Hypothesis:

Based on an assumed principal of compensation, and the importance of dreaming in sleep, we postulate that, after a period of sleep deprivation, a recovery sleep should feature an increased duration of REM sleep.

Our objective is to prove or disprove this hypothesis and to offer explanations for our findings.

Core Questions:

1. Do we find evidence that the duration of REM sleep significantly increases during recovery sleep, as compared to baseline sleep?
2. Do we find evidence that REM sleep is affected in any other way (i.e.: other than duration) between baseline and recovery sleeps?
3. Do we find changes to NREM stages that might indicate a principle of compensation is at work?
4. Do we find any evidence that the principal of compensation applies to sleep at all?

Data Description:

For our analysis, we use the Sleep EEG Data project from the laboratory of Dr. Mary Carskadon. The original Carskadon project recorded the sleep data for 4 subjects during a baseline sleep and a recovery sleep that followed a period of sleep deprivation. So, there are 8 sets of data to be examined: 2 sleep sessions (baseline and recovery) for each of the 4 subjects. Each set of data is stored in a separate (.npz) file.

Each set of data contains 3 subsets of information:
1) The sampling rate, which was the same for all 4 subjects at 128 samples per second.
2) Stage classification data, stored as an integer from 0 to 7 covering 30 second epochs:
   - 0. Awake
   - 1. NREM Stage 1
   - 2. NREM Stage 2
   - 3. NREM Stage 3
   - 4. NREM Stage 4
   - 5. REM
   - 6. Movement Time
   - 7. Unscored
3) Raw sleep data comprised of 9 channels of EEG, EOG and EMG data.

The data that we focused on:
In a group meeting to set the current project objectives and plan our course of action, our team decided that the questions we aimed to answer would be best addressed by using the pre-classified stage data, rather than on re-interpreting the raw data. There were a number of reasons for this choice:
- any effort to answer our selected questions would require us to reproduce the stage classifications and we didn't have a process that we would confidently feel would surpass the values that had been provided.
- we were aware from forum comments from students on similar projects that there were questions about the the raw data provided
- time was not plentiful, so we would avoid taking on tasks that would not likely improve the quality of the results.

Likewise, it was decided to treat the sleep stage data (stages 1 to 5) as well as the waking stage data (stage 0) to be most the relevant stages for our task. This entailed that we would mostly ignore the unscored data, as this did not appear to be relevant to our task. Further review of the stage data provided indicated that the "movement time" stage (stage 6) was only minimally present in the Carskadon results, so it could also be considered insignificant for our purposes.

Unknowns, assumptions and data challenges:
- Little information is available about the 4 subjects: we do not know their age, gender, race, state of health, or even species. We assume they are human, or at least mammalian.
- We do not know anything about the period of sleep deprivation, whether it was a few hours or a matter of days, or even if the period was the necessarily the same for all subjects.
- We do not know under what conditions the data was collected. Our assumption is that it was at least partly in a controlled laboratory setting.
- We do not know how the stage classifications were arrived at. Our assumption is that they were done by manually inspecting the data by experienced and capable researchers, and thus the information will be reliable. Further efforts to confirm the accuracy of the classifications would be desirable at a later stage. The quality of our conclusions depend on the quality of the original classification efforts.
- The fact that there are only 4 subjects limits the strengths of any conclusions that can be arrived at. In all likelihood, we can only expect to achieve suggestive conclusions.

Common terms:
- BSL: Baseline
- REC: Recovery
- S1, S2, S3, S4 (Subject 1, 2, 3, 4)

Analysis: Overview

We will use the stage classification values to determine the amounts of time spent in each of the various phases of sleep and wakefulness.

Since each stage classification was based on a 30 second epoch, we simply need to multiply the number of occurrences of any particular stage classification value by 30 to determine how much time was spent in that stage in seconds.

With this simple strategy in mind, there are a number of ways that we could approach our analysis. The ones we considered:

1. Comparison of total durations for each stage:
- *Objectives:* Determine if the duration of REM sleep (or other stages) is increases in recovery sleep as compared to baseline.
- *General Approach:*
- Compare the total times for each phase combined. In other words, when looking at either the baseline (BSL) data for a given subject, or the recovery data (REC), we can sum every occurrence of a given stage value (say, stage 5 for REM sleep) within that particular data set, then multiply it by 30 to get the total time that subject spent in REM sleep. By performing the same operation on both BSL and REM sleep, we can see how the total times change from stage to stage.
- *Format of supporting analysis:*
- This analysis to be supported by a combination of numerical calculations and graphic representations.

2. Comparison of stages in contiguous blocks:
- *Objectives:* Look for variations in REM or other stages of sleep (and wakefulness) that are not seen in total durations.
- *General Approach:*
- Compare the occurrences of contiguous blocks of particular stages of sleep. In this case, if we were considering, say, stage 5 REM sleep, we might expect that the subject would pass through multiple contiguous periods of REM stage sleep, and that the number, duration and sequencing of those blocks of stage 5 sleep would differ between BSL and REC, so the analysis would require somewhat more complex comparisons. This analysis might lead us to see either increases or decreases in the numbers of contiguous blocks, variations in the durations of those blocks, or variations in the sequences.
- *Format of supporting analysis:*
- Primarily supported with graphic figures showing the progression through the various stage blocks.

