This chapter was originally published in the book *International Review of Neurobiology (Volume 93)*. The copy attached is provided by Elsevier for the author’s benefit and for the benefit of the author’s institution, for non-commercial research, and educational use. This includes without limitation use in instruction at your institution, distribution to specific colleagues, and providing a copy to your institution’s administrator.

All other uses, reproduction and distribution, including without limitation commercial reprints, selling or licensing copies or access, or posting on open internet sites, your personal or institution’s website or repository, are prohibited. For exceptions, permission may be sought for such use through Elsevier's permissions site at:

http://www.elsevier.com/locate/permissionusematerial

From Christian Cajochen, Sarah Chellappa and Christina Schmidt,
What Keeps Us Awake?—the Role of Clocks and Hourglasses, Light, and Melatonin.
In: Angela Clow and Lisa Thorn editors,
International Review of Neurobiology (Volume 93).
ISBN: 978-0-12-381324-4
© Copyright 2010 Elsevier Inc.
Academic Press.
WHAT KEEPS US AWAKE?—THE ROLE OF CLOCKS AND HOURGLASSES, LIGHT, AND MELATONIN

Christian Cajochen*, Sarah Chellappa*, and Christina Schmidt*

*Center for Chronobiology, Psychiatric Hospital of the University of Basel, CH-4012 Basel, Switzerland

I. Introduction

II. Circadian and Homeostatic Impetus for Wakefulness
   A. Timing and Consolidation of the Human Sleep–Wake Cycle: from Basic Arousal States to Controlled Cognitive Behavior
   B. Brain Circuitry Underlying Circadian and Homeostatic Influences on Human Cognition: A Possible Scenario

III. Effects of Light on Human Wakefulness
   A. Light Switches on the Clock and the Hourglass
   B. Alerting Effects of Light
   C. Dose- and Wavelength Response Relationship of Light Exposure on Alertness
   D. Neuroanatomical Underpinnings of the Effect of Light on Alertness and Cognitive Performance
   E. Non-Clinical Applications of Light

IV. Effects of Melatonin on Human Sleep and Wakefulness
   A. Endogenous Melatonin and the Human Circadian Sleep–Wake Cycle
   B. Effects of Exogenous Melatonin on Human Sleep and Wakefulness
   C. Implications for the Treatment of Insomnia and Circadian Rhythm Disorders

References

What is it that keeps us awake? Our assumption is that we consciously control our daily activities including sleep–wake behavior, as indicated by our need to make use of an alarm clock to wake up in the morning in order to be at work on time. However, when we travel across multiple time zones or do shift work, we realize that our intentionally planned timings to rest and to remain active can interfere with an intrinsic regulation of sleep/wake cycles. This regulation is driven by a small region in the anterior hypothalamus of the brain, termed as the “circadian clock”. This clock spontaneously synchronizes with the environmental light–dark cycle, thus enabling all organisms to adapt to and anticipate environmental changes. As a result, the circadian clock actively gates sleep and wakefulness to occur in synchrony with the light–dark cycles. Indeed, our internal clock is our best morning alarm clock, since it shuts off melatonin production and...
boosts cortisol secretion and heart rate 2–3 h prior awakening from Morpheus arms. The main reason most of us still use artificial alarm clocks is that we habitually carry on a sleep depth and/or the sleep–wake timing is not ideally matched with our social/work schedule. This in turn can lead hourglass processes, as indexed by accumulated homeostatic sleep need over time, to strongly oppose the clock. To add to the complexity of our sleep and wakefulness behavior, light levels as well as exogenous melatonin can impinge on the clock, by means of their so-called zeitgeber (synchronizer) role or by acutely promoting sleep or wakefulness. Here we attempt to bring a holistic view on how light, melatonin, and the brain circuitry underlying circadian and homeostatic processes can modulate sleep and in particular alertness, by actively promoting awakening/arousal and sleep at certain times during the 24-h day.

I. Introduction

Despite the fact that humans have invented technologies such as artificial light and online services that allow us to do a certain activity at obviously any time, only a fraction of the humankind is involuntarily awake at night and sleeps during the light phase of the 24-h cycle. This natural synchrony in behavioral states among humans is also surprising because we think that we consciously plan our individual daily activities and thus our bed and wake-up times. There are certainly considerable interindividual and intercultural differences in the timing of sleep and wakefulness (e.g., chronotypes), but as to our knowledge there are no night-active human ethnic groups or cultures. This obviously points to a clear biological basis and an evolutionary adaptive behavior favoring a day-active human species. The neuroanatomical basis of the biological underpinnings of the daily (circadian) regulation of sleep–wake rhythms has been unraveled in the past century, but their physiological functions and implications on our health are still being intensely explored. Thus, how daily rhythms of behavioral states are controlled is an active area of current research. Given its relevance to human health, well-being, and cognitive performance, this is an important challenge to solve, particularly based on the fact that more and more people are forced to be awake at inappropriate or at biologically non-optimal times during shiftwork.

In order to assess the effect of any stimuli either from the environment (e.g., light) or from the body itself (e.g., endogenous melatonin) on the regulation of awakening, a good insight of factors, which regulate sleep and wakefulness, is needed. Sleep and wakefulness are controlled by two primary factors: the
WHAT KEEPS US AWAKE?

The circadian clock and the intrinsic need for sleep reflected in the homeostatic properties of sleep and determined by recent sleep–wake history. In Section II, we describe the neural mechanisms by which the circadian clock influences the sleep–wake system. In particular, we attempt at providing a better grasp of the physiological functions of the circadian clock and their relation to correlates of sleep intensity and its role of actively gating awakening/arousal and sleep at certain times during the 24-h day. We have new evidence from recent electro-physiological and functional magnetic resonance imaging (fMRI) data, to propose a potential brain circuitry underlying circadian and homeostatic influences on human alertness and cognition.

Environmental conditions (e.g., light, sound, temperature, social stimuli) play an important role in the control of sleep and wakefulness as well as their intensity and quality (i.e., spectral composition) respectively. Light is certainly the most regularly occurring stimulus in the environment. The challenge of a daily change of the light–dark (LD) cycle has profound impact on a wide range of biological functions and behavior. Thus, light exerts powerful non-visual effects. In humans, light is intuitively linked with an alert or wakeful state. On the other hand, closing the eyelids or dimming or turning off the lights has a very powerful soporific (i.e., sleep inducing) effect, particularly in children, sleep deprived adults and older people. Compared to the effects of light on human circadian rhythms, little attention has been paid to its acute alerting action. In Section III, we summarize studies from the past two decades, which have defined and quantified the dose (illuminance levels), exposure duration, timing and wavelength of light needed to evoke circadian and/or alerting responses in humans, as well as their temporal relationship to light-induced changes in endocrinological and electrophysiological sequelae of alertness. Furthermore, neuroanatomical and neurophysiological findings from animal and human studies elucidating a potential role of light in the regulation of sleep/wake states and its repercussion on cognitive performance are discussed. A brief outlook of promising non-clinical applications of lights’ alerting properties will be given, and its involvement in the design of more effective lighting at home and in the workplace will be considered.

