Sleep is important: How circadian-timed trauma keeps mice up at day and the societal implications that keep me up at night

Asad I. Beck

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Horacio O. de la Iglesia, Chair

Chet T. Moritz

Timothy E. Brown

Program Authorized to Offer Degree:

Neuroscience

University of Washington

Abstract

Sleep is important: How circadian-timed trauma keeps mice up at day and the societal implications that keep me up at night

Asad I. Beck

Chair of the Supervisory Committee:
Horacio O. de la Iglesia
Department of Biology

This thesis centers on the vital process of sleep and its profound influences on various dimensions of human life, with focus on connections between sleep and aspects of well-being. This thesis adopts a multidisciplinary approach, wherein each chapter examines sleep from a different perspective with the goal of fostering a nuanced dialogue on the multifaceted nature of sleep. Chapter I lays the groundwork for understanding sleep's adaptive nature and its essential role in survival, exploring the neurobiology of sleep timing and its functions in promoting physical and mental well-being. The chapter also delves into the causes of sleep disruption and introduces various research methods, setting the stage for subsequent chapters.

Chapter II, representing the behavioral perspective of sleep, investigates the effects of contextual fear on circadian rhythms. The study reveals that exposure to a fearful context can re-

establish abnormal circadian activity patterns, contributing to an understanding of the intersection between sleep disruptions and conditions like post-traumatic stress disorder. Chapter III, showcasing a computational perspective of sleep, introduces a web-based automated sleep scoring method This tool, utilizing machine learning, exhibits high accuracy in scoring sleep in mice, showcasing its potential to expedite the sleep research process. Chapter IV, using a philosophical perspective of sleep, examines historical biases in sleep research that continue to perpetuate sociocultural disparities. The framework of intersectionality is employed to highlight biases against women and non-White individuals, emphasizing the need of recognizing and accounting for systemic bias in scientific inquiry.

The final chapter synthesizes these perspectives to underscore the unique dialogue generated by the juxtaposition of seemingly disparate viewpoints. The interdisciplinary nature of the thesis allows for a comprehensive exploration of sleep, unveiling insights that may be missed when studied in isolation. This approach emphasizes the importance of adopting diverse perspectives to capture the complexities inherent in sleep, promoting inclusivity and representation in scientific exploration. Such unique dialogues that emerged when concurrently considering different viewpoints during the process of researching sleep subsequently pave the way for unique avenues of inquiry, thereby advocating for increased conscious effort to unite different viewpoints in sleep research.

TABLE OF CONTENTS

List of Figures vii

List of Tables viii

Chapter I. Making the bed: An introduction to sleep and sleep research 1

I.i Neurobiology of sleep timing. 4
I.i.1 Origin of the sleep-wake cycle. 5
I.i.2 Suprachiasmatic nucleus and peripheral clocks. 6
I.ii Functions of sleep, effects of sleep disruption. 8
I.ii.1 Restorative processes of sleep, consequences of sleep disruption and deprivation. 9
I.ii.2 Sleep stages and sleep architecture. 10
I.iii Causes of sleep disruption and deprivation. 13
I.iii.1 Internal factors: genetic predispositions and comorbidities. 14
I.iii.2 External factors: physical and sociocultural environment. 15
I.iii.3 Interactions between internal and external factors. 17
I.iv Methods of studying sleep. 19
I.iv.1 Human studies. 19
I.iv.2 Animal models. 20
I.iv.3 Artificial intelligence. 21
I.v Conclusions and thesis outline. 23
I.vi Figures. 25

Chapter II. Sleep with one eye open: Contextual fear and abnormal circadian rhythms 27

II.i Methods. 29
II.i.1 Contextual fear paradigm. 29
II.i.2 Data analysis. 30
II.ii Results. 32
II.ii.1 Following entrainment to nocturnal fear, re-exposure to fear context is enough to elicit return of abnormal circadian behavior. 32
II.ii.2 Footshock intensity determines speed of entrainment to fear stimuli. 33
II.ii.3 Magnitude of change in activity during contextual recall depends on strength of entrainment for DF but not LF mice. 34
II.ii.4 Activity patterns during contextual recall do not return to baseline after a month. 35
II.ii.5 Males and females display comparable responses during contextual recall. 36
II.iii Discussion. 36
II.iv Figures. 39

Chapter III. Sleeping on the job: Publicly available automated sleep scoring 47

III.i Methods. 49
III.i.1 Animals and housing conditions. 49
III.i.2 Electrocorticographic recordings – de la Iglesia Lab. 50

III.i.3 Telemetric electrocorticographic recordings – de la Iglesia Lab. 51
III.i.4 Electrocorticographic recordings – Kalume Lab. 51
III.i.5 Behavioral experiments and environmental conditions. 52
III.i.6 Signal processing and feature engineering. 52
III.i.7 Software development. 53
III.i.8 Supervised learning algorithm selection. 54
III.i.9 Inter- and intra-scorer reliability. 55
III.i.10 Training dataset 56
III.i.11 Correlation matrix and dendrogram. 56
III.i.12 Dimensionality reduction using sequential feature selection. 56
III.ii Results. 57
III.iii Discussion. 64
III.iv Tables. 69
III.v Figures. 73

Chapter IV. Sleep, with both eyes open: systemic bias, intersectionality, and sleep research 75

IV.i Factors of disparities in health and sleep. 80
IV.i.1 Biological sex and gender identity. 80
IV.i.2 Racial/ethnic identity. 83
Other sources of sleep disparities and intersectionality. 86
IV.ii Steps of research and bias. 88
IV.ii.1 Steps of biomedical research — from research purpose to data collection. 89
IV.ii.2 Human studies. 92
IV.ii.3 Animal models. 94
IV.iii Quantitative data analysis and algorithmic bias. 96
IV.iii.1 Algorithmic bias. 98
IV.iv Accounting for systemic bias. 101
IV.iv.1 Quantifying systemic bias. 102
IV.iv.2 Disciplinary reflexivity and accounting for subjectivity. 102
IV.v Conclusions. 103

Chapter V. Lying in bed: Conclusions. 106

V.i Chapter summaries. 107
V.i.1 Chapter I. Making the bed: An introduction to sleep and sleep research 108
V.i.2 Chapter II. Sleep with one eye open: contextual fear and abnormal circadian rhythms. 108
V.i.3 Chapter III. Sleeping on the job: client-side, publicly available automated sleep scoring. 109
V.i.4 Chapter IV. Sleep, with both eyes open: systemic bias, intersectionality, and sleep research. 110
V.ii Cross-chapter synthesis, contributions, and conclusions. 110

Bibliography 112

Appendix A. Supplemental Materials 129

LIST OF FIGURES

Figure I-i. Two-process model of sleep. 25
Figure I-ii. Representation of typical sleep architecture during a eight-hour sleep bout.. 26
Figure II-i. Experimental conditions and corresponding activity patterns. 39
Figure II-ii. Change in activity during critical time windows over each experimental condition. 41
Figure II-iii. Speed of entrainment and foodshock intensity. 42
Figure II-iv. Linear relationship between strength of entrainment during fear induction and contextual recall activity during ZT0-ZT1. 43
Figure II-v. Extinction of fear entrainment over four weeks, by shock group and footshock intensity. 44
Figure II-vi. Sex differences in activity by shock group and condition..... 46
Figure III-i. Correlation matrix of the features from the complete dataset..... 73
Figure III-ii. Performance under sequential feature selection of the BCRF algorithm trained on the complete training dataset using a one-step approach to classification..... 74

LIST OF TABLES

Table III-i. Inter- and intra-rater reliability metrics for manual scores used in training datasets. 69
Table III-ii. Classification accuracy by algorithm and method. 70
Table III-iii. Sleep stage classification accuracy using subsets of the training data. 71
Table III-iv. Comparison of sleep stage classification accuracy of data from another laboratory and data from another data collection modality. 72

ACKNOWLEDGEMENTS

My deepest, deepest gratitude goes to my advisor, Dr. Horacio de la Iglesia, for being such an amazing mentor these past five years. Truly, without his guidance and support, none of this would have been possible — whether that be through getting the work necessary done and written, or even carrying the sheer stress of being a graduate student. I truly cannot thank him enough.

I would also like to thank each of my committee members, each who have supported me in their own ways, and have contributed so much to my development over the years. I truly believe I could not have asked for a better committee than the one I have had.

An additional thank you to the Neuroethics group at the University of Washington. It truly felt like a second lab/home these past five years, and many of the thoughts shown in this thesis have been greatly influenced by merely being in their presence.

Finally, I would like to thank individuals who made up the de la Iglesia lab — all of whom have similarly supported me and my work over the years, and the seemingly short sentence seen here does not fully capture my gratitude.

There are so many more people I want to acknowledge and thank, but I fear I don't have enough space to write another thesis.

DEDICATION

This thesis is dedicated to the friends I've come to know and love here in Seattle, who have kept me sane, alive, and loved.

This thesis is dedicated to the mentors I have had over the years, who have guided me in ways I did not know I needed. I could never hope to fully express my gratitude.

This thesis is dedicated to my two beautiful cats, Ella and Sophie, without whom I would have lost my mind much sooner.

This thesis is dedicated to those I've had to leave behind.

Chapter I. MAKING THE BED: AN INTRODUCTION TO SLEEP AND SLEEP RESEARCH

Sleep is essential for life. As humans, we spend roughly ⅓ of our lives asleep. Along with air, food, water, shelter, clothing, and human contact, sleep is considered to be a fundamental human necessity.[1] When deprived of any one of these necessities, an individual's general functions — including physical and mental — are drastically impacted, potentially threatening the individual's continued survival.[1–3] Investigations into the role of sleep have gone back centuries, with the Ancient Greek philosopher Aristotle further highlighting the oppositional status sleep appears to hold with wakefulness, and the temporary loss of function that seemingly occurs during sleep.[4] Despite the vulnerable position it places us in, we continue to feel pressure to sleep.[5] From the effects of sleep deprivation and its debilitating impacts on wellbeing,[6–9] to how we individually experience sleep (e.g., a bad night's sleep leaving us with low energy and a bad mood in the morning), or even how we communicate these experiences to others (e.g., sleep and dreams are a frequent topic in contexts ranging from casual conversation to acclaimed poetry), it is clear that sleep has a critical role in our wellbeing.[5]

This holds true for most animal species; while the average amount of time spent asleep differs (humans, ~8 hours per day; mice, ~12 hours; elephants, ~2 hours), the need for and purpose of sleep is generally maintained from species to species.[5,10] Sleep may, in part, serve to balance energy expended with energy acquired during wake.[11] For instance, when an animal is not actively acquiring and consuming food for energy, merely expending this energy without replenishment is a risk to survival.[5,10] Similarly, there may be brain-specific metabolic demands that can only be addressed via restful sleep (as opposed to merely resting while conscious or non-restful

unconsciousness), and not addressing these needs may lead to impaired functioning.[10–13] As part of this, sleep may serve to best support specific metabolic processes and healthy development (e.g., babies tend to require greater amounts of sleep in part due to higher energy demand and to sustain neurodevelopmental processes),[5,11] as well as to serve to bolster restorative processes in the body (apparent during times of illness when the body requires greater amounts of sleep[5]) and in the brain (when the brain clears away accumulated toxins[10,12,13]) note 1.

Another core function of sleep is its promotion of memory consolidation.[14] While memory encoding and recall primarily occur during wake, the brain undergoes processes during sleep to organize and strengthen these new memories.[15] Memory consolidation primarily occurs during stages of deep sleep, during which traces of new memories are reactivated and redistributed throughout the brain, thereby improving memory retention.[14] These periods of deep sleep are separated by periods of a phase called rapid eye movement (REM) sleep – a state characterized by increased neural activity despite continued body immobility (leading to this state previously being referred to as 'paradoxical sleep'). In addition to being the sleep stage primarily associated with the experience of vivid dreams, REM may also promote emotional memory processing.[16] Finally, sleep deprivation acutely impairs memory function (in addition to other cognitive functions, such as attention and learning),[6] further emphasizing the crucial role of sleep in maintaining healthy memory and cognitive function.

Given its accompanying vulnerability, sleep may also interact with and influence emotional wellbeing. A common adaptation in social species that face increased risk of predation is group sleeping, wherein members of the group sleep in close proximity to one another.[17] Occupying the same space as another member in a social group can give a sense of security due to the presence

[1] For further evaluation of sleep in the context of health and development, please refer to section 2, ‘functions of sleep, effects of sleep disruption.’

of other members in the group, concurrently leading to greater chances of surviving the night and better sleep quality.[18,19] Compared to individual sleeping, group sleeping leads to positive changes in sleep quality in a variety of species including mice,[20] doves and mallards,[21] rhesus monkeys,[22] and humans.[23] This is particularly relevant for beings with limited self-care abilities, such as human infants who are dependent on caregivers during their early years. It is likely due to this pressure for security that, in humans, sharing a bed with another may function as a form of psychosocial support [note 2]: For humans, those who have never shared a bed as infants with their parents may grow to be more anxious and have lower self-esteem than those who had shared a bed[24]; for heterosexual couples, individuals report better sleep quality and emotional wellbeing after bed-sharing with their partner, despite objective measures indicating poorer sleep.[19] These examples show that the socio-cognitive benefits of group sleeping are conserved across age and species.

What I hope is clear in this introduction is the multifaceted impact of sleep, both good and bad. Just as sleep serves as a central pillar in physical and mental health, so it does to support psychosocial health. Rather than a period marked primarily by loss of function, as conceptualized by Aristotle, sleep may instead be a period marked by its support of many aspects of living despite or even due to the immobility that generally accompanies the state. Thus, to better understand the function of sleep and to set the stage for investigations described later in this dissertation, I will first introduce the mechanisms of sleep: In section 1, I walk through how the timing of sleep is regulated, with emphasis on circadian rhythms and neural mechanisms underlying these rhythms, as well as the different stages of sleep. In section 2, I explore the interactions between the brain

[2] Crucial to bed-sharing/co-sleeping is trust. Without trust, sharing a space — much less a bed — with another may prove more harmful than helpful. The potential security imparted would be annulled by the potential stress of allowing a stranger access to you when at your most vulnerable.[18] This topic will be more fully explored in section 3 of this chapter, as well as Chapter IV of the dissertation.

and the rest of the body in the context of sleep, as well as give a more thorough exploration of the role of sleep and the effects of sleep disruption. In section 3, I examine potential routes through which sleep disruptions may occur, focusing on internal factors, external factors, and interactions between the two. In section 4, I will introduce three primary methods of studying sleep: human studies, animal models, and the use of artificial intelligence. Finally, in the final section, I will provide an outline for the following four chapters, as well as justification for the structure of this dissertation. While this will require simplifying some details (entire books have been written about sleep, let alone a single function of sleep), my goal is to introduce concepts that will later be further explored in forthcoming chapters, with focus on those concepts that will serve as a consistent undercurrent throughout.

I.i NEUROBIOLOGY OF SLEEP TIMING.

With the wake phase capturing the majority of animal intentional behavior, the sleep phase represents its complementary counterpart in the *sleep-wake cycle*. When we are asleep, we are largely immobile, have higher thresholds to respond to stimuli, and cognitive functioning that occurs during wake (e.g., consciousness) is generally absent.[25] Colloquially, we think of the sleep-wake cycle as being strongly coupled with the day/night (or light/dark, LD) emerging from the solar day. For diurnal species such as humans, this means the animal is awake and active during the day (when it is light) and asleep during the night (when it is dark). Conversely, for nocturnal species such as mice, the animal is most active during the night and sleeps during the day. While there are other patterns of sleep-wake cycles,[26] what is clear across each of these activity patterns is their close relationship with the solar day — a relationship primarily relies on *internal* processes of *endogenous clocks* and *homeostatic sleep pressure*.[3,27]

I.i.1 *Origin of the sleep-wake cycle.*

One of the most significant contributors to the sleep-wake cycle is the organism's circadian system.[4] Circadian rhythms have been found not only in animals, but also in plants,[3] fungi,[3] and even single-cell organisms,[4,28] underscoring the widespread presence of circadian rhythms and their likely adaptive value in many forms of life. These are biological, endogenous rhythms that synchronize numerous physiological functions to a daily 24-hour cycle (Figure 1, red line). In absence of an exogenous time cue (or *zeitgeber*; German for 'time-giver'), animals adopt a pattern of behavior known as *free-running*,[29] wherein the animal's circadian rhythms adopt a period of approximately, but not exactly, 24 hours (with some animals having periods somewhat shorter, such as in nocturnal organisms, and others somewhat longer such as in diurnal organisms).[3,30] This definition is captured by the term's combination of the Latin words *circa* (meaning 'about') and *dian* (meaning 'a day'), signifying that, in absence of exogenous time cues, the animal's endogenous rhythms approximate the length of a day. In order to generate rhythmic outputs that are as close to 24 hours as possible, biological clocks rely on the presence of a cyclic, exogenous cue — a zeitgeber — to *entrain* them, allowing the timing of endogenous rhythms to match the timing of the exogenous cue.[3] This term also captures the close relationship between endogenous clocks and the LD cycle (arguably the most salient *zeitgeber*), but does not preclude the ability for endogenous clocks to entrain to other stimuli so long as these stimuli reliably occur at regular intervals and, crucially, have a period of 24 hours.[30] Consequently, this means that cyclic non-photic stimuli, such as temperature cycles,[29,30] food availability,[29–31] social cycles,[29] or even time-specific fear[32] note 3, may also serve to entrain circadian rhythms[33].

[3] For further discussion on fear and entrainment, as well as on conflicting zeitgebers, please see Chapter II of this dissertation.

Acting relatively independently yet in tandem with the animal's endogenous rhythm, a second contributor to the sleep-wake cycle is homeostatic pressure for sleep (Figure 1, purple line). Unlike circadian rhythms (which run autonomously and are in principle independent of the sleep-wake state), homeostatic pressure can be seen as the accumulation of sleep-promoting substances (e.g., cerebral adenosine monophosphate and extracellular adenosine, both associated with lower available energy for the brain[11]) that occurs during the period of time spent awake and only decreasing when asleep.[4,34] When this pressure is not sufficiently relieved during a bout of sleep (due to not enough sleep or poor sleep quality, as *more* sleep is not necessarily *restorative* sleep[35]), it manifests as what is known as 'sleep debt' – 'left-over' homeostatic pressure from the previous day that may impair functioning during wake.[36,37]

These two factors — the animal's endogenous circadian rhythm and the time spent awake-dependent, homeostatic sleep pressure — make up the two-process model of sleep-wake regulation, and they coalesce to determine the animal's *sleep drive* (Figure 1, blue line).[4] When this sleep drive is at its nadir (often shortly after wake), the animal is at its most alert and awake; when it's at its peak, the animal's energy is at its lowest and most driven to sleep. However, both processes are not necessarily static. For instance, older human adults generally require less sleep than young adults, likely due to decreased rise of homeostatic pressure during wake as well as weaker circadian rhythm amplitudes.[36] While the neural basis for homeostatic regulation of sleep are not fully characterized, the region most responsible for maintenance of circadian rhythms throughout the lifespan is the *suprachiasmatic nucleus* of the hypothalamus.[4,34,36]

I.i.2 *Suprachiasmatic nucleus and peripheral clocks.*

The suprachiasmatic nucleus (SCN) is a small cluster of neurons located in the hypothalamus that serves circadian regulatory functions for most behavioral and physiological

processes, including entrainment to photic stimuli, timing of hormonal secretion, and coordination of 'peripheral clocks' found throughout the body.[4,34] The SCN regulates photic entrainment through projections received from melanopsin-containing intrinsically photosensitive retinal ganglion cells (ipRGCs)[38] via the retinohypothalamic tract, where light activation of ipRGCs increases SCN activity.[4,34] Even in the absence of photic input, the SCN's clock continues to function, regulating the release of the sleep-associated hormone melatonin as a result of its control of pineal gland activity timing (albeit in a near-24-hour rhythm).[4,34,39] In this instance, the pineal gland represents an example of a peripheral clock whose timing is governed by the SCN – the 'central clock'. This role as the central clock becomes evident when typical SCN function is disrupted:[4,34] For instance, when the SCN is isolated or ablated, behavioral and physiological circadian rhythms are lost, and peripheral clocks do not cease to function but instead fall out of sync with one another without the coordinating signals of the SCN.[4,40] Additionally, when transplanting an SCN from a donor animal to a host animal, the endogenous period of circadian output (determined by locomotor activity) matches that of the donor animal.[41] This highlights the hierarchical relationship where the SCN coordinates activity in peripheral clocks and thereby indirectly controls expression of circadian rhythms throughout the body.

While this hierarchy is the case for most situations,[39] it can be altered under specific conditions. As previously mentioned, organisms are capable of entraining to non-photic stimuli,[29–33] and there are primarily two scenarios in which entrainment to non-photic stimuli can exert control of circadian expression above that generated by the SCN: the first involves the photic pathway being disrupted in some way, whether through the absence of an LD cycle (e.g., entrainment to food restriction or behavioral arousal can occur under constant lighting conditions[33]), hampering retinal inputs (e.g., blind mice, bats, and humans are capable of entraining

to social stimuli[33]), or blocking SCN outputs to peripheral clocks (e.g., applying tetrodotoxin to the SCN, preventing its synaptic signaling while preserving cellular rhythmicity[42]). In this scenario, with the output of the central clock effectively dampened, outputs from peripheral clocks may drive behavior. The second scenario is one in which a non-photic zeitgeber is sufficiently potent and physiologically relevant to entrain peripheral clocks and oppose or even supersede control of the SCN even in presence of an LD cycle. For example, rabbits may be capable of entraining to restricted food availability cycle during a standard LD cycle, despite conflicting zeitgebers.[31,43] This suggests that, even in the presence of conflicting zeitgebers, an animal may be able to shift behavioral rhythms to respond to stimuli that may be more biologically relevant note 3, highlighting the plasticity of sleep timing and its role in survival.

I.ii FUNCTIONS OF SLEEP, EFFECTS OF SLEEP DISRUPTION.

Part of the importance of sleep in the physical and mental wellbeing across all animals may lie in its restorative qualities.[10] For example, when ill, others often provide the gentle suggestion, 'get some rest', emphasizing our recognition of the unseen recovery that often occurs when we get quality sleep. Even when otherwise healthy, we still informally recognize sleep hygiene as something we ought to strive for; despite inconclusive research on whether it truly serves as an effective therapeutic practice note 4, we acknowledge that we *ought* to avoid eating, exercising, stressing oneself out, consuming caffeine or other stimulants, and staring at one's phones prior to

[4] As indicated in the referenced article, potential reasons behind these inconclusive results regarding sleep hygiene are: 1, lacking or unclear results from direct evaluations of sleep hygiene; 2, vague and inconsistent recommendations; 3, lack of guidance regarding implementation. They also note the lack of attention given on individual differences (e.g., sleep hygiene recommends not taking naps, despite some potentially needing naps to function or some societies experiencing naps as integral to their cultures[44]) and that some of these behaviors, such as consuming alcohol, may occur as a result from pre-existing disorders.[45] In effect, the concept of sleep hygiene appears to behave as more of a prescriptive ideal than a reality of what is most healthy for the individual. These ideas will be explored more fully in Chapter IV of this dissertation.

bed in order to ensure good and healthy sleep hygiene.[45] The thought is that, by ensuring our sleep is as healthy as possible, the result will be a net benefit for body and mind.[45]

I.ii.1 *Restorative processes of sleep, consequences of sleep disruption and deprivation.*

Homeostasis refers to the physiologically optimal 'balance' of the body's different functions — a 'default' setting of the body, in a sense.[8,46] In this view, a primary goal of sleep is to return the body to this default state, with homeostatic sleep pressure being one example of the internal signaling that occurs to promote this return to baseline.[36,37] Sleep deprivation may cause disruptions in inflammatory processes as well as alter release of growth hormones, hindering the healing process.[47] Similarly, sleep deprivation may lead to elevated levels of cortisol, a hormone commonly associated with stress, suggesting a dysregulating effect on the hypothalamic–pituitary–adrenal axis.[48] Additionally, sleep may be required to recover from physical fatigue, with animals showing changes in neural sleep patterns following exercise.[37] As these findings suggest, not only does sleep deprivation manifest as greater 'wear and tear' on the body, but — as to be expected — in the brain as well.

The role of sleep as recovery similarly manifests in the brain, as during sleep the brain uses cerebrospinal fluid to clean accumulated neurotoxic waste (such as proteins implicated in neurodegenerative diseases, e.g., β-amyloid), thereby acting as a neuroprotective process.[10,13] Sleep deprivation additionally affects socioemotional functioning, with sleep-deprived individuals having poorer emotional regulation, greater emotional reactivity, increased feelings of anxiety, stress, and depression, and resulting increases in interpersonal conflict.[8] Central to both of these examples is the impact of sleep deprivation on cognitive function.[37] Prolonged sleep deprivation leads to intensifying deficits in near global cognitive performance, including specific functions such as attention, decision making, verbal fluency, and executive function.[6] These declines in

cognitive function are likely mediated by changes in synaptic function caused by sleep loss, which alters synaptic plasticity and signaling.[49]

Of the many neurocognitive processes reliant on sleep, two cognitive functions most impacted by the effect of sleep deprivation on synaptic function are learning and memory[14,15,49] note 5. For instance, sleep is required for hippocampus-dependent memory consolidation, as this process entails taking neural representations of memories from temporary storage in the hippocampus and moving them to more stable storage.[14,15] This storage process, as well as the strengthening and integration of these new memories, may be delayed until the animal is able to sleep.[49] However, not all sleep is made equal in its influences of memory function. Better understanding the role of sleep in healthy functioning — especially for learning and memory — means understanding neural dynamics occurring during sleep.

I.ii.2 *Sleep stages and sleep architecture*.

After the initial division between behavioral sleep and wake, sleep in higher vertebrates itself can be broadly divided into two distinct stages: non-REM (NREM) sleep, characterized by neural oscillations with high power in low frequency bands (<4 Hz), thalamic sleep spindles (14-16 Hz), and sharp-wave ripples (100-300 Hz); and REM sleep note 6, characterized by neural oscillations resembling that of neural activity during wake, primarily in the theta band (4-8 Hz), and the presence of ponto-geniculo-occipital (PGO) waves.[15,50,51] The interactions between and roles of each of the two stages in memory consolidation are summarized by the *sequential and active system consolidation hypothesis*, which proposes differential roles of NREM and REM sleep

[5] For further discussion on circadian rhythm-related changes in synapses and plasticity, please read the forthcoming dissertation by, as of this writing, soon-to-be Dr. Alex Neitz.
[6] Named thusly due to the rapid eye movement activity that generally accompanies this phase as measured using electrooculography in humans.[50]

in long-term memory storage.[14] In this hypothesis, *sequential* refers to the tendency of REM sleep to follow NREM sleep, while *active system consolidation* refers to recently encoded memories being distributed throughout the memory-encoding neuronal network.[14]

According to this hypothesis, during NREM sleep, adaptive memories (i.e., information relevant to survival) are strengthened and moved from temporary storage in the hippocampus (a structure located in the temporal lobe of the cortex and implicated in learning and memory formation[52]) to more stable, long-term storage in the neocortex.[9,14,53] Central to this process is the generation of sleep spindles, short bursts of 14-16 Hz neural oscillations that occur during the memory consolidation phase and serve to induce phase-locking between the neocortex, thalamus, and hippocampus.[9,14,52] Optogenetic stimulation resembling spindles, synchronized with slow cortical oscillations, has been shown to promote adaptive memory consolidation, with optogenetic spindle inhibition impairing adaptive memory consolidation,[9] suggesting a role of sleep spindles in memory consolidation.[9,15,16] Additionally, hippocampal sharp-wave ripple activity present during NREM sleep may serve to promote long-term memory storage through its interactions with anterior cingulate cortex neurons, a critical area for long-term memory.[54] In some animals, NREM can be further sub-divided into three stages (where stage 3 is referred to as 'slow-wave sleep' or 'deep-sleep'), with each stage exhibiting relatively more or less of these neural features represented in NREM sleep.[15,50] Through these processes, NREM primarily serves to restructure memory representations in neocortical networks for future retrieval.

During the second stage, REM sleep, adaptive memories stored in the neocortex during NREM sleep are reactivated to increase stability and integration of these memories within previously existing memory networks.[9] The visual similarity of REM neural traces to those of wake is partly attributable to the high power present in the theta frequency band, with inputs from

the basal forebrain innervating the amygdala, the hippocampus, and driving theta activity in wake and REM sleep, but less so in NREM sleep.[51] In wake, these inputs aid in sensory processing and cortical plasticity (among other functions), while during REM sleep they may aid in reinforcing newly stored memories, particularly and possibly driving the vivid dreams that often accompany sleep.[15,51] Additionally, in rodents, PGO wave generation is similarly thought to be linked with memory consolidation in both the hippocampus and the amygdala.[15] These interactions with the amygdala – a neural structure primarily associated with emotion processing – underscore the potential role of REM in processing emotionally salient (i.e., adaptive) memories.[16] While not yet entirely understood, these findings suggest that REM primarily promotes further integration and strengthening of adaptive memories as well as processing particularly for memories with emotional content.[15,16]

Due to the sequential nature of these two stages of sleep, animals have developed a generally conserved temporal arrangement of sleep stages during a bout of sleep known as *sleep architecture*.[15,50] During sleep, healthy adult humans tend to cycle between sleep stages through the night, with NREM sleep being more dominant in earlier sleep cycles of the sleep bout and REM sleep being more dominant later cycles (Figure 2).[10,15,50] Sleep architecture is an important component of sleep hygiene, changing in response to increased homeostatic sleep pressure (e.g., after pulling an all-nighter to write a dissertation) or desynchronization of circadian rhythms (e.g., flying to a city three time zones away).[50] Disruptions in sleep architecture are a common component of many physical and mental disorders,[50] and can occur even without sleep deprivation (e.g., experience of trauma can alter timing of both NREM and REM stages).[55] When examining the various factors capable of disturbing sleep architecture, along with the underlying endogenous

rhythms and homeostatic pressure governing the timing and architecture of sleep, it becomes clear how the mechanisms behind these disruptions may vary depending on their source.

I.iii CAUSES OF SLEEP DISRUPTION AND DEPRIVATION.

Sleep is a remarkably plastic and sensitive process that can readily respond to a range of different influences. Because of its fundamental role in overall wellbeing, this adaptive nature is inclined to promote healthy outcomes but may also lead to both short- and long-term adverse consequences. For instance, sleeping in a new environment can lead to acute sleep disruptions (a phenomenon known as the *first night effect*, where the brain is temporarily more attentive to external stimuli accompanying the new context, metaphorically 'sleeping with one eye open'),[56] while other experiences may lead to longer-lasting, chronic sleep disruptions.[7] The primary factors governing this difference between acute vs. chronic sleep disruptions can be simplified into two categories: The first category is *internal factors*, including *genetic predispositions*, influences that lead an individual to be more- or less-likely to develop sleep disorders, and *comorbid disorders*, where sleep disruptions occur as a result of or in tandem with some other disorder. The second category is *external factors*, particularly the *physical and sociocultural environment*, which captures the ways in which one's physical, built and socioemotional environments interact to promote or prevent sleep. Importantly, these two categories are by no means separate; genetic predisposition often leads to long-term sleep problems,[57] while physical and sociocultural environments can play a role in the development of disorders that lead to sleep disruptions (e.g., depression, anxiety).[58,59] These factors are very closely intertwined as the intensity with which either influence sleep depends on interactions between the two. These interactions can be viewed in the way in which they lead to the emergence of long-standing physiological responses.

