

Article

# Toward Sensor-Based Sleep Monitoring with Electrodermal Activity Measures

William Romine <sup>1</sup>, Tanvi Banerjee <sup>2,\*</sup> and Garrett Goodman <sup>2</sup> 

<sup>1</sup> Department of Biological Sciences, Wright State University, Dayton, OH 45435, USA; romine.william@gmail.com

<sup>2</sup> Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435, USA; garrett@knoesis.org

\* Correspondence: tanvi@knoesis.org

Received: 29 January 2019; Accepted: 19 March 2019; Published: 22 March 2019



**Abstract:** We use self-report and electrodermal activity (EDA) wearable sensor data from 77 nights of sleep of six participants to test the efficacy of EDA data for sleep monitoring. We used factor analysis to find latent factors in the EDA data, and used causal model search to find the most probable graphical model accounting for self-reported sleep efficiency (SE), sleep quality (SQ), and the latent factors in the EDA data. Structural equation modeling was used to confirm fit of the extracted graph to the data. Based on the generated graph, logistic regression and naïve Bayes models were used to test the efficacy of the EDA data in predicting SE and SQ. Six EDA features extracted from the total signal over a night's sleep could be explained by two latent factors, EDA Magnitude and EDA Storms. EDA Magnitude performed as a strong predictor for SE to aid detection of substantial changes in time asleep. The performance of EDA Magnitude and SE in classifying SQ demonstrates promise for using a wearable sensor for sleep monitoring. However, our data suggest that obtaining a more accurate sensor-based measure of SE will be necessary before smaller changes in SQ can be detected from EDA sensor data alone.

**Keywords:** wearable sensor; electrodermal activity; sleep; model search

## 1. Introduction

Sleep is a necessary component to an individual's well-being. Sleep deprivation can cause an increase in day-to-day stress and induce negative emotional responses to routine daily stressors [1,2]. In the long term, continual sleep deprivation can lead to an impaired immune system and increased susceptibility to both chronic and infectious disease [3]. To address these health problems caused by sleep deficiency, we need novel sensing methods for monitoring sleep non-invasively. Traditional measures of sleep quality are self-reported [4]; however, use of sensors to monitor sleep has received increasing attention [5–8]. To this end, electrodermal activity (EDA) may be particularly useful in qualifying sleep activity [9]. In this study, we focus on better understanding the efficacy of EDA data from a wearable sensor called the Empatica E4 (E4) [10] and compare these to self-reported measures of sleep quality (SQ) and sleep efficiency (SE). We focus on better understanding: (1) latencies within features extracted from the EDA signal; (2) the structure of the relationships between the EDA features, SQ, and SE; and (3) the efficacy of the EDA signal in predicting self-reported SQ and SE.

### 1.1. Sleep Monitoring Using Surveys and Sensors

One of the most well-known self-report sleep surveys is the Pittsburgh Sleep Quality Inventory (PSQI) [4]. The PSQI uses self-report questions to calculate SE, sleep disturbances, and daily sleepiness. However, the ubiquity of fitness sensors and smartphone technology allows us to move beyond

self-report data. In Lane et al., the authors designed a smartphone app called “BeWell” to monitor a user’s physical activity, social interactions, and sleep patterns [5]. This is done using the smartphone’s accelerometer while the device is on the bed where the user sleeps. Similarly, Chen et al. discusses a Best Effort Sleep model which is implemented in the “BeWell” smartphone app previously mentioned [6]. Another model devised by Hao et al. proposes an app called “iSleep” [7]. This system uses the microphone of a smartphone to monitor sleep behaviors such as snoring and body movement to measure the SQ of a user.

Sensors like Fitbit [11] and Jawbone [12] are popular and readily available commercial sensors. Muaremi et al. demonstrated sleep monitoring in pilgrims during the Hajj 2013 to detect stress [13]. The authors used two devices: the Zephyr Bioharness 3, and the Empatica E3 which is the previous iteration of the E4 device used in our study. They created a simple two question daily stress questionnaire to monitor sleep and predict stress levels of the pilgrims monitored in the study.

