

Original Paper

# Intensive Longitudinal Data Collection Using Microinteraction Ecological Momentary Assessment: Pilot and Preliminary Results

Aditya Ponnada<sup>1,2</sup>, BDes, PhD; Shirlene Wang<sup>3</sup>, BA; Daniel Chu<sup>3</sup>, PhD; Bridgette Do<sup>3</sup>, MPH; Genevieve Dunton<sup>3</sup>, MPH, PhD; Stephen Intille<sup>1,2</sup>, PhD

<sup>1</sup>Khoury College of Computer Sciences, Northeastern University, Boston, MA, United States

<sup>2</sup>Bouve College of Health Sciences, Northeastern University, Boston, MA, United States

<sup>3</sup>Department of Population and Public Health Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

**Corresponding Author:**

Aditya Ponnada, BDes, PhD  
Khoury College of Computer Sciences  
Northeastern University  
360 Huntington Avenue  
Boston, MA, 02130  
United States  
Phone: 1 6173061610  
Email: [ponnada.a@northeastern.edu](mailto:ponnada.a@northeastern.edu)

## Abstract

**Background:** Ecological momentary assessment (EMA) uses mobile technology to enable in situ self-report data collection on behaviors and states. In a typical EMA study, participants are prompted several times a day to answer sets of multiple-choice questions. Although the repeated nature of EMA reduces recall bias, it may induce participation burden. There is a need to explore complementary approaches to collecting in situ self-report data that are less burdensome yet provide comprehensive information on an individual's behaviors and states. A new approach, microinteraction EMA ( $\mu$ EMA), restricts EMA items to *single, cognitively simple questions* answered on a smartwatch with single-tap assessments using a quick, glanceable microinteraction. However, the viability of using  $\mu$ EMA to capture behaviors and states in a large-scale longitudinal study has not yet been demonstrated.

**Objective:** This paper describes the  $\mu$ EMA protocol currently used in the Temporal Influences on Movement & Exercise (TIME) Study conducted with young adults, the interface of the  $\mu$ EMA app used to gather self-report responses on a smartwatch, qualitative feedback from participants after a pilot study of the  $\mu$ EMA app, changes made to the main TIME Study  $\mu$ EMA protocol and app based on the pilot feedback, and preliminary  $\mu$ EMA results from a subset of active participants in the TIME Study.

**Methods:** The TIME Study involves data collection on behaviors and states from 246 individuals; measurements include passive sensing from a smartwatch and smartphone and intensive smartphone-based hourly EMA, with 4-day EMA bursts every 2 weeks. Every day, participants also answer a nightly EMA survey. On non-EMA burst days, participants answer  $\mu$ EMA questions on the smartwatch, assessing momentary states such as physical activity, sedentary behavior, and affect. At the end of the study, participants describe their experience with EMA and  $\mu$ EMA in a semistructured interview. A pilot study was used to test and refine the  $\mu$ EMA protocol before the main study.

**Results:** Changes made to the  $\mu$ EMA study protocol based on pilot feedback included adjusting the *single-question* selection method and smartwatch vibrotactile prompting. We also added sensor-triggered questions for physical activity and sedentary behavior. As of June 2021, a total of 81 participants had completed at least 6 months of data collection in the main study. For 662,397  $\mu$ EMA questions delivered, the compliance rate was 67.6% (SD 24.4%) and the completion rate was 79% (SD 22.2%).

**Conclusions:** The TIME Study provides opportunities to explore a novel approach for collecting *temporally dense* intensive longitudinal self-report data in a sustainable manner. Data suggest that  $\mu$ EMA may be valuable for understanding behaviors and states at the individual level, thus possibly supporting future longitudinal interventions that require within-day, temporally dense self-report data as people go about their lives.

(JMIR Form Res 2022;6(2):e32772) doi: [10.2196/32772](https://doi.org/10.2196/32772)

## KEYWORDS

intensive longitudinal data; ecological momentary assessment; experience sampling; microinteractions; smartwatch; health behavior research; mobile phone

## Introduction

### Background

Mobile technologies create new opportunities to design personalized health interventions in 2 broad ways. First, mobile technologies such as smartphones and smartwatches can be used to improve the assessment of behavioral, psychological, and contextual states, reducing reliance on known-to-be-problematic retrospective recall surveys that may poorly capture within-day variations [1-3]. Second, real-time measures gathered using mobile technologies can be used to design just-in-time adaptive interventions [4,5] that are sensitive to everyday changes in behaviors and states. State or behavior assessment and personalized interventions both require computational models that can represent the interrelationships between different behaviors and states unique to individuals. Such models could be created by using (1) a hypothetico-deductive approach to study relationships between predictors and outcomes of interest based on prior knowledge [6,7], (2) data-driven discovery from a large number of constructs measured in intensive longitudinal data (ILD) studies [8,9], or (3) an appropriate combination of both [10,11]. Creating models of behavior that capture relationships between behavior and context that can happen quickly and many times in a day requires methods for sustainable data gathering on, and modeling of, within-day changes of variables (eg, physical activity [PA], sleep, sedentary behavior [SB], and affect) [12,13].

Ideally, health-related constructs could be measured using passive sensors from easy-to-wear devices, without requiring any end-user effort other than wearing the devices and charging them. However, *we currently need self-report input* to measure many subjective experiences (eg, perceived pain, hunger, and fatigue) that play an important role in predicting health behavior changes [14]. A popular practice for obtaining self-report data on mobile devices is using experience sampling methods [15], also known as ecological momentary assessment (EMA) [16], to gather participant perspectives on in situ experiences that sensors cannot measure directly. As smartphones have become ubiquitous, EMA has become affordable to deploy with most adult populations [17].

In the most common EMA protocols used to collect ILD, participants' smartphones prompt with audio or vibration several times a day, each time presenting a set of questions that may take a few seconds or minutes to answer, depending upon the number and complexity of the questions. Typically, questions are multiple choice, but sometimes they involve open-ended responses or sliding scales. The sampling frequency in EMA varies widely based on study goals, as does potential burden. Burden results from the annoyance of each EMA interruption and the time and mental effort required to answer the questions presented at the prompt [18,19]. To minimize response burden, researchers often reduce the number of questions in each question set, the complexity of the questions, the frequency of

prompting, or all three, thereby limiting the scope and temporal density of the behaviors and states measured in an ILD study. Although intensive EMA protocols can be sustained for short-term studies (eg, 7-30 days), burden will accumulate in ILD collection studies that last several months or years [20]. This accumulated burden, if not managed, could prevent the conducting of large-scale longitudinal studies or interventions that rely on frequent self-report input.

Although compliance-contingent compensation is a common practice to boost or maintain EMA response rates, it can be expensive for large-scale longitudinal studies involving thousands of participants. Beyond observational studies, future personalized health interventions that may require self-report of participant states and contexts must balance information gathering with burden to ensure that long-term engagement does not wane. For instance, an intervention to reduce stress and increase PA may need to query users regularly about their perceived stress and physical exertion (among other subjective experiences) to adjust the intervention in real time. For such an intervention to be sustainable, the system may have to gather these data frequently at the rate at which stress and activity can change—that is, many times throughout the day—while keeping end-user burden manageable. For such an intervention to be affordable, financially compensating participants to motivate a reasonable response rate is not a viable long-term option. In fact, compliance-contingent compensation may result in poor data quality if the participants are only motivated by money [21]. Thus, new strategies are needed for acquiring temporally dense self-report on multiple constructs with sustainable burden, even for studies or interventions deployed for months or years.

A strategy for acquiring longitudinal self-report data is using microinteraction EMA ( $\mu$ EMA or micro-EMA) [22]. *Microinteractions* are actions lasting for 3-4 seconds (eg, checking the time on a watch or turning on a lamp); microinteractions are so short lived that they can be completed without disrupting ongoing activity [23].  $\mu$ EMA is a type of EMA where all prompts are for *single questions* with *single-tap answers* (eg, *Nervous right now?* with answers *Yes, Sort of, or No*) [22]. Unlike EMA, where questions are prompted in a set, in  $\mu$ EMA a device prompts only a single question per interruption and there is a guarantee that the question can be answered with a microinteraction [23]. Because of this microinteraction guarantee,  $\mu$ EMA may enable self-report data collection at a much higher frequency than EMA. This at-a-glance response property is achieved in 2 ways. First,  $\mu$ EMA is deployed on wearable devices such as smartwatches that permit quick access to question content; unlike mobile phones, the questions can be seen with a flip of the wrist without additional time required to access a mobile phone that may be in a bag or out of sight [24,25]. Prior studies have demonstrated that smartwatches are more suitable for glanceable microinteractions (eg, checking notifications) than smartphones [26,27]. Second,  $\mu$ EMA questions are intended to be cognitively simple to answer (eg, *Feeling stressed?—Yes, Sort of, or No*)

and with a limited answer set that fits on a small smartwatch screen and therefore does not require a scrolling interface. Prior work shows that reducing the number of answers (eg, from a 5-point Likert scale to a 3-point ordinal or nominal scale) improves response time without necessarily changing the perceived item difficulty [28]. The benefit of leveraging microinteractions is also supported by 2 laws in user interface (UI) design. First, Hick's law posits that the more options there are to choose from, the longer the response time on the UI will be [29,30]. Second, Fitt's law shows that the navigation time between 2 targets on a UI is directly proportional to the distance between the targets and inversely proportional to the size of the target [31,32]. However, the downside of restricting all prompted interactions to cognitively simple, glanceable, single-question microinteractions is that information obtained from a single  $\mu$ EMA question is more limited than what can be obtained from unconstrained EMA questions with multiple answer options (eg, *In the past one hour, how stressed did you feel?—Extremely, Quite a bit, Moderately, A little, or Not at all*). Presenting such EMA questions on a smartwatch screen would make smartwatch interaction cumbersome, requiring either a font size that would be too difficult for most people to read, especially for those who need reading glasses, or requiring scrolling. Furthermore, using self-report to capture feelings aggregated over longer time windows (eg, in the past hour or over the past day) introduces recall burden [18,33] and cognitive complexity. Therefore, EMA questions nearly always require adjustment to achieve the microinteraction property we seek.

Prior work has used wearable devices such as smartwatches and heads-up displays (eg, Google Glass) to deploy EMA question sets with the goal of making EMA easier. However, typically, such work has directly adopted surveys (with back-to-back questions) from mobile phone-based EMA surveys that require users to engage in prolonged interactions on the devices (eg, sliding or scrolling) [34-37] and thus do not result in microinteractions. Converting an EMA survey with multiple questions with multiple answers per question to a small smartwatch display has indeed required additional scrolling to make the questions and answers readable, thereby likely slowing down—not speeding up—answer selection and resulting in the smartwatch interaction becoming more burdensome than answering the same surveys on the mobile phone.

In 2 prior 4-week between-subject pilot studies, we found that despite experiencing approximately 4-6 times more interruption when using  $\mu$ EMA instead of mobile phone-based EMA, participants using  $\mu$ EMA on a smartwatch had significantly higher compliance, reported lower perceived burden, and

answered with faster response times than participants when measuring the same constructs using EMA [22,38]. Prior work also demonstrates that for a single construct measured at high frequency,  $\mu$ EMA can yield good criterion validity when compared against a research-grade passive sensor (eg, in the domain of PA measurement) [39]. Table 1 summarizes key differences between the mobile phone-based EMA and smartwatch-based  $\mu$ EMA in a prior study [22] where both EMA and  $\mu$ EMA were used to measure positive and negative affect (using the Positive and Negative Affect Schedule [40] over a period of 4 weeks [41]).

