

Factor Analysis and Structural Equation Modeling of Actigraphy Derived Sleep Variables

J.E. Slaven*¹, M.E. Andrew¹, J.M. Violanti², C.M. Burchfiel¹ and B.J. Vila³

¹*Biostatistics and Epidemiology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Morgantown, WV, USA*

²*School of Public Health and Health Professions, Department of Social and Preventive Medicine, State University of New York at Buffalo, NY, USA*

³*Criminal Justice Program and Sleep and Performance Research Center, Washington State University, Spokane, WA, USA*

Abstract: Actigraphy is a useful tool in the measurement of sleep, with many applicable variables that indicate the quality of a study participant's sleep. Typical actigraphy derived variables from sleep studies include sleep efficiency, sleep latency, total sleep time, and the number of nightly awakenings. Actigraphy can also give parameters for autocorrelation and sinusoidal fitting that are used in circadian rhythm modeling. All of these variables are useful in inferential statistics to determine differences between groups. However, no work has been done that examines a set of relationships among these various sleep variables, which can possibly reduce the dimension of the set into a smaller number of interpretable latent factors, which can aid in characterizing sleep quality. Factor analyses and structural equation modeling were used to assess how these measured sleep variables load into different factors. Analyses showed that these variables load into 3 factors: sleep time, sleep activity and circadian rhythm.

Keywords: Structural equation modeling, actigraphy, sleep models.

INTRODUCTION

Quality of sleep has been shown to affect health and overall performance; moreover, poor sleep can have negative impact on mental and physical characteristics as well as on social factors such as performance during shift work [1,2,3]. Analyzing sleep patterns and sleep characteristics enables researchers to quantify and categorize sleep quality.

One method of gathering sleep data is through actigraphy, which continues to be a popular method of objectively determining sleep characteristics. Actigraphs record participant movement, which in turn allows researchers to determine sleep and wake times. Actigraphs have been shown to have a high correlation with polysomnography (PSG) [4]. Unlike PSG, however, they are able to record data 24 hours a day, over a range of several days, giving long, continuous streams of data for sleep-wake analysis. This allows researchers to not only record data during sleep, but allows for data collection during the non-sleep times, as well as obtaining useful data quality characteristics [5].

Actigraphy provides many sleep related variables, including simple statistics such as total time slept and sleep efficiency, as well as information on the participant's circadian rhythm, including sinusoidal parameters. These large sets of sleep variables provide considerable useful information on sleep quality; however, some of them may explain the same sleep construct.

Structural equation modeling is a combination of multiple regression and factor analyses. The method is a dimension reducing analytic procedure where a large number of observed variables are reduced into a smaller set and the variables in each set are highly correlated with each other, and hence describe the same underlying construct. Structural equation models consist of two parts: a measurement model and a structural model. The measurement model describes the relationship between measured and latent variables. The structural model deals with the relationships between latent variables only.

Factor analysis has been used to evaluate other sleep topics, specifically the Pittsburgh Sleep Quality Index, where it was shown that a 3-factor model may fit PSQI sleep scores better than the 1-factor, global score model [6].

In this study, we examine the factor structure of the sleep variables that are derived from actigraphy. We perform Exploratory Factor Analysis (EFA) in order to determine the number of overall factors the variable set has and how the variables load onto the factors. Structural equation modeling (SEM) is a general extension of factor analysis and path analysis that allows the estimation and testing of linear models that may include both measured and unmeasured (latent) variables and has the ability to model relationships between latent variables in addition to the relationship between measured and latent variables [7]. Confirmatory factor analysis (CFA) is a form of SEM that involves estimation and testing of models where measured variables are thought to provide measurements of the same dimension of a phenomenon. It is used here in order to demonstrate the strength and validity of the EFA results.

