Winter Depression: Integrating mood, circadian rhythms, and the sleep/wake and light/dark cycles into a bio-psycho-social-environmental model

Alfred J. Lewy, MD, PhD,
Oregon Health & Science University, Portland, OR; phone: 503-494-7746, fax: 503-494-5329, lewy@ohsu.edu

Jonathan S. Emens, MD,
Oregon Health & Science University, Portland, OR; phone: 503-494-4041, fax: 503-494-5329, emensj@ohsu.edu

Jeannie B. Songer, BA,
Oregon Health & Science University, Portland, OR; phone: 503-418-4144, fax: 503-494-5329, songerj@ohsu.edu

Neelam Sims, BS,
Oregon Health & Science University, Portland, OR; phone: 503-494-7961, fax: 503-494-5329, simsn@ohsu.edu

Amber L. Laurie, BA,
Oregon Health & Science University, Portland, OR; phone: 503-494-4677, fax: 503-494-5329, lauriea@ohsu.edu

Steven C. Fiala, BA, and
Oregon Health & Science University, Portland, OR; phone: 503-494-8099, fax: 503-494-5329, fialas@ohsu.edu

Allie L. Buti, BS
Oregon Health & Science University, Portland, OR; phone: 503-418-4144, fax: 503-494-5329, butia@ohsu.edu

Synopsis

The phase shift hypothesis (PSH) states that most patients with SAD become depressed in the winter because of a delay in circadian rhythms with respect to the sleep/wake cycle: According to the PSH, these patients should preferentially respond to the antidepressant effects of bright light exposure when it is scheduled in the morning so as to provide a corrective phase advance and restore optimum alignment between the circadian rhythms tightly coupled to the endogenous circadian pacemaker and those rhythms that are related to the sleep/wake cycle. Recent support for the PSH has come from studies in which symptom severity was shown to correlate with the degree of circadian misalignment: it appears that a subgroup of patients are phase advanced, not phase delayed; however, the phase-delayed type is predominant in SAD and perhaps in other disorders as well, such as non-
seasonal unipolar depression. It is expected that during the next few years the PSH will be tested in these and other conditions, particularly since healthy subjects appear to have more severe symptoms of sub-clinical dysphoria correlating with phase-delayed circadian misalignment; critically important will be the undertaking of treatment trials to investigate the therapeutic efficacy of morning bright light or afternoon/evening low-dose melatonin in these disorders in which symptoms are more severe as the dim light melatonin onset (DLMO) is delayed with respect to the sleep/wake cycle (non-restorative sleep should also be evaluated, as well as bipolar disorder). The possibility that some individuals (and disorders) will be of the phase-advanced type should be considered, taking into account that the correct timing of phase-resetting agents for them will be bright light scheduled in the evening and/or low-dose melatonin taken in the morning. While sleep researchers and clinicians are accustomed to phase-typing patients with circadian-rhythm sleep disorders according to the timing of sleep, phase typing based on the DLMO with respect to the sleep/wake cycle may lead to quite different recommendations for the optimal scheduling of phase-resetting agents, particularly for the above disorders and conditions.

Keywords
Melatonin; Light; Dim light melatonin onset (DLMO); Winter depression (SAD); Phase angle difference (PAD); Bio-psycho-social-environmental model