With  $\mu$ EMA, a user is guaranteed that the worst-case interaction required to answer a prompt will always be a single, glance-and-tap microinteraction; thus, it is nearly as easy to answer a microinteraction prompt as it is to manually dismiss it (by swiping on the screen) or ignore it. If researchers develop simple questions appropriate for microinteractions that provide valuable information about constructs of interest,  $\mu$ EMA may support ILD studies that gather dense, within-day information on behavior. So far, researchers have used  $\mu$ EMA to measure only 1-2 constructs per study, such as hyperarousal [42], stress [43], and subjective comfort [44];  $\mu$ EMA has also been used with situated displays in home settings [45,46].  $\mu$ EMA is well suited for constructs that may require frequent self-report, such as chronic pain assessment [47].  $\mu$ EMA can be assessed at such a high temporal density (ie, 4 times per hour) that it can also measure multiple constructs in a single day while maintaining reasonable temporal density of each construct.

Despite the promise of  $\mu$ EMA, much remains to be explored to determine the viability of the technique. In this paper, we present the  $\mu$ EMA protocol for the Temporal Influences on Movement & Exercise (TIME) Study. The TIME Study has a primary protocol that uses EMA, which is presented elsewhere (Wang, S, unpublished data, January 2022). Here, we present the protocol for a secondary, exploratory study within the TIME Study on the viability of using  $\mu$ EMA in an ILD study. The TIME  $\mu$ EMA substudy is the first large-scale ILD study (duration: 1 year) examining how  $\mu$ EMA might be used to measure multiple health behaviors such as PA, SB, and sleep along with time-varying subjective states (eg, stress, fatigue, and happiness). First, we describe the overall goal of the TIME  $\mu$ EMA substudy. Next, we detail the  $\mu$ EMA app we designed for the study, followed by reporting the qualitative feedback received from our pilot study participants. Subsequently, we present details of changes made to the protocol in response to pilot testing. Finally, we present preliminary compliance results from a subset of participants in the main study (81/246, 32.9%).

**Table 1.** Differences between mobile phone-based ecological momentary assessment (EMA) and smartwatch-based micro-EMA ( $\mu$ EMA; from prior work [22]).

	Mobile phone-based EMA	Smartwatch-based $\mu$ EMA
Prompts per day	$\leq 7$	$\leq 48$
Prompting frequency	Once in 2 hours	Four times an hour
Number of questions per prompt	$\leq 6$	Only 1 question
Example of question framing	Over the past hour, how stressed did you feel?	Feeling stressed?
Number of response options per question	$\geq 4$ (including <i>choose all that apply</i> responses)	$\leq 3$ (eg, <i>Yes, Sort of, and No</i> )
Interruption burden	High (must access mobile phone, unlock it, and then start to answer)	Low (smartwatch always accessible with glance)
Response burden	High, because of multiple questions with more answer options	Low, because of only 1 question with limited answer options; aim is cognitive simplicity

### TIME $\mu$ EMA Substudy Objectives

The goal of the TIME Study is to examine daily and within-day microtemporal processes (eg, feeling stressed, increased workload, being with family, and being at home) that may influence PA, SB, and sleep in young adults [48]. EMA is being deployed to capture reflective processes, reactive processes, internal factors, and external factors and assess how they affect health behavior adoption and maintenance (Wang, S, unpublished data, January 2022). Reflective processes are those that are slow and require careful deliberation (eg, intention to engage in healthy behaviors) [49]. In contrast, reactive processes are quick and automatic (eg, being on a regular routine) [49]. Internal factors are physiological and emotional sensations that originate internally (eg, positive and negative affect, pain, and fatigue). External factors are social, situational, and physical settings or events that originate externally to the individual (eg, an increase in workload or meeting a friend). An exploratory aim of the TIME Study is to assess the feasibility of using  $\mu$ EMA to gather similar health behavior data from participants. The fundamental difference between using EMA and  $\mu$ EMA for ILD can be summarized as follows. Traditional EMA is a method that interrupts less often than  $\mu$ EMA but asks for substantially more information with each interruption. In contrast,  $\mu$ EMA is a method that interrupts more often than EMA but asks only 1 simple question with each interruption. The TIME  $\mu$ EMA substudy will permit exploration of what we can learn about health behavior from small amounts of information gathered at high frequency using  $\mu$ EMA; the TIME substudy data will allow investigators to study the following research questions:

1. How sustainable is  $\mu$ EMA for ILD studies compared with mobile phone-based EMA?  
Interrupting participants 4 times more often with a single, simple  $\mu$ EMA question could be perceived as far more burdensome than interrupting less often with a longer, more complex EMA survey. Can participants in a full-year ILD collection study sustain  $\mu$ EMA with high compliance compared with EMA?
2. How do contextual factors influence  $\mu$ EMA compliance? Prior EMA literature has shown that contextual factors such as time of day, day of week, and location influence whether the participant is able to complete the EMA survey [50]. Behavior, such as being active, could also influence

response rates. The TIME  $\mu$ EMA substudy will permit exploration of the contextual factors that may influence  $\mu$ EMA compliance and comparison of such effects between  $\mu$ EMA and EMA.

3. Can intermittent  $\mu$ EMA questions provide information on an individual's overall behavior and state?  
Sustaining intensive mobile phone-based EMA longitudinally may not be realistic, requiring long temporal gaps between measurements and more retrospective recall. The TIME  $\mu$ EMA substudy will permit exploration of whether using  $\mu$ EMA *between* EMA burst periods could provide information on the diurnal patterns of behaviors and states not captured in the EMA bursts alone.
4. How can we use  $\mu$ EMA data *together with* data acquired from the passive sensors and EMA to design predictive models of health behavior change and maintenance?  
Each instance of  $\mu$ EMA self-report provides information on a single variable at a particular time, but  $\mu$ EMA prompts will be temporally dense, measuring different behaviors and states throughout the day. The TIME  $\mu$ EMA substudy will permit exploration of whether the amount and quality of  $\mu$ EMA data acquired in the TIME  $\mu$ EMA substudy is sufficient to build ideographic predictive models of behavior, either alone or in combination with EMA and passive sensing data.

In the remainder of the paper, we describe the protocol for the TIME  $\mu$ EMA substudy, from which ILD will be derived that will support such work.

### $\mu$ EMA Design Overview

Deploying  $\mu$ EMA requires designing a specific smartwatch-based graphical UI and the question scheduling strategy.

### Custom $\mu$ EMA App Interface

We developed a custom  $\mu$ EMA app (called the TIME app or *app*) that runs on a smartwatch (Android Wear OS version 2.0) with a linked app on the paired smartphone (Android OS version  $\geq 7.0$ ). The Android OS was chosen because it provides flexibility in gathering raw sensor data from both smartphone and smartwatch continuously in the background. The apps work together to present  $\mu$ EMA questions on the smartwatch, collect and process passive sensor data, and transfer data to a research server each day. Each  $\mu$ EMA prompt presents only 1 question

at a time that can be answered in a glance and a tap (a microinteraction). Example questions are *Feeling stressed?* and *Feeling productive today?* with three answer options: *Yes*, *Sort of*, and *No*. The smartwatch prompts the participant with a vibration pattern lasting for 3 seconds. The question screen displays at the start of the vibration (Figure 1). The brightness of the screen is determined automatically by the smartwatch based on ambient brightness or the smartwatch settings selected by the participant.

The question stays visible on the screen for 20 seconds, after which the question disappears and a missed response is recorded. If participants answer a question, they are presented with an acknowledgment screen with a short thank-you message and an *Undo?* button (Figure 1). The thank-you messages on the

acknowledgment screen are selected from >250 unique variations designed to reduce repetition and boost engagement—they are all pithy, and some are quirky. The undo option is available for 3 seconds. If participants tap on the undo button, they are taken back to the question screen and they can then change their answer within 20 seconds. If they do not undo within 3 seconds, the app records the final answer and the question disappears. Participants have only 1 chance to undo an answer. If the app is dismissed (by swiping right on the screen) when the original question appears, it is recorded as *never started*, and if the app is dismissed on the undo screen, it is recorded as *completed then dismissed*. Similarly, if the user does not answer the question again after selecting the undo button, then the prompt status is recorded as *partially completed*.

**Figure 1.** (Left) A microinteraction ecological momentary assessment question screen with 3 answer options, which will display for 20 seconds or until an answer is selected. (Right) A thank-you screen with the Undo? button indicating that the user selected the Sort of answer; the screen displays for 3 seconds with a countdown timer before it disappears.



### μEMA Prompt Scheduling

μEMA prompts occur during waking hours, including when the smartwatch is being charged or is off the wrist. Participants self-report their upcoming sleep time and the next day's wake time using EMA on their smartphone (Figure 2); this information is used to compute the waking hours and is sent to the smartwatch in real time. If participants do not respond to the sleep- and wake-time questions on the smartphone or if the smartphone cannot connect to the smartwatch for any reason, the smartwatch uses the previous day's wake and sleep times as the default waking schedule to determine the μEMA prompting schedule. The purpose of the sleep-wake questions is to predetermine the prompting schedule for the day. Once answered, these sleep- and wake-time questions are presented again after 10 hours, in case participants want to change their schedule. Moreover, participants can use the app to manually change the sleep-wake times at any time of the day. μEMA prompting starts 15 minutes after each day's wake time and ends 15 minutes before that day's sleep time. For instance, if the participant plans to sleep at 11 PM and wake up the next day at 6 AM, then μEMA prompting will stop at 10:45 PM and resume at 6:15 AM the next day. Thus, days with longer waking hours will result in more μEMA prompts than days with shorter waking hours. μEMA questions are prompted 4 times an hour at random, with at least eight minutes guaranteed between 2 consecutive prompts using the following formula. The app generates this schedule for the 24-hour period using the following equations and then only prompts during waking hours.