*Address correspondence to this author at the Biostatistics and Epidemiology Branch, Health Effects Laboratory Division, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, 1095 Willowdale Rd, MS 4050, Morgantown, WV, 26501, USA; Tel: 304-285-6299; Fax: 304-285-6112; E-mail: jslaven@cdc.gov

METHODS

Participants

In a study of health outcomes associated with stress among police officers one of the research questions under study is how stress affects sleep quality. To help ascertain this, officers were asked to wear an accelerometer in order to record their movement, which allows us to determine the quantity and quality of their sleep. This investigation uses data from the first 388 of an anticipated 600 officers participating from the Buffalo Police Department. Data from the remaining 212 officers have not yet been collected. Of the participants, 164 had too many missing or inadequate data points, mostly from hardware malfunctions, incomplete data downloads from the accelerometer, and wearers taking off the device more often than instructed [5]. They thus were excluded, giving a final data set of 224 participants for the factor analysis. A general description of the study's design, methods, and participant characteristics has been reported [8]. For the 224 participants in this exploratory methods study, age ranged from 27 to 66 years, with a mean (standard deviation) of 40.8 (6.8) years. 80% of participants were Caucasian, 19% were African-American, and 1% were of Hispanic descent. Males made up 72% of the sample.

Initially, the Motionlogger Actigraphs were to be worn for 15 days without taking them off at any time. It was found that these instruments were not completely water resistant, and the protocol was changed to allow the users to take them off for short periods of time in order to protect them from water damage (e.g., while bathing or swimming). All phases, testing, and reports of the study were approved by the State University of New York at Buffalo Internal Review Board and the National Institute for Occupational Safety and Health Human Subjects Review Board.

The data from the accelerometers were transferred into computer files using Action4 software (Ambulatory Monitoring, Inc). Sleep was scored using the Primary Integration Mode Channel. PIM has been previously shown to work at least as well as the Zero-Crossing Mode and the Time-Above-Threshold methods in scoring sleep [9]. The files were then exported into Excel and finally transformed into SAS data files.

Measures

Actigraphy provides many sleep-related variables. Following standard actigraphy related studies and sleep models, [10,11,12], we chose the following variables for analysis: sinusoidal amplitude, sinusoidal messor (wavelength), 24-hour autocorrelation, time off of 24-hour circadian rhythm, average number of awakenings per night, sleep efficiency, mean activity during sleep, sleep to wake onset, wake within sleep percent, and average total sleep time per day.

The sinusoidal amplitude and messor (wavelength) give the parameters of a fitted sinusoid related to circadian rhythm; 24-hour autocorrelation is the Pearson correlation coefficient based on 24-hour time periods for each participant; sleep efficiency is the total amount of time spent sleeping out of the total amount of time spent in bed; mean activity during sleep is a measure (in volts) of how much a participant moves while having their activity recorded as being asleep; sleep to wake onset is the length of time between

falling asleep and waking up; wake within sleep percent is the percentage of time a participant wakes up during their sleep cycle.

Data Analysis

All analyses were performed using SAS 9.1 (SAS Institute, Cary, NC). To determine the number of factors to retain for EFA, principal factoring analysis, using maximum likelihood estimation, was performed with Proc Factor. Both the proportion criterion and the generated scree plot indicated that three factors should be retained [13].

Proc Factor was then used again to perform EFA on the variable set, using maximum likelihood estimation. The first step of our EFA analysis was to perform an orthogonal rotation on the data in order to maximize the variance of the factor pattern matrix. This was followed by an oblique rotation, which relaxes the orthogonality and maximizes the variance of the structure matrix [14]. Oblique rotations often make sense because forcing latent variables or factors to be orthogonal may not be reasonable where it is expected that factors should be correlated, based on the theory at hand, and can lead to factors that are misleading and have no meaningful interpretation.

Factor loadings were evaluated with the following criteria [15]: 0.71 or greater – excellent; 0.63 to 0.70 – very good; 0.55 to 0.62 – good; 0.45 to 0.54 – fair; 0.32 to 0.44 – poor; less than 0.32 are non-interpretable.

CFA has advantages over EFA in that it allows one to test the hypothesis that a given model is appropriate for the data. CFA also provides a very straightforward description of the structural relationships between measured variables and latent variables thought to be measured by the observed variables (e.g., Fig. 1). This analysis also allows us to compare different models using fit indices estimated by the CFA. SAS's Proc Calis was used to set up the structural equations model, using maximum likelihood estimation. Fit indices were used to determine model acceptability. We used Bentler and Bonnett's Normed Index (NFI), Bentler and Bonnett's Non-normed Index (NNI), and Bentler's Comparative Fit Index (CFI), which are considered acceptable when > 0.90 , as well as the root mean square error of approximation (RMSEA), which is acceptable at < 0.06 . Bozdogan's Corrected Akaike's Information Criterion (CAIC) and Schwarz's Bayesian Criterion (SBC) were used to test different models with each other, where the model with the smaller value indicates a better fit.