The pineal hormone melatonin is probably the most light-sensitive hormone in humans and also in other organisms, so that measuring the 24-h profile of endogenous melatonin levels provides accurate information about the prior light history of an individual. The phase, amplitude, and duration of the active phase of melatonin secretion are all important measures to assess whether somebody has delayed or advanced circadian rhythms or whether somebody lives in dim or brightly lit environments. Thus, there is an intimate transduction of the LD cycle reflecting external time to the endogenous “melatonin cycle” reflecting internal time. Humans are more light sensitive in terms of melatonin suppression than previously thought. Light intensities as low as 40 lux are sufficient to attenuate the
evening increase of melatonin secretion when the light source yields predominance in the blue range of its spectral composition. Interestingly, there is a tight and significant correlation between light’s melatonin suppressing effect and its alerting response, leading some researchers to the speculation that melatonin could act as an internal sleep facilitator. Thus, possible roles of endogenous melatonin in the regulation of sleep and wakefulness are being discussed in Section IV. Furthermore, the use of exogenous melatonin and newly available melatonin agonists to treat sleep disorders such as sleep onset insomnia or premature awakening from sleep are also dealt with in Section IV.

II. Circadian and Homeostatic Impetus for Wakefulness

“There is no animal which is always awake or always asleep, such that all sleep is susceptible of awakening and all wake time beyond the natural time limit is susceptible to sleep” (Aristotle, *On Sleep and Sleeplessness*, 350 BCE). Living organisms are permanently exposed to internal and external changes and the combined action of these dynamics may determine the transition between conscious-controlled to unconscious-automated behavioral states (Tononi and Edelman, 1998). Behavioral or perceptual states continuously vary between the extremes, with on the one hand resting sleep during which consciousness is strongly attenuated and on the other hand a state of wakefulness when we actively interact with the environment, and during which we engage in many cognitive and other activities (Dijk and Archer, 2009). It is nowadays largely accepted that in human beings, homeostatic and circadian sleep–wake regulatory processes are continuously working in harmony or in opposition to each other to allow maintenance of behavioral states such as sleep and wakefulness at appropriate time points within the 24-h LD cycle. However, these states per se seem far from being unitary concepts since their consolidation is achieved by the mutual interaction of multiple brain processes. Even though the interplay between regulatory processes aspires to stability within a given state, there exist fine grained fluctuations in the way we perceive our environment over the waking state. Such slight differences may be exaggerated by inter-individual differences in the orchestration of the underlying processes. A good example of such fluctuations is the discovery by Bodenhausen (1990), who observed that human subjects exhibit stereotypic biases in their judgments to a much greater extent when these were rendered at a time of day reflecting reduced arousal levels for them. Judgments were significantly more affected by stereotypic beliefs in the morning hours for evening types and in the evening hours for morning types. Thus, the quality of judgment fluctuated within the state of wakefulness, which itself showed a differential temporal pattern across the 24-h day in morning and evening types.
WHAT KEEPS US AWAKE?

A. TIMING AND CONSOLIDATION OF THE HUMAN SLEEP–WAKE CYCLE: FROM BASIC AROUSAL STATES TO CONTROLLED COGNITIVE BEHAVIOR

As mentioned above, sleep and wakefulness are periodically occurring at specific times of the 24-h LD cycle. Their consolidation is achieved by the interplay between circadian and homeostatic oscillators, initially conceptualized in the two-process model of sleep and wake regulation (Borbely, 1982; Daan et al., 1984). The homeostatic process represents an hourglass process steadily building up with increasing time awake and exponentially declining during sleep. The circadian process reflects an endogenous, nearly 24 h variation in the propensity for sleep and wakefulness and was originally assumed to be independent of the homeostatic process (i.e., the amount of elapsed time awake) (Borbely, 1982; Daan et al., 1984). This process originates in the suprachiasmatic nuclei (SCNs) of the anterior hypothalamus, an anatomical structure supporting numerous periodic biological functions and considered as the circadian master clock in most living organisms. Findings acquired under a variety of experimental conditions (e.g., internal desynchronization of the sleep–wake cycle, forced desynchrony paradigms, fragmented sleep–wake cycles, sleep deprivation, sleep displacement) point in a remarkably consistent way to the existence of a powerful and active drive for wakefulness at the end of the habitual waking day in humans (Lavie, 2001). Thus, the circadian master clock is tuned such that peak arousal levels in humans are generated in the early evening hours, just before opening the gate for sleep. Accordingly, this time window is characterized by maximal circadian wake promotion and has been called the wake maintenance zone by Strogatz and colleagues (1987). While the endogenous scheduling of the wake maintenance zone to the end of the habitual waking day seems paradoxical at first sight, it takes all sense when one considers it in combination to the temporal evolution of the homeostatic process throughout the habitual 24-h sleep–wake cycle. For instance, it is the very high circadian-based propensity for wakefulness that prevents us falling asleep early in the evening hours when homeostatic sleep pressure is at its highest level and maximally promotes sleep. Thus, during the latter part of the normal waking day, circadian and homeostatic systems work in opposition to ideally ensure a consolidated period of wakefulness. Edgar et al. (1993) have first conceptualized this opponent action based on the framework of the two-process model and data acquired in diurnal squirrel monkeys. SCN-lesioned squirrel monkeys significantly increased total sleep time, which was associated with a 15-fold reduction in the length of wake bouts during the subjective day and no changes in the length of the wake bouts during the subjective night, leading the investigators to suggest that the circadian clock is actively involved in the promotion of wakefulness, by opposing the homeostatic accumulated drive for sleep. Results from human forced desynchrony studies have confirmed the above-mentioned model (Dijk and Czeisler, 1994, 1995) by showing the paradoxical positioning of the circadian alertness peak just before habitual sleep.
time, as indexed by longest sleep latencies and highest amounts of wakefulness within scheduled sleep episodes at this time of the day. Likewise, the SCN also promotes sleep (i.e., circadian increase in sleep tendency) as the biological night progresses (Dijk and Czeisler, 1994, 1995) counteracting the decrease in sleep propensity associated with accumulated sleep, thus allowing us to maintain a consolidated 8-h sleep episode.

Besides sleep and wakefulness, neurobehavioral efficiency seems to be affected by the same paradoxical interplay of circadian and homeostatic sleep–wake promotion over the 24-h cycle such that the wake-dependent deterioration is minimal during the wake-maintenance zone. Data gathered in a constant routine paradigm, which challenged homeostatic sleep pressure conditions by either sleep depriving or sleep satiating study volunteers by regular nap opportunities throughout the circadian cycle, indicate a clear circadian modulation of cognitive performance and subjective sleepiness even in the absence of prominent homeostatic sleep pressure (Fig. 1). This circadian modulation is temporally organized such that neurobehavioral performance (alertness scores and performance lapses) is maximally boosted in the late evening hours. Under sleep deprivation conditions (>16 h of enforced wakefulness), a steep decline on neurobehavioral performance can be observed when the testing is extended into the biological night, i.e., just after the circadian arousal signal has turned off. However, as illustrated in Fig. 1, neurobehavioral performance does not decline linearly with increasing time awake throughout 40 h of sustained wakefulness, but shows a strong improvement coinciding with the biological day, when circadian arousal promotion kicks in again (see also Cajochen et al., 1999b, 2004; Graw et al., 2004; Horowitz et al., 2003).