I.iii.1 *Internal factors: genetic predispositions and comorbidities.*

Given that genome is traditionally regarded as the 'blueprint of life'[60] [note 7], and, as previously described, genetics naturally play a role in sleep regulation and disruption in humans, it follows that genetic predisposition is a crucial factor for various sleep related disorders, shaping not only how and when we sleep, but also our vulnerability to sleep-related health conditions. For instance, family history of insomnia (sleeping too little) or hypersomnia (sleeping too much) can increase an individual's likelihood of developing a sleep disorder themself.[57] In monozygotic twins (i.e., those who share 100% of their genome), if one twin has restless leg syndrome (a sleep disorder marked by the urge to move one's body), the other twin has a higher likelihood of developing it as well.[25] Similarly, about 34% of variability in sleep quality may be attributed to genetics.[61] Certain circadian disorders have been linked to mutations in specific clock genes, altering expression of endogenous clocks.[34]

In mammals, activity of these clocks relies on the interaction of two core clock genes: *Clock* and *Bmal1* (with homologues found in other species).[34,57,62] Simplifying the pathways, *Clock* and *Bmal1* encode for two independent yet functionally related proteins; circadian locomotor output cycles kaput (CLOCK) and brain and muscle Arnt-like protein 1 (BMAL1).[34,62] Both proteins support proper circadian rhythmicity and are consequently expressed in virtually all cells exhibiting circadian oscillations.[62] These proteins heterodimerize to form a transcriptional activator complex that targets downstream core clock genes including *Period* genes (*Per1*, *Per2*, *Per3*), Cryptochrome genes (*Cry1*, *Cry2*), and *PAS* proteins. The protein products of these genes

[7] Note the phrasing 'traditionally regarded.' This conception, while capturing the key role the genome plays in just about all of life, simultaneously casts the genome in a misleadingly expansive role. Some argue it would be more accurate to view the genome as providing a list of materials required for the specific organism, with the actual building of the organism occurring with everything from the individual cell to the environment within which the organism grows.[60]

inhibit the transcriptional activity of the CLOCK:BMAL1 complex and in doing so inhibit their own transcription, ultimately forming a transcriptional feedback loop between the CLOCK:BMAL1 complex and its downstream targets.[3,28,57,62] As suggested in rodent models, *Bmal1* may be particularly crucial for proper circadian rhythm expression, as deletion of this gene results in circadian arrhythmicity and sleep fragmentation.[3,28,57] On the other hand, *Clock* may be crucial for maintenance of circadian rhythms, with mutations of this gene causing loss of rhythmicity in constant conditions.[3,28]

Despite not necessarily arising from mutations in clock genes, sleep disorders are a common comorbidity in many physiological and neuropsychiatric disorders.[63] For instance, individuals with chronic pain often experience insomnia, leading to reduced sleep quantity and quality and changes in sleep architecture.[63] Similarly, individuals with neuropsychiatric disorders such as depression or schizophrenia commonly experience sleep disorders (more often insomnia than hypersomnia).[63] However, the direction of this influence is unclear. Given the role sleep plays in cognitive and emotional wellbeing, lack of sleep due to insomnia may lead to greater likelihood of developing a psychiatric disorder,[35] with its treatment leading improved psychiatric symptoms.[63] Conversely, these psychiatric disorders may instead represent the primary condition, and treating them may in turn treat the comorbid symptoms of sleep disturbance.[25] Further complicating this is not only the genetic influences on the potential development of neuropsychiatric disorders, but also influences of external factors on sleep quality.

I.iii.2 *External factors: physical and sociocultural environment.*

Growing up, a frequent game in my family was listening to the sharp, quick, piercing noises from outside our house and try to discern, ‘was that a gunshot, or a firework?’ While sensitivity to incoming stimuli decreases during sleep,[25] humans still respond to some stimuli deemed relevant

in some way – to the potential detriment of their sleep quality. In terms of the physical environment, multiple levels can determine whether one consistently gets good or bad sleep, with factors such as greater light & air pollution and traffic & related noise contributing to poorer sleep.[64] Such factors can directly disturb one's sleep regulation (e.g., light pollution altering circadian rhythms through the light entrainment pathway) or indirectly impact sleep efficacy through increased environmental stress.[64] As such, many individuals exposed to these sources of sleep disruption often show symptoms of insufficient sleep, insomnia, or sleep apnea.[64] While access to nature has been shown to improve sleep quality, proximity, walkability, and psychosocial feelings such as unsafety may prevent immediate access of tools to increase sleep quality while simultaneously exacerbating existing barriers to good sleep.[59,64]

In some ways another layer of our physical environment, experience of one's sociocultural environment adds to the determination of whether one gets good or bad sleep. In this case, such factors include social cohesion, previously mentioned feelings of safety, or experience of identity-based prejudice and discrimination.[64] In her article, "Does Where We Sleep Matter?" Keisha Ray explores how all of these influences — physical and sociocultural — interact and amalgamate to affect sleep quality specifically in Black people.[59] As she states in the conclusion of the article,

> *"Black people are disproportionately affected by environmental pollution, have lesser access to proper housing than White people, and practice race related vigilance in their everyday lives. Black people have differing access to social determinants of health largely because of the ways their race, gender, and class make them vulnerable to abuse and oppression from the institutions meant to protect citizens and their health. But because of their Blackness, they are not afforded these protections."*[59]

What I hope is clear from this is the many factors influencing sleep — physical, sociocultural, and more — intermingling to cause disproportionately poor outcomes for those with

specific social identities[59] [note 8]. As suggested by Keisha Ray, consistent and prolonged exposure to stressful stimuli may elicit long-term physiological responses inside the individual. This interference can adversely affect their ability and capacity to regulate their sleep, in turn leading to poorer physical and mental health. However, neither these external factors nor the previously discussed internal factors operate independently to exert their effects on sleep quality. Rather, nearly all human characteristics result from the interactions between the two.

I.iii.3 *Interactions between internal and external factors.*

One of the most common interactions between internal and external influences that shape human characteristics is the *gene-environment interaction*. In this dynamic, an individual's genetic background determines their susceptibility to potential disorders, while the environments they encounter throughout their life determines whether or not these disorders actually manifest and, if so, to what degree.[65] This dynamic can be most clearly seen in post-traumatic stress disorder (PTSD), a condition an individual may develop in response to a single, highly traumatizing event, or in response to multiple or sustained traumas in response to which they felt they had no escape (otherwise known as *complex* PTSD).[55,66] Despite the majority of individuals experiencing some traumatic event in their life, those with a higher genetic risk (e.g., having a family member diagnosed with PTSD) are more likely to develop PTSD and corresponding sleep difficulties as a result of such an experience.[65] While adaptive memory formation is crucial for survival (allowing one to remember and learn from prior life-threatening experiences), PTSD may represent a seemingly pathological manifestation of the neural dynamics underlying adaptive memory and learning.[2,55]

[8] For further discussion of sleep and justice, please see Chapter IV.

While presence and severity of PTSD symptoms differ between individuals, a core symptom in many is sleep difficulties, being present in nearly 90% of cases.[55] Added to this is the core tenet of context; exposure to stimuli closely associated with the memory of the traumatic event causes individuals to 're-experience' the traumatic event as if it were happening in the present, leading to accompanying feelings of emotional and physical distress and worsening of PTSD symptoms.[55,66] Another common component in the development of PTSD is, through the fear learning process, 'neutral' stimuli may become indirectly associated with the traumatic memory (e.g., a billboard seen during a life-threatening car accident), allowing this stimuli alone to elicit feelings of distress.[55] As a result, fear-related stimuli (a billboard) that trigger memories of the traumatic event may be enough to elicit psychophysiological responses which, in turn, lead to increased disturbances in sleep architecture[55].

Crucial to the development and maintenance of PTSD is REM, the sleep stage primarily associated with processing and extinguishing fear-related memory.[55] Those with PTSD often experience REM fragmentation, waking up numerous times during REM sleep and thus disrupting emotional memory processing.[67] In fear-learning tasks, spending more time in REM sleep is linked to weaker fear learning and greater fear extinction.[55] Additionally, REM fragmentation following a traumatic experience is associated with the severity of PTSD symptoms six weeks later.[55] Finally, unlike NREM sleep, the timing of REM sleep is controlled by SCN circadian oscillators.[68] Taken together, this findings highlight the remarkably complex interactions between internal and external factors that influence the emergence and severity of PTSD [note 9]. To comprehensively explore the diverse neural mechanisms underlying sleep disturbances in PTSD – and in sleep more broadly – necessitates a variety of methods, each best suited to extract distinct types of information.

[9] For further discussion of sleep architecture and trauma, please see Chapter II, along with additional discussion in Chapter IV.

I.iv METHODS OF STUDYING SLEEP.

While there are arguably countless tools one can use to study the neural mechanisms regulating sleep, I will focus on three broad approaches, each with their own benefits and drawbacks: *human studies*, which, as the name implies, involve studying humans directly[69]; *animal studies*, which instead use different animal species — mice to monkeys to the *Drosophila melanogaster* — to model sleep disorders in humans or to better understand the conserved processes of sleep across species[70]; and artificial intelligence, a broad set of computational tools that allow us to effectively glean relationships and patterns in collected data.[71]

I.iv.1 *Human studies.*

One of the most important inventions in the history of human sleep research is polysomnography (PSG); where once we only saw sleep as a temporary loss of our awareness, we later learned to recognize the stereotypic patterns of activity that characterized each stage of sleep that occurs during sleep, as indicated by the uneven, often overlapping lines mechanically captured on steadily rotating paper.[72,73] Then, as now, sleep scoring was often a manual task; trained polysomnographers visually inspected each equally spaced portion of time (or epoch) in the data and deciding what sleep stage — REM? NREM? Wake? Unremarkable noise? — the lines told them about the patient's sleep.[72,73] Because of the unprecedented insight into the sleeping mind afforded by this technology, electroencephalography (EEG; capturing neural activity from the scalp) was joined by electromyography (EMG; capturing muscle activity) and electrooculography (ECG; capturing eye activity) to become the gold standard procedure for studying sleep architecture in both clinical and experimental contexts.[69,72,73] A disadvantage of PSG, however, is how delicate of a process it is; multiple sensors must be placed in very specific spots on the body, leaving a mess of wires and potentially incomplete circuits. Common caps used to capture EEG

may not consistently connect to hair textures such as afro-textured hair, potentially decreasing quality of data acquired from individuals with such hair types and posing challenges for researchers[74]. While PSG may be the gold standard for studying sleep architecture, there are other, less constraining methods of studying sleep in humans.

One such method is actigraphy — a non-invasive method of capturing activity.[75] In this case, rather than requiring multiple wires, simple rest/activity cycles can be captured with a device resembling a wristwatch.[75] Compared to PSG, this method is capable of capturing longer sessions of data, making longitudinal analyses more straightforward, while this same simplicity prevents the discrimination of individual sleep stages from the data.[75] Additionally, both EEG/EMG and motion-based detection of sleep/wake activity can be recapitulated in animal models. In the case of motion-based detection, use of infrared motion sensors, for instance, can similarly capture animal rest/activity cycles while remaining relatively noninvasive.[76]

I.iv.2 *Animal models.*

Because the function and importance of sleep is generally conserved across species, sleep researchers have utilized different species of animals to glean different pieces of information about sleep in animals broadly.[70] One of the most frequently used mammals for modeling human disease is the mouse, specifically the C57BL/6 strain.[77] There are multiple benefits of using these mice as models of sleep disorders. One of the most prominent uses of animal models in general is their availability for use in a variety of experimental conditions, where some aspect of their physiology can be directly altered to examine the effects said alteration has on a phenotype of interest (compared to humans, in whom attempting this would cross multiple ethical boundaries). Additionally, this strain of mice was one of the first animals in which gene expression could be

manipulated[78] and to have its genome sequenced.[79] This, coupled with the copious research articles using it, has made the C57BL/6 strain one of the more widely used in biomedical research.[77,79]

Due to these advantages, mice can be used to model PTSD, allowing researchers to study neural underpinnings of the disorder in more controlled settings.[80] Different aversive stimuli (e.g., foot shocks, predator scent) can be utilized to induce PTSD-like phenotypes in differing degrees.[80] In context of these models, the PTSD phenotypes are divided into behavior phenotypes (e.g., fear learning, avoidance of trauma-related stimuli, and depression) and biological phenotypes (e.g., inflammation, sleep disturbances, changes in stress hormones).[80] While a benefit of this method is being able to more rigorously and ethically evaluate, for instance, the extent to which PTSD affects sleep and circadian rhythms, there are certain characteristics that may not be available for recapitulation in mouse models. This leads to the question: how do we translate findings from animal models to humans? Although difficult to answer, data gathered from animal models provide valuable insights – insight which can be further expanded upon through the use of computational methods for capturing complex patterns.

I.iv.3 *Artificial intelligence.*

Computational methods using artificial intelligence (AI) are increasingly becoming a staple not only for research in the neuroscience of sleep, but neuroscience research in general.[71,81] While not a 'method of studying sleep' in the classical sense (as captured by previous sub-sections), AI instead serves as a powerful tool to aid in capturing, analyzing, and interpreting sleep-related data.[71] Given its increasingly commonplace status, we can examine how AI can be used to supplement and bolster traditional sleep research methods.[71] Specifically, I will be focusing on machine learning — a type of AI that focuses primarily on pattern recognition in data.[71]

As summarized by Usman et al. (2019),[82] the pipeline of machine learning can be separated into five generally consecutive steps: data collection, pre-processing, feature extraction, classification, and post-processing. The following are key steps involved in using machine learning for sleep stage classification based on neural signals:

- *Data collection* — This represents all steps taken to collect neural signals. In the case of humans, this is often done using EEG,[71] but in certain cases electrocorticography (ECoG) electrodes may be surgically implanted directly on the cortex or deeper inside the brain.[71,82] In the case of mice, implanted ECoG and EMG represent the gold standard sleep data acquisition method.[83]

- *Pre-processing* — This step involves removing noise from the data, as well as any pre-amplification processes including frequency filtering via cutoff or bandpass filters.[82]

- *Feature extraction* — After pre-processing, the signal is run through an algorithm to extract different variables (or features), in both time- and frequency-domains, that are relevant to identifying sleep stages.[82] For instance, sleep spindles have a frequency of about 14-16 Hz[53]; extracting power specifically in this frequency band, in order to capture these NREM-specific oscillations, may aid the algorithm in differentiating sleep stages.

- *Classification* — In this step, all features extracted from the previous step are then passed into a previously trained machine learning algorithm — typically trained with human-scored sleep stages — to find the most statistically likely relationships between the features of interest and sleep stages.[82]

- *Post-processing* — Finally, this step involves validating results from the classification stage, which may involve using cross-validation methods to determine accuracy of the model used or comparing how the model performs relative to chance.[82]

While each step entails greater details than listed here, I hope that this simplified machine learning procedure may serve as a relatively clear example for the power and flexibility of AI-based methods in analyzing sleep [note 10].

I.v CONCLUSIONS AND THESIS OUTLINE.

In the preceding sections, I aimed to underscore the pervasive influence of sleep on various aspects of human life, encompassing physical, mental, and psychosocial well-being. Additionally, these facets are intricately intertwined, indicating a bidirectional relationship wherein sleep quality significantly affects these areas and, in turn, is influenced by them. As explored in section 1, the sleep-wake cycle is heavily regulated by internal circadian rhythms, heavily determining not only sleep timing, but also, consequently, other endogenous rhythms important for overall health. In section 2, I further examined interactions between circadian rhythms, sleep, and health, emphasizing the role of sleep for learning and memory as well as the potentially debilitating effects of sleep deprivation on physical and mental function. In section 3, considered some example causes of sleep deprivation, such as relatively internal factors (e.g., genetics), relatively external factors (e.g., physical environment), and interactions between the two. Finally, in section 4, I gave a brief review of methods commonly used in sleep research, including human studies, animal models, and computational methods. The focus of this introduction was to give a brief introduction to sleep research, setting the stage for the following chapters in this thesis by emphasizing concepts that will be more fully explored.

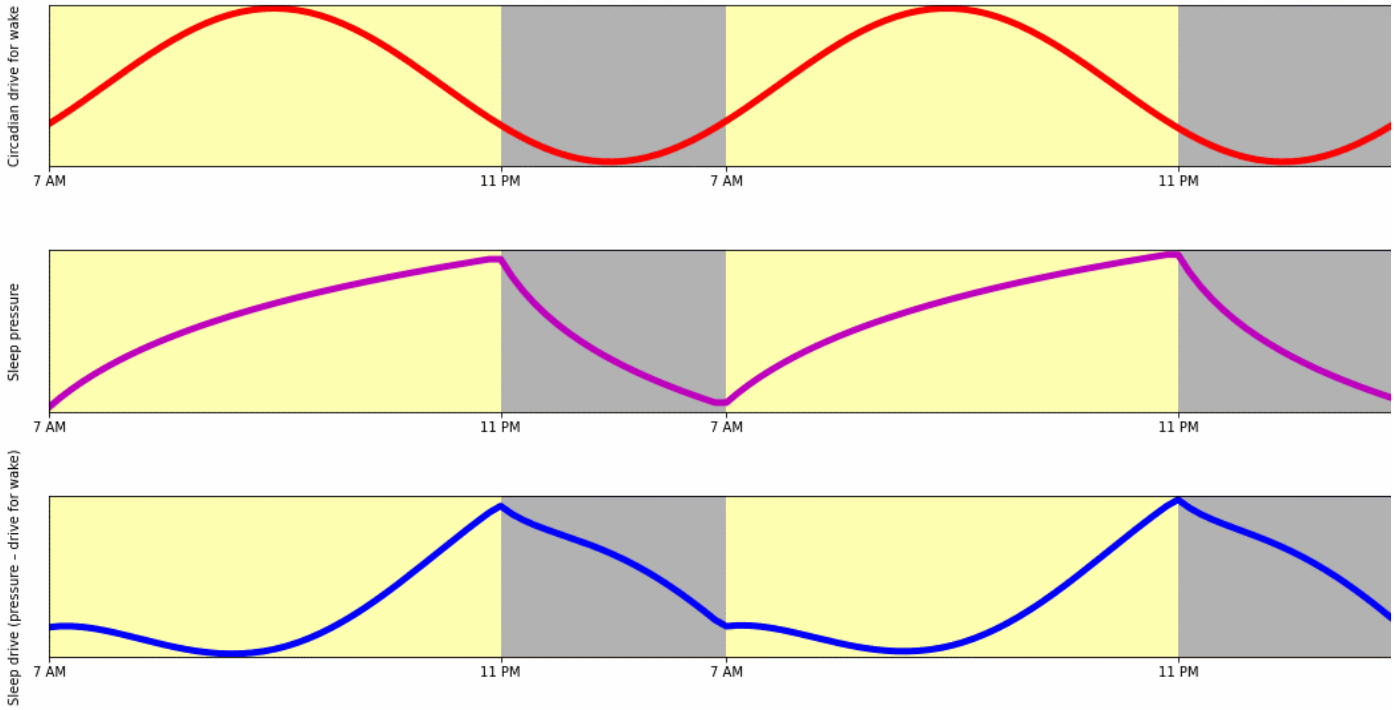
One overarching theme of this thesis is the remarkably multifaceted and adaptable nature of sleep. Among the physiological needs typically considered essential for healthy function (i.e.,

[10] For further discussion of sleep and AI, please see Chapter III, as well as additional discussion in Chapter IV.

air, food, water, shelter, clothing, warmth, sex, and sleep), I am specifically focusing on sleep due to its extensive influence on various health conditions and its susceptibility to external influences. Because of this expansive of role sleep in overall wellbeing, I believe that truly capturing this breadth of influence requires approaching the concept of sleep using ideas and methods from different disciplines. As demonstrated in the upcoming chapters, I will explore this topic through the lenses of behavior & learning, exemplified in Chapter II's exploration of contextual fear and entrainment of circadian rhythms; computational methods, exemplified by Chapter III's use of machine learning for studying sleep architecture; and neuroethics, exemplified by Chapter IV's scrutiny of sleep research within the broader societal context. Through this interdisciplinary lens, this thesis aims to contribute to a more holistic perspective on sleep, bridging gaps between different fields of study and, ideally, contributing to efforts towards a more inclusive and equitable conception of sleep's adaptivity and complexity.

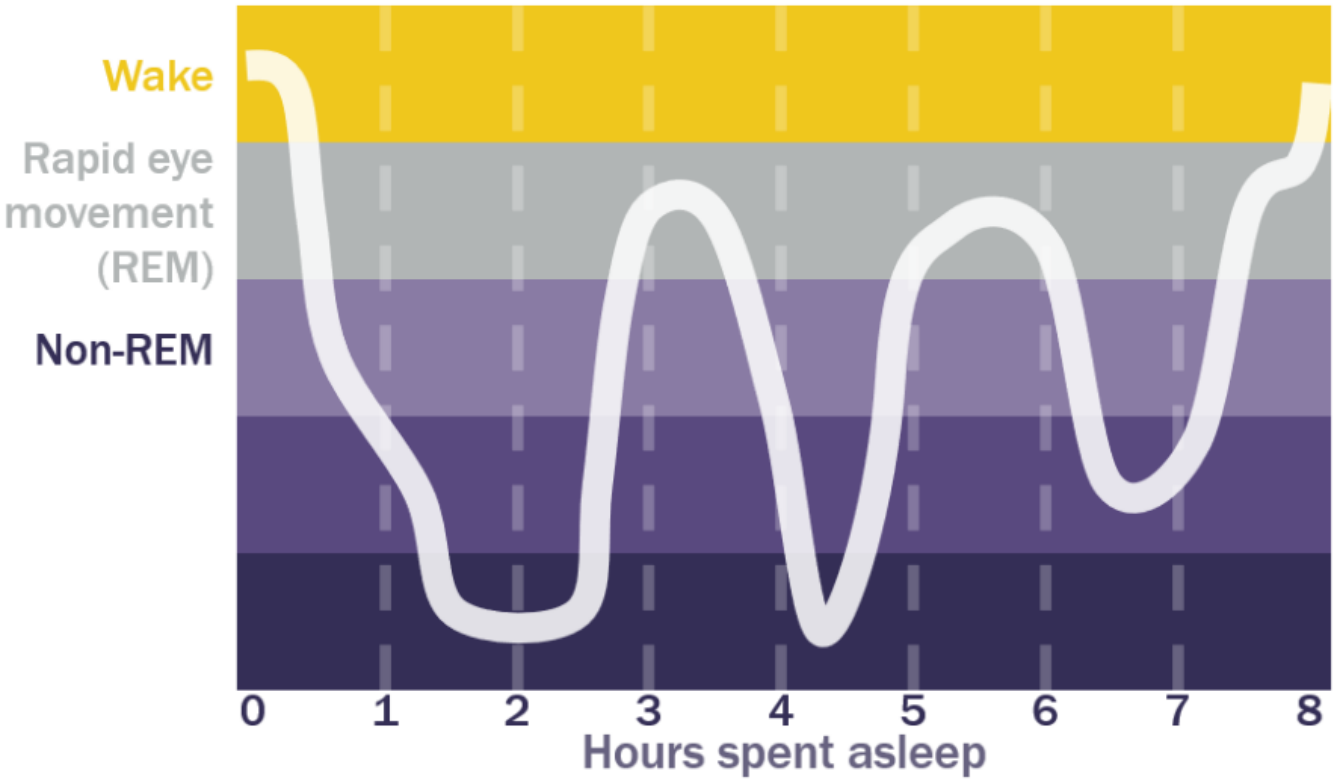
I.vi FIGURES.

Figure I-i. Two-process model of sleep.



Note. An example of circadian rhythm drive for wake (red, top), homeostatic pressure for sleep (purple, middle), and sleep drive (blue, bottom) over two days. Sleep drive represents the combination circadian drive for wake and homeostatic pressure for sleep, reaching its peak near onset of sleep (grey box, 11 pm-7am) and lowest shortly after wake (yellow box, 7am-11pm).[4]

Figure I-ii. Representation of typical sleep architecture during a eight-hour sleep bout.



Note. Shortly upon sleep, animals tend to enter non-REM sleep (representing the stage with the greatest proportion of time spent) and thereafter intermittently enter REM sleep through the night. In humans, non-REM sleep is further divided into three, increasingly deeper stages of sleep, with this sleep getting lighter as the night progresses.

Chapter II. SLEEP WITH ONE EYE OPEN: CONTEXTUAL FEAR AND ABNORMAL CIRCADIAN RHYTHMS

Fear learning is an evolutionarily conserved survival mechanism in which adverse stimuli become mentally associated with a specific context, leading to feelings of anxiety and fear upon re-exposure to fear-related contexts even in the absence of the adverse stimuli.[55] This process, known as fear conditioning, is a form of fear learning in which a neutral environmental cue is repetitively paired with an aversive stimulus, leading to a conditioned fear response when the neutral cue is encountered again.[55] In typical functioning, fear learning serves as a vital tool for animals to distinguish between safe and unsafe cues, playing a crucial role in survival.[2,55] However, if the emotional intensity of the aversive conditioning stimulus is strong enough, this process can lead to maladaptive associations where additional neutral cues unrelated to the original aversive stimulus may provoke heightened fear responses.[55]

This phenomenon is what leads to trauma and is the foundation for post-traumatic stress disorder (PTSD). PTSD is a chronic psychiatric disorder characterized by exposure to a traumatic event (e.g., military combat, physical/sexual assault) and subsequent re-experiencing of the traumatic memory, with contextual cues typically acting as triggers for traumatic memories.[55,84], The emotional memory and learning model of PTSD proposes that the disorder is a product of dysfunction in processes of fear conditioning, extinction learning (the process of learning that a previously fearful cue has ceased to be dangerous), and extinction recall (the ability to remember this learned extinction).[55]

One of the core symptoms of PTSD is sleep disruptions,[85] with individuals most frequently reporting insomnia (80-90%) and nightmares (50-70%).[86] These sleep disruptions are typically associated with circadian disruptions, and may worsen when the individual is re-exposed to

trauma-related triggers, acutely causing later sleep onset and poorer sleep quality.[55] Conversely, poor sleep quality leads to poorer emotional processing and coping in individuals with PTSD,[55] pointing to a bidirectional relationship between PTSD with sleep and circadian disturbances.[85,87] Thus, studying the relationship between sleep regulation, circadian rhythms, and contextual fear conditioning will provide key insight into the neural processes that underly core symptoms of PTSD.

Our laboratory has recently shown that nocturnal rodents can entrain their circadian rhythms to a non-photic stimulus of cyclic fear.[32,88] When rats or mice need to leave a safe nest into a foraging area to obtain all their food and water, they spontaneously do so mainly during the dark phase of the light-dark (LD) cycle. If they experience cyclic fear (footshocks) during the light phase (light fear, LF) their foraging and feeding activity becomes even more nocturnal.[32,88] However, if they experience cyclic fear during the dark phase (dark fear, DF), they shift their foraging and feeding activity predominantly to the light phase.[32,88] Interestingly, this shift into diurnal behavior does not represent an acute response to shocks or a result of associative learning between the LD cycle and the fear stimulus. Instead, it is the result of the entrainment of the circadian rhythms of foraging and feeding to the cyclic fear stimulus, as the timing of feeding and foraging persists for several days after both the LD and the fear cycle are removed.

Our fear entrainment results suggest that cyclic fear stimuli have a strong circadian time stamp, and that circadian and sleep disruptions, which are signature symptoms of PTSD, may in part represent the output of a previously fear-entrained clock. If this is the case, contextual cues that act as triggers for fear and anxiety in PTSD could be simply retrieving a fear-entrained circadian program of behavior. In this chapter, I tested this hypothesis by exposing mice to both the LF and DF groups. After stable entrainment of both groups to the specific cyclic fear stimulus,

we placed mice from both groups in a neutral context where they all eventually displayed normal nocturnal behavior. We then tested whether re-exposure to the fear context but in the absence of aversive stimuli was enough to retrieve aberrant circadian rhythms that represented the output of the previously entrained fear-entrained clock.

II.i METHODS.

All experiments with animals were performed in accordance with animal protocols approved by the Office of Animal Welfare at the University of Washington. Three cohorts of 16 C57BL/6J mice aged 2-3 months, totaling 46 mice (50% female; two mice were removed due to poor acclimation), were purchased from The Jackson Laboratory. Mice were individually housed under a 12:12 LD cycle in either a 'neutral context' or a 'fear context'. The 'neutral context' consisted of a standard mouse cage equipped with a running wheel. The 'fear context' was a fear conditioning chamber featuring a 'nesting area' with corncob bedding connected to a 'foraging area' wherein mice had ad libitum access to food and water but were required to stand on footshock grids. Infrared motion detectors tracked activity within the nesting and foraging areas (termed nesting and foraging activity, respectively), while feeding activity was recorded using break-beam sensors. All data were continuously recorded. For additional information, please see please see Bussi et al., (2023).[88]

II.i.1 *Contextual fear paradigm.*

The contextual fear protocol comprised five phases (refer to the first column in Figure 1): First, mice's pre-fear induction activity was observed in the neutral context ('pre-induction' condition; ~7 days) to establish typical behavior patterns unrelated to aversive stimuli. Subsequently, mice were transitioned to the fear context to acclimate without receiving any

aversive stimulus ('baseline' condition; ~7 days). Afterwards, an Arduino-controlled shocker administered three randomly distributed footshocks per hour in 12-hour daily windows, based on shock group conditions ('fear induction' condition; 10-15 days). Following fear entrainment, mice returned to the neutral context to extinguish fear-entrained activity ('post-induction' condition; 14-19 days). After mice reverted to pre-fear induction activity patterns, they were moved back to the fear context, this time without footshocks ('contextual recall' condition). To study fear entrainment extinction, cohorts two (n=16) and three (n=14) remained in the fear context for up to a month.

Mice were divided equally across sexes into two shock groups: the light fear (LF) group, the 12-h window of shock presentation was paired to the 12-h light phase of the LD cycle; or the dark fear (DF) group, where footshocks were instead paired with the 12-h dark phase. Additionally, to examine effects of stimulus intensity and strength of entrainment, footshock amplitude AUCS was set at 0.2 mA for cohorts one and two, and set to 1.0 mA for cohort three.

II.i.2 *Data analysis.*

All analyses were run in R version 4.3.1,[89] and raw activity counts were z-score normalized by data channel with zero minimum prior to use in analyses. Statistical analyses conducted were either Analysis of Variance (ANOVA) using aov()[90] or Multiple Linear Regression using lm(),[91] and reported statistics used 95% confidence intervals (95% CI). Figures were generated using ggplot2.[92]

Average daily activity patterns ('waveforms') were generated for each phase, utilizing data either from the last seven days of recording (pre-induction, fear-induction, and post-induction) or first seven days (baseline and contextual recall). These data were used to create wind rose plots, representing averaged distribution of activity across different time points, and Rayleigh vectors, representing average circadian activity timing in terms of timing (in ZT) and magnitude (indicated

by |v|).[93] Rayleigh vectors provide a quantification of how fear stimuli impact mice's average circadian activity timing and strength. For instance, if an animal were only active at ZT1, then the average circadian timing would be ZT1 and |v| would be a relatively large value; if the animal were equally active across the dark phase and not active during the light phase, average circadian timing would be ZT18 and |v| would be a relatively small value.

To assess activity differences during critical time windows, trapezoidal sums[94] were used to calculate area under the curve (AUC). Two-way ANOVAs were employed to identify main effects of experimental condition (baseline, fear induction, contextual recall) and either shock group (LF vs. DF) or sex (male vs. female), as well as any interaction effects. In cases with significant interaction terms, post-hoc Tukey tests[95] were conducted to determine specific group differences. To account for multiple comparisons, significance levels were adjusted to $\alpha = 0.0056$ using the Bonferroni method.[96]

To explore relationships between strength of entrainment during fear induction and activity levels during contextual recall (both relative to the baseline activity), AUCs were calculated for each experimental condition within the specified time window. For use in analyses, AUCs equal to 0 were inputted to be 0.01. Multiple linear regression analyses were performed to determine beta estimates for activity during fear induction (relative to baseline activity), sex (dummy-coded, male vs. female), and footshock amplitude (dummy-coded, 1.0 vs. 0.2) in predicting activity during contextual recall (relative to baseline activity). For use in analyses, AUCs that equaled 0 were adjusted to 0.01.

Finally, to assess extinction progression, the month-long extinction phase was divided into four weeks, each compared to baseline and fear induction activity using AUCs extracted for critical time windows. One-Way ANOVAs were performed using condition (baseline, fear induction,

week 1, week 2, week 3, and week 4) to predict AUC. To address the 15 comparisons across conditions, a Bonferroni correction was applied, setting the adjusted α level at 0.0033.

II.ii RESULTS.

II.ii.1 *Following entrainment to nocturnal fear, re-exposure to fear context is enough to elicit return of abnormal circadian behavior.*

During the pre-induction and baseline conditions, wheel-running activity patterns in the neutral context (Figure 1, first row) as well as foraging activity and feeding in the fear context (Figure 1, second row) were undistinguishable between mice assigned to the LF and DF groups. Following fear induction, LF animals showed near-complete abolishment of diurnal activity, while DF animals showed greatly suppressed nocturnal activity and greatly increased diurnal activity, particularly during the first few hours of the light phase; which was confirmed both by the 24-h waveforms and wind rose plots for each group (Figure 1, middle row). When returned to the neutral context, both groups returned to similar wheel running and IR nocturnal activity patterns, comparable to their pre-induction patterns (Figure 1, second-to-last row). Upon return to the fear context but in the absence of footshocks (contextual recall), DF mice once more exhibited increased diurnal activity compared to LF mice, primarily in the first few hours of the light phase (Figure 1, last row).

We examined the first hour of the light phase (ZT0-ZT1), the time of the day at which activity was much higher in DF mice than LF mice during fear induction, finding a significant interaction between shock group and condition (F[2, 132]=53.68, p<0.001). LF mice exhibited no significant changes in activity in the fear cage across the three conditions but displayed a trending significant decrease in contextual recall activity compared to baseline (95% CI [-41.77, 1.44], p=0.082; Figure 2A). In contrast, DF mice exhibited significantly increased activity following fear

induction compared to both other conditions (baseline, 95% CI [71.43, 112.80], $p<0.001$; contextual recall, 95% CI [36.78, 78.15], $p<0.001$) and to LF during fear induction (95% CI [79.11, 121.42], $p<0.001$; Figure 2A). Additionally, while significantly decreased compared to fear induction, DF activity during contextual recall was significantly higher compared to baseline (95% CI [13.96, 55.33], $p<0.001$) and to LF contextual recall (95% CI [26.75, 69.06], $p<0.001$; Figure 2A). These results indicate that the re-exposure to the fear context induced a 24-h pattern of activity that resembled that under each group's specific fear exposure pattern.

