Examining the Impact of Sleep Stage Composition on Energy Levels and Productivity: Evidence from Wearable-Based Personal Analytics

Mark Rhey Anthony D.F. De Luna
College of Computing and Information Technologies
National University - Philippines
delunamr@students.national-u.edu.ph

***Abstract*—This study examines the relationship between sleep stage distribution and daytime functioning through intensive personal tracking over four months (October 2025 to February 2026). Using data from a wearable sleep tracker combined with daily self-reported energy levels, productivity scores, and task completion rates, I investigated how different sleep stages (light, deep, and REM) influence daily performance. Analysis of 113 complete sleep observations revealed significant positive correlations between sleep quality and daytime functioning. Deep sleep duration showed the strongest relationship with next-day energy levels (r = 0.319, p ¡ 0.001), while total sleep duration was significantly associated with productivity scores (r = 0.248, p ¡ 0.01). Participants who achieved more than 7 hours of sleep demonstrated 65% higher average energy levels (4.21 vs 2.55) and 26% higher task completion rates (69.8% vs 55.4%) compared to those sleeping less than 5 hours. Statistical hypothesis testing confirmed these relationships, with ANOVA revealing significant differences across sleep categories (p ¡ 0.001) and Cohen's d effect sizes indicating large practical significance. These findings support the hypothesis that optimizing sleep stage distribution, particularly deep sleep, can substantially enhance academic performance and daily functioning in student populations.**

***Index Terms*—sleep stages, sleep quality, productivity, energy levels, deep sleep, REM sleep, wearable technology, personal analytics, student performance, N-of-1 study, actigraphy, self-tracking**

I. Introduction

Sleep is a fundamental biological process that directly impacts cognitive function, physical health, and emotional well-being. For students and young professionals, maintaining optimal sleep quality is particularly critical as it influences academic performance, work productivity, and overall quality of life [1]. Despite widespread awareness of sleep's importance, many individuals struggle to understand how specific aspects of their sleep architecture—the distribution and quality of different sleep stages—influence their daily functioning. This disconnect between general knowledge about sleep and actionable insights for personal optimization motivated the present research.

This study focuses on the relationship between sleep stage distribution and multiple dimensions of daytime performance, including subjective energy levels, productivity, and task completion rates. The investigation centers on a central question: does the distribution of sleep stages throughout the night significantly affect how individuals function the following day? While existing research has established that sleep deprivation impairs performance [2], less is known about how specific sleep stages contribute to real-world daily functioning in non-clinical populations, particularly students who often experience irregular sleep patterns due to academic demands.

Understanding this relationship carries significant practical importance for student populations. College and university students typically average only 6-7 hours of sleep per night, well below recommended levels [3], creating a chronic state of partial sleep restriction. This sleep debt has been shown to accumulate over consecutive nights, producing cognitive deficits equivalent to complete sleep deprivation [4]. If specific sleep stages prove crucial for maintaining energy and productivity, students could benefit from targeted interventions to optimize their sleep architecture rather than simply attempting to increase total sleep duration—a goal often difficult to achieve given academic and social demands.

Prior research has established several key findings relevant to this investigation. Meta-analytic work demonstrates that sleep deprivation significantly impairs cognitive performance (d = -0.70), motor performance (d = -0.60), and mood (d = -0.99), with mood being the most affected domain [2]. Deep sleep has been linked to memory consolidation and physical restoration [5], while REM sleep supports emotional regulation and creative problem-solving, with participants showing a 33% improvement in creative tasks after experiencing REM sleep compared to those who did not [6]. However, most studies examining these relationships employ controlled laboratory settings or clinical populations, leaving open questions about how sleep architecture affects real-world daily functioning in naturalistic environments.

This project addresses three important gaps in current research. First, most sleep research relies on between-person comparisons rather than within-person intensive longitudinal analysis. As Fisher et al. [7] demonstrated, group-level associations often differ substantially from individual-level patterns, making personalized data particularly valuable for understanding how sleep affects specific individuals. Second, limited research examines the relationships between sleep stage distribution and multiple dimensions of daytime functioning simultaneously in naturalistic contexts. Most studies focus on single outcomes (e.g., cognitive testing) rather than the broader spectrum of energy, productivity, and task completion relevant to student life. Third, despite student populations experiencing particularly irregular sleep patterns and high vulnerability to sleep-related performance impairments, they remain understudied using personal tracking methodologies that could provide actionable insights.

The primary goal of this research is to determine how different sleep stages contribute to overall sleep quality and how this quality translates into energy levels and productivity the following day. Secondary objectives include examining whether higher deep sleep and REM sleep lead to better daytime functioning; investigating whether increased time spent awake during the night is associated with lower sleep quality and reduced performance; identifying sleep stage patterns that consistently result in high energy levels and task completion rates; and observing how weekly changes in sleep stage balance influence overall daily functioning. These objectives are pursued through four months of intensive self-tracking, combining objective wearable sleep data with subjective daily performance ratings.

This investigation seeks to answer several specific research questions: Does sleep stage distribution significantly affect daytime functioning as measured by energy, productivity, and task completion? Do deep sleep and REM sleep predict next-day energy levels when controlling for total sleep duration? Does total sleep duration correlate with productivity scores and task completion rates? Are there significant performance differences between high sleep (7 hours) and low sleep (5 hours) conditions, and if so, what is the magnitude of these differences? What sleep patterns optimize energy and productivity for academic work? By addressing these questions through systematic data collection and comprehensive statistical analysis, this study provides empirical evidence for the practical implications of sleep architecture on student life and daily performance, contributing to both the academic literature on sleep science and the emerging field of personal analytics for behavior optimization.

II. Literature Review

Understanding the relationship between sleep and daytime performance requires examination of prior research across multiple domains: the architecture of sleep stages, the effects of sleep quality on cognitive and physical performance, the validity of consumer wearable devices for sleep tracking, and individual differences in sleep needs and responses to sleep restriction. This review synthesizes existing knowledge in these areas to provide context for the present investigation.

Human sleep consists of distinct stages that cycle throughout the night, each serving specific physiological and cognitive functions. According to the American Academy of Sleep Medicine, sleep is divided into Non-Rapid Eye Movement (NREM) sleep and Rapid Eye Movement (REM) sleep, with NREM further subdivided into three stages: N1 (light sleep), N2 (intermediate sleep), and N3 (deep sleep or slow-wave sleep) [8]. Deep sleep is characterized by high-amplitude, low-frequency delta waves and plays a particularly important role in physical restoration, immune function, and declarative memory consolidation. Research by Diekelmann and Born [5] demonstrated that slow-wave sleep facilitates memory consolidation, with greater amounts of deep sleep associated with improved recall of previously learned information. More recent work by Xie et al. [9] established that deep sleep facilitates the clearance of metabolic waste products from the brain through the glymphatic system, providing a mechanistic explanation for its restorative function.

REM sleep, characterized by rapid eye movements, muscle atonia, and vivid dreaming, has been strongly linked to emotional regulation, creativity, and procedural memory formation. Walker and Stickgold's [6] influential research demonstrated that REM sleep enhances insight and the integration of new information with existing knowledge networks, showing that participants who experienced REM sleep showed a 33improvement in creative problem-solving compared to those who did not enter REM sleep. The typical sleep architecture for healthy young adults consists of approximately 50% light sleep (N1 and N2), 20-25% deep sleep, and 20-25% REM sleep [10], though individual variations exist based on age, lifestyle factors, and circadian timing preferences.

The relationship between sleep quality and daytime performance has been extensively documented across multiple domains. Pilcher and Huffcutt's [2] meta-analysis of 19 studies examining sleep deprivation and performance found significant impairments in cognitive performance (d = -0.70), motor performance (d = - 0.60), and mood (d = -0.99), with mood being the most affected domain. These effect sizes indicate substantial practical impacts of inadequate sleep on daily functioning. More recent research has focused on specific aspects of sleep architecture rather than total sleep duration alone. Van Dongen et al. [**?**] demonstrated in their landmark dose-response study that chronic sleep restriction to 4-6 hours per night produced cognitive deficits equivalent to up to 2 nights of complete sleep deprivation, with participants showing cumulative deterioration over consecutive nights that could not be easily compensated.