Introduction
The diagnosis of winter depression (seasonal affective disorder, or SAD) is based upon its annual pattern of recurrence in the fall/winter and spontaneous remission in the spring/summer. At temperate latitudes, about 5% of the population is estimated to have SAD, with another 15% of the population manifesting less severe symptoms (sub-syndromal SAD) [1]. Females predominate, by a vast majority, at least between the ages of puberty and menopause. In addition to the usual characteristics of depression, particularly what is termed atypical or retarded major depression (in which the hallmark vegetative changes in sleep and appetite are in the direction of increased sleep and appetite), SAD patients crave foods that contain complex carbohydrates (such as pasta, baked goods and sweets) and gain weight in the winter [2]. Fruits and vegetables are preferred in the spring and summer. Weight loss, if it occurs at all, is during these two seasons, often resulting in smaller sized wardrobe in the summer. Fatigue, particularly difficulty getting up in the morning, is omnipresent, despite a tendency to lengthen sleep (up to three hours) in the winter, which is not restorative. While generally not as severe as other types of major affective disorders (for example, suicide is less common in SAD than in bipolar and non-seasonal unipolar major depression), SAD patients socially isolate themselves to a considerable degree; typically, they will say that they withdraw on week-ends and as soon as they get home from work. Many are often quite irritable with family, friends and co-workers. SAD appears to be more common at the higher of the temperate latitudes [3] and affects all ages, sometimes manifesting as school anxiety in young children (during the fall and winter) [4,5]. SAD appears to run in families and therefore is thought to have a strong genetic component. Measurement of symptom severity is accomplished by a number of mood scales, primarily by one or another version of the SIGH-SAD (Structured Interview Guide for the Hamilton Depression Rating Scale, Seasonal Affective Disorders version), which was originally a composite of the standard 21-item Hamilton Depression Scale and eight items thought to be highly representative of SAD [6].

How can understanding the phase shift hypothesis (PSH) for winter depression, or seasonal affective disorder (SAD) be of benefit to sleep researchers and clinicians? One, the same circadian mechanism causing SAD may also be a cause of non-restorative sleep. Two, the melatonin laboratory test useful in SAD, which partially depends on the mid-point between
sleep onset and waketime, may potentially be applicable to non-restorative sleep. Three, the other part of this lab test, the dim light melatonin onset (DLMO) also indicates the times of the light and melatonin phase response curves (PRCs) and therefore provides the optimal schedules for using these phase-resetting agents in treating circadian misalignment disorders (such as SAD), as well as advanced sleep phase syndrome (ASPS) and delayed sleep phase syndromes (DSPS). Four, the bio-psycho-social-environmental model inspired by SAD and the PSH, now appears to be relevant to other disorders, including sleep disorders. It is hoped that this monograph will stimulate further research in disorders that may have a circadian misalignment component (such as non-restorative sleep), as well as canonical circadian phase sleep disorders, ASPS, DSPS, and hypernycthemeral syndrome.

SAD appears to be caused, at least in part, by a mismatch between the sleep/wake cycle and the circadian rhythms that are tightly coupled to the endogenous circadian pacemaker [7]. Phase-resetting agents (such as bright light exposure and low-dose melatonin administration) are the treatments of choice, provided the SAD patient is properly phase typed, so that these agents can be administered at the correct time. It is also possible that these agents can phase shift rhythms too much, causing circadian misalignment in the opposite direction. The prototypical phase-delayed SAD patient has a dim light melatonin onset (DLMO) that is delayed with respect to mid-sleep (the mid-point between sleep onset and sleep offset); a smaller subgroup has a DLMO that is phase advanced with respect to mid-sleep (Fig. 1). This line of thought builds on three decades of research in which endogenous melatonin has been the primary dependent variable and exogenous melatonin the primary independent variable; therefore, by way of introduction, the following review is in order.

Historical Perspectives

The second half of the 1970s was seminal for the field of chronobiology. Dr. Thomas Wehr (working with Dr. Frederick Goodwin) and Dr. Daniel Kripke, and their research teams hypothesized that major affective disorders could be caused by a mismatch between the circadian rhythms associated with core body temperature and those related to the sleep/wake cycle [8–10]. Specifically, the phase advance hypothesis stated that the temperature rhythm (and its related circadian rhythms) was phase advanced with respect to the sleep/wake cycle.

