

Original Paper

# Measuring Criterion Validity of Microinteraction Ecological Momentary Assessment (Micro-EMA): Exploratory Pilot Study With Physical Activity Measurement

Aditya Ponnada<sup>1</sup>, BDes; Binod Thapa-Chhetry<sup>1</sup>, MSc; Justin Manjourides<sup>2</sup>, PhD; Stephen Intille<sup>1</sup>, PhD

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

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

**Corresponding Author:**

Aditya Ponnada, BDes

Khoury College of Computer Sciences

Bouve College of Health Sciences

Northeastern University

360 Huntington Avenue

Boston, MA,

United States

Phone: 1 617 306 1610

Email: [ponnada.a@northeastern.edu](mailto:ponnada.a@northeastern.edu)

## Abstract

**Background:** Ecological momentary assessment (EMA) is an in situ method of gathering self-report on behaviors using mobile devices. In typical phone-based EMAs, participants are prompted repeatedly with multiple-choice questions, often causing participation burden. Alternatively, microinteraction EMA (micro-EMA or  $\mu$ EMA) is a type of EMA where all the self-report prompts are single-question surveys that can be answered using a 1-tap glanceable microinteraction conveniently on a smartwatch. Prior work suggests that  $\mu$ EMA may permit a substantially higher prompting rate than EMA, yielding higher response rates and lower participation burden. This is achieved by ensuring  $\mu$ EMA prompt questions are quick and cognitively simple to answer. However, the validity of participant responses from  $\mu$ EMA self-report has not yet been formally assessed.

**Objective:** In this pilot study, we explored the criterion validity of  $\mu$ EMA self-report on a smartwatch, using physical activity (PA) assessment as an example behavior of interest.

**Methods:** A total of 17 participants answered 72  $\mu$ EMA prompts each day for 1 week using a custom-built  $\mu$ EMA smartwatch app. At each prompt, they self-reported whether they were doing sedentary, light/standing, moderate/walking, or vigorous activities by tapping on the smartwatch screen. Responses were compared with a research-grade activity monitor worn on the dominant ankle simultaneously (and continuously) measuring PA.

**Results:** Participants had an 87.01% (5226/6006)  $\mu$ EMA completion rate and a 74.00% (5226/7062) compliance rate taking an average of only 5.4 (SD 1.5) seconds to answer a prompt. When comparing  $\mu$ EMA responses with the activity monitor, we observed significantly higher ( $P < .001$ ) momentary PA levels on the activity monitor when participants self-reported engaging in moderate+vigorous activities compared with sedentary or light/standing activities. The same comparison did not yield any significant differences in momentary PA levels as recorded by the activity monitor when the  $\mu$ EMA responses were randomly generated (ie, simulating careless taps on the smartwatch).

**Conclusions:** For PA measurement, high-frequency  $\mu$ EMA self-report could be used to capture information that appears consistent with that of a research-grade continuous sensor for sedentary, light, and moderate+vigorous activity, suggesting criterion validity. The preliminary results show that participants were not carelessly answering  $\mu$ EMA prompts by randomly tapping on the smartwatch but were reporting their true behavior at that moment. However, more research is needed to examine the criterion validity of  $\mu$ EMA when measuring vigorous activities.

(JMIR Mhealth Uhealth 2021;9(3):e23391) doi: [10.2196/23391](https://doi.org/10.2196/23391)

## KEYWORDS

ecological momentary assessment (EMA); experience sampling; physical activity; smartwatch; microinteractions; criterion validity; activity monitor;  $\mu$ EMA

## Introduction

Ecological momentary assessment (EMA), also known as the experience sampling method, is used to measure behaviors of people in natural settings [1]. In a typical EMA study, a user's phone is prompted multiple times a day (often 6+ times) with a set of multiple-choice questions measuring behaviors of interest [2,3]. The repeated EMA prompts, which typically ask about momentary behaviors or states, not only reduce recall biases present in retrospective surveys [3,4] but also capture temporal changes in health behaviors unique to each individual [5]. Because of these benefits, EMA is commonly used to measure behaviors in intensive longitudinal studies [6].

The drawback of EMA is participation burden [7-9]. Participants are first interrupted with a beep and/or vibration. They must then find the phone, unlock the device, and respond to a set of complex multiple-choice questions. This repeated effort, which can take tens of seconds for even the shortest surveys and several minutes for many common surveys, can be burdensome, negatively impacting study compliance [7,9,10]. Microinteraction EMA ( $\mu$ EMA or micro-EMA) is a type of EMA that may, for some behaviors (eg, chronic pain or fatigue), enable high frequency self-report data collection with low study burden [11]. In  $\mu$ EMA, rather than using complex multiquestion surveys, each prompt contains only a single question that can be answered with a glanceable microinteraction [12], typically just a tap on a smartwatch. Prior studies have shown that despite approximately 8 times more interruption than EMA,  $\mu$ EMA had a significantly higher response rate and lower perceived burden because all interactions are limited to microinteractions [11,13]. Thus, there is preliminary evidence that  $\mu$ EMA may enable gathering high-frequency self-report with manageable burden, a complementary approach to EMA. Recently,  $\mu$ EMA has been used to gather data on stress [14], hyperarousal [15], and perceived comfort [16], and it has also been used with small pervasive displays [17].

