancy) and in-the-moment factors (road type, average speed, and braking behavior) showed significant associations with the participant's decision to start or cancel an intervention. We also identified several driving behaviors that "negated" the effectiveness of interventions and highlight the potential of using such "negative" driving characteristics to better inform intervention delivery. Finally, we compared trips with and without intervention and found that both interventions employed in our study did not have a negative effect on driving behavior [5]. Based on our analyses, *we provide solid recommendations on delivering interventions to maximize responsiveness and effectiveness and minimize the burden on the drivers.*

3. Future Directions

My dissertation work has laid the groundwork for what is a rich and exciting field of study. Building on my current and past work, *I want to continue working in multi-disciplinary teams to develop new methods and tools to enable the successful delivery of digital interventions for various mental and behavioral health conditions*. In addition, I am interested in the following research directions, which I believe would complement my research goals.

3.1. Moving beyond associations, and inferring causality

Most mental and behavioral-health sensing studies evaluate correlations and associations between the passively collected data and the health outcome of interest. While such associations are useful to show that it is feasible to detect a condition, they might not be sufficient to provide effective interventions and support. For ubiquitous devices to provide impactful support, it is crucial to understand and address the actions or situations that *caused* the negative outcome. Causal reasoning is also important if and when clinicians want to leverage sensing data to determine a course of treatment or understand why a particular intervention was provided. Causal inference using passively collected sensor data, however, is a challenge since there could be several missing confounding variables, thus resulting in a biased causal inference. Conducting a study for causal inference must be meticulously planned to obtain sufficient information either by well-structured self-reported questionnaires or by involving additional sensing modalities, and would involve collaborations with engineers, behavioral scientists, and other computer scientists. Initially, I plan to start with causal inference in a dedicated setting or environment before scaling up to free-living conditions. One example could be my current work on evaluating what causes stress while driving. I use heart-rate variability and accelerometer data from 10 participants during their daily commute. I supplement this data with the CAN data from the car (which includes acceleration, braking, and steering movement) and video recordings of the road and the vehicle's inside to get a sense of what the driver was doing while driving. I also use mood and affect self-reports before and after the drive to account for the driver's internal affective state. I believe *this work would be a stepping stone to evaluating causality in other contexts and eventually to my goal for causal inference in free-living conditions.*

3.2. Expanding on the state-of-receptivity

In my current work, I explored and evaluated receptivity only to specific types of interventions, e.g., chat-bot based physical activity interventions, where all interventions had a similar load on the participants. However, it is possible that some treatments have interventions that require a varying degree of user engagement, e.g., conversing with a chat-bot, taking 10 deep breaths, meditating for 10 seconds, going for a walk, or taking a particular medication. Hence, there could be contexts where users are receptive to a specific type of intervention but be unreceptive to a different kind of intervention. I plan to study how receptivity changes with the intervention type or load. This work would give way to other interesting research questions on what type of interventions to deliver under what contexts and deciding whether to deliver a less effective intervention when a user is receptive to it or to wait for an expected future receptive moment for a more effective intervention. In the long term, *I envision intervention systems that not only detect receptive moments but also determine what is the best intervention to deliver in-that-moment, and then make a decision on whether to deliver the intervention based on the immediate need and expected effectiveness, all while aiming to maximize the long-term effectiveness over the entire treatment duration*.

3.3. Leveraging multiple devices

Several surveys and reports have found that *people are using and interacting with more devices every day*. These devices range from smartphones, smartwatch and fitness wearables, tablets, PCs, smart home devices like Amazon Alexa or Phillips Hue, smart televisions, gaming consoles, and so on. There is an overlap in what the different devices can infer about a person, e.g., microphones on smartphones, wearables, and smart home devices can all detect conversations. Given the evolving nature of technology, relying on only one type of device for making inferences about a person might not be sufficient. *I envision a modular ecosystem of connected devices, all making overlapping inferences about a person. If a particular type of device gets outdated, it can be replaced with the next* *"trendy" device, providing similar data in conjunction with the other devices in the ecosystem*. The goal is that even if an actual device may no longer be used, the data and hence monitoring of health conditions is still available from other sources. The same holds for delivering interventions. Given the wide range of devices demanding and occupying user attention, delivering digital-health interventions using just the smartphone might not be the most optimal solution, both for receptivity towards the intervention or the intervention's effectiveness. To lay a foundation in this direction, I am currently planning a study where participants would receive interventions on multiple devices they might use, like smartphones, wearable devices, tablets, and laptops. The goal is to explore if and how the choice of the device used for interventions varies with context. I also intend to explore the overlap between the state-of-receptivity and the device choice, i.e., do there exist situations where the participant is in a state-of-receptivity on one device and not the other? *This work would lay the groundwork for future research on multi-device interventions.*

As I continue to work in the above core directions, I hope to explore and tackle important research problems and directions along the way, e.g., balancing individual and population-level data and ensuring privacy-sensitive deployments of these systems.

References

[1] **Varun Mishra**, Gunnar Pope, Sarah Lord, Stephanie Lewia, Byron Lowens, Kelly Caine, Sougata Sen, Ryan Halter, and David Kotz. "Continuous detection of physiological stress with commodity hardware," ACM Transactions on Computing for Healthcare (HEALTH), vol. 1, no. 2, 2020.

[2] **Varun Mishra**, Sougata Sen, Grace Chen, Tian Hao, Jeffery Rogers, Ching-Hua Chen, David Kotz. "Evaluating the Reproducibility of Physiological Stress Detection Models," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), vol. 4, no. 4, UbiComp 2021.

[3] Florian Künzler*, **Varun Mishra***, Jan-Niklas Kramer, David Kotz, Elgar Fleisch, and Tobias Kowatsch. "Exploring the state-of-receptivity for mhealth interventions," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), vol. 3, no. 4, UbiComp 2020. (***co-primary authors**)

[4] **Varun Mishra***, Florian Künzler*, Jan-Niklas Kramer, Elgar Fleisch, Tobias Kowatsch and David Kotz. "Detecting Receptivity for mHealth Interventions in the Natural Environment," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), vol. 5, no. 1, UbiComp 2021. (***co-primary authors**)

[5] Kevin Koch*, **Varun Mishra***, Shu Liu, Thomas Berger, Elgar Fleisch, David Kotz, Felix Wortmann. "When Do Drivers Interact with In-vehicle Well-being Interventions? An Exploratory Analysis of a Longitudinal Study on Public Roads," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. (IMWUT), vol. 5, no. 1, UbiComp 2021. (***co-primary authors**)

[6] **Varun Mishra**, Gunnar Pope, Sarah Lord, Stephanie Lewia, Byron Lowens, Kelly Caine, Sougata Sen, Ryan Halter, and David Kotz. "The Case for a Commodity Hardware Solution for Stress Detection," Workshop on Mental Health Sensing and Intervention, UbiComp 2018.

[7] **Varun Mishra**, Tian Hao, Si Sun, Kimberly N. Walter, Marion J. Ball, Ching-Hua Chen, and Xinxin Zhu. "Investigating the Role of Context in Perceived Stress Detection in the Wild," Workshop on Mental Health Sensing and Intervention, UbiComp 2018.