Quantified Sleep: Single-Subject Study on Weekly Variation of Sleep Quality and Quantity in Regard to Well-Being Across one Year

Sebastian Idesis[1]
Supervisors: Frederik Weber[1], Martin Dresler[1]

[1]*Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, The Netherlands*

Having enough sleep is relevant to avoid cognitive and emotional problems. Longitudinal studies are necessary to track how such problems might emerge in relation to sleep. However, the majority of such studies rely on behavioural and self-reported measures of sleep. This study aims to investigate how sleep varies in quality and quantity across time. We took an exploratory approach, using a multitude of measurements, investigating whether these variations in sleep are associated with any other variables such as subjective well-being. Therefore, during a continuous longitudinal study, we followed a single participant across one year on a weekly basis and acquired polysomnography (PSG) sleep recordings, self-reported questionnaires and behavioural tasks. The results show that our participant falls within the normal ranges of his normative group regarding the length of sleep stages but deviated in sleep efficiency and duration of sleep. We found that higher reaction times in the psychomotor vigilance task (PVT) predicted increased feeling of sickness, while higher sleep efficiency predicted decreased feeling of sickness and a more positive mood. Moreover, a longer duration of sleep in minutes during the night predicted a lower level of tiredness. Furthermore, the frequency analysis revealed that certain ranges of frequencies (2-4 Hz and 12-14 Hz) fluctuated more across time compared to other frequencies. Nevertheless, these fluctuations were not associated with behavioural changes. Overall, these results could help developing personalized analysis, diagnostics and treatments. Through their high specificity, individualized interventions could thereby lead to an improvement in life quality.

Keywords: sleep, EEG, longitudinal

Corresponding author: Sebastian Idesis; E-mail: Sebastian.idesis@gmail.com

All humans need sleep. It is as important as food or water. We spend about a third of our lives sleeping. Sleep is defined as "a natural and reversible state of reduced responsiveness to external stimuli and relative inactivity, accompanied by a loss of consciousness" (Rasch & Born, 2013). Even though the amount of research on sleep is continuously increasing, the function of sleep is still not entirely clear. Different functions such as thermoregulation, metabolic regulation and adaptive immune functions have been proposed (Rasch & Born, 2013). In fact, altered sleep behaviour has been linked to various negative outcomes such as mood disorders, obesity, cardiovascular diseases, diabetes, and mortality (de Zambotti, Goldstone, Claudatos, Colrain, & Baker, 2018).

Sleep is classified into two major states: rapid-eye-movement (REM) and non-rapid-eye-movement (Non-REM) sleep. These sleep stages alternate in a cyclic way in humans as well as in other animals. Commonly, more Non-REM sleep is observed during the first half of the night, while REM sleep prevails during the second half of the night. The most widely used technique in human sleep research is electroencephalography (EEG) (De Gennaro, Ferrara, Vecchio, Curcio, & Bertini, 2005). Though other sleep assessment techniques are improving, the gold standard for sleep measurement is polysomnography (PSG) which involves recording electrooculography (EOG), electromyography (EMG) and EEG activity. To this end, multiple electrodes are applied to the head and other body parts. Depending on the placement of the electrodes, they will provide information about eye movement, muscle activity, and EEG rhythms. This information can then be used to visually score wake and sleep stages (N1, N2, N3, and REM sleep) (Berry et al, 2015).

A common problem in sleep research is the high costs of PSG. Furthermore, PSG is rather inconvenient, as it requires an expert technician and preparation time. Therefore, data are usually only acquired during a few nights instead of longitudinally. One way to solve this issue of not having enough data is to assess the phenomena that occurred during sleep behaviourally on the following day. A simple approach to assess the effect of sleep disturbances is by using reaction time (RT) tasks that measure performance during continuous work (Dinges & Powell, 1985). One of the most common RT tasks is the psychomotor vigilance task (PVT) (Graeber, Rosekind, Connell, & Dinges, 1990). Although motor tasks and mood questionnaires are easily assessable and relatively quick, they do not take into account neural data and only focus on short timescales (seconds/minutes). In contrast, there are other measures from diverse fields that approach their research by using long timescales (years/decades) with samples spread over time (non-continuous measurement), to assess the effects after several years (Mischel et al., 2011). Thus far, very limited empirical research has been done on the intermediate level: continuous measurements over an extended period.

There are several caveats for conducting research with continuous, longitudinal data. The lack of motivation and a high rate participants' dropout across time may be the most common one, increasing the complexity of acquiring repeated and frequent measures. Testing for variability over different sessions within single individuals can be inconvenient and costly. Therefore, participants should ideally be aware of the value of the data and have a high intrinsic motivation for successful completion of the study. Therefore, researchers may subject themselves to their experiment. However, the use of self-experimentation may come with expectations regarding the results of the study, which may bias the analysis and interpretation of the data (Poldrack et al., 2015).

A prevalent problem of research, in general, is the generalisability of the results. Clearly, this problem increases with fewer participants. The longitudinal approach solves the inconvenience of statistical conclusions from group studies that cannot be generalised to the intra-individual level. This is only possible in cases in which the processes in question are ergodic (i.e., "if its statistical properties can be deduced from a single, sufficiently long, random sample of the process") (Fisher, Medaglia & Jeronimus, 2018). Cognitive neuroscience studies usually extrapolate the results of single time points as representative of an individual (Poldrack et al., 2015). This is a problem as measures of brain structure and function have been shown to vary as a consequence of, for instance, exercise, water consumption, and caffeine intake. These are just a few of many examples of potential sources of intra-individual variability. As imaging results expressing between-person differences cannot be generalised to within-person time series, intra-individual variability must be studied explicitly, and appropriate methods must be developed for this purpose (Filevich et al., 2017).

Recently, longitudinal studies have benefited from technological improvement and increased access to wearable devices. The use of wearables is increasing at an exponential rate. In the US, a large group of

people uses wearable technology, providing millions and millions of data points (de Zambotti et al., 2018). These immense amounts of data, however, have not been used in continuous longitudinal studies next to neuroimaging techniques. Sleep trackers are usually part of larger devices and only used for validity for a few nights. For example, in one study the authors compared a Fitbit (Fitbit Inc., San Francisco, United States), a consumer multi-sensory wristband wearable, which among other things measures sleep quality, to PSG while assessing sleep/wake state. Nevertheless, the authors assessed 44 healthy adults by measuring only a single night (de Zambotti et al., 2018).