Prior work on  $\mu$ EMA, however, has assumed validity of  $\mu$ EMA responses and not demonstrated it. Because  $\mu$ EMA is designed to gather small amounts of information with each prompt (but with higher frequency), the prompts are both limited to a single question and made cognitively simple to answer with a quick microinteraction (taking only 3-5 seconds). To achieve cognitive simplicity and fit questions on a smartwatch so they can be answered in a single tap without scrolling requires a limited answer set. This calls into question whether  $\mu$ EMA responses could capture behavior similarly to a gold-standard instrument (ie, criterion validity [18]), and validating such a high-frequency self-report requires an instrument that can measure the same behavior continuously in free living, such as a wearable sensor. In some domains where  $\mu$ EMA may be especially useful (eg, chronic pain), such sensors do not yet exist. The purpose of this pilot study is to explore the criterion validity of  $\mu$ EMA self-report, and thus, we used the example of physical activity (PA) measurement, because PA can be estimated continuously using research-grade activity monitors [19].

## Methods

In this pilot study, we compared  $\mu$ EMA self-report on a smartwatch with acceleration data collected using a wearable activity monitor on the dominant ankle to assess criterion validity, such as whether participants are answering the  $\mu$ EMA questions meaningfully in a way that changes as PA changes.

### $\mu$ EMA App

We implemented a  $\mu$ EMA app on an Android Wear OS 2.0 (Figure 1) to measure PA. PA was chosen because (1) it can be estimated continuously using a passive, easy-to-wear sensor (eg, an accelerometer on the ankle) and (2) PA can change frequently within a day, making it suitable for testing a high-frequency  $\mu$ EMA self-report system. Participants were presented with 4 activity intensity options with each  $\mu$ EMA prompt: sedentary, light/standing, moderate/walking, and vigorous.

**Figure 1.** (Left)  $\mu$ EMA interface on a smartwatch with four activity intensity options. (Right) Undo screen to change response, available for 10 s, with a countdown timer.



The  $\mu$ EMA app prompted 6 times an hour between 8 AM to 8 PM (72 expected prompts per day) using vibration on the smartwatch. The question displayed at the start of the vibration consisted of the 4 activity intensity categories, and participants selected the activity intensity they were engaged in at that moment. If the participant did not respond to the prompt within 2 minutes by tapping on a category, the prompt disappeared from the screen and the app recorded a missed response. When a response was selected, the watch displayed an Undo? screen (Figure 1). If participants tapped on the Undo? button, they were returned to the  $\mu$ EMA question to change their responses, otherwise the Undo? screen disappeared after 10 seconds. Data from the watch were sent to the participant's smartphone once per hour. Participants interacted only with the watch, and the phone collected, encrypted, and transferred data from the watch to a remote server.

### Study Design

We conducted a week-long, within-subject pilot study (approved by Northeastern University's institutional review board; project number 14-10-01) to compare  $\mu$ EMA PA responses with a wearable sensor.

### Participant Recruitment

Participants were eligible if they owned a compatible Android smartphone with version 4.3+, were aged 18 to 55 years, were a student or staff at our university (to ensure we could safely recover loaned smartwatches and activity monitors at the end of the study), and were willing to wear a sensor and the

smartwatch for 1 week. The study was advertised using flyers posted on the university campus and also by sending electronic notices to common university announcement portals. Of the 35 people who responded to study advertisements, 20 were eligible to participate based on screening via phone call. Among those, 3 participants dropped out early. One had a wake-period outside of  $\mu$ EMA prompting hours, another had a job that physically made it difficult to answer prompts on the watch, and the third had a malfunctioning phone. This left 17 active participants in the pilot study (11 males and 6 females; aged 19 to 34 years). None of the participants were affiliated to our research group.

### Measurement Tasks

Participants were asked to complete 2 tasks simultaneously between 8 AM to 8 PM every day for one week: answer  $\mu$ EMA prompts on the loaned smartwatch (model Urbane, LG Electronics) and wear an activity monitor on their dominant ankle. Participants were asked to ignore  $\mu$ EMA prompts in unsafe conditions (eg, driving) and charge the watch nightly so that it could be worn next morning.

We used GT9X monitors (35×35 mm, 14 g; ActiGraph LLC) to measure acceleration continuously at 80 Hz [20]. Participants wore the sensor on the dominant ankle above the medial malleolus using an elastic band (Figure 2). The ankle location was chosen because ankle acceleration can reliably capture ambulation activities, more so than the wrist or hip [21,22]. The sensor collected raw acceleration passively; other than wearing it, the participants did not interact with, or charge, this device.