Importantly, compelling data from forced desynchrony studies indicate that circadian and homeostatic processes do not simply add up to characterize daily alertness and performance modulations. It has been observed that the amplitude of the observed circadian modulation in performance depends on homeostatic sleep pressure, such that increasing homeostatic sleep pressure attenuates circadian wake promotion during the subjective evening hours (Dijk and Archer, 2009). Hence, minor changes in the specific interplay between both processes lead to significantly disrupted stability patterns in cognitive states even throughout a normal waking day. This may explain why a series of studies found significant performance fluctuations in cognitive behavior throughout a normal waking day in morning and evening chronotypes differing in circadian and homeostatic sleep–wake regulatory processes throughout the course of a normal waking day (see Schmidt et al., 2007 for a review). Such interindividual differences have recently been used as a tool in order to investigate the functional neuroanatomy subtending modulatory effects of sleep–wake regulation on higher order human behaviors. We will briefly describe these observations within the context of the brain circuitry involved in the circadian control for states of sleep and wakefulness.
Fig. 1. Dynamics of subjective sleepiness on the Karolinska Sleepiness Scale (KSS), objective vigilance on the Psychomotor Vigilance Task (PVT), and core body temperature (CBT) across a 40 h SD (high sleep pressure; filled circles) and NAP protocol (low sleep pressure; open circles). The upper two panels indicate the timing of the naps (black bars) and scheduled episodes of wakefulness (white bars) respectively for the SD and NAP protocol. Data are plotted against the midpoint of the time intervals. Relative clock time represents the average clock time at which the time intervals occurred. Modified from Cajochen et al. (2001).

B. BRAIN CIRCUITRY UNDERLYING CIRCADIAN AND HOMEOSTATIC INFLUENCES ON HUMAN COGNITION: A POSSIBLE SCENARIO

How circadian oscillations in the SCN as well as circuits controlling for states of sleep and wakefulness interact at the cerebral level in order to regulate arousal and cognitive behavior is still an open question. Output of the SCN indirectly reaches target areas implicated in the regulation of sleep and wakefulness (ventro-lateral-preoptic area (VLPO), tuberomammillary nucleus (TMN), lateral hypothalamus (LH), thalamus, and brainstem nuclei via its connections to the
dorsal medial hypothalamus (DMH) (Mistlberger, 2005). Concomitantly, diffuse monoaminergic activating systems are under circadian control and impinge on many thalamo-cortical areas, suggesting that the interaction with sleep homeostasis could take place at many different levels (Dijk and Archer, 2009). Work by Aston-Jones and colleagues (Aston-Jones, 2005; Aston-Jones et al., 2001) has shown that the noradrenergic locus coeruleus (LC) system plays an important role in the circadian regulation of alertness. Within the framework of their model, the SCN indirectly communicates with the LC via projections to the dorsomedial hypothalamic nucleus (DMH). Evidence for that comes from neurophysiological experiments, which revealed circadian variations in LC impulse activity and showed that lesions of the DMH eliminated these circadian changes in LC activity, suggesting a functional significance of the SCN–DMH–LC circuit (Gompf and Aston-Jones, 2008). Through LC activity with its widespread thalamic and cortical connections, this pathway may control a variety of central nervous system functions related to noradrenergic innervations, including alertness and vigilance, and also higher-order cognitive processes. We have recently collected indirect evidence that the circadian arousal signal generated by this circuitry is modulated by homeostatic sleep pressure (Schmidt et al., 2009). In this study, the interaction between these processes at the cerebral level was investigated in chronotypes differing in circadian and homeostatic sleep–wake regulatory processes under normally entrained day–night conditions (Baehr et al., 2000; Bailey and Heitkemper, 2001; Kerkhof, 1991; Kerkhof and Van Dongen, 1996; Mongrain et al., 2004, 2006a, 2006b). Extreme morning and evening chronotypes were examined at different time points within a normal waking day, while performing a sustained attention task in an fMRI environment. The main results of this study are summarized in Fig. 2.

In agreement with previous studies (Kerkhof, 1991; Mongrain et al., 2006a, 2006b; Taillard et al., 2003), we observed that even when the timing of the scheduled testing session was adapted to the specific sleep–wake schedule of the volunteers, morning-type individuals presented higher increases in homeostatic sleep pressure at the end of a normal waking day, as indexed by slow-wave activity (SWA) at the beginning of the night. This effect was paralleled by higher subjective sleepiness and lower objective vigilance levels in the morning than evening types during the evening hours. Interestingly, the fMRI results revealed that maintenance of optimal sustained attention performance in the subjective evening hours was associated with higher cerebral activity in evening than morning chronotypes in a brainstem region compatible with the LC and in an anterior hypothalamic region putatively encompassing the suprachiasmatic area (SCA). Thus, in agreement with the brain circuitry proposed by Aston-Jones and colleagues, our data suggest that activity in these regions contributes to circadian wake promotion in the subjective evening hours. Importantly, we further observed that activity in the SCA decreased with increasing homeostatic sleep pressure, suggesting a direct influence of homeostatic and
WHAT KEEPS US AWAKE?

Morning types
Evening types

Fig. 2. (A) Exponential decay function adjusted on relative SWA in sleep cycles (NREM sleep) measured from the central frontal derivation for all-night EEG of the night preceding the evening scan acquisition (red line: morning types; blue line: evening types. (B) Increased task-related response in the dorsal pontine tegmentum and the anterior hypothalamus, compatible with the LC and SCA, respectively, in evening as compared to morning chronotypes during the subjective evening for optimal sustained attention during the performance of a Psychomotor Vigilance Task. Functional results are displayed at p < 0.001, uncorrected threshold, over the mean normalized structural MRI of the population. Corresponding parameter estimates (arbitrary units) are displayed for event indicators of fast (<percentile 10) reaction times (light grey: morning types; dark grey: evening types). (C) Regression analysis of the relation between estimated BOLD responses during optimal task performance in the SCA region and the amount of SWA during the first sleep cycle in the preceding night (r = 0.54, p < 0.05, n = 27). Crosses: morning types, Triangles: evening types. Modified from Schmidt et al. (2009).

circadian interactions on the neural activity underpinning diurnal variations in human behavior. Our results corroborate findings in the rat, which showed suppression of SCN activity by SWA throughout various vigilance states (Deboer et al., 2003, 2007), and globally speak in favor of the initial assumption that an increase in homeostatic sleep pressure impacts on the circadian wake-promoting signal during the subjective evening hours. Another study used the differential vulnerability to sleep loss according to a polymorphism in the human PER3 clock gene (Viola et al., 2007) to evidence nonlinear interaction patterns between the two basic processes at the cortical level throughout a normal waking day and after a night of sleep deprivation (Vandewalle et al., 2009; see Dijk and Archer, 2009 for a review). In this study, the temporal profile of cortical activity underlying successful performance on an executive task (n-back paradigm) could be tracked by the dynamics predicted by the interplay between circadian and homeostatic processes according to each subject’s specific genotype.

However, the underlying mechanisms by which homeostatic sleep pressure modifies the circadian arousal signal in the evening hours are still unknown. It has been suggested that adenosine is a homeostatic regulator of sleep need (Benington et al., 1995; Landolt et al., 1995; Porkka-Heiskanen et al., 1997; Strecker et al., 2000). During prolonged wakefulness, the energy-producing systems in the brain
run down: brain glycogen reserves are exhausted and ATP levels are depleted. During prolonged wakefulness, as ATP is degraded to ADP, AMP, and eventually adenosine, extracellular adenosine levels rise in some parts of the brain, including the basal forebrain (see Landolt, 2008 for a review). It has been hypothesized that, once adenosine reaches sufficient concentrations after prolonged wakefulness, it has an inhibitory action on the wake-promoting neural circuitry of the basal forebrain and probably activates VLPO neurons by reducing inhibitory Gamma-aminobutyric acid (GABA)ergic inputs. Accordingly, after sleep deprivation, VLPO neurons fire about twice as fast as they do during normal sleep, implying that they are under the influence of homeostatic factors that reflect sleep need (Lu et al., 2002; Saper et al., 2005a; Sherin et al., 1996; Szymusiak et al., 1998). In humans, there is evidence that adenosinergic neurotransmission plays a role in NREM sleep homeostasis. Indeed, a polymorphism in an adenosine-metabolizing enzyme contributes to high interindividual variability in deep SWS duration and intensity (Retey et al., 2005). Furthermore, the adenosine receptor antagonist caffeine has the ability to attenuate electroencephalographic (EEG) markers of NREM sleep homeostasis (Landolt et al., 1995). Accordingly, caffeine administration is effective in counteracting the detrimental performance effects of extended wakefulness (Retey et al., 2006; Wyatt et al., 2004).