Other studies used neuroimaging techniques to acquire within-participant sleep data. While looking at the power spectra, one study found what they called a trait-like consistency in non-rapid eye movement sleep in adults, which involves almost perfect within-subject stability across multiple nights (Ong, Lo, Patanaik, & Chee, 2019). These correlations were lower in the low-frequency range (usually observed in slow-wave sleep stages) and in spindle frequency bins when comparing a sleep altered group compared with a control group. The researchers claimed that these "trait-like" characteristics could be considered as an electrophysiological "fingerprint". Those characteristics stayed consistent even when external manipulations such as changes in sleep schedules and drug administration were performed.

A famous study on longitudinal and single-case research is the so-called "MyConnectome" study (Poldrack et al., 2015). The data from this study has been used for several projects (Filevich et al., 2017; Mirchi, Betzel, Bernhardt, Dagher, & Mišic, 2019). In the original study, the researchers acquired Magnetic Resonance Imaging (MRI) data next to genetic data of a healthy adult over the course of 18 months. In this experiment, the authors claimed that brain function dynamics are associated with psychological and biological variability. This was not the only attempt to study the brain dynamics across time. Another, less prominent, example is single-case study where they repeatedly measured over more than a year involving 51 MRI sessions (Laumann et al., 2015). Using MyConnectome data, another study reported that functional connectivity patterns reliably predicted daily fluctuations in mood that was assessed with the Positive and Negative Affect Schedule (PANAS questionnaire) (Mirchi et al., 2019). Taking notice of the few variables used in the MyConnectome study, other researchers tried to expand the protocol by measuring over 50 covariates next to each session. One project added the measurement of, among others, the hours of sunshine, anxiety, blood pressure and physical activity (Filevich et al., 2017). The researchers also included sleep assessments based on self-report questionnaires, which included sleep quality, at what time the participant went to bed the previous night and how much time was spent in bed the previous night. In the same sense, the majority of sleep studies have focused on group measures that emphasize the differences between subjects rather than those within subjects. In order to solve that, de Gennaro and colleagues (2005) acquired several sleep EEG data from 10 subjects where they found a "fingerprint-like" topographic distribution of the EEG power during non-REM sleep. The authors argued that this stability could be due to genetically determined functional brain anatomy, instead of sleep-dependent mechanisms (de Gennaro et al., 2005). This result was also observed in other studies using a similar procedure (e.g., Finelli, Achermann, & Borbély, 2001).

In summary, until this date, sleep studies failed to generalize results from intra-individual data to a whole population and vice versa. This could be resolved by the implementation of continuous longitudinal studies. Furthermore, none of the studies mentioned above measures sleeping processes physiologically or shows how neuroimaging data and sleep interact in a continuous way across a long period of time. The majority of sleep studies rely only on behavioural and self-reported measurements. Only few of these studies intend to create a consistent measure of EEG data from the same participant and even fewer acquire data during more than a couple of nights per person. In addition, the use of wearables is increasing rapidly but it is still not clear how the information acquired by this technology can be used for bigger projects including neuroimaging techniques.

Our study aims to tackle these problems and investigate how sleep changes over the course of one year. Furthermore, we intend to explore if those sleep measures are associated with any other biological assessments like nutrition, hydration or subjective wellbeing by using wearable devices. To this end, this project was designed to measure a large set of variables from a single participant to allow the assessment of the relationship between psychological, neural and behavioural data. Due to the exploratory approach of the study, no predictions were established in advance. We followed one participant across a whole year and acquired sleep PSG recordings every week along with thirty daily behavioural measurements (including mood, productivity, and stress measurements). This allowed

us to create a rich sleep "phenotype" of one person beyond self-report questionnaires.

Methods

Despite the fact that a tremendous amount of data is available with the current set-up, this thesis mainly focuses on a fraction of the data and possible analyses to comply with text length constraints.

Participant

The experiment assessed a single case: a healthy, right-handed male, aged 29 years at study onset. He did not present any neuropsychiatric disorder prior to or during the study. The participant joined the study after providing written informed consent for EEG usage according to the guidelines of the local ethics committee. The participant offered himself voluntarily and no money rewards was given in exchange for his participation

Design and procedures

PSG recordings were performed on a weekly basis on Monday nights while the behavioural measurements such as mood questionnaires were performed on a daily basis. Collection of EEG started March 20th of 2018 and ended on March 25th of 2019. The aim was to collect weekly EEG sleep data for one calendar year (52 weeks). Due to travel and technical issues, the final amount of 50 sessions was achieved. All statistical analyses were postponed to the end of this period. In order to avoid a bias provoked by self-experimenting, the participant did not have insight into the data throughout the entire collection period (Poldrack et al., 2015). The only information related to sleep known to the participant was sleep self-reports, obtained by phone apps, which indicate sleep quality and how many hours were slept during the previous night. Because the participant was used to observe these phone-acquired reports before the beginning of the study, it was decided that he could continue this routine to increase the ecological validity. When measuring neuroimaging data from participants repeatedly across time, it is useful to keep consistency in data acquisition time of the day to standardize time-of-day effects (Gordon et al., 2017). In our study, data were acquired at the same time and day (Monday night) to keep this consistency. As there were no a priori hypotheses, the analyses approach was exploratory. All the statistical analyses were performed using the commercially available software "R" (R Core Team, 2013).

PSG procedure

The PSG setup was based on a previous master's thesis (Pottkämper, 2019) and was recorded using a wearable device (SOMNOscreen plus, SOMNOmedics GmbH, Randersacker, Germany). The recording included EEG, EMG, EOG and electrocardiogram (ECG). We acquired EEG data from 15 electrodes, in accordance with the American Academy of Sleep Medicine (AASM) rules (Berry et al., 2015) and the 10-20 system (Cz and the two mastoid electrodes were used as a reference, while Fpz was used as a ground electrode). For further details regarding the positioning of the electrodes and the full protocol, see Appendix B-E from Barry et al. (2015).

During recording, the sleep signal was pre-filtered by the recording system. EEG and EOG were high-pass filtered at 0.3 Hz and low-pass filtered at 30 Hz. EMG was high-pass filtered at 10 Hz and low-pass filtered at 100 Hz. Additionally a notch filter around 50 Hz was applied. ECG was high-pass and low-pass filtered at 0.3 Hz and 70 Hz respectively, and additionally with notch filter around 50 Hz. Sleep was scored offline using the SpiSOP toolbox (Weber, 2013, RRID:SCR_015673), in accordance with the AASM standards. Movement artefacts were manually removed, and the analysis was performed on 30-second epochs. Several variables were obtained with the PSG data, from which one of the most relevant for our analysis is the sleep efficiency, which estimates how much time was slept during the night regarding the total time spent in bed.