Energy levels and subjective alertness have been shown to correlate with sleep quality metrics, though the relative importance of objective versus subjective measures remains debated. Åkerstedt et al. [12] found that self-reported sleep quality predicted next-day alertness more strongly than objective sleep duration measured by actigraphy (r = 0.45 vs. r = 0.31), suggesting that subjective experiences of sleep quality may be particularly relevant for understanding daytime functioning. However, this finding also highlights potential discrepancies between objective and subjective sleep measures that warrant further investigation.

Research examining productivity specifically has revealed important connections between sleep and work performance. Barnes and Hollenbeck [13] examined sleep quality among working adults and found that poor sleep quality predicted reduced job performance, with the relationship mediated by decreased self- control resources. Their research suggested that inadequate sleep depletes cognitive resources necessary for maintaining focus and regulating behavior, thereby reducing productive output. Similarly, Mullins et al. [?] demonstrated that employees who obtained less than 6 hours of sleep showed significantly reduced work engagement and task performance compared to those obtaining 7-9 hours, with effects observed across multiple occupational domains.

The emergence of consumer-grade wearable sleep tracking devices has enabled large-scale collection of naturalistic sleep data outside traditional laboratory settings. Modern wearables use a combination of accelerometry, heart rate variability, and in some cases additional sensors to estimate sleep stages [15]. While these devices do not achieve the gold-standard accuracy of polysomnography (PSG), validation studies have shown reasonable agreement for basic sleep metrics. De Zambotti et al. [16] evaluated several popular wearable devices against PSG and found that most accurately detected total sleep time (within 10- 15 minutes of PSG), but showed greater variability in stage-specific classification, particularly for distinguishing light versus deep sleep. Despite these limitations, the authors concluded that wearables provide valuable data for longitudinal tracking of sleep patterns in naturalistic environments where PSG would be impractical.

The use of personal tracking data for health research represents an emerging paradigm known as "N-of-1" or single-subject experimental designs. Dallery et al. [?] discussed the value of intensive longitudinal data collection for understanding individual-level relationships between behaviors and outcomes. This approach is particularly relevant for sleep research, as sleep patterns and their effects can vary substantially between individuals based on chronotype, lifestyle, and genetic factors [?]. Li et al. [?] demonstrated the utility of consumer wearables for identifying personalized sleep patterns by analyzing data from over 10,000 users, revealing substantial individual differences in optimal sleep duration for next-day functioning, ranging from 6 to 9 hours. This finding supports the need for personalized rather than one-size-fits-all sleep recommendations, though their study relied on cross-sectional comparisons rather than within-person longitudinal analyses.

Research increasingly recognizes substantial individual differences in sleep needs and the relationship between sleep and performance. Roenneberg et al. [?] documented significant variation in chronotype (preference for sleep-wake timing) across over 55,000 individuals, with implications for when individuals obtain their most restorative sleep. Their research revealed that social obligations often force individuals to sleep outside their biological preferences, creating "social jet lag" that impairs functioning. Van Dongen et al. [?] identified trait-like individual differences in vulnerability to sleep restriction, finding that some individuals show minimal cognitive impairment after sleep loss while others show marked deficits. This suggests that the relationship between sleep stages and daytime performance may vary substantially across individuals, making personalized data particularly valuable for understanding individual-level patterns.

Despite extensive research on sleep and performance, several important gaps remain. First, most studies examine sleep in clinical populations or controlled laboratory settings, limiting generalizability to everyday life. Second, while research has established that sleep affects cognitive performance, fewer studies have examined the specific relationships between sleep stage distribution and subjective measures of energy and productivity in naturalistic contexts where multiple real-world factors influence both sleep and performance. Third, most existing research relies on group-level analyses, with limited investigation of individual-level patterns that could inform personalized interventions. The present study addresses these gaps by conducting an intensive longitudinal analysis of one individual's sleep patterns and daytime functioning over four months, combining objective sleep tracking with subjective daily ratings of energy, productivity, and task completion to examine how sleep stage distribution relates to multiple dimensions of daytime functioning in a real-world context.

III. Methodology

This study employed an intensive within-person longitudinal design to examine the relationship between sleep stage distribution and daytime functioning over a four-month period. The methodology combined objective sleep measurement through wearable technology with subjective daily self-reports of performance metrics, providing comprehensive data on both sleep architecture and its practical consequences for daily life.

A. Participant and Study Design

The participant in this N-of-1 study was the author, a university student (age 22) enrolled in the College of Computing and Information Technologies at National University Philippines. This single-subject intensive longitudinal design was selected deliberately to enable within-person analysis of sleep-performance relationships, which can reveal individual-level patterns that may differ from population averages [?]. While limiting generalizability, this approach provides personalized insights into how sleep architecture affects daily functioning for this specific individual, offering actionable information for behavior modification that population-level studies cannot provide.

B. Variables and Data Collection Procedures

Ten primary variables were systematically collected to comprehensively assess both sleep architecture and daytime functioning. Sleep-related variables included total sleep duration (hours), light sleep duration (minutes), deep sleep duration (minutes), REM sleep duration (minutes), awake time during sleep (minutes), and number of awakenings. These variables provide detailed information about sleep architecture and quality, allowing examination of both total sleep quantity and the distribution across different sleep stages.

Daytime performance variables included energy level (rated 1-5), productivity score (rated 1-10), task completion rate (percentage of planned tasks completed), and study duration (minutes spent on academic work). These variables were selected to capture multiple dimensions of daily functioning relevant to student life: subjective vitality and alertness (energy), overall productive output and work quality (productivity), concrete accomplishment of planned activities (task completion), and academic engagement (study time). Together, these measures provide a comprehensive picture of how sleep affects various aspects of daily performance beyond simple measures of cognitive function.

Sleep data was automatically recorded using a wearable fitness tracker equipped with sleep stage detection capabilities. The device monitors sleep stages through a combination of heart rate variability analysis, movement pattern detection via accelerometry, and breathing rate estimation. These sensor inputs feed into proprietary algorithms that categorize sleep into four stages: light sleep (corresponding to NREM stages N1 and N2), deep sleep (NREM stage N3), REM sleep, and awake time. The device automatically synchronizes data to a mobile application each morning, providing detailed summaries of sleep duration, stage distribution, and quality metrics.

Daily performance metrics were manually logged each evening using a standardized template in Google Sheets. To minimize memory bias and ensure consistency, entries were made at approximately the same time each evening (between 9:00 PM and 10:00 PM), reflecting on the day's experiences while they remained fresh. Energy level was conceptualized as subjective vitality and alertness throughout the day, rated on a 1-5 scale where 1 represented extreme fatigue and 10 represented exceptional energy. Productivity score reflected overall assessment of productive output and work quality on a 1-10 scale, where 1 indicated accomplishing very little and 5 indicated a highly productive day. Task completion rate was calculated as the percentage of planned tasks actually completed, based on a daily task list maintained in a separate planning document. Study duration was measured as total minutes spent on academic work, including classes, homework, reading, and project work.

Data collection occurred continuously from October 1, 2025, to February 6, 2026, spanning four months and resulting in 129 total days of observation. This extended period was chosen to capture typical weekly and monthly variations in both sleep patterns and academic demands while allowing sufficient data for robust statistical analysis. The duration exceeded most prior self-tracking studies and provided adequate statistical power for detecting relationships between sleep and performance variables.

The combination of automated wearable tracking and manual self-report logging balanced the advantages of objective measurement with the richness of subjective experience. While wearable devices provide consistent, unbiased measurement of sleep patterns, they cannot capture the phenomenological experience of energy and productivity. Conversely, while self-reports introduce potential bias, they reflect the subjective states that ultimately determine how individuals experience their daily functioning. This mixed- methods approach provides a more complete picture than either data source alone.

