Detecting Sleep Disruptions in Adolescents Using Context-sensitive Ecological Momentary Assessment: A Feasibility Study

Rithika Lakshminarayanan[1], Arushi Uppal[1], Ha Le[1], James C. Spilsbury[2] and Stephen Intille[1]

[1] Northeastern University, Boston MA 02115, USA
[2] Case Western Reserve University, Cleveland OH 10900, USA
lakshminarayanan.r@northeastern.edu

Abstract. Adolescents are recommended to sleep at least 8-10 hours per day. Inadequate sleep in adolescents is detrimental to their overall wellbeing and is linked to poor academic performance. Identifying causes of poor sleep in the sleep environment can help researchers and adolescents determine what changes need to be made to improve sleep quality. However, in-situ sleep monitoring is challenging because measurements cannot interfere with sleep, and people are poor at remembering what happens during the night. We report on the feasibility testing of an in-situ sleep monitoring application that uses passive sensing to drive context-sensitive ecological momentary assessments (EMAs) to help participants recall sleep disruptions when they wake up in the morning. Participants answered over 80% of EMAs delivered during the feasibility study and could recall meaningful reasons for over 40% of noise and motion events when they answered context-sensitive questions presented in the morning EMA. We discuss some challenges and future opportunities in sleep disruption detection.

Keywords: sleep disruptions, adolescents, mobile health, EMA, experience sampling, smartphone, smartwatch

1 Introduction

According to recommendations by the American Academy of Sleep Medicine, adolescents aged 13-18 should sleep at least 8-10 hours per day [22]. Adolescents who do not get adequate sleep have a higher risk of obesity, diabetes, injuries, poor mental health, and problems with attention and behavior [21]. Although there has been work on assessing sleep duration and sleep quality in both adults and adolescents in clinical settings with controlled environments [6, 17], there has been limited research on identifying and understanding the causes of sleep disruptions in-situ. In-situ sleep monitoring is challenging because any measurement used should not interfere with the individual's sleep, and people are poor at remembering what happened during wake events during the night [7]. One way to unobtrusively measure sleep disruptions is to collect self-reported information to assess sleep quality [8]. However, self-report may introduce response burden and recall bias [16, 29]. We developed a system designed to advance understanding of in-situ sleep disruptions by delivering ecological momentary assessments (EMAs) that respond to automatically detected events that may result from sleep disruption. The timing of administering EMAs to understand sleep behaviors is important because the longer an individual is awake, the less the person will remember about intermittent wake events that occurred during the night [26]. The goal of the research tool prototype we test in this study is to help scientists understand the relationship between sleep disruption events and sleep quality.

In-situ sleep monitoring allows for the examination of contextual factors that influence adolescent sleep, including environmental stimuli, social interactions, and technological use. For instance, research has shown that exposure to computer or smartphone screens before bedtime can disrupt sleep patterns and contribute to sleep deficiency in adolescents [14]. By integrating environmental and behavioral data with sleep metrics, in-situ monitoring may offer researchers valuable insights into the multifaceted nature of sleep regulation in adolescents. Furthermore, automated in-situ sleep monitoring may facilitate future systems that provide the early detection of sleep disorders and related health issues in adolescents. Adolescents are vulnerable to sleep disorders such as insomnia, sleep apnea, and delayed sleep phase syndrome, which can impair daytime functioning and lead to long-term health consequences [19]. Timely identification of sleep disturbances through in-situ monitoring may enable healthcare providers to intervene promptly and provide tailored treatments, thereby mitigating the adverse effects of sleep disorders on adolescent health and well-being. We present preliminary results from feasibility testing of the SleepMeasurement application we developed that uses data from a smartphone, smartwatch, and environmental sensor to deliver context-sensitive EMAs intended to monitor sleep quality. Feasibility was assessed in a two-week pilot study with 12 adolescents.

2 Related work

Poor sleep quality among adolescents negatively impacts overall well-being and academic performance [2, 4, 11, 20, 25, 27]. One cross-sectional survey involving 150 adolescents evaluated sleep quality in relation to age and sleep-related habitual and environmental factors [20]. The results revealed that 82.0% of participants exhibited poor sleep quality, irrespective of their stage of adolescence. Factors such as later bedtime, longer sleep latency, presence of electronic devices in the bedroom, and engagement with social media before sleep were associated with a higher likelihood of poor sleep quality [25]. Vazsonyi et al. conducted a two-year longitudinal study with 586 adolescents that demonstrated the developmental significance of sleep quality, as opposed to sleep quantity, on various aspects of adolescents' mental health and adjustment. Specifically, poor sleep quality increased depression, anxiety, low self-esteem, and externalizing behaviors over time, highlighting the impact of sleep disturbances on adolescent well-being [27]. These findings underscore the need for interventions that assist adolescents and their caregivers in recognizing sleep disruptions to mitigate the adverse effects of poor sleep quality on their physical and mental health.