While previous work demonstrated various approaches of monitoring sleep and making SQ predictions, only the study by Muaremi et al. [13] used a wearable device in conjunction with a questionnaire for sleep monitoring. In a study by Beecroft et al. [14], using only actigraphy or a questionnaire was found to be insufficient for sleep assessment. However, a key difference with [13] and our proposed study is that the authors were predicting stress measures rather than specific sleep measures such as SE or SQ. Also, the studies using only the mobile devices, accelerometer and microphone did not use the physiological signals provided by a wearable sensor. We seek to understand the efficacy of EDA data collected from the non-intrusive wearable E4 device in predicting SE and SQ derived from the daily PSQI (see Supplementary Materials).

### *1.2. Sleep Monitoring Using a Controlled Environment*

The most accurate method of monitoring sleep is carried out in an enclosed environment using invasive technologies such as the polysomnography (PSG) test. The PSG is an intrusive test which can affect the sleep of the user. Douglas et al. [8] used data from the PSG test and compared other possible methods of sleep monitoring to recommend different approaches to classifying sleep apnea/hypopnea syndrome (SAHS). In Beecroft et al., researchers looked at sleep loss in clinically ill patients and compared PSG results to data gathered from actigraphy and reports from nurses [14]. Studies such as [14] support the need for non-invasive sensors that can monitor sleep without disrupting normal sleep behavior. Moreover, it is also critical to be able to monitor sleep in the homes of the participants to better assess the sleep quality as the measurement of sleep will be affected by the environment as well (such as the lighting in the bedroom, number of bathroom visits). To that end, we describe the use of non-intrusive wearable sensors in users’ homes to evaluate quality of sleep within an authentic, natural context.

## **2. Materials and Methods**

### *2.1. Data Collection*

We used the E4 and the daily-PSQI to gather data for a total of 77 nights of sleep from six participants ages 24 to 36, three male and three female. One participant provided a majority of the collected nights of sleep totaling 38 nights of data, similar to the study performed by Sano et al. [9]. Of the remaining five participants, one participant provided four nights of data, and the remaining four participants provided seven nights each. During the data gathering process, the participants all wore the device to bed at night. When they got up, they would remove the device and answer the questions from the daily-PSQI using an Android application.

### *2.2. Instrumentation*

The device used in this study, the Empatica E4, is an unobtrusive wrist-worn physiological sensor. The E4 is similar to a Fitbit in design and size [11] and is capable of reading Blood Volume Pulse (BVP),

Heart Rate (HR), Interbeat Intervals (IBI), Skin Temperature (TEMP), 3-Axis Accelerometer (ACC), and Electrodermal Activity (EDA) [10]. We focus on the EDA feature which is the measurement of electrical conductance of the skin for detection of sympathetic nervous system arousal [15]. From the daily-PSQI, we calculated sleep efficiency (SE) as a function of the number of minutes spent asleep divided by the number of minutes spent in bed as reported by the participants. The ratings from the SQ range from one (as very poor) to four (as very good). We defined ratings of 3–4 as “good sleep” and 1–2 as “poor sleep”.