**Figure 2.** GT9X activity monitor worn on the dominant ankle.

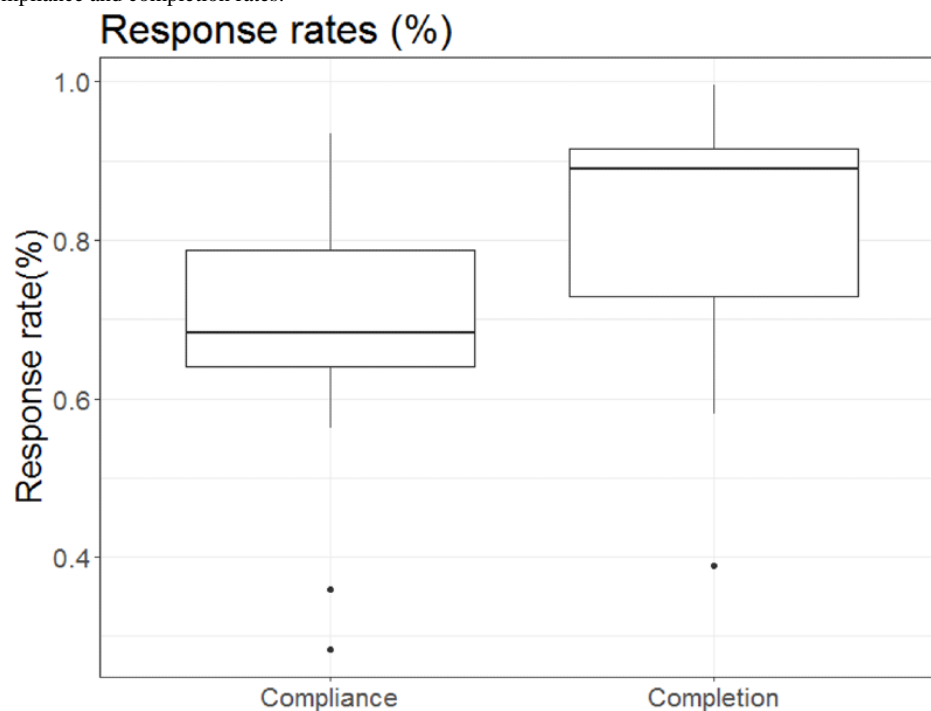
### Procedures

Day 0: Researchers met with participants, obtained informed consent, and loaned the activity monitor and smartwatch with the  $\mu$ EMA app. Research staff then presented participants with some examples of the types of different activities that would fall into each of the 4 target activity categories (Multimedia Appendix 1). Days 1-7: Participants wore the activity monitor on the dominant ankle and answered 6  $\mu$ EMA prompts per hour between 8 AM to 8 PM on the smartwatch for 1 week. Day 8: Researchers recovered the monitor. Participants were not compensated financially to ensure that we measure criterion validity without any external motivation, but participants could use the smartwatch for 4 more days as their personal device for fun if they desired. All participants agreed to use the device for 4 more days.

### Results

We computed participants' (1) compliance rate, the percentage of  $\mu$ EMA prompts answered out of all the scheduled prompts (ie, including when the watch was off); (2) completion rate, the percentage of  $\mu$ EMA prompts answered out of delivered prompts (ie, excluding when the watch was off); and (3) response time, the time taken to answer a prompt, measured from the start of the prompt vibration.

Two participants had low compliance (Figure 3). From the debriefing, we learned that these participants did not charge the smartwatch regularly, receiving fewer scheduled prompts. Their compliance fell below the 1.5 interquartile range (<40%); therefore, they were considered outliers and were excluded from the main analysis of data from the remaining 15 participants (Table 1) [23]. Implications of dropping the outliers is discussed later in our results.

**Figure 3.**  $\mu$ EMA compliance and completion rates.

**Table 1.** Response behavior of pilot study participants.

Characteristics	With outliers (n=17)	Without outliers (n=15)
Expected prompts, n	7591	7062
Delivered prompts, n	6387	6006
Answered prompts, n	5238	5226
Compliance rate (%)	69.00	74.00
Completion rate (%)	82.01	87.01
Response time (s), mean (SD)	5.5 (1.6)	5.4 (1.5)

## Data Preparation

### Computing Activity Counts

ActiGraph activity counts are widely used motion summary metrics computed from raw acceleration for a specified epoch [24]. Activity counts have been used in prior work to compare EMA responses and accelerometer data [25,26]. We first computed activity counts for 1 second epochs of the raw data from the ankle-worn activity monitor. Using this, we calculated the total activity counts 60 seconds prior to the  $\mu$ EMA prompt as the PA level measured using the activity monitor.

### Removing Sensor Nonwear Data

We removed the instances of raw data when the participants were not wearing the activity monitor. Following Choi et al [27], 90+ minutes of continuous zero-valued activity counts computed for the 1 second epochs were considered sensor nonwear times. All  $\mu$ EMA responses recorded during these nonwear times were dropped. This eliminated only 1.3% (70/5226) of the total responses from 15 participants leaving 5156 valid responses with sensor wear. Of these, 25% (18/70) of the responses were from just 1 participant.

