How Sleep and Caffeine Consumption Affect Each Other Over Time: an Experience Sampling Study

Iris D. Bouwer (s2831953)

Department of Psychology, University of Twente

202000375: Bachelor Thesis Health Psychology and Technology

First Supervisor: Dr. Ten Klooster, P.M.

Second Supervisor: Dr. Drossaert, C.H.C.

July 1, 2024

Abstract

To explore how sleep quality and quantity and caffeine consumption are related to each other over time, an experience sampling method (ESM) study with 21 participants (*M*age = 28.11, *SD*age = 14.54; 55% female, 45% male) was performed over 14 days with the use of the app SEMA3. One morning survey measuring sleep and three surveys assessing caffeine intake at semi-random time points in the morning, afternoon and evening, were prompted each day. Longitudinal associations were analyzed using repeated measures linear mixed modelling. It was found that hours of sleep were significantly negatively associated over time with total caffeine consumption the day after [$\beta = -0.30$, $SE = 0.09$, $t = -3.17$, $p = .002$], but not with total caffeine consumption the day before [$\beta = -0.08$, $SE = 0.05$, $t = -1.48$, $p = .140$]. For coffee consumption only, sleep quantity was found to be negatively associated both with consumption the day before [$\beta = -0.19$, $SE = 0.09$, $t = -2.03$, $p = .044$] and consumption the day after [$\beta = -0.13$, $SE = 0.06$, $t = -2.12$, $p = .035$]. No significant associations over time were found between total caffeine or coffee consumption and sleep quality. Overall, this study suggested a reciprocal negative relationship between higher caffeine consumption and lower sleep quantity over time, particularly for coffee consumption. The lack of association with sleep quality could be due to sleep quality being a more subjective rating of sleep compared to sleep quantity.

Key words: Experience Sampling Method (ESM) study, caffeine consumption, coffee consumption, sleep quality, sleep quantity

How Sleep and the Intake of Caffeine Affect Each Other

Some people are full of energy in the morning after seven hours of sleep (morning people), while others would rather stay up late and sleep for at least nine hours before getting up (evening people). The difference in the amount of sleep someone needs and how active one is during a certain point of the day depends, among other things, on the cardician rhythm of that individual. Ciccarelli and White (2021) defined the cardician rhythm as a cycle of bodily rhythm occurring over a period of 24-hours. The circadian sleep-wake cycle of humans, as for most other species, revolves around taking care of our needs when it is light and sleeping when it is dark. However, it is not always possible to stick to one's natural sleep-wake cycle, as many factors nowadays can disrupt this, such as having a jet lag, doing shift work or taking an all-nighter to finish an important task, which could have huge impacts on one's functioning in daily life (Lubbe et al., 2019).

In general, adults need between 7 and 8.5 hours of sleep to fully restore for the next day (Kumar, 2008). Such a typical night of sleep is characterized by a cycle of three stages: the REM (Rapid Eye Movement) stage and two stages of NREM (Non-Rapid Eye Movement) sleep (Kapsi & Drigas, 2020). Going through these cycles while sleeping is important because it may help to process and regulate emotions that have been experienced throughout the day, resulting in a positive contribution to one’s well-being. Additionally, sleep is necessary for cognitive performances and plays an important role in the encoding and storage processes of memories and information. Furthermore, it facilitates physical restoration processes in the body, promoting one’s physical health overall and has a positive influence on metabolic health (Vyazovskiy, 2015). Still, many people worldwide seem to have sleep problems. As found by Léger et al. (2008), 56% of the Americans, 31% of Western Europeans and 23% of Japanese have sleeping problems, creating a big impact on how they function in daily life. Additionally, half of the individuals have not taken action to resolve their sleep problems or either got used to or do not know about having sleep difficulties (Léger et al., 2008).

Sleep is usually measured by looking at both its quality and quantity, using various tools and methods. An objective measurement of sleep is polysomnography (PSG), in which sleep is measured in a sleep laboratory, resulting in a detailed picture of sleep stages and disorders by monitoring physiological parameters like eye movements (EOG) and brain waves (EEG). In contrast, an accessible but less accurate way to objectively measure sleep, is through wearable devices or sleep apps in which heart rate or movements are registered (Van de Water et al., 2011). Yet, most research has been based on cross-sectional retrospective survey studies or with longitudinal sleep diaries, which subjectively measure sleep. However, to be successful, the diary studies heavily depend on having participants immediately fill in the diary after waking up, which not everyone might be consistent with. Therefore, standardized self-reports might be more useful as they can be widely used across different populations and in several settings as they are low cost, straightforward, and no supervision is needed (Fabbri et al., 2021).

Experience Sampling Method (ESM) or Ecological Momentary Assessment (EMA) is another subjective way to intensively measure real-life experiences and behaviors over time. ESM uses self-reports to assess various daily life experiences and behaviors, such as the quality and quantity of one's sleep. To capture these experiences, participants receive questionnaires at certain times during the day over a period of days or weeks on their smartphone, which allows researchers to gain insight into the experiences and emotions of individuals at fixed moments in time (Myin-Germeys & Kuppens, 2022). According to Trull and Ebner-Priemer (2009), a few advantages of using ESM studies are high compliance rates and being able to capture momentary changes. The results often provide a detailed and valid picture of the behaviors, experiences, thoughts or symptoms of individuals in their natural environment. Another advantage is that ESM studies reduce recall bias, because of the collection of real-time data. Collecting data immediately and having straightforward and simple questions, results in less memory distortions, which gives more accurate data (Trull & Ebner-Priemer, 2009). The representativeness of ESM studies is supposed to be high as real-world experiences are measured, but the predictive validity depends on whether experiences that are captured in the moment can be generalized to bigger situations or population groups (Myin-Germeys & Kuppens, 2022).

Several studies have already used ESM to study sleep patterns and correlates of sleep over time. For instance, Triantafillou et al. (2019) conducted a six-week ESM study for which they recruited 208 participants in four equal groups: depressed or anxious only, depressed and anxious, and a control group. Each day, they were asked to report their sleep patterns and mood on their phone. A significant association was found between sleep quality and mood, with the effect of mood on sleep quality being lower than sleep quality on mood.

Many factors can influence sleep or how one is affected by sleep during the day. Even though one's biological rhythm automatically prepares to go to sleep or is still in the process of waking up in the morning, there are certain ways that help one to wake up or stay awake longer. One behavior often assumed to affect sleep (or wakefulness) is through a mild stimulant that is one of the most used worldwide: caffeine. Caffeine is found in many products such as tea, soda, chocolate, energy drinks and coffee (Whalen et al., 2008), and is therefore used by many individuals across the globe (Bolton & Null, 1981), with 75-98% of the youth consuming at least one caffeine containing beverage a day. Coffee, in particular, is the main source of caffeine intake in the United States and Western Europe (Clark & Landolt, 2017) and compared to drinking a cup of tea or soda, coffee contains twice to four times the amount of caffeine (Mitchell et al., 2014). This can have a bigger influence on the reaction time, arousal and attention of the body, which in turn, can affect the quality and duration of sleep more (Whalen et al., 2008).

The intake of caffeine about one hour before bedtime can inhibit the deeper sleep stages and disturb the REM stage (Bolton & Null, 1981). However, people who take in a lot of caffeine throughout the day are more tolerant for these effects and will therefore have less issues with sleep disturbances. As caffeine stays in the body for about three to four hours, with the first effects arising within one hour after consumption, sleep quality can decrease when drinking a cup of coffee before bedtime. Caffeine can cause one to feel less tired and increases attention and clear thinking (Bolton & Null, 1981). This is beneficial when someone wants to increase their performance during prolonged wakefulness or after having a poor night of sleep (Roehrs & Roth, 2008). However, it can also result in the occurrence of insomnia, due to the stimulation of the central nervous system (Bolton & Null, 1981).

As caffeine is similar in structure to the nucleoside adenosine, it can bind to and block the receptors that usually respond to adenosine. Therefore, most of the biological effects caused by caffeine arise because it antagonizes all types of adenosine receptors (ARs). Where adenosine exerts its effects on the glial cells and certain neurons in all areas of the brain, caffeine influences certain brain functions as well, such as cognition and sleep. Besides, reaching the threshold of about 400mg of caffeine a day for healthy adults can increase physical effects including nervousness, increased urination, muscle tremors, restlessness, and insomnia. This is, among other things, due to the effect of caffeine on the central nervous system and its stimulating effects disrupting the sleep-wake cycle of an individual. This inhibits the relaxation and sleep effects of adenosine and makes it more difficult to fall asleep or stay asleep. It specifically reduces REM sleep, crucial for cognitive function and restorative rest, which leads to one feeling more tired when waking up in the morning (Rodak et al., 2021).