$$P_n = \mathbb{R} \in [0, \text{MaxTimeAvailable}_n) + 8 + P_{n-1}$$

$$\text{MaxTimeAvailable}_n = 55 - ((4 - n + 1) \times 8) - P_{n-1}$$

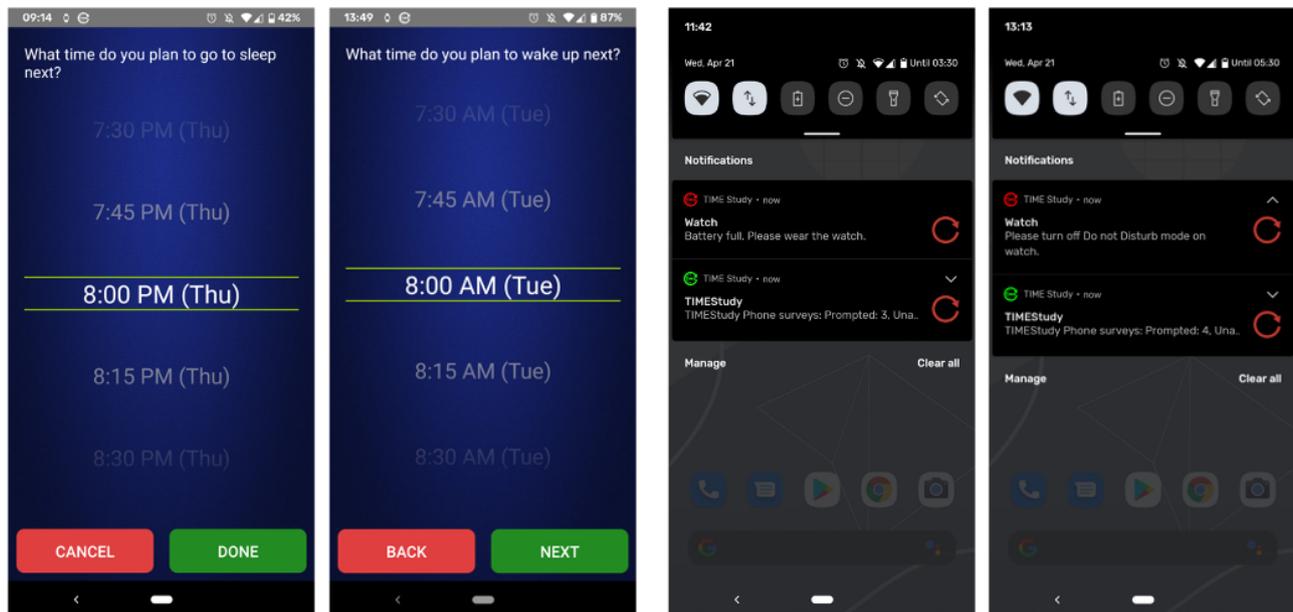
First, for each hour (eg, 8 AM to 9 AM), the app computes the maximum remaining time ( $\text{MaxTimeAvailable}_n$ ) in minutes to schedule a prompt ( $P_n$ , where  $n=1, \dots, 4$ ,  $P_0=0$ ) by subtracting the minimum required time gap (ie, 8-minute gap per prompt) and the time elapsed since the previous prompt within that hour ( $P_{n-1}$ ). Then, the app selects a random time from the  $\text{MaxTimeAvailable}_n$  and adds the 8-minute gap and the prompt time of the previous prompt ( $P_{n-1}$ ) to determine the current prompt time ( $P_n$ ). The μEMA smartwatch app respects the settings on the smartwatch; therefore, prompting is paused when a participant turns on the do-not-disturb (DND) mode on the smartwatch. However, if the smartwatch is in DND mode for >60 consecutive minutes, participants receive a notification on the mobile phone to disable the DND mode on the smartwatch so that prompting can be resumed without further data loss (Figure 2). If the smartwatch is off the wrist, including when it is being charged, μEMA prompting continues normally.

Participants receive a notification on the mobile phone and smartwatch to (1) connect the smartwatch if it is disconnected from the mobile phone (through Bluetooth) for >60 consecutive minutes, (2) wear the smartwatch if the system detects that the smartwatch is fully charged but on the charger, (3) charge the smartwatch when the smartwatch battery reaches ≤15% capacity (Figure 2), and (4) update the software when the mobile phone or smartwatch software requires an update. Mobile phone and

smartwatch notifications disappear as soon as the problem that was flagged is resolved. In addition, the mobile phone app presents a persistent notification indicating that the TIME app

is running in the background collecting data and highlighting the countdown of when the next burst period on the mobile phone will start.

**Figure 2.** (Left) A mobile phone survey asking for prospective sleep and wake times. These times are used to adjust the prompt scheduling. (Right) Example notifications: (1) a reminder to wear the smartwatch prompted when it is still on the charger and is 100% charged and (2) a reminder to turn off the do-not-disturb mode on the smartwatch, which appears if that mode is active for >60 minutes consecutively.



## System Components

The system has three components: mobile phone, smartwatch, and a remote server. The mobile phone prompts users with EMA surveys and end-of-day surveys. In addition, the mobile phone passively collects sensor data such as acceleration, location, and mobile phone use (eg, number of apps used). The details of mobile phone-related functionality are beyond the scope of this paper and are reported elsewhere (Wang, S, unpublished data, January 2022). The smartwatch prompts  $\mu$ EMA questions and passively collects raw accelerometer data (described in the next section). The data from the smartwatch are sent to the mobile phone either (1) once every 4 hours or (2) whenever the smartwatch is on the charger. The mobile phone then encrypts the data and sends it to a remote server daily. When the smartwatch or mobile phone are not transferring data automatically as expected, they each provide an option that permits participants to force the data transfer to the mobile phone; this functionality is used to resolve data transfer problems working remotely with the research staff.

## Methods

### Pilot Study

With the  $\mu$ EMA system described in the previous sections, we conducted a pilot study with 15 participants who answered  $\mu$ EMA questions for up to 3 consecutive weeks.

### Pilot Study Participants

Participants for the pilot study were recruited by means of flyers posted on university campuses in the Los Angeles metropolitan area and through targeted ads on Facebook. Participants were eligible to participate if they (1) were aged 18-24 years, (2)

owned an Android smartphone with OS version  $\geq 6.0$ , (3) were fluent in English, (4) were not planning to change their smartphone in the next 1 month, (4) were willing and able to wear a smartwatch and answer questions on their smartphone for a period of 1 month, and (5) were currently engaged in recommended levels of PA or intending to do so in the next 12 months.

### Pilot Study Design

The research staff met with the eligible participants in person (in 2019) and installed the TIME app on their smartphones. The app prompted participants with EMA surveys in 2 measurement burst periods. Each burst period lasted 4 days (called burst days). If participants completed at least eight EMA surveys per day in the first burst period, they were loaned a Fossil Gen 4 smartwatch (Fossil Group, Inc) and instructed on how to install the TIME smartwatch app. For the remaining 3 weeks, participants also collected data on the smartwatch. On the nonburst days, the smartwatch prompted with  $\mu$ EMA questions about a myriad of behaviors, including positive and negative affect, location, and PA. With each prompt, the smartwatch TIME app selected a question at random and presented it on the screen. In addition, questions about momentary PA, SB, and location (ie, *Physically active now?*, *Sedentary right now?*, and *At home right now?* with answer options *Yes*, *Sortof*, and *No*) were also selected randomly throughout the day. Thus, it was possible for a question to repeat consecutively when sampled at random. The goal of the pilot study is to assess the app's performance, improve the study protocol, and fine-tune question wording and the overall user experience. At the end of the pilot study, participants took part in a structured interview about their experience of answering questions on the smartphone and smartwatch.

## Pilot Study $\mu$ EMA Response Behavior

A total of 15 participants completed 1 month of data collection (3 weeks with  $\mu$ EMA prompting on the smartwatch). With 15,120  $\mu$ EMA questions delivered, the  $\mu$ EMA response rate was 76.4% (SD 22.3%), with a mean  $\mu$ EMA response time of 4.3 (SD 1.0) seconds.

## Pilot Study Qualitative Feedback

At the end of the 1-month pilot study, participants completed a structured interview lasting for approximately 30 minutes to enable us to gather feedback on how to improve the user experience with the TIME app and to adjust the main study protocol accordingly. When asked about positive experiences with  $\mu$ EMA, participants highlighted the ease of access of the smartwatch and the simplicity of answering just 1 question. A participant described the experience as follows:

*Honestly, it was the easiest to answer surveys on the watch because it's right on your wrist and it vibrates [and] there is a survey, I answer the question and I am done...The watch questions are actually less burdensome than the phone questions. Watch questions are pretty convenient because you can just touch [tap] on the watch and you're done basically.* [F, aged 24 years]

However, several participants felt that the  $\mu$ EMA question did not stay available for long enough. A participant stated as follows:

*I like how it was really simple, it wasn't like a series of questions. They were just "yes" or "no." What made it difficult was the timing...it seems like it only lasts for five seconds and it disappears. Sometimes I may be doing something like resting and then survey comes up. Then I have to stick my arm out of the blanket to answer but then it goes away before I could.* [F, aged 22 years]

Initially, the prompt only remained active for 15 seconds, which we later updated during the pilot to 20 seconds. Participants generally noticed the smartwatch vibration on their wrists with a delay, despite a relatively aggressive vibration pattern. In fact, pilot participants described the  $\mu$ EMA vibration pattern (lasting for 11 seconds, intensifying progressively) as "too loud." Although the smartwatch vibration creates a quiet buzzing sound when on the arm, if the smartwatch is sitting on a surface (eg, charging), the resulting surface vibration can make a surprisingly loud sound that is difficult to ignore.

Participants reported difficulty answering  $\mu$ EMA questions when both hands were occupied, especially while driving. However, they were instructed to only answer  $\mu$ EMA questions when it was safe to do so, and they were given the specific example of driving as a situation in which they should ignore the prompts. Participants also reported difficulty answering

$\mu$ EMA questions while doing other cognitively demanding activities such as writing. In fact, a participant mentioned missing prompts while being in an examination room, stating as follows:

*Sometimes if I was taking a quiz or test in class, I didn't want to be caught using my watch, so I didn't touch [ie, answer] it.* [M, aged 19 years]

Participants found the ability to undo their responses useful, especially when there were accidental taps on the smartwatch. A participant stated as follows:

*Yes [Undo] was useful because sometimes on the watch...maybe because of the clothes I'm wearing...but [if] an option would be chosen so I'd be able to go "undo" and choose the right answer.* [F, aged 24 years]

Participants also found the *Sort of* answer option helpful in instances when they were not sure of their experience at the moment and would have had difficulty entering a more limited *Yes or No* response. A participant stated as follows:

*"Sort of" made me think about my answer a little more rather than an outright yes or no, which is how I kind of determine how I'm going to answer yes or no...if you're feeling such and such, [then] I do like the "sort of" option because you're not feeling fully something. And you are not "not" feeling it completely either. So ["Sort of"] is a pretty easy option for people who are feeling like in the middle.* [F, aged 24 years]

In response to such observations with the goal of achieving cognitive simplicity to support microinteraction, we designed all the questions to have a *middle* answer such as *sort of*.

In the pilot study, each question presented was randomly sampled from the pool of questions. This resulted in questions either being overrepresented or repeated too consecutively for some individuals, adding monotony, which some participants noted. A participant stated as follows:

*I think there were questions based on time or place. Questions like how was your day?...It would ask multiple times a day. I don't know if it was intentional, but it feels like I'm answering the same question over and over again. Same thing for places. It would ask me multiple times "is this your home?" or something like this.* [M, aged 18 years]

## Changes to the Main Study Protocol

### Overview

**Table 2** summarizes the changes made to the  $\mu$ EMA app and protocol based on the pilot participants' feedback. In summary, there were 4 major changes to the main protocol based on the data and participant feedback in the pilot study. These changes are explained in detail in the next sections.

**Table 2.** Changes made to the microinteraction ecological momentary assessment ( $\mu$ EMA) protocol or technology before the main study in response to the pilot participants' feedback.

	Pilot study	Main study
Physical activity and sedentary behavior questions	Asked at random	Asked as sensor-triggered questions based on wrist-accelerometer activity
Location-based questions	Asked at random	Removed from $\mu$ EMA and included as part of mobile phone EMA to obtain semantic location labels
$\mu$ EMA question selection	At random	Using filter-based sampling algorithm
Person-level characteristics and validation questions	Not included in pilot study	Included in main study to assess validity and add variety to questions on $\mu$ EMA days
Thank-you messages after answering a $\mu$ EMA question	Messages selected at random from a bank of 10 messages	Messages selected sequentially from a bank of >250 messages
$\mu$ EMA prompt wait time	15 seconds	Increased to 20 seconds
$\mu$ EMA prompt vibration	11 seconds of progressively intense vibration pattern	6 seconds (2-second vibration with 1-second pause)

### Changed $\mu$ EMA Question-Selection Algorithm

In the pilot study, each  $\mu$ EMA prompt selected a question at random. As a result, the app had less control over preventing overrepresentation of questions within the day. On the basis of the interviews, we learned that participants found this repetitiveness of the questions monotonous. Thus, for the main study we developed a new question-selection algorithm that (1) guarantees that a given question is never asked consecutively and (2) ensures that a question is not answered more than a predetermined number of times.

### Added Sensor-Triggered Questions

In the pilot study, we randomly sampled questions on PA, SB, and the participant's current location. However, these questions also felt repetitive to the pilot participants. Thus, in the main study protocol, these questions were asked when the passive sensors (using accelerometer and GPS) detected a relevant event, gathering information *only* when it was contextually relevant.