We also examined the last hour of the dark phase (ZT23-ZT24), when the activity of LF was higher than that of DF mice during fear induction. There was a significant interaction between shock group and condition ($F[2, 132]=5.24$, $p=0.006$); LF exhibited significantly increased activity following fear induction compared to baseline (95% CI [10.80, 67.94], $p=0.002$) and to DF fear induction activity (95% CI [8.62, 64.55], $p=0.003$; Figure 2B). DF mice exhibited no significant changes in activity at this time of the day. Together, these results show that time-specific exposure to an aversive stimulus entrained the circadian rhythm of foraging in the LF and DF groups, and that the re-exposure to the fear context induced 24-h patterns of activity that were specific to each group.

II.ii.2 *Footshock intensity determines speed of entrainment to fear stimuli.*

A key condition that defines circadian entrainment is a stable phase relationship between the zeitgeber and the entrained circadian rhythm. Before circadian rhythms reach a stable phase of entrainment they go through transient phases that reflect the progressive adjustment of the circadian pacemaker phase. These "transients" determine the speed of entrainment of a circadian clock and the rhythms it sustains, and typically depend on the strength of the zeitgeber. That is, higher amplitude environmental cycles will typically lead to faster entrainment. If fear entrainment

shows similar properties, we reasoned that higher intensity footshocks (1 mA) would achieve faster entrainment than lower intensity footshocks (0.2 mA).

The LF and DF groups reach a final phase relationship with the 12-h fear window that is characterized by an onset of activity shortly after shocks offset in LF mice and 90 minutes prior to shock offset in DF mice, both relative to the offset of shocks for the respective groups. When measuring the number of days needed to reach this phase over the two-week period, both LF and DF mice that experienced 1.0 mA footshocks entrained more quickly and strongly compared to mice experiencing 0.2 mA footshocks (Figure 3A & 3B, right column). Furthermore, as indicated by Rayleigh vectors in day 19, final average circadian activity timing was centered around ZT17 for 0.2 mA (|v| = 1.19) and ZT16 for 1.0 mA (|v| = 1.56) in LF mice (Figure 3A, right column); in DF mice, the average circadian activity timing was was centered at approximately ZT22 for 0.2 mA (|v| = 0.48) and ZT3 for 1.0 mA (|v| = 0.55; Figure 3B, right column).

II.ii.3 *Magnitude of change in activity during contextual recall depends on strength of entrainment for DF but not LF mice.*

We reasoned that the induction of ZT0-ZT1 activity during the contextual recall could be associated with the strength with which the fear stimulus induced activity at the same time during the fear induction phase of the experiment, both relative to baseline activity. Hence, we used multiple regression predicting activity during contextual recall from activity during fear induction, footshock intensity, and sex. This model was significant in LF mice (F[3, 18]=4.71, p=0.013, adjusted R^2 =0.346); activity during contextual recall showed a positive association with activity during fear induction (b=0.167, 95% CI [0.20, 0.69], p=0.029) and negative association with footshock intensity (b=-0.360, 95% CI [-0.70, -0.024], p=0.037; Figure 4A). This model was also significant in DF mice (F[3, 20]=52.59, p<0.001, adjusted R^2 =0.871), with activity during contextual recall showing a positive association with activity during fear induction (b=0.733, 95%

CI [0.593, 0.875], $p<0.001$), and male mice showing a stronger response during contextual recall compared to female mice ($b=2.662$, 95% CI [0.16, 5.18], $p=0.039$; Figure 4B). Together, these results show that mice that displayed a stronger response (reducing and increasing early morning activity in LF and DF mice, respectively) during fear induction had a stronger response during the contextual recall.

II.ii.4 *Activity patterns during contextual recall do not return to baseline after a month.*

Contextual fear conditioning is characterized both by the induction of fear responses by cotextual cues and by the extincton of contextual fear responses in the absence of the aversive stimulus. Our results show that the contextual fear recall retrieves a previously fear-entrained circadian program and we wondered if this extintion would also progressively erase the 24-h pattern of locomotor activity that emerged in LF and DF animals.

In LF mice, no significant effect of condition was observed for low intensity during ZT0-ZT1 ($F[5, 42]=1.88$, $p=0.118$), and although an effect was found for high intensity ($F[5, 30]=4.914$, $p=0.002$), no significant differences in paired comparisons were detected after applying Bonferroni correction for multiple comparisons (for 15 comparisons, $\alpha = 0.0033$; Figure 5A & B, left bar graphs). Conversely, during ZT23-ZT24, there was a significant effect of condition for both low intensity ($F[5, 42]=10.20$, $p<0.001$) and high intensity ($F[5, 30]=8.18$, $p<0.001$). For both shock intensities, activity in the last hour of the dark phase was higher than baseline activity during weeks three and four, with this difference becoming more pronounced over the four-week period (especially in the case of high intensity; Figure 5A & B, right bar graphs).

The opposite pattern was present in DF mice: An effect of condition was found for both low intensity ($F[5, 42]=6.65$, $p<0.001$) and high intensity ($F[5, 42]=9.77$, $p<0.001$) during ZT0-ZT1. At low intensity, activity patterns across the four weeks resembled baseline activity, while at

high intensity, activity patterns in each week generally showed differences with activity in both baseline and fear induction (but were no longer significant after correction for multiple comparisons; Figure 5C & D, left bar graphs). During ZT 23-24, there were no differences at low intensity (F[5, 42]=0.85, p=0.524), and while there was an effect of condition for high intensity (F[5, 42]=4.45, p=0.002), no significant differences remained after correction for multiple comparisons (Figure 5C & D, right bar graphs).

II.ii.5 *Males and females display comparable responses during contextual recall.*

Because male and female mice differentially respond to fear conditioning (where male mice display stronger initial responses and female mice show slower extinction over time[97]), we investigated the presence of sex-dependent change in ZT0-ZT1 activity across the three conditions. There was no significant interaction between sex and condition in LF (F(2, 60)=0.79, p=0.457) or DF mice (F(2, 66)=1.89, p=0.159; Figure 6).

II.iii DISCUSSION.

In this study, we show that the re-exposure to a fear context in the absence of aversive stimulus has the ability of reactivating a previously existing circadian program of behavior that was the result of cyclic fear entrainment. Mistimed sleep-wake patterns and disrupted circadian rhythms are signature features of PTSD.[87] That nocturnal fear can entrain circadian behavioral outputs and lead to abnormal diurnal behavior in mice after exposure to contextual triggers suggests that these symptoms could represent the output of a circadian oscillator that was previously entrained to aversive stimuli with a clear 24-hour structure. Furthermore, our current results show that contextual triggers in animals that have regained normal circadian rhythmicity in a neutral context have the ability of rekindling this abnormal circadian program of behavior. Our results may be important to further understand the mechanisms behind core symptoms of PTSD,

as they suggest that circadian time-stamped traumatic memories may be central to these symptoms and to how these symptoms are reactivated by contextual triggers.

Contextual fear learning is characterized by the association of neutral contextual cues (conditioned stimulus, CS) with an aversive stimulus (unconditioned stimulus, US),[55] such as footshocks in laboratory rodents or traumatic events (e.g., military combat, physical/sexual assault) in humans. Evidence of this learning is the presence of behavioral and physiological responses (i.e., fear responses) to the CS alone that would normally be elicited only by the US.[55] In our behavioral paradigm, the CS is represented by our fear context in which animals live in more naturalistic conditions where they can experience threats while foraging but not in their safe nesting area. In this setup, the response to the US is changes in 24-hour behavioral patterns that involves the entrainment of circadian oscillator by cyclic fear. The fact that the same pattern of behavior is elicited several weeks later by the CS only indicates that the fear-entrained oscillator is downstream of the contextual fear association. In other words, the CS response is constituted by the activation of a previously entrained fear-entrained oscillator.

The LF and DF groups were returned to the fear cages in the absence of footshocks after they had become undistinguishably nocturnal in the neutral context. The most remarkable result in our study is that the DF group displayed a considerable amount of foraging and feeding during the light phase that can only be accounted for by their history of nocturnal fear exposure. Nevertheless, it is important to highlight that LF mice displayed a clear decrease in daytime activity when they were re-exposed to the fear context. This result indicates that 24-hour cyclic fear has the ability of changing the temporal pattern of behavior after the aversive stimulus is no longer present even if the fear was perceived at a time when mice are typically inactive. This suggests that cyclic 24-hour fear can entrain circadian clocks regardless of the timing of the aversive

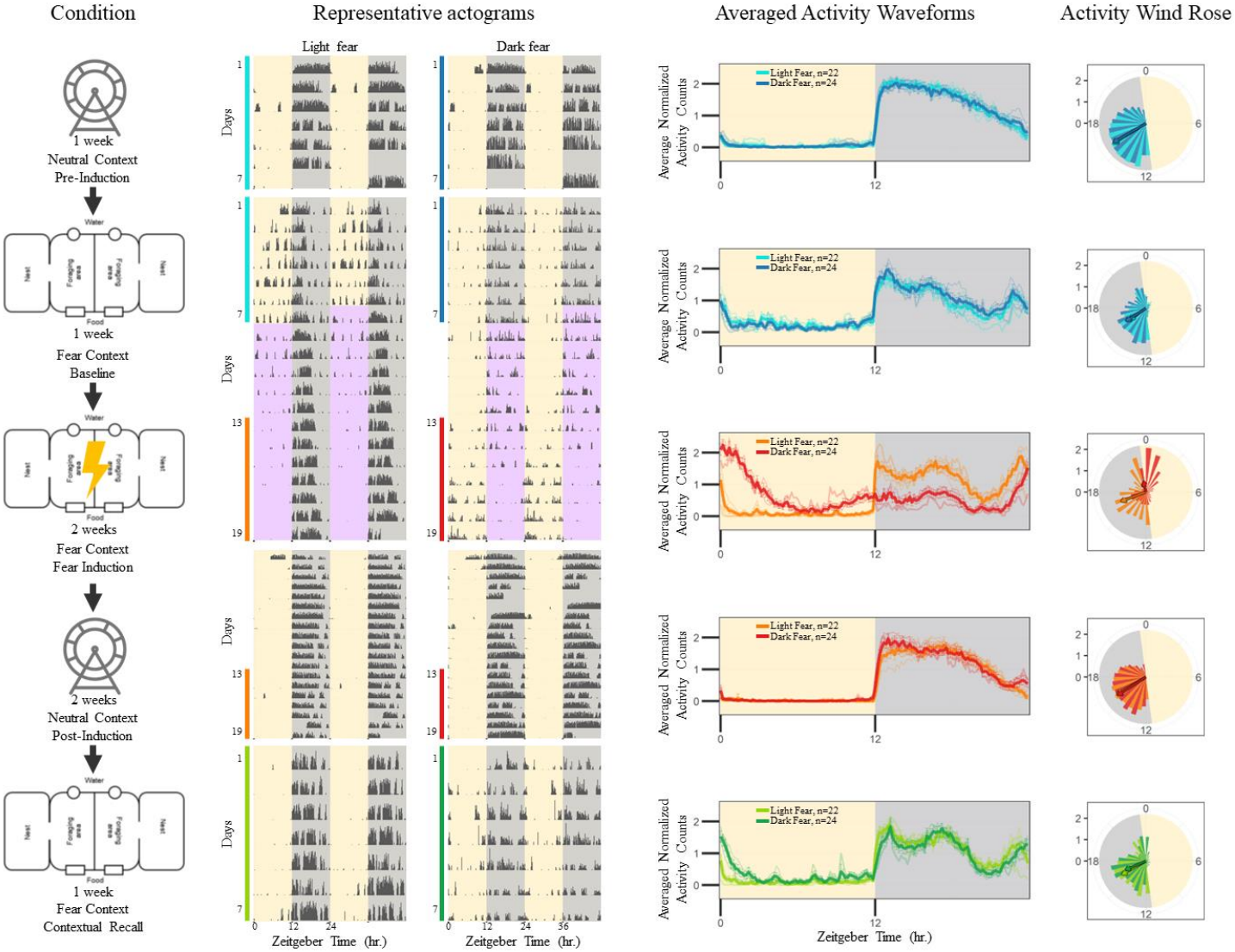
stimulus relative to spontaneous pattern of activity of the species. The result also suggests that temporal programs for behavior that are specific to the previous timing of fear can be retrieved by exposure to the fear context.

Compared to males, female mice have been previously shown to exhibit relatively lower fear responses (e.g., lower levels of freezing behavior) to threatful stimuli, as well as lower fear recall.[97] These findings emphasize the existence of sex differences in fear processing and learning. On the other hand, compared to men, women are more likely to be diagnosed with fear and anxiety disorders including as PTSD.[98] Importantly, although the diagnosis rate of these disorders is higher for women, the strength with which symptoms manifest in each disorder is not necessarily higher in women.[98] Despite these clear differences, the cause of sex-based differences in PTSD is not yet fully understood.[98] Critical sex-related factors that could interact to shape the development and symptomatology of PTSD include: biological factors (e.g. genetics, hormones), sex-specific vulnerability factors (e.g. sex differences in negative affect and anxiety traits), and sociocultural differences that affect gender-specific exposure to stress and trauma, as well as response expectations.[98] In PTSD, these factors could influence the strength of the traumatic experience and of the fear salience the experience has, as well all the steps in the contextual fear memory consolidation, recall and extinction.

In conclusion, we show that the abnormal diurnal behavior, which is the result of fear entrainment of a circadian clock after exposure of mice to nocturnal fear, reverts to normal nocturnal behavior when the animals are returned a neutral, fear-free context. We also show that the abnormal circadian behavior re-emerges several weeks after if animals are exposed to the initial fear context. These findings suggest that the worsening of PTSD sleep dysregulation in response to triggers may emerge as the output of a previously entrained fear-entrained oscillator.

II.iv FIGURES.

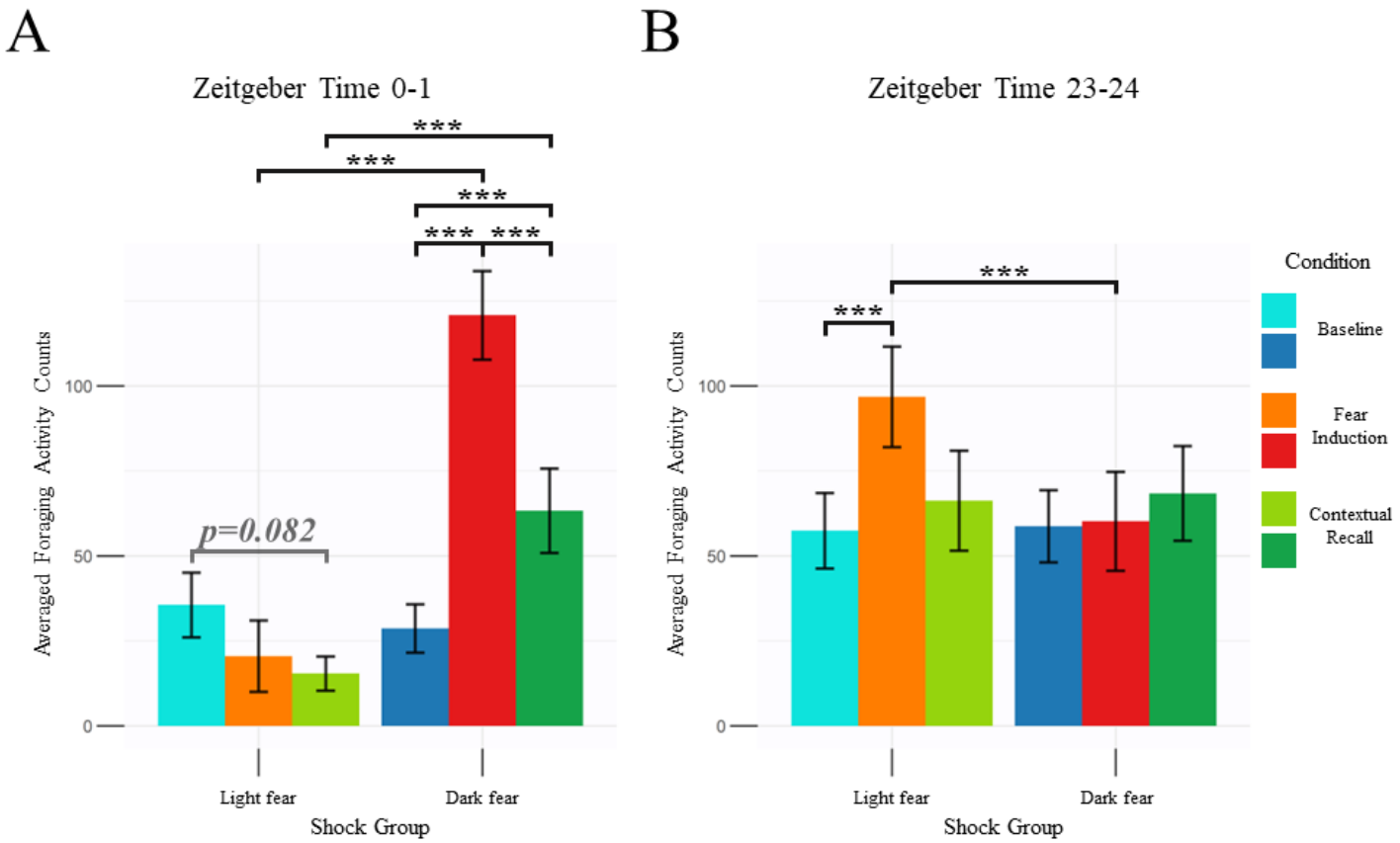
Figure II-i. Experimental conditions and corresponding activity patterns.



Note. Each row of data represents the corresponding experimental condition shown in the left-most column. Representative actograms show data acquired from one mouse for each shock group (light fear, left; dark fear, right) across each experimental condition. Actograms are double-plotted, with each row showing two days. Colored bars adjacent to actograms denote data used to generate activity waveforms in respective conditions. Activity waveforms depict data averaged by shock group across all three cohorts (solid line), accompanied by 95% standard error of the mean (SEM) error bars, as well as with activity averages by cohort (transparent lines). Wind rose plots

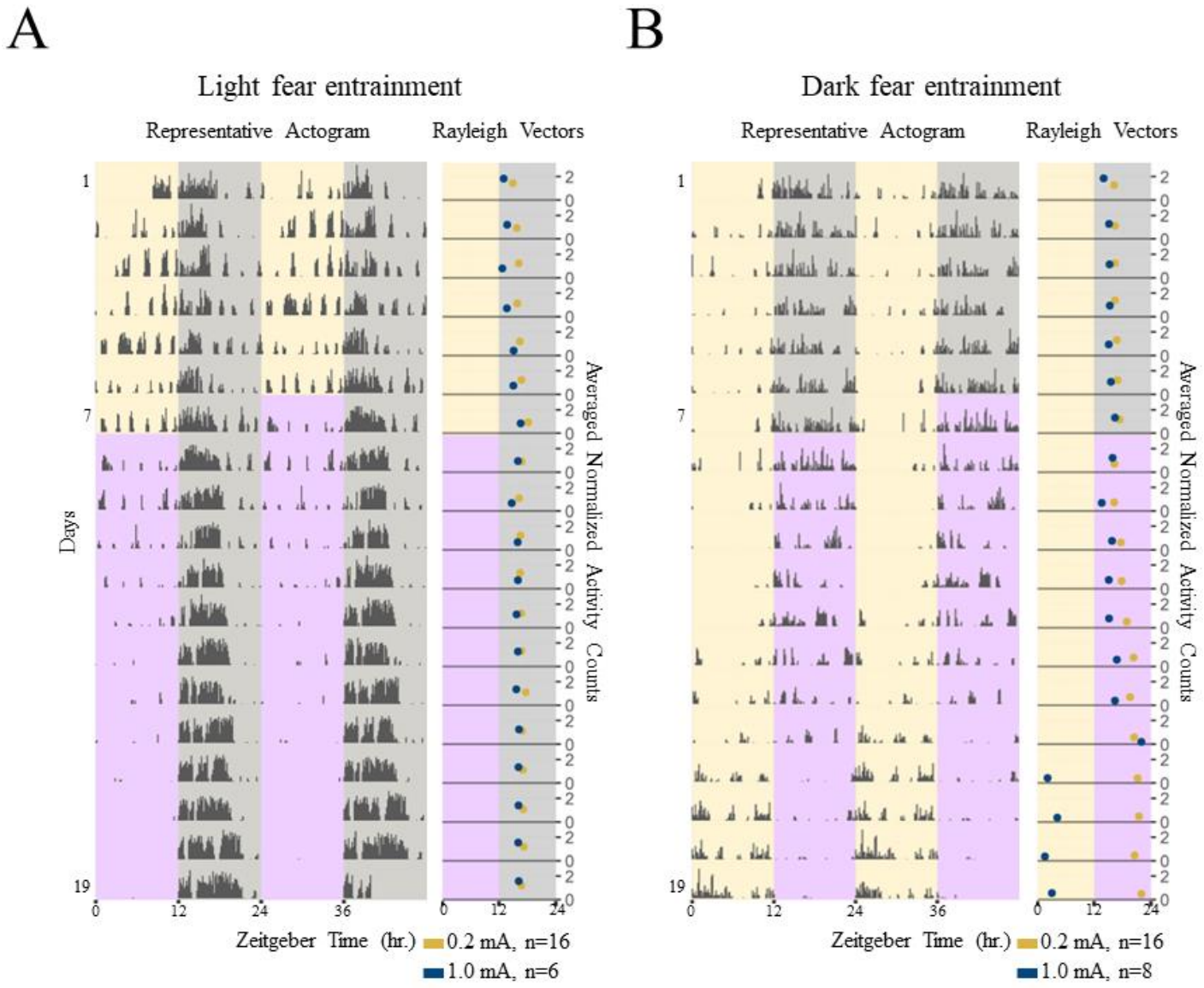
include overlayed Rayleigh vectors calculated by shock group. Light yellow denotes light phase; dark grey denotes dark phase; and purple denotes shock phase.

Figure II-ii. Change in activity during critical time windows over each experimental condition.



Note. Averaged activity in each condition for both shock groups are shown for zeitgeber time 0-1 (A) and 23-24 (B). Error bars indicate 95% SEM. ***indicates $p<0.001$. Light grey coloring indicates non-significance.

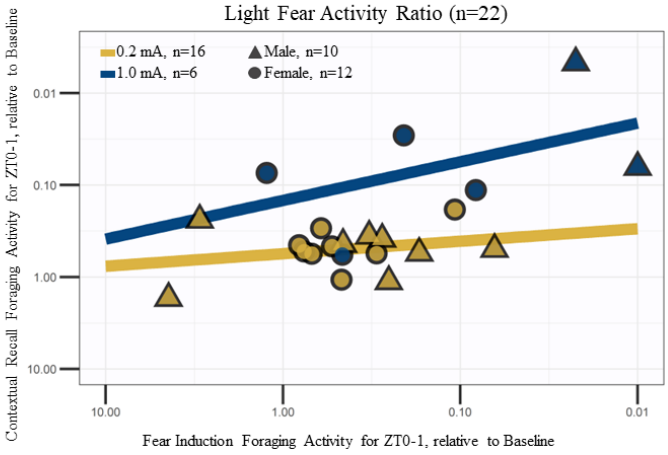
Figure II-iii. Speed of entrainment and foodshock intensity.



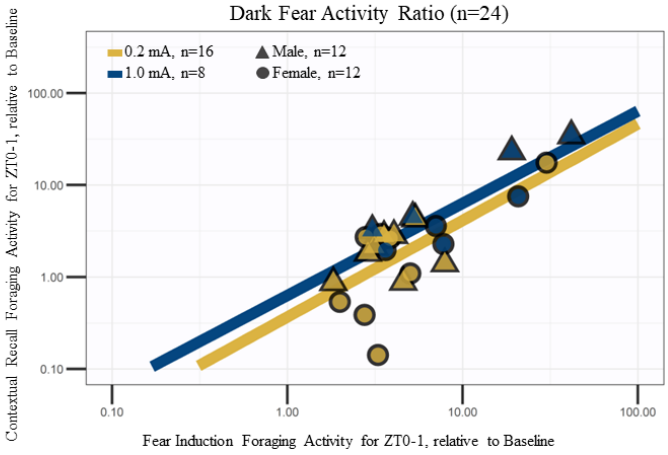
Note. Representative actograms from an individual mouse (left column) and Rayleigh vectors averaged across all mice (right column) over the course of fear induction paradigm in the light fear (A) or dark fear (B) group. Each row represents a day (double-plotted for actograms, with each row showing two days). Rayleigh vectors (indicating average circadian activity timing and magnitude) for each day are separated and displayed by footshock intensity (0.2, mA 1.0 mA). Light yellow denotes light phase; dark grey denotes dark phase; and purple denotes shock phase.

Figure II-iv. Linear relationship between strength of entrainment during fear induction and contextual recall activity during ZT0-ZT1.

A

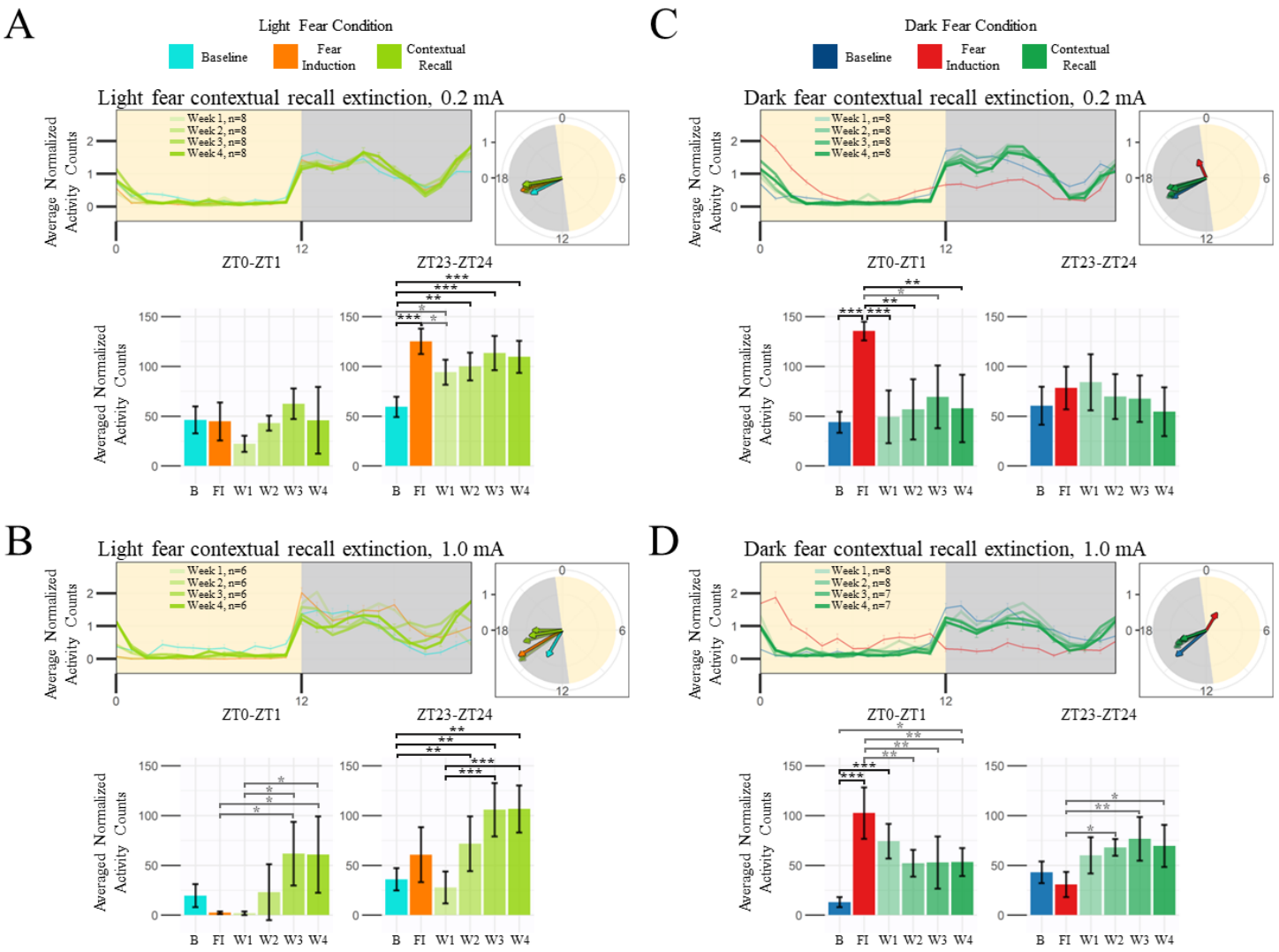


B



Note. Each graph displays linear relationship of activity in ZT0-ZT1 during fear induction (x-axis) and during contextual recall (y-axis). Both axes are in logarithmic scale, with higher numbers indicating greater activity during this time window relative to baseline, and lower numbers indicating lower activity relative to baseline. For light fear (A), axes range from 10.00 (bottom/left) to 0.01 (top/right); for dark fear (B), the axes range from 0.01 to 10.00. Each dot represents an individual mouse, is color-coded according to footshock intensity group, and its shape is determined by sex of the mouse. ZT = zeitgeber time.

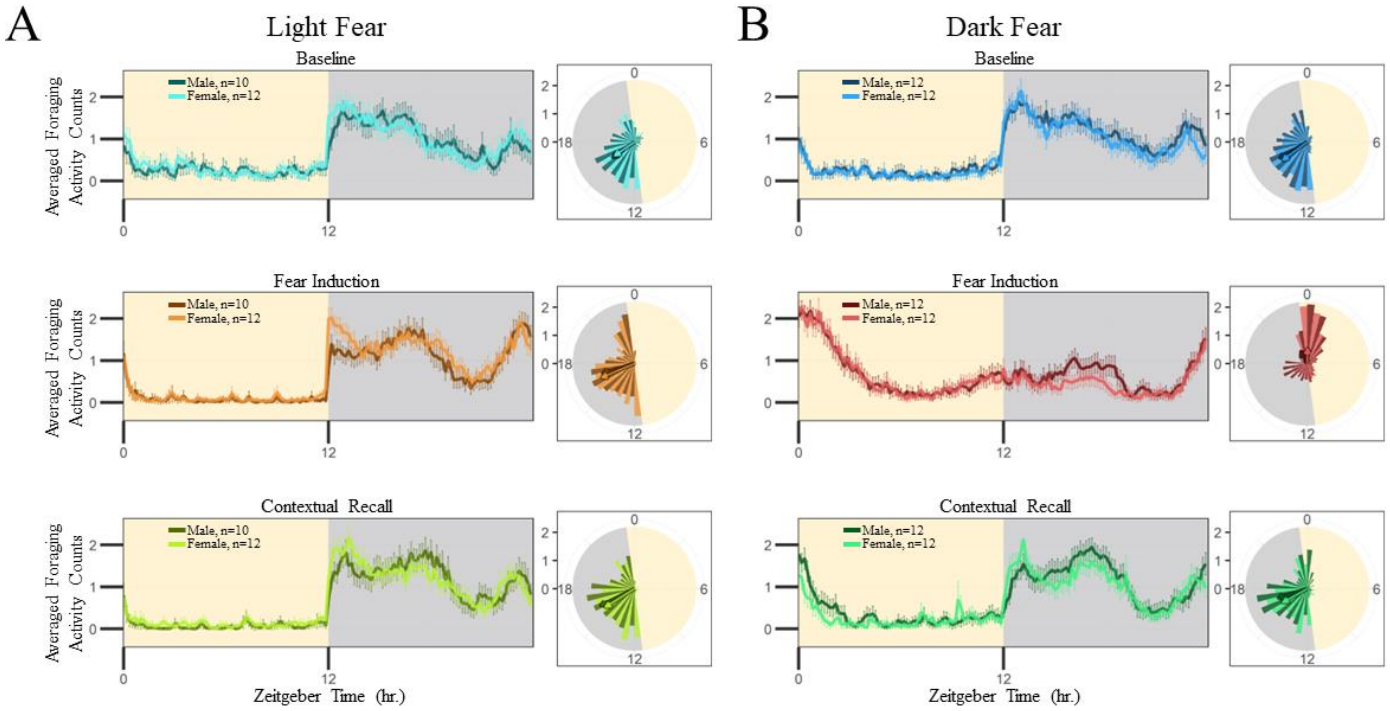
Figure II-v. Extinction of fear entrainment over four weeks, by shock group and footshock intensity.