To sum up, sleep and wakefulness are determined by the multiple interplay between circadian and homeostatic oscillators. Active circadian wake promotion during the subjective evening hours attempts the achievement of stability of cognitive states throughout a normal waking day, by opposing the increasing homeostatic sleep pressure at this time of the day. Likewise, circadian sleep promotion takes place in the early subjective morning hours to oppose the decreasing homeostatic sleep pressure allowing a consolidated bout of sleep. However, fine-grained interindividual differences in the complex interplay between these processes may result in significant modulations in cognitive behavior even throughout a normal waking day. A couple of studies recently took advantage of such interindividual differences for the investigation of the cerebral correlates underlying circadian and homeostatic influences on human cognition (Schmidt et al., 2007; Vandewalle et al., 2009). Together with data gathered in the animal domain, their results point into the direction that the circadian arousal signal and accumulated homeostatic sleep pressure directly interact at the cerebral level in order to control cognitive behavior throughout wakefulness. In one possible scenario (Fig. 3), the efficacy of the circadian arousal signal, generated by the indirect communication of the circadian master clock to the brainstem LC and thereby to widespread cortical areas, may be modified through adenosine, a putative mediator of sleep homeostasis. Importantly, these assumptions should now be investigated in the framework of protocols more systematically manipulating the interaction between both processes and allowing tracking their interaction throughout the entire 24-h cycle.
WHAT KEEPS US AWAKE?

Circadian wake promotion
Cortical networks regulating cognition
Adenosine (diffuse system)

Fig. 3. A possible simplified scenario of the circadian and homeostatic interplay to regulate alertness and cognitive performance over the 24 h cycle. With increasing time awake, homeostatic sleep pressures accumulates throughout the day and may affect cortical activation by mechanisms including synaptic potentiation in several circuits. This mechanism is tied to homeostatic regulation of sleep slow-wave activity during the night. Information about the amount of accumulated homeostatic sleep pressure is transferred to hypothalamic structures including the circadian master clock which in turn feeds back in a "time-of-day" specific manner by sending signals to wake-promoting brainstem as well as thalamic structures. From there on, the information is transferred to cortical areas in order to allow the maintenance of an adequate cognitive state (Aston-Jones, 2005).

III. Effects of Light on Human Wakefulness

To be of functional significance, circadian rhythms must be entrained to the 24-h LD cycle. Thus, it is not surprising that light plays a powerful role on behavior and physiology. In fact, a change in the timing of the external LD cycle leads to a shift in endogenous phase of circadian rhythms (Brainard et al., 1997). Besides these long-term effects on circadian phase, many acute effects of light have been consistently shown for a wide range of physiological processes, such as hormonal secretion, heart rate, sleep propensity, alertness, body temperature, pupillary constriction, and gene expression (Aalto and Hilakivi, 1986; Badia et al., 1991; Berson, 2003; Cajochen et al., 1992, 1996, 2005, 2006; Lavoie et al., 2003; Munoz et al., 2005). Both long-term and acute effects of light are usually referred to as non-visual (or non-image forming, NIF) effects, since they drift apart from the classical involvement of rod and cone photopigments in the visual responses to light. These NIF responses were firstly demonstrated in mice devoid of classic photoreceptors, since light still had the capacity to elicit circadian phase-shifting responses (Freedman et al., 1999) and melatonin suppression (Lucas et al., 1999). In humans, the fact that visually blind people still exhibit light-induced melatonin suppression
(Czeisler et al., 1995) and that the spectral sensitivity of NIF responses differed from visual responses (Brainard et al., 2001; Thapan et al., 2001) challenged the classical involvement of rod and cone photopigments in responses to light. Furthermore, since Berson and coworkers (2002) detected intrinsic photosensitive retinal ganglion cell (ipRGC) in the retina of mammals, it began to emerge that the eye performs a dual role in detecting light for a range of behavioral and physiological responses distinct from the classical visual responses. Melanopsin-containing ipRGCs have a specialized non-visual retino-hypothalamic tract which provides direct neuronal connection to the SCN, as well as direct and indirect (via SCN) projections to brain areas implicated in the regulation of arousal (Gooley et al., 2003). Furthermore, the SCN has connections to the pineal gland, which is responsible for the regulation of melatonin, as well as to many areas that share an input from the visual photoreceptor system, such as the lateral geniculate nucleus, pretectum and superior colliculus (Lockley and Gooley, 2006). The brain areas implicated in the non-visual effects of light beyond these ipRGC projections are still unknown. Nevertheless, if one considers the number of brain areas that are just one synapse away from ipRGCs, and the numerous projections of just one key target of ipRGCs, the SCN, it becomes evident that non-visual responses to light could affect many brain functions, including cognitive functions.

In this section, we will address the following points: (1) how light (timing, dose, and wavelength) impinges on human wakefulness; (2) how light modulates cognition, in particular in tasks associated with sustained attention, and (3) the importance of lights effect in non-clinical settings.

A. Light Switches on the Clock and the Hourglass

Even in the absence of an LD cycle, the rest-activity rhythm persists with a periodicity of approximately 24 h, instead of redistributing across the 24-h day. The synchronization to LD cycles is obtained through variation in the response of the circadian pacemaker in the SCN to light pulses, whereby light exposure late in the biological day delays sleep onset in humans, while exposure early in the biological day (dawn) advances activity onset (Czeisler and Gooley, 2007). Thus, light acts as a synchronizer (Zeitgeber) by transmitting the information about external time (LD cycle) to the organisms’ internal timing system and as consequence marginally influences the switch between behavioral states such as sleep and wakefulness.

Wakefulness requires a certain alertness level to actively interact with the environment. Thus, alertness is a construct associated with high levels of environmental awareness, which can be operationalized through many converging measurements, including subjective responses, behavior, and brain activity (Buysse et al., 2003). Alertness is associated with self-reported high levels of wakefulness and low
levels of fatigue, short response times, fast and more accurate cognitive performance, and lower levels of theta activity (4.75–7.75 Hz) in the electroencephalogram (EEG), particularly in the frontal cortex (Badia et al., 1991; Cajochen et al., 1995, 1999a; Daurat et al., 2000). Subjective perception of alertness heavily depends on time-of-day, to the extent that the circadian modulation of alertness has a strikingly similar temporal association with the circadian rhythm of core body temperature (CBT) with its maximum in the evening and nadir in the early morning (see also Section III; Kleitman, 1987). Considering the temporal dynamics of these processes on alertness, one can hypothesize that light exerts its alerting effects most strongly when the circadian drive for sleep is at its maximum (i.e., in the early morning at the CBT minimum) and under high homeostatic sleep pressure conditions (i.e., after more than 16 h of wakefulness). However, besides the temporal occurrence of a light pulse relative to the circadian and homeostatic system, factors such as the intensity of light, light stimulus duration, and its wavelength play a crucial role in determining the impact on alertness and cognitive performance.