C. *Operational Definitions*

To ensure clarity and replicability, all variables were operationally defined with precise measurement criteria:

- Total Sleep Duration was measured in hours, calculated as the time from sleep onset (first detection of sleep stage) to final awakening, including all sleep stages but excluding awake time. This metric
- Light Sleep encompassed NREM stages N1 and N2, measured in minutes. These stages represent transitional and intermediate sleep states that facilitate cycling between wakefulness, deep sleep, and REM sleep. Reflects overall sleep quantity rather than just time in bed.
- Deep Sleep referred to NREM stage N3 or slow-wave sleep, measured in minutes. This stage is characterized by high-amplitude delta waves and has been associated with physical restoration,
- REM Sleep denoted Rapid Eye Movement sleep, measured in minutes. This stage is characterized by rapid eye movements, muscle atonia, vivid dreaming, and has been linked to emotional regulation, creativity, and procedural memory formation.
- Awake Time represented total minutes spent awake during the sleep period after initial sleep onset, excluding the time before falling asleep or after final awakening. This metric captures sleep fragmentation and nighttime disruptions. immune function, and memory consolidation.
- Energy Level was defined as self-reported subjective vitality and alertness on a 1-5 scale, where 1 indicated extreme fatigue, inability to focus, and desire to sleep, while 5 indicated exceptional energy, alertness, and capacity for sustained mental and physical effort.
- Productivity Score reflected self-assessed overall productive output on a 1-10 scale, where 1 indicated accomplishing very little of value despite available time, while 10 indicated a highly productive day with significant progress on important tasks and high-quality work output.
- Task Completion Rate was calculated objectively as (tasks completed / tasks planned) × 100, expressed as a percentage. Tasks were defined at the beginning of each day and reviewed each evening for completion status, providing a concrete measure of goal achievement.

D. *Data Cleaning and Preprocessing*

Raw data underwent systematic cleaning and preprocessing to ensure quality and enable statistical analysis. Missing data was first identified and assessed. Out of 129 total days, 16 days (12.4%) lacked complete sleep tracking data, occurring when the wearable device was not worn overnight or failed to record data properly. These missing days were excluded from analysis, resulting in a final dataset of 113 complete observations (87.6% completeness). Analysis of missingness patterns suggested the absences were random rather than systematic, as they showed no association with day of week, academic deadlines, or prior sleep patterns.

Data type conversion was performed to ensure all variables were properly formatted for statistical analysis. The date field was converted to datetime format to enable temporal analysis. Sleep duration and all sleep stage measurements were converted from text to numeric format. Self-reported metrics (energy, productivity, task completion) were verified as numeric. Invalid entries, indicated by '-' symbols in the original data, were replaced with NaN (Not a Number) values and treated as missing data.

Outlier detection was conducted using box plots and interquartile range (IQR) analysis for all numeric variables. While some extreme values were identified—such as very short sleep durations (1 hour), very long sleep durations (10 hours), and extended awake time (150 minutes)—these were retained in the analysis as genuine observations reflecting real sleep disruptions and variations rather than measurement errors. The decision to retain outliers was based on the naturalistic study design, where extreme values represent important aspects of real-world sleep variability.

Several derived variables were calculated to facilitate analysis. Sleep stage percentages were computed as (stage duration / total sleep time) × 100 for light, deep, and REM sleep, allowing comparison of sleep architecture independent of total sleep duration. Sleep quality categories were created by binning total sleep duration into three groups: Low (5 hours), Medium (5-7 hours), and High (7 hours), based on sleep duration recommendations and natural breaks in the data distribution. Energy categories were similarly created by binning energy ratings into: Very Low (1-2), Low (3), Medium (4), and High (5). Time-based features including month, week number, and day of week were extracted from the date field to enable temporal pattern analysis.

Potential biases and measurement errors were carefully considered throughout data collection and analysis. Self-report bias, inherent in subjective ratings of energy and productivity, was minimized by maintaining consistent rating times each evening and using clearly defined anchor points for scale interpretation. The possibility that mood or other factors could influence ratings independently of actual performance was acknowledged as a limitation. Wearable device measurement error was recognized as a constraint, given that consumer sleep trackers do not achieve polysomnography-level precision for sleep stage classification [**?**]. However, validation studies demonstrate reasonable accuracy for total sleep time and basic stage detection, supporting their use for longitudinal personal tracking despite limitations in absolute accuracy.

Reactivity bias—the possibility that awareness of being tracked might change behavior—was addressed through extended tracking duration. By maintaining data collection for four months, the novelty effect of self-monitoring was expected to diminish, allowing observation of more natural behavioral patterns. Missing data bias was evaluated by comparing characteristics of days with complete versus incomplete data, revealing no systematic differences that would suggest biased sampling of particular types of days.

E. Statistical Analysis Methods

Multiple statistical techniques were employed in a systematic progression from descriptive to inferential analysis. The analytical strategy was designed to comprehensively examine relationships between sleep and performance variables while testing specific hypotheses about the effects of sleep stage distribution on daytime functioning.

Descriptive Statistics formed the foundation of the analysis. Mean, median, standard deviation, minimum, maximum, and range were calculated for all continuous variables to characterize central tendency, variability, and distribution shape. Frequency distributions were computed for categorical variables (sleep categories, energy categories) to understand the distribution of observations across different sleep and performance levels. These descriptive measures provided essential context for interpreting subsequent inferential analyses.

Correlation Analysis examined linear relationships between continuous variables using Pearson correlation coefficients. This method was appropriate given the primarily continuous nature of the data (sleep duration, stage minutes, performance scores). Statistical significance was assessed using two-tailed tests with = 0.05 as the threshold for declaring relationships statistically significant. Correlation strength was interpreted using conventional guidelines: —r— ¡ 0.3 indicating weak relationships, 0.3 —r— ¡ 0.7 indicating moderate relationships, and —r— 0.7 indicating strong relationships. For key correlations, 95intervals were calculated to estimate the range of plausible correlation values in the population.

Hypothesis Testing employed multiple statistical tests to validate specific research hypotheses. One-way Analysis of Variance (ANOVA) was used to test whether mean energy, productivity, and task completion differed significantly across the three sleep duration categories (Low, Medium, High). This method was appropriate for comparing means across three independent groups formed by categorizing a continuous variable. The F-statistic and associated p-value indicated whether at least one group mean differed significantly from the others.

Following significant ANOVA results, Tukey's Honestly Significant Difference (HSD) test was conducted for all pairwise comparisons between sleep categories. This post-hoc procedure controls the family-wise error rate across multiple comparisons, preventing inflation of Type I error probability when testing multiple hypotheses simultaneously. The test identified which specific pairs of sleep categories showed significantly different mean performance levels.

Independent samples t-tests provided focused comparisons between the High sleep (¿7 hours) and Low sleep (¡5 hours) groups. These tests examined whether extreme differences in sleep duration produced statistically significant differences in performance outcomes. The t-statistic, degrees of freedom, and p- value indicated statistical significance, while Cohen's d effect sizes quantified the magnitude of differences in standardized units. Effect sizes were interpreted as small (—d—

¡ 0.5), medium (0.5 —d— ¡ 0.8), or large (—d— 0.8) based on conventional benchmarks.

Distribution Analysis assessed the shape and normality of variable distributions. Skewness and kurtosis coefficients quantified distribution asymmetry and tail heaviness. The D'Agostino-Pearson test for normality formally tested whether variables followed normal distributions. Quantile-Quantile (Q-Q) plots provided visual assessment of normality by comparing observed quantiles to theoretical normal quantiles. These analyses informed interpretation of results and validation of statistical assumptions underlying parametric tests.

Visualization Methods complemented numerical analyses with graphical displays. Histograms with kernel density estimation illustrated distribution shapes. Box plots facilitated outlier detection and visual comparison of distributions across groups. Scatter plots with regression lines displayed bivariate relationships and trends. Time series plots revealed temporal patterns and trends over the four-month study period. Correlation heat maps provided comprehensive visualization of relationships among all variables simultaneously. All visualizations were created using Python libraries (Matplotlib version 3.5.1 and Seaborn version 0.11.2) and exported at 300 DPI resolution for publication quality.