RESULTS

EFA results are displayed in Table 1, where each sleep variable is given its loading value for each of the three factors. Factor 1 is labeled "Sleep Time," given its high loadings from the variables total sleep time and sleep to wake onset. Factor 2 is labeled "During Sleep Activity" due to its high loadings from the variables wake within sleep percent, mean activity during sleep, and sleep efficiency. Factor 3 is labeled "Circadian Rhythm," because of its high loadings from the sinusoid parameters (amplitude and messor), 24-hour autocorrelation, and time away from 24-hour cycle.

Factors 1 and 2 both account for 28% of the common variance, while factor 3 accounted for just under 9% of the variance.

Table 1. Factor Matrix and Loading Scores for 3-Factor Model

Sleep Variable	Sleep	During Sleep	Circadian
	Time	Activity	Rhythm
Sinusoidal messor	0.46	0.3	0.55*
Sinusoidal amplitude	0.05	0.21	0.92*
Daily autocorrelation	0.07	0.15	0.88*
Time off 24-hour cycle	0.09	-0.06	-0.31*
Total sleep time	-0.94*	-0.52	0.08
Wake within sleep %	0.55	0.95*	0.05
Activity during sleep	0.55	0.94*	0.01
Sleep to wake onset	0.95*	0.53	-0.1
Sleep efficiency	-0.17	-0.52*	-0.24

*Indicates the dominant variables for each factor.

The three-factor model had good fit statistics (Table 2), except for the RMSEA value of 0.16 (although the NNI also failed, but barely, at 0.88). Using Lagrange multipliers to determine if there are unmodeled parameter constraints, we developed a 4-factor model in order to see if it would provide a better fit than the 3-factor model.

EFA was then performed again, specifying four loading factors, instead of three. Factor loadings are displayed in Table 3. Factors 1 and 2 do not change; however, factor 3 has split into two separate factors. The new factor 3 is labeled "Circadian Sinusoid" from its high loadings with the sinusoidal amplitude and messor and factor 4 is labeled "Circadian Strength" due to its high loadings with the 24-hour autocorrelation and time off of a 24-hour rhythm.

Both factors 1 and 2 each account for 31% of the common variance, while factor 3 accounts for 10% and factor 4 accounts for 8%.

The 4-factor model had good fit statistics (Table 2), but did not show improvement over the 3-factor model, with the CAIC and SBC both increasing, indicating a worse fit. However, one of the two variables loading on factor 4 had a very low loading value, indicating that it is potentially unnecessary.

Another 3-factor model analysis was performed, this time with neither of the two variables from factor 4 (24-hour autocorrelation and the time off of a 24-hour cycle), as they appear to work together, but do not aid in modeling sleep.

This new 3-factor model has the same variables loading on the same factors as the first 3-factor model, with the exception that the two excluded variables are not present in factor 3 (Table 4).

CFA shows this new 3-factor model has a better fit than either of the two previous models. Although the RMSEA is still higher than would be desirable, all of the other fit indices met the recommended criteria. The model fit comparison indices also indicate that this 3-factor model is a better fit than the other two.

Standardized path (regression) coefficients for this factor ranged from 0.11 to 4.98. Correlations between the factors ranged from 0.05 (low effect) to -0.68 (large effect) (Fig. 1).

DISCUSSION

Even with objective measures of sleep, such as those obtained through actigraphy, sleep quality can be difficult to

Table 2. Fit Indices from Confirmatory Factor Analysis

	Fit Index						
	NFI	NNI	CFI	CAIC	SBC	RMSEA	RMSEA 90% CI
Model 1	0.91	0.88	0.92	13.65	37.65	0.16	(0.14, 0.19)
Model 2	0.92	0.87	0.93	16.84	37.84	0.17	(0.14, 0.19)
Model 3	0.96	0.94	0.97	-11.72	-0.72	0.14	(0.11, 0.18)

Model 1: 3-Factor model with all sleep variables.

Model 2: 4-Factor model with all sleep variables.

Model 3: 3-Factor model without 24-hour autocorrelation and time off of 24 hours.