Polysomnography (PSG), a procedure that records brain activity using an electroencephalogram (EEG) and other sensors, is the gold standard for sleep monitoring in clinical settings. However, acquiring PSG data requires attaching many sensors to the head and body. For this reason, PSG is most often only performed in clinical sleep centers with trained technicians [15]. Alternative, less-invasive techniques for evaluating sleep quality have been developed [13] for in-situ measurements. Self-report surveys provide a practical means of gathering information about sleep. The Munich ChronoType Questionnaire (MCTQ), for example, can effectively identify irregular sleep patterns influenced by socio-demographic factors [10]. EMAs have also been used to monitor sleep quality [9, 12, 28]. However, self-report surveys rely on participant memory. Passively sensed sleep data, alternatively, eliminate recall bias. Beattie et al. used optical pulse plethysmography (PPG) and accelerometers to collect overnight sleep data from adults, achieving 69% accuracy in classifying the four sleep stages [5]. Zhang et al. [30] proposed a multi-level feature learning technique to classify sleep stages based on actigraphy and heart rate, achieving classification accuracies of 64.0% and 60.5% for different groups. Additionally, Nakamura et al. introduced Hearables [18], using in-ear sensors to classify sleep stages with 74% accuracy. Using passively sensed data alone, however, will fail to capture subjective sleep quality assessments, and body-worn passive sensing may not provide rich contextual data about events taking place in an environment that impact sleep-related behavior and sleep quality. Our system addresses the limitations of using EMA, such as reducing recall bias [24] in participants' responses, by administering context-sensitive EMA [23] that relies on passively sensed data to facilitate memory recall and help participants identify causes of motion and noise disruption events.

3 System Design

Upon consenting to participate in our study, participants were loaned a Moto G Play smartphone (Motorola, Inc.), Fossil Gen 6 smartwatch (Fossil, Inc.) and an Omron 2JCIE-BU01 environmental sensor (Omron Electronics, Inc.); the sleep measurement application we developed was installed on the smartphone and smartwatch. The environmental sensor was connected to a USB wall charger and plugged into a wall outlet in the participant’s bedroom at the beginning of the study, and participants were instructed not to move this sensor during the study. The phone and smartwatch chargers were also plugged into the wall near the bed and set up so the phone and the watch could be left charging next to the bed (Fig. 1). Research staff brought portable bedside trays to homes and extensions cords to ensure this arrangement would be possible. Staff tried to ensure that the environmental sensor was not directly next to a loud device (e.g., fan) with the sensor taped to the bedside table to best measure the ambient light in the room from the perspective of the in-bed adolescent.

The environmental sensor recorded noise level (dB), light intensity (lm), temperature (°C), barometric pressure (Pa), relative humidity (%), and CO_2 levels (e_{CO_2}), each at 1 Hz, and transmitted aggregated data every minute to the smartphone via Bluetooth. The smartphone and smartwatch ran an Android application we developed. The phone was set up to receive data from the environmental sensor every minute as an average of the 1 Hz measurements over the minute. The phone also received motion data from the smartwatch every minute. These data were used to estimate possible sleep disruptions.

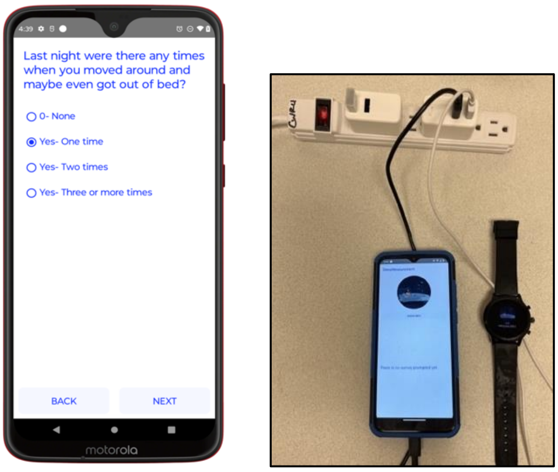

Fig. 1. Example setup of the smartwatch, smartphone, and Omron environmental sensor at a participant's home

Participants used the system for two weeks of the school year while participants attended school daily. On the first day of the study, a member of the research staff talked with participants and their families to assess typical sleep behaviors and input approximate sleep, wake, and dinner times into our application for each day of the upcoming two weeks. These sleep, wake, and dinner times could be adjusted remotely during the study by the research team if necessary. The research staff member also showed participants what a sample morning and evening survey would look like, instructed participants on how to charge the devices, and explained to participants when they would receive surveys. The phone had parental control software installed on it so that participants would not misuse study phones. The watch was plugged in by the bed and participants were instructed to take it off the charger only around dinnertime or afterwards, before going to bed; they were told to place it back on the charger in the morning. Each night, starting one hour after the set dinner time, participants were prompted on the phone (which remained in the bedroom) to remember to answer a bedtime survey. They were instructed to answer the bedtime survey right before going to sleep. This reminder was prompted every 15 minutes starting at dinner time until the participant answered the survey before sleeping. The bedtime survey administered 16-19 questions each day; the questions asked about caffeine intake, technology use, noise around the house that might keep a participant awake past bedtime, and a question for females that asked about period cramps if the participant indicated they were menstruating. Algorithms on the phone and watch apps detected significant motion and noise disruptions during the night, which were used in a morning EMA survey. To enable remote data collection, all data collected by the SleepMeasurement application were uploaded to our servers over Wi-Fi, if available, or a cellular data connection that we provided on the research phone.