### $\mu$ EMA Response Distribution

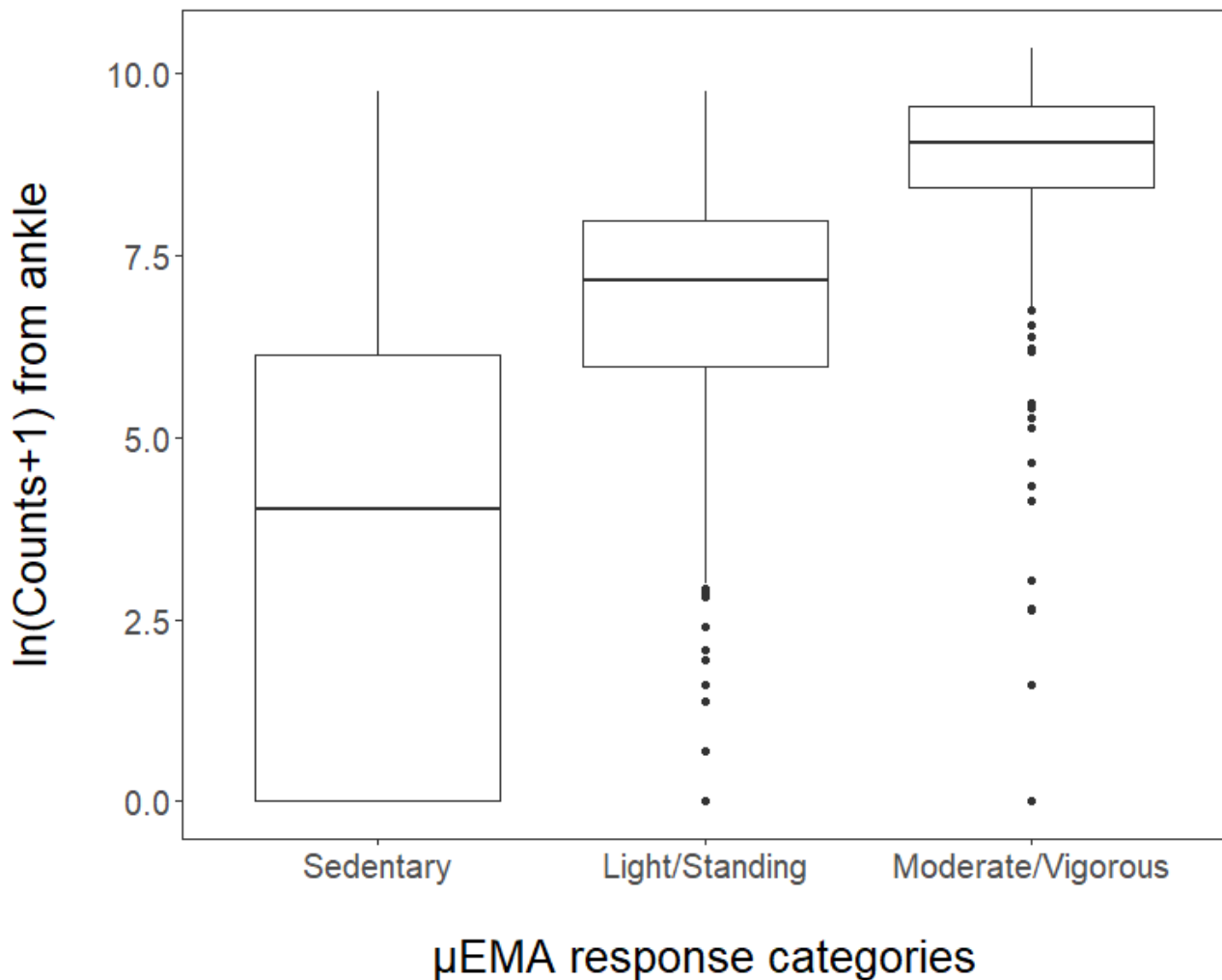
We received more sedentary responses (3619/5156, 69.99%) than light/standing (978/5156, 18.97%) and moderate/walking (544/5156, 10.56%) from  $\mu$ EMA, which is consistent with general physical (in)activity trends [28,29]. However, we received few vigorous responses (15/5156, 0.29%). Thus, we combined the moderate/walking and vigorous categories into a single category of moderate+vigorous. Hence, we compare the 3 PA intensities (sedentary, light/standing, and moderate+vigorous) from  $\mu$ EMA with activity count (60 seconds prior to the prompt) from the ankle-worn activity monitor (Table 2). These activity counts were within the ranges recorded previously in young adults for sedentary, light, and moderate+vigorous activities using ankle-worn accelerometers [30-33].

Activity counts computed 60 seconds before the prompt ranged from 0 (for sedentary) to >15K (for moderate/vigorous), and resulted in a right-skewed distribution. Thus, we log-transformed these activity counts into  $\ln(\text{Counts} + 1)$ , where 1 is the smallest nonzero count recorded in this pilot study [34]. Figure 4 presents the final distribution of these log-transformed counts corresponding to  $\mu$ EMA categories (sedentary, light/standing, and moderate+vigorous).

**Table 2.**  $\ln(\text{Counts} + 1)$  measured on ankle for each  $\mu$ EMA category.

Category	Value
<b>Sedentary</b>	
Mean (SD)	3.56 (2.94)
Median (IQR <sup>a</sup> )	4.01 (6.13)
<b>Light/standing</b>	
Mean (SD)	6.66 (2.10)
Median (IQR)	7.16 (2.01)
<b>Moderate+vigorous</b>	
Mean (SD)	8.72 (1.52)
Median (IQR)	9.04 (1.11)

<sup>a</sup>IQR: interquartile range.

**Figure 4.**  $\mu$ EMA responses versus  $\ln(\text{Counts} + 1)$ .

### Criterion Validity of $\mu$ EMA on Smartwatch

We applied a linear mixed-effects model with a random intercept (using the lme4 package [35]):

$$\ln(\text{Counts}_{ij} + 1) = \beta_0 + \beta_1 (\mu\text{EMA}_{ij}) + u_i$$

Here,  $\text{Counts}_{ij}$  is the activity count from the ankle-worn monitor measured for an individual  $i$  computed for 60 seconds before  $\mu$ EMA prompt  $j$ ,  $\beta_0$  is the fixed-effect intercept,  $\mu\text{EMA}_{ij}$  is the ordinal self-report (0 = sedentary, 1 = light/standing, and 2 = moderate/vigorous) on the smartwatch by participant  $i$  at prompt  $j$ , and  $u_i$  is the random-intercept for the participant  $i$ . Although we were interested in the fixed-effects part of the model, the random intercept is included to account for repeated measures within participants. The momentary PA levels measured on the activity monitor,  $\ln(\text{Counts}_{ij} + 1)$ , were significantly different ( $P < .001$ ) for sedentary, light/standing, and moderate+vigorous activity categories captured using  $\mu$ EMA self-report (ie,  $\mu\text{EMA}_{ij}$ ). The final model fit for this pilot study was:

$$\ln(\text{Counts}_{ij} + 1) = 6.28 + 3.57 (\mu\text{EMA}_{ij}) + u_i$$

We then included data from the two outliers to explore the sensitivity of the model fit and the activity levels,  $\ln(\text{Counts}_{ij} + 1)$  were significantly different ( $P < .001$ ) for the 3  $\mu$ EMA

response categories ( $\mu\text{EMA}_{ij}$ ). However, we observed that these participants showed continuous sensor nonwear for 2 to 3 days at once in addition to not charging the smartwatch regularly as required by the study protocol. In fact, removing the sensor nonwear for these participants eliminated 27.78% (30/108) and 43.04% (65/151) of the answered  $\mu$ EMA prompts, respectively. As a result, these participants were excluded from our final model fit. For the remaining 15 participants, the pair-wise, post hoc comparison of the 3  $\mu$ EMA response categories with Tukey adjustment revealed that  $\ln(\text{Counts}_{ij} + 1)$  for sedentary responses were significantly less than light/standing, which were significantly less than the moderate+vigorous ( $P < .001$ ), the expected order of activity intensity (Table 3). In other words, when participants self-reported (using  $\mu$ EMA) being in moderate+vigorous instead of light/standing or sedentary activities, they were also more likely to have passively measured higher PA levels in those moments.

For exploratory purposes, we simulated random  $\mu$ EMA responses for each participant; instead of analyzing their actual responses, we compared the random data as if participants were randomly tapping on the watch screen only to dismiss the prompt with motion data recorded from the ankle. The model fit did not yield any significant differences in the counts for different  $\mu$ EMA categories. This further suggests that

participants were answering  $\mu$ EMA questions carefully, not randomly (or carelessly), despite the intensive sampling rate.

**Table 3.** Pairwise comparison of  $\ln(\text{Counts} + 1)$  from activity monitor with  $\mu$ EMA responses.

$\mu$ EMA <sup>a</sup> response (i)	$\mu$ EMA response (j)	Mean difference (i-j)	SE	95% CI
Sedentary	Light/standing	-3.12	0.28	-3.34 to -2.90
Light/standing	Moderate+vigorous	-1.93	0.13	-2.24 to -1.16
Moderate+vigorous	Sedentary	5.05	0.11	4.78 to 5.32

<sup>a</sup> $\mu$ EMA: microinteraction ecological momentary assessment.

## Discussion

### Principal Findings

In this pilot study, we explored criterion validity of  $\mu$ EMA on a smartwatch by measuring PA intensity of free-living individuals. For 1 week, 15 participants answered 6  $\mu$ EMA prompts per hour (between 8 AM to 8 PM each day) reporting their PA intensity while simultaneously wearing an ankle-worn research-grade activity monitor. The activity monitor gathered raw accelerometer data continuously at 80 Hz. When comparing the  $\mu$ EMA self-report and activity monitor, we observed significantly higher momentary PA levels on the activity monitor when participants self-reported engaging in moderate+vigorous activities compared with when they reported sedentary or light/standing activities. Similarly, we observed significantly higher PA levels on the activity monitor when participants self-reported engaging in light/standing activities compared with when they reported being sedentary. Thus, we observed the expected order of intensity of these activity categories, suggesting criterion validity of  $\mu$ EMA self-report. Holistically, this result is aligned with the prior work where PA researchers have compared phone-based EMA with body-worn research-grade activity monitors measuring raw acceleration and found EMA to measure PA similar to the objective sensor [25,36,37]. However, the frequency of EMA self-report in these studies varied from once a day [25] to once an hour [37] to minimize interruption burden. In our pilot study, we extend these findings to an alternate self-report approach ( $\mu$ EMA) that allowed for 6 times more temporal density in measurement, yielding high response rate but without burdening participants as much as the phone based EMA.

Overall, this preliminary result shows that when measuring PA, participants'  $\mu$ EMA self-report (at a high temporal density) could capture the PA levels consistent with a continuous high-frequency sensor. It appears, for example, that participants answered  $\mu$ EMA prompts meaningfully—not just tapping on an answer to dismiss the prompt—and sustained this answering at a rate of 6 times per hour for 12 hours per day for an entire week (approximately 504 prompts per person). Our preliminary findings suggest that the  $\mu$ EMA prompts achieve the goal of being so easy to answer that they can be sustained, instead of being ignored or dismissed—all while recording the true behavior at that moment. Participants know that every  $\mu$ EMA prompt just requires a single, 1-tap response on the easily accessible smartwatch that can be completed in a

microinteraction; this may contribute to high compliance and valid data entry. EMA protocols, alternatively, often require answering multiple, sometimes complex, questions that can be time consuming and feel burdensome. In fact, the effort needed to dismiss an  $\mu$ EMA prompt is roughly equivalent to the effort required to answer the prompt, thus encouraging survey completion.