Mood measurements

A set of questionnaires were administered during the day by using phone apps. The questionnaires contained information about exercise, activities and mood between others. As a mood assessment, the PANAS Scale was used (Crawford & Henry, 2004). PANAS is a self-report 10 items measuring both positive and negative affect. All data were exported by the participant, without him looking at their contents and then sent to the experimenter for analysis. The data were included as a regressor in the statistical models next to the EEG data in order to assess the relationship between other variables. For the mood measurements, a correlation was assessed with sleep data variables (i.e., sleep efficiency) in

order to assess the relation between mood variables and sleep variables. Of all the subjective measures of well-being measured for the participant's own interest, we focused only on sickness, productivity, happiness, sadness, and tiredness.

PVT (psychomotor vigilance task)

Reaction time was assessed using the PVT (Graeber et al., 1990). The 2-minutes version of the PVT was performed every morning right after waking up to assess objective vigilance by means of the participant's average reaction time.

Correlational analysis

For the correlational analysis, several variables were incorporated to assess their relationship. Two of these variables were obtained during the night recordings (sleep efficiency and total sleep time in minutes) as objective indicators of sleep quality and quantity, while the others were obtained on the following day by means of questionnaires and behavioural tasks (PANAS scores, stress, sickness, tiredness, and PVT reaction time) to indicate subjective and objective measures of well-being. Correlation strength was assessed by using Pearson's r-value. A p-value of 0.05 or lower was considered significant.

Spectral EEG analysis

In this study, we performed a similar analysis to asses "trait-like" characteristics as Ong and colleagues (2019). In brief, this assessed which frequencies vary most across the year while comparing weekly EEG sleep recordings. EEG analyses were performed using the SpiSOP tool (Weber, 2013, RRID:SCR_015673) based on MATLAB 2013b (Mathworks, Natick, USA) and FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011, https://fieldtriptoolbox.org, RRID:SCR_004849). In order to obtain the results, several analysis steps were performed. These include segmentation (epoching data into time segments), calculation (transformation of the signal in each segment using Fast Fourier Transformation), averaging over bands (averaging the data points in order to display them) and normalization (normalize the power estimates to obtain power density estimates; for more information, see SpiSOP web documentation on https://www.spisop.org/documentation/).

Power spectra were calculated on consecutive artefact-free 5-s intervals of Non-REM sleep, which overlapped in time by 4 s. Each interval was tapered by a single Hanning window before applying Fast Fourier Transformation that resulted in interval power spectra with a frequency resolution of 0.2 Hz. Power spectra were then averaged across all

Table 1. Sleep table: Representation of sleep stage time, both in minutes and percentages of the total sample (N = 50). SD refers to standard deviation; SE refers to standard error; Min refers to the minimum value and Max to the maximum value. The difference (Diff) between our participant and the normative group is indicated. Values denoting NA indicate that data were not available. The normative group was taken from Boulos et al. (2019). Asterisks (*) refer to the presence of three atypical data points which could be modifying the outcome.

Sleep stage parameters	Mean	SD	Min	Max	Range	SE	Normative	Diff
Sleep efficiency (%)	97.31	1.36	93.58	99.29	5.71	0.19	89	8.31
Total sleep time (min)	376.45	36.19	289.50	437.00	147.50	5.12	410.6	-34.15
Sleep onset time (min)	6.06	2.25	2.50	11.00	8.50	0.32	14.3	-8.24
Wake after sleep onset (min)	4.46	4.94	0.00	23.00	23.00	0.70	32.1	-27.64
Stage 1 (min)	8.68	5.79	2.50	22.00	19.50	0.82	NA	NA
Stage 2 [min]	236.01	36.69	168.00	312.00	144.00	5.19	NA	NA
Stage 3 [min]	51.48	16.31	16.00	92.00	76.00	2.31	NA	NA
REM (min)	74.49	24.46	17.00	116.00	99.00	3.46	NA	NA
REM without MA (min)	67.77	22.08	14.50	108.50	94.00	3.12	NA	NA
Stage 1 (%)	2.13	1.34	0.68	5.47	4.79	0.19	6	-3.87
Stage 2 (%)	55.16	9.42	23.69	73.06	49.37	1.33	51.3	3.86
Stage 3 (%)	13.46	4.46	3.45	24.02	20.57	0.63	21.4	-7.94
REM (%)	19.64	5.92	5.21	29.87	24.65	0.84	19.8	-0.16
REM without MA (%)	17.89	5.41	4.45	27.93	23.48	0.76	NA	NA