Participants were prompted using audio and vibration to answer a seven-question morning survey on their phone starting 10 minutes after the wake time set for the day. Prompting continued until the participant either completed the survey or until 11:59 AM of the day. The morning survey always contained questions about when the participant went to bed the previous night, when the participant woke up that morning, and perceived quality of the participant's sleep. If the SleepMeasurement application detected times when there was a loud noise or a significant motion event, up to six additional questions could be asked about motion or noise disruptions between the participant's sleep and wake events. The motion and noise detection algorithms are described in Subsections 3.1 and 3.2.

3.1 Motion detection algorithm

We deployed a motion-detection algorithm on the Fossil Gen 6 smartwatch. The smartwatch collected raw three-axis accelerometer data at 50 Hz (i.e., the 'raw signal') with a range of ±8 g. The raw signal was smoothed using a moving average filter with a window size of 0.5 s, resulting in the 'filtered signal.' This window size was picked to remove large magnitude noise with near real-time (0.5 s delay) detection of motion events, resulting in the best tradeoff for the purpose of the SleepMeasurement application. For each axis (e.g., x), the app then computed the area under the filtered curve (AUC) at timestamp 't' as $\text{AUC}_{x_t} = |\,\text{raw}_{x_t} - \text{filtered}_{x_t}|$ to approximate removing the DC component of the signal, compensating for the effect of gravity on the axis. To conserve battery and minimize CPU use on the smartwatch, we calculated a simple motion summary for all three axes as $\text{AUC}_{(x+y+z)_t} = \text{AUC}_{x_t} + \text{AUC}_{y_t} + \text{AUC}_{z_t}$. An alternative approach could be to calculate the norm of the $\text{AUC}_{(x+y+z)_t}$ vector, but this would be more computationally intensive on the smartwatch. Data were aggregated in a 10 s window (i.e., approximately 500 values) to obtain an orientation-independent motion summary: $AUC_{10s} = \sum_{t=0}^{500\ samples\ (10\ s)} \text{AUC}_{(x+y+z)_t}$.

Then, we used thresholds to decide if a period of AUC_{10s} should be classified as a non-motion or motion event. The algorithm used two bounds for the AUC values: a lower bound *l* indicating any AUC_{10s} values that are too low to be motion events, and an upper bound *u* for any AUC_{10s} values that we were certain would be caused by a significant motion event. The AUC_{10s} values were passed through a weighting function, which mapped the AUC_{10s} value to a real number between [0,1] using linear interpolation. This continuous value, AUC_{score}, indicated the significance of the AUC_{10s} value. An AUC_{score} of 1 indicates strong certainty that the high AUC_{10s} is caused by a significant motion, while a AUC_{score} of 0 indicates certainty that a low AUC_{10s} value that is caused by non-wear or sleep.

After all the AUC_{score} values were computed, we computed if the participant was awake at timestamp *t*. A motion event was defined as sustained motions for at least five minutes. We passed a window size $w_d = 5$ to the algorithm, averaged the AUC_{score} for the past w_d minutes, denoted by AUC_{avg}, and compared to a threshold $T_w \in (0,1)$. We marked an event over a period of w_d minutes, denoted by E_w, using the following function: $E_W = \begin{cases} non-motion\ event, & AUC_{avg} < T_w \\ motion\ event, & AUC_{avg} \geq T_w \end{cases}$

After tuning the detection algorithm on five participants (three females, two males) from the research team, we set the parameters of the algorithm as follows: l = 50, u = 500, T_w = 0.5, w_d = 5. All computations for the motion algorithm were performed on the Fossil Gen 6 smartwatch, and identified motion event timestamps were sent to the smartphone every minute until the smartphone acknowledged receipt of the event timestamp. These timestamps were used to present *specific motion questions* in the morning survey.

3.2 Noise detection algorithm

The environmental sensor sent data to the phone every minute for processing. Based on prior work that studied the effect of noise levels on patients' sleep in an intensive care unit [3] and CDC thresholds for loud noises that could be disruptive [1], we classified a timestamp as noisy during sleep if the aggregated sound pressure level for the minute was above 70 dB. The phone identified the three most significant noise events during the night in the morning before the morning survey was answered. It did this by picking the highest noise value above 70 dB from the time that the participant completed the bedtime survey, or one hour after the set sleep time if the participant had not answered the bedtime survey, until 10 min before the morning survey was answered, and marking that as a noise event. The algorithm then removed all other noise measurements within 15 min of the identified noise event, and then repeated the selection, looking for the second 70+ dB event. The procedure was then repeated a third time, resulting in zero to three noise events with timestamps for the evening. These timestamps were used to present *specific noise questions* in the morning survey.