### Limitations

This exploratory pilot study provides preliminary findings on criterion validity of  $\mu$ EMA, but more research is needed. We merged the vigorous and moderate/walking activities into moderate+vigorous because we received only 15 responses of vigorous activities. One reason may be that most individuals engage in significantly more sedentary than vigorous activities [38]. Another may be that during the exit debriefing, 2 participants reported that when they engaged in vigorous activities like outdoor cycling, they could not respond to prompts within 2 minutes, thus resulting in fewer vigorous activity responses. This type of response behavior during vigorous activities has also been observed in a prior EMA versus activity monitor validation study where more missing responses were found during vigorous activity [37]. Nevertheless, this potential for a bias when reporting vigorous activity using  $\mu$ EMA should be explored in future work. If a reporting bias for vigorous activity is observed in future studies, one remedy could be to explore sensor-triggered  $\mu$ EMA, where the  $\mu$ EMA prompts might be presented based on real-time processing of PA and then delivered not during vigorous PA but rather right after it is confidently estimated to have been completed [39]. This also highlights the need to rethink question wording, where instead of asking about behavior in the moment (eg, Doing vigorous PA now?),  $\mu$ EMA will have to ask about the recent past (eg, Vigorous PA 2 min ago?) but without compromising the cognitive simplicity required for a microinteraction. Notably, EMA delivered via a smartphone would have likely been even more difficult to respond to during vigorous PA due to the difficulty of accessing and interacting with the phone device (versus the comparative simplicity of completing a microinteraction on the smartwatch) [40]. Future studies with larger sample sizes including individuals recruited specifically because they are known to regularly engage in vigorous activities could provide more insights on  $\mu$ EMA validity when measuring vigorous activity.

Being an exploratory pilot study, we had a small sample size. We were limited by available equipment, and recruiting was

more challenging than in typical studies because we did not offer financial compensation for study participation (contrary to most phone EMA studies); our intent was to measure  $\mu$ EMA compliance and data quality in a situation without any external monetary incentives because such incentives may not be viable for future longitudinal measurement or intervention studies that might use  $\mu$ EMA.

Because validity assessments are domain-dependent, our findings with PA do not necessarily generalize to other behaviors. PA was chosen in this work because we have research-grade activity monitors to compare  $\mu$ EMA responses against at a high temporal density, not necessarily because PA intensity is best measured with  $\mu$ EMA if passive sensors are available. However, in some domains where  $\mu$ EMA might be useful (eg, chronic pain), passive sensors that continuously monitor the behavior are not available yet, and methods such as direct or physiological observations that require laboratory conditions are not practical for multiday free-living studies. Nevertheless, validation studies generally rely on imperfect comparisons, and so confidence in validity of  $\mu$ EMA, just like EMA, will require multiple cross-domain and longitudinal experiments from different research teams, of which this may be the first of many.

## Conclusion

The  $\mu$ EMA implemented on a smartwatch addresses the common Achilles' heel of traditional phone-based EMA, the participation

burden of accessing, unlocking, and answering multiple multichoice questions on a smartphone. Despite significantly higher interruption rates than phone-based EMA,  $\mu$ EMA is able to yield significantly higher response rates with manageable participation burden. This is achieved by keeping the single questions in  $\mu$ EMA cognitively simple to answer allowing for high frequency of prompting (like a continuous sensor). However, this makes  $\mu$ EMA vulnerable to careless tapping on the smartwatch to dismiss the prompts, potentially compromising the validity of self-report responses. Thus, in this pilot study, we explored the criterion validity of  $\mu$ EMA self-report, comparing it with a continuous sensor to assess if participants submit their responses based on their true behavior in the moment. We used PA as an example domain because PA can be measured using a gold-standard sensor (research-grade accelerometers). We conducted a 1-week exploratory pilot study with 15 participants answering 72  $\mu$ EMA prompts each day while measuring continuous PA using an ankle-worn accelerometer. We found that participants were able to correctly report their sedentary, light/standing, and moderate+vigorous activities. This highlights that participants were not carelessly tapping on the smartwatch only to dismiss the prompt but were providing accurate information about their behavior comparable to the continuous sensor data from the ankle suggesting criterion validity. However, more research is needed to explore criterion validity of  $\mu$ EMA in other behavioral domains of interest including vigorous PA measurement.

## Acknowledgments

AP led development of the  $\mu$ EMA data collection app and experiment design and wrote the first draft of the manuscript. BTC assisted with raw accelerometer data processing from the activity monitors. JM provided guidance on statistical analysis and interpreting the results. SI was the principal investigator for the project responsible for securing research funding and working closely with AP on project design and manuscript development. We thank Krystal Huey for assisting in participant recruitment and data collection and our participants for their time. Research reported in this paper was supported, in part, by a Google Glass Award and by the National Heart, Lung, and Blood Institute of the National Institutes of Health under award number R21HL108018. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Onboarding instructions.

[\[DOCX File , 14 KB-Multimedia Appendix 1\]](#)

## References

1. Stone AA, Shiffman S. Ecological momentary assessment: measuring real world processes in behavioral medicine. *Ann Behav Med* 1994;16:199-202. [Medline: [28298264](#)]
2. Smyth J, Smyth JM. Ecological Momentary Assessment research in behavioral medicine. *J Happiness Studies* 2003;4(1):35-52 [FREE Full text] [doi: [10.1023/a:1023657221954](#)]
3. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol* 2008;4:1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](#)] [Medline: [18509902](#)]
4. Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. *Psychol Assess* 2009 Dec;21(4):486-497 [FREE Full text] [doi: [10.1037/a0017074](#)] [Medline: [19947783](#)]
5. Chandra S, Scharf D, Shiffman S. Within-day temporal patterns of smoking, withdrawal symptoms, and craving. *Drug Alcohol Depend* 2011 Sep 01;117(2-3):118-125 [FREE Full text] [doi: [10.1016/j.drugalcdep.2010.12.027](#)] [Medline: [21324611](#)]