Note. Data were separated by shock group and footshock intensity (A, light fear and 0.2 mA; B, light fear and 1.0 mA; C, dark fear and 0.2 mA; and D, dark fear and 1.0 mA). For each subfigure, top rows contain waveforms averaged across animals are displayed by condition and corresponding wind rose graphs (including overlayed Rayleigh vectors for each condition), and bottom rows contain bar graphs comparing activity across each condition in critical time windows. ZT = zeitgeber time; B = baseline; FI = fear induction, and W1, W2, W3, and W4 are weeks 1 through 4. Error bars in waveforms and bar graphs indicate 95% SEM. *indicates $p<0.05$,

**indicates $p<0.01$, and ***indicates $p<0.001$. Light grey coloring indicates non-significance after Bonferroni correction for multiple comparisons ($\alpha = 0.0033$).

Figure II-vi. Sex differences in activity by shock group and condition.



Note. Activity waveforms depict data averaged by sex across all three cohorts and are organized by shock group (A, light fear; B, dark fear) and experimental condition (top row, baseline; middle row, fear induction; bottom row, contextual recall). Each waveform includes 95% SEM error bars. Wind rose plots include overlayed Rayleigh vectors calculated by sex. Light yellow denotes light phase; dark grey denotes dark phase.

Chapter III. SLEEPING ON THE JOB: PUBLICLY AVAILABLE AUTOMATED SLEEP SCORING

Sleep is a highly adaptive process that plays an important role in numerous physiological functions. Sleep generally occurs in two stages: rapid eye movement (REM) and non-rapid eye movement (NREM). These stages are characterized by distinct electrophysiological signatures, and the temporal distribution of these stages during sleep defines a 'sleep architecture' that can have important consequences for critical physiological and behavioral processes including memory consolidation, mood, as well as immune and cognitive function.

The study of sleep architecture in the context of disease is an area of great interest. Sleep disorders are extremely common, impacting over 100 million individuals in the United States alone, and contribute to or are co-morbid with a wide range of neurological, metabolic and psychiatric conditions.[99,100]. Sleep disorders are usually diagnosed using polysomnography, which measures multiple physiological signals including electroencephalography (EEG), electrooculography (EOG), chin and leg electromyography (EMG), electrocardiography (ECG), breathing effort, oxygen saturation, and airflow.[101] However, EEG and EMG are essential and sufficient signals for the classification of behavioral states as either wake, NREM or REM.

The characteristics of each stage are currently standardized and defined by the American Academy of Sleep Medicine (AASM) Scoring Manual.[102] In humans NREM can be further subdivided into three distinct stages (N1 to N3), but this distinction is lost or ignored in studies of rodent models, with very few exceptions.[103] Despite this standardization, trained professionals exhibit an overall inter-scorer agreement of 82.6% for human polysomnographic recordings, which decreases to around 70% for N1 and N3 stages, and sleep spindles.[104–106] Recent findings suggest that this variability can be largely attributed to epochs that could legitimately be assigned to

multiple sleep stages,[107] and accounting for this discrepancy can largely decrease interscorer variability.[108] Unlike human sleep data, there is currently no "gold standard" for the manual classification of rodent sleep stage and manual scoring of sleep recordings is often error-prone, time-consuming, and generally must be done offline.

Attempts to develop automated algorithms for sleep staging date back to 1969[109] and have incorporated a wide range of approaches with varying degrees of success.[110] Previous protocols have been developed and validated using both human and rodent data,[111] often with good results but low adoption by the field. One reason for this limited reach could be because many methods were developed and described assuming the user has expertise in mathematics and computer programming. While more user-friendly approaches have been developed,[112] potential flaws of these methods include: validation on data acquired only from wild-type (WT) animals; near-exclusive use of standard laboratory conditions; requiring the purchase of a commercial license; and little to no information about the underlying algorithm or training data used to validate the approach.

Long-term chronic polysomnographic recordings are becoming increasingly valuable for studying the role of the circadian clock in regulating sleep/wake cycles. Accordingly, we have developed Sleep Identification Enabled by Supervised Training Algorithms (SIESTA): an open-source, automated sleep stage classification web application. We first generated a training dataset consisting of 20 days of ECoG and EMG recordings obtained from both WT and transgenic mice with disturbed sleep under multiple experimental lighting conditions. For every 10 seconds of recording, we extracted 54 total features from ECoG and EMG signals and used these data to evaluate a battery of supervised machine learning algorithms. We selected the bagging classifier using random forest, which showed the highest performance as determined by the F_1 score. We

aimed to reach an accuracy of at least 83% across behavioral states, reflecting the average inter-scorer reliability reported by the AASM.[106] When evaluated on test data withheld from the training set, this classifier performed well, achieving F_1 scores of 0.92 for wakefulness, 0.81 for REM, and 0.93 for NREM. We further validated our approach using data collected from the same transgenic mouse line in another laboratory and under different experimental conditions and found comparable levels of classification accuracy.

SIESTA is an open-source platform developed in Python. Our user-friendly graphical interface provides an intuitive workflow from data selection to final scored output and includes SIESTA's own training database as well as an optional training module where users can train SIESTA on their own data or contribute their own data to the preexisting training database.

III.i METHODS.

III.i.1 *Animals and housing conditions.*

All experiments with animals were performed in accordance with animal protocols approved by the Office of Animal Welfare at the University of Washington and Seattle Children's Research Institute. Mice with a heterozygous deletion of the *Scn1a* gene (*Scn1a*$^{+/-}$), which model an intractable form of epilepsy called Dravet syndrome (DS; hereafter referred to as DS mice), were generated by targeted deletion of the last exon, encoding domain IV from the S3 to S6 segment and the entire C-terminal tail of $Na_V1.1$ channel, as previously described.[113] The mice used in this study were generated by crossing heterozygous mutant mice of C57BL/6 background with wild-type (WT) C57BL/6 mice (both males and females of each genotype), resulting in only WT or heterozygous *Scn1a* mutant offspring.[114] The DS mouse has been previously described to have deficits in both the circadian and homeostatic regulation of sleep.[114–116]

III.i.2 *Electrocorticographic recordings – de la Iglesia Lab.*

Sleep was recorded as previously described.[117] Briefly, mice were anesthetized with isoflurane and placed into a stereotaxic device where isoflurane anesthesia continued throughout surgery. Each mouse was implanted with ECoG electrodes, consisting of dental screws (Pinnacle Technology, Lawrence, KS; No. 8209: 0.10-in.). Using aseptic technique, a midline incision was made above the skull. Recording electrodes were screwed through cranial holes as follows: over the left frontal cortex (1.5 mm lateral and 2 mm anterior to bregma) and over the right parietal cortex (1.5 mm lateral and 2 mm posterior to bregma), a ground electrode was placed over the visual cortex (1.5 mm lateral and 4.0 mm posterior to bregma), and a reference electrode was placed over the cerebellum (1.5 mm lateral and 6 mm posterior to bregma). EMG signals were obtained by placing a pair of silver wires into the neck muscles. The screws were connected, through silver wires, to a common 6-pin connector compatible with the Pinnacle recording device. The screws and connector were fixed to the skull with dental cement. Mice were implanted between 3 and 4 months of age to account for the long duration of circadian sleep experiments. After surgery, mice were housed in single recording cages under a 12:12 light-dark (LD) cycle. Mice had a recovery period of 1 week, were then fitted with a tether and a preamplifier (100x signal gain) connected to the Pinnacle Technology recording system, and they were allowed 1 day to acclimate before recording started. The ECoG and EMG signals were sampled at 400 Hz with low-pass filters of 80 Hz and 100 Hz, respectively. All recordings were saved and exported for processing using the European Data Format (EDF), an open-source file format commonly used in polysomnographic recording.

III.i.3 *Telemetric electrocorticographic recordings – de la Iglesia Lab.*

Surgery procedure for telemetric recordings followed the above protocol, but included changes to account for different requirements: only two ECoG electrodes were screwed through cranial holes (left frontal cortex, reference) and were connected to a Stellar Dual-Biopotential Transmitter (TSE Systems, Inc, Chesterfield, Montana; E-430001-IMP-60, Type BBTA-XS-C) before being insulated with dental cement. EMG signals were obtained by suturing leads from the Stellar Transmitter to neck muscles and the incision was sutured closed. Following recovery, data were recorded in two periods of 24 hours, spaced a week apart, at a sampling rate of 250 Hz using a Notochord-hem (version 3.5, Notochord Systems). Recordings were saved, exported as EDF files and post-processed (a 0.5 Hz high-pass filter; a 100 Hz low-pass filter, a 60 Hz notch-filter, and amplification by 10^{-3} were applied) prior to being used in analyses.

III.i.4 *Electrocorticographic recordings – Kalume Lab.*

Mice underwent survival surgery to implant ECoG and EMG electrodes under isoflurane anesthesia with additional subcutaneous bupivacaine (1mg/kg) for analgesia as described previously.[114] Using aseptic technique, a midline incision was made anterior to posterior to expose the cranium. Micro-screw ECoG electrodes, connected to fine (diameter: 0.130 mm bare; 0.180 mm coated) silver wires, were inserted through cranial burr holes drilled with a small drill bit. The ECoG electrodes were placed at visually identified locations – left and right frontal cortices, approximately 1 mm anterior to the bregma and 3 mm lateral. EMG electrodes were placed in back muscles. A reference electrode was placed at the midline cerebellum and a ground electrode was placed subcutaneously over the back and the skin was closed with sutures. All electrodes connected to a micro-connector that permitted the interface with the recording system and had a typical impedance < 10 kΩ. The implant was fixed to the cranium using dental cement and the surgical

incision was closed with sutures. Mice were allowed to recover for two to three days with carprofen MediGel put in the animal cages for 48 hours. ECoG patterns were evaluated prior to recording to ensure complete recovery from anesthesia.

Twenty-four to 48 hours after electrode implantation surgery, simultaneous video-ECoG-EMG records were collected in conscious mice on a PowerLab 8/35 data acquisition unit using LabChart 8.1 software (AD Instruments, Colorado Spring, Co). All bioelectrical signals were acquired at 1-KHz sampling rate and using a 100X amplification. The ECoG signals were processed offline with a 1-70 Hz bandpass filter and the EMG signal with a 3-Hz high pass filter.

III.i.5 *Behavioral experiments and environmental conditions.*

Our recordings came from experiments maintained in a ventilated, light-tight room under either a 12:12 LD cycle with 200-lux intensity or constant darkness (DD). To determine the effect of abrupt LD phase shifts on the phase of the circadian rhythm of sleep, we used a delay and advance "jetlag" protocol. After mice were entrained to a 12:12 LD cycle, a delay jetlag was initiated by extending the light phase by 6 hours (i.e., delaying the time of lights off by 6 hours). The mice stayed under this new 12:12 LD cycle until they re-entrained to the new cycle. The advance jetlag was initiated by shortening the light phase by 6 hours (i.e., advancing the time of lights off by 6 hours). To assess the endogenous circadian period of sleep, we released mice into DD for at least 14 days after they had synchronized to a 12:12 LD cycle.

III.i.6 *Signal processing and feature engineering.*

We used a total of 54 features extracted for each 10-second epoch, calculated from raw ECoG and EMG data. ECoG and EMG signals were prefiltered using a Hanning filter with a moving window of one minute.

From ECoG signals, we calculated power, energy and amplitude values for a series of spectral bands as described in previous studies of the ECoG and EMG signatures of sleep stages.[103] We defined the alpha band as 8-13 Hz, beta as 20-40 Hz, sigma as 11-15 Hz, gamma as 35-45 Hz,[118] and a frequency band indicating the occurrence of sleep spindles as 12-14 Hz. Because there is little consensus on the definition of mouse theta in the literature, we incorporated several previously described definitions as features (4-12 Hz,[119] 6-9 Hz[120], 5.5-8.5 Hz,[103] 7-10 Hz[119]). Next, we calculated three different power ratios between frequency bands. The first is the theta-delta ratio, which is commonly incorporated into heuristics used to discriminate sleep and wake states.[103] The other two have been used for sleep spindle detection[121] and are defined as the ratio of power values of frequencies between 0.5-20 Hz to those between 0.5-50 Hz; and the ratio of those between 11-16 Hz to those between 0.5-40 Hz. Additional features included the 90% spectral edge, 50% spectral mean, mean and median amplitude, the root mean square, variance, skewness, and kurtosis of the signal. Finally, we incorporated the number of zero crossings, the peak-to-peak range, and the spectral entropy of the signal.

From the EMG signal, we calculated the amplitude, signal variance, skewness, kurtosis, root mean square amplitude and the spectral entropy. A list of all the features and their descriptions can be found in Table S1 (Appendix A).

III.i.7 *Software development.*

All code was written and tested in Python 3.7.3, using Spyder 3.3.2 and Jupyter 5.7.2 on Windows, MacOS and Linux operating systems. The scikit.learn package version 0.20.1 was used for validating the machine learning algorithms.[122] We used Pandas 0.23.4, Pickle 3.0, Numpy 1.15.4 and Scipy 1.1.0 for data management and signal processing.

III.i.8 *Supervised learning algorithm selection.*

Using our manually-scored sleep data (Table 1), we tested the ability of several algorithms to identify each sleep stage. We used methods of one-step classification, wherein an epoch was classified as either wake, REM or NREM in a single step using one classifier; and two-level hierarchical classification, wherein one classifier was used to score an epoch as either wake or sleep, and a second classifier was used to further divide each sleep bout as either NREM or REM. The full list of classifiers tested is given in Table 2.

Because the random forest algorithm is at the core of SIESTA, we provide a brief discussion of its implementation as follows. The random forest algorithm is an ensemble learning method that takes into account the classification results of multiple decision trees.[123] Although effective, decision trees often generalize poorly and underperform on classification problems with many input variables.[124] To compensate for this, random forest classifiers utilize a combination of decision trees such that each tree depends on a random, independently sampled vector of the dataset that has the same distribution for all decision trees. After several trees are created, the random forest algorithm "votes" on the most popular class. A margin function is used to determine by how much the average number of votes from decision trees for the right class exceeds the vote total for any of the other classes (e.g., wake, NREM or REM).[124]

SIESTA employs two other similarly structured ensemble learning methods called bagging and gradient boosting,[125] using random forest as the base classifier for each. A bagging classifier[126] is an ensemble meta-estimator method that creates random individual results by training each classifier on a random redistribution of the training dataset, and then aggregating their individual predictions (by a vote) to form a final prediction. This meta-estimator is typically used to reduce the variance of the core method (in this case random forest). Similarly, gradient boosting is an

ensemble learning method that boosts the performance of "weak learners" such as decision trees using a loss function, which determines how well the model fits the training data.[124]

In the screening for all algorithms, we performed 20-fold cross validation with shuffle and used F_1 score to assess performance. The F_1 score is the harmonic mean of the precision and recall of the test. Precision is given by the number of true positives (TP) divided by the sum of false positives and TP, and recall is calculated as the number of TPs divided by the sum of false negatives and TPs. Assessing test performance using F_1 score compensates for the imbalance in the occurrences of wake, NREM and REM bouts, and as such is widely used in machine learning-based approaches to automatic sleep scoring.[110] All results are shown as mean with standard deviation.

III.i.9 *Inter- and intra-scorer reliability.*

To calculate inter-scorer reliability, we randomly selected a 24-hour recording from our training dataset and compared the agreement between 4 experienced manual scorers using both percent agreement and Fleiss' kappa.[127] In accordance with previous reports[128] of inter-scorer reliability for sleep scoring between multiple scorers, we set the following levels of agreement for evaluating: 0-0.20 = slight agreement, 0.21-0.40 = fair agreement, 0.41-0.60 = moderate agreement, 0.61-0.80 = substantial agreement, 0.81-1.0 = near perfect agreement. values are reported as being within a 95% confidence interval (CI).

We then calculated intra-scorer reliability, or how consistently each manual scorer classified a single epoch as being the same state when presented with the epoch multiple times. We wrote a custom R script to randomly select 600 10-second epochs from the same 24-hour recording session, duplicate each of these epochs 5 times, randomly insert them into the original data file, and finally randomize the order of all epochs in the data file. This new file was then

scored by 4 experienced manual scorers, and the score consistency between duplicate epochs was evaluated using percent agreement for each individual scorer (Table 1).

III.i.10 *Training dataset.*

The training dataset consisted of 20 total days of recording from 20 different mice, scored by 4 different experts blind to experimental conditions. Each recording was a day long (24 hours), under several different environmental lighting conditions including LD (12h light:12h dark), Constant Darkness (DD), and LD in the days immediately following induced delay or advance jet lag. These conditions covered the most common light cycle paradigms used in the study of behavioral circadian rhythms, with DD and jet lag known to induce changes in sleep timing and quantity.[129] Data from these experiments were selected to make the dataset more generalizable to recordings obtained under diverse experimental conditions.

III.i.11 *Correlation matrix and dendrogram.*

We generated a correlation matrix to visualize the Pearson correlation coefficients calculated for pairs of features. We used single linkage clustering to generate the feature order.[130] The dendrogram measures the pairwise distance between the features using a threshold and metric to divide the measurements called the cophenetic correlation coefficient. This allowed us to cluster the variables according to the average distance between each subset of merged features.

III.i.12 *Dimensionality reduction using sequential feature selection.*

Sequential Feature Selection belongs to a family of greedy search algorithms that are used to reduce an initial d-dimensional feature space (where d is the number of features extracted from the EDF file) to a k-dimensional subset (where k is the target number of features, and $k < d$). These methods select an initial size for the subset of features, and sequentially find the features that are

most informative at this timestep.[131] The process is repeated at the next timestep, choosing the next feature subset depending on the previously selected features and the metric of the classifier. In the sequential forward selection (SFS), a small feature subset is selected, and the algorithm adds a single feature at a time to the initial subset. This process is repeated until all features are added back in. For the sequential backward selection (SBS), the initial subset is of size d or near d, and the algorithm removes one feature at each timestep.

III.ii RESULTS.

III.ii.1 *Inter- and intra-scorer reliability.*

To evaluate the consistency of manual scoring and ensure the robustness of our training dataset, we first calculated measures for both inter-scorer and intra-scorer reliability. Four experienced manual scorers from our laboratory were asked to score a total of 12,608 10-second epochs as wake, NREM or REM, and inter-scorer reliability was determined using both Fleiss' kappa and percent agreement. Across all three stages, manual scorers displayed high Fleiss' kappa agreement—ranging between 0.73 and 0.78—particularly in epochs scored as wake and REM, with NREM Fleiss' kappa agreement being slightly lower (Table 1). Wake and NREM showed the highest percent agreement, with the lowest for REM bouts (Table 1).

Next, we evaluated intra-rater reliability using percent agreement. We found that, on average, our four manual scorers were highly consistent in their scoring across all three states (Table 1). This agreement was slightly higher for wake bouts compared to NREM and REM bouts (Table 1).

III.ii.2 *One-step classifier with complete dataset.*

Once we established consistency and high inter-rater reliability in our manual scoring, we examined which commonly used supervised learning algorithms performed with the highest accuracy in sleep stage classification when trained on our dataset (see Table 2 and Methods section for complete list). We tested each algorithm with the feature values both without scaling (Table 2) and with scaling (Table S2, Appendix A), as the performance of some algorithms changed depending on scaling. Although several algorithms produced an F_1 score that was better than chance using the one-step approach, classification using the gradient boosting classifier (GBC) and bagging classifier, random forest (BCRF) resulted in the highest F_1 scores.

Because both GBC and BCRF do not require scaled data as a prerequisite to their use[126,132] they did not show a change in performance for one-step classification whether the data were scaled or unscaled.

III.ii.3 *Hierarchical classifier with complete dataset.*

Using a hierarchical approach, we observed improved detection of NREM and REM states using most of the algorithms we tested, with a F_1 score of 0.907 ± 0.032 in the case of BCRF, an improvement from the F_1 score of 0.827 ± 0.059 obtained using the one-step approach (Table 2). This improvement was not of equal magnitude in all three states scored. Table 2 and Table S2 (Appendix A) illustrate the performance of each classifier in first distinguishing sleep and wake, and second distinguishing between NREM and REM. The BCRF algorithm was the highest performing in terms of classification speed and accuracy (Tables S3 and S4, Appendix A) and was subsequently selected as the classifier for SIESTA.

III.ii.4 *Generalization of algorithm across genotypes and training datasets.*

We validated the performance of SIESTA using a test dataset excluded from the training dataset. This test data included WT and DS mice, manually scored by the same experts that labeled the original training dataset.

First, we tested SIESTA with only the WT data, training the algorithm with all but one WT mouse, using the scores from these new data as targets to validate the scoring method. This process was repeated 20 times, with both the one-step and hierarchical classifiers. Complete results of the scoring can be seen in Table 3. Finally, we added the first 2 hours of the manually scored data of the target mice to the dataset to reduce the interindividual variability of the individual mice in the training process,[112] and then re-trained the BCRF in the hierarchical classifier with this new dataset. Using this approach, we increased classification performance for all three stages (Table 3).

These three different approaches (one-step classifier, hierarchical classifier, and hierarchical classifier with the first 2 hours to the dataset) were repeated using a dataset consisting of only DS mice. Compared to WT data, these F_1 scores were lower than for the WT database in the first two methods but increased in the last method (Table 3). This discrepancy could be due to the known circadian rhythm and sleep disturbances present in DS mice[15,16].

Lastly, we attempted to use a training dataset consisting of only data from WT mice to classify data from DS mice. Once again, the best method was the hierarchical classifier using the first 2 hours of the target mice (Table 3). Based on these results, we concluded that although algorithm performance is best when the target data and the data in the training dataset are similar, adding scores for 2 hours of the target data increases performance in all cases (Table 3).

Although scoring sleep in 10-second increments is common, this temporal resolution can result in the misclassification of boundary epochs, as sleep stage transitions can happen on rapid timescales. Therefore, we also tested the performance of the algorithm when tasked with scoring data in 5-second epochs. We normalized the database and validated the scores with a WT and a DS mouse. The F_1 scores for these mice were similar to those obtained using 10-second epochs, with a score for REM of 0.75 in WT and 0.83 in DS using the One-step classifier; and 0.76 in WT and 0.85 in DS with the hierarchical classifier (Table S5, Appendix A). These results validated the utility of this method even when changing the epoch size.

III.ii.5 *Generalization of algorithm to datasets collected under different conditions.*

Next, we assessed the performance of SIESTA in classification of sleep stages from data recorded in a different laboratory. We trained both BCRF and GBC algorithms on either WT data, DS data or both from the de la Iglesia laboratory dataset (IL). We then validated performance on a dataset consisting of 8 to 10-hour recordings from 4 DS mice recorded in the Kalume laboratory (KL) using the hierarchical scoring approach in the same process as described above (Table 4).

In all tests, the GBC and BCRF methods performed comparably well in both steps of classification, but slightly below the classification accuracy of the scoring when both training and testing data were IL data (Table 3). Notably, both classification methods performed better on distinguishing between NREM and REM in KL data than in IL data (Tables 3 and 4). Additionally, both methods performed better distinguishing wake/sleep in data gathered using a telemetric system, and distinguished between NREM and REM with moderate accuracy (Table 4). Together, these results validate the utility of SIESTA, and its original training data set, to score data recorded in a different lab using a different recording system, and data gathered with a telemetric data collection system.

III.ii.6 *Feature evaluation and reduction.*

Next, we decided to further characterize the features that we extracted from our raw data for training. Most features were empirically chosen based on those used to identify sleep stages according to the AASM, as well as previously developed methods for automatic sleep scoring in rodents. The aim of this feature characterization was to find a subset of the most important features for successful classification. Reducing the number of features used for scoring would allow for faster and more flexible classification that could work in a real-time scoring paradigm without significantly reducing classification accuracy.

We computed a correlation matrix to determine the extent of the similarity between features used in training, and found highly correlated features ($r>0.9$) that appeared to cluster primarily in two main groups: one with features describing the amplitude of the signal (in several frequency bands of both the ECoG signal and EMG signals), and a second cluster that included the relative power of all but the delta band of the ECoG signal, along with other features that describe the frequency domain of the signal (the spectral mean and spectral entropy of ECoG) (Fig. 1).

To further characterize the clustering and correlation of the features used for training, we constructed a dendrogram (Fig. S1, Appendix A). We found 3 clusters and 1 independent variable. The first cluster (light blue) consisted primarily of variables related to the amplitude of the signals, similar to the correlation matrix. The second cluster (red), containing the fewest features, included the relative power and amplitude of the delta band and the spectral characteristics of the EMG. The third cluster (green) contained a mix of the relative power and the energy of most frequency bands. The only feature that did not belong to any cluster was the second ECoG index ('ECOGrel2', the ratio of the relative power of the 0.5-20 Hz band and the 0.5-50 Hz band), which has been previously identified as being useful in automated approaches to sleep scoring.[133]

After identifying subsets of features that were highly correlated, we next sought to find the optimal subset of features for classifying sleep stages. We used both SFS and SBS to identify the features that were the most critical for sleep stage identification. In both selection algorithms the inflection point (the number of features at which increasing the size of the feature subset does not significantly improve the performance of the classifier[134]) of the performance curve was around 6 features (Fig. 2). Both algorithms were run with the one-step classifier using BCRF three times, each with 5-fold cross-validation, to evaluate the robustness of the feature selection (Fig. S2, Appendix A). In most cases the same features were identified as being the most important for classification performance, and the inflection point of the performance curve always occurred at 6-8 features. The list of all features found by the sequential selection algorithm can be found in Fig. S2 (see Appendix A). The top 5 features consistently identified by both SFS and SBS as being most important were: ECoG Zero crossing, EMG Amplitude, ECoG Sleep Spindle Hanning window (listed as 'ECoG Spindlehan' in Fig. S2 and Table S1, Appendix A), ECoG Delta power and EMG Spectral Entropy. Most of the features identified have been already reported to be relevant for classifying sleep stages, either by being used to score sleep (ECoG Delta[135]) or being present mainly in certain stages (sleep spindles[136]).

As expected, several of the highest-rated features in all the iterations of the sequential search method were the same. We found 5 features that were present in all cases, and these features were spread across the three clusters identified in the dendrogram (Fig. S1, Appendix A) as well as the two most correlated clusters of the correlation matrix. Similarly, the correlation matrix revealed one feature identified by SFS and SBS in each cluster while the rest were not highly correlated with any other.

III.ii.7 *Graphical interface.*

Given the success of SIESTA in sleep stage classification, we sought to ease the access of our code to the broader sleep and circadian biology community. We developed a web-based graphical interface that allows researchers to use our method, from pre-processing to scoring (https://siesta.azurewebsites.net/). SIESTA takes European Data Format (EDF) files as input, an open-source file format commonly used in both human and animal polysomnographic recordings.[137] The interface performs three basic functions. 1) Feature extraction: this function extracts the features from the EDF file (test file) to be scored as well as from an EDF file (train file) that the user can provide to train a new scoring algorithm. 2) Model fitting: this function fits the training model with the features extracted from the train file. 3) Data scoring: this function scores sleep stages using the extracted features from the test file. SIESTA returns both a file with the extracted features as well as the sleep stages for each epoch. Users have the option to input parameters that are specific to their experimental needs, including the sampling frequency used in their recording, their desired epoch length and information about signal amplification. Both the features and final score files are output in .csv format. The stored database, which consists of both data from the de la Iglesia lab (IL) and the Kalume lab (KL) data, uses the Pickle library of Python to store the training results in a non-binary file so that users do not need to re-train the algorithm each time they want to score new data. A flow diagram of the complete training process and the actions available in the graphical interface is illustrated in Figure S3 (see Appendix A).

We provide the source code freely for the graphical interface, a user's manual, and the database that we used to train our algorithm, as well as scripts used to validate SIESTA on GitHub® [specific link provided after acceptance of manuscript].

III.iii DISCUSSION.

Using open-source coding and data management tools, we developed a novel automated sleep stage classification system for rodent polysomnographic data we call SIESTA. First, we compared the performance of a set of commonly used classifiers in identifying sleep stages from a training dataset containing data from both WT and transgenic mice with disturbed sleep. From this comparison, we identified the highest-performing classifier: a bagging classifier using random forest as the base classifier. We demonstrated that the accuracy of several classifiers was heavily influenced by the pre-processing of the signals, but the chosen method was not affected by this process. While other groups have reported using the random forest classifier to score sleep,[138–140] fewer have used the bagging[141] or gradient boosting ensemble classification methods.[129] To our knowledge, even fewer approaches have used the bagging classifier with random forest as the base classifier. We obtained high classification accuracy on par with AASM inter-rater reliability standards using a one-step classifier for all stages, and this score was further improved using a hierarchical classifier. We were able to improve classification accuracy even further when using a small amount (2 hours) of manually scored epochs from the data needing to be scored. Finally, we validated SIESTA using data collected from another laboratory and under different behavioral and recording protocols and obtained classification accuracy that was similar.

We have used this method to score data obtained in two laboratories, including continuous long-term polysomnographic recordings lasting over 3 weeks. Scoring this data manually is normally a tedious, error prone and time-consuming process. The run time for the feature extraction with the SIESTA code is less than a minute. In other words, we can score 24 hours of recording in under 1 minute if we use our complete training data set. If the algorithm is trained on new data, scoring 24 hours of recording takes approximately 5 minutes. Information on

computational times associated with training and scoring with the different methods and subset of data can be found in Tables S3 and S4 (see Appendix A).

We welcome users to contribute both data and code to SIESTA to improve its performance and better suit their own needs. Accordingly, we provide the manually scored data we used to train the algorithm, as well as the source code of the classifier and feature extraction methods [specific link provided after acceptance of manuscript]. We hope that SIESTA is used in a wider range of experimental conditions and mouse lines, which could eventually contribute to a more comprehensive training dataset. A key feature of this code is that we do not use a time-dependent model, meaning the score of each epoch is independent from the previous epoch, so data from experimental animals with altered sleep architecture can be scored in an unbiased way. This characteristic allows for the training dataset to be replaced, modified, or expanded with ease, allowing for more flexible scoring. Even though scoring can be improved when taking the previous epoch into account[137] this can lead to mistakes when the sleep architecture is atypical.

Scoring with SIESTA can be done without extensive coding knowledge, and our goal is that our simple interface and transparent analysis pipeline can make a novel contribution to the field of automatic sleep scoring and facilitate studies involving the chronic monitoring of sleep in the research community.[142] Additionally, SIESTA dramatically reduces the time needed to obtain sleep scores from raw data compared to manual scoring. Our code is modular and separated into 3 main processes that comprise our analysis pipeline: feature extraction, training of the algorithm with the dataset and scoring of the experimental data file. All of the input and output files are in standardized open-source formats, making the manipulation of data using SIESTA easy. The motivation behind this open-source approach and the free release of SIESTA is to provide a

trustworthy application that can be understood and validated by the community, encouraging any comment or modification that could improve the scoring process.

One of the main limitations of our approach is that we are likely missing the classification boundary epochs (epochs that include more than one state or a state change). These epochs can be difficult for both human and machine scorers to classify, and as such may decrease the efficiency of our algorithm. Additionally, neither our manual nor automatic scorers have a criterion for the rejection of an epoch, so all data is stored in one of the predetermined labels (including epochs containing artifacts and transition epochs), which could also lower the performance of the algorithm. We opted for this approach to simplify the use of the software and manual scoring of the recording. Even with this limitation, we obtain F_1 scores of between 0.83 and 0.93 in all states.

Notably, SIESTA performed comparably when validated on data collected in a second laboratory using only IL data, although both datasets contained data from DS mice. SIESTA additionally performed well when validated on data collected using telemetric data collection system, despite differing appearance of artifacts in telemetric vs. tethered data collection systems (data not shown). In the future, we aim to expand the training dataset to include more data from additional mouse lines with altered sleep phenotypes. Additionally, we hope to test our approach with data obtained from humans and other animal models, but there is recent work using similar approaches that lead us to believe our method can be effective in other models.[143,144] Indeed, a recent study demonstrated that another supervised learning algorithm, deep convolutional neural networks, can be used to predict sleep stages from manually scored data in narcoleptic mice with a comparable degree of success as our own approach.[145]

An additional value of this work is the characterization of the features we use to train our model and score recordings. Through feature correlation and reduction, we specify a small subset

of features needed to maintain high accuracy of SIESTA. Most of these features have been described in the literature to have a biological meaning (e.g. the relative power of the delta frequency band in the ECoG signal, EMG amplitude, and sleep spindles), which further validated our approach when they were found to be key in successful scoring. Other features are usually more prominently associated with other neural phenomena, as in the case of ECoG Zero-Crossing in the context of seizure and interictal spike detection.[146]This might point to underlying biological processes that we have yet to fully characterize and are being identified by an unbiased machine learning algorithm. These supervised classifiers could also be finding success in scoring using the interaction between different features in the subset identified by SFS and SBS, as an emergent property of the complex ECoG and EMG features that by itself might not have a direct biological interpretation.[147] Finally, the reduced subset of features we identified can be useful in the development of a real-time scoring paradigm, allowing for easier long-term closed-loop manipulations of sleep stages in rodent models with both normal and pathological sleep, a potentially valuable tool for the sleep and circadian biology communities.