### 2.3. Creating Models for Sleep Efficiency (SE) and Sleep Quality (SQ)

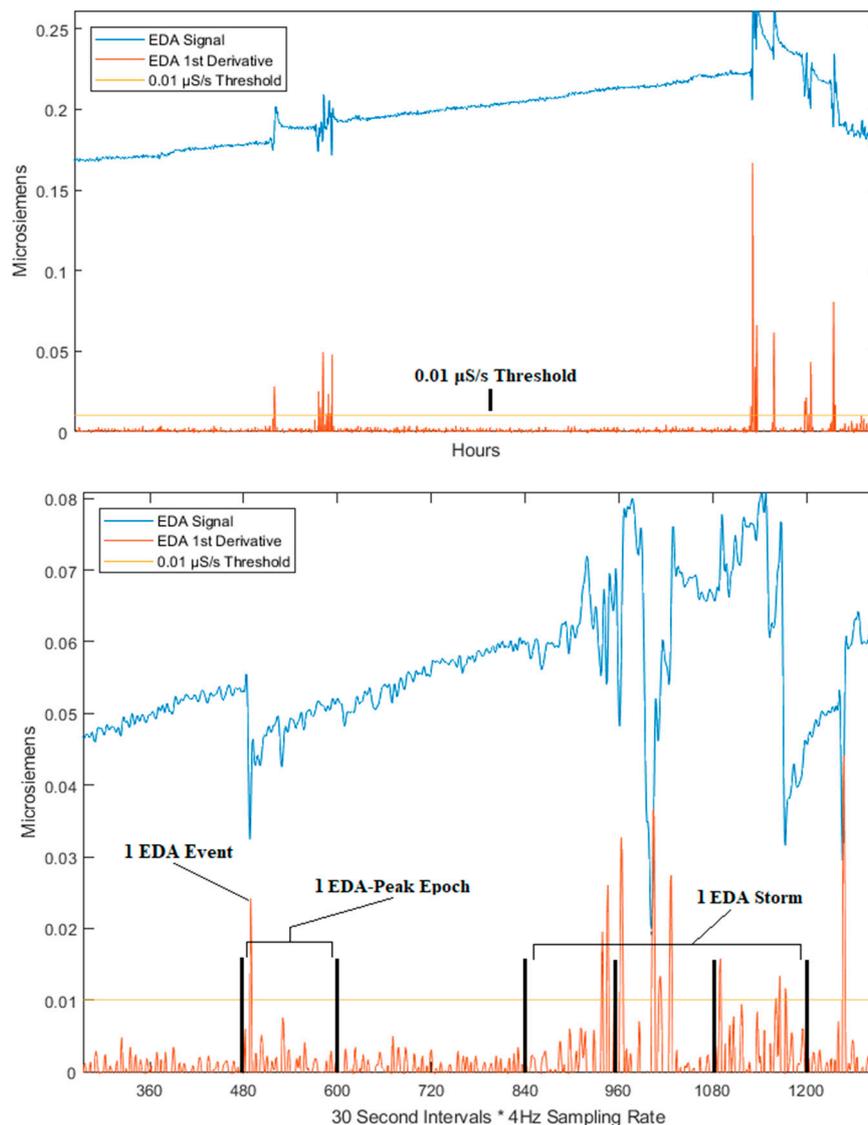
#### 2.3.1. Feature Extraction

We extracted six EDA features (Table 1, Figure 1 using the automatic feature extraction method described by Sano et al. [9]. We calculated the awake segments for derivation of SE using the E4’s 3-axis accelerometer and the function described by Cole et al. [16] but found that these did not correlate well with self-report measures ( $r = 0.17$ ). Given this low correlation and that Cole et al.’s derivation has not been validated clinically, we proceeded to use the self-report derivation for SE in this study.

The six EDA features (Table 1) are derived from three primary aspects of the EDA signal: (1) EDA peaks, (2) EDA epochs, and (3) EDA storms [9] (visual representations in Figure 1).

**Table 1.** EDA features extracted from the Empatica E4 and self-report features drawn from the daily-PSQI.

<i>Feature</i>	<i>Type</i>	<i>Description</i>
Amount of Sleep Minutes	PSQI Feature	The number of minutes the participant was asleep.
Amount of Wake Minutes	PSQI Feature	The number of minutes the participant was awake.
Number of EDA Epochs	EDA Feature	The number of EDA epochs over the entire duration of sleep in a single night
Number of EDA Storms	EDA Feature	The number of EDA storms over the entire duration of sleep in a single night
Average Size of EDA Storms	EDA Feature	Mean number of epochs within an EDA storm across all EDA storms over a single night of sleep.
Standard Deviation of EDA Storms	EDA Feature	Standard deviation in the number of peak epochs within an EDA storm across all EDA storms over a single night of sleep.
Largest EDA Storm	EDA Feature	The number of peaks within the largest EDA storm in the signal for a single night of sleep.
Number of EDA Events	EDA Feature	Number of EDA events (peaks) in the signal over the entire duration of sleep in a single night.
SE	PSQI Feature	SE calculated using the daily-PSQI questions 1, 2, and 3.
SQ	PSQI Feature	SQ deduced using SQ rating scale from the daily-PSQI, creating a binary system where poor is 1–2 rating and good is 3–4 rating.



**Figure 1.** Visual representation of our data illustrating an EDA event, EDA epoch, and EDA storm. The top image shows the peaks with respect to the 0.01  $\mu\text{S/s}$  threshold. The bottom image depicts a focused portion of the signal with the x-axis showing 30 s time intervals (compensating for the 4 Hz sampling rate of the E4 device) for determining EDA Epochs and EDA Storms.

