WatchOver: Using Apple Watches to Assess and Predict Substance Co-use in Young Adults

Sahiti Kunchay
College of Information Sciences and Technology
Pennsylvania State University
State College, USA
sahiti@psu.edu

Saeed Abdullah
College of Information Sciences and Technology
Pennsylvania State University
State College, USA
saeed@psu.edu

ABSTRACT
Simultaneous alcohol and marijuana (SAM) use can significantly impact young adults’ physical and mental well-being. While SAM use is becoming increasingly prevalent in this population, there has been much work to monitor and understand related behaviors and contexts. We aim to address this gap by using smartwatches to collect ecological momentary assessments (EMAs) and sensor data. In this paper, we describe the design and development of the smartwatch framework focusing on SAM use. We also collected pilot data from an n=1 deployment over 7 days using the framework. Our findings indicate that EMAs on smartwatches can be completed with lower perceived burden, which is important for longitudinal SAM use data collection. We also provide design guidelines and rationale for future work aiming to use smartwatches.

CCS CONCEPTS
• Human-centered computing → Empirical studies in ubiquitous and mobile computing. • Applied computing → Health informatics.

KEYWORDS
Mental Health, Smartwatches, Substance use

1 INTRODUCTION
Substance abuse is a mental illness that is characterized by an overindulgence or a dependence on substances to a degree that hinders physical, mental and social well-being [11]. In 2018, approximately 20.3 million people aged 12 or older had a substance use disorder (SUD) associated with their use of alcohol or illicit drugs in the past year [1]. These SUDs not only increase the burden of disease in the country, but also have great economic and personal costs.

Specifically, the simultaneous alcohol and marijuana (SAM) use disproportionately affects young adults – 22.5% of young adults aged 19 – 20 engaged in SAM use in the US [26]. Furthermore, the prevalence of SAM use has been increasing [26, 39]. It causes a myriad of negative consequences for adolescents’ and young adults’ mental and physical well-being [7, 23, 27]. SAM use, when compared to singular substance use (alcohol or marijuana use alone), has been associated with academic issues [7], symptoms of substance use disorders [23], and has also been associated with an increased likelihood of heavy alcohol consumption [27]. Given the high prevalence of SAM use and its associated risks, especially among young adults, it becomes increasingly important to understand the motivations and contextual factors that underlie this phenomenon. However, a significant portion of research within this domain has relied on infrequent surveys and cross-sectional studies. As such, there remains a serious gap in our understanding of SAM use in real-world contexts.

Recently, ubiquitous technologies have been making breakthroughs in terms of enabling efficient monitoring of illnesses and deploying novel intervention techniques to manage and treat these issues [2, 6, 29, 41]. There also has been recent work in understanding singular substance use with a specific focus on alcohol consumption [4, 5]. However, no such work has looked into SAM use, especially among young adults.

To address this gap, we propose the use of smartwatches to collect relevant behavioral, contextual and Ecological Momentary Assessment (EMA) data. Our rationale behind choosing smartwatches as data collection instruments stems from the novel user experience it offers: intuitive notifications, interactions without disrupting social contexts, acceptability of device among youth, and the inherently personal nature of the device.

The main contribution of this work are as follows:

(1) Establishing the design rationale behind the key elements of the smartwatch application.

(2) Development of the smartwatch and the analogous smartphone applications using EMAs and passive sensor data.

(3) An n=1 deployment aimed towards examining the usability of the smartwatch towards capturing substance use data. Further, we aim to compare both smartwatches and smartphones in this aspect.

The rest of the paper is organized as follows: Section 2 details the prior work conducted in the areas on SAM use, as well as the developments in ubiquitous technology that attempt to address tracking
and predicting substance use. Section 3 describes the design of the applications in detail, with Section 4 detailing the results of the n=1 deployment of the aforementioned application. Section 5 elaborates our plan for a longitudinal study in the future.

2 RELATED WORK

2.1 Young Adults and SAM Usage

It is pertinent to examine and establish the reasons as to why understanding the dynamics of SAM use is an important problem to study. Towards this effort, in this sub-section, we present various studies that, first, establish SAM use as a growing concern among adolescents and young adults in the US and second, demonstrate associations between SAM usage and various physical, psychological, and behavioral factors.