Although deep learning approaches have shown recent success in automated sleep scoring in both human patients and animal models (with large datasets and promising results),[148–151] we opted for a shallow-learning approach without dimensionality reduction, allowing for greater interpretability of the results. The knowledge that the algorithm is using features with known biological relevance makes these results easier to approach and understand by the medical and research communities. Still, future studies should continue to directly compare the efficacy of supervised, unsupervised and deep learning approaches to scoring rodent polysomnographic data from diverse mouse lines and experimental conditions.

Finally, the study of the circadian regulation of sleep can produce long time series data that are often cumbersome to analyze. Our hope for SIESTA is not only to create an open-source, community-driven tool for sleep and circadian biologists, but also to encourage other researchers in the field not to shy away from performing experiments that would produce otherwise unwieldy datasets. These approaches are extremely valuable in furthering our understanding of the relationship between sleep, the circadian system and neurological and psychiatric conditions. Long-term monitoring of sleep in pre-clinical models of disease is an indispensable tool in understanding and leveraging these interactions to address the impact of sleep on health.

III.iv TABLES.

Table III-i. Inter- and intra-rater reliability metrics for manual scores used in training datasets.

	Inter-Rater Reliability			
	Total	Wake	NREM	REM
Fleiss Kappa [95% CI]	0.75 [0.744-757]	0.76 [0.758-0.770]	0.73 [0.726-0.738]	0.78 [0.773-0.786]
Agreement	87.34%	93.55%	89.04%	84.73%
# Epochs (% of Total)	12608	5934 (47.07%)	6047 (47.96%)	627 (4.97%)
	Intra-Rater Reliability			
Rater	Total	Wake	NREM	REM
A	98.53%	98.89%	98.62%	95.14%
B	91.60%	93.62%	87.71%	95.00%
C	96.17%	96.21%	96.32%	93.85%
D	96.90%	97.47%	97.12%	89.33%
Average	95.80%	96.55%	94.94%	93.33%

Note. Cells under intra-rater reliability represent % agreement.

Table III-ii. Classification accuracy by algorithm and method.

Algorithm	One-step	Hierarchical	Hierarchical + 2 h
Logistic Regression	0.78 (0.081)	0.35 (0.074)	0.86 (0.060)
Linear Discriminant Analysis	0.77 (0.090)	0.85 (0.077)	0.85 (0.077)
K-nearest Neighbors Classifier	0.76 (0.074)	0.51 (0.009)	0.84 (0.060)
Decision Tree Classifier	0.72 (0.076)	0.84 (0.043)	0.84 (0.038)
Gaussian Naïve Bayes	0.52 (0.18)	0.35 (0.87)	0.60 (0.149)
Passive Aggressive Classifier	0.71 (0.079)	0.48 (0.055)	0.77 (0.0113)
Ridge Classifier	0.70 (0.082)	0.85 (0.077)	0.85 (0.077)
Logistic Regression, cross validation	0.79 (0.077)	0.36 (0.086)	0.86 (0.062)
Bernoulli Naïve Bayes	0.60 (0.163)	0.50 (0.032)	0.75 (0.118)
Nearest Centroid	0.54 (0.129)	0.49 (0.013)	0.67 (0.135)
Random Forest Classifier	0.80 (0.061)	0.89 (0.030)	0.89 (0.028)
Ada Boost Classifier	0.79 (0.066)	0.88 (0.040)	0.88 (0.040)
Bagging Classifier, Decision Tree	0.79 (0.061)	0.89 (0.030)	0.88 (0.033)
Extra Trees Classifier	0.80 (0.063)	0.89 (0.031)	0.89 (0.033)
Gradient Boosting Classifier	**0.83 (0.057)**	**0.90 (0.035)**	**0.90 (0.035)**
Perceptron	0.72 (0.112)	0.49 (0.006)	0.78 (0.077)
Bagging Classifier, Random Forest	**0.83 (0.060)**	**0.91 (0.032)**	**0.91 (0.032)**
Bagging Classifier, Extra Trees	0.82 (0.062)	0.90 (0.034)	0.90 (0.032)

Note. Values represent F_1 scores (with standard deviation in parentheses). Hierarchical + 2 h refers to the method of adding two hours of manually scored data from the target mouse to the training set for the classifier. All data are unscaled.

Table III-iii. Sleep stage classification accuracy using subsets of the training data.

	Wild Type Training/Wild Type Validation			
Method	Wake	Sleep	NREM	REM
One-Step	0.94 (0.015)		0.90 (0.017)	0.77 (0.038)
Hierarchical	0.94 (0.017)	0.92 (0.021)	0.94 (0.019)	0.81 (0.036)
Hierarchical + 2 h	0.96 (0.031)	0.94 (0.038)	0.95 (0.039)	0.84 (0.054)
	Dravet Syndrome Training/Dravet Syndrome Validation			
Method	Wake	Sleep	NREM	REM
One-Step	0.85 (0.059)		0.82 (0.094)	0.60 (0.048)
Hierarchical	0.85 (0.061)	0.82 (0.115)	0.91 (0.054)	0.70 (0.084)
Hierarchical + 2 h	0.94 (0.013)	0.93 (0.028)	0.95 (0.016)	0.81 (0.054)
	Wild Type Training/Dravet Syndrome Validation			
Method	Wake	Sleep	NREM	REM
One-Step	0.84 (0.091)		0.86 (0.048)	0.68 (0.081)
Hierarchical	0.85 (0.097)	0.88 (0.058)	0.88 (0.072)	0.75 (0.082)
Hierarchical + 2 h	0.88 (0.049)	0.89 (0.030)	0.89 (0.037)	0.76 (0.061)

Note. Cells contain F_1 accuracy scores (and standard deviation in parentheses). Hierarchical + 2 h refers to the method of adding two hours of manually scored data from the target mouse to the training set for the classifier. All data is unscaled. WT database n= 14, Dravet syndrome database n=6.

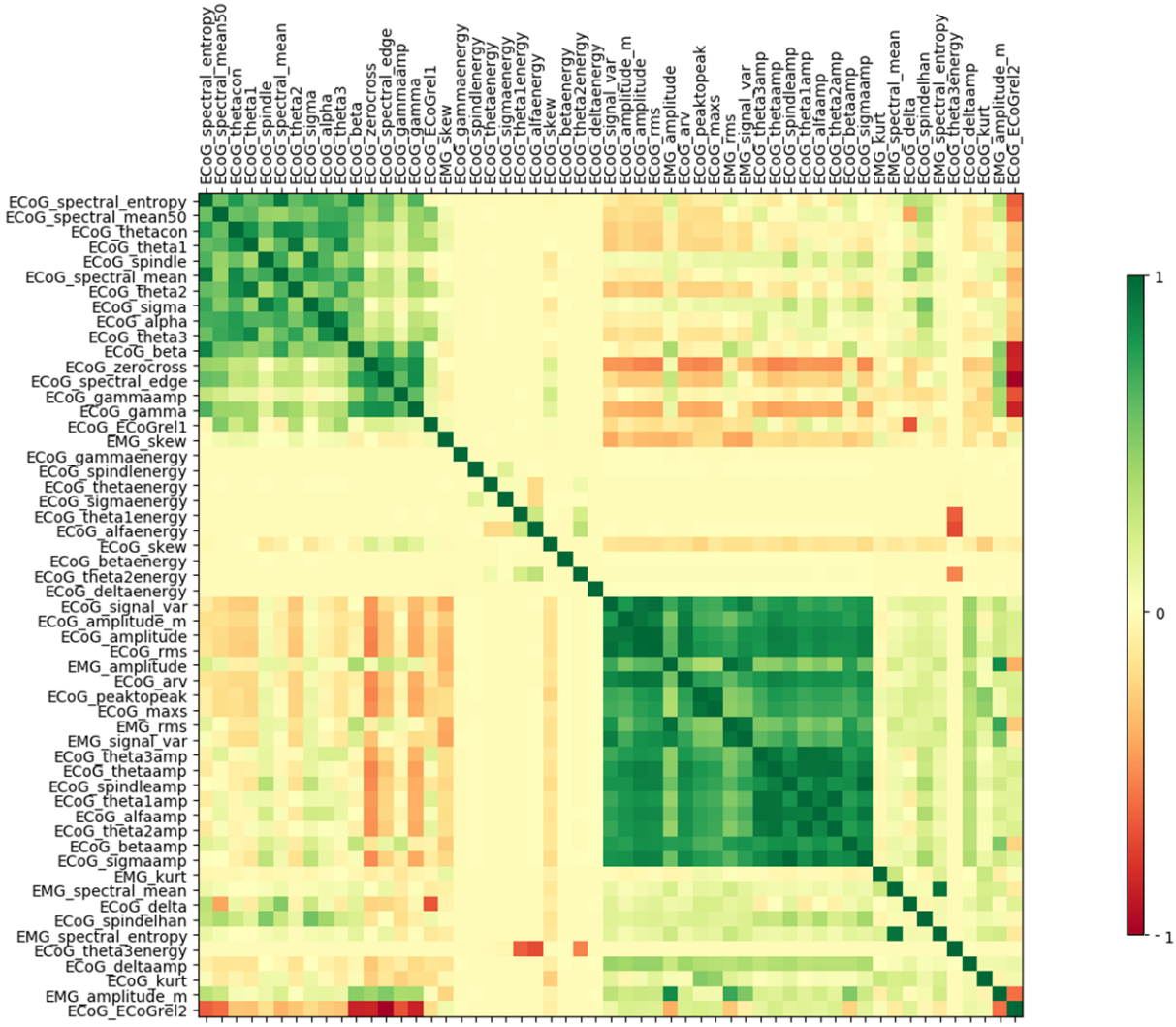
Table III-iv. Comparison of sleep stage classification accuracy of data from another laboratory and data from another data collection modality.

IL Wild Type Training/KL Validation				
Method	Wake	Sleep	NREM	REM
GBC	0.77	0.86	0.95	0.80
BCRF	0.74	0.78	0.94	0.78
IL Dravet Syndrome Training/KL Validation				
Method	Wake	Sleep	NREM	REM
GBC	0.68	0.73	0.94	0.83
BCRF	0.69	0.73	0.94	0.82
IL Full Dataset Training/KL Validation				
Method	Wake	Sleep	NREM	REM
GBC	0.77	0.89	0.91	0.77
BCRF	0.79	0.88	0.91	0.77
Full Tethered Dataset Training/Telemetric Validation				
Method	Wake	Sleep	NREM	REM
GBC	0.88	0.84	0.82	0.67
BCRF	0.89	0.86	0.85	0.70

Note. Cells contain F_1 accuracy scores obtained using the hierarchical classification approach. All data is unscaled. de la Iglesia lab Wild type database n= 14, Dravet syndrome n=6, Telemetric n = 2 (wild type). Kalume lab Dravet syndrome database n = 4. IL = de la Iglesia lab, KL = Kalume lab, GBC = gradient boosting with random forest classifiers, BCRF = bagging random forest classifiers.

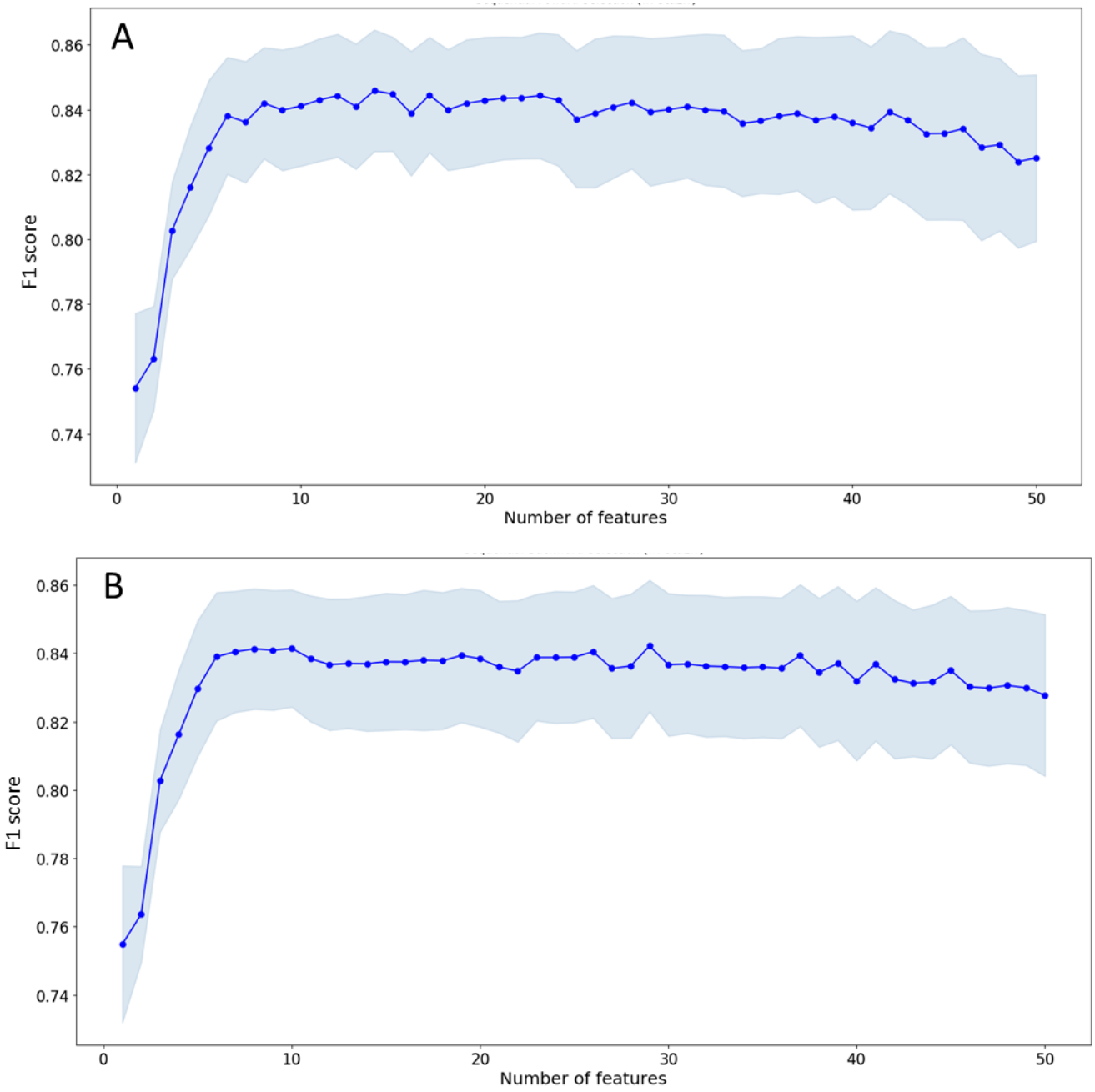
III.v FIGURES.

Figure III-i. Correlation matrix of the features from the complete dataset.



Note. The data was reorganized via the clustering algorithm described in the Methods section to identify highly correlated subsets of features (in both positive [green] and negative [red] correlation). Each of the above features is described in detail in Table S1, Appendix A.

Figure III-ii. Performance under sequential feature selection of the BCRF algorithm trained on the complete training dataset using a one-step approach to classification.



Note. The blue dots indicate the mean F_1 values from 3 runs of the sequential feature selection method, and the light blue shaded area represents the standard deviation in each case. **A)** Sequential Forward Selection, **B)** Sequential Backward Selection.

Chapter IV. SLEEP, WITH BOTH EYES OPEN: SYSTEMIC BIAS, INTERSECTIONALITY, AND SLEEP RESEARCH

As demonstrated over the past three chapters, there exist a myriad of ways in which sleep can be studied. Each method highlights different components implicated in the systems controlling sleep that can, in turn, inform decisions on how sleep disruptions can be most effectively discerned and addressed in humans. In Chapter II, for example, I showed how we may understand the origin of sleep disruptions in individuals with post-traumatic stress disorder (PTSD) through the potential interactions between neural systems differentially governing circadian rhythm and the innate fear response. In Chapter III, I demonstrated how we may use computational methods to study and analyze sleep patterns in mice, with the goal of both making this process as user-friendly as possible and setting up potential translational methods for humans. While the core methodology between the two chapters differ — the former being primarily concerning behavioral data; the latter, computational analyses — both are focused on the concept of sleep, serving as two examples of ways one may study sleep. In this way, the act of studying sleep is akin to looking inside of a kaleidoscope; shift your gaze — your method of inquiry — even slightly, and the observations you work with — your variables — will shift accordingly. That is to say, when studying sleep, we are never *just* studying sleep; there is likely not a singular, concrete concept of 'sleep' we are studying, but rather a collage of different, very closely tied concepts that, when pieced together, allow us to better explain what makes for better or worse sleep. As stated by Glenn Begley and John Ioannidis (2015) when speaking of the research process in general, "Our knowledge is fragmentary. Many, perhaps most of our models are naive and our constructs are often rough approximations of the truth."[152]

Core to ameliorating some of the difficulties that arise from this fragmentary nature of the research process is the obligation to make science as reproducible and replicable as possible. That

is, for research to be trustworthy, have a recognizable impact on human health, and make an effective use of funding, it is important to ensure the research is done as carefully and meticulously as possible with proper justification from prior research. This has led to the form of a general research article; scientists clearly state the methods taken (everything from foundational ideas and data collection to statistical analysis and interpretation) to reproduce or replicate the results reported. Akin to saying, 'if you look at the kaleidoscope from this specific angle, you will see what I see', attempts by other scientists to reproduce reported results will not yield exactly the same results, but what is important is that the major conclusions are reproduced.[152] In other words, science reproducibility and replicability primarily serve as methods to account for the fragmentary nature of knowledge generation, where those findings that can be reproduced or replicated are those more likely to be used as foundation for further research.[153]

This obligation towards scientific reproducibility and replicability (conducting a similar study with some differences in methodology) can lead to conflict with experimental exploration, where investigations based on less scientifically established phenomena are more likely to be groundbreaking but are simultaneously less likely to be replicated.[153] In other words, proposed research cannot be so novel such that it is completely unrelated to what is established in science, but cannot be so replicable that it effectively re-treads established theories.[153] The practice of research necessarily involves grappling with this inherent conflict; due to the sheer vastness of what is unknown, decisions regarding what information is important to the research at hand and what information is extraneous must be made to limit the complexity of the topic. Otherwise, if the topic studied involves highly complex associations and interactions, it runs the risk of being largely non-replicable and, thereby, less scientifically trustworthy than those smaller yet more replicable studies.[153]

It is this action — the act of deciding what information is or is not important to the present research topic — that has important effects for the shape of the research project. For instance, when studying sleep quality, how many factors should be studied at once? If the research project is focused on trauma, should it also account for all the other factors that are known to play a role in trauma, including demographic or genetic influences? Should knowledge of sleep gathered in one social context be applied to a different social context, or should they be treated as separate phenomena? And, perhaps most importantly, should all such decisions made regarding what information is deemed relevant (i.e., the underlying conceptual framework) be stated as explicitly as possible? Again, ensuring trustworthy knowledge is the goal of reproducible science, and this goal is primarily fulfilled by making scientific method utilized in the study as transparent and accessible as possible. That said, even with the most diligent precautions taken to ensure transparency of the entire research process, the reality is that methodology cannot be applied exactly the same as before, especially given that each researcher necessarily brings parts of themself into the research process.[154] As a result, science necessarily cannot be fully transparent as subjectivity (i.e., the impact of one's beliefs, feelings, and experiences) may be "an inevitable part of the research process".[154] One's standpoint (including their social identity, background, and personal beliefs) necessarily plays a role in research[155] — whether or not one is aware.

In more qualitative methods-based research fields (e.g., social sciences), the act of attempting to account for researcher subjectivity during the research process is known as 'reflexivity', wherein the researcher recognizes their own social background and prior assumptions and acknowledges the ways in which these subjectivities influence the research findings.[154] While evidently more pertinent for quantitative data (e.g., when performing interviews regarding sleep quality, an interviewer who has experienced poor sleep may elicit differing responses from the

interviewee compared to one who has not), the goals of reflexivity may similarly hold utility in improving the downstream efficacy of quantitative research. Given that the goal of reflexivity is for the researcher to demonstrate self-awareness regarding the ways in which their subjective beliefs and experiences shape their research, such self-awareness may aid in furthering research transparency and, resulting, research trustworthiness.

Research transparency and trustworthiness hold important positions in regards to socioeconomic and cultural factors that contribute to health inequities[156] — including the quality and experience of sleep.[157] Lack of transparency regarding subjectivities influencing decisions made throughout the research project may, inadvertently or otherwise, serve to mask throughlines between pre-existing biases and reproduction of these biases through a structured system of research. Correspondingly, mistrust in medical research — an institution founded on a history of supporting or even strengthening discriminatory beliefs and actions — continues to harm the same populations the longer the underlying beliefs and histories remain unaddressed.[158] Thus, while not a total salve for systemic bias, what may serve to alleviate some of these harms is 1) greater transparency regarding judgments made at each step of research, and 2) greater recognition and acknowledgement of one's subjectivity and how it relates to the research topic.

In this chapter, I will focus on the motivations behind incorporating reflexivity to sleep research processes, what this may look like, and the ideal outcomes. To do this, I will examine how quantitative research is used to recreate pre-existing, historical sociocultural disparities that inevitably harm already underrepresented populations. By relying on the analytical framework of intersectionality,[159] I hope to make clear the ways in which pre-existing sociocultural disparities may find their way through implicit beliefs (either from the researcher or the institution of research itself) to ensuing research into sleep and touching various steps of the process — from data

planning collection to clinical studies and beyond. I argue that current sleep research and clinical guidelines fail to adequately acknowledge the ways in which sleep disparities arise throughout the scientific and clinical process, and that there are ways to address and possibly ameliorate these harms at each step.

I've structured this chapter to capture the general steps of the scientific method: In section 1, I will define the scope of this chapter, providing examples of factors which contribute to sleep disparities (focusing on sex, race, and socioeconomic status) and highlight the ways in which these factors may interact with and compound upon one another to promote unique effects on sleep efficacy. In section 2, I will examine the structure of sleep research and, examine the steps taken by researchers to try to account for these disparities, and the ways that, despite these attempts, pre-existing disparities still find their way through to the research project. In section 3, I will delve deeper into the process of experimental data analysis, using the lens of artificial intelligence (AI) methods to further highlight the ways these previously identified disparities arise in data analysis and interpretation. In section 4, I will examine the output of this research in relation to its intended goals — not only where those goals were met, but also where & how those goals were not met — as well as discussing potential methods to account for pre-existing disparities at each step of the scientific process and how the concept of reflexivity may provide some aid in accounting for these disparities.

The goal of this chapter is to provide a framework to think about justice and equity in the process of researching sleep and sleep disorders, as well as the role of the researcher's standpoint in the entire process. I will be focusing on sleep because of the core role it plays in general wellbeing. Given how remarkably adaptive sleep is, I believe the study sleep deserves to be conducted in a way that truly respects this adaptivity.

IV.i FACTORS OF DISPARITIES IN HEALTH AND SLEEP.

The term 'health' broadly encompasses physical, mental, and interpersonal well-being. The most prominent survey for health status, the 36-item short-form (SF-36), captures health in the components of physical health (including physical function, bodily pain, and general physical health) and mental health (including feeling energized, social functioning, emotional functioning and general mental health).[160] As discussed in previous chapters, sleep quality is a major determinant of health and, as such, is necessarily a core component of health; poor health tends to accompany poor sleep quality.[161]

Sources that aid in creating disparities in health and sleep come from variety of environmental, psycho/physiological, and, importantly, sociocultural factors.[161,162] To illustrate how sociocultural factors impact health and sleep, I will examine how biases and discriminatory beliefs based on biological sex/gender identity and racial/ethnic identity have had a recognizable effect both in terms of general health as well as specifically for sleep. While I will initially discuss these as though they are entirely separate factors, I will discuss how these two factors — as well as other sources of health disparities — interact to generate unique circumstances that influence sleep.

IV.i.1 *Biological sex and gender identity.* *note 11*

Biological sex and gender identity naturally play a fundamental role in all of life — so much so that, in 1993, the National Institutes of Health (NIH)'s Revitalization Act mandated that

[11] Biological sex refers to a construct based on genetics and anatomy, while gender identity refers to one's internal sense of their gender often constructed in relation to societal gender roles. In most cases these two are congruent (i.e., those assigned male at birth tend to identify as men; those assigned female at birth tend to identify as women), but there are important nuances that I am unable to fully explore here. While it may appear that I generally conflate the two in this section, I wanted to make clear that these are two separate constructs — not synonyms for the same concept.

women be adequately represented in clinical trials.[163] Later, in 2015, the NIH published policy guidelines necessitating that all NIH-funded projects examining vertebrate animals or humans must include data from both male and female subjects unless otherwise justified.[164] Looking at how the historical exclusion of women from the clinical research trials has, at best, inadvertently led to disproportionate harms,[165] this decision makes complete sense: absence of disease-related data in women has led to disparities in recognizing symptoms of a problem, making an accurate diagnosis, and, consequently, giving relevant and effective solutions to help address the disease.[166]

This problem can be most clearly exemplified by the sex-based disparities in cardiovascular disease (CVD): despite being the cause of death in 35% of women, the exclusion of women from CVD clinical trials has led to the perception of CVD being a 'man's disease'. As a result, when seeking treatment for CVD, physicians often assume that the primary symptoms highlighted by the majority of prior research — most prominently, chest pain — are uniform for both men and women.[166] For instance, when being evaluated for a potential heart attack, a majority of men and women present with chest pain (89.5% and 87.0%, respectively).[167] However, women are more likely to present with *additional* symptoms, potentially contributing to both the patient and the physician being more likely to misinterpret the cause of the symptoms in women compared to men.[166,167] This all culminates in women being less likely to be accurately diagnosed with CVD, subsequently leading to poorer treatment and worse clinical outcomes.[166,167] What I hope is clear from this example is that, due to the relative dearth of clinical research on CVD in women, the historical bias of 'male being the norm population'[165] and the subsequent harms caused by this bias are effectively recreated in the context of clinical research and treatment of CVD in women.

Sex-related disparities in sleep have been found, with women generally sleeping longer and having greater sleep quality compared to men across the lifespan.[168] These findings are

primarily based on objective measurements using polysomnography (PSG), with women more frequently exhibiting slow wave sleep (SWS; indicative of deeper sleep) as well as less total wake time during a night of sleep.[168] Paradoxically, women are more likely to report sleep disruptions (such as insomnia) across the lifespan.[168] As highlighted in a review by Mong and Cusmano (2016), two factors potentially explaining this apparent paradox are 1) sex-based differences in circadian rhythms, which are daily rhythms that heavily regulate sleep timing, and 2) the impact of sex hormones on sleep.[168]

In both humans and animal models, sex-based differences manifest in multiple facets of circadian rhythms, such as differences in the shape of the suprachiasmatic nucleus (SCN), the 'central pacemaker' which regulates circadian rhythms throughout the body; differences in neural pathways to and from the SCN; and differences in expression of circadian rhythms in clocks located throughout the body as well as behavioral expression of circadian rhythms.[169] Furthermore, while it is unclear if sex hormones (particularly testosterone) has any effect on sleep in men, estrogen and progestin have been shown to affect sleep quality and timing in women, with frequency of sleep complaints generally accompanying fluctuations in presence of hormones in the body.[168] Finally, unlike men, women naturally undergo substantial changes hormone production during their midlife — a process known as menopause, which is associated with increased sleep disruptions[170] as well as greater risk of CVD.[171]

Despite these findings, and just like CVD research, women have been historically underrepresented in studies of sleep, and only recently are these disparities being properly addressed. For instance, in 2013, the U.S. Food and Drug Administration issued the first sex-specific guideline for treatment of any sleep disorder.[172] It is difficult to determine the extent to which such historical underrepresentation in research has impacted the health of women

experiencing CVD or sleep disruptions. Despite the potential role of hormonal fluctuations in influencing sleep quality, scientific research into menopause primarily began in earnest during the mid-1990s[173] — and even then, much of this work excluded minoritized racial/ethnic populations.[171] What is clear is that historical biases against women, and the idea of 'male is the norm population', has had rippling consequences on how sleep research is done, how its findings are applied, and what sex stands to benefit more from the research.

IV.i.2 *Racial/ethnic identity.* note 12

Another major source of sleep disparities is racial/ethnic identity. Like biological sex/gender identity, racial/ethnic identity has wide-ranging effects on many aspects of health.[175] Alongside mandating increased inclusion of women in clinical trials, NIH's Revitalization Act of 1993 also mandated increased inclusion of those from underrepresented racial/ethnic backgrounds. This was in part an attempt to account for the historical underrepresentation in research, caused by years of discriminatory beliefs, practices, and structures that not only prevented participation of underrepresented individuals in research, but actively supported prejudiced beliefs regarding these identities.[158]

One example of such discriminatory beliefs being actively supported by research institutions is the 19th century practice of craniometry, where the size of the cranium was purported to be correlated with intelligence and was used to argue that Black people were lesser humans than White people.[176] Another salient example of discriminatory practices causing lasting harm to an

[12] While some may argue the presence of a biological/genetic basis of race, I would argue that it is primarily conceived as a social construct founded on historical sociopolitical processes that, while still being a construct, has and continues to affect lives of countless individuals.[174] Ethnicity, on the other hand, is primarily founded on *sociocultural* ancestry, with shared beliefs and norms being the core basis.[174] Lee (2009) captures the nuanced differences between the two that is often missed when discussed in the context of biomedical research, especially when both used as a proxy for some underlying biological difference.[174] While I will generally use the two terms interchangeably, nuances in the meaning and conceptualization of 'race' or 'ethnicity' are complex — far too complex to fully capture here.

underrepresented population is the 40-year Tuskegee syphilis experiment (from 1932 to 1972), where 399 Black men were deliberately and covertly denied treatment for their syphilis diagnosis.[177] Even now, the legacy of the Tuskegee experiment has had lasting effects on interactions between Black individuals and research institutions, with individuals from these populations expression greater doubt regarding whether they would truly stand to benefit from the results of health research.[158] To these populations, the research institutions are seen as active agents in continuing to re-create discriminatory beliefs through research.

Once again, similar to women, racial/ethnic identity-related disparities in sleep have been found, but, due to similarly poor inclusion of underrepresented populations in sleep studies, evidence of these differences are limited — especially regarding sleep disparities for Hispanic, Asian, and other racial-ethnic groups.[175] That said, underrepresented populations (Black, Hispanic, and Asian) tended to sleep less than White populations[59,178] and were more likely to present with sleep disturbances,[179] with Black populations particularly having poorer sleep efficiency and tending to remain in lighter stages of sleep.[175] As reported by Giddens et al. (2022), when examining specific income groups (low, middle, or high), White children generally have higher total sleep time (TST) compared to children of the same income group who identified as Asian, Black, Hispanic, or 'Other'.[178] Furthermore, in some cases, TST in *low*-income White children surpassed that of *high*-income children from other racial/ethnic backgrounds.[178] While there is some evidence of racial-ethnic differences in timing of circadian rhythms,[180] the larger explaining factor for these disparities in sleep duration is likely due to differing exposure to socioenvironmental stressors.

As Keisha Ray examines in her book Black Health (2023), Black people are more likely to live in areas with greater light and noise pollutants, are more likely to experience hypervigilance

during the night, and are more likely to experience effects of environmental racism that negatively impact health and sleep.[59] For many Black people, conditions like these are a fact of life, and their sleep has seemingly adapted to these difficulties; despite sleeping less (as determined by objective measurements of sleep), both Black and Asian individuals are less likely to report sleep complaints compared to White individuals.[175] At the same time, when studying sleep, such socioenvironmental barriers to sleep are not entirely captured, with many studies often studying the dynamics of sleep under ideal circumstances — free from the influence of noise pollution or emotional strain of hypervigilance. As a result, advice for good sleep taking the form as 'sleep hygiene', too, assumes the individual has the ideal conditions for the recommendations to have the greatest impact on sleep quality. As stated by Ray (2023), sleep hygiene is ultimately inconsequential advice for underrepresented communities, as it "does not account for people's lack of access to what they need for sleep health".[59]

Embedded in these disparities is a cultural element of sleep recommendations that can be clearly seen when we examine recommendations for infant sleep. For instance, the act of 'co-sleeping' (wherein a parent and their child shares the same bed)[181] is often opposed in many Western, typically American parenting advice books.[182] Rather than sharing a bed, American parenting advice tends to advocate for 'sleep training' the child by letting them sleep independently and, when they vie for attention, letting them 'cry-it-out'.[182] This is despite the historical importance of co-sleeping in human existence and culture,[181] partly due to role of physical touch during sleep in an infant's development.[183] This discrepancy is best captured by Elaine Barry, when she writes, "Co-sleeping is a complex familial phenomenon that has yet to be well understood by Western scientists."[183] As she later notes, co-sleeping is more common in racial/ethnic minoritized groups, as well as in families with lower socioeconomic status. Like those

represented by the term sleep hygiene, what I hope is clear is how these recommendations are primarily coming from unstated goals from a particular perspective — that is, a middle- to upper-class White Western perspective — that are then overextended to circumstances outside of its reach. Despite these recommendations being clear benefits of sleep research, it is clear those who stand to gain are those who are most represented in research, while underrepresented populations may similarly try and likely fail in the attempt or, even if they achieve the sleep hygiene goals, they may see less benefit; these findings aren't *for* them, and attempts to benefit from specific findings may, ultimately, do more harm than good.