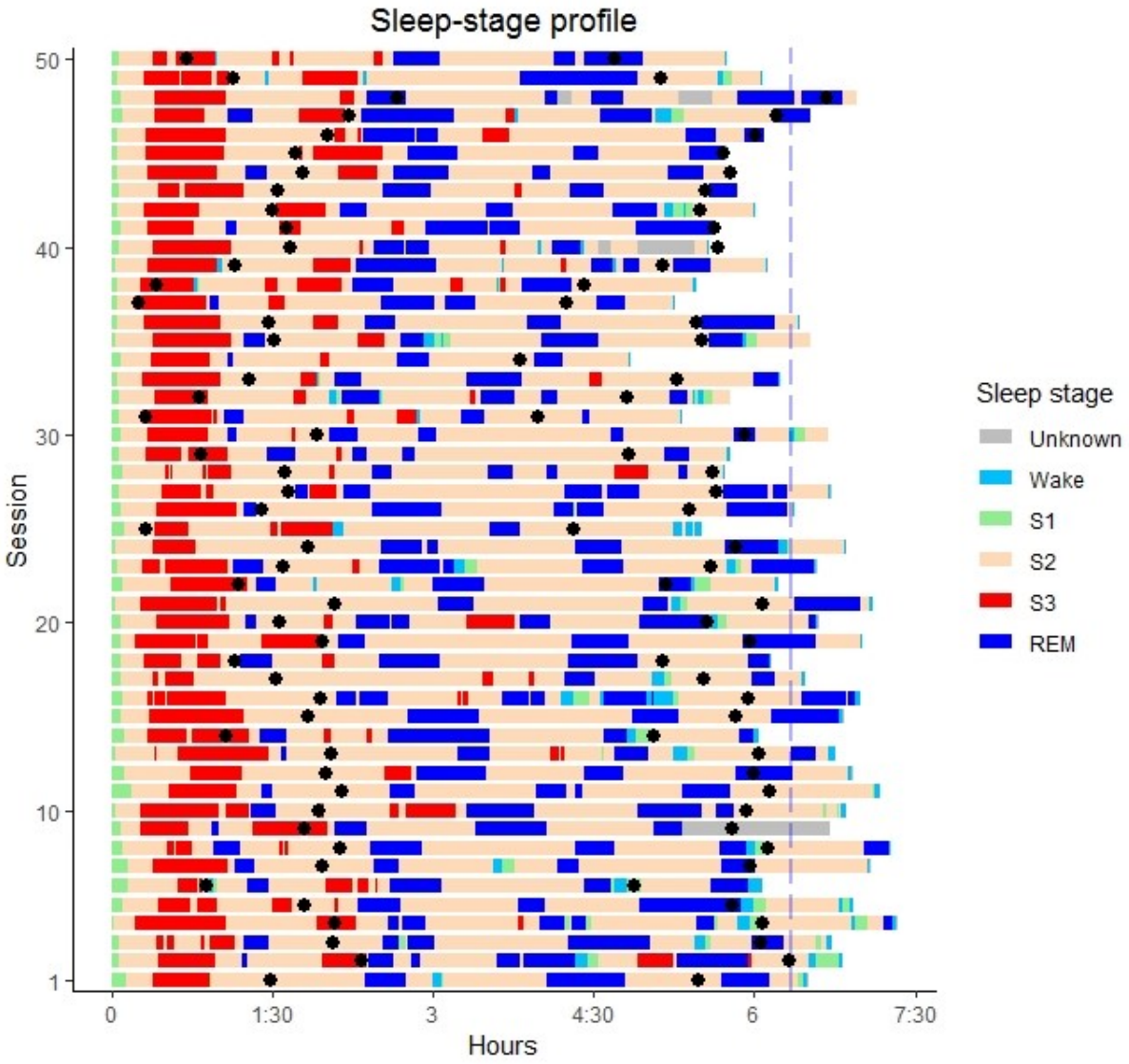

Figure 1. Sleep-stage profile. Each line represents a different session with each stage is coloured differently. The black dots indicate the time of the clock (2 and 6 AM respectively). Vertical dashed line represents the session's average.

blocks (Welch's method) and normalized by the effective noise bandwidth to obtain power spectral density estimates for the whole data (Wang, Weber, Zinke, Inostroza, & Born, 2018). This study follows an exploratory approach; therefore, all frequencies between the ranges of 0.3-30 were included in the analysis. The analysis was performed for all the EEG night recordings in order to compare them. Variation in each frequency band was assessed by calculating the standard error of the mean. For power spectrum analysis the sessions 1, 4, 5, 19, 21, 28, 29, 30, 32, 47, 48, 49 and 50 (final $N = 37$) were excluded because the absolute power values indicated technical problems (e.g., erroneous placement of electrodes).

Results

Sleep composition

In order to summarize the data obtained by the sleep EEG recordings, a descriptive analysis was performed including information about sleep stages, both in minutes and percentages. The results were compared to a normative group, matching the participant's age (healthy adults between 18 and 34 years old) that was taken from a meta-analysis using normal polysomnography parameters (Boulos et al., 2019). Through the comparison of the normative group against the case data, it is visible that our participant fits the expected range of values from the normative group regarding sleep stages length and proportion. Nevertheless, our participant had a high sleep efficiency, together with a short sleep onset duration and small amount of wake after sleep

onset. These results remain even when the lower total sleep time is considered (Table 1).

Sleep-stage profile

Hypnograms were obtained and listed in chronological order for each night of the 50 sessions with colours representing the different sleep stages (Figure 1). Appendix Figure A1 displays the same information ordered based on sleep onset time in order to achieve a different visualization of the data. The start of the sessions was aligned to the sleep onset. This descriptive summary denotes the participant's typical sleep-stage profile across the year. Also, information about when the participant fell asleep, woke up, his sleep length, the number of sleep cycles during the night and the distribution of the sleep stages are marked.

Correlational analysis of sleep parameters and behavioural measures

Association between selected variables of interest was tested with a correlation analysis including sleep efficiency, total sleep time, positive affect score, stress, tiredness, sickness, and PVT reaction time average (Figure 2). The negative affect score did not reveal any strong association with any other variable and therefore was not included in the figure. The detailed correlations are displayed in Figure 3.

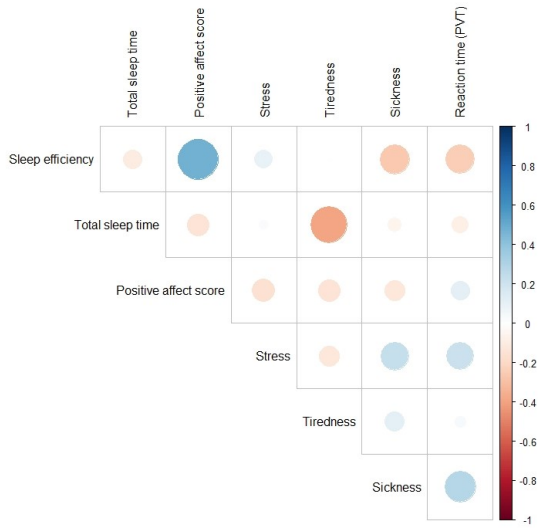

Figure. 2. Sleep quality and quantity association with behavioural measures. Pearson's correlations between the selected variables of interest. Blue dots indicate a positive correlation while the red dots indicate a negative correlation. Higher colour saturation and greater dot size indicate stronger correlation coefficients.

Higher reaction time in the morning from the PVT correlated with higher self-reported sickness of that day ($r(49) = .28, p = .047$; Figure 3A), but surprisingly not significantly related with sleep efficiency ($r(49) = -.24, p = .095$; Figure 3B). Showing that sleep quality and mood are related, the nights with higher sleep efficiency predicted higher positive affect scores in the self-reported (PANAS) questionnaire the next day ($r(49) = .49, p = .021$; Figure 3C). However, sleep was at most weakly associated with less sickness the following day. ($r(49) = -.26, p = .065$; Figure 3D).

As expected, mere higher sleep quantity, as measured by the total time of sleep, predicted subjective tiredness the next day ($r(49) = -.4, p = .0042$; Figure 3E). Other sleep parameters and behavioural measures were uncorrelated.

Spectral composition of non-REM and REM sleep

Power spectral density was calculated for all non-REM and REM stages separately for every session. Then variability between sessions (standard error of the mean, SEM) for each frequency bin was calculated. In non-REM sleep EEG activity displayed marked higher variability between 2–4 Hz and 12–14 Hz which are typically associated with sleep slow waves/K-complexes and sleep spindles, respectively. This variability was stable after exclusion of outliers due to erroneous recording sessions (Figure 4A). Surprisingly, power-density variability in REM sleep was also higher in the 2–4 Hz range (Figure 4B). By assessing a correlation between power obtained in both frequency ranges with the aforementioned behavioural variables, no significant correlation was found.