SAM usage has been steadily growing in prevalence among adolescents and college students - 22.5% of young adults aged 19 - 20 engaged in SAM use in the US, and has been increasing [26, 39]. A recent study of high school seniors in the US revealed that among students who used alcohol and/or marijuana in the past 12 months, 33.8% participated in SAM use. Subbaraman et al. also reported that the prevalence of SAM use is almost twice that of concurrent (non-overlapping) use [34]. Further, when compared to concurrent or only alcohol users, SAM users were more likely to participate risky behaviors such as truancy, use of other illicit drugs, drunk driving, violence and self-harm [19, 25, 34, 38]. SAM use has also been linked to an increased likelihood of heavy alcohol consumption [27].

Brière et al. reported that SAM use was associated with wide range of individual, family, and peer-related factors, including but not limited to alcohol intoxication, depressive symptoms, poor academic performance, drug use by peers, and household/parental complications [7]. Research by Lipperman-Kreda et al. suggests that location and social-contexts have a significant impact on alcohol, marijuana, and tobacco use and co-use [20, 37]. Further, the perceived percentage of intoxicated peers in proximity was associated with greater likelihood of substance use, while formal establishments (bars, restaurants) had significantly less alcohol and marijuana use. Among adolescents, greater adult supervision and absence of underage drinkers was associated with significantly lower risk of SAM use [19]. These studies show that there exists a need to, first, assess correlations of SAM use with mood, affect, location, and situational as well as social contexts and second, provide informed interventions based on the results of such studies – which is what we hope to achieve in the long term through our efforts.

2.2 Ubiquitous Technology and Substance Use

Ubiquitous technology in personal health is a recent but rapidly growing domain. However, the challenges and issues associated with supporting substance co-use and abuse (and related subsequent concerns) through such systems are quite unexplored. Existing research, however, can be classified into two broad categories: applications that seek to collect data, analyze and predict drinking episode through various modalities, and those that seek to support individuals manage problematic substance use behaviors.

To detect drinking episodes, smartphones have been the primary device of choice: several studies have investigated the efficacy of inferring drinking behaviors through a smartphone user’s gait [3, 15], with encouraging results - yielding 56% accuracy on the training set and 70% accuracy on the validation set - that merit further investigation. Another approach to understanding behavior post alcohol consumption from smartphone sensors is presented by Suffoletto et al., where they detect changes in gait during drinking episodes through ANNs [35]. Bae et al. assessed the feasibility of employing smartphone sensors and machine learning – using an application that continuously collects sensor data and self-reported alcohol consumption levels [4, 5]. This work reported time of day, movement, and device usage as the most informative features in predicting consumption, with the resulting model having a 96.6% accuracy at identifying non-drinking, drinking and heavy drinking episodes. Mariakakis et al. present yet another alternative to detecting an individual’s alcohol consumption: Drunk User Interfaces (DUIs), which use machine learning models to assess a person’s coordination and cognition through a set of tasks and estimate the individual’s blood alcohol level (BAL) [21].

In terms of gathering data on or detecting marijuana use episodes, the Substance Abuse Research Assistant App (SARA) is one of the few applications that utilize smartphones as a data collection instrument, rather than rely on cross-sectional or diary-based surveys [30, 31]. SARA collects information regarding the user’s daily feelings and activities, as well as weekly accounts of substance use frequency, while also focusing on identifying key strategies that will improve the users’ engagement with the application. However, there has been no prior work that has used passive sensor data to investigate and predict SAM use. Within this work, we hope to extend beyond the work done on the SARA application through the usage of both self reports through Ecological Momentary Assessments (EMAs) and passive sensor data obtained through smartphones and smartwatches.

In terms of assistive technologies, there is considerably more work - several applications focus on treating alcohol dependence by facilitating improved communication with family members [43], self-monitoring systems [13, 42], and peer-support based social tools [17]. Similar systems are available to those recovering from drug addictions that employ strategies such as developing coping skills through phone-based support systems [18], routine drug and addiction monitoring through self-administered tests and diaries [42, 43]. The success of these applications demonstrates the importance of having support mechanisms and strategies as an important part of the larger solutions targeting substance co-use and abuse. As with the previous category of systems, there is a pronounced lack of applications that support individuals experiencing or recovering from SAM use.