To get these, we first take the first derivative (red line in Figure 1) of the signal (blue line in Figure 1). A threshold of 0.01 microsiemens per second ( $\mu\text{S/s}$ ) [9] (orange line in Figure 1) is then applied to this transformed signal. Any point in time which exceeds the threshold is labeled as an EDA event. Next, an EDA epoch is defined as any 30 s section of the signal in which at least one EDA event exists. An EDA storm is defined as a string of two or more consecutive EDA epochs. Average size of the EDA storms was calculated as the mean number of EDA epochs within the EDA storms occurring over a single night of sleep. Standard deviation of the size of EDA storms represents the variation from the mean in the number of EDA epochs contained in the EDA storms. The largest EDA storm was calculated as that which contained the most consecutive EDA epochs over the course of a single night of sleep.

### 2.3.2. Model Selection

Obtaining a proper understanding of how PSQI and E4 sensor readings relate to each other required specification of a directed acyclic graph (DAG) that captures accurately the causal connections

between these variables. DAGs allow derivation of causal connections between self-report and sensor sleep data without confounding factors and type 1 error inflation which often persist in ordinary linear stepwise regression methods [17]. In light of the problem described, our inferential process was fourfold: (1) search for latent factors in the extracted E4 features, (2) learn the structure of the causal connections (the DAG) between the observed or latent variables, (3) verify that this DAG reproduces the data adequately, and (4) use the DAG to derive statistical inferences which align with our research questions.

*Latent Variable Search:* We used exploratory factor analysis (EFA) with maximum likelihood estimation and promax rotation [18] to search for latent factors in the EDA signal. We retained all factors with a variance greater than that provided by a single feature (i.e., eigenvalue > 1) [19]. Regression coefficients of the promax-rotated latent factor scores on the features were used to define the meaning of the latent factors qualitatively.

*Learn the DAG from the Data:* We then utilized causal model search to derive a DAG explaining causal links between PSQI and E4 observed or latent features that are best supported by our data. We used the Fast Greedy Search (FGS) algorithm within the TETRAD software package [20] to extract the most likely DAG given the data using the features from the EDA signal, sleep quality, and sleep efficiency. As a score-based approach, the FGS selects the most probable model given the data. We used the Conditional Gaussian Bayesian Information Criterion (BIC) score with a structure prior of 1 [21].

*Verify that the DAG Explains the Data:* We then used a structural equation modeling (SEM) framework to evaluate statistically the extent to which the DAG explains the data as well as the strength of edges in the model [22]. We used Pearson's chi-square, the root mean square error of approximation (RMSEA) and comparative fit index (CFI) as measures of fit. RMSEA is a measure of absolute fit, with measures of 0.07 or below indicating adequate fit [23]. The CFI is a measure of relative fit with respect to the independence model, where a value of 1 indicates perfect fit, and a value of 0 indicates a level of fit no better than that offered by treating the features as independent. CFI values of 0.90 or above indicate adequate fit [24]. Estimation of path coefficients was done using the diagonally weighted least squares (WLSMV) estimator in Mplus 7 [25].

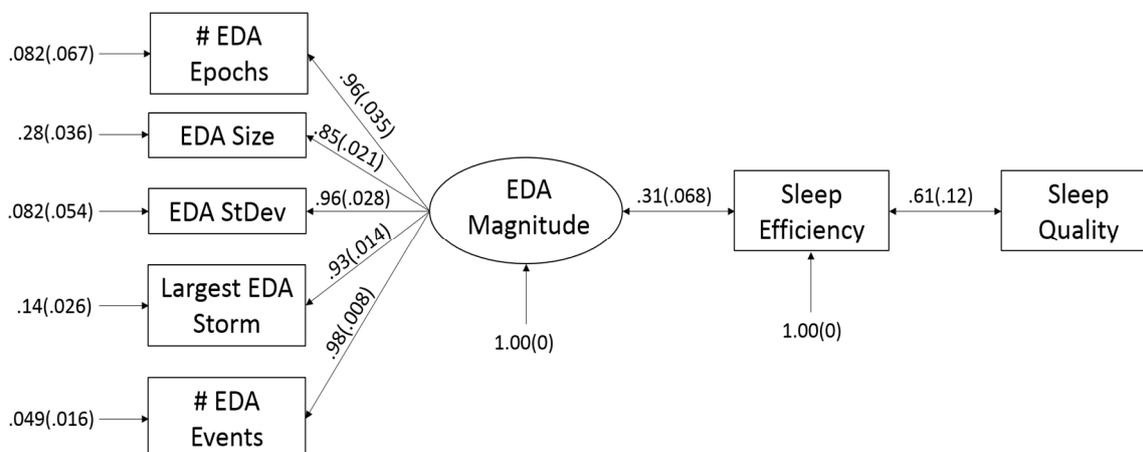
We selected the set of relevant predictors by taking the Markov blanket for the nodes of sleep efficiency and sleep quality, respectively, on the DAG [26]. Upon selecting the features based on the Markov blanket, ordinary least squares linear regression was used to evaluate the efficacy of the selected sensor features ( $\alpha = 0.05$  level) in predicting SE, and binary logistic regression was used to evaluate SQ ( $\alpha = 0.05$  level). In addition, we used the naïve Bayes classifier as a generative counterpart to gain a more complete picture of how well SE and the EDA features extracted from the E4 sensor predict SQ.