It was difficult to test this hypothesis, because the only available phase-resetting treatment at the time was to shift the timing of the sleep/wake cycle [10]; when sleep was scheduled earlier, so as to theoretically correct the misalignment, the resulting clinical benefit was only transient. Shifting sleep was the only treatment available because chronobiologists had concluded that, unlike all other animals, humans did not make primary use of the light/dark for synchronizing their biological rhythms; instead, social cues were thought to be more important [11]. In animals, seasonal rhythms were known to be cued to the time of the year via night-length as it affected the duration of nighttime melatonin production [12]; the duration of melatonin production in the summer was shortened due to the acute suppressant effect of light on melatonin production in the morning and in the evening. However, the acute suppressant effect of light was not demonstrated in humans until after the phase advance hypothesis was formulated (as described below), which in turn depended on the development of a sufficiently accurate and sensitive assay for measuring melatonin in humans.

The gas chromatographic – negative ion mass spectrometric (GCMS) assay for melatonin met the specifications and became the gold standard for laboratory analysis and measurement of melatonin in human plasma [13]. The other assays that were available to researchers in the 70s lacked adequate specificity and/or sensitivity; some reported high daytime circulating levels and relatively little difference between day and night [14–17], even though it was known that whole pineal glands contained large amounts of melatonin at night compared to the day. The GCMS assay was used in some of the most influential studies of human melatonin physiology,
as well as directly and indirectly enabled development of less costly and more convenient radioimmunoassays (RIAs) with sufficient specificity and sensitivity.

The GCMS assay was used to validate the use of circulating levels of melatonin as a measure of pineal output [18,19], unaffected by so-called extrapineal contributions [20] that turned out to be immunoreactive substances that were not melatonin. In 1980, it was reported that human nighttime melatonin could be suppressed by light exposure, providing it was sufficiently intense [21] (see Fig. 2). Previous studies had apparently used room light of ordinary intensity [22–24], which was not sufficiently intense to suppress melatonin. Chronobiologists immediately understood the implication that humans might therefore have biological rhythms cued to the natural cycle of (brighter) sunlight and darkness relatively unperturbed by exposure to ordinary-intensity room light. Furthermore, this finding directly led to the use of bright artificial light to experimentally, and therapeutically, manipulate biological rhythms.

Dr. Kripke was the first investigator to treat non-seasonal depressives with morning bright light exposure. In December, 1980 Lewy and co-workers at the NIMH were afforded the opportunity to treat Mr. Herbert Kern, the first self-identified patient with SAD, using 2500 lux exposure between 6 –9 a.m. and 4–7 p.m., based on animal models in which seasonal rhythms respond to the time interval between the twilight transitions (and the duration of nighttime melatonin production). Mr. Kern’s depression began to remit in a few days, and the response was complete within two weeks, a time course that continues to hold for patients with SAD. The first controlled study of light therapy in SAD was conducted by Rosenthal and co-workers, in which relatively dim yellow light was used as a placebo[2].

At OHSU, the thinking of the Lewy and Sack research team began to diverge from that of the Kripke and NIMH groups. Dr. Kripke chose morning as the best time to schedule light for treating non-seasonal depressives. While this choice may turn out to be valid, it was based on the idea of a critical photosensitive interval for light exposure at that time consistent with a photoperiodic model. Similarly, the NIMH group focused on a seasonal/photoperiodic, rather than a circadian, approach (although the two are interrelated), as well as on a “photon counting” hypothesis in which light at any time of the day (preferably at the most convenient time in the evening) would be therapeutic in SAD patients provided that the exposure was of sufficient intensity and duration. On the other hand, the OHSU team focused on the circadian phase-resetting effects of light, and hypothesized that for most SAD patients morning light would provide more of an antidepressant effect than evening light, because the OHSU team thought that most seasonal depressives were phase delayed [25,26]. Specifically, the phase shift hypothesis (PSH) states that most patients with SAD become depressed in the winter because of the later dawn which causes their circadian rhythms to drift out of phase with the sleep/wake cycle [25,26]. The PSH was inspired by the work of Drs. Kripke and Wehr, but in this case the pertinent affective disorder was SAD. Furthermore, although the PSH left open the possibilities of both a phase-advanced and phase-delayed type of patient, SAD was thought to be primarily of the phase-delayed type and not of the phase-advanced type as hypothesized by Drs. Kripke and Wehr for non-seasonal depression.