3.3 Participant recruitment

This study was approved by the IRB of Case Western Reserve University (protocol STUDY20201856). Adolescents were recruited using flyers posted at recreational centers, libraries, after-school programs, information tables at health fairs, and neighborhood block parties. Adolescents were recruited if (1) the participant was 11 to 14 years old; (2) the participant's parent was 18+ years old; and (3) the participant and caregiver were willing to provide access to the adolescent's bedroom or sleeping area to ensure study equipment (chargers, environmental sensor) were appropriately situated. Additionally, only one adolescent and caregiver per family could participate in the study. Adolescents were excluded from participation based on the parent/caregiver's report if the adolescent had a diagnosed neurodevelopmental disorder, a diagnosed serious chronic disease, or a diagnosed sleep disorder; if the adolescent had an inability to speak or read English; if the child and parent would not be sleeping at home during the two-week study period; and if the family was residing in a shelter.

3.4 Data Analysis

Each participant had to answer one EMA survey before going to bed and one EMA survey the next morning after waking up, resulting in a total of 14 morning survey prompts and 14 bedtime survey prompts per participant, if the participant answered every survey during the study period. We analyzed data from the 12 participants to compute compliance and completion rates and other descriptive statistics about answering patterns. *Completion rate* is defined as: $\frac{Prompts\ answered}{Prompts\ delivered} \times 100$.
Surveys could be scheduled but not delivered due to reasons such as the phone not being charged. To measure the number of surveys answered during the two-week study period, we define *Compliance rate* as: $\frac{Prompts\ answered}{Prompts\ scheduled} \times 100$.

4 Preliminary results

Initially, 13 participants consented to participate in the study. However, one participant did not have electricity for the duration of the study, making it impossible to keep the devices charged, and thus was removed. Participant demographics for the remaining 12 participants from the feasibility study are reported in Table 1.

Table 1. Participant demographics

Variable	n
Sex	9 male, 3 female
Age (years)	Avg 12.9 years old (SD=1.1)
Race	4 black, 5 white, 3 biracial

4.1 Survey statistics

Overall, participants had a survey completion rate of 82.6% and compliance rate of 61.3%. One participant did not charge the smartphone or smartwatch for the second week of the study despite multiple attempts to contact the participant by the research team; we report on only the first week of data from this participant.

Table 2 shows the questions asked in the morning and the distribution of the responses. Overall, we received 101 morning survey responses and 105 responses from the bedtime survey. If a participant started answering a survey but did not complete it at that time, the smartphone continued to prompt the participant to complete the survey, and the survey was available to answer until the end of the prompting period. The median response duration for the morning survey was 54.3 s (SD = 563.9 s, range = 22.9 – 5,720.6 s). The median response duration for the evening survey was 83.4 s (SD = 959.7 s, range = 36.9 – 9,528.4 s). Figure 2 shows the summary of motion and noise events detected, and specific questions answered for each participant. Overall, 313 noise and 210 motion events were detected; 121 specific noise and 62 specific motion questions were answered; and 54 meaningful noise and 23 meaningful motion answers were collected (answers other than “Don’t know”).

Table 2. Results from the morning survey in order of appearance of question in survey. The morning survey questions that asked about sleep and wake time are not included here.

Questions	Distribution of responses
“How good was your sleep last night?”	0 “Very bad” (0%), 42 “Bad” (41.6%), 26 “So-so” (25.7%), 33 “Good” (32.7%), 0 “Very good”
“Last night were there any times when you moved around and maybe even got out of bed?” (general motion question)	78 “None” (77.2%), 17 “One time” (16.8%), 4 “Two times” (3.9%), 2 “Three or more times” (2.1%)
“Last night were there any loud noises that might have disturbed you?” (general noise question)	94 “None” (93.1%), 7 “One time” (6.9%), 0 “Two times” (0%), 0 “Three or more times” (0%)
“Last night around [time] you moved around a lot in bed and maybe even got up. Do you remember what happened?” (specific motion question)	39 “Don’t know” (62.9%), 10 “Lying awake in bed” (16.1%), 4 “Went to the bathroom” (6.4%), 3 “Something else” (4.8%), 2 “Got a snack or drink” (3.2%), 2 “Used electronics” (3.2%), 1 “Answered a text message” (1.7%), 1 “Answered a phone call” (1.7%)
“Last night around [time] there was a loud noise that might have disturbed you. Do you remember what it was?” (specific noise question)	67 “Don’t know” (55.4%), 30 “Something else” (24.8%), 15 “I was doing something” (12.4%), 5 “Voices from other rooms in the house” (4.1%), 2 “TV or radio” (1.7%), 1 “Voices outside the house” (0.8%), 1 “Traffic” (0.8%)