### Added $\mu$ EMA Validation and Engagement Questions

We added 2 additional types of questions for  $\mu$ EMA. First, we included *validation questions* to check whether the participants are paying attention to the  $\mu$ EMA prompts. Second, we added questions related to the slowly changing characteristics of the participants; these questions add variety to the question pool while also gathering useful information. Finally, on the  $\mu$ EMA undo screen, we added more variety in the thank-you messages that participants receive after answering a  $\mu$ EMA prompt. As the main study would last for 12 months per participant, we deemed reducing burden and increasing engagement to be important considerations.

### Optimized $\mu$ EMA Interface

We increased the time duration that a  $\mu$ EMA question stays visible on the smartwatch face from 15 seconds to 20 seconds. To address participant concerns about the loud vibration, we changed the pattern from an 11-second-long vibration to a 6-second pattern of 2-second vibrations separated by 1-second pauses.

### Main TIME $\mu$ EMA Substudy

The protocol described here is based on changes made to the  $\mu$ EMA app after the pilot study feedback.

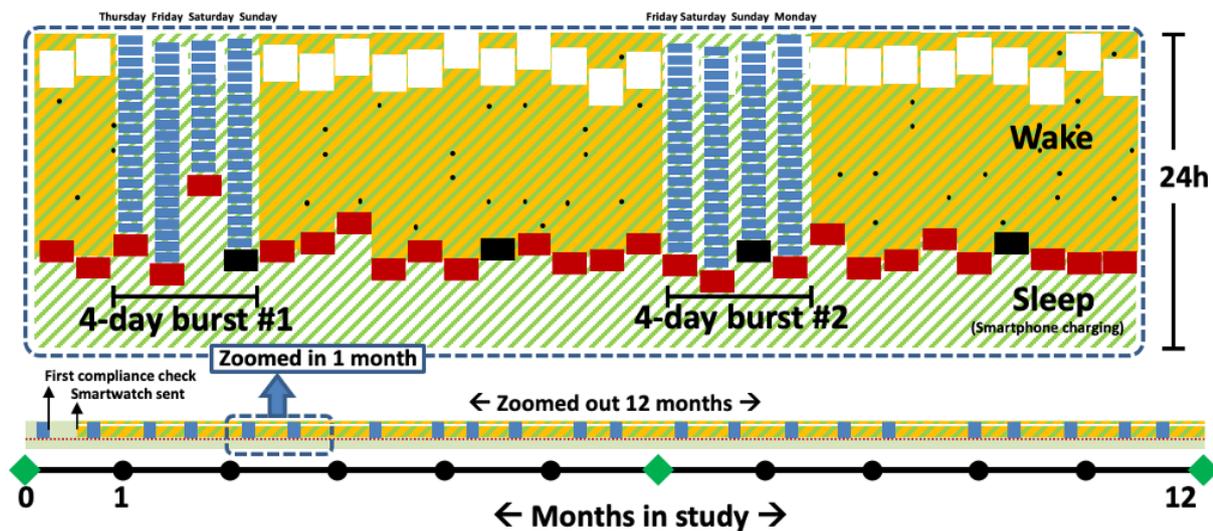
### TIME Study Design

The TIME Study uses a nested design with multiple measurement bursts of smartphone-based EMA across a 12-month study period (Figure 3). Each measurement burst occurs over 4 days and always includes Saturday and Sunday. Measurement bursts occur every 2 weeks. The app ensures that there is at least a one-week gap between 2 consecutive burst periods. Thus, for a 12-month period (ie, 52 weeks), we expect up to 26 EMA burst periods (104 EMA burst days). The remaining nonburst days (or  $\mu$ EMA days) are reserved for  $\mu$ EMA prompting on the smartwatch; there can be up to 261 such days, with the number per participant dependent upon the receipt date of the smartwatch that is mailed only after the first successful burst period. The smartphone automatically sets the study schedule and controls the smartwatch so that  $\mu$ EMA questions are only prompted on the appropriate days and at the appropriate times. End-of-day surveys are prompted on all days, irrespective of measurement bursts. No  $\mu$ EMA prompts occur on the EMA burst days, and no other smartphone-based EMA prompts occur on  $\mu$ EMA days except for the end-of-day survey and sensor-triggered location surveys. Details of the smartphone-based EMA protocol are beyond the scope of this paper; they can be found in the main protocol paper (Wang, S, unpublished data, January 2022). Participants in the TIME Study are exposed to more interruptions from  $\mu$ EMA than from EMA, both within a day (4  $\mu$ EMA prompts vs 1 EMA prompt per hour with possible additional reprompts) and across months (approximately 261 days for  $\mu$ EMA vs approximately 104 days for EMA for a year). Participants are also asked to complete baseline, 6-month, and 12-month surveys on the web using REDCap (Research Electronic Data Capture) [51]; the 6-month and 12-month surveys include questions about the perceived burden [52] of responding to  $\mu$ EMA and EMA questions. After the completion of the study, participants undergo a semistructured interview designed to gather additional information about their overall experiences answering questions on the devices. Participants who voluntarily withdraw from the

study or who are asked to withdraw because of low EMA compliance are also asked to complete an exit survey to enable

us to learn more about their experiences and reasons for withdrawal or poor compliance.

**Figure 3.** Temporal Influences on Movement & Exercise Study protocol with nested ecological momentary assessment (EMA) bursts and microinteraction EMA ( $\mu$ EMA) on nonburst days; CS: context-sensitive.



- EMA (full surveys; 1-2 minutes each; approximately 60 minutes apart; reflective variables; 1536 prompts maximum)
- EMA (end-of-day diaries; full surveys; approximately 2 minutes each; reflective variables; 313 prompts maximum)
- EMA (end-of-day diary+Sun weekly survey; full surveys; approximately 3 minutes each; reflective variables; 52 prompts maximum)
- CS  $\mu$ EMA and EMA (single context-triggered questions; reactive variable and behavioral outcomes; approximately 6 per day)
- /// Sensors (reactive variables and behavioral outcomes; continuous data from smartwatch and smartphone)
- $\mu$ EMA (single questions; 4 s each; reactive variables and behavioral outcomes; 4 per hour, approximately 15,000 prompts maximum)
- Smartwatch charging, smartphone sensors only (approximately 45-90 minutes)
- ◆ Baseline, 6-month, and end-of-Study surveys administered through web (approximately 90 minutes)

## Measures

Here we describe the self-report measures, question-selection algorithms, and simulation results to validate the filtering algorithm to select questions for  $\mu$ EMA.

### Self-report $\mu$ EMA Measures

#### Core-Construct Questions

Core-construct questions measure different behaviors, momentary states, and contexts that cannot be objectively captured using sensors and are expected to change throughout the day. The questions are selected to capture reflective processes, reactive processes, and internal and external processes that can predict or explain behavior adoption and maintenance [49]. Example constructs include feelings of pain, stress, procrastination, happiness, and positive and negative affect. These questions are sampled from a question bank (based on a procedure defined later) consisting of 30 questions (Multimedia Appendix 1); they are presented with signal-contingent prompts based on the aforementioned schedule. Each day during waking hours, 90% (55/62) of the scheduled prompts are reserved for these core-construct questions. In the question bank of 30

core-construct questions, 6 (20%) measure reflective processes, 5 (17%) measure reactive processes, 6 (20%) measure external factors, and 13 (43%) measure internal factors.

#### Engagement and Validation Questions

In addition to the core-construct questions, 8% (5/62) of the scheduled prompts during waking hours are used to present questions about person-level characteristics. These are questions that gather data on attributes such as whether a participant owns a car, cares for a pet, self-describes as a morning person, or has other specific personality characteristics. The questions serve two purposes: gathering information that changes relatively slowly and breaking monotony (thus, perhaps helping with participant engagement) by *surprising* participants with approximately 4 novel questions a day. These questions follow the wording requirements of other  $\mu$ EMA questions (eg, *Own a pet?* with answers *Yes* or *No* or *Bike to work?* with answers *Frequently*, *Sometimes*, and *Rarely*). The system uses a question bank of >280 questions so that they will not repeat for approximately 3½ months (assuming 60  $\mu$ EMA prompts per day).

The remaining 2% (1/62) of the waking-day prompts are used to present validation questions designed to assess whether participants are paying attention to the  $\mu$ EMA questions and answering them carefully (ie, not randomly tapping on the screen to dismiss a question without looking at it, carelessly answering prompts incorrectly, or accidentally pressing buttons on the smartwatch without realizing it). The validation questions are like other  $\mu$ EMA questions—cognitively simple single questions with 2-3 answers. However, these questions always have an unambiguously correct answer to verify response accuracy. Validation questions include simple math problems ( $1 + 3 = 5?$ —*Yes* or *No*), trivial trivia questions (*Sun rises in the east?*—*Yes* or *No*), intentionally silly questions (*Do pigs fly?*—*Yes* or *No*), and simple attention-assessment questions (*Please press A:*—*A, B, or C*). To break the question monotony and enhance engagement, validation questions rarely repeat. Our item bank has >300 validation questions for an estimated 7½ months of question prompting without repetition.

### PA and SB Questions

Sensor-triggered questions are prompted to measure PA and SB, which are indirectly, passively measured using sensors but may require self-report for validation. Participants are asked to report whether they are currently engaging in PA or SB based on their activity level as inferred from the smartwatch raw acceleration data (collected continuously at 50 Hz). In real time, we compute the area under the curve (AUC) of the high-passed raw accelerometer signal for 10-second epochs [53]. The AUC is then used to determine the prompting of sensor-triggered questions. If the AUC exceeds an empirically determined threshold in a continuous 10-minute window, a PA question is eligible to be asked. If the AUC falls below a threshold for a continuous 60-minute window, an SB question is eligible to be asked [53].

Sensor-triggered questions are prompted *only* when a specific sensor event of interest occurs that suggests contextual relevance. The questions are *Physically active [ $\Delta$ ] min ago?* and *Sedentary [ $\Delta$ ] min ago?*, with answer options *Yes*, *Sort of*, and *No*, where  $\Delta$  is the time difference from the current time to the end of the activity plus 25% of the activity bout length (ie, 2½ minutes for a 10-minute PA bout or 15 minutes for a 60-minute sedentary bout). This additional 25% of bout length is used to ensure that the question is asked at approximately the moment in time when the behavior was likely taking place but not at the behavioral transition.

At each scheduled  $\mu$ EMA prompt time, if there is an eligible sensor-triggered question, the system will trigger the question only 50% of the time. This is done to avoid continuous and monotonous sensor-triggered prompting for prolonged sedentary and physical activities. When the sensor-triggered questions are eligible to be prompted, they take priority over the core-construct, validation, and engagement questions. For instance, if the next scheduled prompt has a core-construct question queued but the app has identified a PA event, then 50% of the time, that core-construct question is replaced by the PA confirmation question. If both PA and SB are identified, the question about the most recent event is presented.

### Question Selection for Prompting

A way to gather self-report data on a comprehensive set of behaviors using  $\mu$ EMA is to randomly select a question to ask the user—as was done in our pilot study. However, it may be useful to measure internal factors (eg, perceived stress, fatigue, and nervousness) more frequently within a day rather than external factors (eg, being with a friend). The app uses a strategic sampling algorithm to maximize the value of data collected about some constructs by limiting presentation of some questions, ensure that questions differ within an hour, and guarantee that each day includes questions from each of four broad construct categories (ie, internal and external factors and reflective and reactive processes).