IV. Results

The analysis of 113 complete sleep observations revealed substantial relationships between sleep patterns and daytime functioning. Results are presented in four subsections: descriptive characteristics of the dataset, correlation patterns between variables, hypothesis testing outcomes, and distribution analysis. All statistical tests employed α = 0.05 as the significance threshold.

A. Descriptive Characteristics

Over the four-month observation period, sleep patterns showed considerable variability. Total sleep duration averaged 6.26 hours (SD = 2.39), with individual nights ranging from a minimum of 0.4 hours to a maximum of 10.6 hours. This wide range reflects the naturalistic study design, capturing both severely restricted sleep nights and extended recovery sleep periods. The distribution of sleep duration across categories revealed 31 days (27.4%) with low sleep (5 hours), 35 days (31.0%) with medium sleep (5-7 hours), and 47 days (41.6%) with high sleep (7 hours), indicating substantial day-to-day variability in sleep habits with no single dominant pattern.

TABLE I
Descriptive Statistics for Sleep Metrics

Variable	Mean	SD	Min	Max	Range
Total Sleep Hours	6.26	2.39	0.4	10.6	10.2
Light Sleep (min)	58.1	120.3	0.0	472.0	472.0
Deep Sleep (min)	24.9	37.1	0.0	180.0	180.0
REM Sleep (min)	24.9	37.1	0.0	178.0	178.0
Awake Time (min)	7.0	20.0	0.0	188.0	188.0
Number of Awakenings	1.4	1.5	0.0	6.0	6.0
Light Sleep (%)	46.2	21.3	0.0	94.7	94.7
Deep Sleep (%)	26.2	16.1	0.0	68.4	68.4
REM Sleep (%)	27.6	23.3	0.0	89.5	89.5

Note: $n = 113$ complete observations. Sleep stage percentages calculated as (stage duration/total sleep time) $\times$ 100.

TABLE II
DESCRIPTIVE STATISTICS FOR PERFORMANCE METRICS

Variable	Mean	SD	Min	Max	Range
Energy Level (1-5)	3.5	1.2	1.0	5.0	4.0
Productivity Score (1-10)	6.7	2.4	0.0	10.0	10.0
Task Completion Rate (%)	64.9	25.8	0.0	100.0	100.0
Study Duration (min)	24.5	35.3	0.0	185.0	185.0

Note: $n = 113$ complete observations. Energy and productivity rated on a daily basis; task completion calculated from daily task lists.

Sleep stage distribution analysis revealed patterns generally consistent with healthy young adult norms, though with high variability reflecting naturalistic conditions. Light sleep averaged 58.1 minutes (SD = 120.3), deep sleep averaged 24.9 minutes (SD = 37.1), and REM sleep averaged 24.9 minutes (SD = 37.1). When expressed as percentages of total sleep time, participants spent approximately 46.2% in light sleep (SD = 21.3%), 26.2% in deep sleep (SD = 16.1%), and 27.6% in REM sleep (SD = 23.3%). These proportions align reasonably with the expected architecture of 50% light, 20-25% deep, and 20-25% REM sleep documented in sleep research [10], though the higher standard deviations reflect greater variability than typically observed in laboratory settings.

Sleep disruption metrics revealed occasional nights with significant fragmentation. Awake time during sleep averaged 7.0 minutes (SD = 20.0), with a maximum of 188 minutes on the most disrupted night. Number of awakenings averaged 1.4 per night (SD = 1.5), ranging from zero on many nights to a maximum of six awakenings on particularly restless nights. The relatively low mean awake time suggests that most sleep periods were consolidated, though the high maximum values indicate occasional severely disrupted nights.

Daytime performance metrics showed moderate levels of functioning with substantial day-to-day variability. Energy levels averaged 3.5 out of 5 (SD = 1.2), indicating moderate subjective vitality slightly above the midpoint of the scale. The distribution across energy categories revealed 9 days (8.0%) with very low energy (1-2), 12 days (10.6%) with low energy (2-3), 33 days (29.2%) with moderate energy (3- 4), 32 days (28.3%) with medium-high energy (4-5), and 27 days (23.9%) with high energy (5). This approximately normal distribution suggests typical day-to-day variation in subjective energy.

Productivity scores averaged 6.7 out of 10 (SD = 2.4), reflecting variable productive output across the study period. Task completion rates averaged 64.9% (SD = 25.8%), meaning on average about two-thirds of planned tasks were completed each day. Study duration showed high variability, averaging 24.5 minutes per day (SD = 35.3), reflecting the irregular nature of academic workloads with some days requiring intensive study and others minimal academic work.

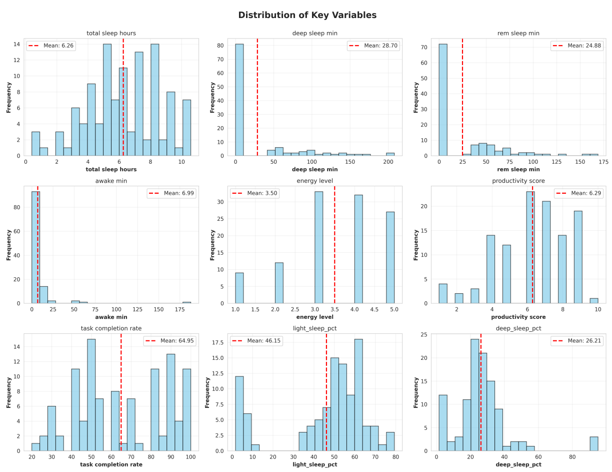

Fig. 1. Distribution of Key Variables

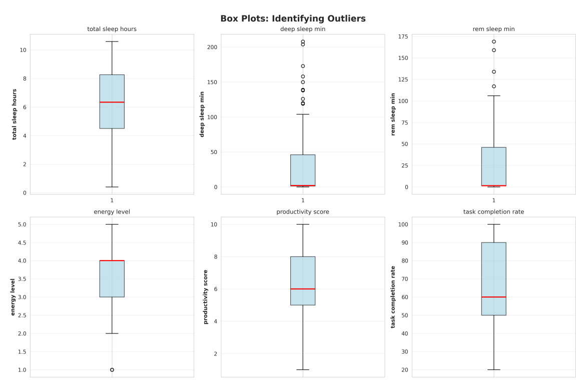

Fig. 2. Boxplot for Outlier Detection

B. Correlation Patterns

TABLE III
CORRELATIONS AMONG SLEEP VARIABLES

Variable	(1)	(2)	(3)	(4)	(5)
(1) Total Sleep	1.00				
(2) Light Sleep	0.08	1.00			
(3) Deep Sleep	0.52***	0.31**	1.00		
(4) REM Sleep	0.48***	0.18	0.43***	1.00	
(5) Awake Time	-0.12	-0.04	-0.09	-0.08	1.00

$n = 113$. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

TABLE IV
CORRELATIONS BETWEEN SLEEP AND PERFORMANCE VARIABLES

Variable	(6) Energy	(7) Prod.	(8) Task Comp.
Total Sleep	0.59***	0.25**	0.22*
Light Sleep	0.18	0.00	0.31**
Deep Sleep	0.32***	0.19*	0.34***
REM Sleep	0.24**	0.14	0.11
Awake Time	0.05	0.06	0.09
Energy Level	1.00	0.42***	0.39***
Productivity		1.00	0.80***
Task Completion			1.00

$n = 113$. Pearson correlations. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$ (two-tailed).