4.2 Context-sensitive EMA to assist morning recall

Throughout the study, participants reported 888.7 hours of sleep (calculated based on the responses on the morning survey). Our system collected 595.3 hours of AUC data from the smartwatch and 775.9 hours of environment data from the Omron sensor. There were 62 nights overall when the watch was not charged; we received no AUC data on those nights as a result. Figure 3 shows the percentage of sensing data collected using our system for each participant. The percentage of data collected per night was calculated as the number of hours of data collected by the watch/Omron sensor, divided by the total hours of sleep reported by the participants in the morning survey. We omitted the data from P15 because P15 did not charge the devices on many days of the study. P12 only charged the smartwatch once, resulting in 1.5 days of AUC data. Overall, our system collected 43.2% of AUC data from the smartwatch (SD = 37.7), and 78.3% of environmental sensing data from the Omron sensor per night (SD = 32.7). The low percentage of the overall AUC data was due to technical issues caused by unreliable Bluetooth connection between the smartwatch and the smartphone, and battery issues, discussed in Section 5.

Our system was able to capture motion and sound signals from participants, despite the technical difficulties. Figure 4 shows graphs of AUC and sound data collected from the system for two participants, with markers indicating the self-reported sleep and wake time, time when the bedtime and morning survey were completed, and the motion/noise events asked about during the morning survey. P19 was able to recall most of the noise and motion events detected by the system (Fig. 4 (a)). For P17, even though a significant amount of AUC data from the watch was missing due to low battery issues (the smartwatch only had 35% battery at 10 PM), the system was able to capture disruptive noise events using the Omron sensor and collect meaningful labels for the noise events, shown in Fig. 4 (b).

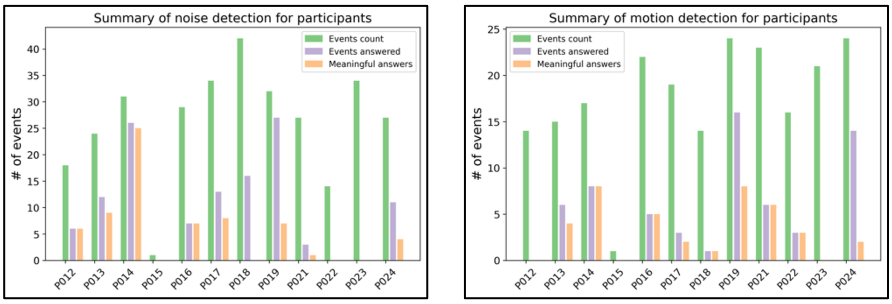

Fig. 2. Plots showing the overall number of events detected, questions answered, and questions with meaningful answers (i.e., any answers other than the “Don’t know” option) per participant for both noise and motion detection over the duration of the feasibility study.

5 Discussion

The SleepMeasurement application used sensor data from a smartwatch and Omron environmental sensor to provide contextual information for morning surveys. The Fossil Gen 6 smartwatch that we used for the study was the best model available at the time for continuous accelerometer data collection and real-time data processing. However, the smartwatch had issues with low battery life when running our software (~11 hours on a single charge), and unreliable Bluetooth connectivity. Low battery life led us to lose more than 50% of AUC data across all participants over the study duration. The battery life issue was exacerbated by the smartwatch charger being sensitive to placement, with vibrations or other movements leading to the watch being displaced from the correct charging position, and subsequent discharging. Additionally, the smartwatch went into battery saver mode when the battery dropped below 30%, resulting in intermittent data as shown in Fig. 4(b). In response, we made the decision to have participants don the smartwatch before going to bed and charge the smartwatch during the day.

The unreliability of the Bluetooth connection led to many specific motion questions not being asked in the morning survey because the smartwatch was unable to communicate with the smartphone promptly. The smartwatch often lost connection with the smartphone and sometimes did not automatically reconnect unless one of the devices was restarted. This could have been caused due to a participant rolling over and sleeping on top of the watch, restricting the Bluetooth connection. We had no programmatic control over the state of the Bluetooth connection due to restrictions placed by the Android operating system on the smartphone and smartwatch. The connection between the smartphone and Omron was more reliable; we did, however, still lose data when there was interference with the Bluetooth connection from Wi-Fi, or when the Omron sensor was out of range from the phone. We only received real-time data from the Omron sensor, resulting in lost data when the sensor was not in connection with the smartphone.