### 3. Results

The six EDA features could be described adequately with two latent factors. The first factor contained 4.5 features of variance, accounting for 75% of the total variation in the set of features. Five features (number of EDA epochs, average size of the EDA storms, standard deviation of the size of the EDA storms, the largest EDA storm, and the number of EDA events) had high loadings (0.87–1) onto this first factor, suggesting that these all describe a single latent factor: magnitude of EDA activity through the night, which we call “EDA Magnitude” for the remainder of the paper. Number of EDA storms had a small loading (−0.09) onto the EDA Magnitude factor, but instead loaded highly (0.83) onto the second factor, which we call “EDA Storms” from here on. This second factor contained 1.2 features of variance for an additional 20% of the total variance in the data. In total, the 2-factor solution described 95% of the variance in the six EDA features.

The FGS yielded the model displayed in Figure 2. SEM suggested that this model replicated the covariance between the observed features well (RMSEA = 0.064, CFI = 0.97) and could not be rejected by the data ( $\chi^2_{df=22} = 28.69, p = 0.15$ ). Paths from the latent variable of EDA Magnitude to the observed

sensor features were strong (0.85–0.98) corroborating the EFA that these measure a single latent factor. We found a significant connection between EDA Magnitude and reported SE ( $b = 0.31$ ,  $SE = 0.068$ ,  $z = 4.6$ ,  $p < 0.001$ ), and between reported SE and SQ ( $b = 0.61$ ,  $SE = 0.12$ ,  $z = 5.3$ ,  $p < 0.001$ ). These constitute moderate-to-large effects [27], indicating that the features extracted from the E4 sensor may be useful in predicting both SE and SQ.



**Figure 2.** Path diagram of the causal relationship between EDA features extracted from the E4 sensor, reported sleep efficiency, and reported sleep quality. Standardized coefficients are reported. All are significant at the 0.05 alpha level.

The Markov blanket for SE on the DAG in Figure 2 suggests that EDA Magnitude and SQ can be used as direct predictors of SE. The Markov blanket on SQ suggests that SE may be sufficient to predict SQ without need for EDA data, but that EDA features may also be useful for prediction of SQ through their ability to predict SE.

The significant relationship between EDA Magnitude extracted from the E4 sensor and SE after controlling for reported SQ ( $r_{\text{partial}} = 0.28$ ,  $p = 0.018$ ) provides support for the utility of EDA sensor data in addition to self-report data for sleep monitoring applications. A linear regression model using only the sensor data with 10-fold cross-validation models SE with a root mean square error (RMSE) of 0.99, a mean absolute error (MAE) of 0.75. This model is statistically significant ( $F_{df=1,73} = 7.59$ ,  $p = 0.007$ ,  $r^2_{\text{adj}} = 0.082$ ,  $SE_{\text{reg}} = 0.96$ ). This indicates that using only the sensor data without self-reported information may provide a promising model to predict SE.