3. Comparison of stage transitions:
- *Objectives:* Continue to seek variations in REM or other stages of sleep (and wakefulness) that are not seen in total durations.
- *General Approach:*
- Compare the transitions from one stage to another. We may see that BSL and REC sleep show different tendencies when it comes to moving from one sleep stage to another. This is challenging to accomplish with the previous approaches.
- *Format of supporting analysis:*
- Primarily supported with graphic figures showing the number of combinations of transitions from one stage block to another.

Each of the 3 approaches above will constitute a section in this document. Conclusions will be supplied at the end of each section, then a summary with final conclusions will be presented at the end.

Section 1: Comparison of total durations for each stage

Objectives for this section:
In this section, we want to compare total durations for each stage of sleep to see if any are consistently longer, particularly REM sleep.

Approach:
The data is extracted by subject so we can compare the BSL and REC durations for each stage. After, we resort the comparison results so that we can examine each set of stages together. Comparison will take into account the absolute durations for each stage to see how the time spent in each stage may have increased or decreased after sleep deprivation. The comparison will also take into account the difference in time spent in each stage of slumber, relative to the overall difference in sleep times for all stages of sleep.

Challenges:
Sleep vs. wakefulness:
We wanted to focus on sleep and waking stages, largely ignoring the Movement Time and Unscored data. It makes sense to compare how periods of wakefulness will interact with periods of sleep before and after sleep deprivation. Nonetheless, it seems prudent to make distinctions as to when that wakefulness occurred: at the start, middle or end of the recorded sleep period. If the recorded wakefulness occurred before a sleep stage was encountered, or afterward, it may have a different significance than levels of wakefulness that were achieved within sleep.

Calculating relative duration differences
When we calculate the relative differences of each stage we want to understand whether the difference is simply in proportion to any lengthening or shortening of the overall sleep period, or whether it appears to differ noticeably from the general increase or decrease in sleep time. For example, if the overall sleep time happens to increase by 1.25%, and each stage individually also increases by somewhere in the general vicinity of 1.25%, then we might be tempted to conclude that the increase in any stage is just a function of the overall increase.

So, we want to pay attention to how any given stage of sleep increases or decreases with the overall sleep time. So, one of our first acts as we process data for each subject, will be to calculate a value for the ratio of overall sleep time between BSL and REC session. Then we will consider that ratio as an indicator of whether a given stage has increased or decreased by a significant amount relative to the whole sleep session.

It should be noted that, for the purpose of the relative ratio just mentioned, we have only taken sleep stages (stages 1-5) into consideration, and not wakefulness (stage 0). The intent was to ensure that we are comparing apples to apples. The validity of this approach may need to be considered.

Pre-analysis biases:
Prior to carrying out this analysis, the following general biases applied:
- We generally suspected that recovery sleep would generally be longer than baseline sleep.
- We generally suspected that the overall duration of REM sleep would longer in recovery sleep than baseline sleep.

Total Duration by Subject: Wakefulness stage 0 only

Subject		Pre-sleep Wakefulness (epochs)	Intermittent Wakefulness During Sleep (epochs)	Post-sleep Wakefulness (epochs)	Total Wakefulness: (epochs)
S1	Baseline	35	92	0	127
	Recovery	1	21	121	143
	Difference	-34	-71	121	16
S2	Baseline	84	23	0	107
	Recovery	2	239	94	335
	Difference	-82	216	94	228
S3	Baseline	51	58	1	110
	Recovery	5	115	0	120
	Difference	-46	57	-1	10
S4	Baseline	33	42	1	76
	Recovery	5	9	68	82
	Difference	-28	-33	67	6

Analysis of Total Wakefulness stage 0

The spreadsheet above shows a consistent pattern of a consistent decrease in wakefulness during the pre-sleep phase. The effect on bouts of wakefulness during sleep is not consistent. Post-sleep wakefulness is noticeably longer for 3 out of 4 subjects. In all cases, we saw an overall increase in total wakefulness, but in all but one case, the total increases seemed minor.

Total Duration By Subject: Sleep stages 1 to 5 only:

Subject	Baseline (epochs)	Recovery (epochs)	Difference (epochs)	Difference (seconds)	Difference (hours)	Ratio REC:BSL (where BSL = 1)
S1	1073	1895	822	24660	6.85	**1.77**
S2	1093	1828	735	22050	6.12	**1.67**
S3	1087	2039	952	28560	7.93	**1.88**
S4	1125	1546	421	12630	3.51	**1.37**

Analysis of Total Sleep stages 1 to 5

One main purpose of the spreadsheet above is to calculate a ratio of the durations of recovery sleep to baseline sleep. We will use this number later to help compare the relative durations of each sleep stage.

This number was arrived at by simply dividing the total duration of recover sleep stages by the total duration of baseline stages. The number provided in the rightmost column can be thought of as X:1. So, S1 had a recovery to baseline ration of 1.77 to 1, and so on.

From the above results, we can see that all of the subjects experienced a significant increase in the duration of recovery sleep over baseline sleep. From the ratio field, it seems noteworthy that 3 out of 4 subjects experienced more than a 66% increase in recovery sleep over baseline.

Table 1

S1	96	REC:	133	REC > BSL	Diff:	*37*	(bins)	1110	(sec)	0.31	(hrs)
S2	117	REC:	56	REC < BSL	Diff:	-61	(bins)	-1830	(sec)	-0.51	(hrs)
S3	132	REC:	157	REC > BSL	Diff:	25	(bins)	750	(sec)	0.21	(hrs)
S4	63	REC:	42	REC < BSL	Diff:	-21	(bins)	-630	(sec)	-0.17	(hrs)