Among the ESM studies that have focused on sleep, most have concentrated on how sleep influences mood or one's well-being, specifically in people with certain conditions, such as depression or anxiety. A few studies have investigated how sleep and caffeine consumption are related to each other over time, specifically how caffeine consumption can influence one's sleep. For example, one EMA study by Whalen et al. (2008) in youth with a major depression disorder (MDD), found that youth with MDD reported more caffeine use and sleep problems and that caffeine use among youth with MDD decreased across treatment, but sleep complaints remained elevated. Though, no ESM studies have been found that focused on the relationship between sleep and caffeine consumption the following day. Although a few articles, like Roehrs and Roth (2008), state that sleep disturbances lead to one feeling sleepier the next day, which can consequently result in an increase of caffeine consumption.

As, to date, not many studies have specifically examined how sleep and caffeine consumption affect each other over time, an ESM study was conducted over the course of two weeks to get a better understanding of how sleep is associated with caffeine consumption the day before and the day after. The study's aim was to get more insight into how the total consumption of caffeine and the duration and quality of sleep affect each other in everyday life. As coffee is one of the most consumed stimulants worldwide and often contains more caffeine than most beverages (Mitchell et al., 2014), the consumption of this specific variable is compared to sleep duration and quality as well. It was hypothesized that both total caffeine and total coffee consumption are reciprocally related to sleep quality and quantity. This was examined by addressing the following research questions:

1. What is the relationship between the quantity and quality of sleep and the consumption of caffeine the next day?
2. What is the relationship between caffeine consumption on the day before and the quantity and quality of sleep the following night?
3. What is the relationship between the quantity and quality of sleep and the consumption of coffee the next day?
4. What is the relationship between coffee consumption on the day before and the quantity and quality of sleep the following night?

Methods

Design

This study entailed an ESM study design to explore how sleep and caffeine are associated with each other over time. In this intensive longitudinal design, participants were asked to fill in two surveys at the start of the study once and four short surveys each day for two weeks: one survey in the morning regarding the quantity (or hours of sleep) and quality of their sleep the past night and one survey, which focused on caffeine consumption and energy levels, on three semi-random times throughout the day. The caffeine consumption included coffee, tea, soda and energy drinks. Next to the total caffeine consumption, coffee was analyzed separately regarding sleep quality and quantity. The study period was the same for each participant as the data collection started on Wednesday, the 11th of April and ended on Thursday, the 25th of April. This study has been approved by the BMS Ethical Committee of the University of Twente (request number: 240445).

Participants

The participants were recruited through convenience sampling in the researcher's network and snowball sampling. The aim was to recruit at least 20 people to participate in the study as this corresponds to the medium number of participants in a previous review of ESM studies by Berkel et al. (2017). The inclusion criteria of the study were for the participants to have a sufficient level of English, be 18 years of age or older, and drinking at least one cup of caffeine a day, either coffee, tea, soda or energy drinks. A total of 23 people volunteered to participate and fit with the inclusion criteria. Each participant was able to withdraw from the study at any moment.

ESM Protocol

The study had a duration of two weeks, which is common practice for ESM studies (Berkel et al., 2017). A 14-day study allows for an even distribution of the days of the week, as days of the week can make a difference in the amount of sleep someone gets or how much caffeine someone consumes. Participants with a response rate lower than 30% were excluded from the study (Myin-Germeys & Kuppens, 2022). The decision to have four surveys each day, is because Morren et al. (2009) found that compliance rates were higher with shorter diaries. It allows participants to quickly fill in the surveys after receiving the notification, lowering the number of careless responses and the possibility of participants dropping out as the burden might be perceived as higher. The scales of the slider questions that ranged from 0-10 were self-composed, partly based on the continuous slider scale of Myin-Germeys and

Kuppens (2022) that ranged from 0 to 100. Even though Likert scales ranging from 1 to 7 are frequently used in ESM research, it was impractical to use it as a scale to measure units of caffeine consumed, as it would have been rare that all participants would have had at least one unit of each type of caffeine before each time slot. The same scale was used for the other variables to make it more straightforward for the participants.

Materials

To conduct this ESM study, the SEMA3 researcher dashboard and associated smartphone app was used (O'Brien et al., 2023). A total of four different surveys were programmed on the dashboard. The first survey consisted of three questions, explaining the study and asking for informed consent from the participants after signing up to the study in the app (Appendix A). The second survey entailed questions regarding demographics and general information about average, daily caffeine consumption (the baseline survey), as found in Appendix B. These two surveys were triggered on the first day and second day of the study. The last two surveys were focused on the data needed to answer the research questions of this study. One was the 'sleep survey' (Appendix C) appearing at a random time in the morning. The sleep survey contained three short questions, consisting of one open question to fill in the number of hours one had slept the past night and two slider questions to rate one's sleep and energy level in the morning on a scale of 1 to 10. Finally, caffeine consumption was assessed with a 5-item questionnaire (Appendix D), which appeared at three semi-random timeslots each day. The five questions of the daily survey regarding caffeine intake were all slider questions with a scale of 1 to 10. These questions focused on five different variables: sleepiness, coffee, tea, soda, and energy drinks. RStudio was used to calculate the total daily mean value for each of the variables. These values, except for sleepiness, were put together for a total daily value for caffeine.

Procedure

Two days before the start of the study, all participants were invited via the SEMA3 website, by entering their name and email address, and adding a start and end date for the data collection. The data collection was done over two weeks. Each participant started on April 11 (00:00) and received the last surveys on April 25 (before 23:59). After inviting the participants, each of them received an email with their participant number and instructions of how to install the SEMA3 app. As a researcher, it was not possible to identify the participants anymore after inviting them, as the information changed into an anonymized personal participant number.

The participants were asked to fill in six surveys in total in SEMA3. After the start of the study, participants had two full days to fill in the informed consent survey and the baseline questions. The other four surveys appeared daily at semi-random moments and took about one minute to complete. The first one, 'the sleep survey' appeared at a random time between 7am and 10am, with an expiry time of 240 minutes (until 2pm). The survey regarding the morning and evening questions appeared at random times of three timeslots throughout the day, each with an expiry time of 30 minutes. The morning timeslot was set between 10am and 11am, the afternoon timeslot between 3pm and 4pm, and the evening timeslot between 8pm and 9pm.

Data Analysis

The final data was exported as a CSV file from SEMA3 into the statistical program RStudio for further analysis (Appendix E). Statistical tests were considered to be significant at an alpha level of .05 or lower. A total of three participants were excluded from the study as their response rate was below 30%. Therefore, final analyses were conducted with a sample of 21 participants, who had an average response rate of 68.10% (*SD*= 18.48, range = 35% - 95%).

To get an overview of the data, descriptive statistics were computed for all variables of interest. A daily total value was computed for the number of coffee, tea, soda, and energy drinks units that were consumed in the morning, afternoon, and evening for all participants. A daily total coffee consumption score was also computed.

The longitudinal associations between daily caffeine or coffee consumption and sleep quality and quantity scores were first visualized as mean scores over time using the ggplot2 package. To statistically test the research questions, random intercept Linear Mixed Model (LMM) analyses were performed with the NLME package. LMM has the advantage that it can take into account the hierarchical structure of the data with measurements nested within participants, and that missing values (N/As) on the dependent data can be handled in the model, using maximum likelihood estimation (Myin-Germeys & Kuppens, 2022). Restricted maximum likelihood estimation was used with an unstructured covariance matrix for the random effects. Eight separate LMMs were performed. The first two models examined the association between total caffeine consumption (DV) and either sleep quality or quantity the next day as fixed covariate. To examine whether caffeine consumption was associated with sleep quantity and quality the next day, caffeine consumption was lagged by one day for each participant. Next, two LMMs with lagged caffeine consumption as fixed covariate and either sleep quantity or quality were performed. The same four analyses were done for coffee only consumption.

Results

Sample and Descriptives

Among the 21 participants included in the final analysis (*M*age = 28.11, *SD*age = 14.54; 55% female, 45% male) were 12 Dutch and 4 Spanish participants, with the other three participants being from Slovakia, Peru, and Argentina. Two participants did not answer the baseline questions, but were kept in the study, as they gave consent and had a response rate above 30%.