Binary logistic regression using both EDA Magnitude and SQ as predictors ( $\chi^2_{df=2} = 16.4$ ,  $p < 0.001$ , Cox & Snell  $r^2 = 0.20$ ) demonstrates the need for including SE as a predictor for SQ. After accounting for SE, a participant's EDA activity has no significant effect on SQ (OR = 1.2,  $\chi^2_{df=1} = 0.31$ ,  $p = 0.58$ ). SE is in itself useful for predicting SQ (OR = 3.7,  $\chi^2_{df=1} = 10.6$ ,  $p = 0.001$ ). The logistic regression model only using SE predicts SQ with an AUC of 0.76 and an F1 score of 0.73 (precision = 0.73, recall = 0.76). The predictive power of the logistic regression model remains relatively unchanged when the sensor data are used in addition to SE, with an AUC of 0.77 and an F1 score of 0.73 (precision = 0.73, recall = 0.76). The naïve Bayes classifier with both EDA Magnitude and SE as predictors also performs well, with an AUC of 0.85 and an F1 score of 0.83 (precision = 0.82, recall = 0.83). When only SE is used as a predictor of SQ (without the EDA sensor data), the predictive power of the naïve Bayes model decreases slightly: AUC of 0.83 and F1 score of 0.80 (precision = 0.80, recall = 0.80). The increased utility of the sensor data in the naïve Bayes model may be due to our relatively small number of training samples relative to many machine learning applications. Generative models approach their asymptotic error more quickly than discriminative models, explaining why naïve Bayes outperforms logistic regression in this study and benefits from addition of EDA Magnitude to the model [28].

#### 4. Discussion

Our data suggest that EDA measures show promise toward sleep monitoring, but more work remains for a deeper understanding of physiological changes occurring during the different sleep stages (such as the differences in REM vs slow wave sleep). After controlling for reported SQ, EDA Magnitude had a significant relationship with SE, supporting the usefulness of EDA measures toward predicting SE. However, our data suggest that only a coarse (high/low) measure of SE can be attained through the use of sensor data alone. This implies that a sensor using these EDA features may be useful for detecting large anomalies in time asleep which may be induced by changes in medication, major lifestyle changes, or large changes in living environment. That said, more work is needed before such a device could be used to detect fine deviations in SE. This would indicate the use of additional features such as heart rate variability, along with the current EDA measures.

Figure 2 suggests that SE is a mediator between measures of EDA Magnitude and SQ; however, insufficient evidence is available to tease out the directionality of these relationships i.e., whether EDA causes SQ or changes in SE and SQ result in EDA changes. It suffices to say, however, that when SE is accounted for, the relationship between EDA Magnitude and SQ is non-significant. It is not surprising, then, that omitting the sensor data and just using SE as a predictor for SQ yields no significant difference in the classification performance when logistic regression is used. Naïve Bayes, however, tells a slightly different story, generating a small but noticeable deterioration in classification performance (F1 score decreasing from 0.83 to 0.80) when EDA Magnitude is removed from the model. This indicates that having both EDA data and a measure for SE may be helpful for a higher performance when a generative modeling framework is used. Generative models like naïve Bayes approach their asymptotic error more quickly than discriminative models like logistic regression [28], and may therefore prove advantageous toward modeling our 77 data samples. Further, including the simple EDA features does not add much to the computational complexity of the model while slightly improving model performance. Whether a discriminative or generative modeling approach is used, a useful classifier for high vs low SQ is obtainable so long as an accurate and repeatable measure of SE exists.

Toward understanding the efficacy of EDA data in informing us about SE and SQ on a nightly basis, a reader may ask why we did not use a sensor-related measure of SE such as that described in [16]. We had considered using Cole's function [16] for calculation of SE that utilized the accelerometer data from the  $x$ ,  $y$ , and  $z$  directions. However, this showed poor correlation with the self-reported SE values used in this study ( $r = 0.17$ ,  $F_{df=1,73} = 2.03$ ,  $p = 0.16$ ), meaning that they only shared ~3% of their variance in common. Further, this measure was not informative from a modeling perspective given that it showed no relationship to EDA Magnitude ( $F_{df=1,73} = 0.43$ ,  $p = 0.52$ ) and EDA Storms ( $F_{df=1,73} = 0.15$ ,  $p = 0.70$ ) based on linear regression, and was a non-significant predictor of SQ based on logistic regression ( $\chi^2_{df=1} = 0.078$ ,  $p = 0.78$ ). These results illustrate that any relationship found with the sensor-derived measure of SE in our data is likely due to chance. We consider self-reported SE as a ground truth, so given this negligible relationship, we decided not to use this sensor-derived measure in our current modeling. This computation using the accelerometer was reported in the study by Sano et al. [9] but not validated (they used an earlier version of the E4 device). The poor performance from using the movement (in  $x$ ,  $y$ , and  $z$  directions) alone highlights the differences in individual sleep habits such as certain participants being more active in their sleep as compared to others, which a reliable automated sensor must be able to capture. Clearly, more customized methods are needed to cater Cole's function to the habits of specific individuals.