Discussion

We measured a single person across one year in order to see how sleep features change across time and how they interact with a diverse set of behavioural and psychological variables. The results show that our participant falls within the normal ranges of his normative group concerning the length of sleep stages. This, however, is not the case with sleep efficiency and duration of sleep. Here, the participant deviates from his normative group; interestingly, we found several relations between the variables that could be assessed in a causal way in future studies: Higher reaction times in the PVT task during the morning seems to predict feeling more sick. In addition, higher sleep efficiency seems

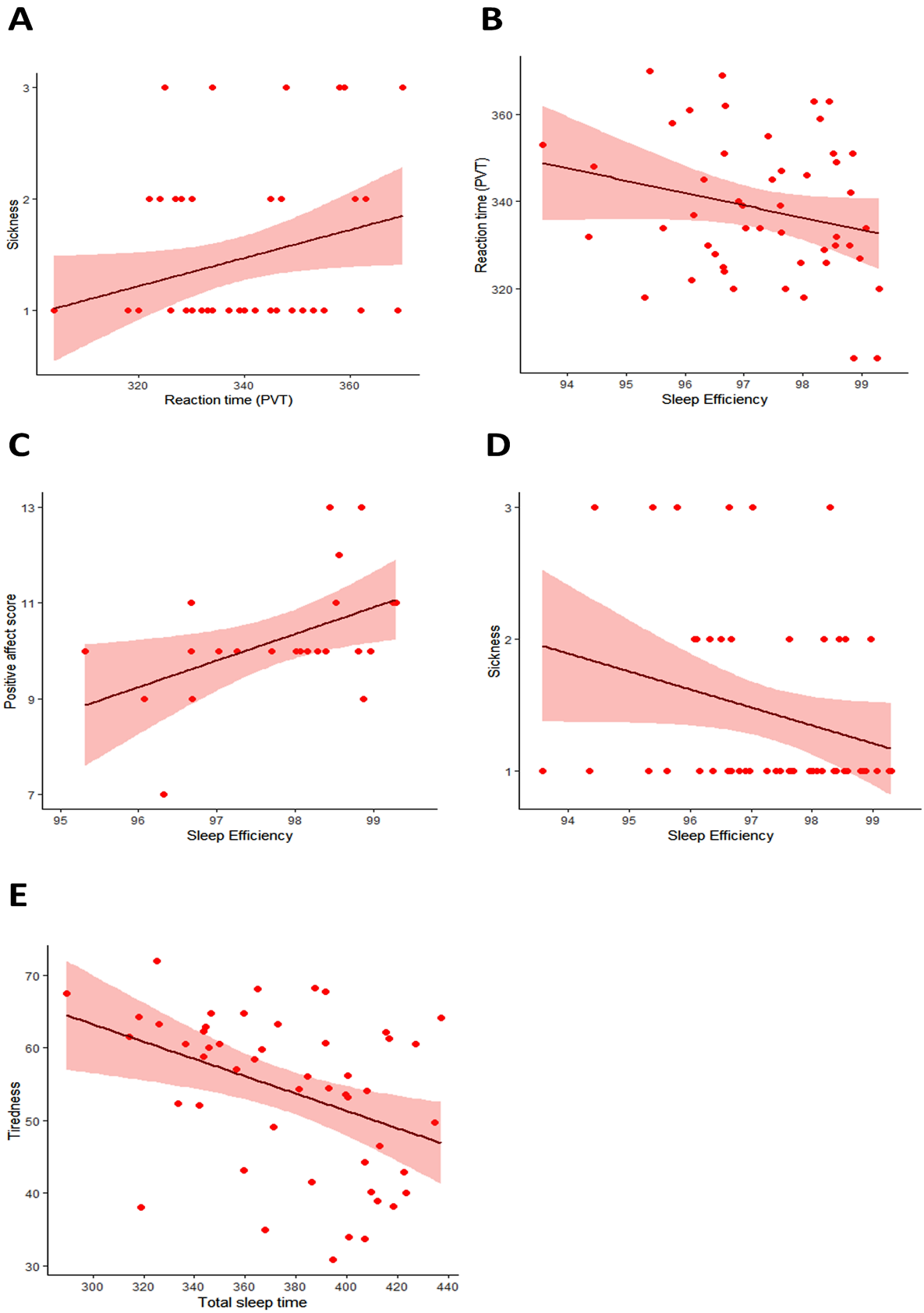

Figure 3. (A) Linear regression between the reaction time (RT) average provided by the PVT and the self-reports of sickness at the day (r(49) = .28, p = .047). (B) Linear regression between sleep efficiency and reaction time average provided by the PVT (r(49) = - .24, p = .095); (C) the Positive Affect score of the PANAS questionnaire (r(49) = .49, p = .021) (D) self-reported sickness on the day after (r(49) = -.26, p = .065), (E) the total sleep time and self-report of tiredness on the day after (r(49) = -.4, p = .004)

to predict reduced feeling of sickness and a more positive mood. Moreover, a longer duration of sleep in minutes during the night seems to predict a lower level of tiredness.

Lastly, a frequency analysis revealed that certain ranges of frequencies (2-4 Hz and 12-14 Hz during Non-REM and only the first range during REM) fluctuated more across time compared to the other frequencies. Nevertheless, these fluctuations were not associated with behavioural changes.

Sleep features

The sleep features found in the current participant are similar to those found in studies with multiple subjects (Mitterling et al, 2015; Boulos et al, 2019). Although the obtained data support most of the results, some findings contradict our expected outcomes based on previous literature. These contradictions involve high sleep efficiency, short sleep onset, low amounts of wake after sleep onset and total sleep time. Our participant displayed a significantly higher sleep efficiency compared to the normative group. He furthermore, did not use an electrode cap system, previously reported to be less comfortable as it restricts participants' movements during the night (Šušmáková, 2004). Another possible explanation is the fact that our participant slept the nights at his house and not in the lab. Sleeping in a lab is likely more stressful and therefore compromises sleep quality (Johns & Doré, 1978).

Moreover, our participant slept with the EEG device on repeated occasions. He therefore likely habituated to perturbances by wearing the device. Lastly, the sleeping duration during the nights was lower than in the normative group, while the proportion of sleep stages was similar. This could indicate sleep deprivation (Philip et al., 2004). Our participant may have increased his sleep efficiency through higher sleep pressure at sleep onset which also can explain the short sleep onset times.