(NFI: Bonnett's Normed Index; NNI: Bonnett's Non-normed Index;

CFI: Bentler's Comparative Fit Index; RMSEA: Root Mean Square

Error of Approximation; CAIC: Bozdogan's Corrected Akaike's

Information Criterion; SBC: Schwarz's Bayesian Criterion.

Table 3. Factor Matrix and Loading Scores for 4-Factor Model

Sleep Variable	Sleep	During Sleep	Circadian	Circadian
	Time	Activity	Sinusoid	Strength
sinusoidal messor	0.48	0.31	0.72*	0.41
sinusoidal amplitude	-0.05	0.13	0.9*	0.64
daily autocorrelation	-0.02	0.08	0.52	0.89*
time off of 24-hour cycle	0.11	-0.03	-0.06	-0.33*
total sleep time	-0.96*	-0.61	-0.38	0.09
wake within sleep %	0.63	0.96*	0.34	0.02
activity during sleep	0.64	0.96*	0.34	-0.02
sleep to wake onset	0.96*	0.62	0.29	-0.11
sleep efficiency	-0.2	-0.51*	-0.15	-0.25

*Indicates the dominant variables for each factor.

Table 4. Factor Matrix and Loading Scores for 3-Factor Model, without 24-Hour Autocorrelation and Time off of 24-Hour Cycle

Sleep Variable	Sleep	During Sleep	Circadian
	Time	Activity	Rhythm
sinusoidal messor	0.51	0.3	0.63*
sinusoidal amplitude	-0.01	0.13	0.71*
total sleep time	-0.96*	-0.58	-0.09
wake within sleep %	0.66	0.96*	0.12
activity during sleep	0.66	0.95*	0.09
sleep to wake onset	0.97*	0.58	0.07
sleep efficiency	-0.22	-0.51*	-0.28

*Indicates on which factor a variable loads.

characterize, as there are many sleep variables and sleep models from which to choose.

Factor analysis has shown, though, that this large set of sleep variables can be reduced into a smaller set of factors, where the variables in the same factor help in explaining the same sleep construct. The analyses performed here show that actigraphy-derived sleep variables can characterize sleep with three main concepts: total sleep time, activity during sleep, and circadian rhythm.

This can aid investigators in the determination of what variables are optimal to use when studying sleep with actigraphy. Total sleep time is a specific indicator of overall sleep quality, especially as the human body requires a certain amount of sleep per night, with eight hours being the recommended amount. Activity levels during sleep can give insight into how restful a person's sleep truly is and may also be an indicator that a potential sleep disorder is affecting a study participant, although it should be noted that actigraphy is not a useful tool in differentiating specific types of sleep

disorders. Circadian rhythm is a standard model of sleep-wake patterns, and can act as an indicator of irregular sleep patterns, such as might be found with people who work variable shifts. Knowing what the outcome of interest is and identifying it with its corresponding sleep factor can reduce the number of sleep variables necessary to measure. It can also aid in reducing the number of variables used even if all three sleep characterizations are of interest.

One limitation of this study with respect to generalizability is that all subjects were police officers, a group of people who regularly work long hours and rotating shifts. The outcomes of another population may be different; however, the methods would remain the same and remain a valuable tool for analysis in any population of interest.

Another general limitation is that of assumed structural invariance. There may be structural differences between different subgroups, including, but not limited to, shift-work, age, or gender. However, because all participants were members of the same profession and have many of the same