### Selecting Day-Level Subset for Core-Construct Questions

Each  $\mu$ EMA day, from our bank of 30  $\mu$ EMA questions of core constructs, the app selects 4 (13%) questions related to internal factors, 2 (7%) to external factors, 2 (7%) to reflective processes, and 2 (7%) to reactive processes. More internal factors than other categories are selected because they are expected to vary throughout the day [54]. Rather than selecting any question at random, this procedure ensures that at least one question measuring each of the different factors is presented each day, with a temporally dense presentation. Each question is expected to be presented  $\geq 4$  times in a day, allowing measurement of within-day changes.

### Filters Applied for Core-Construct Question Selection

#### Maximum Allowable Prompting per Question

Each  $\mu$ EMA question is assigned a maximum number of presentations per day (Multimedia Appendix 1). This ensures that some questions are not overrepresented in a day at the expense of others. For instance, questions capturing frequently changing constructs are assigned a higher limit (eg, *Feeling stressed?*) than questions that assess more stable constructs (eg, *Slept well yesterday?*). If the number of answers received for a question reaches the question's limit, that question is no longer presented that day.

#### Minimum Time Gap for Question Repetition

Once a particular  $\mu$ EMA question has been answered, it will not be presented again until at least 60 minutes have elapsed. For instance, if the question *Feeling stressed?* is answered, it is not asked again for at least one hour. However, if the question is not answered, it could be presented at the next prompt, within the hour.

#### Filtering Algorithm When Selecting a $\mu$ EMA Question

The app uses the following algorithm to select  $\mu$ EMA questions (Textbox 1). For each  $\mu$ EMA day ( *$\mu$ EMA\_prompting\_day*, line 1), the app first generates a day-level subset (*day\_level\_subset*, line 2) of 10 questions from the pool of 30  $\mu$ EMA questions (*all\_μEMA\_questions*). Next, at each  $\mu$ EMA prompt, day-level questions that have reached their limit (*remove\_maxed\_out\_questions*, line 4) or have been asked recently (*remove\_questions\_answered\_within\_one\_hour*, line 5) are filtered. From the remaining list, a question is selected at random (*select\_question\_at\_random*, line 6). If the question

prompted is answered, the question's prompt count is updated (*update\_prompt\_count*, line 8), as is the last-answered time (*update\_answer\_time*, line 9). If the question is not answered,

then it is available for selection in the next scheduled  $\mu$ EMA prompt.

**Textbox 1.** Microinteraction ecological momentary assessment ( $\mu$ EMA) question-filtering algorithm at each prompt on nonburst days.

#### Algorithm used to select $\mu$ EMA questions

1. For each  $\mu$ EMA\_prompting\_day:
2. day\_level\_subset=subset (all\_  $\mu$ EMA \_questions)
3. For each  $\mu$ EMA prompt:
4. filtered\_questions=remove\_maxed\_out\_questions (day\_level\_subset)
5. filtered\_questions=remove\_questions\_answered\_within\_one\_hour (filtered\_questions)
6. question\_to\_ask=select\_question\_at\_random (filtered\_questions)
7. if question\_to\_ask is answered:
8. update\_prompt\_count (question\_to\_ask)
9. update\_answer\_time (question\_to\_ask)

### $\mu$ EMA Question-Selection Simulation

To verify our filtering algorithm (Textbox 1), we simulated question sampling and compared the distributions of questions asked through random sampling (used in the pilot study) with our proposed protocol.

### Question-Sampling Algorithm Simulation

We simulated question selection for 365 days for 300 users, where 261 days were reserved for  $\mu$ EMA prompting, resulting in 78,300 total  $\mu$ EMA days. We assumed 8 hours of sleep and 16 hours of wake time per day, resulting in 15½ hours available for  $\mu$ EMA prompting. Targeting 4 prompts an hour results in 62  $\mu$ EMA prompting opportunities. Of these, 90% (55/62) were available for core constructs, resulting in 55 slots for core-construct questions each day. We assumed that 75% (46/62) of the scheduled prompts would be answered (ie, compliance).

### Algorithm Simulation Results

Compared with random sampling (Multimedia Appendix 1), the filtering algorithm increased the frequency of presentation of core-construct questions in a day. For instance, when using random sampling, certain items such as sleep satisfaction are repeated at the expense of other variables of interest (eg, internal factors). Although with random sampling a question is prompted on more days (approximately 75% vs approximately 33%), the questions are repeated only twice (median frequency) on a given day. This presentation is insufficient to study within-day changes in slopes when we need at least three observations. The filtering algorithm ensures that on days when questions are presented, they are repeated 4 times (median frequency) a day. Later, we validate this algorithm against the real-world data from our main study participants (Multimedia Appendix 2).

### TIME Study Main Trial Participants and Recruitment

Participants were eligible for the TIME Study main trial if they (1) owned an Android smartphone running Android version  $\geq 6.0$  as their only personal mobile phone with no intention to switch to a non-Android smartphone, (2) did not wear a smartwatch already, (3) were aged 18-29 years and living in the United States, (4) were currently engaged in recommended levels of PA (or intended to within the next 12 months) [55], (5) spoke and read English, (6) resided in an area with Wi-Fi connectivity, (7) did not have any physical or cognitive limitations that prevented participation, and (8) were able wear a smartwatch and answer real-time smartphone and smartwatch surveys. All study procedures were approved by the institutional review board at the University of Southern California (USC; HS-18-00605). Participants were recruited using several strategies: (1) sending emails to individuals enrolled in the Happiness & Health Study, a USC longitudinal cohort study of young adults [56]; (2) posting flyers in the greater Los Angeles metropolitan area; (3) purchasing web and social media advertisements; (4) sending emails to addresses on file from other institutional review board-approved USC studies; and (5) contacting participants identified using ResearchMatch [57]. As of June 2021, 246 participants had been enrolled, among whom 81 (32.9%) had completed at least 6 months of data collection from their respective start dates (Table 3). We recruited 246 participants who received a smartwatch after the initial run-in period. This sample size was determined based on the objectives of the main TIME Study protocol (Wang, S, unpublished data, January 2022). The  $\mu$ EMA substudy was added (Figure 3) for exploratory purposes.

**Table 3.** Participant demographic survey for those who had completed 6 months as of June 2021 (N=81).

Demographics	Values
Age (years), mean (SD)	21.7 (2.4)
<b>Sex, n (%)</b>	
Male	45 (55)
Female	36 (45)
<b>Ethnicity, n (%)</b>	
Non-Hispanic	51 (63)
Hispanic	30 (37)
<b>Race<sup>a</sup>, n (%)</b>	
White	41 (54)
Asian or Pacific Islander	35 (46)
Black	7 (9)
American Indian or Alaska Native	5 (7)
<b>Education, n (%)</b>	
High school	12 (15)
Some college	47 (58)
College graduate	22 (27)
<b>Work status<sup>b</sup>, n (%)</b>	
Employed	36 (45)
Out of work	16 (20)
Student	53 (66)
Unable to work	4 (5)

<sup>a</sup>Participants could select >1 answer option from American Indian or Alaska Native, Hawaiian or Pacific Islander, Black or African American, White, Asian, Unknown, and Would prefer not to answer.

<sup>b</sup>Participants could select >1 answer option; for example, “Student” and “Employed.”

## Study Procedures

Because of COVID-19 restrictions on in-person recruitment for human subjects research, all participant recruitment and onboarding was conducted remotely. After screening, research staff individually met with each participant remotely (through Zoom) to obtain informed consent. Staff then guided the participants through the TIME app installation on their personal smartphones. Researchers then observed EMA burst compliance for the first 4-day EMA burst (ie, run-in period) during the next 2 weeks. If the compliance for the run-in period was <8 surveys per day on all days, the participants were withdrawn from the study. Otherwise, they could continue in the study and were mailed a smartwatch (Fossil Gen 4 or Gen 5 model). Once the participants received the smartwatch, a staff member scheduled a second remote orientation session with the participant to guide them through the TIME smartwatch setup. Expectations for smartwatch wear (including wearing it during sleep), charging it, and answering  $\mu$ EMA questions were explained. Participants were instructed that to remain compliant they should wear the smartwatch for 23 hours a day and charge it every day. They were also instructed to wear the smartwatch while sleeping. For  $\mu$ EMA prompting, participants were instructed as follows:

*“You will be prompted on the smartwatch with single questions of the type Are you walking right now? Yes / Sort of / No. These questions will take only 2-4 seconds to respond to—just like checking time on the smartwatch. There may be up to six such question prompts in an hour, on average, and the frequency of question prompts may vary at different times of your waking day. All the smartwatch questions will be outside of the 4-day smartphone survey burst periods. Each watch prompt will wait for up to 20 seconds for you to respond; if the question is not answered within 20 seconds, the question may disappear from the watch. Unlike phone surveys, watch questions are not re-prompted.”*

After the smartwatch orientation, participants could use the smartwatch as they deemed fit (if it did not interfere with TIME app functionality), while answering  $\mu$ EMA questions when they appeared. Participants receive US \$20 per month for wearing the smartwatch for 23 hours on at least 24 days of the month. Participants could earn up to an additional US \$80 per month for EMA burst compliance (Wang, S, unpublished data, January 2022). In addition, if at the end of the study, participants have >50% compliance with  $\mu$ EMA, they were able to keep the

smartwatch as their personal device. However, no monthly or regular compensation was provided for  $\mu$ EMA compliance. Participants do not receive feedback on their  $\mu$ EMA compliance, and their withdrawal from the study does not depend on  $\mu$ EMA compliance.

### Response Behavior Measures

We measure compliance rate, completion rate, undo rate, and validation rate to characterize participant response behavior when answering  $\mu$ EMA questions.

#### Compliance Rate

Compliance rate is measured as the percentage of  $\mu$ EMA questions answered out of all the scheduled questions, including those not prompted because of the device being turned off or in DND mode.

$$\text{Compliance rate (\%)} = \frac{\text{\#Questions answered}}{\text{\#Questions scheduled}} \times 100$$

#### Completion Rate

Completion rate is measured as the percentage of  $\mu$ EMA questions answered out of all the delivered questions, excluding those not prompted because of the device being turned off or in DND mode.

$$\text{Completion rate (\%)} = \frac{\text{\#Questions answered}}{\text{\#Questions delivered}} \times 100$$

#### Undo Rate

Undo rate is measured as the total number of  $\mu$ EMA questions answered when the *Undo?* option was selected after providing an initial answer.

$$\text{Undo rate (\%)} = \frac{\text{\#Undo count}}{\text{\#Questions answered}} \times 100$$

#### Validation Rate

Validation rate is measured as the percentage of  $\mu$ EMA validation questions answered correctly out of all the validation questions answered.

$$\text{Validation rate (\%)} = \frac{\text{\#Validation correctly answered}}{\text{\#Validation answered}} \times 100$$

## Results

### Compliance, Completion, and Validation Rates

Overall, the compliance rate ( $n=81$  at  $\geq 6$  months) for  $\mu$ EMA was 67.7% (SD 13.7%) and the completion rate was 80.23% (SD 13.3%). The  $\mu$ EMA validation rate for the participants ( $n=81$ ) was 92.9% (SD 23.7%)—indicating that they were paying attention to the  $\mu$ EMA question content and not carelessly tapping to dismiss them.

### Response Behavior

Table 4 presents a summary of the response behavior. We also verified the filter-based question-selection algorithm's performance against the real-world data from 2% (2/81) of the study participants who completed 6 months in the study ([Multimedia Appendix 2](#)). We found that no core-construct question was asked more than its maximum allowable number of times. In addition, the median number of times a core-construct question was asked was consistent with our simulation results ([Multimedia Appendix 1](#)).

**Table 4.** Main study microinteraction ecological momentary assessment ( $\mu$ EMA) response behavior for participants completing 6 months as of June 2021 (N=81).