6. van Berkel N, Ferreira D, Kostakos V. The experience sampling method on mobile devices. *ACM Comput. Surv* 2018 Jan 12;50(6):1-40 [FREE Full text] [doi: [10.1145/3123988](https://doi.org/10.1145/3123988)]
7. Collins RL, Kashdan TB, Gollnisch G. The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. *Exp Clin Psychopharmacol* 2003 Feb;11(1):73-78. [doi: [10.1037//1064-1297.11.1.73](https://doi.org/10.1037//1064-1297.11.1.73)] [Medline: [12622345](https://pubmed.ncbi.nlm.nih.gov/12622345/)]
8. 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/)]
9. 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 [FREE Full text] [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/)]
10. 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/)]
11. 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/)]
12. Ashbrook D. Enabling Mobile Microinteractions [Dissertation]. Atlanta: Georgia Institute of Technology; 2010.
13. 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) [FREE Full text] [doi: [10.1145/3130957](https://doi.org/10.1145/3130957)] [Medline: [30198012](https://pubmed.ncbi.nlm.nih.gov/30198012/)]
14. 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) [FREE Full text] [doi: [10.1145/3351249](https://doi.org/10.1145/3351249)] [Medline: [32432212](https://pubmed.ncbi.nlm.nih.gov/32432212/)]
15. Larsen J, Eskelund K, Christiansen T. Active self-tracking of subjective experience with a one-button wearable: a case study in military PTSD. 2017 Presented at: International Conference on Human Factors in Computing Systems; 2017; Denver URL: <https://arxiv.org/abs/1703.03437>
16. 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 [FREE Full text] [doi: [10.3390/buildings10100174](https://doi.org/10.3390/buildings10100174)]
17. 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 [FREE Full text] [doi: [10.1145/3264942](https://doi.org/10.1145/3264942)]
18. Cronbach LJ, Meehl PE. Construct validity in psychological tests. *Psychol Bull* 1955 Jul;52(4):281-302. [Medline: [13245896](https://pubmed.ncbi.nlm.nih.gov/13245896/)]
19. Meijer GA, Westerterp KR, Verhoeven FM, Koper HB, ten Hoor F. Methods to assess physical activity with special reference to motion sensors and accelerometers. *IEEE Trans Biomed Eng* 1991 Mar;38(3):221-229. [doi: [10.1109/10.133202](https://doi.org/10.1109/10.133202)] [Medline: [2066134](https://pubmed.ncbi.nlm.nih.gov/2066134/)]
20. Miller J. Accelerometer technologies, specifications, and limitations: a presentation by ActiGraph at ICAMPAM 2013. Actigraph, LLC. 2013 Jun 17. URL: <https://actigraphcorp.com/wp-content/uploads/2019/12/ActiGraph-ICAMPAM-2013.pdf> [accessed 2021-03-02]
21. LaMunion S, Bassett D, Toth L, Crouter S. The effect of body placement site on ActiGraph wGT3X-BT activity counts. *Biomed. Phys. Eng. Express* 2017 Jun 23;3(3):035026 [FREE Full text] [doi: [10.1088/2057-1976/aa777c](https://doi.org/10.1088/2057-1976/aa777c)]
22. Mannini A, Sabatini A, Intille S. Accelerometry-based recognition of the placement sites of a wearable sensor. *Pervasive Mob Comput* 2015 Aug 01;21:62-74 [FREE Full text] [doi: [10.1016/j.pmcj.2015.06.003](https://doi.org/10.1016/j.pmcj.2015.06.003)] [Medline: [26213528](https://pubmed.ncbi.nlm.nih.gov/26213528/)]
23. Pagano M, Gauvreau K. Principles of Biostatistics. London: Chapman and Hall; Nov 2018.
24. What are counts? Actigraph. 2018 Nov 08. URL: <https://actigraphcorp.force.com/support/s/article/What-are-counts> [accessed 2021-03-02]
25. Knell G, Gabriel KP, Businelle MS, Shuval K, Wetter DW, Kendzor DE. Ecological momentary assessment of physical activity: validation study. *J Med Internet Res* 2017 Jul 18;19(7):e253 [FREE Full text] [doi: [10.2196/jmir.7602](https://doi.org/10.2196/jmir.7602)] [Medline: [28720556](https://pubmed.ncbi.nlm.nih.gov/28720556/)]
26. Bruening M, van Woerden I, Todd M, Brennhofers S, Laska MN, Dunton G. A mobile ecological momentary assessment tool (devilSPARC) for nutrition and physical activity behaviors in college students: a validation study. *J Med Internet Res* 2016 Jul 27;18(7):e209 [FREE Full text] [doi: [10.2196/jmir.5969](https://doi.org/10.2196/jmir.5969)] [Medline: [27465701](https://pubmed.ncbi.nlm.nih.gov/27465701/)]
27. Choi L, Ward SC, Schnelle JF, Buchowski MS. Assessment of wear/nonwear time classification algorithms for triaxial accelerometer. *Med Sci Sports Exerc* 2012 Oct;44(10):2009-2016 [FREE Full text] [doi: [10.1249/MSS.0b013e318258cb36](https://doi.org/10.1249/MSS.0b013e318258cb36)] [Medline: [22525772](https://pubmed.ncbi.nlm.nih.gov/22525772/)]
28. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, et al. Amount of time spent in sedentary behaviors in the United States, 2003-2004. *Am J Epidemiol* 2008 Apr 1;167(7):875-881 [FREE Full text] [doi: [10.1093/aje/kwm390](https://doi.org/10.1093/aje/kwm390)] [Medline: [18303006](https://pubmed.ncbi.nlm.nih.gov/18303006/)]