An early refinement of the PSH for SAD was to include the possibility of a phase-advanced type [25,26], thought at first to constitute a very small and therefore a negligible subgroup that could be included in studies without violating the integrity of the experimental design (however, see below). In the first major publication of morning vs. evening light treatment, one of eight SAD patients seemed to have a better antidepressant response on evening light [26]. This 1/8th proportion was initially deemed to be relatively high, and some researchers began to regard the PSH as the phase-delay hypothesis (for SAD).
Dr. David Avery made an important contribution to this field by emphasizing the importance of hypersomnia in these presumably phase-delayed individuals [27,28]. Initially, waketime was regarded as a good way to phase type individuals, and those with hypersomnia were considered to be phase delayed based on a late waketime, even if there was an early bedtime. Although consensus was growing about the preferred use of bright light in the morning [26–30], some investigators remained skeptical [31,32]. For example, the NIMH group focused on testing a melatonin/photoperiod hypothesis. At first the results were non-supportive [33], but over the last decade this became the preferred hypothesis to test at the NIMH [34]. Other investigators dismissed the importance of the timing of the light based on their studies showing no difference between morning and evening light [32], studies thought by some investigators to be confounded by the profound placebo response accompanying bright light exposure, documented by Dr. Charmane Eastman [35].

Consensus on the preferential benefit of morning vs. evening light was finally achieved in 1998, when large studies were published by three independent groups [36–38]. However, this did not necessarily validate the PSH, since it could be argued that morning is a time of increased light sensitivity. It became clear that another type of test of the PSH was needed. Fortunately, a second phase response curve (PRC) to melatonin was obtained that replicated and extended the findings in the first melatonin PRC study in humans [39], providing a way to use low-dose daytime melatonin to cause phase shifts (Fig. 3) and to test the antidepressant effects of phase shifts to melatonin, in which a placebo control group is possible.

The history of testing the phase-resetting effects of melatonin in humans has already been reviewed [40]. Suffice it to say that many investigators think that the demonstration in a mammalian (rodent) species under free-running conditions was the inspirational landmark study by the Armstrong team [41], although credit should be given to Dr. Herbert Underwood for his work in lizards [42] as well as to the numerous studies in birds [43,44]. Drs. Sack and Lewy chose to investigate blind people initially [45], so as to follow up on the Armstrong study as closely as possible. Before and at the same time of the studies of blind people, the Arendt and Claustrat teams were testing the circadian phase-shifting effects of melatonin in sighted people [46,47].

We later showed that the seventh blind free runner (BFR) could be entrained to 0.5 mg of melatonin [48]. Based on our PRC to melatonin [49,50], we surmised that when melatonin is administered on the advance zone of the melatonin PRC there will always be a phase advance; however, its magnitude may be reduced if there is too much spillover on the delay zone of the melatonin PRC. In addition to avoiding spillover [48], another heuristically useful pharmacokinetic principle for optimizing melatonin pharmacodynamics is to ensure that there is overlap between the exogenous melatonin pulse and either the onset or the offset of the endogenous melatonin profile so as to optimize the magnitude of the desired phase advance or phase delay, respectively [40].

The dose-response curve for the phase-shifting effects of melatonin indicates a log-linear relationship for doses in the physiological range [51]. An added benefit of using low doses of melatonin is that they are less likely to acutely result in sleepiness that would be undesirable when melatonin is given during the day or early evening. Although the latter has not been systematically examined. [S2]This side effect seems to occur in about a third of the population, and there is no way to predict who will be sensitive to it, although the higher the dose, the greater the proportion of individuals who will be affected and the greater will be the magnitude of the soporific effect.
Testing the PSH with Low-Dose Melatonin Administration