2.3 Smartwatches as Platforms for Behavioral Data Collection

The uptake of smartwatches by consumers has propelled research into how smartwatches can be used as instruments of behavioral health related studies. Hänsel et al. collected mood (emotion self-assessments), wrist movement, ambient noise, heart rate, workouts through Apple Watch, in an effort to conduct large-scale mood sensing through the Apple Watch [9, 10]. However, there was no analysis of the collected data, which would have been useful in
determining whether smartwatch users are inclined towards record-
ing affect through this device, and whether this resulted in higher
engagement when compared to smartphones.

There have been several efforts into investigating whether ill-
nesses and disorders could be recorded/managed through this de-
vice — Dibia et al. proposed FOQUS, a basic prototype deployed
on Samsung gear 2 that encouraged behavioral management tech-
niques for individuals with conditions like ADHD [8]. This work
utilized the unique capabilities of the smartwatch to provide feed-
back, using haptic feedback to notify start and end of meditation
sessions. Several other smartwatch-based efforts within this domain
include an application to track PTSD symptoms [16], a system to
aid students with intellectual and developmental disabilities (IDDs)
[44], and an evaluative study that explores the usage of smart-
watches to track and manage chronic disorders, such as tinnitus
[32]. Further, there has also been interesting work in terms of de-
tecting harmful behaviors such as smoking using these devices
— Skinner et al. use the accelerometer and gyroscope data in An-
droid LG-G-Watch to detect signature hand movements of cigarette
smoking [33].

The aforementioned work provide a brief overview of the ad-
vances of recent research in utilizing smartwatches to gauge bio-
behavioral features, however, there has also been interesting work
in terms of smartwatch usage that motivates the proposed applica-
tions [40]: Jeong et al. conducted usage studies among students
attending university full time and their work revealed that average
wearing hours for these devices are generally high within this pop-
ulation [14]. Further, users recalled that these devices were more
appropriate for use in social contexts as compared to smartphones.
McMillian et al. focused on analyzing smartwatch usage sessions,
and found that users are more likely to engage in shorter ‘peek’
sessions, which supports our design of utilizing micro-interactions
(simple data-capturing EMAs answerable in very short duration
of time, typically under 5 seconds) to assess factors such as mood
and alcohol consumption context [22].

The aforementioned work presents a strong case for designing ex-
periments that focus on analyzing whether smartwatches would
be a successful instrument in terms of capturing substance use
data, and in the future, to deploy interventions that seek to prevent
problematic substance use behaviors.

3 SMARTWATCH AND SMARTPHONE
APPLICATIONS
In this section, we describe the guiding principles behind the design
of the smartwatch and smartphone applications, as well as the
different modules that they comprise of.

3.1 Baseline Application
In order to effectively compare users’ opinions and experiences, it
is important to have a baseline application that is representative of
the designer’s goals and anticipated user requirements. For the pur-
poses of this work, the baseline application (with common elements
across devices) is designed as follows: an application available on
both the smartphones and smartwatches that collects the following
information through regular EMAs: (a) Mood and general affect, (b)
Substances used since last EMA, (c) Amount consumed since last
EMA, (d) Feelings of intoxication/inebriation, (e) Context: location,
social environment. Additionally, the application also collects pas-
sive sensor data that includes accelerometer and gyroscope data,
location from GPS sensor stream, physical activity and health data
streams that includes heart-rate, sleep duration, type of exercise(s),
and active minutes.

3.2 Application Design
Our fundamental rationale for choosing smartwatches as a data col-
lection instrument stems from the unique user experience it offers:
subtle/haptic notifications, allowing quick/glanceable interactions
without violating social norms, acceptability of device among youth,
and the overall private nature of the device. We posit that these
features make it ideal for capturing data that is inherently personal
and private.