29. Sisson SB, Camhi SM, Church TS, Martin CK, Tudor-Locke C, Bouchard C, et al. Leisure time sedentary behavior, occupational/domestic physical activity, and metabolic syndrome in U.S. men and women. *Metab Syndr Relat Disord* 2009 Dec;7(6):529-536 [FREE Full text] [doi: [10.1089/met.2009.0023](https://doi.org/10.1089/met.2009.0023)] [Medline: [19900152](https://pubmed.ncbi.nlm.nih.gov/19900152/)]
30. Ozemek C, Kirschner MM, Wilkerson BS, Byun W, Kaminsky LA. Intermonitor reliability of the GT3X+ accelerometer at hip, wrist and ankle sites during activities of daily living. *Physiol Meas* 2014 Feb;35(2):129-138. [doi: [10.1088/0967-3334/35/2/129](https://doi.org/10.1088/0967-3334/35/2/129)] [Medline: [24399138](https://pubmed.ncbi.nlm.nih.gov/24399138/)]
31. Rawson ES, Walsh TM. Estimation of resistance exercise energy expenditure using accelerometry. *Med Sci Sports Exerc* 2010 Mar;42(3):622-628. [doi: [10.1249/MSS.0b013e3181b64ef3](https://doi.org/10.1249/MSS.0b013e3181b64ef3)] [Medline: [19952824](https://pubmed.ncbi.nlm.nih.gov/19952824/)]
32. Stec MJ, Rawson ES. Estimation of resistance exercise energy expenditure using triaxial accelerometry. *J Strength Cond Res* 2012 May;26(5):1413-1422. [doi: [10.1519/JSC.0b013e318248d7b4](https://doi.org/10.1519/JSC.0b013e318248d7b4)] [Medline: [22223228](https://pubmed.ncbi.nlm.nih.gov/22223228/)]
33. Rhudy MB, Dreisbach SB, Moran MD, Ruggiero MJ, Veerabhadrapa P. Cut points of the Actigraph GT9X for moderate and vigorous intensity physical activity at four different wear locations. *J Sports Sci* 2020 Mar;38(5):503-510. [doi: [10.1080/02640414.2019.1707956](https://doi.org/10.1080/02640414.2019.1707956)] [Medline: [31865845](https://pubmed.ncbi.nlm.nih.gov/31865845/)]
34. McCune B, Urban DL. Data transformations. In: McCune B, Grace JB, Urban DL, editors. *Analysis of Ecological Communities*. Glendened Beach: MjM Software Design; 2002.
35. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
36. Dunton GF, Whalen CK, Jamner LD, Henker B, Floro JN. Using ecologic momentary assessment to measure physical activity during adolescence. *Am J Prev Med* 2005 Nov;29(4):281-287. [doi: [10.1016/j.amepre.2005.07.020](https://doi.org/10.1016/j.amepre.2005.07.020)] [Medline: [16242590](https://pubmed.ncbi.nlm.nih.gov/16242590/)]
37. Dunton GF, Liao Y, Kawabata K, Intille S. Momentary assessment of adults' physical activity and sedentary behavior: feasibility and validity. *Front Psychol* 2012;3:260 [FREE Full text] [doi: [10.3389/fpsyg.2012.00260](https://doi.org/10.3389/fpsyg.2012.00260)] [Medline: [22866046](https://pubmed.ncbi.nlm.nih.gov/22866046/)]
38. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. *Am J Prev Med* 2011 Apr;40(4):454-461. [doi: [10.1016/j.amepre.2010.12.016](https://doi.org/10.1016/j.amepre.2010.12.016)] [Medline: [21406280](https://pubmed.ncbi.nlm.nih.gov/21406280/)]
39. 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/)]
40. Ashbrook D, Clawson J, Lyons K, Starner T, Patel N. Quickdraw: the impact of mobility and on-body placement on device access time. 2008 Presented at: Proc SIGCHI Conf Human Factors Comput Syst; 2008; Florence URL: <https://doi.org/10.1145/1357054.1357092> [doi: [10.1145/1357054.1357092](https://doi.org/10.1145/1357054.1357092)]

## Abbreviations

**EMA:** ecological momentary assessment

**PA:** physical activity

**μEMA:** microinteraction ecological momentary assessment

*Edited by L Buis; submitted 11.08.20; peer-reviewed by M Newman, A Miller; comments to author 06.10.20; revised version received 01.12.20; accepted 22.02.21; published 10.03.21*

*Please cite as:*

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;9(3):e23391

URL: <https://mhealth.jmir.org/2021/3/e23391>

doi: [10.2196/23391](https://doi.org/10.2196/23391)

PMID: [33688843](https://pubmed.ncbi.nlm.nih.gov/33688843/)

©Aditya Ponnada, Binod Thapa-Chhetry, Justin Manjourides, Stephen Intille. Originally published in JMIR mHealth and uHealth (<http://mhealth.jmir.org>), 10.03.2021. 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 mHealth and uHealth, is properly cited. The complete bibliographic information, a link to the original publication on <http://mhealth.jmir.org/>, as well as this copyright and license information must be included.