The results of the baseline questions regarding sleep quality and quantity showed that almost half of the participants ($n = 10$) needed around 16 to 30 minutes to fall asleep and that around three quarter of the participants ($n = 15$) slept an average of 9-10 hours each night. Among the four types of caffeine holding drinks, tea and coffee were the most consumed stimulants, whereas soda and energy drinks were the least consumed stimulants, with a mean value of 1 to 1.5 lower than found for coffee and tea (Table 1).

Regarding the outcomes of the daily questionnaires, the results of the sleep survey showed a mean of 7.74 ($SD = 1.31$) hours for sleep quantity, a mean value of 7.06 ($SD = 1.43$) for sleep quality, and a mean energy level of 6.39 ($SD = 1.37$).

The mean values for each variable of the three semi-random timeslot surveys during the day are found in Table 2. For sleepiness, people tended to feel slightly less energetic at the end of the day, with a mean difference of 0.42 between the morning and evening surveys. Coffee was the most consumed type of caffeine in the morning but peaked with a mean of 0.57 in the afternoon. Tea was the most consumed type of all four stimulants, peaking in the afternoon with a mean of 0.63. Soda and energy drinks were barely consumed in the morning, but consumption increased slightly, especially soda increasing by 0.25 between morning and evening.

Table 1

Descriptive results of the baseline questions for the descriptives and different types of caffeine consumed in units (n=21).

	Mean	SD	Minimum	Maximum
Coffee	1.47	2.04	0	8
Tea	1.47	1.58	0	5
Soda	0.47	0.77	0	3
Energy Drinks	0.05	0.23	0	1

Table 2

Mean units of caffeine consumed and mean value of sleepiness of the daily semi-random questions (n=21).

	Mean Morning	Mean Afternoon	Mean Evening
Sleepiness	6.64	6.64	6.22
Coffee	0.49	0.57	0.28
Tea	0.37	0.63	0.55
Soda	0.01	0.16	0.26
Energy Drinks	0.00	0.02	0.03

Daily Caffeine Consumption and Sleep

In Figure 1A and 1B the total amount of caffeine consumed during each day of the study is visualized (see the bottom line in both figures). The total amount of caffeine is the mean of the sum of the total amount of daily coffee, tea, soda, and energy drinks consumption. The participants drank an average of 2.04 units of caffeinated drinks each day. The consumption is highest on day 1 (Wednesday), day 5 (Sunday), day 8 (another Wednesday), and day 14 (Tuesday). On day 3 (Friday), day 6 (Monday), and day 13 (another Monday), consumption is lowest.

Relation Between Daily Caffeine Consumption and Sleep Quantity

In Figure 1A the total daily number of drinks containing caffeine is visualized against the average amount of hours the participants slept each previous night. A relational pattern can be seen here, in that on days participants had less hours of sleep, they tended to consume more caffeine the next day. On days they did sleep a fair number of hours, the caffeine consumption was lower. Specifically, on day 3 the average amount of hours of sleep is highest, whereas the lowest amount of caffeine consumption is on day 6. On both day 1 and 5, the average amount of sleep is lowest, and the amount of caffeine consumed is highest.

The LMM analysis with daily sleep as an independent variable confirmed that the average hours of sleep at night were significantly and negatively associated with caffeine consumption the next day [$\beta = -0.30$, $SE = 0.09$, $t = -3.17$, $p = .002$]. Each additional hour of sleep was associated with 0.30 less units of caffeine the following day.

The LMM model in which caffeine consumption was lagged with one day to examine the relationship between caffeine consumption the day before and hours of sleep the night after did not show a significant association [β = -0.08, *SE* = 0.05, *t* = -1.48, *p* = .140], indicating that total caffeine consumption was not associated with sleep quantity the night after.

Relation Between Caffeine Consumption and Sleep Quality

The results of the daily total amount of caffeine consumed and the quality of sleep are visualized in Figure 1B. Compared to caffeine consumption and hours of sleep, and sleep quality and caffeine consumption, no clear association pattern between sleep quality and caffeine consumption the next day was visible. The LMM analysis for the quality of sleep and total caffeine consumption the next day showed a slightly positive, but non-significant association over time [β = 0.02, *SE* = 0.07, *t* = 0.25, *p* = .802]. The lagged LMM model, indicating the relation between coffee consumption and sleep quality the next night, showed a similar non-significant relation [β = 0.01, *SE* = 0.06, *t* = 0.26, *p* = .796].

Daily Coffee Consumption and Sleep

The average daily total coffee consumption and daily sleep across the timespan of 14 days is visualized in Figure 2A and 2B (see the bottom line in both figures). The mean of total coffee consumption was 0.81 units per day, which is 0.66 units less than found in the baseline questions (Table 1). The highest means of coffee consumption were observed on day 5 (a Sunday), day 8 (a Wednesday), and on day 14 (a Tuesday). On day 7 (a Tuesday) and 11 (a Saturday) participants drank the least amount of coffee.

Relation Between Daily Coffee Consumption and Sleep Quantity

In Figure 2A the average total amount of coffee and the average total hours of sleep for each day are visualized. On a few days similar patterns as with caffeine consumption were seen, with more coffee consumption on days one had less hours of sleep and vice versa.

The LMM analysis confirmed that the average hours of sleep at night were indeed significant and negatively associated with coffee consumption the next day [β = -0.13, *SE* = 0.06, *t* = -2.12, *p* = .035]. On average, this association did however appear weak, as each hour of sleep more was associated with around 0.13 less units of coffee the next day. As opposed to total caffeine consumption, the one-day lagged model of coffee and hours of sleep showed that hours of sleep were also significant and negatively associated with coffee consumption the previous day [β = -0.19, *SE* = 0.09, *t* = -2.03, *p* = .044].

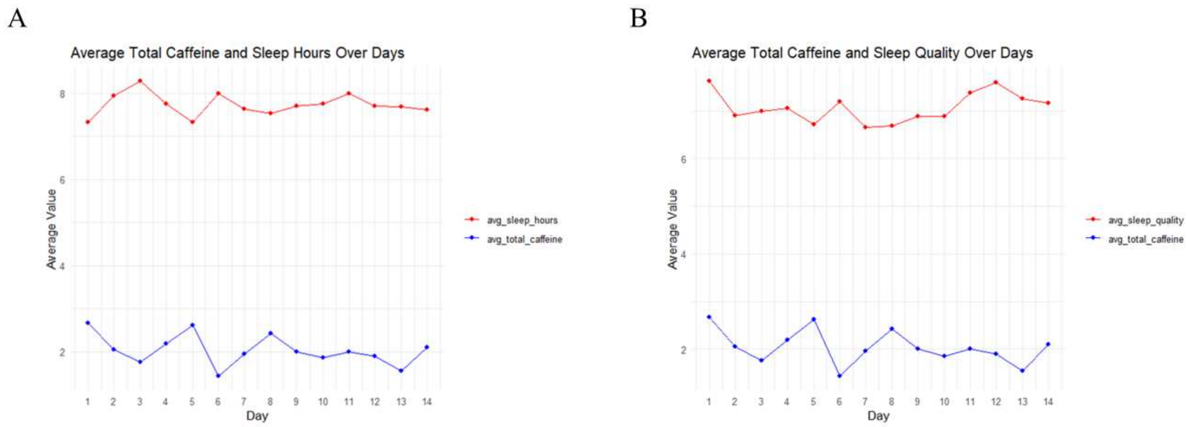
Relation Between Daily Coffee Consumption and Sleep Quality

Figure 2B shows the relationship between sleep quality and coffee consumption. For some days a clear difference is seen between a bad quality of sleep and an increase in coffee consumption or vice versa. However, most days do not give very clear distinctions between the variables, as seen with sleep quantity.

LMM analysis between sleep quality and coffee consumption the following day showed a negative but non-significant relation [β = -0.03, *SE* = 0.45, *t* = -0.58, *p* = .562]. The lagged model also showed a negative and non-significant output regarding sleep quality and coffee consumption the day before [β = -0.28, *SE* = 0.09, *t* = -0.30, *p* = .763].