Our primary aim in creating both applications was to provide
a user interface that is conducive to high response rates, engage-
ment, and compliance rates. We wanted to provide an experience
that incorporated all the above to get granular data, while keep-
ing the perceived burden of interaction low. Hence, we chose
micro-interactions as our primary method of data collection. Micro-
interactions focus on capturing simple data in very short duration
of time, typically under 5 seconds. Micro-interactions have been
demonstrated to reduce perceptions of user burden when compared
to regular interactions, even with higher rates of interruptibility
[12]. Micro-EMAs (μEMA) have been shown to result in higher
compliance, completion, and first-prompt response rates in smart-
phones [12]. Prior literature also reports that watch μEMA results
in higher response rates than just regular watch EMA [28].

To extend findings from prior work, we aim to understand whether
watch μEMAs result in better compliance, completion, and first-
prompt response rates than phone μEMAs in the context of SAM
use. We hypothesize that it would, given the accessibility, glance-
ability, private nature, and the highly personal experience provided
by the smartwatch. If the results support this hypothesis, we can
leverage the smartwatch as a novel health behavior data collection
tool within the domain of substance use.

In terms of the sensor data, our primary sensor streams com-
prise of: GPS, data from the AWARE application1, and data from
the HealthKit application. The literature highlighted in Subsection
2.2 points out the importance of sensor data to infer alcohol con-
sumption. As such, we decided to use the AWARE framework and
application to capture sensor data. Furthermore, the literature high-
lighted in Subsection 2.1 details how social and situational contexts,
along with location are strongly associated with SAM usage. Follow-
ing these findings, we focused on capturing GPS and contextually
relevant location information. In terms of the GPS, we collect the
data with an accuracy of up to 100m to preserve the privacy of the
participant. We also ask about location contexts in the EMA (e.g.,
“Home”, “Friend’s House”, “Bar/Restaurant”). The rationale behind
collecting the HealthKit data stems from the substantial body of
literature associating sleep and exercise with substance use among
young adults [24, 36, 37].

---

1 AWARE Framework. https://awareframework.com/
3.3 Data Modules

3.3.1 Daily μEMA Module. Our framework deploys notifications at 11 AM, 4 PM, 7 PM, 10 PM, and 1 AM everyday. The 11 AM EMA (once daily EMA) asks the user what has happened since last night: the substances they consumed, amount of substances, how they consumed the substances, their mood and behaviors during and immediately after consumption, feelings of stress and affect. It also asks questions relating to substance use context and behaviors: where they were, who they were with, how intoxicated the people around them were, whether they felt safe with the people they were with, and whether they felt a sense of belonging with the people they were with. The once daily EMA expires at 3 PM to avoid overlapping with subsequent EMAs.

The 4 PM, 7 PM, 10 PM and 1 AM EMAs (periodic EMAs) seek to collect data of what happened since the last time the user answered an EMA. These EMAs cover similar topics to those covered in once daily EMAs. The periodic EMAs expire 2 hours after the users are notified. The crucial difference between the once daily EMA and the periodic EMAs is that we do not record their behaviors during and immediately after consumption in the periodic EMAs. This is done to keep the periodic EMAs short and succinct.

3.3.2 HealthKit Data Module. This module comprises of the following data collected from Apple’s HealthKit:

- total number of steps walked per day
- total number of stand hours and stand minutes
- total number of exercise minutes
- walking + running distance
- heartbeat at regular intervals
- number and type of workouts done in a day
- active and resting calories burned in a day
- number of hour slept, sleep time, and wake time

3.3.3 Sensor + Location Module. We collect the following data using the AWARE Framework’s iOS client application:

- Accelerometer, Magnetometer, and Gyroscope
- Application usage data
- Ambient noise
- Battery levels
- Communication (Calls, Messages)
- Device usage
- Location
- Network description
- Screen sensor data
- Weather

3.4 Data Storage

The data from the Daily EMAs is stored using Google’s Cloud Firestore, using the Google Firebase SDK. In the smartwatch application, the data is first transferred to the iPhone through the WatchConnectivity Module. The data in then transferred from the iPhone to Cloud Firestore. In the smartphone application, the data is directly uploaded to Cloud Firestore. EMA data uploading happens right after a user completes it. In case of network failure, these records are cached in the iPhone till network connectivity is restored.

Sensor data (see 3.3.3) is stored in an SQL instance on Google Cloud. This data is uploaded every hour from the iPhone on to the remote server.