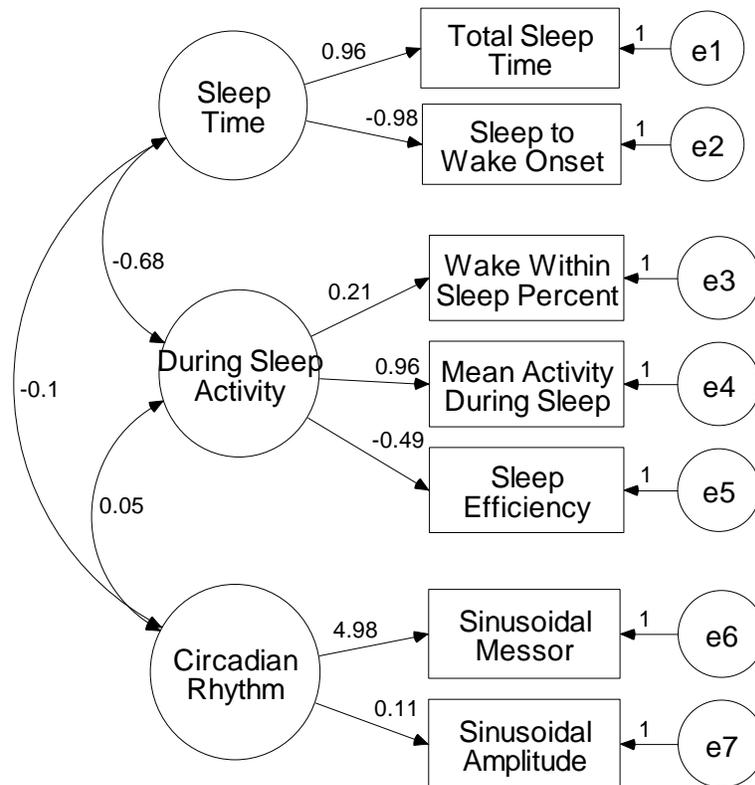


Fig. (1). Structural equation model with standardized path (regression) coefficients between the factor solution and sleep variables, and the correlations between factors.

overall characteristics and lifestyles, any effect of variance differences are expected to be minimal.

In summary, actigraphy-derived sleep variables are an objective way of obtaining sleep information. Factor analysis has shown that these sleep variables can be described by three broad characterizations of sleep: total sleep time, activity during sleep, and circadian rhythm. These three factors can aid in giving investigators a tool for determining which sleep variables are optimal for their sleep study, depending on their hypothesis and sleep variable of interest, as well as aiding in the assessment of health outcomes and sleep quality associations.

This work was supported by the National Institute for Occupational Safety and Health (NIOSH) and the National Institute of Justice (NIJ), contract no. 200-2003-018580. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

- [1] Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. *Semin Neurol* 2005; 25: 117-29.
- [2] Caruso CC, Bushnell T, Eggerth D, *et al.* Long working hours, safety, health: Toward a national research agenda. *Am J Ind Med* 2006; 49: 930-42.
- [3] Vila B. Impact of long work hours on police officers and the communities they serve. *Am J Ind Med* 2006; 49: 972-80.
- [4] Cole R, Kripke D, Gruen W, Mullaney D, Gillin J. Automatic sleep/wake identification from wrist activity. *Sleep* 1992; 15: 461-9.
- [5] Slaven JE, Andrew ME, Violanti JM, Burchfiel CM, Vila BJ. A statistical test to determine quality of accelerometer data. *Physiol Meas* 2006; 27: 413-423.
- [6] Cole JC, Motivala JS, Buysse DJ, Oxman MN, Levin MF, Irwin MR. Validation of a 3-factor scoring model for the Pittsburgh Sleep Quality Index in older adults. *Sleep* 2006; 29: 112-116.
- [7] Arminger G, Clogg CC, Sobel ME. *Handbook of statistical modeling for the social behavioral sciences* Plenum Press, New York, 1995.
- [8] Violanti JM, Burchfiel CM, Miller DB, *et al.* Buffalo Cardio-Metabolic Occupational Police Stress (BCOPS) pilot study: methods and participant characteristics. *Ann Epidemiol* 2006; 16: 148-56.
- [9] Gerardin J, Kripke D, Mason W, Elliott J, Youngstedt S. Sleep estimation from wrist movement quantified by different actigraphic modalities. *J Neurosci Methods* 2001; 105: 185-91.
- [10] Carter P. Brief behavioral sleep intervention for family caregivers of persons with cancer. *Cancer Nurs* 2006; 29: 95-103.
- [11] Hastings P, Vazir A, O'Driscoll D, Morrell M, Simonds A. Symptom burden of sleep-disordered breathing in mild-to-moderate congestive heart failure patients. *Eur Respir J* 2006; 27: 748-755.
- [12] Achermann P, Borbély A. Mathematical models of sleep regulation. *Front Biosci* 2003; 8: s683-693.
- [13] Hatcher L. *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*, SAS Institute Inc, Cary, NC, 1994.
- [14] Rennie KM. Exploratory and confirmatory rotation strategies in exploratory factor analysis; Southwest Educational Research Association annual meeting, 1997.
- [15] Comrey AL, Lee HB. *A First Course in Factor Analysis*. Lawrence Erlbaum, Hillsdale, NJ, (second edition) 1992.