Associations between sleep features and other variables

As mentioned the association between PVT reaction times and sleep quality is supported by an extensive literature (Graeber et al., 1990; Loh, Lamond, Dorrian, Roach, & Dawson, 2004; Belenky et al, 2003). It is therefore surprising that we found no strong association between PVT and sleep efficiency. This is possibly due to unusually low variability in the measures of sleep quality, which

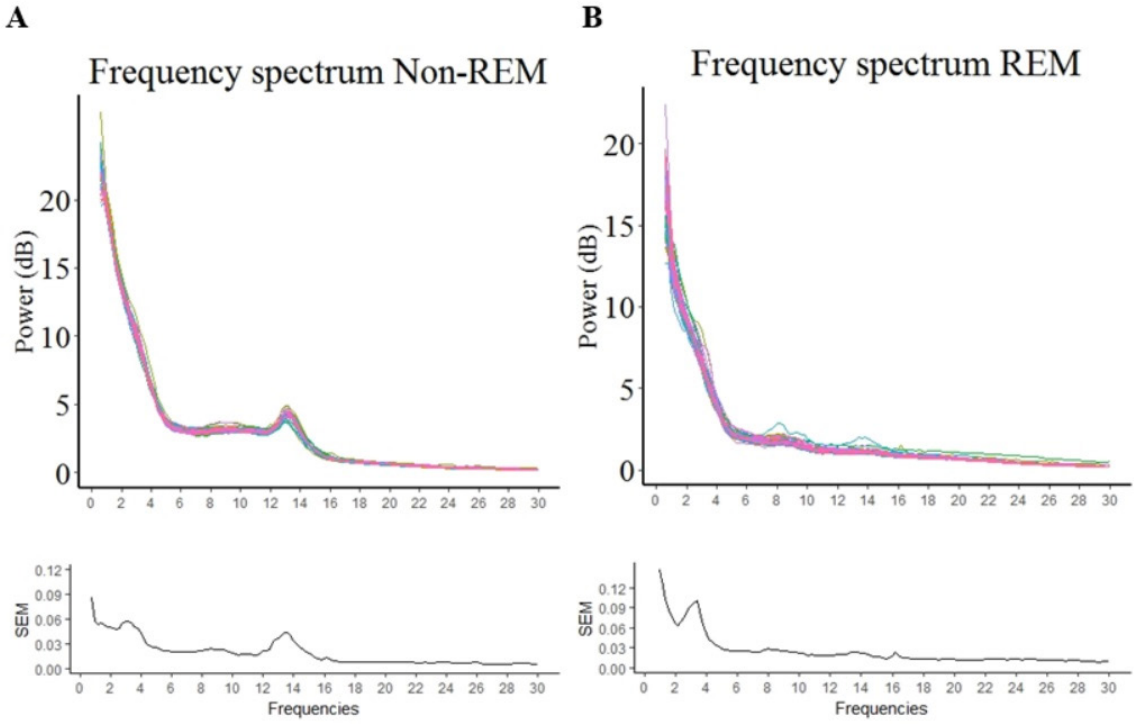

Figure 4. Power analysis of Non-REM stages (A) and REM stages (B). Power spectral density (above) indicating the presence of the different frequency ranges with the corresponding power. Each line corresponds to a different night of recording. The standard error of the mean (SEM; below) was calculated to assess which frequency bins were more variable across time. Average power was calculated from C3 and C4 electrodes as deviations to the linked, averaged mastoids.

was also relatively high. Additionally, we used a 2-minute version of the PVT task when typically, 5- to 10-minute versions are more sensitive to sleepiness during the day. Our subject did clearly not show any improvement or detriment at the PVT task throughout the first year (Figure A2). Thus, the contradiction with the literature gives rise to the question of the sensitivity of the validity of our tasks and assessment of sleep quality and whether they are suitable as valid measures in all populations. Further studies should analyse similar cases to see if the PVT task is useful for this kind of participants, who are possibly confounded by sleep deprivation states. Nevertheless, in the current study, PVT reaction times were at least sensitive to reported sickness as in previous articles (Dinges et al, 1997).

Interestingly, several studies have tried to look for the relation between sleep and mood, indicating a weak correlation between sleep quality and emotional features such as anxiety, anger, etc. (Thomsen, Mehlsen, Christensen, & Zachariae, 2003; Akashiba et al., 2002). The findings reported in the aforementioned studies align with the ones presented in the current study with the difference of the time-frame of data acquisition. Nevertheless, the current study found no correlations while assessing negative affect scores assessed by the PANAS questionnaire.

Most of the studies focused on short timescales such as days/weeks while our study acquired data throughout a whole year. To our knowledge, only one study looked into the sleep-mood relation on a large time scale. In this study, patients with bipolar disorder were measured across 18 months, assessing how sleep duration affected mood changes (Leibenluft, Albert, Rosenthal, & Wehr, 1996). The authors found a relationship between sleep and mood by claiming that decreased sleep duration was the best predictor of mood fluctuations present in mania or hypomania the next day. Nevertheless, the aforementioned study only relies on self-report questionnaires in order to measure sleep features ignoring sleep quality. In the current study, we use the same questionnaires but by adding PSG measurements, we improved sleep assessment while adding an objective and not-biased quantification.

Spectral EEG analysis

Although previous studies included repeated sessions of sleep EEG, they only assessed a few nights per participant (de Genaro et al., 2015; Ong et al., 2019). Based on previous work by Ong et al. (2019), we aimed to assess a trait-like pattern of sleep EEG power spectra. By doing this, further analysis can evaluate which behavioural/environmental variables are associated with frequency changes in order to establish causal relationships. The aforementioned study assessed sleep during three different nights to see coherent and incoherent patterns in three different conditions. In the current study, we approached this topic in an exploratory way in order to see which frequencies have a high variability throughout a year. Previous research has investigated frequency stability. A high temporal stability of gamma-band activity was found (Keil, Stolarova, Heim, Gruber, & Müller 2003).

However, as the aforementioned study focused on mechanisms underlying complex cognitive tasks, the other frequency (lower than 30 Hz) bands were ignored from the analysis. Taking an exploratory approach, we looked into a broader spectrum of frequencies. Our findings indicate higher variability in lower ranges of frequencies (2-4 Hz) during non-REM and REM stages and the sleep spindle band (12-14 Hz). Furthermore, the analysis made by Keil and colleagues (2003) was based on cognitive task by looking at specific time-windows (moments before and after behavioural response). The same happened with the work published by Zheng, Zhu and Lu (2017) where they find a relatively stable within and between sessions EEG pattern. In both mentioned studies, the analysis was task-related and therefore not comparable to the ones obtained in the current study. Here, we take a new approach by analysing the frequencies during sleep without cognitive tasks.

By assessing the correlation between the overrepresentation of these frequencies in the sleep EEG variability and behavioural measures, no strong association was found. This indicates that these variations may not transfer to subjective and objective measures of well-being (positive affect score, stress, tiredness, sickness and reaction time) during the day. It remains open if they relate to another behaviour at all. Therefore, further studies should look for the variables affected by these variabilities across time.

Additionally, as the current study is a single-case study, it is plausible that our results are highly biased, specifically considering that our participant fell outside of his normative group on several sleep-related variables. Moreover, the effects of gender and age cannot be assessed here. Further studies are needed to determine whether the same results (including the factor of time and variability of the features) are obtained with different populations and in different parts of the world. In order to do that, it would be advisable to consider this study as a pilot

and replicate the methodology proposed here but optimizing certain aspects. For instance, behavioural measurements, which were recorded on a daily basis, could instead be measured on a weekly basis.