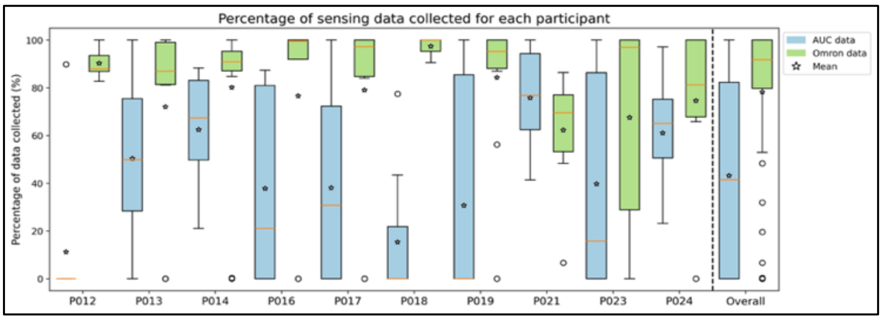

Fig. 3. Distribution of percentages of sensing data collected from the watch (AUC data) and the Omron sensor per day for each participant over the study duration.

The timing of the morning survey prompt was crucial to ensure that participants answered the EMA as close as possible to the waketime for best recall, but not early enough to disrupt sleep, and not so late that they missed answering the EMA and left for school. We initially planned to have the smartwatch detect a participant's wake time in real-time and communicate that information to the smartphone to present the morning survey accordingly. However, based on testing with our research team, we found automatic waketime detection to be unviable due to the unreliable Bluetooth connection between the smartwatch and smartphone. We thus decided to minimize reliance on communication between the smartwatch and smartphone by having the smartphone prompt the morning survey at a set waketime depending on the participant's schedule.

In the morning surveys, we provided context to participants in the form of times during the night when a significant motion or loud noise event was detected; the goal was to improve recall. Although recall improved on the provision of additional context for both noise and motion events (Section 4.2), participants were sometimes unable to recall the disruption event the next morning. Recording loud noise events and playing them in the morning survey might help provide participants with additional context that could help identify causes of disruption. This could, however, raise privacy concerns about recording sensitive conversations or other information that the participant or family members of the participant may not want to share.

The version of the noise detection algorithm described in this paper did not filter out constant sources of loud noises such as fans, heating or cooling systems, snoring, or television in the background. If these are consistent noises, participants may not find them disruptive, but the system can still mark the events as disruptive noise. Future systems should use a more advanced noise filtering algorithm to avoid flagging eliminate such constant sources of noise in the background as noise events that trigger EMA questions. The Omron environmental sensor recorded data in addition to sound levels, including temperature, humidity, and barometric pressure levels; we did not ask participants about these data in the morning survey in this version of the SleepMeasurement application. Future systems could incorporate these values to potentially improve recall of disruptive events.

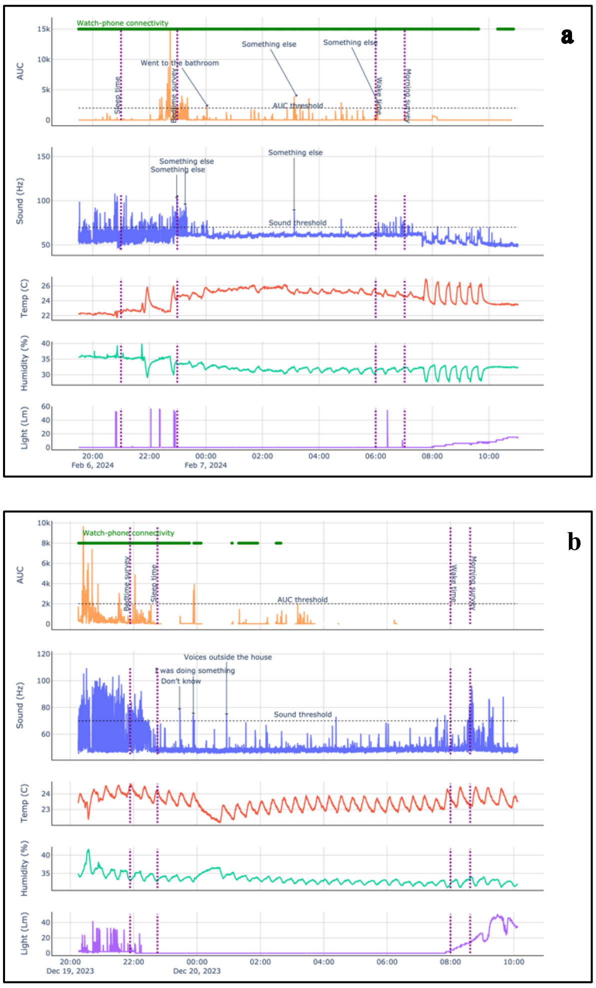

Fig. 4. Each subplot shows a night of sensing data (AUC and sound) collected overlaid with watch-phone connectivity for two participants. The smartwatch was connected to the smartphone if there is a green dot/line for that timestamp. Dashed purple lines mark participant events. Arrows mark answers to specific motion and noise questions.