4 PRELIMINARY RESULTS

The applications were deployed for use by a single participant. The participant was a graduate student in the School of Engineering, Pennsylvania State University. The applications were deployed on an iPhone 11 with iOS 13.5.1, and Apple Watch Series 3 with watchOS 6.2.6. The deployment period was for a week, with the participant using both the smartwatch and the smartphone applications. The participant was encouraged to use the applications in any manner of their choosing, to identify any usability issues or flaws in the system that might come up with unrestricted use.

This work has been approved by Pennsylvania State University’s Institutional Review Board.

The deployment was followed by a semi-structured interview that covered a variety of themes including overall experiences using the smartwatch application, perceived time burden, usability and interruptability of the notifications, barriers to completing the survey, and privacy considerations.

In terms of the overall experience, the participant reported that “It was not difficult to complete the questions, but some of the questions were a little vague”. The participant reported that the question
depicted in Figure 3 would be difficult to answer, suggesting instead that the app ask about the number of intoxicated people instead of the percentage.

The perceived burden of the application in terms of time was low. When the participant was asked whether they felt if the surveys were time consuming, they answered, "Not at all!" However, they did note that the frequency of the notifications were high, indicating that while the survey might be easy and quick to complete, having to repeat the process 4 times a day might be burdensome for a user. The participant also specifically noted that "It was notifying me more on the phone than on the watch", even though both the watch and the phone notify the user the same number on times. This might indicate that the notifications on the phone are perceived to be more interruptive than notifications on the watch. While we do want to ensure data quality and garner responses of high granularity, we also do not want to disrupt the social contexts under which people usually drink and consume substances. While further investigation is warranted, we believe that this theme highlights the subtlety of the smartwatch notifications. We plan to conduct a larger study (as outlined in Section 5), to compare the effectiveness of smartwatch and smartphone notifications in eliciting responses.

When asked about the relative ease of completing the surveys across devices, the participant responded that it was easier on the phone, 'I think, for me that the phone was easier, maybe I'm just more used to the phone'. This highlights an important theme that we seek to explore in a future study: device preference. Device preference can drastically impact the quality of the data we collect. If we pair a smartphone with a person who prefers smartwatches to capture such data, the resulting dataset might not be as rich. Future studies seeking to both collect data and deploy interventions should take device preference into account.

In terms of what the participant would change about the study, they noted that more questions about indicators of depression might be useful. "I really like that the questions talk about your emotional state and ask you to give a range of what you’re feeling right now, there was a question about how stressful you were, and I think there’s definitely a link to this, but maybe you could ask more about depression indicators, like what made you sad, or feelings of misery". Currently, we have the following items in the EMA that we are using as indicators of affect that might predict substance use: Tension, Stress, Wellness, Feelings of Safety and Belonging, Calmness, Energy Levels, and Contentment.

Finally, We note that the perceived burden of the application might be low because the deployment window was only 7 days. We do realize that perceived burden and engagement might vary within a larger, more diverse set of participants if the study were to be run for a longer duration. We understand that an n=1 study does not lead to conclusive results, however, we find that important themes were discovered which would help shape the discussion around improving the applications and designing future studies. Taking the findings into account, we have elaborated our design for a future study in the following section.

5 FUTURE STUDY DESIGN

Our future pilot study has 2x2 design i.e., smartwatches vs smartphones x 2 weeks of data capture vs. 4 weeks of data capture. Each condition will have approximately 20 participants, with a total of 80 participants. Eligibility is based on whether the person participated in binge drinking (4 or more alcoholic drinks [if woman] or 5 or more alcoholic drinks [if man]) in the past 1-2 weeks. This specific study design was chosen to serve several goals. First, we want to understand whether varying the duration of data capture window (2 weeks vs. 4 weeks) would have an impact on the model thus generated; secondly, we want to understand if interface type (smartwatches vs smartphones) has any effect on the measures of engagement and compliance. Third, we want to be able to collect enough behavioral and contextual data to train and evaluate our predictive models.

Through this workshop, we hope to elicit feedback on our current system, as well as our future study design in order to significantly improve the applications’ usability while meeting the study goals of understanding the contexts under which individuals over-indulge in SAM use.

REFERENCES


on Ubicomp/ISWC ’20 Adjunct, September 12–16, 2020, Virtual Event, Mexico
Kunchay and Abdullah, et al.