Other sources of sleep disparities and intersectionality.

Although touched on in the two preceding sub-sections, there are many other sources of disparities for sleep. Two such examples are 1) socioenvironmental factors, capturing the effects of mental and emotional wellbeing[184]; and 2) socioeconomic status (SES), capturing the effects of one's social position and degree of access to resources.[185] As suggested when discussing hypervigilance and sleep disturbances, perceptions of safety in the current environment have influences on emotional wellbeing and sleep quality.[184] Living in a community that feels safe, or even sleeping near someone you trust, can help improve sleep quality.[19,184] Similarly, being a major determinant of health, lower SES has been associated with poorer sleep quality based on objective measurements.[186] Specifically, lower SES is associated with less overall sleep, greater arousal during sleep, and more variability in sleep timing.[186] These two examples highlight the role of social context in determining sleep quality, among many other factors.

Across all sources of disparities mentioned so far, I want to be clear that all of these are not factors we can study in a vacuum. Instead, they all interact with and compound upon one another, necessitating a more holistic approach to understand how these different factors interact to generate

unique conditions for sleep disruptions. To make these interactions more explicit, I will be using the analytical framework of *intersectionality*, introduced by black feminism theorist Kimberlé Crenshaw. Intersectionality is a framework used to analyze how an individual's different sociopolitical identities — including biological sex and gender identity, racial and ethnic identity, SES, and more — interact to produce unique conditions of discrimination and privilege.[187]

In the context of biomedical research, the historically dominant perspective (and thus primary target for much of the findings of biomedical research) has been middle- to upper-class white men.[187] Even when considering potentially implicit identities that confer additional privileges (e.g., able-bodied, neurotypical, living in more urban contexts), this perspective already posits a population situated in the intersection of sex, race, and socioeconomic position that stands to benefit most from biomedical research.[187] As previously discussed, any 'deviation' in any one of these three axes — being female, being part of an underrepresented population, or having lower SES — is enough to impact health.

However, when viewing biomedical research from the perspective of intersectionality, while deviation on any one axis can lead to disadvantages, deviation on *multiple* axes can lead to unique advantages or disadvantages. For instance, when considering sex-based discrepancies in CVD, socioeconomic and sociocultural factors play a stronger role in the development of CVD in women compared to men.[166] Similarly, not only do sex and race both effect sleep independently, but they also interact to produce even harsher effects on sleep such that, when examining average sleep duration in Black populations, Black men tend to sleep less compared to Black women.[59] Finally, despite experiencing previously discussed disadvantages due to the 'male as the norm population' bias in biomedical research, middle- to upper-class White women still draw greater benefit from CVD and sleep research compared to Black women.[166,172]

This leads to a core question: how should we, as sleep researchers, properly account for such disparities that arise through interactions of various identities underrepresented in research — let alone from a single underrepresented identity? Properly answering this question warrants a deeper interrogation of the structure of sleep research, focusing on how the implicit yet dominant perspective in sleep research can lead to potentially harmful assumptions for each step of the research process.

IV.ii STEPS OF RESEARCH AND BIAS.

Regardless of the finer details of the exact methods described in a research project, the overall process generally follows the eight steps: problem statement, research question, conceptual approach, research design, subject inclusion, data collection, data analysis, and conclusions.[188] As Rogers & Kelly (2011) argue, biases like the ones described in the previous section may manifest at any step in this process.[187] Furthermore, because biomedical research studies generally do not directly address relationships between ethical issues of social justice and factors leading to poor health outcomes (particularly in context of intersectionality),[187] each of these steps have unique susceptibilities for the potential of biases to be effectively recreated.

In this section, I will focus on the first six steps, examining the last two steps of data analysis and conclusions in further detail in the following section. I will describe the goals and intents of each of the six steps, safeguards used to ensure each step is as scientifically robust as possible, and the ways in which unconscious biases — primarily those originating from an implicit perspective laden in biomedical research — may nevertheless appear despite such efforts. From there, I will give examples of how such biases may manifest in sleep research using 1) humans and 2) animal models. The goal of this section is to fully explicate the problem of bias in biomedical research, describing the ways in which biases may be recreated anywhere in the

research process. The examples I will serve to clarify how such biases may present in the designs of human and animal research.

IV.ii.1 *Steps of biomedical research — from research purpose to data collection.*

As previously described, the research process can be summarized into eight steps: problem statement, research question, conceptual approach, research design, subject inclusion, data collection, data analysis, and conclusions.[188] What I will show is that, for each of these steps, there are crucial decisions made regarding what information is relevant to the research project, and what information is extraneous. It is in these decisions at each step that the researcher's subjectivity — their social background, experiences, and beliefs[154] — may (unintentionally) influence what is decided. To illustrate this, I will focus on the first six of these steps (with the last two being the focus in the following section):

1. *Problem statement and research question* — Addressing the gap in knowledge outlined by the problem statement is the fundamental goal of a research project. The problem statement defines the primary topic of interest for the research project, thereby serving as guidance for all following steps in the process.[188] Conversely, where the problem statement defines the topic of interest, a research question defines the purpose and focuses of the study. This research step fully articulates the question the research project will seek to answer, thereby (ideally) addressing the gap in knowledge highlighted by in the problem statement.[188] Given these definitions, one can quickly identify how a researcher's subjectivity may play a direct role in this step. Research is sometimes (derogatorily) referred to say 'me-search' due to the ways in which the researchers' subjectivity sometimes drive the subject matter they research.[189,190] It is often difficult to fully dissociate a researcher's scientific interest from their personal interests

— especially for those whose scientific interests are spurred by a related personal experience.[189] Put simply, researchers' subjectivities (and any biases that accompany them) cannot easily be extricated from the problems they focus on in their research, and this necessarily extends to decisions regarding research problem and question.

2. *Conceptual approach and research design* — These two steps concern the planning and design of the study; the former specifies what factors are relevant to capture, what associations are anticipated, and effectively guides interpretation of the data acquired.[188] Conversely, the latter specifies how the study will be conducted, describing what participants will take part in the study and what methods will used.[188] These two steps heavily rely on findings from prior research, using potential theories and explanations described in other studies to help refine the current study (e.g., helping determine which variables are most likely to be relevant, what methodology would be most appropriate for the research topic).[188] Another goal of using prior research as foundation is to ensure robust intellectual thought; by justifying the study by compiling evidence in the form of (relevant) prior studies, not only does this give justification for the current study, but peers can also critique the reasoning in order to improve the study's conceptual framework. However, unlike the previous two steps where the source of bias was more likely to originate from the researchers' own subjectivities, here biases may arise due to historical-rooted discriminations and inequities in biomedical sciences. As mentioned above, the historically dominant perspective in health research is that of a middle- to upper-class white man[187] and, as such, the bulk of biomedical research implicitly holds this assumption. In this case, decisions of what prior studies are relevant to the current study will primarily draw from those studies with these biases. Thus, if left unaddressed and unaccounted for, follow-up

studies that use these studies as conceptual foundation are liable to continue recreate this perspective, inadvertently allowing these biases to continue.

3. *Subject inclusion and data collection* — Where the previous two steps concerned study planning, these steps concern study execution. For instance, in human subjects research, subject inclusion is based on a specific target population (relevant to the study), inclusion/exclusion criteria, and selection procedure.[188] Once a part of the study, subjects will undergo methods outlined by the research design step and all data collected will be stored for analysis and interpretation.[188] Hence, these steps are intended to perform the study and collect data as outlined in the previous steps. Given that this occurs after rigorous literature review and rounds of peer review, potential problems of biases or flawed reasoning are intended to be ironed out by the time the study is finally executed and data is collected. Despite this, and as discussed before, researchers necessarily bring a part of themselves into the research process,[154] thereby leading to some differences between conceptual framework of the study and the actual execution of the study. Furthermore, while generally using prior literature as justification (which alone, as discussed above, runs the risk of recreating biases), inclusion/exclusion criteria are often insensitive to disparities in certain populations caused by sociopolitical problems. For instance, when selecting studies for meta-analyses, one of the most common exclusion criteria is the study being reporting in a language other than English[191] — a language that, despite being the native language of 7.3% of the world's population and spoken by less than 20%, is the *de facto lingua franca* in the world of the science and academia.[192] In addition, according to a report examining 74 countries, children who attended higher SES schools were exposed to more foreign-language classes than those in lower SES schools,[193] highlighting the relationship between foreign language learning and socioeconomic

privilege. This all suggests that the decisions regarding what subjects are included or excluded from the study, while relying on prior research, may recreate the same perspectives implicit in biomedical research.

What I hope is clear is that, just as highlighted by Rogers & Kelly (2011), ensuring biomedical research adheres to principles of equity and justice requires researchers to take a more informed approach throughout their research process.[187] The above represent examples of some of the various ways in which the researcher's subjectivity or biased perspective in research literature may continue to be recreated in new research projects. That said, so far I have examined this appearance of bias in research in a more general sense. To further clarify this point, I will now turn to examples of data collection in human subjects and animal models.

IV.ii.2 *Human studies.*

As described in Chapter I, one of the most important inventions in the history of human sleep research is the PSG — a technique that combined data acquired via electroencephalography (EEG; capturing neural activity from the scalp), electromyography (EMG; capturing muscle activity) and electrooculography (ECG; capturing eye activity) to examine neural processes occurring in an unconscious individual.[194] This method is often done during in-lab sleep studies, where the individual comes into a laboratory, is connected to all of the data collection modalities, and sleeps in the lab overnight while data is collected (with in-home sleep studies being a viable alternative that is more comfortable for the participants yet less controlled).[194] As the gold-standard for human sleep research, PSG is capable of diagnosing a variety of sleep disorders (including excessive daytime sleepiness, sleep apnea, narcolepsy, sleep movement disorders, parasomnias, and even nocturnal seizures).[194] As noted by Markun and Sampat (2020), a notable limitation of

PSG is what is known as the ‘first night’ effect, where individuals tend to have worse sleep when sleeping in a unfamiliar environment (e.g., a laboratory).[194]

This leads to two underrecognized but pertinent matters related to justice and PSG-based sleep studies. The first has to do with the interaction between the first night effect and how, for example, Black individuals experience sleep and how they regard research institutions. As stated before, compared to White individuals, Black individuals are more likely to experience hypervigilance during the night,[175] and are less trustful of biomedical institutions.[158] This likely culminates in the finding that Black individuals tend to have lighter sleep for in-lab sleep studies compared to in-home (with the opposite being true for White individuals), possibly reflecting the greater degree of discomfort experienced by Black people who take part in sleep studies.[195] The second matter is the head cap used to capture EEG, which was not designed with afro-textured hair in mind.[74] Because dark, afro-textured hair is not a common hair type of the populations in mind when the cap was designed, data collected from those with such hair types are more likely to be of poorer quality.

These two issues emphasize how prior research may inadvertently be used to harm underrepresented populations. In the first case, based on much of the prior research, one could reasonably conclude that in-lab PSG would be the ideal method for all participants, not attending to the differing social contexts between Black and White individuals that may impact their sleep. In the second case, poor data quality means these data are more likely to be discarded — leading to continued underrepresentation, this time due to (likely unintentionally) discriminatory hardware.[196]

IV.ii.3 *Animal models.*

Mice (specifically the C57BL/6 strain) are one the most frequently used mammals for modeling human sleep disorders.[77] Given the comparative advantages of utilizing mice, it's clear to see why: mice have a much shorter lifespan and gestation period (about 3 weeks, compared to about 9 months in humans), meaning hereditary disease can be examined relatively quickly; the mouse genome has been fully sequenced, allowing for a variety of genetic tools for developing models for various human disorders; additionally, mice share about 95% of their genes with humans; finally, mice generally take up less space due to their small size, making relatively inexpensive to maintain colonies.[197] On top of all this, not only is the function and importance of sleep generally conserved in mice,[70] but certain sleep-related neural features have analogies in mice.[198] These facts make mice an ideal model for studying circadian rhythms and sleep disorders. To better discuss this use of mice in studying sleep, I will be referring to the data I presented in Chapter II regarding contextual fear and the re-emergence of abnormal circadian behavior. However, I will first discuss the value of using mice as a model for sex differences in human sleep.

Mice are a common model for studying potential sex differences in various physiological functions[199] — including sleep.[200] Similar to women, changes in circadian rhythms, sleep, and sleep architecture (i.e., the temporal organization of each sleep stage) are associated with fluctuations in sex hormones in female mice.[200,201] Contrary to women, however, female mice tend to spend less time asleep and have less sleep fragmentation compared to male mice.[200] Despite these relatively mixed comparative results, it is clear that the relationship between sex (including the role of hormones) and sleep is increasingly being seen as a vital aspect that may aid in translating findings from animal models to humans.[202] In a national survey of 1234 American biomedical researchers, 47.51% of respondents agreed that sex as a biological variable will

improve translation to humans (with 34.77% being unsure and 17.72% disagreeing),[203] further emphasizing this increased awareness of including sex as a biological variable.

In my chapter on cyclic contextual fear, one of the sections included was a discussion on the apparent sex differences in behavior. Based on what I knew in humans (i.e., women are more likely to develop PTSD[204]), I initially anticipated female mice to exhibit a greater response in the contextual recall phase (where mice were returned to the context in which they experienced fear). However, as discussed in the chapter, male mice instead exhibited a greater change in behavior in response to the context, suggesting interesting sex disparities in development of PTSD-like symptomology. Despite this, and keeping in theme, few studies on fear learning and PTSD utilize female mice in addition to male mice.[205] This underrepresentation mirrors broader challenges seen in human studies data acquisition has been historically centered on white males, not only perpetuating biases but, in the process, impeding progress examining how biological sex impacts development of PTSD.

While biological sex plays an important role in the development of PTSD in humans, other factors, such as being a member of an underrepresented population[206] or having lower SES,[207] also increase one's susceptibility to developing PTSD. Unsurprisingly, however, there is no mouse model equivalent for racial/ethnic identity or SES (as far as I can tell) note 13, meaning the concept of intersectionality generally cannot be applied in the context animal model research. While there

[13] One could imagine how create a mouse model of racial trauma, but an immediate concession to make is that, like all models, there will inevitably be some conceptual facets that will be lost in translation. That said, there are some procedures that may approximate aspects of racial trauma using mouse models. For instance, social defeat stress (SDS) is a social/psychological stress procedure commonly used as a model for depression and post-traumatic stress disorder (PTSD)[208] — disorders which are more prevalent in Black populations.[209] For a given experimental mouse, SDS can take the form of the mouse being exposed to and attack by a dominant mouse; the mouse witnessing another mouse undergoing SDS; or the mouse being exposed to the sensory information of another mouse, and potentially being attacked.[208] While each of these variants may approximate different experiences of racial trauma, I would posit that a combination of all of these variants of SDS, in addition to a salient differentiating characteristic of the dominant, attacking mice (e.g., white fur), might make it a more exacting model of racial trauma.

are very clear benefits of using animal models, one of the inevitable realities is that any findings in animal models are necessarily devoid of socioenvironmental factors that strongly impact populations generally underrepresented in health-related research. Consequently, this parallel issue of underrepresentation in biomedical data, as seen in human studies, risks shaping conceptual models of health that neglect the impacts of societal influences on health, thereby perpetuating a biased perspective. These biases can extend even to *computational models*, further exacerbating the manifestation of biases within sleep research.

IV.iii Quantitative data analysis and algorithmic bias.

While more commonly associated with qualitative methods, the impacts of subjectivity can similarly arise in quantitative methodology — where the goal is to be as objective as possible, the presence of subjectivity can be seen as evidence of 'bad science'.[190] However, as previously discussed, some decisions necessarily rely on the researcher's beliefs. For instance, data collection and cleaning rely on the summarizing the concept to some degree to better capture the core phenomena of interest (e.g., sleep quality can be reduced to a single numerical value, to the potential detriment of some of its nuances).[210] While this act of data summarization is necessary, it simultaneously relies on some pre-existing structure of thought to determine, for example, what information is relevant and important to collect as data or what methods are best suited to follow for analyzing the data. Be it unintentional or otherwise, decisions made in the process of analyzing quantitative data can be motivated by some subjectivity.

In the context of methods such as artificial intelligence (AI), however, the focus shifts from the subjectivity of the researcher to the implicit perspectives laden in the data being analyzed. A clear example of this is Amazon's attempt to use AI to review job applications and find a means of recruiting the best applicants. If done by a human, for example, certain unconscious biases may

lead them to rejecting applicants based on irrelevant information (e.g., those with names commonly viewed as Black are less likely to be hired compared to those with names commonly viewed as White).[211] Hence, the intent of Amazon's AI was clear; bypass human unconscious bias and create an objective method of selecting candidates. However, this is not what happened; instead, the AI recruiting tool learned to value applications from men, devaluing any applications it recognized as coming from a women.[212] Importantly, the engineers behind the algorithm attempted to prevent such biases from appearing in the model, but to no avail. This is likely due to the way AI algorithms are trained; in this case, it used pre-existing data on applicants that were hired and those that were not and found associations within these data. In other words, the AI did not invent its sexist decision making process; it *learned* it from the patterns that already existed in the hiring data they had. The data had an implicit perspective, based on prior work, that was recapitulated using AI.[213]

To analyze how this likely occurred and what it might mean for biomedical uses of AI — particularly for studying sleep — I will discuss the issue of algorithmic bias, focusing on three components: representation, implementation, and transparency. From there, I will shift to examining the use of computational models to analyze different sleep-related concepts (i.e., circadian rhythms, neural signals, and trauma). To be clear, there are numerous tools one can use to analyze data; from the relatively simple linear regression (attempting to fit a line between two independent variables) to the relatively complex neural network (a type of machine learning algorithm inspired by the human brain, using connections between nodes to perform complex, non-linear transformations). I will be focusing on the use of neural networks for clarity, but the ideas discussed will still apply to less complex algorithms (albeit likely to a lesser degree compared to neural network-based models).

IV.iii.1 *Algorithmic bias.*

I first want to define what algorithmic bias entails, as there are important differences between algorithmic bias and the social psychology concept of bias. Throughout this piece, I have been using the term 'bias' to refer to evaluations where one group is deemed better in some metric compared to another.[214] While this often takes the form of unconscious/implicit associations between the group and the metric (e.g., one may unconsciously yet reflexively associate 'research' with 'male'), it should go without saying that one can be conscious of their bias.[214] Conversely, algorithmic bias refers to a circumstance where the outputs of a computational model consistently benefit certain populations more than others.[213]

The two terms clearly show some similarities, but there are two important nuances to note: First, in the case of algorithmic bias, not only is bias integral to the algorithm's decision making, but also attempts to correct these biases (as seen in the Amazon example) will likely not work because of how intertwined the bias is to the algorithm.[213] Secondly, given that algorithms cannot be conscious, algorithmic biases cannot be either implicit or explicit. Both points underscore a recommendation that researchers demonstrate awareness of how pre-existing biases may reemerge in the computational model. To best illustrate how these biases may be recreated in a model, I will be using an example of generating a neural network model capable of autonomously scoring sleep stage (akin to the website presented in Chapter III) based on EEG data captured in humans. I have roughly summarized the data analysis and interpretation process into three steps: data collection, model fitting, and model implementation. While I will focus on one element for each of these steps (representation, transparency, and interpretation, respectively), there are undoubtedly more elements to consider.

1. *Data collection and representation* — As described in the previous section, there are a myriad of factors leading to imbalanced representation, with those historically underrepresented in biomedical science research (e.g., Black individuals) continuing to be underrepresented in present studies on sleep.[215] Added to this are the limitations of hardware leading to poorer quality in underrepresented populations[196] and the apparent influences of race/social context potentially generated differences in sleep quality when undergoing sleep studies.[195] Even when considering just these three factors — historical underrepresentation in data, relatively poorer data quality, and data collection inadvertently capturing influences from social context — its clear that the data being used to create the computational model are already showing the results of disparities laden in sleep research. If we attempted to create a sleep-scoring algorithm based on these data, it would likely work better for populations properly represented in the data, but worse for populations whose representation in the data is limited by various factors. An often used term in computer science is 'GIGO': garbage in, garbage out. Just like the sex-biased recruitment algorithm developed by Amazon, the goal of computational models is to find associations in the data. If the data going into the model is biased, the model will likely learn, recreate, and output the same biases.
2. *Model fitting and transparency* — If biases in the data go undetected, model fitting is where the algorithm will begin to learn these biases. For AI methods such as neural networks, 'learning' is really a process of finding the simplest yet most robust statistical relationships between the input data and the output of interest.[71] It is this step where the use of more complex AI models run the risk of making the process of identifying and accounting for bias more difficult. This is because, in simpler models (e.g., linear regression) the step of feature extraction (where different variables are extracted from the data) is done manually, meaning

the researchers have to define what variables they want to extract from the data.[82] In more complex models like neural networks, this step is done autonomously, meaning it is difficult to determine what features of the data are being used by the algorithm in its decision making processes — effectively behaving akin to a 'black box'.[71,213] Because of the difficulty of peering into a black box algorithm to determine its inner workings, transparency of the algorithm becomes a bigger issue, again, particularly with more complex models. If the data used to train the model is biased, it will be harder to probe the model to determine where, why, and how the pre-existing biases may be recreated by the model.

3. *Model implementation and interpretation* — The previous two (general) steps culminate in the implementation of a model, in this case applying the algorithm to score sleep in humans. In typical use cases of neural network algorithms, if there were any biases in the data that were then learned during model fitting, this would be when these algorithmic biases will be most evident (e.g., Amazon's recruitment algorithm's bias against women applicants,[212,213] Apple's algorithm giving lower credit limits to women compared to their spouses[213]). This is because, in cases like these, the discrepancy between what would be expected (e.g., a mix of man and woman candidates, equal credit limits for both genders) and what is observed (women candidates are devalued, women receive lower credit limits) suggests something is wrong with the algorithm. This is the case because the output *can* be observed. In the case of sleep scoring, these errors may not be as evident because of the use of EEG data — abstract numbers that do not readily hold much meaning to humans. While bias can be probed based on the output of the algorithm (e.g., comparing manual scores to the algorithms scores and to determine how accurate the algorithm is), it doesn't change the fact that the algorithm is still likely biased, and as such will still benefit one population over another.

Again, despite being a limited discussion on algorithmic bias, what I hope to have illustrated in these examples is how quantitative/computation tools may, similarly, recreate pre-existing biases and disparities that will go on to recreate systems of advantage and disadvantage based on information regarding social context. This leads to the question: how might we consciously quantify issues of justice and intersectionality for use in our models, rather than letting the algorithm learn and recreate biases during the model fitting process?

IV.iv ACCOUNTING FOR SYSTEMIC BIAS.

As I have hopefully shown thus far, historical biases in health research not only have had lasting disadvantages on populations that deviate from the dominant perspective (i.e., that of a middle- to upper-class white man), but these biases continue to be effectively recreated in current studies — leading further harms to these underrepresented populations. Drawing from the framework of intersectionality, deviation in any one identity axis leads to disadvantages, while further deviations on multiple axes interact to create unique positions of disadvantages that manifest in health disparities.[159,187] Importantly, even without explicit acts of discrimination and exclusion are stopped, systemic bias, in place, "is often self-perpetuating, with persistently damaging effects on health even after the explicitly discriminatory measures are no longer in effect."[216] This places the burden of justice primarily in two places: the first is the institution in which the systemic bias exist, while the second is those who benefit most from the presence of the systemic bias, and thus (generally unwitting) continue to support the structures keeping the biases in place.[216]

Accounting for systemic bias in the context of sleep research would thus entail addressing and deconstructing the institutional structures that continue to support its laden systemic biases. Simultaneously, those predominantly represented in the institution of sleep research would hold

some burden of acknowledging and accounting for how their (generally unstated) beliefs and perspectives influence each step of their research — often in support of these systemic biases. In what follows, I will give an example of how one might start working towards accounting for systemic bias in these two avenues: quantifying systemic bias, and the concept of reflexivity, respectively.

IV.iv.1 *Quantifying systemic bias.*

While systemic bias has existed in various forms over the centuries, efforts of quantifying systemic bias in related to its effects on health are still sparce.[217] One might naturally ask; how might you quantify the effects of systemic bias, given that they operate in ways that are "covert, [and] are embedded in normal operations of institutions"[216]? As demonstrated by Groos et al. (2018), variables capturing different sociopolitical domains associated with systemic bias (e.g., perceived racism in social institutions, or residential neighborhood/housing) are associated with a variety of individual- and population-level health outcomes.[217] Similarly, LaFave et al. (2022) reported a method of quantifying based on capturing indicators of systemic bias across nine discrimination contexts (civics, education, employment, environment, healthcare, income/wealth, media/marketing, neighborhood factors, and policing).[218] While yet to be more deeply incorporated into biomedical research — let alone research on sleep — these two methods demonstrate quantitative methods of accounting for the influences from social contexts that are often not fully acknowledged in biomedical research.

IV.iv.2 *Disciplinary reflexivity and accounting for subjectivity.*

On the level of the individual, one idea I hope I have made clear is how obliviousness to systemic biases maintained during the research process can lead to harm for individuals from

populations underrepresented in biomedical research. Rogers & Kelly (2011) provide some points for researchers to consider during the research process to ensure motives of social justice stay consistent throughout.[187] Some of these include considering the social context of the research, determining whether an intersectional approach can be used in the research design, and ensuring data analysis and interpretation reflect the intersectionality of disadvantage where possible.[187]

These actions would pair well with what feminist philosopher Sue Wilkinson calls as 'disciplinary reflexivity".[154] Disciplinary reflexivity would involve clearly stating institutional traditions which have had an important role in shaping the overall form of the research project — with the goal of addressing and accounting for systemic biases.[154] An extreme hypothetical case of disciplinary reflexivity may take the appearance of the following: gathered data would be understandable in terms of what foundational beliefs are present for all interactions with data, including collection, analysis, and interpretation; researcher motivations, beliefs, and biases are clearly defined and stated as part of research communication; subjective beliefs used to motivate foundations and goals of the research project are made as explicit as possible. While this is, admittedly, a very unlikely form for the research process to take any time soon, what I hope to demonstrate is that there exist ways to lessen the extent to which opacity exists in each step of the research process and, in doing so, to lessen the extent to which pre-existing sociocultural disparities can propagate through sleep research unchecked.

IV.v Conclusions.

In this chapter, I examined the multifaceted landscape of sleep research, exploring factors contributing to sleep disparities as they relate to wider concepts related to societal justice and equity, and in doing so highlighting potentially challenges faced by sleep researchers in accounting

for these disparities. The overall goal of this chapter was to show how systemic biases are recreated in sleep research and to discuss potential ways to address these biases.

In section 1, I explored various factors contributing to sleep disparities, highlighting the key idea that these disparities are likely to emerge partly (and unintentionally) due to the recreation of social and systemic biases within sleep research. When conducting this analysis, I applied Kimberlé Crenshaw's concept of intersectionality to provide further clarity on how different social identities can interact to amplify sleep disparities. Using a simplified version of the research process, in section 2, I illustrated potential avenues through which systemic biases may be recreated in 1) the study's problem statement and research question; 2) conceptual approach and research design; and 3) subject inclusion and data collection. I used examples in human studies (using EEG to study sleep stages) and animal models research (modeling PTSD in mice) to show how recreation of biases may manifest in study design and data collection. This comes to a head in section 3, where I examined how biases can be perpetuated in data analysis and computational models in ways that may make them hard to detect. The three avenues I discussed were 1) poor inclusion of underrepresented populations in the training data; 2) how choice of algorithm can make it difficult to probe the algorithm's decision-making; and 3) if such a model were implemented, interpretation issues may arise for those underrepresented populations. Finally, in section 4, I discussed two potential methods for accounting for systemic bias arising in sleep research: quantifying systemic bias for use in the study, and disciplinary reflexivity where institutional influences on the study are clearly stated.

I hope to make it clear that, for systemic bias (especially in the context of sleep research), the concept of 'blaming' certain populations for these disparities does not make much sense. It is referred to as 'systemic' for a reason; the system relies on institutional structures to maintain the

biases, rather more than the actions of any one individual. To using the kaleidoscope example from the introduction: it would be akin to learning that, when using a kaleidoscope, you must ensure that a certain side of the kaleidoscope faces up. This has various benefits, one of which being that it ensures consistency and reproducibility of certain views in the kaleidoscope, easing communication and (to stretch the metaphor a little) collaboration between those who have used the kaleidoscope. A drawback, however, is the glaring oversight of the rest of the kaleidoscope's orientations — all of which may hold additional important information. Like systemic bias, it would not make sense to find a particular individual responsible for this oversight, but there is still the reality that this oversight exists. My hope is that the ideas expressed in this chapter will encourage greater awareness and critical examination of systemic biases in sleep research. By doing so, we may more forthrightly ensure justice and equity in sleep research.

Chapter V. LYING IN BED: CONCLUSIONS.

The overall focus of this thesis has been on the concept of sleep, a process that is essential for life. Specifically, my primary goal was to not only illustrate how many aspects of life relate with sleep in some way, but also perform sleep research in a way that is ground in this fact. The method I chose to best capture this multifaceted nature of sleep was to probe the concept sleep from three different disciplinary viewpoints: behavioral, computational, and philosophical, represented in Chapters II, III, and IV, respectively. While these are far from the only disciplines through one can study sleep (others include, for instance, molecular/cellular neurobiology,[219] social psychology,[220] or anthropology[221]), what I hope is that this multidisciplinary assembly of studies on sleep can represent an attempt to highlight the inherent interconnectedness of sleep.

Each chapter in this thesis possesses its own intrinsic value, yet it is the juxtaposition of seemingly disparate approaches across these three chapters that generates what I see as unique dialogue on the concept of sleep. When viewed primarily through a behavioral lens (particularly when utilizing animal models), certain nuances such as those seen when considering broader social disparities in sleep might be overlooked. Similarly, when viewed primarily through a philosophical lens, certain norms in the realm of sleep research (as contentious some of them can be) may go unacknowledged in philosophical analysis. By closely intertwining findings derived from these varied perspectives, I believe they inform and enrich one another in ways that lead to novel lines of inquiry, distinct from those typically explored within the general confines of their respective disciplines. This interdisciplinary approach allows for a comprehensive exploration of sleep, unveiling insights that might be missed when studied in isolation.

It is this dialogue that I see as one of the most significant contributions from this thesis. Again, I believe that more fully capturing the concept of sleep in research means simultaneously

approaching the concept from different disciplinary viewpoints. To make this point more clearly, in the following section I will provide concise summaries of each preceding chapter in this thesis, emphasizing key findings and concepts. Subsequently, I will synthesize the unique ideas explored in each chapter and underscore recurring themes, acting as a conclusion of this chapter and thesis. Throughout this concluding chapter, I aim to underscore the invaluable insights gained through the application of interdisciplinary methods to the study of sleep.

V.i CHAPTER SUMMARIES.

As previously outlined, each preceding chapter delved into the study of sleep from distinct perspectives, with Chapter I providing a comprehensive overview of sleep research and its multifaceted nature. The subsequent three chapters approached sleep research from different disciplinary viewpoints: behavioral, viewing sleep primarily in terms of actions done by the animal; computational, viewing sleep primarily in terms of quantitative analysis mathematical modeling; and philosophical, critically viewing sleep as a concept to explore assumptions and values. Chapter II represented a behavioral study of sleep, investigating the effects of contextual fear and abnormal circadian rhythms, as indicated by changes in circadian activity and sleep/wake timing. Chapter III represented a computational study of sleep, examining the use of computational methods (i.e., machine learning) to analyze neurophysiological data and gain insights into sleep architecture. Finally, Chapter IV represented a philosophical study of sleep, examining sleep research within the context of concepts from social justice literature and inspecting ways sleep research incorporates these ideas into its typical functioning. For the remainder of this section, each of these four chapters will be more comprehensively summarized, emphasizing their key findings and concepts.

V.i.1 *Chapter I. Making the bed: An introduction to sleep and sleep research*

The goal of Chapter I was to provide an overview of the remarkably multifaceted and adaptable nature of sleep. As the introductory chapter for this thesis, Chapter I served to highlight core ideas to serve as foundation and conceptual justification for following chapters. I opened this chapter by emphasizing two key concepts: 1) the evolutionarily conserved function of sleep and its vital role in wellbeing, and 2) its highly adaptive nature, both necessary for survival.[5] Delving into the mechanisms underlying sleep, I explored the neurobiology of sleep timing, highlighting the core role of the suprachiasmatic nucleus in regulating endogenous circadian rhythms which, in turn, regulate sleep timing. I then further explored the functions of sleep in promoting physical and mental wellbeing, while also exploring the debilitating effects of sleep disruptions. I highlighted three general causes of sleep disruptions: internal factors, external factors, and interactions between the two. Finally, I concluded the chapter by discussing three major avenues of sleep research: human studies, animal models, and computational methods, each providing unique insights into sleep processes.