Figure 1

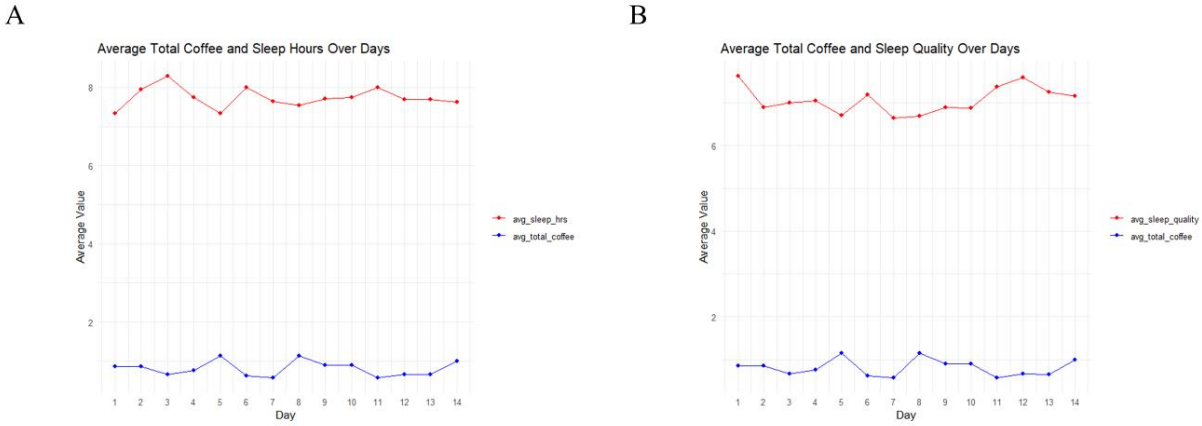
Average number of total consumed units of caffeine and sleep quantity and quality for each day of the study.



Note. The maximum amount of caffeine consumption is 4, whereas the maximum, average amount of hours of sleep is 10.

Figure 2

Average number of total consumed units of coffee and sleep quality and quantity for each day of the study.



Note. The maximum amount of caffeine consumption is 4, whereas the maximum, average amount of hours of sleep is 10.

Discussion

The aim of this ESM study was to get more detailed insight into the reciprocal relationship between sleep quantity and quality and the consumption of caffeine containing drinks over time. The results showed that there was a significant negative relationship between the quantity of sleep and total caffeine consumption the day after, but no relationship was found regarding caffeine consumption the day before and sleep quantity. For coffee consumption, both consumption the day before and the day after had a significant negative relationship with sleep quantity. No relationship was found between sleep quality and the amount of coffee or caffeine consumed the day before or after. Sleep quantity, but not sleep quality, thus seems to have a predictive relationship with caffeine consumption and a reciprocal relationship with coffee consumption.

Although no specific previous studies have focused on sleep quality and quantity and caffeine consumption both on the day before and after, based on literature it was expected that sleep and caffeine consumption would be reciprocally related to each other. This expectation was based on the earlier mentioned studies of Whalen et al. (2008), Bolton and Null (1981), and Rodak et al. (2021), in which caffeine was found to influence sleep. On the other hand, Roehrs and Roth (2008) stated that sleep disturbances often lead to people feeling sleepier the following day, which consequently led to an increase in caffeine consumption. Besides, O'Callaghan et al. (2018) introduced the term 'coffee cycle' in which people who feel tired in the morning increase their caffeine intake, which consequently leads to disturbed sleep patterns and feeling tired again the next day. Therefore, it was expected that sleep quality and quantity would have a reciprocal relationship with both caffeine and coffee.

However, sleep quality was found not to be significantly associated with the amount of caffeine or coffee consumed either the day before or after. One reason for this could be that sleep quality entails a more subjective rating, in which each participant might have had a different opinion on what determines the quality of their sleep. This rating can be based on how that person normally sleeps, such as waking up during the night or being restless. There is no standard of which rating applies to certain characteristics of how one slept (Kohyama, 2021). In comparison, sleep quantity is measured more straightforwardly by simply counting the hours of sleep. This might have created a difference in the outcomes of how sleep quantity and quality are related to caffeine and coffee consumption.

Interestingly, out of the four LMM analyses focusing on sleep quantity, only its relationship with total consumed caffeine the day before was not significant, whereas the relationship between caffeine consumption and sleep quantity the day after and the relationship between coffee consumption and sleep quantity the day before and after were found to be significant. A possible reason for this difference between total caffeine consumption and coffee consumption could be that, compared to coffee itself, the total caffeine calculation also included tea, soda, and energy drinks containing caffeine. As tea only contains a quarter to half of the amount of caffeine that coffee contains, the effect on sleep might have been lower. Besides, energy drinks and soda were barely consumed during the study. Even though energy drinks can contain the same or even higher amount of caffeine as coffee, it might not have been consumed enough in the current sample to affect sleep. Moreover, just like tea, soda contains about one fourth to half of the amount of caffeine found in coffee (Mitchell et al., 2014). Additionally, during this ESM study, coffee was the most consumed stimulant, and it was mostly consumed in the morning only, whereas tea was mostly consumed in the afternoon and evening. This corresponds with a study from Pelau and Radulescu (2021) in which they found that people mostly drink coffee in the morning and during lunch time but tend not to drink coffee anymore after lunch or dinner, and at night. Instead, people prefer to drink warm milk or tea before bedtime as it has a calming effect, especially for those with insomnia (Pelau & Radulescu, 2021). People in the current study might therefore have purposely chosen to only drink coffee in the morning and tea in the evenings, which might have had less influence on their sleep quantity. Besides, some people might have built up a tolerance for the effects of caffeine, allowing them to consume caffeine a few hours before bedtime, without their bodies being as sensitive to the effects of caffeine anymore, resulting in these people possibly having less sleep problems during the night (Bolton & Null, 1981). Tolerance, low consumption, and drinking certain types of caffeine at certain times of the day, might therefore explain the non-significant outcome.

However, this does not yet explain why the relation between sleep quantity and caffeine consumption the day after was significant. One reason for this could be that, as pointed out by Bolton and Null (1981), regardless of how one slept, caffeine consumption is a stable habit for many people. They found that habitual coffee drinkers who did not have their usual morning coffee, experienced headaches and nervousness. Additionally, O'Callaghan et al. (2018) highlighted that medical professionals who felt tired in the morning increased their caffeine intake. Similarly, Mahoney et al. (2019) found that students drank caffeine to feel more awake during certain times of the day. However, caffeine consumption may then more likely be related to sleepiness in the morning instead of sleep quantity, which could be a point of focus in a future study. As stated before, not many ESM studies have focused yet on the relationship between sleep and caffeine consumption the following day. Still, the significant outcome between sleep quantity and caffeine consumption the day after, correspond with the expectation based on the article of Roehrs and Roth (2008), in which a disturbed night of sleep would lead to one feeling more tired the next day, which resulted in an increase in the consumption of caffeine. Overall, consumption habits, especially in the morning, disturbed sleep or the sleepiness of a person could explain the significance of sleep quantity and caffeine consumption the day after.

Strengths and Limitations

Although missing values were observed in the final dataset of the study, due to not getting notifications or not having time when the survey appeared, the overall response rate of 68.10% was still within the range of other ESM studies. A 6-day ESM study from Rintala et al. (2019) had a response rate of 70%, which was in line with other studies that had compliance rates within the range of 66% to 86%. Considering that their study was almost twice as short and compliance rates might have lowered with the increase of study days, the response rate of this current study appears very acceptable for a typical ESM study.

It should be mentioned that data on caffeine consumption might have been missed because of the specified sampling protocol, as the evening survey appeared between 8pm and 9pm. With the morning survey specifically asking about caffeine intake after waking up, potential caffeine intake between the evening survey and waking up has not been recorded. It would therefore be recommended to extend the time of the evening survey or to add a(n) (extra) survey in the morning to ask about caffeine consumption the past evening.

Many other factors that could have had an influence on the quantity and quality of sleep should also be taken into account. For example, Morioka et al. (2013) indicated that increased alcohol consumption led to more sleep disturbances. During social gatherings, participants could have consumed more (soda mixed with) alcohol, which might have resulted in a lower sleep quality and quantity the night after. This could have then resulted in an increase in caffeine consumption the next day. Another factor that could have influenced caffeine consumption could be the weather, as people tend to drink more warm drinks when it is cold and vice versa (Keleş et al., 2018; Yohannes & Matsuda, 2016). Moreover, an article by Mahoney et al. (2019) found that students consume caffeine mostly because it makes them feel more awake, it tastes good, or it comes with social aspects. These are a few factors that could have had an influence on caffeine consumption and hours of sleep.

A final interesting factor to mention is the measurement reactivity that may have been caused by the intensive measurement of caffeine consumption in this study. Measurement reactivity, a phenomenon for which French and Sutton (2010) found evidence that it can affect behavior, emotion or cognition, is a serious concern in ESM (Eisele et al., 2023). When comparing the baseline average of cups of daily coffee consumption with the average amount consumed during the study, participants appeared to drink less coffee during the study. Possibly because they were more aware of their consumption behavior during the study.