B. ALERTING EFFECTS OF LIGHT

The vast majority of light studies have been conducted at night (Badia et al., 1991; Cajochen et al., 2000; Campbell and Dawson, 1990; Foret et al., 1996; Lockley et al., 2006) during a time when one would expect most pronounced alerting effects in humans. Indeed light at night significantly enhances subjective alertness and reduces objective markers of sleepiness, such as EEG theta activity and the incidence of slow-eye movements as assessed by the electrooculogram (EOG). However, also during the biological day, when melatonin is at minimal level, light does impact alertness. In an “in-lab” study, individuals who were exposed to polychromatic white light with levels >7000 lux for 20 min during daytime exhibited an enhancement in cortical activity during an oddball task and subjective alertness improved in a dynamic manner, such that these alerting effects declined within minutes after the end of the light stimulus, following various region-specific time courses, such as enhanced responses in the posterior thalamus, including the pulvinar nucleus, which has been implicated in visual attention and alertness regulation (Vandewalle et al., 2006). This suggests that light may modulate activity of subcortical structures involved in alertness, thereby promoting cortical activity in networks involved in ongoing non-visual cognitive processes. Further evidence in support of time independency of alertness builds up from a study in which participants were exposed to either bright light (5000 lux) or dim light (<10 lux) (control condition) either between 12:00 and 16:00 h or between 00:00 and 04:00 h. Bright light had a
time-dependent effect on heart rate and CBT, such that bright light exposure at night, but not during daytime, increased heart rate and CBT (Rüger et al., 2006). On the other hand, nighttime and daytime bright light reduced sleepiness and fatigue significantly and similarly and thus was independent of its timing (Rüger et al., 2006).

The aforementioned studies used polychromatic bright light above 1000 lux. It could very well be that light with this high intensity does not exhibit time-dependent alerting responses.

C. Dose- and Wavelength Response Relationship of Light Exposure on Alertness

Although it is clearly recognized that bright light (≥ 1000 lux) is an effective Zeitgeber and alerting factor in humans (Badia et al., 1991; Daurat et al., 2000; Foret et al., 1996; Myers and Badia, 1993), one could assume that the human circadian pacemaker is insensitive to lower levels of light illumination (<100 lux). However, it has been shown that the relationship between the resetting effect of light and its intensity follows a compressive nonlinear function, such that exposure to lower illuminances still exerts a robust effect (Boivin et al., 1996). For instance, the dose–response function to a single episode of light in the phase delay region (light prior to temperature nadir) can be characterized by a logistic function with a high sensitivity, such that half of the maximal resetting and melatonin suppression achieved in response to bright light (9100 lux) can be obtained with 1% of this light (dim room light of ~100 lux) (Cajochen et al., 2000; Zeitzer et al., 2000b) (Fig. 4). Interestingly, the illuminance response function for alertness is similar to that of the dose–response function reported for the magnitude of suppression of plasma melatonin concentrations as a function of light intensity, as well as the dose–response function reported for the circadian phase resetting effects of light (Cajochen et al., 2000). This suggests that nighttime exposure to typical room light (90–180 lux) can exert an alerting effect in humans, regardless of whether alertness is quantified by subjective ratings or by analysis of the EOG (i.e., incidence of slow-eye movements) and the EEG (activity in the theta and alpha range). Surprisingly, humans were able to maintain stable circadian entrainment to a 24-h cycle in which ambient room light was about 1.5 lux, suggesting that even candlelight can induce small shifts of the human circadian system (Duffy and Wright, 2005). Taken together, this suggests a saturation point for light’s impact on alertness, and that this relatively high sensitivity may explain why in some previous studies a direct effect of light was not observed as the effects of “bright light” were compared to “dim light”
conditions of sufficient intensity to elicit near maximal effects (Dollins et al., 1993; Myers and Badia, 1993). In contrast to the intensity dose-response relationships of light and the circadian system and alertness, very little is known about the duration dependence of the circadian resetting responses to light. However, in analyses of the human phase-response curve (”response to light”), maximum phase shifts to 1 h of bright white light (~10,000 lux) were about 40% as effective as phase shifts measured in response to 6.5 h of white light (~10,000 lux), despite representing only 15% of the stimulus strength (1 h/6.5 h) (Khalsa et al., 2003). Exposure to intermittent light also seems to be highly effective at resetting the human circadian system. The phase-resetting effect of 6.5 h of continuous bright white light (~10,000 lux) is comparable to a 6.5-h intermittent exposure consisting of six cycles of 15 min of bright light (~10,000 lux) and 60 min of dim light (<3 lux) (Rimmer et al., 2000). Despite representing only 23% of continuous bright-light exposure conditions, the intermittentlight regimen elicited comparable phase shifts. Thus, a single sequence of intermittent bright-light pulses has a greater resetting efficacy on a per-minute basis than does continuous light exposure. In a subsequent study, exposure to two 45-min pulses of bright light in the early subjective evening entrained the circadian system to a non-24-h day,
indicating that intermittent pulses are highly efficient at resetting human circadian rhythms (Gronfier et al., 2007), and can significantly contribute to efficient wakefulness.

The relationship between the wavelength of light and its alerting response yielded clear superiority of short wavelength light (470 nm and lower) in comparison to other wavelengths (Cajochen et al., 2005; Lockley et al., 2006; Münch et al., 2006; Revell et al., 2006a). For instance, exposure to 460-nm monochromatic light for 6.5 h during the biological night attenuated subjective sleepiness (Fig. 5) and waking EEG power density in the delta–theta frequency range, with concomitant increase in the high-frequency alpha range, in comparison to light exposure to an equal photon density of 555-nm monochromatic light (Lockley et al., 2006). Given that greater responses were elicited

![Graph showing the effects of light exposure on melatonin levels and subjective sleepiness.](image)

**Fig. 5.** Effects of a 2-h light exposure at 460 nm (Dark gray circles), 550 nm (Light gray circles), and no light (Black squares) in the evening under constant posture conditions (i.e., supine in bed) on salivary melatonin levels and subjective sleepiness (mean values ($n=9$) and SEM). For clarity, the SEM values for the 550-nm light condition were not plotted. Significant post hoc comparisons ($p<0.05$; Duncan’s multiple range test corrected for multiple comparisons) are indicated by the following symbols: *, 460-nm light vs. no light; ▼, 550-nm light vs. no light; and ○, 460-nm light vs. 550-nm light. The pre-light exposure episode represents a 2-h dark adaptation episode under 0 lux, whereas the light level in the 1.5-h post-light exposure was 2 lux. Taken from Cajochen et al. (2006) with permission.
following exposure to an equal number of photons of 460-nm light, as compared to 555-nm light, it is very likely that photoreceptors mediating the acute effects of light on subjective and objective correlates of alertness are blue shifted relative to the visual photopic system. This blue-shift response was similarly observed in a study that compared a 2-h evening exposure to monochromatic light of two different wavelengths (460 and 550 nm) at very low intensities, whereby subjects were more alert during the 460-nm than the 550-nm light (Cajochen et al., 2005). These findings corroborate to a wide range of non-visual light responses in humans, such as melatonin suppression (Lewy et al., 1980; Zeitzer et al., 2000b), circadian phase shifting (Czeisler et al., 1986), nocturnal decline in EEG SWA (Cajochen et al., 1992; Münch et al., 2006), and circadian gene expression (PER2) in oral mucosa (Cajochen et al., 2006). Common to these responses is that they are all more sensitive to short wavelength light. However, very recent findings suggest that cone photoreceptors contribute substantially to non-visual responses at the beginning of a light exposure and at low irradiances, whereas melanopsin appears to be the primary circadian photopigment in response to long-duration light exposure and at high irradiances (Gooley et al., 2010).