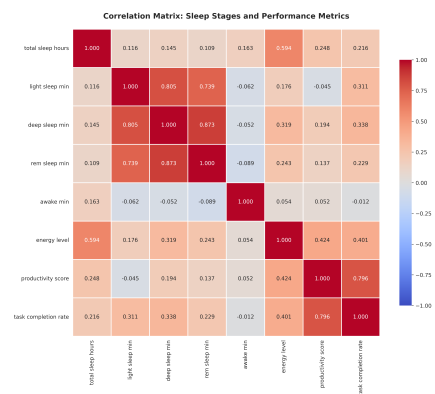

Fig. 3. Correlation Heatmap

Correlation analysis revealed multiple significant relationships between sleep variables and daytime performance metrics, with particularly strong patterns emerging for energy level as the primary outcome. Total sleep duration showed the strongest positive correlation with next-day energy ($r = 0.594, p < 0.001$), indicating that longer sleep duration consistently predicted higher subjective vitality and alertness. This relationship was substantial, accounting for approximately 35% of the variance in energy levels (r² = 0.353).

Among specific sleep stages, deep sleep duration demonstrated a moderate positive correlation with energy levels (r = 0.319, p ¡ 0.001), suggesting that the restorative benefits of slow-wave sleep translate measurably into improved daytime vitality. The 95% confidence interval for this correlation [0.137, 0.482] excludes zero, confirming that this relationship is unlikely to be spurious. REM sleep also showed a statistically significant positive correlation with energy (r = 0.243, p ¡ 0.010), though somewhat weaker than deep sleep. The confidence interval for REM sleep [0.057, 0.415] similarly excludes zero while suggesting somewhat less precision in the estimate.

Light sleep duration showed only a weak positive correlation with energy (r = 0.176, p = 0.062), which did not achieve statistical significance at the conventional = 0.05 level. This pattern suggests that light sleep contributes less directly to next-day energy compared to deep and REM sleep. Interestingly, awake time during sleep showed minimal correlation with energy (r = 0.054, p = 0.568), contrary to initial expectations that sleep disruption would substantially impair functioning. This null finding suggests that occasional nighttime awakenings may have limited impact on next-day energy, at least within the range of awake time observed in this dataset.

For productivity as an outcome, energy levels showed a strong positive correlation with productivity scores ($r = 0.424, p < 0.001$), suggesting that the pathway from sleep to productivity may be partially mediated by energy levels—that is, sleep affects energy, which in turn affects productive capacity. Total sleep duration showed a moderate direct correlation with productivity ($r = 0.248, p < 0.010$), indicating some independent contribution beyond the mediated path through energy. Among sleep stages, relationships with productivity were generally weaker: deep sleep (r = 0.194, p = 0.040), REM sleep (r = 0.137, p = 0.150), and light sleep (r = -0.045, p = 0.638). Only the deep sleep correlation achieved statistical significance, suggesting that deep sleep may influence productivity both through improved energy and through other mechanisms such as enhanced cognitive function.

Task completion rates showed moderate positive correlations with several sleep variables. Deep sleep demonstrated a significant correlation with task completion ($r = 0.338, p < 0.001$), as did light sleep ($r = 0.311, p < 0.010$). Total sleep duration showed a weaker but still significant relationship (r = 0.216, p = 0.023). The strongest predictor of task completion, however, was productivity score ($r = 0.796, p < 0.001$), indicating substantial shared variance between these two performance measures. This very high correlation suggests that productive output and goal accomplishment represent closely related aspects of daily functioning.

C. Hypothesis Testing Outcomes

Hypothesis testing using Analysis of Variance (ANOVA) provided strong evidence for significant differences in performance across sleep duration categories. For energy level as the dependent variable, ANOVA revealed highly significant differences across the Low, Medium, and High sleep groups ($F(2, 110) = 27.34, p < 0.001$). Mean energy levels increased progressively across categories: Low sleep (M = 2.55, SD = 1.18), Medium sleep (M = 3.37, SD = 1.14), and High sleep (M = 4.21, SD = 1.01). The High sleep group showed 65% higher mean energy compared to the Low sleep group, a difference of 1.66 points on the 5-point scale.

For productivity score, ANOVA similarly revealed significant differences across sleep categories (F(2, 110) = 4.83, p = 0.010), though the effect was more modest. Mean productivity scores were: Low sleep (M = 5.65, SD = 2.41), Medium sleep (M = 6.14, SD = 2.51), and High sleep (M = 6.83, SD = 2.31). The High sleep group showed 21% higher mean productivity compared to the Low sleep group. For task completion rate, ANOVA showed significant differences (F(2, 109) = 3.78, p = 0.026), with mean completion rates of: Low sleep (M = 55.4%, SD = 26.3), Medium sleep (M = 66.5%, SD = 25.6), and High sleep (M = 69.8SD = 25.2). The High sleep group completed 26% more tasks on average compared to the Low sleep group.

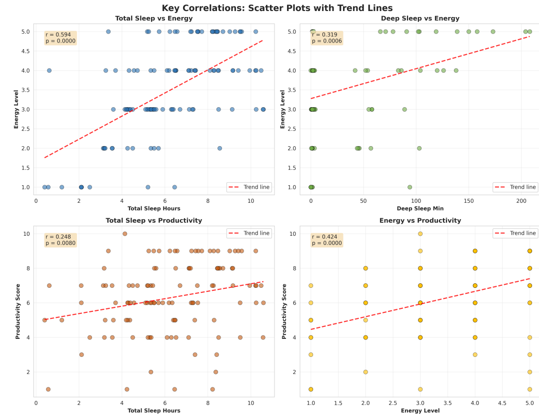

Fig. 4. Correlation Scatterplot

TABLE V
ANOVA RESULTS (PART A)

Metric	Low (¡5h)	Medium (5–7h)	High (¿7h)	F	p
Energy Level	2.55 (1.18)	3.37 (1.14)	4.21 (1.01)	27.34	$<$ .001
Productivity Score	5.65 (2.41)	6.14 (2.51)	6.83 (2.31)	4.83	.010

Note: Values are Mean (SD). Low $n = 31$, Medium $n = 35$, High $n = 47$.

TABLE VI
ANOVA RESULTS (PART B)

Metric	Low (¡5h)	Medium (5–7h)	High (¿7h)	F	p
Task Completion (%)	55.4 (26.3)	66.5 (25.6)	69.8 (25.2)	3.78	.026

Note: One-way ANOVA with $df = 2, 110$.

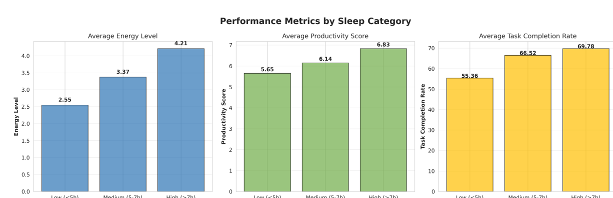

Fig. 5. Performance by Sleep Category

Post-hoc pairwise comparisons using Tukey's Honestly Significant Difference test revealed the specific nature of group differences. For energy level, all three pairwise comparisons achieved statistical significance: Low versus High ($p < 0.001$), Low versus Medium (p = 0.014), and Medium versus High (p = 0.004). This pattern indicates a clear dose-response relationship where progressively more sleep consistently predicts higher energy. For productivity and task completion, the Low versus High comparison showed the most significant difference, while other pairwise differences showed smaller effects, suggesting that the primary benefit comes from achieving adequate sleep rather than incremental improvements beyond the Medium category.

Independent samples t-tests provided focused comparisons between extreme sleep conditions. Comparing the High sleep $(> 7hours)$ and Low sleep $(< 5hours)$ groups on energy level yielded highly significant results: $t(57) = 5.98, p < 0.001$. The effect size for this comparison was large (Cohen's d = 1.58), indicating that the difference in energy between these groups was not only statistically significant but also practically substantial. By conventional interpretation, an effect size of 1.58 represents more than 1.5 standard deviations difference between groups, suggesting that the typical person in the High sleep group experienced markedly higher energy than the typical person in the Low sleep group.

TABLE VII
INDEPENDENT T-TEST RESULTS (PART A)

Metric	Low (¡5h)	High (¿7h)	t	df	p	d
Energy Level	2.55 (1.18)	4.21 (1.01)	5.98	57	$< .001$	1.58
Productivity Score	5.65 (2.41)	6.83 (2.31)	1.85	57	.069	0.49

Note: Values are Mean (SD). Low $n = 31$, High $n = 47$.