	Values
$\mu$ EMA days, n	13,415
Expected $\mu$ EMA questions, n	790,388
Delivered $\mu$ EMA questions, n, %	662,397 (83.81)
Answered $\mu$ EMA questions, n, %	535,430 (80.83)
Mean daily $\mu$ EMA compliance rate, % (SD)	67.6 (24.4)
Mean daily $\mu$ EMA completion rate, % (SD)	78.5 (22.2)
Mean participant $\mu$ EMA compliance rate, % (SD)	67.4 (13.7)
Mean participant $\mu$ EMA completion rate, % (SD)	80.2 (13.3)
Mean $\mu$ EMA response time, seconds (SD)	4.8 (1.4)
Total $\mu$ EMA undos (% of total $\mu$ EMA questions answered)	22,202 (4.2)
Mean $\mu$ EMA question validation rate, % (SD)	92.9 (23.7)

## Discussion

Our work with  $\mu$ EMA demonstrates that the technique may enable temporally dense ILD collection with manageable burden to support longitudinal studies or interventions, although some limitations of this work will require further investigation.

### Summary and Strengths

Traditional mobile phone-based EMA can lead to interruption and response burden. To prevent this burden, researchers often compromise on the temporal density of prompts (ie, by prompting less frequently) and reduce the number of questions or constructs being measured. Complementary forms of EMA are needed where both the researchers' needs for comprehensive

understanding of health behaviors and users' concern for burden are taken into account.  $\mu$ EMA provides such an opportunity, where the quick microinteractions not only make it less burdensome for the users to answer questions on the smartwatch, but also enable data gathering at a higher temporal density.

The TIME  $\mu$ EMA substudy combines smartphone-based EMA with low-burden  $\mu$ EMA on a smartwatch to gather real-time self-report data from naturalistic settings using the personal smartphones of participants. Participants are encouraged to use their smartphone and the loaned smartwatch normally with only a few limitations (ie, avoiding apps that interfere with TIME app functionality, such as fitness trackers). Participants are incentivized to wear the smartwatch, but they can pause smartwatch survey prompting as needed. Participants are not directly compensated for high  $\mu$ EMA compliance, although they are told that they can keep the smartwatch if they have >50%  $\mu$ EMA compliance at the end of 1 year. EMA compliance, by contrast, is reinforced with an explicit and compliance-contingent reward provided at regular monthly intervals.

Our preliminary results from the pilot and ongoing TIME  $\mu$ EMA substudy show that despite a high interruption rate,  $\mu$ EMA could be sustainable for gathering self-report data longitudinally. Controlled filtering of  $\mu$ EMA questions enables measurement of a comprehensive set of behaviors repeatedly during waking hours. The validation questions make it possible to identify careless responding on both  $\mu$ EMA and smartphone-based EMA.

The TIME  $\mu$ EMA substudy can be used to assess the utility and sustainability of  $\mu$ EMA. Comparing response rates with the qualitative data from the exit interviews, for example, suggests that  $\mu$ EMA can be used to gather temporally dense self-report information. In addition, passively collected data on wrist motion (using accelerometers) and location (using GPS) may help explore contextual factors that affect  $\mu$ EMA compliance. The filtering-based  $\mu$ EMA question-selection algorithm may be used to capture diurnal patterns of different constructs during waking hours. These diurnal patterns could be used to explore (1) different clusters of individuals who have similar diurnal patterns for variables of interest [58] and (2) how similar, or different, diurnal patterns are that were captured using  $\mu$ EMA versus those captured using EMA. With 4 prompts per hour (or 70 prompts for an 18-hour wake period),  $\mu$ EMA may enable denser measurement than EMA. Combined with passive sensing,  $\mu$ EMA may support study of diurnal patterns, which might be compared with EMA at the end of the day.

Overall, the purpose of  $\mu$ EMA is not to replace traditional EMA but to complement it. The TIME Study's  $\mu$ EMA substudy protocol provides an opportunity to explore how small amounts of information gathered at high frequency can help us learn about health behavior change and maintenance, even as constructs related to behavior change deviate by the hour. Beyond observation studies, TIME Study data could enable exploration of the viability of  $\mu$ EMA for longitudinal interventions (eg, just-in-time adaptive interventions), where self-report on many different behaviors may be required. For instance, when optimizing mobile health interventions as part of microrandomized trials [59],  $\mu$ EMA may provide a

low-burden self-report interface that can gather information about behaviors that sensors cannot measure yet (eg, fatigue, procrastination, and pain), especially when such measurements are needed in close time intervals (eg, within an hour) and across changing contexts.

## Limitations and Opportunities

There will be several limitations of the TIME  $\mu$ EMA substudy that will provide opportunities for future research. Findings about the  $\mu$ EMA methodology may not generalize to other population age groups because the main protocol is designed to study health behavior adoption and maintenance in young adults (age 18-29 years). There has been an increased interest in exploring the acceptance of wearable technologies in older adults [60-62], youth [63], and children [64] and more research is needed to explore the acceptance of  $\mu$ EMA as a data collection method for both longitudinal observational and intervention studies with these demographic groups. As  $\mu$ EMA is deployed on a wearable smartwatch, our data collection excludes participants who may not be allowed to wear, or be comfortable wearing, the smartwatch during work hours; in some professions, answering  $\mu$ EMA or EMA surveys may not be appropriate (eg, professions that require being in intensive care units, using cleanrooms, extended driving, operating heavy machinery, and working in construction).

We had to restrict our data collection to Android users because our research app used advanced sensor capabilities not programmatically available on iOS devices; thus, the TIME  $\mu$ EMA substudy excludes iOS users. Nevertheless, we observed more demographic diversity in the Android users who participated in our screening surveys (of the 746 participants, 393, 52.7%, were women; 372, 49.9%, were White; 216, 28.9%, were Asian; 119, 15.9%, were Black; and 184, 24.6%, were Hispanic; 630, 84.5%, engage in PA) versus the iOS users (of the 395 participants, 296, 75.1%, were women; 177, 44.9%, were White; 99, 25.1%, were Asian; 71, 17.8%, were Black, and 130, 32.9% were Hispanic; 365, 92.4%, engage in PA). We have no reason to believe that the platform itself would affect  $\mu$ EMA response patterns.

Although  $\mu$ EMA may keep burden more manageable than EMA despite intensive prompting, each question provides a limited answer set (*Yes*, *Sort of*, and *No*) that may be less sensitive to construct variance than mobile phone-based EMAs with more response options. Yet, with a median of  $\geq 4$  measurements per construct within a day (Multimedia Appendix 1),  $\mu$ EMA provides opportunities to use mixed effects location scale models to study changes in variances and slopes of different constructs [65,66]. This also provides opportunities to explore approaches beyond multilevel models, for instance, using dynamic Bayesian networks [67,68] or temporal networks [69,70] to build individual-specific models.

Unlike EMA, where multiple questions are asked back to back, thus allowing for the simultaneous comparison of different behaviors,  $\mu$ EMA presents only 1 question at a time to ensure that each prompt can be responded to as a microinteraction. Therefore, different construct measurements are spread throughout waking hours at different times. Although this approach does not allow for testing concurrent associations

among constructs, there are alternative ways of exploring temporally lagged associations of different  $\mu$ EMA measures (eg, Granger Causality [71]).

### Conclusions

The TIME  $\mu$ EMA substudy provides an opportunity to explore a new method of collecting temporally dense self-report data. Although EMA provides information on temporal dynamics of behavior, the burden of accessing the mobile phone, unlocking it, and then answering multiple complex questions compromises engagement. In this formative research, we explored the feasibility of using  $\mu$ EMA to measure multiple constructs in an ILD study. When deploying  $\mu$ EMA, each prompt is a cognitively simple single question that can be answered with a quick

glanceable microinteraction with an always-accessible smartwatch. Because of this simplicity,  $\mu$ EMA permits a higher interruption rate than EMA without a corresponding increase in perceived burden. The TIME Study is the first ILD study to deploy  $\mu$ EMA for an entire year; thus, it could provide insights on methodological properties of  $\mu$ EMA—especially the ability to sustain participant engagement. Data from the study could be used to study what influences  $\mu$ EMA compliance, what can be learned about an individual's behavior using  $\mu$ EMA, and how  $\mu$ EMA might be used along with EMA and passive sensors to richly characterize human behavior, especially how it may change throughout the day in response to rapidly changing contexts.

### Acknowledgments

This work is supported by the National Institutes of Health (NIH)/NHLBI U01HL146327. The content is solely the responsibility of the authors and does not represent the official views of the NIH. Krit Goyal, Adeeb M Arif, Binod Thapa Chhetry, Jixin Li, and Qu Tang contributed to the development of the TIME app and associated data processing algorithms. We thank the pilot and main study participants for their time.

### Authors' Contributions

AP led the development of the TIME app, the microinteraction ecological momentary assessment ( $\mu$ EMA) components of the study design,  $\mu$ EMA deployment and troubleshooting, and the first draft of this manuscript. SW, DC, and BD designed the EMA components of the TIME study, baseline and follow-up survey designs, participant recruitment, and analysis of participant demographics and contributed to manuscript writing. GD and SI are the co-primary investigators of the research grant and contributed to research study design, technology design, research oversight, and manuscript writing.

### Conflicts of Interest

None declared.

### Multimedia Appendix 1

Random versus filter-based question sampling.

[\[DOCX File , 19 KB-Multimedia Appendix 1\]](#)

### Multimedia Appendix 2

Filter-based sampling simulation versus main study participant results.

[\[DOCX File , 23 KB-Multimedia Appendix 2\]](#)

### References

1. Nusser S, Intille S, Maitra R. Emerging technologies next generation intensive longitudinal data collection. In: Models for Intensive Longitudinal Data. Oxford, United Kingdom: Oxford University Press; 2006.
2. Intille S. The precision medicine initiative and pervasive health research. IEEE Pervasive Comput 2016 Jan;15(1):88-91. [doi: [10.1109/mprv.2016.2](#)]
3. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015 Feb 26;372(9):793-795 [[FREE Full text](#)] [doi: [10.1056/NEJMp1500523](#)] [Medline: [25635347](#)]
4. Intille SS. Ubiquitous computing technology for just-in-time motivation of behavior change. In: Studies in Health Technology and Informatics. Amsterdam, Netherlands: IOS Press; 2004.
5. Nahum-Shani I, Smith S, Spring B, Collins L, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2018 May 18;52(6):446-462 [[FREE Full text](#)] [doi: [10.1007/s12160-016-9830-8](#)] [Medline: [27663578](#)]
6. Yang C, Maher J, Ponnada A, Dzibur E, Nordgren R, Intille S, et al. An empirical example of analysis using a two-stage modeling approach: within-subject association of outdoor context and physical activity predicts future daily physical activity levels. Transl Behav Med 2021 Apr 26;11(4):912-920 [[FREE Full text](#)] [doi: [10.1093/tbm/ibaa107](#)] [Medline: [33159452](#)]