V.i.2 *Chapter II. Sleep with one eye open: contextual fear and abnormal circadian rhythms.*

The 'behavioral viewpoint' of the thesis, in this chapter I presented results of a study examining how re-exposure a fearful context in the absence of aversive stimuli was enough to re-establish aberrant circadian rhythms in mice. This was done by 1) establishing baseline circadian activity in mice in a 'neutral context' and, subsequently, a 'fear context'; 2) randomly administering footshocks, either during the light phase (LF) or dark phase (DF) until mice entrained to the cyclic fear stimuli; 3) returning mice to the neutral context and allowing their

activity to return to baseline; and, finally 4) moving mice back to the fear context, without footshocks, to study the effect of contextual recall on circadian activity pattern.

I found that mice that entrained to cyclic fear in the fear context, allowed to return to baseline activity in the neutral context, and subsequently moved back to the fear context without the aversive stimuli, they exhibit activity patterns akin to those when they were entrained to the cyclic fear stimuli. As reported, this suggests the context is enough to re-establish abnormal circadian activity patterns that do not return to baseline even after a month. Additionally, this effect is modulated by intensity of the aversive stimuli and may differ depending on sex of the mouse. Given that one of the core symptoms of post-traumatic stress disorder (PTSD) in humans is sleep disruptions,[85], I hypothesized that there may be a potential role of neural systems underlying the circadian system in the development and maintenance of PTSD.

V.i.3 *Chapter III. Sleeping on the job: client-side, publicly available automated sleep scoring.*

The 'computational viewpoint' of the thesis, in this chapter I presented results for a publicly available web-based automated method of sleep scoring, Sleep Identification Enabled by Supervised Training Algorithms (SIESTA), intended to aid sleep researchers by reducing the amount of time spent manually scoring sleep in mice. In this study, electrocorticographic (ECoG) data obtained from both wild-type mice and a mouse model of epilepsy[222] were processed and used to train machine learning algorithms. Of the different models examined, gradient boosting classifier (GBF) and bagging classifier using random forest (BCRF) demonstrated the highest accuracy. Additionally, these models exhibited high accuracy in generalizing across different genotypes, data sources, and environmental conditions of data acquisition. These findings

underscore the potential utility of SIESTA in expediting the sleep scoring process for researchers, highlighting its potential utility in the field of sleep research.

V.i.4 *Chapter IV. Sleep, with both eyes open: systemic bias, intersectionality, and sleep research.*

The 'philosophical viewpoint' of the thesis, in this chapter I explored how historical biases in sleep research have perpetuated sociocultural disparities, continuing to adversely affect underrepresented populations. I examined factors contributing to health and sleep disparities, particularly focusing on the historical biases against women and non-White individuals (among other disadvantaged groups) — a point further emphasized by utilizing Kimberlé Crenshaw's framework of intersectionality.[159] Given the existence of these disparities, I walked through general steps of sleep research to highlight how systemic biases can emerge at each stage of the research process, from the formulation of a research question to the interpretation of results. This exploration prompted a critical question: I proposed two key recommendations: First, researchers should consider quantifying systemic bias by incorporating measures that account for social context influences. Secondly, embracing disciplinary reflexivity by critically examining institutional practices and influences that may have contributed to the perpetuation of systemic biases. The overarching objective of this chapter was to connect sleep research more closely with the broader social context of sleep, thereby highlighting the need to address systemic biases in scientific inquiry.

V.ii CROSS-CHAPTER SYNTHESIS, CONTRIBUTIONS, AND CONCLUSIONS.

Taken together, these chapters approach the central theme of this thesis from a variety of angles — including biological, environmental, technological, and societal angles. While it feels

almost like a cop out to say, the central theme for all these angles is the concept of sleep itself. While each chapter utilized different core ideas and methodologies, the core theme was this multifaceted nature of sleep. I believe this may be most clearly shown in Chapter IV, where I used concepts from Chapters II and III to further illustrate the importance of considering social context to add nuance to the results and conclusions presented in those chapters. As mentioned before, this added nuance can lead to unique questions, such as 'how might we model social disadvantage in animal models?', or, as touched on in chapter IV, 'how might we account for the effects of social disadvantage in computational models of sleep?' While currently difficult to answer, these sorts of questions may not only lead to greater representation in biomedical research, but will simultaneously begin to acknowledge and account for historical systemic biases in sleep research. In other words, the juxtaposition of these different viewpoints as exemplified in the current thesis can lead to unique avenues of inquiry.

As research into sleep continues, I hope my research underscores the importance of adopting a multidisciplinary approach to fully capture the complexities inherent in sleep. As shown, while behavioral, computational, and philosophical perspectives individually contribute to our understanding of sleep, close dialogues between the different perspectives can lead to unique avenues of inquiry. I believe in the need for close collaboration between disciplines to encourage a more holistic understanding of sleep that connects its neurobiological underpinnings to its sociocultural influences. In my view, influences from social context and systemic biases on sleep are often not given the recognition they deserve, and thus I believe fostering such inclusivity and representation in scientific exploration will only ever be, simply put, a good thing. I believe that this can be done by making a conscious effort to unite different perspectives on sleep. Sleep is essential for life; just as life is multifaceted, so is sleep.

BIBLIOGRAPHY

1. McLeod, S. Maslow's hierarchy of needs. *Simply psychology* **1**, (2007).
2. Nairne, J. S. & Pandeirada, J. N. Adaptive memory: The evolutionary significance of survival processing. *Perspectives on Psychological Science* **11**, 496–511 (2016).
3. Vitaterna, M. H., Takahashi, J. S. & Turek, F. W. Overview of Circadian Rhythms. *Alcohol Res Health* **25**, 85–93 (2001).
4. Moore, R. Y. Suprachiasmatic nucleus in sleep–wake regulation. *Sleep Medicine* **8**, 27–33 (2007).
5. Anafi, R. C., Kayser, M. S. & Raizen, D. M. Exploring phylogeny to find the function of sleep. *Nature Reviews Neuroscience* **20**, 109–116 (2019).
6. Alhola, P. & Polo-Kantola, P. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatr Dis Treat* **3**, 553–567 (2007).
7. Basner, M., Rao, H., Goel, N. & Dinges, D. F. Sleep deprivation and neurobehavioral dynamics. *Current Opinion in Neurobiology* **23**, 854–863 (2013).
8. Ben Simon, E., Vallat, R., Barnes, C. M. & Walker, M. P. Sleep Loss and the Socio-Emotional Brain. *Trends in Cognitive Sciences* **24**, 435–450 (2020).
9. Lee, M. L. *et al.* Fragmentation of Rapid Eye Movement and Nonrapid Eye Movement Sleep without Total Sleep Loss Impairs Hippocampus-Dependent Fear Memory Consolidation. *Sleep* **39**, 2021–2031 (2016).
10. Siegel, J. M. Sleep function: an evolutionary perspective. *The Lancet Neurology* **21**, 937–946 (2022).
11. Schmidt, M. H. The energy allocation function of sleep: A unifying theory of sleep, torpor, and continuous wakefulness. *Neuroscience & Biobehavioral Reviews* **47**, 122–153 (2014).
12. Eugene, A. R. & Masiak, J. The Neuroprotective Aspects of Sleep. *MEDtube Sci* **3**, 35–40 (2015).
13. Xie, L. *et al.* Sleep Drives Metabolite Clearance from the Adult Brain. *Science* **342**, 373–377 (2013).
14. Born, J. & Wilhelm, I. System consolidation of memory during sleep. *Psychological Research* **76**, 192–203 (2012).
15. Diekelmann, S. & Born, J. The memory function of sleep. *Nat Rev Neurosci* **11**, 114–126 (2010).

16. Ackermann, S. & Rasch, B. Differential Effects of Non-REM and REM Sleep on Memory Consolidation? *Curr Neurol Neurosci Rep* **14**, 430 (2014).
17. Wichlinski, L. J. Adaptive Solutions to the Problem of Vulnerability During Sleep. *Evolutionary Psychological Science* **8**, 442–477 (2022).
18. Richter, K., Adam, S., Geiss, L., Peter, L. & Niklewski, G. Two in a bed: The influence of couple sleeping and chronotypes on relationship and sleep. An overview. *Chronobiology International* **33**, 1464–1472 (2016).
19. Andre, C. J., Lovallo, V. & Spencer, R. M. C. The effects of bed sharing on sleep: From partners to pets. *Sleep Health* **7**, 314–323 (2021).
20. Febinger, H. Y., George, A., Priestley, J., Toth, L. A. & Opp, M. R. Effects of Housing Condition and Cage Change on Characteristics of Sleep in Mice. *J Am Assoc Lab Anim Sci* **53**, 29–37 (2014).
21. Lima, S. L., Rattenborg, N. C., Lesku, J. A. & Amlaner, C. J. Sleeping under the risk of predation. *Animal Behaviour* **70**, 723–736 (2005).
22. Berro, L. F. *et al.* Influence of Pair-housing on Sleep Parameters Evaluated with Actigraphy in Female Rhesus Monkeys. *j am assoc lab anim sci* **61**, 165–172 (2022).
23. Samson, D. R. The human sleep paradox: The unexpected sleeping habits of Homo sapiens. *Annual Review of Anthropology* **50**, 259–274 (2021).
24. McKenna, J. J. & McDade, T. Why babies should never sleep alone: A review of the co-sleeping controversy in relation to SIDS, bedsharing and breast feeding. *Paediatric Respiratory Reviews* **6**, 134–152 (2005).
25. Chokroverty, S. Overview of sleep & sleep disorders. *Indian Journal of Medical Research* **131**, 126–140 (2010).
26. Phillips, A. J. K., Fulcher, B. D., Robinson, P. A. & Klerman, E. B. Mammalian Rest/Activity Patterns Explained by Physiologically Based Modeling. *PLoS Comput Biol* **9**, e1003213 (2013).
27. Van Reeth, O. *et al.* PHYSIOLOGY OF SLEEP (REVIEW)–Interactions between stress and sleep: from basic research to clinical situations. *Sleep Medicine Reviews* **4**, 201–219 (2000).
28. Hastings, M. H., Maywood, E. S. & O'Neill, J. S. Cellular Circadian Pacemaking and the Role of Cytosolic Rhythms. *Current Biology* **18**, R805–R815 (2008).

29. Sharma, V. K. & Chandrashekaran, M. K. Zeitgebers (time cues) for biological clocks. *Current Science* **89**, 1136–1146 (2005).
30. Hirschie Johnson, C., Elliott, J. A. & Foster, R. Entrainment of Circadian Programs. *Chronobiology International* **20**, 741–774 (2003).
31. Jilge, B. & Stahle, H. Restricted food access and light-dark: impact of conflicting zeitgebers on circadian rhythms of the rabbit. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **264**, R708–R715 (1993).
32. Pellman, B. A. *et al.* Time-Specific Fear Acts as a Non-Photic Entraining Stimulus of Circadian Rhythms in Rats. *Sci Rep* **5**, 14916 (2015).
33. Caldelas, I., Chimal-Monroy, J., Martínez-Gómez, M. & Hudson, R. Non-photic circadian entrainment in mammals: A brief review and proposal for study during development. *Biological Rhythm Research* **36**, 23–37 (2005).
34. Sanchez, R. E. A., Kalume, F. & de la Iglesia, H. O. Sleep timing and the circadian clock in mammals: Past, present and the road ahead. *Semin Cell Dev Biol* **126**, 3–14 (2022).
35. Kohyama, J. Which Is More Important for Health: Sleep Quantity or Sleep Quality? *Children (Basel)* **8**, 542 (2021).
36. Skeldon, A. C., Derks, G. & Dijk, D.-J. Modelling changes in sleep timing and duration across the lifespan: Changes in circadian rhythmicity or sleep homeostasis? *Sleep Medicine Reviews* **28**, 96–107 (2016).
37. Vyazovskiy, V. V. Sleep, recovery, and metaregulation: explaining the benefits of sleep. *Nat Sci Sleep* **7**, 171–184 (2015).
38. Do, M. T. H. Melanopsin and the Intrinsically Photosensitive Retinal Ganglion Cells: Biophysics to Behavior. *Neuron* **104**, 205–226 (2019).
39. Hastings, M. H., Maywood, E. S. & Brancaccio, M. The Mammalian Circadian Timing System and the Suprachiasmatic Nucleus as Its Pacemaker. *Biology (Basel)* **8**, 13 (2019).
40. Yoo, S.-H. *et al.* PERIOD2::LUCIFERASE real-time reporting of circadian dynamics reveals persistent circadian oscillations in mouse peripheral tissues. *Proc Natl Acad Sci U S A* **101**, 5339–5346 (2004).
41. Ralph, M. R., Foster, R. G., Davis, F. C. & Menaker, M. Transplanted Suprachiasmatic Nucleus Determines Circadian Period. *Science* **247**, 975–978 (1990).

42. Michel, S., Colwell, C. S. & Colwell, C. S. Cellular Communication and Coupling Within the Suprachiasmatic Nucleus. *Chronobiology International* **18**, 579–600 (2001).
43. Mistlberger, R. E. Food as circadian time cue for appetitive behavior. *F1000Res* **9**, F1000 Faculty Rev-61 (2020).
44. Naska, A., Oikonomou, E., Trichopoulou, A., Psaltopoulou, T. & Trichopoulos, D. Siesta in Healthy Adults and Coronary Mortality in the General Population. *Archives of Internal Medicine* **167**, 296–301 (2007).
45. Irish, L. A., Kline, C. E., Gunn, H. E., Buysse, D. J. & Hall, M. H. The Role of Sleep Hygiene in Promoting Public Health: A Review of Empirical Evidence. *Sleep Med Rev* **22**, 23–36 (2015).
46. Cannon, W. B. Organization for physiological homeostasis. *Physiological Reviews* **9**, 399–431 (1929).
47. Christian, L. M., Graham, J. E., Padgett, D. A., Glaser, R. & Kiecolt-Glaser, J. K. Stress and Wound Healing. *Neuroimmunomodulation* **13**, 337–346 (2006).
48. Hirotsu, C., Tufik, S. & Andersen, M. L. Interactions between sleep, stress, and metabolism: From physiological to pathological conditions. *Sleep Sci* **8**, 143–152 (2015).
49. Cirelli, C. Sleep and synaptic changes. *Curr Opin Neurobiol* **23**, 841–846 (2013).
50. De La Iglesia, H. O. & Lee, M. L. A Time to Wake, a Time to Sleep. in *Mechanisms of Circadian Systems in Animals and Their Clinical Relevance* (eds. Aguilar-Roblero, R., Díaz-Muñoz, M. & Fanjul-Moles, M. L.) 197–217 (Springer International Publishing, 2015). doi:10.1007/978-3-319-08945-4_11.
51. Scammell, T. E., Arrigoni, E. & Lipton, J. O. Neural Circuitry of Wakefulness and Sleep. *Neuron* **93**, 747–765 (2017).
52. Beenhakker, M. P. & Huguenard, J. R. Neurons that Fire Together Also Conspire Together: Is Normal Sleep Circuitry Hijacked to Generate Epilepsy? *Neuron* **62**, 612–632 (2009).
53. Latchoumane, C.-F. V., Ngo, H.-V. V., Born, J. & Shin, H.-S. Thalamic Spindles Promote Memory Formation during Sleep through Triple Phase-Locking of Cortical, Thalamic, and Hippocampal Rhythms. *Neuron* **95**, 424-435.e6 (2017).
54. Wang, D. V. & Ikemoto, S. Coordinated Interaction between Hippocampal Sharp-Wave Ripples and Anterior Cingulate Unit Activity. *J. Neurosci.* **36**, 10663–10672 (2016).

55. Colvonen, P. J., Straus, L. D., Acheson, D. & Gehrman, P. A Review of the Relationship Between Emotional Learning and Memory, Sleep, and PTSD. *Curr Psychiatry Rep* **21**, 2 (2019).
56. Agnew, H. W., Webb, W. B. & Williams, R. L. The first night effect: An EEG study of sleep. *Psychophysiology* **2**, 263–266 (1966).
57. Hamet, P. & Tremblay, J. Genetics of the sleep-wake cycle and its disorders. *Metabolism* **55**, S7–S12 (2006).
58. Alvaro, P. K., Roberts, R. M. & Harris, J. K. A Systematic Review Assessing Bidirectionality between Sleep Disturbances, Anxiety, and Depression. *Sleep* **36**, 1059–1068 (2013).
59. Ray, K. Does Where We Sleep Matter? in *Black Health: The Social, Political, and Cultural Determinants of Black People's Health* (Oxford University Press, 2023).
60. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Science* **28**, 1947–1951 (2019).
61. Genderson, M. R. *et al.* Genetic and Environmental Influences on Sleep Quality in Middle-Aged Men: A Twin Study. *J Sleep Res* **22**, 519–526 (2013).
62. Takahashi, J. S. Molecular components of the circadian clock in mammals. *Diabetes, Obesity and Metabolism* **17**, 6–11 (2015).
63. Chokroverty, S. M. D. Diagnosis and treatment of sleep disorders caused by co-morbid disease. *Neurology Effect of Co-morbid Disease and Neuropharmacology on Sleep Architecture* **54**, (2000).
64. Johnson, D. A., Billings, M. E. & Hale, L. Environmental Determinants of Insufficient Sleep and Sleep Disorders: Implications for Population Health. *Curr Epidemiol Rep* **5**, 61–69 (2018).
65. Koenen, K. C., Nugent, N. R. & Amstadter, A. B. Gene-environment interaction in posttraumatic stress disorder. *Eur Arch Psychiatry Clin Neurosc* **258**, 82–96 (2008).
66. Brewin, C. R. *et al.* A review of current evidence regarding the ICD-11 proposals for diagnosing PTSD and complex PTSD. *Clinical Psychology Review* **58**, 1–15 (2017).
67. Lipinska, G. & Thomas, K. G. F. The Interaction of REM Fragmentation and Night-Time Arousal Modulates Sleep-Dependent Emotional Memory Consolidation. *Frontiers in Psychology* **10**, (2019).

68. Lee, M. L., Swanson, B. E. & De La Iglesia, H. O. Circadian Timing of REM Sleep Is Coupled to an Oscillator within the Dorsomedial Suprachiasmatic Nucleus. *Current Biology* **19**, 848–852 (2009).
69. Worley, S. L. The Extraordinary Importance of Sleep. *P T* **43**, 758–763 (2018).
70. Zimmerman, J. E., Naidoo, N., Raizen, D. M. & Pack, A. I. Conservation of sleep: insights from non-mammalian model systems. *Trends in Neurosciences* **31**, 371–376 (2008).
71. Watson, N. F. & Fernandez, C. R. Artificial intelligence and sleep: Advancing sleep medicine. *Sleep Medicine Reviews* **59**, 101512 (2021).
72. Deak, M. & Epstein, L. J. The History of Polysomnography. *Sleep Medicine Clinics* **4**, 313–321 (2009).
73. Vaughn, B. V. & Giallanza, P. Technical review of polysomnography. *Chest* **134**, 1310–1319 (2008).
74. Choy, T., Baker, E. & Stavropoulos, K. Systemic Racism in EEG Research: Considerations and Potential Solutions. *Affec Sci* **3**, 14–20 (2022).
75. Sadeh, A. & Acebo, C. The role of actigraphy in sleep medicine. *Sleep Medicine Reviews* **6**, 113–124 (2002).
76. Park, J. *et al.* Prediction of Locomotor Activity by Infrared Motion Detector on Sleep-wake State in Mice. *Clin Psychopharmacol Neurosci* **19**, 303–312 (2021).
77. Bryant, C. D. The blessings and curses of C57BL/6 substrains in mouse genetic studies. *Ann N Y Acad Sci* **1245**, 31–33 (2011).
78. Bockamp, E. *et al.* Of mice and models: improved animal models for biomedical research. *Physiological Genomics* **11**, 115–132 (2002).
79. Sarsani, V. K. *et al.* The Genome of C57BL/6J "Eve", the Mother of the Laboratory Mouse Genome Reference Strain. *G3 (Bethesda)* **9**, 1795–1805 (2019).
80. Deslauriers, J., Toth, M., Der-Avakian, A. & Risbrough, V. B. Current Status of Animal Models of Posttraumatic Stress Disorder: Behavioral and Biological Phenotypes, and Future Challenges in Improving Translation. *Biological Psychiatry* **83**, 895–907 (2018).
81. Wang, X.-J. *et al.* Computational neuroscience: a frontier of the 21st century. *National Science Review* **7**, 1418–1422 (2020).
82. Usman, S. M. *et al.* Using scalp EEG and intracranial EEG signals for predicting epileptic seizures: Review of available methodologies. *Seizure* **71**, 258–269 (2019).

83. Fisher, S. P. *et al.* Rapid assessment of sleep/wake behaviour in mice. *J Biol Rhythms* **27**, 48–58 (2012).
84. Kelmendi, B. *et al.* PTSD: from neurobiology to pharmacological treatments. *Eur J Psychotraumatol* **7**, 31858 (2016).
85. Slavish, D. C., Briggs, M., Fentem, A., Messman, B. A. & Contractor, A. A. Bidirectional associations between daily PTSD symptoms and sleep disturbances: A systematic review. *Sleep Medicine Reviews* **63**, 101623 (2022).
86. Koffel, E., Khawaja, I. S. & Germain, A. Sleep Disturbances in Posttraumatic Stress Disorder: Updated Review and Implications for Treatment. *Psychiatr Ann* **46**, 173–176 (2016).
87. Agorastos, A. & Olff, M. Traumatic stress and the circadian system: neurobiology, timing and treatment of posttraumatic chronodisruption. *Eur J Psychotraumatol* **11**, 1833644 (2020).
88. Bussi, I. L. *et al.* The circadian molecular clock in the suprachiasmatic nucleus is necessary but not sufficient for fear entrainment in the mouse. 2023.06.26.546624 Preprint at https://doi.org/10.1101/2023.06.26.546624 (2023).
89. R Core Team. R: A Language and Environment for Statistical Computing. (2023).
90. Chambers, J. M., Freeny, A. E., Heiberger, R. M. & Hastie, T. J. Analysis of variance; designed experiments. in *Statistical models in S* 145–193 (Routledge, 2017).
91. Chambers, J. M. Linear models. in *Statistical models in S* (ed. Hastie, T. J.) (Routledge, 2017).
92. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. (2016) doi:https://doi.org/10.1007/978-0-387-98141-3.
93. Landler, L., Ruxton, G. D. & Malkemper, E. P. Circular data in biology: advice for effectively implementing statistical procedures. *Behavioral ecology and sociobiology* **72**, 1–10 (2018).
94. Yeh, S.-T. Using trapezoidal rule for the area under a curve calculation. *Proceedings of the 27th Annual SAS® User Group International (SUGI'02)* 1–5 (2002).
95. Abdi, H. & Williams, L. J. Newman-Keuls test and Tukey test. *Encyclopedia of research design* **2**, 897–902 (2010).

96. Armstrong, R. A. When to use the Bonferroni correction. *Ophthalmic and Physiological Optics* **34**, 502–508 (2014).
97. Clark, J. W., Drummond, S. P. A., Hoyer, D. & Jacobson, L. H. Sex differences in mouse models of fear inhibition: Fear extinction, safety learning, and fear–safety discrimination. *Br J Pharmacol* **176**, 4149–4158 (2019).
98. McLean, C. P. & Anderson, E. R. Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review* **29**, 496–505 (2009).
99. Sutton, E. L. Insomnia. *Ann Intern Med* **174**, ITC33–ITC48 (2021).
100. Schulz, H. & Salzarulo, P. The Development of Sleep Medicine: A Historical Sketch. *J Clin Sleep Med* **12**, 1041–1052 (2016).
101. Lim, D. C. *et al.* Reinventing polysomnography in the age of precision medicine. *Sleep Med Rev* **52**, 101313 (2020).
102. AASM scoring Manual - American academy of sleep medicine. *American Academy of Sleep Medicine – Association for Sleep Clinicians and Researchers* (2020).
103. Katsageorgiou, V.-M., Lassi, G., Tucci, V., Murino, V. & Sona, D. Sleep-stage scoring in mice: The influence of data pre-processing on a system's performance. in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 598–601 (IEEE, 2015).
104. Wendt, S. L. *et al.* Inter-expert and intra-expert reliability in sleep spindle scoring. *Clin Neurophysiol* **126**, 1548–1556 (2015).
105. Rosenberg, R. S. & Van Hout, S. The American Academy of Sleep Medicine Inter-scorer Reliability Program: Sleep Stage Scoring. *J Clin Sleep Med* **9**, 81–87 (2013).
106. Danker-Hopfe, H. *et al.* Interrater reliability between scorers from eight European sleep laboratories in subjects with different sleep disorders. *J Sleep Res* **13**, 63–69 (2004).
107. Younes, M., Raneri, J. & Hanly, P. Staging Sleep in Polysomnograms: Analysis of Inter-Scorer Variability. *J Clin Sleep Med* **12**, 885–894 (2016).
108. Gerardy, B. *et al.* An Approach for Determining the Reliability of Manual and Digital Scoring of Sleep Stages. *Sleep* zsad248 (2023) doi:10.1093/sleep/zsad248.
109. Smith, J. R., Negin, M. & Nevis, A. H. Automatic analysis of sleep electroencephalograms by hybrid computation. *IEEE transactions on systems science and cybernetics* **5**, 278–284 (1969).

110. Faust, O., Razaghi, H., Barika, R., Ciaccio, E. J. & Acharya, U. R. A review of automated sleep stage scoring based on physiological signals for the new millennia. *Computer methods and programs in biomedicine* **176**, 81–91 (2019).
111. Miladinović, Đ. *et al.* SPINDLE: End-to-end learning from EEG/EMG to extrapolate animal sleep scoring across experimental settings, labs and species. *PLoS computational biology* **15**, e1006968 (2019).
112. Allocca, G. *et al.* Validation of 'Somnivore', a machine learning algorithm for automated scoring and analysis of polysomnography data. *Frontiers in neuroscience* **13**, 207 (2019).
113. Yu, F. H. *et al.* Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. *Nature neuroscience* **9**, 1142–1149 (2006).
114. Kalume, F. *et al.* Sleep impairment and reduced interneuron excitability in a mouse model of Dravet Syndrome. *Neurobiology of disease* **77**, 141–154 (2015).
115. Han, S. *et al.* NaV1. 1 channels are critical for intercellular communication in the suprachiasmatic nucleus and for normal circadian rhythms. *Proceedings of the National Academy of Sciences* **109**, E368–E377 (2012).
116. Sanchez, R. E. *et al.* Circadian regulation of sleep in a pre-clinical model of Dravet syndrome: dynamics of sleep stage and siesta re-entrainment. *Sleep* **42**, zsz173 (2019).
117. Hsu, Y.-W. A. *et al.* The dorsal medial habenula minimally impacts circadian regulation of locomotor activity and sleep. *Journal of biological rhythms* **32**, 444–455 (2017).
118. Marrosu, F. *et al.* Beta and gamma range EEG power-spectrum correlation with spiking discharges in DBA/2J mice absence model: role of GABAB receptors. *Epilepsia* **47**, 489–494 (2006).
119. Franken, P., Malafosse, A. & Tafti, M. Genetic variation in EEG activity during sleep in inbred mice. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **275**, R1127–R1137 (1998).
120. Feng, Z.-X., Dong, H., Qu, W.-M. & Zhang, W. Oral Delivered Dexmedetomidine Promotes and Consolidates Non-rapid Eye Movement Sleep via Sleep–Wake Regulation Systems in Mice. *Frontiers in pharmacology* **9**, 1196 (2018).
121. Lacourse, K., Delfrate, J., Beaudry, J., Peppard, P. & Warby, S. C. A sleep spindle detection algorithm that emulates human expert spindle scoring. *Journal of neuroscience methods* **316**, 3–11 (2019).

122. Pedregosa, F. *et al. Scikit-learn: machine learning in python J Mach Learn Res 12: 2825–2830*. (2011).
123. Hastie, T., Tibshirani, R., Friedman, J. H. & Friedman, J. H. *The elements of statistical learning: data mining, inference, and prediction*. vol. 2 (Springer, 2009).
124. Breiman, L. Random forests. *Machine learning* **45**, 5–32 (2001).
125. Dietterich, T. G. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine learning* **40**, 139–157 (2000).
126. Breiman, L. Bagging predictors. *Machine learning* **24**, 123–140 (1996).
127. Landis, J. R. & Koch, G. G. The measurement of observer agreement for categorical data. *biometrics* 159–174 (1977).
128. Magalang, U. J. *et al.* Agreement in the scoring of respiratory events and sleep among international sleep centers. *Sleep* **36**, 591–596 (2013).
129. Hasan, S., Foster, R. G., Vyazovskiy, V. V. & Peirson, S. N. Effects of circadian misalignment on sleep in mice. *Scientific Reports* **8**, 15343 (2018).
130. Murtagh, F. & Contreras, P. Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **2**, 86–97 (2012).
131. Ferri, F. J., Pudil, P., Hatef, M. & Kittler, J. Comparative study of techniques for large-scale feature selection. in *Machine intelligence and pattern recognition* vol. 16 403–413 (Elsevier, 1994).
132. Friedman, J. H. Greedy function approximation: a gradient boosting machine. *Annals of statistics* 1189–1232 (2001).
133. Libourel, P.-A., Corneyllie, A., Luppi, P.-H., Chouvet, G. & Gervasoni, D. Unsupervised online classifier in sleep scoring for sleep deprivation studies. *Sleep* **38**, 815–828 (2015).
134. Mayer, J., Rahman, R., Ghosh, S. & Pal, R. Sequential feature selection and inference using multi-variate random forests. *Bioinformatics* **34**, 1336–1344 (2018).
135. Silber, M. H. *et al.* The visual scoring of sleep in adults. *Journal of clinical sleep medicine* **3**, 121–131 (2007).
136. Jiang, X., Gonzalez-Martinez, J. & Halgren, E. Coordination of human hippocampal sharpwave ripples during NREM sleep with cortical theta bursts, spindles, downstates, and upstates. *Journal of Neuroscience* **39**, 8744–8761 (2019).

137. Kemp, B., Värri, A., Rosa, A. C., Nielsen, K. D. & Gade, J. A simple format for exchange of digitized polygraphic recordings. *Electroencephalography and clinical neurophysiology* **82**, 391–393 (1992).
138. Cooray, N. *et al.* Detection of REM sleep behaviour disorder by automated polysomnography analysis. *Clinical Neurophysiology* **130**, 505–514 (2019).
139. Klok, A. B. *et al.* A new fully automated random-forest algorithm for sleep staging. in *2018 40th annual international conference of the ieee engineering in medicine and biology society (EMBC)* 4920–4923 (IEEE, 2018).
140. Şen, B., Peker, M., Çavuşoğlu, A. & Çelebi, F. V. A comparative study on classification of sleep stage based on EEG signals using feature selection and classification algorithms. *Journal of medical systems* **38**, 1–21 (2014).
141. Gao, V., Turek, F. & Vitaterna, M. Multiple classifier systems for automatic sleep scoring in mice. *Journal of neuroscience methods* **264**, 33–39 (2016).
142. Fiorillo, L. *et al.* Automated sleep scoring: A review of the latest approaches. *Sleep medicine reviews* **48**, 101204 (2019).
143. Gangstad, S. W. *et al.* Automatic sleep stage classification based on subcutaneous EEG in patients with epilepsy. *Biomedical engineering online* **18**, 1–17 (2019).
144. Skorucak, J. *et al.* Automatic detection of microsleep episodes with feature-based machine learning. *Sleep* **43**, zsz225 (2020).
145. Exarchos, I. *et al.* Supervised and unsupervised machine learning for automated scoring of sleep–wake and cataplexy in a mouse model of narcolepsy. *Sleep* **43**, zsz272 (2020).
146. Zandi, A. S., Tafreshi, R., Javidan, M. & Dumont, G. A. Predicting epileptic seizures in scalp EEG based on a variational Bayesian Gaussian mixture model of zero-crossing intervals. *IEEE Transactions on Biomedical Engineering* **60**, 1401–1413 (2013).
147. Garcia-Cairasco, N. Learning about brain physiology and complexity from the study of the epilepsies. *Brazilian Journal of Medical and Biological Research* **42**, 76–86 (2009).
148. Choo, B. P. *et al.* Benchmarking performance of an automatic polysomnography scoring system in a population with suspected sleep disorders. *Frontiers in Neurology* **14**, 1123935 (2023).
149. Liu, Y., Yang, Z., You, Y., Shan, W. & Ban, W. An attention-based temporal convolutional network for rodent sleep stage classification across species, mutants and experimental

environments with single-channel electroencephalogram. *Physiological Measurement* **43**, 085002 (2022).