Although we tested the current version of the SleepMeasurement application with adolescents, this system could be extended to other populations, including adults interested in understanding the effect of environmental factors on their sleep quality. To reduce data loss, we recommend using newer smartwatches that have improved battery life and Bluetooth connectivity. A smartwatch that can continuously collect motion data for over 15 hours without needing to be charged would be ideal to minimize data loss. We conducted preliminary tests where we ran the SleepMeasurement application on Pixel Watch 2 smartwatches and found that the Pixel Watch 2 had a battery life of ~20 hours and improved Bluetooth connectivity with the smartphone, leading to lower data loss. This version of the system did not provide any feedback to participants; future versions of this system could provide actionable insights on what an individual might change to reduce environmental disruptions in one's sleep environment.

6 Conclusion

In-situ sleep monitoring can help researchers understand sleep behaviors in a natural setting. In this manuscript, we described technical details of a system we developed that uses EMA and passive sensing to identify causes of sleep disruptions in adolescents. Participants had an overall completion rate >80% for EMAs. Participants could recall over 40% of the motion and noise events present as part of the morning EMA. Addressing challenges such as filtering out background noises in the environment, and overcoming smartwatch battery and Bluetooth connectivity issues, remain challenges to address in future sleep monitoring system development and research.

Acknowledgement

This study was supported by grant R01HD104601 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD (Spilsbury).

References

[1] *What Noises Cause Hearing Loss?* November 3, 2022 (Accessed on 2024 April 30); Available from: https://www.cdc.gov/hearing-loss/causes/.

[2] Saad Mohammed AlShareef. 2022. The impact of bedtime technology use on sleep quality and excessive daytime sleepiness in adults. *Sleep Science*, **15**(S 02): p. 318-327.

[3] Aylin Aydın Sayılan, Nurşen Kulakaç, and Samet Sayılan. 2021. The effects of noise levels on pain, anxiety, and sleep in patients. *Nursing in Critical Care*, **26**(2): p. 79-85.

[4] Younghwa Baek, Kyungsik Jung, Hoseok Kim, and Siwoo Lee. 2020. Association between fatigue, pain, digestive problems, and sleep disturbances and individuals'

health-related quality of life: A nationwide survey in South Korea. *Health and Quality of Life Outcomes*, **18**: p. 1-9.

[5] Zachary Beattie, Yang Oyang, A Statan, Atiyeh Ghoreyshi, Alexandros Pantelopoulos, Andrew Russell, and CJPM Heneghan. 2017. Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals. *Physiological Measurement*, **38**(11): p. 1968.

[6] Iddo Z Ben-Dov, Jeremy D Kark, Drori Ben-Ishay, Judith Mekler, Liora Ben-Arie, and Michael Bursztyn. 2007. Predictors of all-cause mortality in clinical ambulatory monitoring: Unique aspects of blood pressure during sleep. *Hypertension*, **49**(6): p. 1235-1241.

[7] Michael H Bonnet. 1983. Memory for events occurring during arousal from sleep. *Psychophysiology*, **20**(1): p. 81-87.

[8] D. J. Buysse, L. Yu, D. E. Moul, A. Germain, A. Stover, N. E. Dodds, K. L. Johnston, M. A. Shablesky-Cade, and P. A. Pilkonis. 2010. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. *Sleep*, **33**(6): p. 781-92.

[9] Ahuti Das-Friebel, Anita Lenneis, Anu Realo, Adam Sanborn, Nicole K. Y. Tang, Dieter Wolke, Adrian von Mühlenen, and Sakari Lemola. 2020. Bedtime social media use, sleep, and affective wellbeing in young adults: An experience sampling study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **61**(10): p. 1138-1149.

[10] Eva Farkova, Jan M Novak, Denisa Mankova, and Jana Kopřivová. 2020. Comparison of Munich chronotype questionnaire (MCTQ) and morningness-eveningness questionnaire (MEQ) Czech version. *Chronobiology International*, **37**(11): p. 1591-1598.

[11] Prayas Gautam, Maginsh Dahal, Kushalata Baral, Rohit Acharya, Sudip Khanal, Aastha Kasaju, Raj Kumar Sangroula, Koshish Raj Gautam, Kabita Pathak, and Anu Neupane. 2021. Sleep quality and its correlates among adolescents of Western Nepal: A population-based study. *Sleep Disorders*, **2021**: p. 1-8.

[12] Nancy A. Hamilton, Delwyn Catley, and Cynthia Karlson. 2007. Sleep and the affective response to stress and pain. *Health Psychology* **26**(3): p. 288-95.

[13] Zawar Hussain, Quan Z. Sheng, Wei Emma Zhang, Jorge Ortiz, and Seyedamin Pouriyeh. 2022. Non-invasive techniques for monitoring different aspects of sleep: A comprehensive review. *ACM Trans. Comput. Healthcare*, **3**(2): p. 1-26.

[14] Mari Hysing, Ståle Pallesen, Kjell Morten Stormark, Reidar Jakobsen, Astri J Lundervold, and Børge Sivertsen. 2015. Sleep and use of electronic devices in adolescence: Results from a large population-based study. *BMJ Open*, **5**(1): p. e006748.