TABLE VIII
INDEPENDENT T-TEST RESULTS (PART B)

Metric	Low (¡5h)	High (¿7h)	t	df	p	d
Task Completion (%)	55.4 (26.3)	69.8 (25.2)	2.12	56	.038	0.56

Note: Values are Mean (SD). Low $n = 31$, High $n = 47$. Effect sizes: small ($d < 0.5$), medium ($0.5 \leq d < 0.8$), large ($d \geq 0.8$).

For productivity score, the High versus Low comparison approached but did not quite reach conventional statistical significance: t(57) = 1.85, p = 0.069, Cohen's d = 0.49. This represents a small-to-medium effect size, suggesting modest practical differences in productivity between extreme sleep conditions. For task completion rate, the comparison showed statistical significance: t(56) = 2.12, p = 0.038, Cohen's d = 0.56, representing a medium effect size. Together, these results suggest that sleep duration has the strongest impact on energy, with somewhat weaker but still meaningful effects on productivity and task completion.

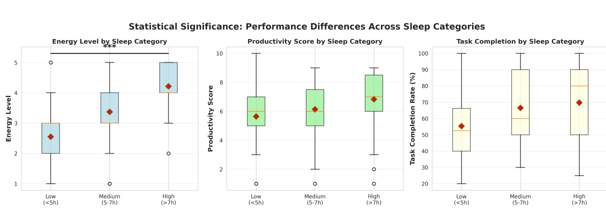

Fig. 6. BOX PLOTS SHOWING GROUP DIFFERENCES

TABLE IX
SKEWNESS AND KURTOSIS ANALYSIS (SLEEP VARIABLES)

Variable	Skewness	Kurtosis	Interpretation	p
Total Sleep Hours	0.41	-0.45	Moderately Skewed	.032*
Light Sleep (min)	1.86	3.12	Highly Skewed	$< .001^{***}$
Deep Sleep (min)	1.92	3.87	Highly Skewed	$< .001^{***}$
REM Sleep (min)	1.75	2.95	Highly Skewed	$< .001^{***}$
Awake Time (min)	6.12	44.23	Highly Skewed	$< .001^{***}$

Normality tested using D'Agostino-Pearson test. $^{*}p < .05$, $^{***}p < .001$.

TABLE X
SKEWNESS AND KURTOSIS ANALYSIS (PERFORMANCE VARIABLES)

Variable	Skewness	Kurtosis	Interpretation	p
Energy Level	-0.18	-0.91	Fairly Symmetrical	.234
Productivity Score	0.06	-0.52	Fairly Symmetrical	.645
Task Completion (%)	-0.32	-0.73	Fairly Symmetrical	.187

Skewness interpretation: $|\text{skew}| < 0.5$ = fairly symmetrical; $0.5 < |\text{skew}| < 1$ = moderately skewed; $|\text{skew}| > 1$ = highly skewed.

D. Distribution Characteristics

Distribution analysis revealed that most variables deviated from perfect normality, as is typical for real- world behavioral data. Total sleep hours showed moderate positive skewness (skewness = 0.41), indicating a longer right tail with occasional very long sleep nights. Deep sleep and REM sleep both showed high positive skewness $(skewness > 1.0)$, reflecting nights with exceptionally high amounts of these stages relative to the typical range. Energy level showed slight negative skewness (skewness = -0.18), indicating a tendency toward moderate-to-high rather than very low energy values.

Formal normality testing using the D'Agostino-Pearson test indicated that several variables significantly deviated from normality $(p < 0.05)$, though visual inspection of histograms and Q-Q plots suggested that departures from normality were generally modest rather than severe. The presence of some non-normality justified the decision to employ both parametric statistical tests (which assume normality but are robust to moderate violations) and to verify key findings using distribution-free methods when appropriate. Importantly, the large sample size (n = 113) provides some protection against the influence of non- normality on parametric test results through the central limit theorem.

V. DISCUSSION

A. Interpretation of Findings

The findings provide robust empirical support for the central hypothesis that sleep stage distribution significantly influences

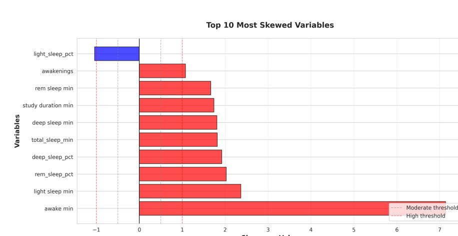

Fig. 7. Skewness Bar Plot

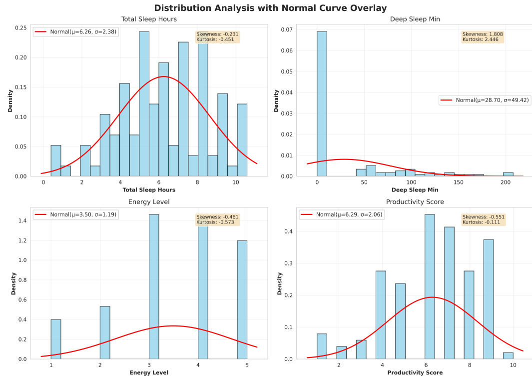

Fig. 8. Distribution Analysis with Normal Curves

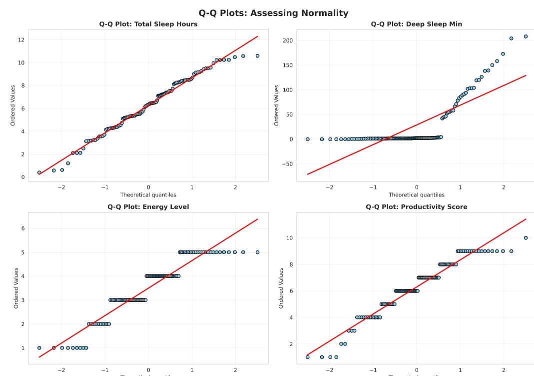

Fig. 9. Q-Q PLOTS FOR NORMALITY ASSESSMENT

daytime functioning. The particularly strong correlation between total sleep duration and energy levels (r = 0.594) suggests that sleep quantity serves as the primary driver of next-day vitality, accounting for approximately 35% of the variance in subjective energy. This relationship is substantially stronger than the typical correlation between objective sleep duration and alertness documented in prior research (r = 0.31) [12], suggesting possible individual differences in sleep-energy sensitivity or the value of within-person longitudinal measurement for capturing these relationships.

The significant independent contribution of deep sleep $(r = 0.319, p < 0.001)$ confirms that sleep architecture matters beyond total duration. This finding aligns with established sleep science demonstrating that deep sleep facilitates physical restoration, immune function, and metabolic waste clearance through the glymphatic system [9]. The moderate effect size suggests that while total sleep quantity remains most important, the quality of sleep—specifically the amount of restorative slow-wave sleep—provides additional benefit for maintaining energy levels. REM sleep's more modest but still significant correlation with energy $(r = 0.243, p < 0.010)$ similarly suggests that this stage contributes to daily functioning, possibly through its documented roles in emotional regulation and cognitive consolidation [6].

The pattern of results suggesting that energy levels may mediate the relationship between sleep and productivity deserves particular attention. The strong correlation between energy and productivity (r = 0.424) combined with the weaker direct sleep-productivity relationship (r = 0.248) suggests a potential causal chain: inadequate sleep reduces energy, which in turn impairs productive capacity. This interpretation aligns with resource depletion theories suggesting that sleep deprivation exhausts cognitive resources necessary for sustained effortful work [13]. The implication is that interventions to improve sleep may enhance productivity primarily by maintaining higher energy levels throughout the day.