Wrist-worn EDA devices are most useful for collecting measures in the context of activities of daily living such as cooking, cleaning, eating, or sleeping. However, these measures are limited in that they are less sensitive than finger transducers such as the Moodmetric EDA ring which measure EDA on the palm [29]. We use a wireless EDA device, but this consequently requires a wrist-worn measure where there is less EDA activity, and it therefore may exhibit less sensitivity [9]. This limitation should be noted when interpreting our findings in that a more sensitive device with less noise will approximate

more closely the participant's true EDA measure. Lack of precision tends to attenuate the observable relationship between features [30]. We would therefore theoretically expect the connection between EDA Magnitude and SE to be stronger, and that EDA would in turn provide better performance for classifying SQ, if a palmar EDA device were used. Practically, however, a finger-worn wired device is more likely to interfere with sleep, thereby invalidating the context of the data. As development of more sensitive and less invasive devices continues, our ability to use EDA to monitor sleep will continue to improve.

## 5. Conclusions

Our data indicate that a combination of sensor and self-report data be used to generate a useful metric to both quantify and qualify sleep. However, we still have a long way to go towards measuring sleep with sensors alone. That said, we have taken a promising first step towards understanding how to quantify sleep measures (SE and SQ) using EDA sensor data. We hope this will be used for continuous assessment of sleep behavior when large changes or deteriorations in individual sleep patterns over time are of interest. The next step involves including more physiological features to enable detection of more subtle temporal aspects of sleep behavior.

**Supplementary Materials:** The daily-PSQI is available online at <http://www.mdpi.com/1424-8220/19/6/1417/s1>.

**Author Contributions:** W.R. conducted the data analysis and drafted the Introduction, Methods, Results, and Discussion sections. T.B. conceptualized the study and drafted the Literature Review and Discussion Sections. G.G. collected the data, extracted the features, and contributed to the Literature Review.

**Funding:** This work was supported in part by the NIH under grant K01 LM012439-02.

**Acknowledgments:** We thank the participants who made this study possible. We would like to thank the Center for Causal Discovery, supported by grant U54HG008540, for providing open access to its software TETRAD.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Human Subjects:** All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Wright State University (06155).

## References

1. Minkel, J.D.; Banks, S.; Htaik, O.; Moreta, M.C.; Jones, C.W.; McGlinchey, E.L.; Simpson, N.S.; Dinges, D.F. Sleep deprivation and stressors: Evidence for elevated negative affect in response to mild stressors when sleep deprived. *Emotion* **2012**, *12*, 1015–1020. [[CrossRef](#)] [[PubMed](#)]
2. Daniela, T.; Alessandro, C.; Giuseppe, C.; Fabio, M.; Cristina, M.; Luigi, D.G.; Michele, F. Lack of sleep affects the evaluation of emotional stimuli. *Brain Res. Bull.* **2010**, *82*, 104–108. [[CrossRef](#)] [[PubMed](#)]
3. Irwin, M.R. Why Sleep Is Important for Health: A Psychoneuroimmunology Perspective. *Annu. Rev. Psychol.* **2015**, *66*, 143–172. [[CrossRef](#)] [[PubMed](#)]
4. Buysse, D.J.; Reynolds, C.F., III; Monk, T.H.; Berman, S.R.; Kupfer, D.J. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res.* **1989**, *28*, 193–213. [[CrossRef](#)]
5. Lane, N.; Mohammad, M.; Lin, M.; Yang, X.; Lu, H.; Ali, S.; Doryab, A.; Berke, E.; Choudhury, T.; Campbell, A. BeWell: A Smartphone Application to Monitor, Model and Promote Wellbeing. In Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare, Dublin, Ireland, 23–26 May 2011.
6. Chen, Z.; Lin, M.; Chen, F.; Lane, N.D.; Cardone, G.; Wang, R.; Li, T.; Chen, Y.; Choudhury, T.; Campbell, A.T. Unobtrusive Sleep Monitoring using Smartphones. In Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare, Venice, Italy, 5–8 May 2013; pp. 145–152.
7. Hao, T.; Xing, G.; Zhou, G. iSleep: Unobtrusive SQ Monitoring using Smartphones. In Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys'13), Roma, Italy, 11–15 November 2013.
8. Douglas, N.; Thomas, S.; Jan, M. Clinical value of polysomnography. *Lancet* **1992**, *339*, 347–350. [[CrossRef](#)]
9. Sano, A.; Picard, R.W.; Stickgold, R. Quantitative analysis of wrist electrodermal activity during sleep. *Int. J. Psychophysiol.* **2014**, *94*, 382–389. [[CrossRef](#)] [[PubMed](#)]