[15] Anastasi Kosmadopoulos, Charli Sargent, David Darwent, Xuan Zhou, and Gregory D Roach. 2014. Alternatives to polysomnography (PSG): A validation of wrist actigraphy and a partial-PSG system. *Behavior Research Methods*, **46**(4): p. 1032-1041.

[16] Leslie Vaughan , Susan Redline, Katie Stone, Julianne Ulanski, Michael Rueschman, Heather Dailey, Stephen R. Rapp, Beverly M. Snively, Laura D. Baker, and Sally A. Shumaker.2016. Feasibility of self-administered sleep assessment in older women in

the Women's Health Initiative (WHI), in *Sleep and Breathing*. Springer Berlin Heidelberg. p. 1-13.

[17] Adult Obstructive Sleep Apnea Task Force of the American Academy of Sleep Medicine. 2009. Clinical guideline for the evaluation, management and long-term care of obstructive sleep apnea in adults. *Journal of Clinical Sleep Medicine*, **5**(3): p. 263-276.

[18] Takashi Nakamura, Yousef D Alqurashi, Mary J Morrell, and Danilo P Mandic. 2019. Hearables: Automatic overnight sleep monitoring with standardized in-ear EEG sensor. *IEEE Transactions on Biomedical Engineering*, **67**(1): p. 203-212.

[19] Maurice M Ohayon, Robert E Roberts, Jurgen Zulley, Salvatore Smirne, and Robert G Priest. 2000. Prevalence and patterns of problematic sleep among older adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, **39**(12): p. 1549-1556.

[20] Faith Orchard, Alice M Gregory, Michael Gradisar, and Shirley Reynolds. 2020. Self-reported sleep patterns and quality amongst adolescents: Cross-sectional and prospective associations with anxiety and depression. *Journal of Child Psychology and Psychiatry*, **61**(10): p. 1126-1137.

[21] J. Owens, Group Adolescent Sleep Working, and Adolescence Committee on. 2014. Insufficient sleep in adolescents and young adults: An update on causes and consequences. *Pediatrics*, **134**(3): p. e921-32.

[22] Shalini Paruthi, Lee J Brooks, Carolyn D'Ambrosio, Wendy A Hall, Suresh Kotagal, Robin M Lloyd, Beth A Malow, Kiran Maski, Cynthia Nichols, and Stuart F Quan. 2016. Consensus statement of the American Academy of Sleep Medicine on the recommended amount of sleep for healthy children: Methodology and discussion. *Journal of Clinical Sleep Medicine*, **12**(11): p. 1549-1561.

[23] Alejandro Porras-Segovia, Rosa María Molina-Madueño, Sofian Berrouiguet, Jorge López-Castroman, Maria Luisa Barrigón, María Sandra Pérez-Rodríguez, José Heliodoro Marco, Isaac Díaz-Oliván, Santiago de León, and Philippe Courtet. 2020. Smartphone-based ecological momentary assessment (EMA) in psychiatric patients and student controls: A real-world feasibility study. *Journal of Affective Disorders*, **274**: p. 733-741.

[24] Tiffani Razavi. *Self-report measures: An overview of concerns and limitations of questionnaire use in occupational stress research*. (Discussion Papers in Accounting and Management Science, 01-175) Southampton, UK. University of Southampton 23pp.

[25] Yasemin Şimşek and Nurdan Tekgül. 2019. Sleep quality in adolescents in relation to age and sleep-related habitual and environmental factors. *Adolescence*, **2**: p. 5-7.

[26] Giulio Tononi and Chiara Cirelli. 2003. Sleep and synaptic homeostasis: A hypothesis. *Brain Research Bulletin*, **62**(2): p. 143-150.

[27] Alexander T Vazsonyi, Dan Liu, Magda Javakhishvili, Julia J Beier, and Marek Blatny. 2021. Sleepless: The developmental significance of sleep quality and quantity among adolescents. *Developmental Psychology*, **57**(6): p. 1018.

[28] W-L. Wang, D. Hedeker, T.B. Mason, E. Dzubur, S. Intille, A. Ponnada, C. Naya, S.G. O'Connor, and G.F. Dunton. 2021. Daily coupling of negative affect and sleep

predict longitudinal changes in children's mental health: An ecological momentary assessment study. in *Annals of Behavioral Medicine (Vol. 55, pp. S594-S594)*.

[29] M. T. Williams, H. Lewthwaite, F. Fraysse, A. Gajewska, J. Ignatavicius, and K. Ferrar. 2021. Compliance with mobile ecological momentary assessment of self-reported health-related behaviors and psychological constructs in adults: Systematic review and meta-analysis. *Journal of Medical Internet Research*, **23**(3): p. e17023.

[30] Xin Zhang, Weixuan Kou, I Eric, Chao Chang, He Gao, Yubo Fan, and Yan Xu. 2018. Sleep stage classification based on multi-level feature learning and recurrent neural networks via wearable device. *Computers in Biology and Medicine*, **103**: p. 71-81.