7. Maher JP, Dzibur E, Nordgren R, Huh J, Chou C, Hedeker D, et al. Do fluctuations in positive affective and physical feeling states predict physical activity and sedentary time? *Psychol Sport Exerc* 2019 Mar;41:153-161 [FREE Full text] [doi: [10.1016/j.psychsport.2018.01.011](https://doi.org/10.1016/j.psychsport.2018.01.011)] [Medline: [30853854](https://pubmed.ncbi.nlm.nih.gov/30853854/)]
8. Marsch LA. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacology* 2021 Jan;46(1):191-196 [FREE Full text] [doi: [10.1038/s41386-020-0761-5](https://doi.org/10.1038/s41386-020-0761-5)] [Medline: [32653896](https://pubmed.ncbi.nlm.nih.gov/32653896/)]
9. Wongvibulsin S, Martin SS, Saria S, Zeger SL, Murphy SA. An individualized, data-driven digital approach for precision behavior change. *Am J Lifestyle Med* 2020;14(3):289-293 [FREE Full text] [doi: [10.1177/1559827619843489](https://doi.org/10.1177/1559827619843489)] [Medline: [32477031](https://pubmed.ncbi.nlm.nih.gov/32477031/)]
10. Yang X, Ram N, Gest S, Lydon-Staley D, Conroy D, Pincus A, et al. Socioemotional dynamics of emotion regulation and depressive symptoms: a person-specific network approach. *Complexity* 2018;2018:5094179 [FREE Full text] [doi: [10.1155/2018/5094179](https://doi.org/10.1155/2018/5094179)] [Medline: [30613129](https://pubmed.ncbi.nlm.nih.gov/30613129/)]
11. Oyibo K, Adaji I, Vassileva J. Social cognitive determinants of exercise behavior in the context of behavior modeling: a mixed method approach. *Digit Health* 2018;4:2055207618811555 [FREE Full text] [doi: [10.1177/2055207618811555](https://doi.org/10.1177/2055207618811555)] [Medline: [30479828](https://pubmed.ncbi.nlm.nih.gov/30479828/)]
12. Dunton G, Rothman A, Leventhal A, Intille S. How intensive longitudinal data can stimulate advances in health behavior maintenance theories and interventions. *Transl Behav Med* 2021 Feb 11;11(1):281-286 [FREE Full text] [doi: [10.1093/tbm/ibz165](https://doi.org/10.1093/tbm/ibz165)] [Medline: [31731290](https://pubmed.ncbi.nlm.nih.gov/31731290/)]
13. Spruijt-Metz D, Hekler E, Saranummi N, Intille S, Korhonen I, Nilsen W, et al. Building new computational models to support health behavior change and maintenance: new opportunities in behavioral research. *Transl Behav Med* 2015 Sep;5(3):335-346 [FREE Full text] [doi: [10.1007/s13142-015-0324-1](https://doi.org/10.1007/s13142-015-0324-1)] [Medline: [26327939](https://pubmed.ncbi.nlm.nih.gov/26327939/)]
14. Baldwin W. Information no one else knows: the value of self-report. In: *The Science of Self-report: Implications for Research and Practice*. NJ, USA: Lawrence Erlbaum Associates Publishers; 2000.
15. Csikszentmihalyi M, Larson R. Validity and reliability of the Experience-Sampling Method. *J Nerv Ment Dis* 1987 Sep;175(9):526-536. [doi: [10.1097/00005053-198709000-00004](https://doi.org/10.1097/00005053-198709000-00004)] [Medline: [3655778](https://pubmed.ncbi.nlm.nih.gov/3655778/)]
16. Stone A, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. *Annals Behav Med* 1994;16(3):199-202. [doi: [10.1093/abm/16.3.199](https://doi.org/10.1093/abm/16.3.199)]
17. van Berkel N, Ferreira D, Kostakos V. The experience sampling method on mobile devices. *ACM Comput Surv* 2018 Nov 30;50(6):1-40. [doi: [10.1145/3123988](https://doi.org/10.1145/3123988)]
18. Fuller-Tyszkiewicz M, Skouteris H, Richardson B, Blore J, Holmes M, Mills J. Does the burden of the experience sampling method undermine data quality in state body image research? *Body Image* 2013 Sep;10(4):607-613. [doi: [10.1016/j.bodyim.2013.06.003](https://doi.org/10.1016/j.bodyim.2013.06.003)] [Medline: [23856302](https://pubmed.ncbi.nlm.nih.gov/23856302/)]
19. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. *Psychol Assess* 2012 Sep;24(3):713-720. [doi: [10.1037/a0026733](https://doi.org/10.1037/a0026733)] [Medline: [22250597](https://pubmed.ncbi.nlm.nih.gov/22250597/)]
20. Burke LE, Shiffman S, Music E, Styn MA, Kriska A, Smailagic A, et al. Ecological momentary assessment in behavioral research: addressing technological and human participant challenges. *J Med Internet Res* 2017 Mar 15;19(3):e77 [FREE Full text] [doi: [10.2196/jmir.7138](https://doi.org/10.2196/jmir.7138)] [Medline: [28298264](https://pubmed.ncbi.nlm.nih.gov/28298264/)]
21. Stone A, Kessler R, Haythornthwaite J. Measuring daily events and experiences: decisions for the researcher. *J Pers* 1991 Sep;59(3):575-607. [doi: [10.1111/j.1467-6494.1991.tb00260.x](https://doi.org/10.1111/j.1467-6494.1991.tb00260.x)] [Medline: [1960643](https://pubmed.ncbi.nlm.nih.gov/1960643/)]
22. Intille S, Haynes C, Maniar D, Ponnada A, Manjourides J.  $\mu$ EMA: microinteraction-based Ecological Momentary Assessment (EMA) using a smartwatch. *Proc ACM Int Conf Ubiquitous Comput* 2016 Sep;2016:1124-1128 [FREE Full text] [doi: [10.1145/2971648.2971717](https://doi.org/10.1145/2971648.2971717)] [Medline: [30238088](https://pubmed.ncbi.nlm.nih.gov/30238088/)]
23. Ashbrook D. *Enabling Mobile Microinteractions*. Atlanta, GA: College of Computing, Georgia Institute of Technology; 2010.
24. Ashbrook D, Clawson J, Lyons K, Starner T, Patel N. Quickdraw: the impact of mobility and on-body placement on device access time. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2008 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 5 - 10, 2008; Florence, Italy. [doi: [10.1145/1357054.1357092](https://doi.org/10.1145/1357054.1357092)]
25. Dey A, Wac K, Ferreira D, Tassini K, Hong J, Ramos J. Getting closer: an empirical investigation of the proximity of user to their smartphones. In: *Proceedings of the 13th international conference on Ubiquitous computing*. 2011 Presented at: Proceedings of the 13th international conference on Ubiquitous computing; Sep 17 - 21, 2011; Beijing, China. [doi: [10.1145/2030112.2030135](https://doi.org/10.1145/2030112.2030135)]
26. Chun J, Dey A, Lee K, Kim S. A qualitative study of smartwatch usage and its usability. *Hum Factors Man* 2018 Feb 16;28(4):186-199. [doi: [10.1002/hfm.20733](https://doi.org/10.1002/hfm.20733)]
27. Pizza S, Brown B, McMillan D, Lampinen A. Smartwatch in vivo. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016 Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, California, USA. [doi: [10.1145/2858036.2858522](https://doi.org/10.1145/2858036.2858522)]

28. Schneid SD, Armour C, Park YS, Yudkowsky R, Bordage G. Reducing the number of options on multiple-choice questions: response time, psychometrics and standard setting. *Med Educ* 2014 Oct;48(10):1020-1027. [doi: [10.1111/medu.12525](https://doi.org/10.1111/medu.12525)] [Medline: [25200022](https://pubmed.ncbi.nlm.nih.gov/25200022/)]
29. Hick WE. On the rate of gain of information. *Q J Exper Psychol* 2018 Jan 01;4(1):11-26. [doi: [10.1080/17470215208416600](https://doi.org/10.1080/17470215208416600)]
30. Proctor RW, Schneider DW. Hick's law for choice reaction time: a review. *Q J Exp Psychol (Hove)* 2018 Jun;71(6):1281-1299. [doi: [10.1080/17470218.2017.1322622](https://doi.org/10.1080/17470218.2017.1322622)] [Medline: [28434379](https://pubmed.ncbi.nlm.nih.gov/28434379/)]
31. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol* 1954 Jun;47(6):381-391. [Medline: [13174710](https://pubmed.ncbi.nlm.nih.gov/13174710/)]
32. Bi X, Li Y, Zhai S. FFitts law: modeling finger touch with fitts' law. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2013 Presented at: CHI '13: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; Apr 27- May 2, 2013; Paris, France. [doi: [10.1145/2470654.2466180](https://doi.org/10.1145/2470654.2466180)]
33. Schembre SM, Liao Y, O'Connor SG, Hingle MD, Shen S, Hamoy KG, et al. Mobile ecological momentary diet assessment methods for behavioral research: systematic review. *JMIR Mhealth Uhealth* 2018 Nov 20;6(11):e11170 [FREE Full text] [doi: [10.2196/11170](https://doi.org/10.2196/11170)] [Medline: [30459148](https://pubmed.ncbi.nlm.nih.gov/30459148/)]
34. Hawkley LC, Burleson MH, Bertson GG, Cacioppo JT. Loneliness in everyday life: cardiovascular activity, psychosocial context, and health behaviors. *J Pers Soc Psychol* 2003 Jul;85(1):105-120. [doi: [10.1037/0022-3514.85.1.105](https://doi.org/10.1037/0022-3514.85.1.105)] [Medline: [12872887](https://pubmed.ncbi.nlm.nih.gov/12872887/)]
35. Kikuchi H, Yoshiuchi K, Yamamoto Y, Komaki G, Akabayashi A. Does sleep aggravate tension-type headache?: an investigation using computerized ecological momentary assessment and actigraphy. *Biopsychosoc Med* 2011 Aug 12;5:10 [FREE Full text] [doi: [10.1186/1751-0759-5-10](https://doi.org/10.1186/1751-0759-5-10)] [Medline: [21835045](https://pubmed.ncbi.nlm.nih.gov/21835045/)]
36. Hernandez J, McDuff D, Infante C, Maes P, Quigley K, Picard R. Wearable ESM: differences in the experience sampling method across wearable devices. In: *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 2016 Presented at: MobileHCI '16: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services; Sep 6 - 9, 2016; Florence, Italy. [doi: [10.1145/2935334.2935340](https://doi.org/10.1145/2935334.2935340)]
37. Timmermann J, Heuten W, Boll S. Input methods for the Borg-RPE-scale on smartwatches. In: *Proceedings of the 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. 2015 Presented at: 2015 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth); May 20-23, 2015; Istanbul, Turkey. [doi: [10.4108/icst.pervasivehealth.2015.259220](https://doi.org/10.4108/icst.pervasivehealth.2015.259220)]
38. Ponnada A, Haynes C, Maniar D, Manjourides J, Intille S. Microinteraction ecological momentary assessment response rates: effect of microinteractions or the smartwatch? *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017 Sep;1(3):1-16 [FREE Full text] [doi: [10.1145/3130957](https://doi.org/10.1145/3130957)] [Medline: [30198012](https://pubmed.ncbi.nlm.nih.gov/30198012/)]
39. Ponnada A, Thapa-Chhetry B, Manjourides J, Intille S. Measuring criterion validity of microinteraction ecological momentary assessment (Micro-EMA): exploratory pilot study with physical activity measurement. *JMIR Mhealth Uhealth* 2021 Mar 10;9(3):e23391 [FREE Full text] [doi: [10.2196/23391](https://doi.org/10.2196/23391)] [Medline: [33688843](https://pubmed.ncbi.nlm.nih.gov/33688843/)]
40. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988 Jun;54(6):1063-1070. [doi: [10.1037//0022-3514.54.6.1063](https://doi.org/10.1037//0022-3514.54.6.1063)] [Medline: [3397865](https://pubmed.ncbi.nlm.nih.gov/3397865/)]
41. Dunton GF, Liao Y, Intille S, Huh J, Leventhal A. Momentary assessment of contextual influences on affective response during physical activity. *Health Psychol* 2015 Dec;34(12):1145-1153. [doi: [10.1037/hea0000223](https://doi.org/10.1037/hea0000223)] [Medline: [26053885](https://pubmed.ncbi.nlm.nih.gov/26053885/)]
42. Larsen J, Eskelund K, Christiansen T. Active self-tracking of subjective experience with a one-button wearable: a case study in military PTSD. In: *Proceedings of the CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2017 Presented at: CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems; May 6 - 11, 2017; Denver Colorado USA.
43. King ZD, Moskowitz J, Egilmez B, Zhang S, Zhang L, Bass M, et al. micro-Stress EMA: a passive sensing framework for detecting in-the-wild stress in pregnant mothers. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2019 Sep;3(3):91 [FREE Full text] [doi: [10.1145/3351249](https://doi.org/10.1145/3351249)] [Medline: [32432212](https://pubmed.ncbi.nlm.nih.gov/32432212/)]
44. Jayathissa P, Quintana M, Abdelrahman M, Miller C. Humans-as-a-sensor for buildings—intensive longitudinal indoor comfort models. *Buildings* 2020 Oct 01;10(10):174. [doi: [10.3390/buildings10100174](https://doi.org/10.3390/buildings10100174)]
45. Paruthi G, Raj S, Gupta A, Huang C, Chang Y, Newman M. HEED: situated and distributed interactive devices for self-reporting. In: *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. 2017 Presented at: UbiComp '17: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers; Sep 11 - 15, 2017; Maui, Hawaii. [doi: [10.1145/3123024.3123162](https://doi.org/10.1145/3123024.3123162)]
46. Paruthi G, Raj S, Baek S, Wang C, Huang C, Chang Y, et al. Heed: exploring the design of situated self-reporting devices. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018 Sep 18;2(3):1-21. [doi: [10.1145/3264942](https://doi.org/10.1145/3264942)]
47. Ono M, Schneider S, Junghaenel DU, Stone AA. What affects the completion of ecological momentary assessments in chronic pain research? An individual patient data meta-analysis. *J Med Internet Res* 2019 Feb 05;21(2):e11398 [FREE Full text] [doi: [10.2196/11398](https://doi.org/10.2196/11398)] [Medline: [30720437](https://pubmed.ncbi.nlm.nih.gov/30720437/)]

48. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. *JAMA* 2018 Aug 21;320(7):639-640 [FREE Full text] [doi: [10.1001/jama.2018.6621](https://doi.org/10.1001/jama.2018.6621)] [Medline: [29852046](https://pubmed.ncbi.nlm.nih.gov/29852046/)]
49. Rothman AJ, Sheeran P, Wood W. Reflective and automatic processes in the initiation and maintenance of dietary change. *Ann Behav Med* 2009 Dec;38 Suppl 1:S4-17. [doi: [10.1007/s12160-009-9118-3](https://doi.org/10.1007/s12160-009-9118-3)] [Medline: [19787308](https://pubmed.ncbi.nlm.nih.gov/19787308/)]
50. van Berkel N, Goncalves J, Hosio S, Sarsenbayeva Z, Velloso E, Kostakos V. Overcoming compliance bias in self-report studies: a cross-study analysis. *Int J Human Comput Stud* 2020 Feb;134:1-12. [doi: [10.1016/j.ijhcs.2019.10.003](https://doi.org/10.1016/j.ijhcs.2019.10.003)]
51. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
52. Suh H, Shahriree N, Hekler E, Kientz J. Developing and validating the user burden scale: a tool for assessing user burden in computing systems. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 2016 Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7 - 12, 2016; San Jose, CA. [doi: [10.1145/2858036.2858448](https://doi.org/10.1145/2858036.2858448)]
53. Dunton GF, Dzibur E, Intille S. Feasibility and performance test of a real-time sensor-informed context-sensitive ecological momentary assessment to capture physical activity. *J Med Internet Res* 2016 Jun 01;18(6):e106 [FREE Full text] [doi: [10.2196/jmir.5398](https://doi.org/10.2196/jmir.5398)] [Medline: [27251313](https://pubmed.ncbi.nlm.nih.gov/27251313/)]
54. Nosakhare E, Picard R. Toward assessing and recommending combinations of behaviors for improving health and well-being. *ACM Trans Comput Healthcare* 2020 Mar 02;1(1):1-29. [doi: [10.1145/3368958](https://doi.org/10.1145/3368958)]
55. Piercy KL, Troiano RP. Physical activity guidelines for Americans from the US department of health and human services. *Cardiovascular Q Outcomes* 2018 Nov;11(11):e005263. [doi: [10.1161/circoutcomes.118.005263](https://doi.org/10.1161/circoutcomes.118.005263)]
56. Leventhal AM, Strong DR, Kirkpatrick MG, Unger JB, Sussman S, Riggs NR, et al. Association of electronic cigarette use with initiation of combustible tobacco product smoking in early adolescence. *JAMA* 2015 Aug 18;314(7):700-707 [FREE Full text] [doi: [10.1001/jama.2015.8950](https://doi.org/10.1001/jama.2015.8950)] [Medline: [26284721](https://pubmed.ncbi.nlm.nih.gov/26284721/)]
57. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. *Academic Med* 2012;87(1):66-73. [doi: [10.1097/acm.0b013e31823ab7d2](https://doi.org/10.1097/acm.0b013e31823ab7d2)]
58. Ram N, Brinberg M, Pincus AL, Conroy DE. The questionable ecological validity of ecological momentary assessment: considerations for design and analysis. *Res Hum Dev* 2017;14(3):253-270 [FREE Full text] [doi: [10.1080/15427609.2017.1340052](https://doi.org/10.1080/15427609.2017.1340052)] [Medline: [30613195](https://pubmed.ncbi.nlm.nih.gov/30613195/)]
59. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol* 2015 Dec;34S:1220-1228 [FREE Full text] [doi: [10.1037/hea0000305](https://doi.org/10.1037/hea0000305)] [Medline: [26651463](https://pubmed.ncbi.nlm.nih.gov/26651463/)]
60. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. *JMIR Mhealth Uhealth* 2016 Jan 27;4(1):e7 [FREE Full text] [doi: [10.2196/mhealth.4225](https://doi.org/10.2196/mhealth.4225)] [Medline: [26818775](https://pubmed.ncbi.nlm.nih.gov/26818775/)]
61. Farivar S, Abouzahra M, Ghasemaghaei M. Wearable device adoption among older adults: a mixed-methods study. *Int J Inf Manage* 2020 Dec;55:102209 [FREE Full text] [doi: [10.1016/j.ijinfomgt.2020.102209](https://doi.org/10.1016/j.ijinfomgt.2020.102209)] [Medline: [32834339](https://pubmed.ncbi.nlm.nih.gov/32834339/)]
62. Wu A, Munteanu C. Understanding older users' acceptance of wearable interfaces for sensor-based fall risk assessment. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018 Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; Apr 21 - 26, 2018; Montreal, QC, Canada. [doi: [10.1145/3173574.3173693](https://doi.org/10.1145/3173574.3173693)]
63. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. *JMIR Mhealth Uhealth* 2016 Nov 23;4(4):e129 [FREE Full text] [doi: [10.2196/mhealth.6540](https://doi.org/10.2196/mhealth.6540)] [Medline: [27881359](https://pubmed.ncbi.nlm.nih.gov/27881359/)]
64. Schaefer SE, Van Loan M, German JB. A feasibility study of wearable activity monitors for pre-adolescent school-age children. *Prev Chronic Dis* 2014 May 22;11:E85 [FREE Full text] [doi: [10.5888/pcd11.130262](https://doi.org/10.5888/pcd11.130262)] [Medline: [24854236](https://pubmed.ncbi.nlm.nih.gov/24854236/)]
65. Hedeker D, Nordgren R. MIXREGLS: a program for mixed-effects location scale analysis. *J Stat Softw* 2013 Mar;52(12):1-38 [FREE Full text] [doi: [10.18637/jss.v052.i12](https://doi.org/10.18637/jss.v052.i12)] [Medline: [23761062](https://pubmed.ncbi.nlm.nih.gov/23761062/)]
66. Dzibur E, Ponnada A, Nordgren R, Yang C, Intille S, Dunton G, et al. MixWILD: a program for examining the effects of variance and slope of time-varying variables in intensive longitudinal data. *Behav Res Methods* 2020 Aug;52(4):1403-1427 [FREE Full text] [doi: [10.3758/s13428-019-01322-1](https://doi.org/10.3758/s13428-019-01322-1)] [Medline: [31898295](https://pubmed.ncbi.nlm.nih.gov/31898295/)]
67. Xu S, Thompson W, Kerr J, Godbole S, Sears DD, Patterson R, et al. Modeling interrelationships between health behaviors in overweight breast cancer survivors: applying Bayesian networks. *PLoS One* 2018;13(9):e0202923 [FREE Full text] [doi: [10.1371/journal.pone.0202923](https://doi.org/10.1371/journal.pone.0202923)] [Medline: [30180192](https://pubmed.ncbi.nlm.nih.gov/30180192/)]
68. Chih M, Patton T, McTavish FM, Isham AJ, Judkins-Fisher CL, Atwood AK, et al. Predictive modeling of addiction lapses in a mobile health application. *J Subst Abuse Treat* 2014 Jan;46(1):29-35 [FREE Full text] [doi: [10.1016/j.jsat.2013.08.004](https://doi.org/10.1016/j.jsat.2013.08.004)] [Medline: [24035143](https://pubmed.ncbi.nlm.nih.gov/24035143/)]
69. Wright AG, Woods WC. Personalized models of psychopathology. *Annu Rev Clin Psychol* 2020 May 07;16:49-74. [doi: [10.1146/annurev-clinpsy-102419-125032](https://doi.org/10.1146/annurev-clinpsy-102419-125032)] [Medline: [32070120](https://pubmed.ncbi.nlm.nih.gov/32070120/)]

70. Beltz AM, Wright AG, Sprague BN, Molenaar PC. Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment* 2016 Aug;23(4):447-458 [FREE Full text] [doi: [10.1177/1073191116648209](https://doi.org/10.1177/1073191116648209)] [Medline: [27165092](https://pubmed.ncbi.nlm.nih.gov/27165092/)]
71. Molenaar PC. Granger causality testing with intensive longitudinal data. *Prev Sci* 2019 Apr;20(3):442-451. [doi: [10.1007/s11121-018-0919-0](https://doi.org/10.1007/s11121-018-0919-0)] [Medline: [29858760](https://pubmed.ncbi.nlm.nih.gov/29858760/)]

## Abbreviations

**AUC:** area under the curve  
**DND:** do not disturb  
**EMA:** ecological momentary assessment  
**ILD:** intensive longitudinal data  
**PA:** physical activity  
**REDCap:** Research Electronic Data Capture  
**SB:** sedentary behavior  
**TIME:** Temporal Influences on Movement & Exercise  
**UI:** user interface  
**μEMA:** microinteraction ecological momentary assessment

*Edited by A Mavragani; submitted 09.08.21; peer-reviewed by D Fulford, J Lai; comments to author 29.09.21; revised version received 24.11.21; accepted 17.12.21; published 09.02.22*

*Please cite as:*

*Ponnada A, Wang S, Chu D, Do B, Dunton G, Intille S*

*Intensive Longitudinal Data Collection Using Microinteraction Ecological Momentary Assessment: Pilot and Preliminary Results*

*JMIR Form Res* 2022;6(2):e32772

URL: <https://formative.jmir.org/2022/2/e32772>

doi: [10.2196/32772](https://doi.org/10.2196/32772)

PMID:

©Aditya Ponnada, Shirlene Wang, Daniel Chu, Bridgette Do, Genevieve Dunton, Stephen Intille. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 09.02.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.