10. Real-time Physiological Signals. E4 EDA/GSR Sensor. Empatica. Available online: <https://www.empatica.com/en-eu/research/e4/> (accessed on 6 November 2018).
11. Designer. Fitbit Blaze™ Smart Fitness Watch. Available online: <https://www.fitbit.com/blaze> (accessed on 20 January 2019).
12. Times, T.N.Y.; Verge, T.; CNET. UP by Jawbone™. Find the Tracker That's Right for You. Jawbone. Available online: <https://jawbone.com/up/trackers> (accessed on 20 January 2019).
13. Muaremi, A.; Bexheti, A.; Gravenhorst, F.; Arnrich, B.; Troster, G. Monitoring the impact of stress on the sleep patterns of pilgrims using wearable sensors. In Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Valencia, Spain, 1–4 June 2014; pp. 185–188.
14. Beecroft, J.M.; Ward, M.; Younes, M.; Crombach, S.; Smith, O.; Hanly, P.J. Sleep monitoring in the intensive care unit: Comparison of nurse assessment, actigraphy and polysomnography. *Intensive Care Med.* **2008**, *34*, 2076–2083. [[CrossRef](#)] [[PubMed](#)]
15. Boucsein, W. *Electrodermal Activity*; Plenum: New York, NY, USA, 1992.
16. Cole, R.J.; Kripke, D.F.; Gruen, W.; Mullaney, D.J.; Gillin, J.C. Automatic Sleep/Wake Identification from Wrist Activity. *Sleep* **1992**, *15*, 461–469. [[CrossRef](#)] [[PubMed](#)]
17. Henderson, D.A.; Denison, D.R. Stepwise regression in social and psychological research. *Psychol. Rep.* **1989**, *64*, 251–257. [[CrossRef](#)]
18. Hendrickson, A.E.; White, P.O. Promax: A quick method for rotation to oblique simple structure. *Br. J. Stat. Psychol.* **1964**, *17*, 65–70. [[CrossRef](#)]
19. Costello, A.B.; Osborne, J.W. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Pract. Assess. Res. Eval.* **2005**, *10*, 1–9.
20. Glymour, C.; Scheines, R.; Spirtes, P.; Ramsey, J. TETRAD [Computer Software]. Center for Causal Discovery. 2016. Available online: <http://www.phil.cmu.edu/tetrad/current.html> (accessed on 12 October 2018).
21. Ramsey, J.; Glymour, M.; Sanchez-Romero, R.; Glymour, C. A million variables and more: The Fast Greedy Equivalence Search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images. *Int. J. Data Sci. Anal.* **2017**, *3*, 121–129. [[CrossRef](#)]
22. Klem, L. Path analysis. In *Reading and Understanding Multivariate Statistics*; Grimm, L.G., Yarnold, P.R., Eds.; American Psychological Association: Washington, DC, USA, 1995; pp. 65–97.
23. Steiger, J.H. Understanding the limitations of global fit assessment in structural equation modeling. *Pers. Individ. Differ.* **2007**, *42*, 893–898. [[CrossRef](#)]
24. Bentler, P.M. Comparative Fit Indexes in Structural Models. *Psychol. Bull.* **1990**, *107*, 238–246. [[CrossRef](#)] [[PubMed](#)]
25. Muthén, B.O.; Muthén, L.K. *Mplus 7 Base Program*; Muthén & Muthén: Los Angeles, CA, USA, 2012.
26. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*; Representation and Reasoning Series; Morgan Kaufmann: San Mateo, CA, USA, 1988.
27. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge Academic: New York, NY, USA, 1988.
28. Ng, A.Y.; Jordan, M.I. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2002.
29. Torniaainen, J.; Cowley, B.; Henelius, A.; Lukander, K.; Pakarinen, S. Feasibility of an electrodermal activity ring prototype as a research tool. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 6433–6436.
30. Spearman, C. The proof and measurement of association between two things. *Am. J. Psychol.* **1904**, *15*, 72–101. [[CrossRef](#)]