In turn, with this rather small sample, the study's generalizability should be considered carefully. In an article from Barone and Roberts (1984) a mean of 2.6mg/kg body weight of daily caffeine intake was found. These authors suggested that instant coffee contains about 60mg of caffeine, and tea about 30mg. O'Callaghan et al. (2018) found that 90% of United States citizens consume an average of 200mg of caffeine each day. The participants of these studies thus consumed at least three (coffee) to eight (tea) cups of caffeine each day. This is twice to four times the amount of what was consumed by the sample in the current ESM study. Thus, even though this study showed certain relationships between caffeine consumption and sleep, the low levels of coffee consumption may limit the findings' generalizability to the general population (Barone & Roberts, 1984). Besides, to refer back to the point of missing out on data after the evening survey, Myin-Germeys and Kuppens (2022) state that not having extra evening assessments could also affect the generalizability of the study and the final conclusions. Overall, the extent to which an ESM study can be generalized is complex, as it relies on its representativeness and its predictive validity (Myin-Germeys & Kuppens (2022).

This study had an acceptable compliance rate, but it might be recommended to extend the time of expiry of each survey. This allows participants to have more time to fill in the survey. Moreover, it allows the researcher to receive more data. As, for example, the evening survey might appear at 8.30pm, participants would still have an extra hour to answer the questions. This might also contribute to a previous suggestion to extend the time of the evening survey to not miss out on data after 9pm.

Recommendations

For future research, it could be interesting to dive more into caffeine consumption at different specific times of the day and how this might influence sleep quality or quantity. Based on the article of Rodak et al. (2021) caffeine consumption before bedtime should most likely result in lower sleep quality and quantity, as, for example, coffee needs at least four hours to break down and for the effects to wear off. Furthermore, it could be interesting to look at what days caffeine consumption is higher and why. As this study showed high levels of caffeine consumption on certain weekdays, especially Wednesdays. Additionally, an interview study could be done to get insight in why people drink certain types of caffeine on certain days or what their beliefs behind caffeine consumption are, regarding habits (Bolton & Null, 1981) or drinking certain types at certain times of the day (Pelau & Radulescu, 2021).

Besides, Rintala et al. (2019) found that participation compliance differs among gender, age, and time, with participants having a lower compliance during the weekend, and females and older people usually being more compliant. If a future researcher would like to know more about compliance of participants during an ESM study, it could be interesting to look more into when and which participants are most compliant and why.

Finally, relating back to the introduction, a study could be done regarding caffeine consumption among either morning or evening people, with the expectation that evening people who have to wake up rather early, drink more caffeine in the morning to stay awake during the day and morning people drinking more caffeine towards the end of the day when they have certain commitments in the evening. A similar study could be done focusing on specific jobs that may affect sleep, such as shift workers or people who are on call, and their caffeine consumption (Ciccarelli & White, 2021; Lubbe et al., 2019).

Conclusion

Overall, the findings of this ESM study suggest that sleep quantity and caffeine consumption may be reciprocally related over time. No significant relationship was found between sleep quality and caffeine consumption over time, but as this is one of the first ESM studies that has focused on the reciprocal relationship between caffeine consumption and sleep, it would be interesting to continue similar studies to see if there are comparable results with bigger sample sizes.

References

Barone, J. J., & Roberts, H. (1984). Human consumption of caffeine. In *Caffeine: Perspectives from recent research* (pp. 59-73). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69823-1_4

Berkel, N. van, Ferreira, D., & Kostakos, V. (2017). The Experience Sampling Method on Mobile Devices. *ACM Computing Surveys, 50*(6), 93:1-93:40. https://doi.org/10.1145/3123988

Bolton, S., & Null, G. (1981). Caffeine: Psychological effects, use and abuse. *Orthomolecular Psychiatry, 10*(3), 202-211. JOM_1981_10_3_09_Caffeine_Psychological_Effects_Use_and_Abuse.pdf (isom.ca)

Ciccarelli, S.K. & White, J.N. (2021). *Psychology, Global Edition* (6th ed.). Pearson.

Clark, I., & Landolt, H. P. (2017). Coffee, caffeine, and sleep: A systematic review of epidemiological studies and randomized controlled trials. *Sleep medicine reviews, 31*, 70-78. https://doi.org/10.1016/j.smrv.2016.01.006

Eisele, G., Vachon, H., Lafit, G., Tuyaerts, D., Houben, M., Kuppens, P., ... & Viechtbauer, W. (2023). A mixed-method investigation into measurement reactivity to the experience sampling method: The role of sampling protocol and individual characteristics. *Psychological Assessment, 35*(1), 68. https://psycnet.apa.org/doi/10.1037/pas0001177

Fabbri, M., Beracci, A., Martoni, M., Meneo, D., Tonetti, L., & Natale, V. (2021). Measuring subjective sleep quality: a review. *International journal of environmental research and public health, 18*(3), 1082. https://doi.org/10.3390/ijerph18031082

French, D. P., & Sutton, S. (2010). Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it?. *British journal of health psychology, 15*(3), 453-468. https://doi.org/10.1348/135910710X492341

Kapsi, S., Katsantoni, S., & Drigas, A. (2020). The Role of Sleep and Impact on Brain and Learning. *Int. J. Recent Contributions Eng. Sci. IT, 8*(3), 59-68. https://doi.org/10.3991/ijes.v8i3.17099

Keleş, B., Gómez-Acevedo, P., & Shaikh, N. I. (2018). The impact of systematic changes in weather on the supply and demand of beverages. *International Journal of Production Economics, 195*, 186-197. https://doi.org/10.1016/j.ijpe.2017.08.002

Kohyama, J. (2021). Which is more important for health: sleep quantity or sleep quality?. *Children, 8*(7), 542. https://doi.org/10.3390/children8070542

Kumar, V. M. (2008). Sleep and sleep disorders. *Indian Journal of Chest Diseases and Allied Sciences, 50*(1), 129. https://www.researchgate.net/publication/5241469_Sleep_and_sleep_disorders

Léger, D., Poursain, B., Neubauer, D., & Uchiyama, M. (2008). An international survey of sleeping problems in the general population. *Current medical research and opinion, 24*(1), 307-317. https://doi.org/10.1185/030079907X253771

Lubbe, R., Verwey, W. & Borsci, S. (2019). *Brain and Cognition*. Pearson.

Mahoney, C. R., Giles, G. E., Marriott, B. P., Judelson, D. A., Glickman, E. L., Geiselman, P. J., & Lieberman, H. R. (2019). Intake of caffeine from all sources and reasons for use by college students. *Clinical nutrition, 38*(2), 668-675. https://doi.org/10.1016/j.clnu.2018.04.004

Mitchell, D. C., Knight, C. A., Hockenberry, J., Teplansky, R., & Hartman, T. J. (2014). Beverage caffeine intakes in the US. *Food and Chemical Toxicology, 63*, 136-142. https://doi.org/10.1016/j.fct.2013.10.042

Morioka, H., Itani, O., Kaneita, Y., Ikeda, M., Kondo, S., Yamamoto, R., ... & Ohida, T. (2013). Associations between sleep disturbance and alcohol drinking: a large-scale epidemiological study of adolescents in Japan. *Alcohol, 47*(8), 619-628. http://dx.doi.org/10.1016/j.alcohol.2013.09.041

Myin-Germeys, I., & Kuppens, P. (2022). *The open handbook of experience sampling methodology: A step-by-step guide to designing, conducting, and analyzing ESM studies* (2nd ed.). REAL. https://ppw.kuleuven.be/okp/esmhandbook.php

O'Brien, S. T., Dozo, N., Hinton, J. D. X., Moeck, E., Susanto, R., Jayaputera, G., … Koval, P. (2023). SEMA3: A Free Smartphone Platform for Daily Life Surveys. PsxArXiv. https://doi.org/10.31234/osf.io/cnar8

O'Callaghan, F., Muurlink, O., & Reid, N. (2018). Effects of caffeine on sleep quality and daytime functioning. *Risk management and healthcare policy*, 263-271. https://doi.org/10.2147/RMHP.S156404

Pelau, C., & Radulescu, L. (2021). Coffee consumption patterns in a Hurried society. *Collaborative Research for Excellence in Economics and Social Sciences*, 240-248. http://doi.org/10.2478/9788366675322-028

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2019). Response compliance and predictors thereof in studies using the experience sampling method. *Psychological assessment, 31*(2), 226. https://psycnet.apa.org/doi/10.1037/pas0000662

Rodak, K., Kokot, I., & Kratz, E. M. (2021). Caffeine as a factor influencing the functioning of the human body—Friend or foe?. *Nutrients*, *13*(9), 3088. https://doi.org/10.3390/nu13093088