150. Kwon, K., Kwon, S. & Yeo, W.-H. Automatic and accurate sleep stage classification via a convolutional deep neural network and nanomembrane electrodes. *Biosensors* **12**, 155 (2022).
151. Cho, J. H. *et al.* Validation study on automated sleep stage scoring using a deep learning algorithm. *Medicina* **58**, 779 (2022).
152. Begley, C. G. & Ioannidis, J. P. A. Reproducibility in Science. *Circulation Research* **116**, 116–126 (2015).
153. National Academies of Sciences, Engineering & Medicine. *Reproducibility and Replicability in Science*. (The National Academies Press, 2019). doi:10.17226/25303.
154. *Reflexivity*. (John Wiley & Sons, Ltd, 2003). doi:10.1002/9780470776094.
155. Toole, B. Demarginalizing Standpoint Epistemology. *Episteme* **19**, 47–65 (2022).
156. Adler, N. E. & Ostrove, J. M. Socioeconomic Status and Health: What We Know and What We Don't. *Annals of the New York Academy of Sciences* **896**, 3–15 (1999).
157. Grandner, M. A., Williams, N. J., Knutson, K. L., Roberts, D. & Jean-Louis, G. Sleep disparity, race/ethnicity, and socioeconomic position. *Sleep Medicine* **18**, 7–18 (2016).
158. Scharff, D. P. *et al.* More than Tuskegee: Understanding Mistrust about Research Participation. *J Health Care Poor Underserved* **21**, 879–897 (2010).
159. Crenshaw, K. W. *On intersectionality: Essential writings*. (The New Press, 2017).
160. Ware, J. E. J. SF-36 Health Survey Update. *Spine* **25**, 3130 (2000).
161. Nelson, K. L., Davis, J. E. & Corbett, C. F. Sleep quality: An evolutionary concept analysis. *Nursing Forum* **57**, 144–151 (2022).
162. National Academies of Sciences, E. *et al.* The State of Health Disparities in the United States. in *Communities in Action: Pathways to Health Equity* (National Academies Press (US), 2017).
163. Chen Jr, M. S., Lara, P. N., Dang, J. H. T., Paterniti, D. A. & Kelly, K. Twenty years post-NIH Revitalization Act: Enhancing minority participation in clinical trials (EMPaCT): Laying the groundwork for improving minority clinical trial accrual. *Cancer* **120**, 1091–1096 (2014).

164. Miller, L. R. *et al.* Considering sex as a biological variable in preclinical research. *The FASEB Journal* **31**, 29 (2017).
165. Liu, K. A. & Mager, N. A. D. Women's involvement in clinical trials: historical perspective and future implications. *Pharm Pract (Granada)* **14**, 708 (2016).
166. Vogel, B. *et al.* The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. *The Lancet* **397**, 2385–2438 (2021).
167. Lichtman, J. H. *et al.* Sex Differences in the Presentation and Perception of Symptoms Among Young Patients With Myocardial Infarction. *Circulation* **137**, 781–790 (2018).
168. Mong, J. A. & Cusmano, D. M. Sex differences in sleep: impact of biological sex and sex steroids. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150110 (2016).
169. Nicolaides, N. C. & Chrousos, G. P. Sex differences in circadian endocrine rhythms: Clinical implications. *European Journal of Neuroscience* **52**, 2575–2585 (2020).
170. Shaver, J. L. & Woods, N. F. Sleep and menopause: a narrative review. *Menopause* **22**, 899–915 (2015).
171. Reeves, A., Elliott, M. R., Karvonen-Gutierrez, C. A. & Harlow, S. D. Systematic exclusion at study commencement masks earlier menopause for Black women in the Study of Women's Health Across the Nation (SWAN). *International Journal of Epidemiology* **52**, 1612–1623 (2023).
172. Mallampalli, M. P. & Carter, C. L. Exploring Sex and Gender Differences in Sleep Health: A Society for Women's Health Research Report. *Journal of Women's Health* **23**, 553–562 (2014).
173. Harlow, S. D., Sievert, L. L., LaCroix, A. Z., Mishra, G. D. & Woods, N. F. Women's midlife health: the unfinished research agenda. *Womens Midlife Health* **9**, 7 (2023).
174. Lee, C. "Race" and "ethnicity" in biomedical research: How do scientists construct and explain differences in health? *Social Science & Medicine* **68**, 1183–1190 (2009).
175. Ahn, S. *et al.* A scoping review of racial/ethnic disparities in sleep. *Sleep Medicine* **81**, 169–179 (2021).
176. Perrin, C. & Anderson, K. Reframing craniometry: human exceptionalism and the production of racial knowledge. *Social Identities* **19**, 90–103 (2013).

177. Gamble, V. N. Under the shadow of Tuskegee: African Americans and health care. *Am J Public Health* **87**, 1773–1778 (1997).
178. Giddens, N. T., Juneau, P., Manza, P., Wiers, C. E. & Volkow, N. D. Disparities in sleep duration among American children: effects of race and ethnicity, income, age, and sex. *Proceedings of the National Academy of Sciences* **119**, e2120009119 (2022).
179. Chen, X. *et al.* Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA). *SLEEP* (2015) doi:10.5665/sleep.4732.
180. Egan, K. J., Knutson, K. L., Pereira, A. C. & von Schantz, M. The role of race and ethnicity in sleep, circadian rhythms and cardiovascular health. *Sleep Medicine Reviews* **33**, 70–78 (2017).
181. Thoman, E. B. Co-sleeping, an ancient practice: issues of the past and present, and possibilities for the future. *Sleep Medicine Reviews* **10**, 407–417 (2006).
182. Ramos, K. D. & Youngclarke, D. M. Parenting Advice Books About Child Sleep: Cosleeping and Crying It Out. *Sleep* **29**, 1616–1623 (2006).
183. Barry, E. S. Co-sleeping as a proximal context for infant development: The importance of physical touch. *Infant Behavior and Development* **57**, 101385 (2019).
184. Billings, M. E., Hale, L. & Johnson, D. A. Physical and Social Environment Relationship With Sleep Health and Disorders. *Chest* **157**, 1304–1312 (2020).
185. Navarro-Carrillo, G., Alonso-Ferres, M., Moya, M. & Valor-Segura, I. Socioeconomic Status and Psychological Well-Being: Revisiting the Role of Subjective Socioeconomic Status. *Frontiers in Psychology* **11**, (2020).
186. Etindele Sosso, F. A., Holmes, S. D. & Weinstein, A. A. Influence of socioeconomic status on objective sleep measurement: A systematic review and meta-analysis of actigraphy studies. *Sleep Health* **7**, 417–428 (2021).
187. Rogers, J. & Kelly, U. A. Feminist intersectionality: Bringing social justice to health disparities research. *Nurs Ethics* **18**, 397–407 (2011).
188. Kviz, F. J. *Conducting health research: Principles, process, and methods*. (SAGE Publications, 2019).
189. Altenmüller, M. S., Lange, L. L. & Gollwitzer, M. When research is me-search: How researchers' motivation to pursue a topic affects laypeople's trust in science. *PLOS ONE* **16**, e0253911 (2021).

190. Gardner, S. K. *et al.* “Me-search”: Challenges and opportunities regarding subjectivity in knowledge construction. *SGPE* **8**, 88–108 (2017).
191. Meline, T. Selecting Studies for Systemic Review: Inclusion and Exclusion Criteria. *CICSD* **33**, 21–27 (2006).
192. Bahji, A., Acion, L., Laslett, A.-M. & Adinoff, B. Exclusion of the non-English-speaking world from the scientific literature: Recommendations for change for addiction journals and publishers. *Nordic Studies on Alcohol and Drugs* **40**, 6–13 (2023).
193. Salinas, D. The socio-economic gap in foreign-language learning. (2021).
194. Markun, L. C. & Sampat, A. Clinician-Focused Overview and Developments in Polysomnography. *Curr Sleep Medicine Rep* **6**, 309–321 (2020).
195. Stepnowsky Jr, C. J., Moore, P. J. & Dimsdale, J. E. Effect of ethnicity on sleep: complexities for epidemiologic research. *Sleep* **26**, 329–332 (2003).
196. Frontiers | Demographic reporting and phenotypic exclusion in fNIRS. https://www.frontiersin.org/articles/10.3389/fnins.2023.1086208/full.
197. Bryda, E. C. The Mighty Mouse: The Impact of Rodents on Advances in Biomedical Research. *Mo Med* **110**, 207–211 (2013).
198. Lacroix, M. M. *et al.* Improved sleep scoring in mice reveals human-like stages. (2018) doi:10.1101/489005.
199. Arnold, A. P. Conceptual frameworks and mouse models for studying sex differences in physiology and disease: Why compensation changes the game. *Exp Neurol* **0**, 2–9 (2014).
200. Dib, R., Gervais, N. J. & Mongrain, V. A review of the current state of knowledge on sex differences in sleep and circadian phenotypes in rodents. *Neurobiol Sleep Circadian Rhythms* **11**, 100068 (2021).
201. Yaw, A. M., Duong, T. V., Nguyen, D. & Hoffmann, H. M. Circadian rhythms in the mouse reproductive axis during the estrous cycle and pregnancy. *J Neurosci Res* **99**, 294–308 (2021).
202. Karp, N. A. & Reavey, N. Sex bias in preclinical research and an exploration of how to change the status quo. *British Journal of Pharmacology* **176**, 4107–4118 (2019).
203. Waltz, M., Saylor, K. W., Fisher, J. A. & Walker, R. L. Biomedical Researchers’ Perceptions of the NIH’s Sex as a Biological Variable Policy for Animal Research: Results from a U.S. National Survey. *Journal of Women’s Health* **30**, 1395–1405 (2021).

204. Haag, K. *et al.* The emergence of sex differences in PTSD symptoms across development: evidence from the ALSPAC cohort. *Psychol Med* **50**, 1755–1760 (2020).
205. Shansky, R. M. Sex differences in PTSD resilience and susceptibility: Challenges for animal models of fear learning. *Neurobiology of Stress* **1**, 60–65 (2015).
206. Roberts, A. L., Gilman, S. E., Breslau, J., Breslau, N. & Koenen, K. C. Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. *Psychol Med* **41**, 71–83 (2011).
207. Webb, E. K. *et al.* The role of pain and socioenvironmental factors on posttraumatic stress disorder symptoms in traumatically injured adults: A 1-year prospective study. *J Trauma Stress* **35**, 1142–1153 (2022).
208. Verbitsky, A., Dopfel, D. & Zhang, N. Rodent models of post-traumatic stress disorder: behavioral assessment. *Transl Psychiatry* **10**, 1–28 (2020).
209. Alegría, M. *et al.* Prevalence, risk, and correlates of posttraumatic stress disorder across ethnic and racial minority groups in the United States. *Medical care* **51**, 1114–1123 (2013).
210. Lupia, A. & Alter, G. Data Access and Research Transparency in the Quantitative Tradition. *APSC* **47**, 54–59 (2014).
211. Cotton, J. L., O'Neill, B. S. & Griffin, A. The "name game": affective and hiring reactions to first names. *Journal of Managerial Psychology* **23**, 18–39 (2008).
212. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters* (2018).
213. Kordzadeh, N. & Ghasemaghaei, M. Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems* **31**, 1–22 (2021).
214. Blair, I. V., Steiner, J. F. & Havranek, E. P. Unconscious (Implicit) Bias and Health Disparities: Where Do We Go from Here? *Perm J* **15**, 71–78 (2011).
215. Johnson, D. A., Jackson, C. L., Williams, N. J. & Alcántara, C. Are sleep patterns influenced by race/ethnicity–a marker of relative advantage or disadvantage? Evidence to date. *Nature and science of sleep* 79–95 (2019).
216. Braveman, P. A., Arkin, E., Proctor, D., Kauh, T. & Holm, N. Systemic And Structural Racism: Definitions, Examples, Health Damages, And Approaches To Dismantling. *Health Affairs* **41**, 171–178 (2022).

217. Groos, M., Wallace, M., Hardeman, R. & Theall, K. P. Measuring inequity: a systematic review of methods used to quantify structural racism. **11**, (2018).
218. LaFave, S. E. *et al.* Quantifying older Black Americans' exposure to structural racial discrimination: how can we measure the water in which we swim? *Journal of Urban Health* **99**, 794–802 (2022).
219. Pace-Schott, E. F. & Hobson, J. A. The Neurobiology of Sleep: Genetics, cellular physiology and subcortical networks. *Nat Rev Neurosci* **3**, 591–605 (2002).
220. Gordon, A. M., Mendes, W. B. & Prather, A. A. The Social Side of Sleep: Elucidating the Links Between Sleep and Social Processes. *Curr Dir Psychol Sci* **26**, 470–475 (2017).
221. Williams, S. & Wolf-Meyer, M. Longing for sleep: Assessing the place of sleep in the 21st century. *Somatosphere: Science, Medicine and Anthropology* **19**, 19–22 (2013).
222. Catterall, W. A., Kalume, F. & Oakley, J. C. NaV1.1 channels and epilepsy. *The Journal of Physiology* **588**, 1849–1859 (2010).

APPENDIX A. SUPPLEMENTAL MATERIALS

Table S1. List of Signal Features used by SIESTA

Feature	Description
ECoG_delta	Relative power of the 0.5-4 Hz frequency band
ECoG_delta_energy	Energy of the 0.5-4 Hz frequency band
ECoG_delta_amp	Mean amplitude of the signal in the 0.5-4 Hz frequency band
ECoG_thetacon	Relative power of the 4-12 Hz frequency band
ECoG_thetaenergy	Energy of the 4-12 Hz frequency band
ECoG_thetaenergy	Energy of the 4-12 Hz frequency band
ECoG_thetamp	Mean amplitude of the 4-12 Hz frequency band
ECoG_theta1	Relative power of the 6-9 Hz frequency band
ECoG_theta1_energy	Energy of the 6-9 Hz frequency band
ECoG_theta1amp	Mean amplitude of the 6-9 Hz frequency band
ECoG_theta2	Relative power of the 5.5-8.5 Hz frequency band
ECoG_theta2energy	Energy of the 5.5-8.5 Hz frequency band
ECoG_theta2amp	Mean amplitude of the 5.5-8.5 Hz frequency band
ECoG_theta3	Relative power of the 7-10 Hz frequency band
ECoG_theta3energy	Energy of the 7-10 Hz frequency band
ECoG_theta3amp	Mean amplitude of the 7-10 Hz frequency band
ECoG_beta	Relative power of the 20-40 Hz frequency band
ECoG_betaenergy	Energy of the 20-40 Hz frequency band
ECoG_betamp	Amplitude of the 20-40 Hz frequency band
ECoG_alpha	Relative power of the 8-13 Hz frequency band
ECoG_alphaenergy	Energy of the 8-13 Hz frequency band
ECoG_alphaamp	Mean amplitude of the 8-13 Hz frequency band
ECoG_sigma	Relative power of the 11-15 Hz frequency band
ECoG_sigmaenergy	Energy of the 11-15 Hz frequency band
ECoG_sigmaamp	Mean amplitude of the 11-15 Hz frequency band
ECoG_spindle	Relative power of the 12-14 Hz frequency band, the frequency at which NREM sleep spindles typically occur
ECoG_spindlenergy	Energy of the 12-14 Hz frequency band
ECoG_spindleamp	Mean amplitude of the 12-14 Hz frequency band
ECoG_gamma	Relative power of the 35-45 Hz frequency band
ECoG_gammaenergy	Energy of the 35-45 Hz frequency band

ECoG_gammaamp	Mean amplitude of the 35-45 Hz frequency band
ECoG_ECoGrel1	Ratio of the relative power value of the thetacon band (4-12 Hz) to that of the delta band (0.5-4 Hz)
ECoG_ECoGrel2	Ratio of the relative power value of the 0.5-20 Hz band to that of the 0.5-50 Hz band
ECoG_Spindlehan	Ratio of 11-16 Hz power to 0.5-40 Hz power smoothed with a 12-point Hanning filter
ECoG_spectral_edge	90% spectral edge of the ECoG signal
ECoG_spectral_mean50	50% spectral mean of the ECoG signal
ECoG_zerocross	Counts of the number of times the amplitude of the ECoG signal falls above or below the mean of the ECoG signal for any given epoch
ECoG_maxs	Maximum amplitude of the raw ECoG signal
ECoG_peaktopeak	Peak-to-peak amplitude of the ECoG signal
ECoG_arv	Arithmetic mean of the absolute values of the ECoG signal in a given epoch
ECoG_rms	Root mean square value of the ECoG signal
ECoG_amplitude	Mean amplitude of the ECoG signal in a given epoch
ECoG_amplitude_m	Median amplitude of the ECoG signal in a given epoch
ECoG_signal_var	Spectral variance of the ECoG signal
ECoG_skew	Skewness of the ECoG signal
ECoG_kurt	Kurtosis of the ECoG signal
ECoG_spectral_mean	Mean of the spectral power distribution of the ECoG signal for a given epoch
ECoG_spectral_entropy	Entropy of the spectral power distribution of the ECoG signal for a given epoch
EMG_amplitude	Mean amplitude of the EMG signal
EMG_signal_var	Spectral variance of the EMG signal
EMG_skew	Skewness of the EMG signal
EMG_kurt	Kurtosis of the EMG signal
EMG_spectral_mean	Mean of the spectral power distribution of the EMG signal for a given epoch
EMG_spectral_entropy	Entropy of the spectral power distribution of the EMG signal for a given epoch
EMG_amplitude_m	Median amplitude of the EMG signal in a given epoch

Table S2. Classification accuracy by algorithm and method when data are scaled.

Algorithm	**One-step**	**Hierarchical**	**Hierarchical + 2 h**
Logistic Regression	0.80 (0.067)	0.47 (0.005)	0.87 (0.054)
Linear Discriminant Analysis	0.77 (0.090)	0.86 (0.056)	0.86 (0.056)
K-nearest Neighbors Classifier	0.76 (0.075)	0.48 (0.003)	0.85 (0.050)
Decision Tree Classifier	0.72 (0.078)	0.80 (0.061)	0.80 (0.063)
Gaussian Naïve Bayes	0.52 (0.18)	0.48 (0.005)	0.79 (0.136)
Passive Aggressive Classifier	0.68 (0.11)	0.46 (0.033)	0.84 (0.052)
Ridge Classifier	0.70 (0.083)	0.78 (0.072)	0.78 (0.072)
Logistic Regression, cross validation	0.79 (0.077)	0.48 (0.005)	0.87 (0.062)
Bernoulli Naïve Bayes	0.60 (0.163)	0.48 (0.009)	0.69 (0.123)
Nearest Centroid	0.54 (0.129)	0.39 (0.012)	0.79 (0.125)
Random Forest Classifier	0.80 (0.063)	0.86 (0.051)	0.86 (0.051)
Ada Boost Classifier	0.79 (0.066)	0.86 (0.05)	0.86 (0.050)
Bagging Classifier, Decision Tree	0.79 (0.06)	0.85 (0.054)	0.85 (0.054)
Extra Trees Classifier	0.80 (0.062)	0.86 (0.053)	0.86 (0.053)
Gradient Boosting Classifier	0.83 (0.057)	0.86 (0.051)	0.86 (0.051)
Perceptron	0.72 (0.11)	0.40 (0.01)	0.84 (0.057)
Bagging Classifier, Random Forest	**0.83 (0.060)**	**0.87 (0.055)**	**0.87 (0.055)**
Bagging Classifier, Extra Trees	0.82 (0.061)	0.87 (0.052)	0.87 (0.054)

Note. Values represent F_1 scores (with standard deviation in parentheses). Hierarchical + 2 hrs. methods refers to the method of adding two hours of manually scored data to the training set for the classifier. All data are scaled.

Table S3. Training time (min) for each classification method using the complete training dataset.

Method	**One-step Training**	**Hierarchical classifier Training (min)**		
		Sleep vs. Wake	**NREM vs. REM**	**Total**
Logistic Regression	1.31	0.57	0.57	1.13
Linear Discriminant Analysis	0.33	0.44	0.44	0.88
K-nearest Neighbors Classifier	0.49	0.65	0.65	1.29
Decision Tree Classifier	3.09	3.36	3.36	6.73
Gaussian Naïve-Bayes	0.17	0.22	0.22	0.45
Passive Aggressive Classifier	0.33	0.25	0.25	0.51
Ridge Classifier	0.22	0.28	0.28	0.57
Logistic Regression with cross-validation	2.73	1.11	1.11	2.22
Bernoulli Naïve Bayes	0.19	0.24	0.24	0.48
Nearest Centroid	0.14	0.19	0.19	0.38
Random Forest Classifier	2.20	2.46	2.46	4.92
Ada Boost Classifier	13.15	15.80	15.80	31.60
Bagging Classifier with decision tree	19.74	20.96	20.96	41.91
Extra Trees Classifier	0.65	0.78	0.78	1.55
Gradient Boosting Classifier	50.63	19.06	19.06	38.11
Perceptron	0.41	0.22	0.22	0.44
Bagging Classifier with Random Forest Classifier	**14.22**	**13.94**	**13.94**	**27.89**
Bagging Classifier with Extra Trees Classifier	4.15	4.19	4.19	8.38

Table S4. Scoring time (min) of the BCRF algorithm on subsets of the training and scoring data.

Database	Classifier	Sleep Stages	Train Time	Score Time (24 h)	Train Time with 2 h of manual score	Score Time of 24 h with 2 h of manual score	Database
WT	One-Step	Awake/NREM /REM	0.82	0.01			
DS	One-Step	Awake/NREM /REM	0.35	0.01	-	-	-
WT+DS	One-Step	Awake/NREM /REM	1.53	0.01	-	-	-
WT	Hierarchical	Sleep/Awake	0.93	0.01	0.90	0.01	0.90
		NREM/REM	0.97	0.01	0.99	0.01	0.99
		Total	1.90	0.02	1.89	0.02	1.89
DS	Hierarchical	Sleep/Awake	0.28	0.01	0.29	0.01	0.29
		NREM/REM	0.30	0.01	0.33	0.01	0.33
		Total	0.58	0.02	0.62	0.02	0.62
WT+DS	Hierarchical	Sleep/Awake	1.44	0.01	1.43	0.01	1.43
		NREM/REM	1.66	0.01	1.54	0.01	1.54
		Total	3.09	0.02	2.97	0.02	2.97

Table S5. SIESTA performance using data binned in 5-second epochs.

One-Step Approach

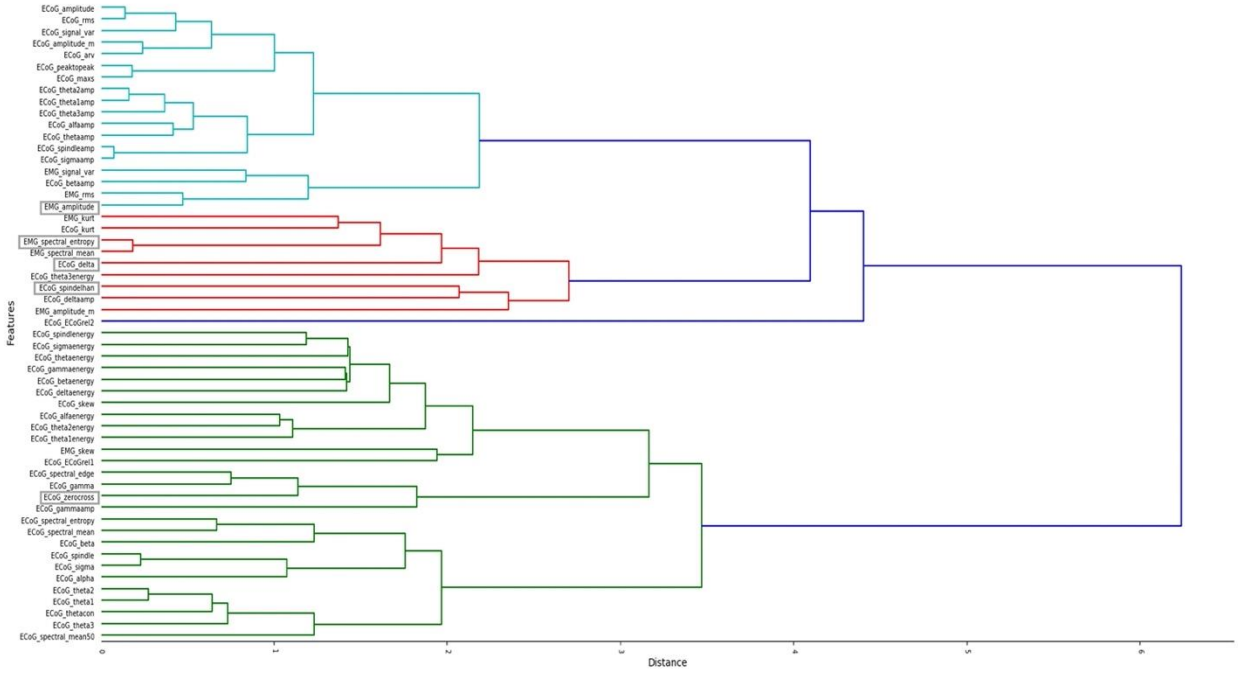
Using complete database to score WT mice		Using complete database to score DS mice	
Awake	0.94	**Awake**	0.91
NREM	0.90	**NREM**	0.87
REM	0.75	**REM**	0.83

Hierarchical Approach

Using complete database to score WT mice		Using complete database to score DS mice	
Awake	0.93	**Awake**	0.93
Sleep	0.91	**Sleep**	0.91
NREM	0.91	**NREM**	0.84
REM	0.76	**REM**	0.85

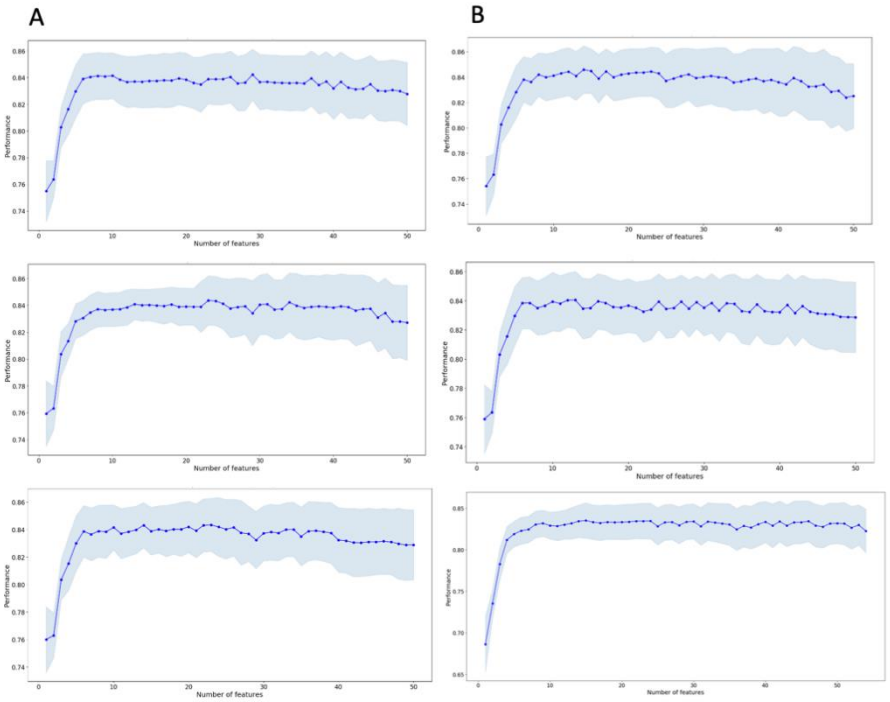
Note. These tests use the complete dataset (WT n=14 , DS n=6) to score one mouse of each genotype.

Figure S1. Cluster dendrogram of the pair-wise distance of the features from the complete dataset



Note. In light blue, red and light green are the branches of the identified clusters. The most important features identified by SFS and SBS are boxed in grey on the y-axis.

Figure S2. Sequential feature selection test using the BCRF algorithm trained on the complete dataset, with a One-Step approach.



C

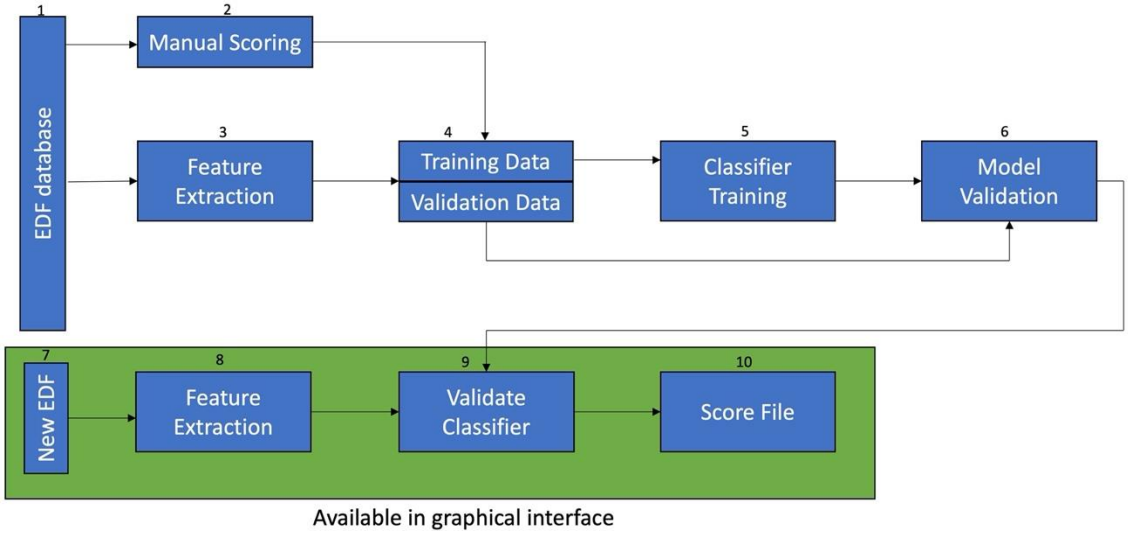
Sequential Forward Selection		
ECoG_zerocross	ECoG_zerocross	ECoG_zerocross
ECoG_spindelhan	EMG_amplitude	ECoG_spindelHan
EMG_amplitude	ECoG_spindelhan	ECoG_spectral_mean
ECoG_delta	ECoG_delta	ECoG_delta
EMG_spectral_entropy	EMG_spectral_entropy	EMG_sepctral_entropy
ECoG_thetacon	ECoG_beta	ECoG_theta3
ECoG_spectral_edge	EMG_signal_var	ECoG_spectral_edge
ECoG_deltaenergy	ECoG_spectral_edge	ECoG_theta2

D

Sequential Backward Selection.		
ECoG_zerocross	ECoG_zerocross	ECoG_zerocross
EMG_amplitude	ECoG_spindelhan	EMG_amplitude
ECoG_spindelhan	ECoG_theta3amp	ECoG_spindelhan
ECoG_delta	EMG_spectral_entropy	ECoG_delta
EMG_spectral_entropy	ECoG_rms	ECoG_theta3amp
ECoG_theta2	ECoG_sigma	ECoG_theta2
ECoG_betaenergy	ECoG_theta2amp	ECoG_theta3
ECoG_sigmaenergy	ECoG_signal_var	ECoG_sigmaamp

Note. The blue dots are the mean performance, and the light blue shaded area is the STD in each case. **A)** Three repetitions of the Sequential Forward Selection. **B**) Three repetitions of the Sequential Backward Selection. **C)** Table with the top 8 features of each run of the Sequential Forward Selection in order of importance. **D**) Table with the top 8 features of each run of the Sequential Backward Selection in order of importance. Features highlighted in green occurred in 2 or more runs of any Sequential Selection Algorithm.

Figure S3. Overview of the SIESTA workflow.



Note. *To* train the model, we start with raw ECoG/EMG recordings in EDF file format **(1)**. Our dataset is manually scored by one of four human experts **(2)**. From data in the same EDF files, we calculated features used for scoring **(3).** Next, we split the data into training and validation sets **(4).** Using only the training set, we trained the BCRF classifier and obtain our model **(5).** We then evaluate model performance on the validation set **(6).** When a user wants to score data from their own experiment, they first upload their raw recording in EDF format **(7).** Feature extraction occurs as in Step 2, which generates a timestamped file containing feature values for every epoch in the recording. **(8).** Next, users can load the working model from our dataset. They will be informed of the validation results of the model as well as information about the underlying dataset (number and genotype of the animals included) **(9).** Users will also have the option to train a new classifier on their own manually scored dataset or combine their data with ours. Finally, users can upload the newly generated feature data to create a final output file containing timestamped scores and several metrics commonly used in sleep analysis, including ECoG delta power and theta power **(10).**

VITA

Asad I. Beck was born in San Diego, California, on August 26, 1996, the sixth child of seven. They attended California Virtual Academy for all their primary and secondary schooling and graduated in June 2014. In the following August, they enrolled in San Diego State University where they were accepted into the Advancing Diversity in Aging Research (ADAR) program, through which they would work in the Vietnam Era Twin Study of Aging (VETSA) lab at University of California, San Diego, under Drs. Carol Franz and Bill Kremen. In June 2018, they graduated with a Bachelor of Arts in Psychology, earning summa cum laude honors, and completed double minors in Philosophy and in English. They were accepted into the Graduate Program in Neuroscience at the University of Washington for the 2018 cohort, where they rotated in and later joined the lab of Horacio de la Iglesia. They received a Doctor of Philosophy in Neuroscience in December 2023. After graduating, they plan to expand their baking repertoire and finish the solo album they have been working on for the past two years.