D. NEUROANATOMICAL UNDERPINNINGS OF THE EFFECT OF LIGHT ON ALERTNESS AND COGNITIVE PERFORMANCE

The neuroanatomical structures and the concomitant neurophysiology that mediate the capacity of light to enhance alertness and cognitive performance are currently under intensive investigation. It is known that ipRGCs project to a range of targets, including the SCN, subparaventricular zone, and pretectal area that are implicated in mediating NIF responses (Hattar et al., 2002). Furthermore, these cells also project directly to the VLPO that also receives secondary afferents from the SCN, subparaventricular zone, and DMH (Hattar et al., 2002). The VLPO innervates all of the major nuclei of the ascending monoaminergic and in particular the histaminergic pathways, which are thought to play a key role in wakefulness and EEG arousal (Aston-Jones et al., 1999; Lin et al., 1996; Saper et al., 2005b). Direct photic input to this nucleus may therefore alter VLPO activity and waking arousal levels. The LC is also involved in the regulation of the sleep–wake cycle (Saper et al., 2005b), regulating the amplitude of the sleep–wake circadian rhythm set by the SCN by increasing wakefulness during the active period (see also Section III, Gonzalez and Aston-Jones, 2006).

Light impacts on cognitive performance through its synchronizing/phase-shifting effects on the circadian clock or acutely via its alerting effects, as
performance (in tasks such as digit recall, serial addition–subtraction and simple reaction time tasks) can immediately improve after the onset of light exposure at night (Badia et al., 1991; Campbell and Dawson, 1990; Foret et al., 1996; Lockley et al., 2006) and also during the day (Phipps-Nelson et al., 2003; Ruger et al., 2006). EEG and ocular correlates of alertness can vary with cognitive performance, such that EEG alpha (8–12 Hz) and beta (12–20 Hz) activities show a pronounced circadian rhythm with a peak in the second half of the biological day (Cajochen et al., 2002). Light exposure reduces alpha, theta, and low frequency EEG activity, and also the incidence of slow-eye movements, which are correlates of sleepiness, and thus good indicators of inattention that increase as a result of extended wakefulness particularly during the biological night. Lights’ performance enhancement, however, does not occur in a similar manner for all subcortical and cortical regions. Light-induced modulations of cortical activity during auditory cognitive tasks occur for alertness-related subcortical structures, such as the brainstem (LC—compatible region) (Vandewalle et al., 2007b); the hypothalamus, in a location encompassing the SCN (Perrin et al., 2004), and dorsal and posterior parts of thalamus (Vandewalle et al., 2006, 2007a), in long-term memory and emotion-related areas, such as the hippocampus (Vandewalle et al., 2006) and amygdala (Vandewalle et al., 2007b). Taken together, these responses indicate that wide-range subcortical and cortical areas are activated by non-visual effects of light, during specific cognitive tasks. Since cognitive performance can exhibit a circadian modulation, the next logical question is whether these cortical responses are wavelength dependent. Blue light (460 nm) appears to be more effective in sustaining performance in a simple vigilance reaction time task compared to green light (550 nm) (Lockley et al., 2006). fMRI-assessed brain responses undergo a wavelength dependency for higher executive task (2-back task), such that blue light enhances modulations of higher executive tasks in the brainstem (in an LC-compatible location), in the thalamus and insula, in relation to green (550 nm) and violet exposures (430 nm). In this case, the effect of blue light occurs before 1 min after the start of the exposure (Vandewalle et al., 2007b) and last for nearly 20 min (Vandewalle et al., 2007a). However, the magnitude, time course, and regional brain distribution of non-visual effects of light heavily depend on the dose, duration, and intensity of the light exposure. Indeed, longer durations and higher intensities can elicit long-lasting and widespread task-related responses (Perrin et al., 2004). While subcortical regions are activated faster and show short-lasting responses to light, cortical activity requires stronger and longer stimulations, as indicated in a study (Vandewalle et al., 2006), in which 20 min of bright white light induced both thalamic and cortical modulations that steadily declined after light exposure, albeit its rather lasting effects (responses were observed several minutes after the end of the light exposure). Moreover, when the duration of light exposure was reduced to less than a minute, the effects were mostly restricted to subcortical structures such as the dorso-
What keeps us awake?

The posterior thalamus and the brainstem (LC-compatible area), and cortical modulations were sharply reduced (Vandewalle et al., 2007b).

The importance of LC areas in this case is due to the fact that this region projects to numerous cortical sites and is, therefore, well placed to mediate light-induced changes in alertness and cognition (Gonzalez and Aston-Jones, 2006). The thalamus, in particular its dorsal and posterior nuclei (i.e., pulvinar), is a key structure involved in the interaction between alertness and cognition (Portas et al., 1998). Thus, light-induced changes in thalamic activity can be directly implicated in enhanced alertness during light exposure. Given that the thalamus plays a critical role in the relay of information to the cortex, it can regulate information flow in the brain, and an effect of light on the thalamus may thus lead to widespread cortical effects.

E. NON-Clinical APPLICATIONS OF Light

The application of light in non-clinical settings, such as intercontinental travel (jet-lag), shift work, and even non-shift working environments, is under intense scrutiny. The main assumption for the first two cases is the misalignment between the internal circadian pacemaker and the external environment. As a consequence, this circadian deregulation may contribute to health problems in the long term such as sleep disorders, cardiovascular disease, and diabetes (Rajaratnam and Arendt, 2001). Previous strategies to reduce jet-lag have focused on shaping the perceived LD cycle after arrival, in order to facilitate a phase shift in the appropriate direction. In one study, phase advancements of habitual sleep–wake schedules and light exposure in the morning were investigated in order to test the idea that if travelers could phase-advance their circadian rhythms prior to eastward flight, they would arrive with their circadian rhythms already partially re-entrained to local time. For this three treatments were used, in which habitual sleep schedule was advanced by 1 h/day for 3 days, together with morning light exposure for the first 3.5 h after waking on each of the 3 days. This exposure was either continuous bright light (>3000 lux), or intermittent bright light (>3000 lux, 0.5 h on, 0.5 off, etc.), or ordinary dim indoor light (<60 lux). Dim light melatonin onset (DLMO) phase advance was higher in the continuous light exposure (nearly 2 h), although it did not drastically differ from the intermittent light exposure. Importantly, in both cases, alertness was significantly higher under light exposure (Burgess et al., 2003).

With respect to shift work, it is unambiguous that the circadian misalignment between the endogenous circadian signal and the imposed rest-activity cycle is one of the main sources of sleep, performance, and health troubles in night-shift workers (Lamond et al., 2003). Timed bright light exposure during night work can
reduce circadian misalignment in night workers. As an illustration, shift workers under bright light exposure (7000–12,000 lux) during the night (and darkness during the day) had a temperature nadir shifted after 4 days of treatment to a significantly later, mid-afternoon hour (compared to the previous 03:00 h), indicating a successful circadian adaptation to daytime sleep and nighttime work. Similarly, there were concomitant shifts in subjective assessment of alertness and cognitive performance, both of which improved substantially under this light exposure (Czeisler et al., 1990). However, despite the fact that scheduled bright light and darkness can phase shift the circadian clocks of night workers for complete adaptation to a night work with day sleep schedule, few night workers would rather be out of phase with the diurnal world on their days off. Similarly in other situations, such as rapidly rotating shifts and the normal office environment, it is more appealing to time light exposures toward improving alertness without phase shifting (Horowitz and Tanigawa, 2002). However, given that there is no dead zone for phase shifting the circadian system in humans (Khalsa et al., 2003), it is not conceivable to enhance alertness with light without affecting circadian phase. Thus, a “compromise” circadian phase position for permanent night-shift work in which the sleepiest circadian time is delayed out of the night work period and into the first half of the day sleep episode would seem a feasible alternative (Smith et al., 2009). In a recent study, the target compromise phase position was a DLMO of 3:00 h, which puts the sleepiest circadian time at approximately 10:00 h. This was predicted to improve night-shift alertness and performance while permitting sufficient daytime sleep after work as well as late-night sleep on days off. For such, intermittent four 15 min of bright light pulses were conducted during each night-shift, together with recommendations such as use of dark sunglasses during the day, sleep in dark bedrooms at scheduled times, and outdoor afternoon light exposure, all of which to keep rhythms from delaying too far. Interestingly, subjects who phase delayed close to the target phase (3:00 h) performed better and were more alert during night shifts. This suggests that light application in night shift workers is both a feasible and promising intervention (Smith et al., 2009).