Roehrs, T., & Roth, T. (2008). Caffeine: sleep and daytime sleepiness. *Sleep medicine reviews, 12*(2), 153-162. https://doi.org/10.1016/j.smrv.2007.07.004

Triantafillou, S., Saeb, S., Lattie, E. G., Mohr, D. C., & Kording, K. P. (2019). Relationship between sleep quality and mood: ecological momentary assessment study. *JMIR mental health, 6*(3), e12613. https://doi.org/10.2196/12613

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: introduction to the special section. https://doi.org/10.1037/a0017653

Van de Water, A. T., Holmes, A., & Hurley, D. A. (2011). Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography–a systematic review. *Journal of sleep research, 20*(1pt2), 183-200. https://doi.org/10.1111/j.1365-2869.2009.00814.x

Vyazovskiy, V. V. (2015). Sleep, recovery, and metaregulation: explaining the benefits of sleep. *Nature and science of sleep*, 171-184. https://doi.org/10.2147/NSS.S54036

Whalen, D. J., Silk, J. S., Semel, M., Forbes, E. E., Ryan, N. D., Axelson, D. A., ... & Dahl, R. E. (2008). Caffeine consumption, sleep, and affect in the natural environments of depressed youth and healthy controls. *Journal of pediatric psychology, 33*(4), 358-367. https://doi.org/10.1093/jpepsy/jsm086

Yohannes, M. F., & Matsuda, T. (2016). Weather effects on household demand for coffee and tea in Japan. *Agribusiness, 32*(1), 33-44. https://doi.org/10.1002/agr.21434

Appendices

Appendix A
Informed Consent

General information about the study

As sleep is very important for everyone to function in daily life, you might be interested in getting more insight in what factors can influence your sleep. In this study we will specifically focus on how caffeine intake and physical activity are related to sleep.

Yes, I have read the introduction and am interested in participating *1*
No, I am not interested and want to withdraw from this study *0*

Inclusion criteria

In order to prevent having to exclude people from the study after they have participated, we kindly ask you to only continue with the study if you...

- ...are 18 years or older
- ...have a proficient understanding of English
- ...drink at least one of the following stimulants at least once a week: coffee, tea, soda* or energy drink.
- ...engage in physical activity** at least 3 times a week

Do you fit with these criteria (please answer below)?

Yes, I identify with all 4 criteria *1*
No, I do not identify with all and cannot continue with the study *0*

* Soda includes: coca cola, fanta, sprite, 7-up, bitter lemon, ginger ale, ice tea, tonic or root beer.

**Physical activity includes: participating in activities and/or sports, which increase your heart rate, for at least 3 times a week (such as biking to the store, going for a walk at a moderate/fast pace or engaging in weekly sport training

Informed Consent

The data gathered in this study will be anonymous and deleted after completion of this thesis. Only the researchers and their supervisor will have access to your answers.

You are asked to participate for 2 weeks (14 days) by using the app 'SEMA3'. Via this app we kindly ask you to fill in the demographics survey within 2 days after the start of the survey. Furthermore, for the next two weeks, you will receive a morning survey about how you slept the past night, followed by a few short survey questions at 3 fixed moments of each day.

Be aware that you can withdraw from the study at any time.

For questions, please contact: i.d.bouwer@student.utwente.nl OR j.i.piconrattel@student.utwente.nl.

Please answer below: I agree with the terms and give consent to participate in this study.

Yes, I agree and want to continue with the study	*1*
No, I do not agree and withdraw from the study	*0*

We thank you for participating. Make sure to allow the app to send you notifications when installing the app on your phone. Or check if notifications are allowed, by turning on the notifications of the app. This in order to have less possibilities of missing surveys at certain timeslots. If you are not sure how to put on the notifications of an app, please refer to How to Enable and Disable Push Notifications on Android - Make Tech Easier for Android and to Use notifications on your iPhone or iPad – Apple Support (UK) for iOS.

Appendix B
Baseline Questions

1. What is your age?
    a. Slider with minimum value of 18 and maximum value of 70.
2. What do you identify as?
    a. Male (0)
    b. Female (1)
    c. Other (2)

3. What is your occupation?
    a. Employed (0)
    b. Unemployed (1)
    c. Student (2)
    d. Other (3)
4. If you selected 'employed' in the previous question, please specify in which sector you work. If not selected, leave an 'x' or another small icon to continue.
5. What is your nationality?
6. How long did it usually take you to fall asleep during the past week?
    a. 0-15 minutes (0)
    b. 16-30 minutes (1)
    c. 31-45 minutes (2)
    d. 46-60 minutes (3)
    e. More than 60 minutes (4)
7. On average, how many hours did you sleep each night during the past week?
    a. 3 to 4 hours (0)
    b. 5 to 6 hours (1)
    c. 7 to 8 hours (2)
    d. 9 to 10 hours (3)
    e. More than 10 hours (4)
8. How would you, on average, rate the quality of you sleep over the past week?
    a. Slider with a minimum of 0 (really bad) and a maximum of 10 (excellent)
9. How many cups of coffee do you normally drink during one day?
    a. Slider between 0 and 10.
10. How many cups of tea do you normally drink during one day?
    a. Slider between 0 and 10.
11. How many glasses of soda do you normally drink during one day?
    a. Slider between 0 and 10.
12. How many cans of energy drink do you normally drink during one day?
    a. Slider between 0 and 10.

Appendix C

Sleep survey

1. How many hours did you sleep last night?
    a. Open question.
2. On a scale from 0-10, how would you rate your sleep?
    a. Slider between 0 and 10.
3. How would you rate your energy level right now?
    a. Slider from 0 (very sleepy) to 10 (very energetic).

Appendix D

Daily questions

1. How do you feel right now?
    a. Slider from 0 (very tired) to 10 (full of energy).
2. How many cups of coffee have you had since you woke up/ since the last survey?
    a. Slider from 0 to 10.
3. How many cups of tea have you had since you woke up/ since the last survey?
    a. Slider from 0 to 10.
4. How many glasses of soda have you had since you woke up/ since the last survey?
    a. Slider from 0 to 10.
5. How many cans of energy drink have you had since you woke up/ since the last survey?
    a. Slider from 0 to 10.

Appendix E

General R Studio Script

```
library(readr)
library(dplyr)
library(tidyr)
library(ggplot2)
library(scales)
library(nlme)
library(lme4)
```

```
library(performance)
library(psych)

# Download the data from the the different surveys as separate .csv files from SEMA3
# Define a list of the PARTICIPANT_ID values you want to keep, for instance, based on
compliance rate of >33% in SEMA3
selected_ids <- c("s027726465", "s077062546", "s143806477", "s159381842",
            "s230084949", "s244004245", "s283444806", "s330222955", "s392553353",
            "s595575588", "s662830324", "s683002070", "s694160992", "s728977899",
            "s762447320", "s775955172", "s822401278", "s828965218", "s870035860",
            "s963624552", "s993509399")
```

Load the separate surveys from SEMA3

Informed consent survey

```
### Read data from CSV file
ic <- read_delim("Inf.Cons.csv",
           col_types = cols(
             CREATED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             SCHEDULED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             STARTED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             COMPLETED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             EXPIRED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             UPLOADED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
             TOTAL_RT = col_double(),
             RAND_PROB = col_integer(),
             GEN.INFO = col_integer(),
             GEN.INFO_RT = col_double(),
             INCL.CRIT = col_integer(),
             INCL.CRIT_RT = col_double(),
             INF.CONS = col_integer(),
             INF.CONS_RT = col_double()
           )) %>%

# Sort by participant ID and scheduling
```

```
arrange(PARTICIPANT_ID, SCHEDULED_TS) %>%

# Group by participant ID and keep only the first entry for each
group_by(PARTICIPANT_ID) %>%
        slice(1) %>%

# Rename columns except for PARTICIPANT_ID
rename_with(~ paste0("ic_", .x), .cols = -c(PARTICIPANT_ID))

#  Filter the dataframe to keep only the cases with the defined PARTICIPANT_ID values
ic_filtered <- ic %>%
        filter(PARTICIPANT_ID %in% selected_ids)

# Define the relevant column names you want to keep
relevant_columns <- c("PARTICIPANT_ID", "ic_STUDY_ID", "ic_SURVEY_ID",
               "ic_COMPLETED_TS", "ic_GEN.INFO", "ic_INCL.CRIT",
                  "ic_INF.CONS")

# Select only the relevant columns
ic_filtered <- select(ic_filtered, relevant_columns)
```