A meta-analysis by Gradisar, Gardner and Dohnt (2011) summarized the average sleep patterns for each age range but also classified it by region, country and city. Future studies could compare EEG during sleep and during wakefulness (repeated times) in order to look for consistency within participants. By taking all of these variables into consideration, the development of personalized analysis, diagnostics and treatments are possible. Through their high specificity, individualized interventions would benefit the effectiveness of treatments and could thereby lead to an improvement in life quality.

Future directions

Having sleep features extracted objectively from the sleep EEG recordings (such as sleep efficiency), instead of solely relying on self-report measures, allowed us to successfully look into the relationship between these concepts and other variables such as self-reported tiredness and sickness. Moreover, further analysis could establish causal relations between these variables in order to predict sleep quality based on behavioural variables and vice versa. This could also be achieved by applying specific interventions (such as changing the sleep schedule or length) in order to see the consequences in the behavioural measurements. By doing this, directional relations could be assessed and used for diagnosis and treatments. This approach could allow using sleep to improve behaviour and vice versa.

Additionally, the frequency analysis proposed in the current study could lead to the development of new analyses. With the data of the current study, the relationship between specific frequency bins and behavioural data can be addressed by using the time as a variable factor. To establish causal relationships, future studies could assess whether behavioural interventions can modify these changes in the frequency power to improve sleep quality and therefore, quality of life. Although several software tools are already available for the pre-processing and statistical analysis of neural data, new tools are also being developed. The data provided from the current study can also be used to guide the validity and reliability tests of measures on brain structure and function obtained through newly developed tools. Ultimately, by analysing the relation between sleep quality and mood reports with sufficient data opens the potential for the development of algorithms to predict mood states based on sleep recordings, and vice versa. Finally, the data and results collected and analysed in the current project will be part of an online open-source data. This will allow researchers to validate our analyses as well as perform additional ones.

Conclusions

This study investigated vital subjective and objective parameters of sleep and behaviour from the most extensive longitudinal study of this kind. Sleep quality and quantity were compared to a normative group indicating that our participant revealed similar length of sleep stages but different levels of sleep efficiency and duration of sleep.

The relation between sleep quality and several variables over the course of one year (i.e., physiological and behavioural) showed how sleep efficiency predicted decreased feeling of sickness and a positive mood. Besides, it was revealed that higher reaction times predicted feeling sicker and higher duration of sleep predicted a decreased feeling of tiredness. The current study constitutes the first attempt to create a dense longitudinal sleep phenotype and adds data to the growing open database of the analysed study that can be used for more future interdisciplinary studies.

The current study provides a one-year dataset to enrich this idea by creating a longer electrophysiological "fingerprint". Therefore, ideas related to how consistent are certain electrophysiological features, that were based on theory, can now be assessed empirically. The dependency on behavioural reports can now be modified and complemented with EEG data providing richer and well-founded information. The current study provides new findings about sleep frequency analysis by revealing that certain frequencies vary more than others (2-4 and 12-14 Hz fluctuate more than the others across time) in one individual during the course of a year and in a habituated sleep situation.

References

Akashiba, T., Kawahara, S., Akahoshi, T., Omori, C., Saito, O., Majima, T., & Horie, T. (2002). Relationship between quality of life and mood or depression in patients with severe obstructive sleep apnea syndrome. *Chest*, *122*(3), 861-865.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., ... & Balkin, T. J. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent

recovery: A sleep dose-response study. *Journal of sleep research*, *12*(1), 1-12.

Berry, R. B., Brooks, R., Gamaldo, C. E., Harding, S. M., Lloyd, R. M., Marcus, C. L., & Vaughn, B. V. (2015). The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications, version 2.2. *American Association of Sleep Medicine*.

Boulos, M. I., Jairam, T., Kendzerska, T., Im, J., Mekhael, A., & Murray, B. J. (2019). Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. *The Lancet Respiratory Medicine*, *7*(6), 533-543.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non□ clinical sample. *British Journal of Clinical Psychology*, *43*(3), 245-265.

De Gennaro, L., Ferrara, M., Vecchio, F., Curcio, G., & Bertini, M. (2005). An electroencephalographic fingerprint of human sleep. *NeuroImage*, *26*(1), 114–122.

De Zambotti, M., Goldstone, A., Claudatos, S., Colrain, I. M., & Baker, F. C. (2018). A validation study of Fitbit Charge 2TM compared with polysomnography in adults. *Chronobiology International*, *35*(4), 465–476.

Dinges, D. F., Pack, F., Williams, K., Gillen, K. A., Powell, J. W., Ott, G. E., ... & Pack, A. I. (1997). Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night. *Sleep*, *20*(4), 267-277.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655.

Filevich, E., Lisofsky, N., Becker, M., Butler, O., Lochstet, M., Martensson, J., … & Kühn, S. (2017). Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. *BMC Neuroscience*, *18*(1), 65.

Finelli, L., Achermann, P., & Borbély, A. A. (2001). Individual "Fingerprints" in Human Sleep EEG Topography. *Neuropsychopharmacology*, *25*(5), S57–S62.

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, *115*(27), E6106–E6115.

Fitbit Official Site for Activity Trackers and More. (2019). Retrieved 18 July 2019, from https://www.fitbit.com/nl/home

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., … & Dosenbach, N. U. F. (2017). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791–807.e7.

Graeber, R. C., Rosekind, M. R., Connell, L. J., & Dinges, D. F. (1990). Cockpit napping. Retrieved from https://ntrs.nasa.gov/search.jsp?R=19910043203

Gradisar, M., Gardner, G., & Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: a review and meta-analysis of age, region, and sleep. *Sleep medicine*, *12*(2), 110-118.

Johns, M. W. & Doré, C. (1978). Sleep at home and in the sleep laboratory: Disturbance by recording procedures. *Ergonomics*, *21*(5), 325-330.

Keil, A., Stolarova, M., Heim, S., Gruber, T., & Müller, M. M. (2003). Temporal stability of high-frequency brain oscillations in the human EEG. *Brain Topography*, *16*(2), 101-110.

Laumann, T. O., Gordon, E. M., Adeyemo, B., Snyder, A. Z., Joo, S. J., Chen, M.-Y., … & Petersen, S. E. (2015). Functional system and areal organization of a highly sampled individual human brain. *Neuron*, *87*(3), 657–670.

Leibenluft, E., Albert, P. S., Rosenthal, N. E., & Wehr, T. A. (1996). Relationship between sleep and mood in patients with rapid-cycling bipolar disorder. *Psychiatry research*, *63*(2-3), 161-168.