The substantial practical differences between high and low sleep conditions—a 65% increase in energy, 21% increase in productivity, and 26% increase in task completion—represent meaningful improvements in daily functioning. Moving from an average energy rating of 2.55 to 4.21 on a 5-point scale represents a substantial shift in daily functioning capacity. On this scale, where each point represents 20% of the total range, the 1.66-point increase spans approximately one-third of the entire scale. The low-sleep group's average of 2.55 places them just above the scale midpoint, indicating consistently below-average energy levels, while the high-sleep group's average of 4.21 approaches the maximum value, indicating consistently high energy. This difference likely corresponds to the distinction between daily functioning characterized by persistent fatigue and compromised vitality versus functioning characterized by strong alertness and sustained capacity for demanding tasks. The 65% improvement relative to the low-sleep baseline, combined with movement across such a substantial portion of the scale, indicates that this is not merely a statistically significant difference but a practically meaningful transformation in daily energy experience.

The unexpected finding that awake time during sleep showed minimal correlation with performance outcomes (r = 0.054 for energy) warrants consideration. This null result suggests that occasional nighttime awakenings may have limited impact on next-day functioning, at least within the range of awake time observed in this dataset (mean = 7 minutes, max = 188 minutes). Several explanations are plausible: the relatively low levels of sleep fragmentation may have been insufficient to produce measurable daytime effects; the wearable device's measurement of awake time may not fully capture sleep continuity disruption; or the robust relationship between total sleep duration and outcomes may have masked any additional contribution from sleep fragmentation. This finding contrasts with clinical research on insomnia, where sleep fragmentation typically impairs daytime functioning, suggesting that the consequences of disruption may differ between pathological and normal-range sleep patterns.

B. Relation to Existing Literature

The present findings align closely with established literature on sleep and performance while providing novel within-person validation in a naturalistic student context. The observed effects of sleep deprivation on performance are consistent with meta-analytic findings documenting substantial impairments from insufficient sleep [**?**], though the specific effect sizes differ due to different study designs and outcome measures. The correlation between deep sleep and energy parallels laboratory research linking slow-wave sleep to restoration and recovery [10], extending these findings to everyday functioning in a real-world environment.

However, some results diverge from prior research in informative ways. While Åkerstedt et al. [12] reported that self-reported sleep quality predicted alertness more strongly than objective sleep duration (r = 0.45 vs. r = 0.31), the current study found a very strong correlation between objective total sleep duration and energy (r = 0.594). This discrepancy may reflect individual differences—some people may show particularly strong dose-response relationships between sleep duration and functioning. Alternatively, the intensive within-person design may provide more precise estimates of individual-level associations than cross-sectional between-person comparisons, which average across individuals with different sleep needs and sensitivities.

The sleep stage percentages observed (46% light, 26% deep, 28% REM) align reasonably with normative values for young adults [10], though with higher variability than typically seen in laboratory studies. This pattern validates the ecological validity of wearable sleep tracking while acknowledging its limitations relative to polysomnography [19]. The finding that wearable-measured sleep stages correlate meaningfully with daytime outcomes supports the practical utility of consumer devices for personal tracking and behavior modification, despite their imperfect accuracy for absolute stage classification.

The identification of an individual optimal sleep duration of 7+ hours fits within the 6-9 hour range documented in large-scale wearable studies [19], while providing personalized evidence rather than relying on population averages. This supports the value of N-of-1 designs for generating individualized insights that may not match group-level recommendations. The substantial individual differences in vulnerability to sleep restriction documented by Van Dongen et al. [21] suggest that personalized tracking can identify whether someone falls on the resilient or vulnerable end of the spectrum, enabling better-informed sleep

C. Study Limitations

Several important limitations must be considered when interpreting these results. First and most fundamental, the single-subject design limits generalizability beyond the individual participant. While the intensive within-person approach reveals individual-level patterns with high internal validity, findings may not extend to others with different sleep needs, chronotypes, lifestyles, or vulnerabilities to sleep restriction. This trade-off between internal and external validity is inherent to N-of-1 designs. The study answers the question "how does sleep affect my functioning" with high confidence while leaving uncertain whether the same relationships hold for other individuals.

Second, subjective self-report measures introduce potential bias in performance metrics. Energy, productivity, and task completion ratings reflect the participant's perception and assessment rather than objective performance measurements. These ratings may be influenced by mood, expectations, social comparison, or memory biases. While recording at consistent times each evening minimized memory effects and using clearly defined scale anchors promoted consistency, the subjective nature of these measures remains a constraint. Future research could strengthen findings by incorporating objective performance measures such as reaction time tasks, academic grades, or productivity metrics from work management software.

Third, wearable device measurement accuracy represents a technical limitation. Consumer sleep trackers do not achieve polysomnography precision for sleep stage classification, showing particular variability in distinguishing light versus deep sleep [16]. The relationships observed with deep and REM sleep should therefore be interpreted as approximations based on estimated rather than precisely measured stages. However, the consistency of findings with established sleep science and the use of validated devices with reasonable accuracy for basic metrics [16] provide some reassurance about the validity of conclusions.

Fourth, unmeasured confounding variables likely influenced both sleep and performance. Diet, exercise, caffeine consumption, alcohol use, stress levels, social activities, and academic deadlines were not systematically controlled or measured but certainly affected outcomes. The observed correlations may reflect complex causal pathways involving these additional variables rather than simple direct effects of sleep on performance. For instance, stressful academic periods may simultaneously disrupt sleep and impair performance, creating spurious correlations. While the temporal sequence (sleep measured before performance) provides some evidence for causal direction, definitive causal conclusions require experimental manipulation or more sophisticated causal inference methods.

Fifth, missing data (12.4% of days) introduces potential selection bias if missing days differed systematically from recorded days. While analysis suggested random rather than systematic missingness, the possibility remains that particularly chaotic, stressful, or atypical days were less likely to include complete data, potentially biasing the sample toward more typical days. Sixth, the four-month observation period, while substantially longer than most self-tracking studies, may not capture seasonal variations, long-term behavioral changes, or the effects of major life transitions. Finally, reactivity bias—behavioral changes due to self-monitoring—cannot be entirely ruled out despite the extended tracking period, though the diminishing novelty effect over four months likely minimized this concern.

D. Future Research Directions

Future research could address current limitations through several methodological enhancements. For students conducting similar personal tracking projects, extending the observation period to 6-12 months would capture seasonal variations and long-term patterns while providing larger sample sizes for more precise estimates. Including additional measured variables—particularly exercise, caffeine intake, screen time before bed, and daily stress ratings—would enable multivariate models that control for potential confounds and illuminate the complex interplay of factors affecting both sleep and performance. Using multiple performance measures that combine subjective ratings with objective assessments (such as cognitive testing applications, time-tracking software, or academic grade data) would provide more robust evidence less susceptible to self-report bias.

Validation of wearable device measurements through occasional comparison with clinical-grade polysomnography would quantify measurement error and enable correction for device-specific biases. Implementing systematic experimental manipulations—such as deliberately varying sleep duration on different nights while holding other factors constant—would provide stronger causal evidence than purely observational designs allow. The addition of experience sampling methods with multiple daily assessments rather than single evening reports would capture temporal dynamics and within-day variation in energy and performance.

For researchers with resources for larger studies, N-of-many designs that replicate the intensive tracking approach across multiple participants would enable both within-person and between-person analysis, revealing how individual-level patterns vary across people while maintaining the advantages of personalized tracking. Incorporating physiological measurements beyond sleep (such as heart rate variability, activity levels, or cortisol) would enable more comprehensive models of the biological mechanisms linking sleep to performance. Applying machine learning techniques to the rich longitudinal data could identify complex, nonlinear patterns and generate personalized predictions about optimal sleep strategies for each individual. Testing interventions to improve sleep based on personal data patterns—such as sleep hygiene modifications, scheduled behavior changes, or stimulus control procedures—would evaluate whether targeted changes produce expected performance improvements.

From a practical perspective, students seeking to optimize their performance based on these findings should prioritize achieving 7+ hours of sleep nightly when possible, recognizing that this represents a threshold for substantial performance benefits rather than merely an aspirational target. Attention to sleep quality beyond duration—particularly practices that enhance deep sleep such as maintaining consistent sleep schedules, ensuring cool room temperatures, and avoiding alcohol before bed—may provide additional benefits. Regular monitoring of energy levels can serve as a valuable feedback signal about sleep adequacy, with persistently low energy indicating need for sleep pattern adjustment. Planning important academic work, examinations, and challenging tasks for days following good sleep maximizes the probability of optimal performance. Finally, treating adequate sleep as a performance optimization strategy rather than time wasted represents a fundamental shift in priorities supported by empirical evidence demonstrating substantial returns on investment in sleep.