Baseline survey

```
# Read data from CSV file
bl <- read_delim("baseline_questions.csv",
            col_types = cols(CREATED_TS = col_datetime(format = "%d-%b-%Y
                 %H:%M"),
                        # other columns defined here as before
                        SL._QLTY_RT = col_double()
                 ))
# Remove redundant columns up to 'SPRT/ACTVTY_RT'
bl <- bl %>%
        select(1:which(colnames(bl) == "SPRT/ACTVTY_RT"))

# Set <no-response> to NA for all character columns
```

```
bl <- bl %>%
  mutate(across(where(is.character), ~ifelse(. == "<no-response>", NA, .))) %>%
  arrange(PARTICIPANT_ID, COMPLETED_TS) %>%
  group_by(PARTICIPANT_ID) %>%
  slice(1) %>%
  rename_with(~ paste0("bl_", .x), .cols = -c(PARTICIPANT_ID))

#  Filter the dataframe to keep only the cases with the defined PARTICIPANT_ID values
bl_filtered <- bl %>%
  filter(PARTICIPANT_ID %in% selected_ids)

# Define the relevant column names you want to keep
relevant_columns <- c("PARTICIPANT_ID", "bl_COMPLETED_TS", "bl_AGE", "bl_ID",
                      "bl_JOB_0", "bl_JOB_1", "bl_JOB_2", "bl_JOB_3", "bl_SECTOR",
                      "bl_NAT.", "bl_SLP.WK", "bl_SL.HOURS", "bl_SL.QLITY",
                      "bl_DAY_COF", "bl_DAY_TEA", "bl_DAY_SODA", "bl_DAY_ENDR")

# Select only the relevant columns
bl_filtered <- select(bl_filtered, relevant_columns)
```

Sleep Survey

```
sleep <- read_csv("sleep.survey.csv",
                  col_types = cols(CREATED_TS = col_datetime(format = "%d-%b-%Y
                                   %H:%M"),
                              SCHEDULED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                              STARTED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                              COMPLETED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                              EXPIRED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                              UPLOADED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                              TOTAL_RT = col_double(),
                              RAND_PROB = col_integer(),
                              SL.HRS.NIGHT = col_double(),
                              SL.HRS.NIGHT_RT = col_double(),
                              QLTY = col_double(),
```

```
                              QLTY_RT = col_double(),
                              ENRGYLVL = col_double(),
                              ENRGYLVL_RT = col_double()
                  ))%>%
  #Sort by participant ID and scheduling
  arrange(PARTICIPANT_ID, SCHEDULED_TS)

# Set <no-response> to NA for all character columns
char_cols <- sapply(sleep, is.character)
sleep <- sleep %>%
  mutate_if(char_cols, ~ifelse(. == "<no-response>", NA, .))%>%
  rename_with(~ paste0("sleep_", .x), .cols = -c(PARTICIPANT_ID))
#  Filter the dataframe to keep only the cases with the defined PARTICIPANT_ID values
sleep_filtered <- sleep %>%
  filter(PARTICIPANT_ID %in% selected_ids)

# Define the relevant column names you want to keep
relevant_columns <- c("PARTICIPANT_ID", "sleep_CREATED_TS",
"sleep_SCHEDULED_TS",
                      "sleep_STARTED_TS", "sleep_COMPLETED_TS",
"sleep_SL.HRS.NIGHT",
                      "sleep_QLTY", "sleep_ENRGYLVL")

# Select only the relevant columns
sleep_filtered <- select(sleep_filtered, relevant_columns)

# Count unique PARTICIPANT_IDs in the dataframe
unique_participants <- n_distinct(sleep_filtered$PARTICIPANT_ID)
```

AM survey

```
am <- read_delim("Am questions.csv",
                 col_types = cols(CREATED_TS = col_datetime(format = "%d-%b-%Y
%H:%M"),
                              SCHEDULED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
```

```
                STARTED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                COMPLETED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                EXPIRED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                UPLOADED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                TOTAL_RT = col_double(),
                RAND_PROB = col_integer(),
                SLEEPINESS = col_double(),
                SLEEPINESS_RT = col_double(),
                COFFEE_1 = col_double(),
                COFFEE_1_RT = col_double(),
                TEA_1 = col_double(),
                TEA_1_RT = col_double(),
                SODA_1 = col_double(),
                SODA_1_RT = col_double(),
                ENERGY_DRINK_1 = col_double(),
                ENERGY_DRINK_1_RT = col_double(),
                MVNT_MIN.1_RT = col_double(),
                SP.ACT.1_RT = col_double(),
                PLACE.1 = col_integer(),
                PLACE.1_RT = col_double()
        ))%>%

#Sort by participant ID and scheduling
  arrange(PARTICIPANT_ID, SCHEDULED_TS)

# Remove redundant (old?) columns
am <- am %>%
  select(1:which(colnames(am) == "PLACE.1_RT"))

# Set <no-response> to NA for all character columns
char_cols <- sapply(am, is.character)
am <- am %>%
  mutate_if(char_cols, ~ifelse(. == "<no-response>", NA, .))%>%
  rename_with(~ paste0("am_", .x), .cols = -c(PARTICIPANT_ID))
```

```
#  Filter the dataframe to keep only the cases with the defined PARTICIPANT_ID values
am_filtered <- am %>%
  filter(PARTICIPANT_ID %in% selected_ids)

# Define the relevant column names you want to keep
relevant_columns <- c("PARTICIPANT_ID", "am_CREATED_TS",
"am_SCHEDULED_TS",
                      "am_STARTED_TS", "am_COMPLETED_TS", "am_SLEEPINESS",
                      "am_COFFEE_1", "am_TEA_1", "am_SODA_1",
                      "am_ENERGY_DRINK_1")

# Remove unnessesary columns
am_filtered[ , c ('SP.ACT.1_RT', 'SP.ACT.1')] <- list (NULL)
am_filtered[ , c ('SLEEPINESS_RT', 'COFFEE_1_RT')] <- list (NULL)
am_filtered[ , c ('TEA_1_RT', 'SODA_1_RT', 'ENERGY_DRINK_1_RT')] <- list (NULL)
am_filtered[ , c ('MVMNT_MIN.1', "MVNT_MIN.1", 'MVNT_MIN.1_RT',
'MVMNT_MIN.1_RT', 'PLACE.1', 'PLACE.1_RT')] <- list (NULL)
am_filtered[ , c ('SLEEPINESS_2', "SLEEPINESS_2_RT", "COFFEE_2",
"COFFEE_2_RT", "TEA_2", "TEA_2_RT", "SODA_2", "SODA_2_RT",
               "ENERGY_DRINK_2", "ENERGY_DRINK_2_RT", "SLEEPINESS_3",
"SLEEPINESS_3_RT", "COFFEE_3", "COFFEE_3_RT",
               "TEA_3", "TEA_3_RT", "SODA_3", "SODA_3_RT", "ENERGY_DRINK_3",
"ENERGY_DRINK_3_RT")] <- list (NULL)
am_filtered[ , c ("AFMOVMT.2", "AFMOVMT.2_RT", "SP_ACT.2", "SP_ACT.2_RT",
"PLACE_2", "PLACE_2_RT", "PLACE_3", "PLACE_3_RT",
               "EVMOVMT.3", "EVMOVMT.3_RT", "ACT.3", "ACT.3_RT")] <- list (NULL)
am_filtered[ , c ("START_END", "STARTED_TS", "EXPIRED_TS", "TOTAL_RT",
"RAND_PROB", "STUDY_NAME", "SURVEY_NAME")] <- list (NULL)

# Select only the relevant columns
am_filtered <- select(am_filtered, relevant_columns)
```