Loh, S., Lamond, N., Dorrian, J., Roach, G., & Dawson, D. (2004). The validity of psychomotor vigilance tasks of less than 10-minute duration. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 339–346.

Mirchi, N., Betzel, R. F., Bernhardt, B. C., Dagher, A., & Mišic, B. (2019). Tracking mood fluctuations with functional network patterns. *Social Cognitive and Affective Neuroscience*, *14*(1), 47–57.

Mischel, W., Ayduk, O., Berman, M. G., Casey, B. J., Gotlib, I. H., Jonides, J., … & Shoda, Y. (2011). 'Willpower' over the life span: decomposing self-regulation. *Social Cognitive and Affective Neuroscience*, *6*(2), 252–256.

Mitterling, T., Högl, B., Schönwald, S. V., Hackner, H., Gabelia, D., Biermayr, M., & Frauscher, B. (2015). Sleep and Respiration in 100 Healthy Caucasian Sleepers—A Polysomnographic Study According to American Academy of Sleep Medicine Standards. *Sleep*, *38*(6), 867–875.

Ong, J. L., Lo, J. C., Patanaik, A., & Chee, M. W. L. (2019). Trait-like characteristics of sleep EEG
power spectra in adolescents across sleep opportunity manipulations. *Journal of Sleep Research*, e12824.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 1.

Philip, P., Taillard, J., Sagaspe, P., Valtat, C., Sanchez-Ortuno, M., Moore, N., ... & Bioulac, B. (2004). Age, performance and sleep deprivation. *Journal of sleep research*, *13*(2), 105-110.

Poldrack, R. A., Laumann, T. O., Koyejo, O., Gregory, B., Hover, A., Chen, M.-Y., …, & Mumford, J. A. (2015). Long-term neural and physiological phenotyping of a single human. *Nature Communications*, *6*(1), 8885.

Pottkämper, J. (2019) *The role of rapid eye movement sleep in emotional schema processing* (Master's thesis).

Rasch, B. & Born, J. (2013). About Sleep's Role in Memory. *Physiological Reviews*, *93*(2), 681–766.

Šušmáková, K. (2004). Human sleep and sleep

EEG. *Measurement science review*, *4*(2), 59-74.

Thomsen, D. K., Mehlsen, M. Y., Christensen, S., & Zachariae, R. (2003). Rumination—relationship with negative mood and sleep quality. *Personality and Individual Differences*, *34*(7), 1293-1301.

Wang, J. Y., Weber, F. D., Zinke, K., Inostroza, M., & Born, J. (2018). More effective consolidation of episodic long-term memory in children than adults—Unrelated to sleep. *Child development*, *89*(5), 1720-1734.

Weber, F. D. (2013). SpiSOP tool(box). https://www.sipsop.org

Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*.

Appendix

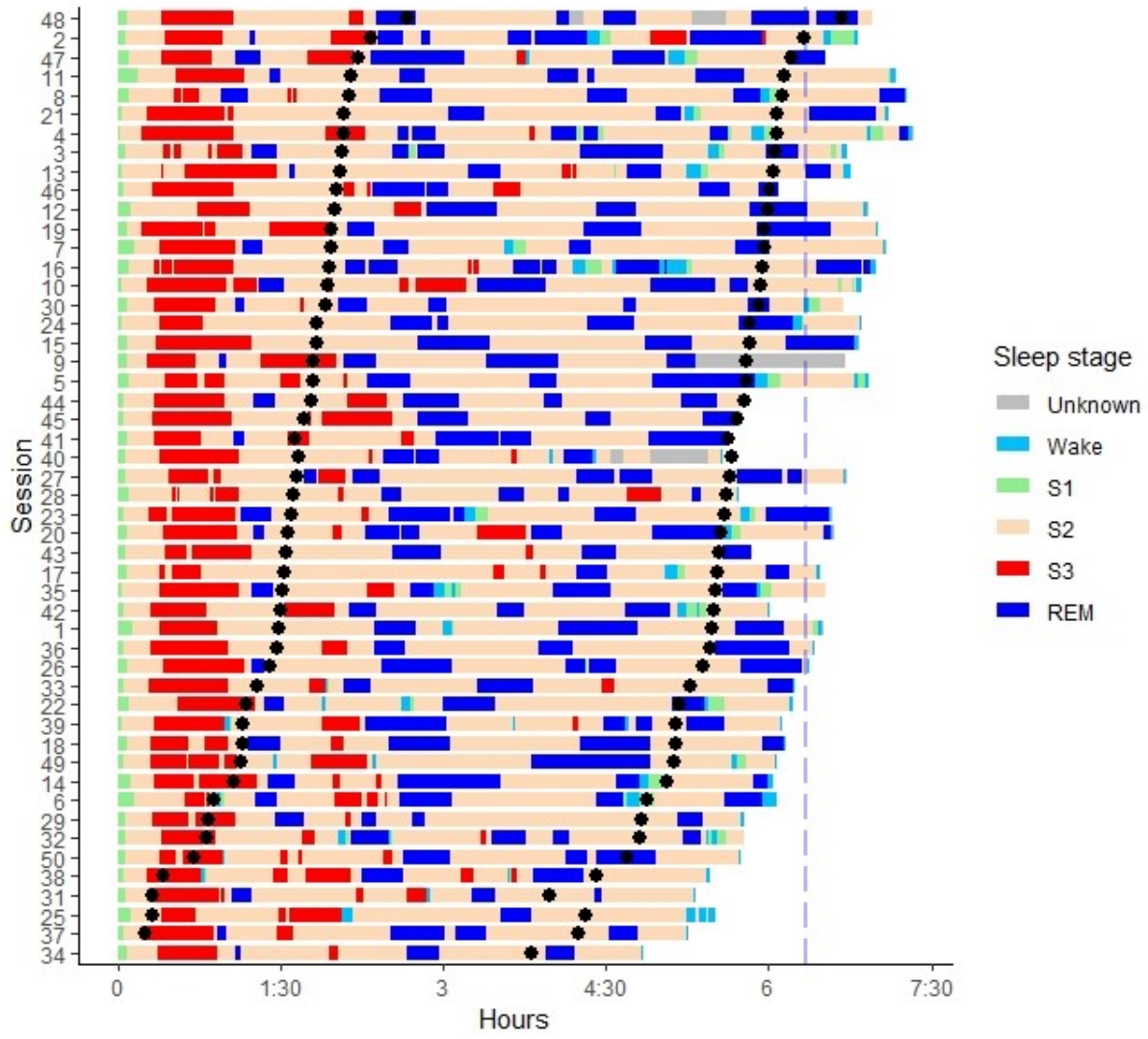

Figure A1 above, A2 below.