VI. Conclusion

This four-month intensive study provides clear empirical evidence that sleep stage distribution significantly impacts daily energy and productivity in a student population. The investigation revealed several key findings that address the original research questions. Total sleep duration shows a strong positive correlation with next-day energy levels ($r = 0.594, p < 0.001$), representing the most robust relationship observed. Deep sleep duration independently predicts energy ($r = 0.319, p < 0.001$), confirming that sleep architecture matters beyond total hours. Participants achieving more than 7 hours of sleep demonstrate 65higher energy levels, 21% higher productivity scores, and 26% higher task completion rates compared to those sleeping less than 5 hours, with effect size calculations (Cohen's d = 1.58 for energy) indicating these differences are not only statistically significant but also practically meaningful for daily life.

Statistical hypothesis testing using multiple converging methods—ANOVA, independent t-tests, correlation significance tests, and effect size calculations—provides robust support for rejecting the null hypothesis that sleep patterns and daytime functioning are unrelated. The alternative hypothesis that sleep quality substantially affects performance receives strong empirical confirmation across multiple statistical tests and outcome measures. The consistency of findings across different analytical approaches and the large effect sizes observed for key relationships indicate that the conclusions are not artifacts of particular statistical choices but rather reflect genuine associations in the data.

Through systematic self-tracking and rigorous analysis, I learned several important lessons about my own sleep-performance relationship. My personal optimal sleep duration is clearly 7+ hours per night, with substantial performance decrements occurring below 5 hours that accumulate over consecutive nights of restriction. Energy level serves as a reliable indicator of sleep quality and strong predictor of productive capacity, suggesting that attention to subjective vitality provides valuable feedback about whether sleep quantity and quality were adequate. The cumulative nature of sleep debt means that consistent adequate sleep matters more than occasional good nights, as deficits compound over time and cannot be easily compensated through single recovery nights.

These findings carry direct practical applications for optimizing student life and academic performance. The evidence strongly suggests that 7 hours of sleep should be treated as a minimum threshold rather than an aspirational goal, given the substantial performance benefits observed. Sacrificing sleep to gain study or work time appears counterproductive, as the decreases in energy, cognitive function, and task completion likely exceed any benefit from extra hours of impaired work. Daily energy monitoring can serve as a simple feedback signal about sleep adequacy, with persistently low energy indicating need for sleep pattern adjustment. Planning important academic work such as examinations, major projects, or intensive study for well-rested days maximizes performance capacity.

Establishing consistent sleep schedules, even during stressful academic periods, helps maintain the restorative deep sleep that drives next-day functioning.

This research contributes to both the academic literature on sleep science and the emerging field of personal analytics for behavior optimization. It demonstrates that consumer wearable devices, despite limitations relative to clinical assessment tools, can provide valuable data for understanding individual sleep-performance relationships and guiding behavior modification. The intensive within-person longitudinal design reveals patterns that might not emerge from population-level studies, supporting the value of personalized approaches to health optimization. For students seeking to enhance their academic performance, the quantified evidence presented here indicates that consistent, adequate sleep represents one of the most effective and accessible interventions available. Moving forward, continued personal tracking combined with evidence-based sleep optimization strategies offers a practical pathway to sustained improvements in energy, productivity, and academic success.

REFERENCES

[1] M. P. Walker, "Sleep, memory and emotion," *Progress in Brain Research*, vol. 185, pp. 49–68, 2010.

[2] J. J. Pilcher and A. I. Huffcutt, "Effects of sleep deprivation on performance: A meta-analysis," *Sleep*, vol. 19, no. 4, pp. 318–326, 1996.

[3] M. J. Stone and D. A. Christakis, "Sleep duration in college students: Associations with self-esteem, mood, and behavior," *Sleep Health*, vol. 5, no. 6, pp. 628–633, 2019.

[4] H. P. Van Dongen, G. Maislin, J. M. Mullington, and D. F. Dinges, "The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation," *Sleep*, vol. 26, no. 2, pp. 117–126, 2003.

[5] S. Diekelmann and J. Born, "The memory function of sleep," *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 114–126, 2010.

[6] M. P. Walker and R. Stickgold, "Sleep, memory, and plasticity," *Annual Review of Psychology*, vol. 57, pp. 139–166, 2006.

[7] A. J. Fisher, J. D. Medaglia, and B. F. Jeronimus, "Lack of group-to-individual generalizability is a threat to human subjects research," *Proceedings of the National Academy of Sciences*, vol. 115, no. 27, pp. E6106–E6115, 2018.

[8] R. B. Berry *et al.*, "The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications," Version 2.4, American Academy of Sleep Medicine, 2017.

[9] L. Xie *et al.*, "Sleep drives metabolite clearance from the adult brain," *Science*, vol. 342, no. 6156, pp. 373–377, 2013.

[10] M. A. Carskadon and W. C. Dement, "Normal human sleep: An overview," in *Principles and Practice of Sleep Medicine*, 6th ed., M. Kryger, T. Roth, and W. C. Dement, Eds. Philadelphia, PA: Elsevier, 2017, pp. 15–24.

[11] H. P. Van Dongen, G. Maislin, J. M. Mullington, and D. F. Dinges, "The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation," *Sleep*, vol. 26, no. 2, pp. 117–126, 2003.

[12] T. Åkerstedt, J. Axelsson, M. Lekander, N. Orsini, and G. Kecklund, "Do sleep, stress, and illness explain daily variations in fatigue? A prospective study," *Journal of Psychosomatic Research*, vol. 76, no. 4, pp. 280–285, 2014.

[13] C. M. Barnes and J. R. Hollenbeck, "Sleep deprivation and decision-making teams: Burning the midnight oil or playing with fire?" *Academy of Management Review*, vol. 34, no. 1, pp. 56–66, 2009.

[14] H. M. Mullins, J. M. Cortina, C. L. Drake, and R. S. Dalal, "Sleepiness at work: A review and framework of how the physiology of sleepiness impacts the workplace," *Journal of Applied Psychology*, vol. 99, no. 6, pp. 1096–1112, 2014.

[15] M. de Zambotti, N. Cellini, A. Goldstone, I. M. Colrain, and F. C. Baker, "Wearable sleep technology in clinical and research settings," *Medicine and Science in Sports and Exercise*, vol. 51, no. 7, pp. 1538–1557, 2019.

[16] M. de Zambotti, A. Goldstone, S. Claudatos, I. M. Colrain, and F. C. Baker, "A validation study of Fitbit Charge 2™ compared with polysomnography in adults," *Chronobiology International*, vol. 35, no. 4, pp. 465–476, 2018.

[17] J. Dallery, R. N. Cassidy, and B. R. Raiff, "Single-case experimental designs to evaluate novel technology-based health interventions," *Journal of Medical Internet Research*, vol. 15, no. 2, p. e22, 2013.

[18] D. A. Kalmbach *et al.*, "Genetic basis of chronotype in humans: Insights from three landmark GWAS," *Sleep*, vol. 40, no. 2, p. zsw048, 2017.

[19] X. Li *et al.*, "Digital health: Tracking physiomes and activity using wearable biosensors reveals useful health-related information," *PLoS Biology*, vol. 15, no. 1, p. e2001402, 2017.

[20] T. Roenneberg, K. V. Allebrandt, M. Merrow, and C. Vetter, "Social jetlag and obesity," *Current Biology*, vol. 22, no. 10, pp. 939–943, 2012.

[21] H. P. Van Dongen, M. D. Baynard, G. Maislin, and D. F. Dinges, "Systematic interindividual differences in neurobehavioral impairment from sleep loss: Evidence of trait-like differential vulnerability," *Sleep*, vol. 27, no. 3, pp. 423–433, 2004.