PM surveys

This dataset contains both the afternoon and evening survey!

```
# double means that it is numeric and therefor all non-numeric variables of that column will be removed
pm <- read_delim("PM SEMA3.csv",
        col_types = cols(CREATED_TS = col_datetime(format = "%d-%b-%Y
%H:%M"),
                SCHEDULED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                STARTED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                COMPLETED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                EXPIRED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                UPLOADED_TS = col_datetime(format = "%d-%b-%Y %H:%M"),
                TOTAL_RT = col_double(),
                RAND_PROB = col_integer(),
                FEEL.PM = col_double(),
                FEEL.PM_RT = col_double(),
                COF.PM = col_double(),
                COF.PM_RT = col_double(),
                TEA.PM = col_double(),
                TEA.PM_RT = col_double(),
                SODA.PM = col_double(),
                SODA.PM_RT = col_double(),
                EN.DR.PM = col_double(),
                EN.DR.PM_RT = col_double(),
                PM_MVMT = col_double(),
                PM_MVMT_RT = col_double(),
                COF.PM = col_double(),
                COF.PM_RT = col_double(),
                SP_PM_ACTVTY = col_character(),
                SP_PM_ACTVTY_RT = col_double(),
                PLACE_PM = col_integer(),
                PLACE_PM_RT = col_double(),
        ))%>%
```

```
  #Sort by participant ID and scheduling
  arrange(PARTICIPANT_ID, SCHEDULED_TS)

# Set <no-response> to NA for all character columns
char_cols <- sapply(pm, is.character)
pm <- pm %>%
  mutate_if(char_cols, ~ifelse(. == "<no-response>", NA, .))%>%
  rename_with(~ paste0("pm_", .x), .cols = -c(PARTICIPANT_ID))

#  Filter the dataframe to keep only the cases with the defined PARTICIPANT_ID values
pm_filtered <- pm %>%
  filter(PARTICIPANT_ID %in% selected_ids)

# Define the relevant column names you want to keep
relevant_columns <- c("PARTICIPANT_ID", "pm_CREATED_TS",
"pm_SCHEDULED_TS",
                      "pm_STARTED_TS", "pm_COMPLETED_TS", "pm_FEEL.PM",
                      "pm_COF.PM", "pm_TEA.PM", "pm_SODA.PM", "pm_EN.DR.PM")

# Select only the relevant columns
pm_filtered <- select(pm_filtered, relevant_columns)

# Convert pm_CREATED_TS to a POSIXct object if it is not already
pm_filtered <- pm_filtered %>%
  mutate(pm_CREATED_TS = as.POSIXct(pm_CREATED_TS, format = "%Y-%m-%d
%H:%M:%S"))

# Extract the date part and create a column to distinguish between afternoon and evening
measurements
pm_filtered <- pm_filtered %>%
  mutate(
    Date = as.Date(pm_CREATED_TS),
    TimeOfDay = ifelse(format(pm_CREATED_TS, "%H") < 18, "Afternoon", "Evening")
```

```
  )

# Split the data into afternoon and evening measurements
afternoon_data <- pm_filtered %>%
  filter(TimeOfDay == "Afternoon") %>%
  select(-TimeOfDay)

evening_data <- pm_filtered %>%
  filter(TimeOfDay == "Evening") %>%
  select(-TimeOfDay)

# Rename the columns in afternoon and evening data to distinguish them
afternoon_data <- afternoon_data %>%
  rename_with(~ paste0(.x, "_Afternoon"), -c(PARTICIPANT_ID, Date))

evening_data <- evening_data %>%
  rename_with(~ paste0(.x, "_Evening"), -c(PARTICIPANT_ID, Date))

# Join the afternoon and evening data on PARTICIPANT_ID and Date
pm_wide <- left_join(afternoon_data, evening_data, by = c("PARTICIPANT_ID", "Date"))

# Joining the am & pm dataframes: convert am_CREATED_TS and
pm_CREATED_TS_afternoon to POSIXct objects if they are not already
am_filtered <- am_filtered %>%
  mutate(am_CREATED_TS = as.POSIXct(am_CREATED_TS, format = "%Y-%m-%d
%H:%M:%S"))

pm_wide <- pm_wide %>%
  mutate(pm_CREATED_TS_Afternoon = as.POSIXct(pm_CREATED_TS_Afternoon,
format = "%Y-%m-%d %H:%M:%S"))

# Extract the date part from the timestamps
am_filtered <- am_filtered %>%
  mutate(Date = as.Date(am_CREATED_TS))
```

```
pm_wide <- pm_wide %>%
  mutate(Date = as.Date(pm_CREATED_TS_Afternoon))

# Join the dataframes on PARTICIPANT_ID and Date
combined_df <- left_join(am_filtered, pm_wide, by = c("PARTICIPANT_ID", "Date"))

 # Join combined_df with the sleep_filtered survey: convert sleep_CREATED_TS to
POSIXct if it is not already
sleep_filtered <- sleep_filtered %>%
  mutate(sleep_CREATED_TS = as.POSIXct(sleep_CREATED_TS, format = "%Y-%m-%d
%H:%M:%S"))

# Extract the date part from the sleep_CREATED_TS timestamp
sleep_filtered <- sleep_filtered %>%
  mutate(Date = as.Date(sleep_CREATED_TS))

# Join the combined_df with sleep_filtered on PARTICIPANT_ID and Date
final_df <- left_join(combined_df, sleep_filtered, by = c("PARTICIPANT_ID", "Date"))

# Join combined_df with bl_filtered on PARTICIPANT_ID
final_df <- left_join(final_df, bl_filtered, by = "PARTICIPANT_ID")

# Move the columns from bl_filtered to the front of the final_df dataframe
bl_columns <- names(bl_filtered)
other_columns <- setdiff(names(final_df), bl_columns)
final_df <- final_df[, c(bl_columns, other_columns)]

# Make a new variable for the day number within each participant & make both a numerical
(Day) and a factor variable (fDay)
final_df <- final_df %>%
  group_by(PARTICIPANT_ID) %>%
  arrange(PARTICIPANT_ID, Date) %>%  # Ensure data is sorted by PARTICIPANT_ID
and Date
```

```
  mutate(Day = row_number()) %>%     # Create the numerical Day variable
  ungroup() %>%
  mutate(fDay = as.factor(Day))      # Create the factor fDay variable

# Move PARTICIPANT_ID, Day, and fDay to the front
final_df <- final_df %>%
  select(PARTICIPANT_ID, Day, fDay, everything())

# Remove rows where Day > 14
final_df <- final_df %>%
  filter(Day <= 14)

# Frequency table for Day using dplyr
day_freq_table <- final_df %>%
  count(Day) %>%
  arrange(Day)

print(day_freq_table)
```

~ Now you have your final dataset for analyses. You can make your own variables per day and make plots to study associations over time with LMM models, for example: ~

Plot Total Caffeine & Sleep Hours

```
# Example total amount of caffeine per day, add the total_caffeine variable, accounting for missing values
final_df <- final_df %>%
  mutate(total_caffeine = rowSums(across(c(am_COFFEE_1, pm_COF.PM_Afternoon,
pm_COF.PM_Evening, am_TEA_1, pm_TEA.PM_Afternoon, pm_TEA.PM_Evening,
am_SODA_1, pm_SODA.PM_Afternoon, pm_SODA.PM_Evening,
am_ENERGY_DRINK_1, pm_EN.DR.PM_Afternoon, pm_EN.DR.PM_Evening)), na.rm =
TRUE))

summary(avg_total_caffeine)
```

```r
# Plot the average amount of caffeine per day - calculate the average total_caffeine for each Day
average_caffeine <- final_df %>%
  group_by(Day) %>%
  summarize(avg_total_caffeine = mean(total_caffeine, na.rm = TRUE))

# Plot the average total_caffeine over Days with labels for each Day
ggplot(average_caffeine, aes(x = Day, y = avg_total_caffeine)) +
  geom_line() +
  geom_point() +
  labs(title = "Average Total Caffeine Over Days",
       x = "Day",
       y = "Average Total Caffeine") +
  scale_x_continuous(breaks = average_caffeine$Day) +
  theme_minimal()

# Or plot number of hours sleep together with total caffeine in one figure - calculate the average total_coffee and sleep_SL.HRS.NIGHT for each Day
average_values <- final_df %>%
  group_by(Day) %>%
  summarize(
    avg_total_caffeine = mean(total_caffeine, na.rm = TRUE),
    avg_sleep_hours = mean(sleep_SL.HRS.NIGHT, na.rm = TRUE)
  )

# Reshape the dataframe to long format for plotting
average_values_long <- average_values %>%
  pivot_longer(cols = c(avg_total_caffeine, avg_sleep_hours),
               names_to = "variable",
               values_to = "value")

# Plot the values with the correct labels for Total Caffeine and Sleep Hours
ggplot(average_values_long, aes(x = Day, y = value, color = variable)) +
```

```
geom_line() +
geom_point() +
labs(title = "Average Total Caffeine and Sleep Hours Over Days",
    x = "Day",
    y = "Average Value") +
scale_x_continuous(breaks = average_values$Day) +
theme_minimal() +
scale_color_manual(values = c("avg_total_caffeine" = "blue", "avg_sleep_hours" = "red"),
                labels = c("Total Caffeine" = "blue", "Sleep Hours" = "red")) +
theme(legend.title = element_blank())
```

~ You can run the same analysis for sleep quality and caffeine or sleep quality or sleep quanity and coffee/tea/soda/energy drink. Just make sure you change the variables as they are named in your final dataset in each of the codes. ~