Controlled light and dark exposure during the daytime also has a significant impact on circadian phase and could be an easier alternative to implement in real-life situations. In a recent field study (Viola et al., 2008), the effects of exposure to blue-enriched white light (17,000 K) were investigated in comparison to another white light (4000 K) during daytime work hours in an office setting. Blue-enriched white light substantially improved subjective measures of alertness, mood, performance, evening fatigue, concentration, and dramatically reduced daytime sleepiness (Fig. 6). This suggests that blue-enriched white light can enhance self-reported measures of alertness, performance, and fatigue after daytime exposure in a “real-life” setting for people who work normal office hours without any abnormal sleep–wake schedule being imposed, which makes it an appealing alternative to enhance alertness.
Fig. 6. Exposure to blue-enriched white light at 17000 k (Dark gray bars) during daytime work hours improves subjective alertness, performance, evening fatigue, and sleep quality, in comparison to white light at 4000 K (White bars). Modified from Viola et al. (2008).

To sum up light exerts powerful non-visual effects on a wide range of physiological, behavioral, and subjective parameters, ranging from alertness to complex behavioral processes like cognition. However, in order to achieve optimal alerting response to light, several factors including dose, duration, timing, and wavelength should definitely be taken into account. Novel evidence points to a potential role of the non-image forming system in the regulation of alertness. This opens an exciting area of investigations that may unravel how the retinal and suprachiasmatic networks are involved in the regulation of circadian rhythms and sleep–wake homeostasis.

IV. Effects of Melatonin on Human Sleep and Wakefulness

Pineal melatonin is primarily a neuroendocrine transducer of external time (LD cycle) promoting an increased propensity for “dark appropriate” behavior. The most unequivocal characteristic of endogenous melatonin is its utility to be used alone or in combination with CBT as a phase marker of the endogenous circadian pacemaker located in the SCN. However, there are three major reasons, which imply that melatonin could also play an important role in the regulation human sleep–wake behavior:
1. The endogenous melatonin rhythm exhibits a close temporal association with the endogenous circadian component of the sleep propensity rhythm.

2. There is evidence that exogenous melatonin (oral intake) is able to induce sleep when the homeostatic drive to sleep is insufficient, to inhibit the drive for wakefulness emanating from the circadian pacemaker and to induce phase shifts in the circadian clock such that the circadian phase of increased sleep propensity occurs at a new desired time.

3. Light’s acute alerting response depends on its capacity to suppress endogenous melatonin levels during the biological night.

A. ENDOGENOUS MELATONIN AND THE HUMAN CIRCADIAN SLEEP-WAKE CYCLE

Melatonin (N-acetyl-5-methoxytryptamine) is a major secretory product of the pineal gland, and its production is under circadian control by the SCN. Because it is produced exclusively at night, it has been referred to as “a chemical code of darkness” (Arendt, 2006). The relationship between external LD cycles and melatonin production can be explained via a multisynaptic pathway beginning with photic transduction of light at the level of the retina; transmission of this LD information via the retinohypothalamic tract (RHT) to the SCN; a descending pathway from the SCN through the superior cervical ganglion in the spinal cord; and, finally, an ascending pathway to the level of the pineal (Vollrath, 1984). Data in tetraplegic patients, whose melatonin production was absent, support the hypothesis that the human pineal must be stimulated by the sympathetic nervous system to produce melatonin (Zeitzer et al., 2000a). At a functional level, bright light acts through this pathway to acutely suppress melatonin production in the pineal (Levy et al., 1980).

The increase in melatonin secretion in the evening correlates with an increase in sleep propensity (Cajochen et al., 1999b; Tzischinsky et al., 1993). This latter phenomenon has been referred to as “the opening of the sleep gate” (Lavie, 1997) and is most likely related to an inhibitory effect of melatonin on SCN activity (Liu et al., 1997). In parallel, the entire thermoregulatory cascade (i.e., decrease in heat production and increase in heat loss leading to decrease in CBT) starts with the rise in endogenous melatonin levels in the evening (Kräuchi et al., 2000). As a consequence, alertness levels start to decline, sleepiness kicks in, and sleep is eventually commenced. The association of sleep with the melatonin rhythm has been confirmed in blind people in whom the circadian pacemaker is not entrained (Lockley et al., 1997; Nakagawa et al., 1992) and in sighted subjects with non-24-h sleep–wake cycle syndrome (Uchiyama et al. 2000). Results obtained from studies using the forced desynchrony protocol to separate out circadian- and wake-dependent components of behavior clearly show that the circadian increase in melatonin secretion
coincides with a decrease in wake episodes during scheduled sleep episodes (Dijk and Cajochen, 1997). Sleep consolidation gradually deteriorates during that phase of the circadian cycle with low melatonin production, and EEG activation during wakefulness is also timed at a specific phase relative to the circadian melatonin rhythm (Cajochen et al., 2002).

Despite the close temporal association between endogenous melatonin levels and sleep–wake rhythms, it is still a matter of debate whether endogenous melatonin is causally implicated in the regulation of sleep and wakefulness (van den Heuvel et al., 2005; Zhdanova, 2005), since the ability to sleep is still possible in the absence of detectable endogenous melatonin during the day, or in tetraplegic patients (Scheer et al., 2005), and only a moderate incidence of sleep disturbance has been reported in pinealectomized patients (Macchi and Bruce, 2004). Furthermore, absolute melatonin production (which varies enormously between individuals) does not correlate with sleep quality in the elderly (Youngstedt et al., 1998) or elderly sleep-maintenance insomniacs (Hughes et al., 1998). However, several lines of evidence suggest that endogenous melatonin levels may still play a role in consolidated sleep and/or wakefulness. Acute suppression of the nighttime melatonin surge—either by light or beta-blockers—compromises sleep quality, which can be reversed by melatonin supplementation (Cajochen et al., 1998; Van Den Heuvel et al., 1997). In a survey of 13 adult pineal surgery patients, over half the patients (54%) reported nighttime wake periods lasting 1 h or longer, 31% reported total nighttime sleep durations of less than 6 h, and 38% complained of experiencing poor or disturbed sleep every night (Macchi et al., 2002). The extent to which these disturbances are directly attributable to pineal dysfunction rather than to a general effect of brain surgery *per se* is not entirely clear, but points to compromised sleep under chronic absence of nighttime melatonin secretion. Similarly, in the study of Scheer et al. (2005), all subjects with a complete cervical spinal cord injury, which interrupts the neural pathway required, had chronically impaired sleep efficiency and quality (Scheer et al., 2005). Furthermore, in a study investigating the effect of bright light and melatonin on neurocognitive function and sleep in elderly residents, long-term bright light (5 years) significantly increased endogenous melatonin levels at night concomitant with an improvement in subjective and objective sleep quality (Riemersma-van der Lek et al., 2008).

**B. Effects of Exogenous Melatonin on Human Sleep and Wakefulness**

The first evidence that exogenous melatonin affects wakefulness was provided by the work of Aaron Lerner, who discovered melatonin in 1958 (Lerner et al., 1958). When he started to treat patients suffering from vitiligo, a human
pigmentation disease, he noted that many of his patients became sleepy and fell asleep. Since then numerous laboratory studies under stringent conditions clearly demonstrated that administration of melatonin acutely affects sleep and wakefulness in humans. Exogenous melatonin elicits all the physiological effects which occur in the evening during endogenous melatonin secretion (for a review, see Cajochen et al., 2003). Indeed, exogenous melatonin is most effective when endogenous levels are low during the biological day. It elicits time-dependent soporific effects, which have been corroborated with electrophysiological measures of sleepiness such as (EEG) theta activity during wakefulness (Cajochen et al., 1997b) and with brain correlates of sleepiness in an fMRI study, which highlighted the role of melatonin in priming sleep-associated brain activation patterns in anticipation of sleep (Gorfine et al., 2006). In an experiment where we blocked the natural evening increase in heat loss, subjective sleepiness, and melatonin secretion by light exposure, we could show that melatonin replacement (5 mg) acutely recovered the evening increase in heat loss, subjective sleepiness, and also theta activity in the waking EEG (Cajochen et al., 1998; Kräuchi et al., 1997). Nighttime melatonin administration does not affect sleep consolidation or sleep efficiency (Cajochen et al., 1997a), whereas, during day­time, an improvement in sleep efficiency could be found (Dijk et al., 1995). More recent data from a forced desynchrony protocol, where melatonin was given to healthy young adults across a full range of circadian phases, confirm that exogenous melatonin can only increase sleep efficiency outside the time window of its normal production (Fig. 7; Wyatt et al., 2006).

Similar findings come from an extended sleep protocol. Chronic administration of melatonin in a slow-release formulation during a 16-h sleep opportunity beginning at 16:00 h resulted in a redistribution of sleep so that sleep efficiency during the first half of the sleep opportunity was substantially higher during melatonin treatment compared to placebo (Rajaratnam et al., 2004). These two studies provide strong support for the hypothesis that exogenous melatonin attenuates the wake-promoting signal of the endogenous circadian pacemaker, allowing for increased sleep efficiency at circadian phases corresponding to the habitual wake episode.

C. IMPLICATIONS FOR THE TREATMENT OF INSOMNIA AND CIRCADIAN RHYTHM DISORDERS

Melatonin’s soporific and chronobiotic properties make it an optimal candidate for treating sleep, in addition to circadian rhythm disorders. In our view, the most successful attempt to treat insomnia and changes in circadian phase position by melatonin has been carried out in free-running blind people. Optimal
WHAT KEEPS US AWAKE?

Fig. 7. Mean sleep efficiency levels during a forced-desynchrony protocol, folded at the intrinsic circadian period derived from core body temperature. The figure shows the circadian rhythm of endogenous sleep propensity (percentage of sleep of recording time, placebo group in black, melatonin 0.3 mg group in red), as well as the endogenous melatonin levels in the placebo condition (gray area). Modified from Wyatt et al. (2006).

Melatonin treatment in those people should utilize not only its soporific effects by administration close to the desired bedtime, but also its chronobiotic properties, in order to entrain sleep–wake behavior (Lockley et al., 2007). Another promising patient group are elderly patients with insomnia. The results of melatonin treatment administered before bedtime in elderly insomniacs were not consistent (for a review, see Olde Rikkert and Rigaud, 2001). However, Olde Rikkert and Rigaud concluded that melatonin is most effective in elderly insomniacs who chronically use benzodiazepines and/or with documented low melatonin levels during sleep.

Abnormal timing of sleep with respect to circadian phase occurs in the delayed sleep phase syndrome (DSPS), in which sleep occurs at a delayed clock time relative to the LD cycle, social, work, and family demands. In the first use of melatonin in patients with DSPS, it was found that when administered 5 h before sleep onset for a period of 4 weeks, melatonin (5 mg) advanced sleep onset and wake times compared with placebo (Dahlitz et al., 1991), which was later confirmed by Nagtegaal et al. (1998) and Mundey et al. (2005), and is most effective in DSPS patients with shorter habitual sleep time and later clinical onset (Kamei et al., 2000).

The first application of melatonin using chronobiological principles was to alleviate the perceived effects of jet-lag. There have been many placebo-
controlled and placebo-uncontrolled studies that have been summarized in a Cochrane (Herxheimer and Petrie, 2002). This stringent analysis concludes that 9 of 10 trials of melatonin, taken close to the target bedtime at destination, decreased jet-lag symptoms arising after flights crossing five or more time zones. One difficulty in using melatonin for jet-lag is that its use requires administration at times when it will have undesired soporific properties.

There is also a great interest in whether melatonin can facilitate phase-shifting in night-shift workers; however, few studies have measured such phase shifts. In two laboratory studies, circadian rhythms were measured before and after a large shift in the sleep–wake schedule (Dawson et al., 1995; Samel et al., 1991). Melatonin (5 mg) was administered during the phase-advance portion of the phase response curve (PRC) and produced larger circadian phase shifts than placebo (Samel et al., 1991). In the other study, subjects took a 4 mg melatonin (or placebo) before and during their daytime sleep (Dawson et al., 1995) and melatonin did not produce a larger phase delay than placebo. In a night-shift field study, melatonin produced larger circadian phase shifts than placebo in only 7 of the 24 subjects studied (Sack and Lewy, 1997). Overall, these studies do not provide strong evidence that melatonin can help phase shift the circadian rhythms of night-shift workers, in particular, when comparing its action as being less strong than exposure to light. One problem has been the lack of control over time of melatonin administration and of the subjects’ sleep schedules. In a recent study where the timing of melatonin administration, the sleep–wake schedule and, to some extent, the LD cycle could be controlled in a field setting, melatonin clearly produced larger phase advances than placebo in the circadian rhythms of melatonin and CBT (Sharkey and Eastman, 2002). Moreover, significantly larger phase advances with 0.5 and 3.0 mg melatonin compared with placebo have been reported in a study to determine if phase advances induced by morning light could be increased with afternoon melatonin (Revell et al., 2006b). Additional caution is required in this setting to avoid the soporific effects of melatonin during work requiring vigilance, or driving home after the shift.

In an attempt to take advantage of the therapeutic opportunities of melatonin, several melatonin agonists with improved properties in comparison to melatonin have been developed. Some of these agents are selective for specific melatonin receptors (MT1, MT2). Results from animal studies suggest that MT1 and MT2 receptors have distinct functional roles in the SCN, albeit with some overlapping function (for a review see Turek and Gillette, 2004). These distinct roles provide great potential for receptor-specific pharmacological agents to affect specific aspects of the sleep–wake cycle and/or circadian rhythmicity. It may be possible to develop specific agents that promote sleep without phase-shifting the circadian clock, or the converse. The three more prominent examples of melatonin receptor agonists that are the furthest along in clinical development are Agomelatine, Ramelteon, and Tasimelteon. All of them appear to be efficacious in the treatment of circadian rhythm sleep disorders and some types of insomnia.
WHAT KEEPS US AWAKE?

(for a review see Ferguson et al., 2010). An important point for the effects of melatonin analogues is to understand that they are not hypnotic drugs that resemble benzodiazepines and their derivatives. Melatonin-like compounds amplify day–night differences in alertness and sleep quality.

References


WHAT KEEPS US AWAKE?