The use of low-dose daytime melatonin to cause phase shifts in blind and sighted people based on the melatonin PRC [49,50] makes possible a critical test of the PSH. SAD patients were given 3–4 doses of melatonin (0.75-0.3 mg) every two hours, beginning at waketime or in the afternoon to cause phase delays or phase advances, respectively [7]. After the data were collected, they were initially analyzed under the assumption that nearly all subjects would be phase delayed and that the inclusion of any phase-advanced subjects would be discounted: therefore, for the group as a whole, afternoon/evening melatonin should be more antidepressant than morning melatonin. This comparison did not reach statistical significance. Hence, subjects were retrospectively phase typed according to the number of hours between their DLMO and mid-sleep, or their phase angle difference (PAD). PAD 6 is the average in historical healthy controls: subjects with PAD > 6 were designated as phase advanced (the DLMO is relatively advanced with respect to mid-sleep) and patients with PAD < 6 were designated as phase delayed (the DLMO is relatively delayed with respect to mid-sleep) (see Fig. 1). Surprisingly, one-third of the subjects were phase-advanced at baseline, before they entered the treatment phase of the study. Remarkably, for both phase-advanced and phase-delayed subjects, the more baseline PAD deviated from 6, the greater was the depression rating. That is, in the phase-advanced group, the more phase advanced the DLMO relative to mid-sleep, the greater the depression ratings, and in the phase-delayed group, the more phase delayed the DLMO relative to mid-sleep, the greater the depression ratings.

Sleep Disorders and Circadian Misalignment

Previously, phase typing was usually only possible in extreme cases of misaligned sleep; for example, people with advanced sleep phase syndrome (ASPS) and people with delayed sleep syndrome (DSPS). PAD 6 offers a way to also phase type individuals with conventional sleep times. Furthermore, both DLMOs and sleep times are required for PAD phase typing. While sleep times are appropriate for diagnosing ASPS and DSPS and for determining the correct scheduling of phase-resetting treatments, they do not take into consideration that there may be internal circadian misalignment that may require a different treatment schedule. This issue is discussed further below.

In fact, there are at least three ways in which circadian misalignment can cause sleep and alertness difficulties. One, circadian misalignment can lead to ASPS and DSPS, because of an unconventional time for sleep propensity. Two, circadian misalignment might affect PSG measures of sleep quality, even if sleep times are conventional. Three, other measures of non-restorative sleep and daytime alertness might be affected by circadian misalignment, even if sleep times and PSG measures are not.

The PSH and SAD: More Recent Findings

Treatment response was evaluated after retrospectively subgrouping subjects in the SAD study who happened to be randomly assigned to the correct treatment vs. the incorrect treatment vs. placebo [7]. The correct treatment was afternoon/evening melatonin for subjects who were prototypically phase delayed before treatment (at baseline) and morning melatonin for subjects who were phase advanced before treatment. The incorrect treatment was morning melatonin for subjects who were phase delayed before treatment and afternoon/evening melatonin for subjects who were phase advanced before treatment. One-third of the subjects were assigned to each treatment regimen.

The placebo response was about 13%, which may have been due to the fact that photoperiod was increasing over the course of the four-week study [7]. This was expectedly low, certainly in comparison to studies in which subjects who are exposed to light (which has a strong placebo
effect). Another reason it was low was because of the instructions we gave to patients and raters: we did not expect to see large changes in either direction in how they felt. These instructions were compatible with the fact that the study was not designed to optimize melatonin treatment efficacy but rather to test the antidepressant mechanism of action for light treatment; that is, all that was needed was a statistically significant difference between treatment groups: in fact, we found about a 20% separation in depression ratings between treatment groups and the most conservative effect size was 0.61, both of which are impressive when compared to fixed-dose antidepressant drug-treatment studies. Thus, in addition to establishing the PSH for SAD, this study establishes therapeutic efficacy for appropriately timed low-dose daytime melatonin treatment.

This study was the first in which symptom severity in a psychiatric disorder was shown to correlate before and in the course of treatment in the same subjects. In fact, the circadian misalignment component was shown to be causal, in that treatment efficacy change scores depended on the degree that circadian misalignment was corrected. Among the 10 phase-delayed (correctly treated) subjects who received afternoon/evening melatonin and 12 phase-advanced (incorrectly treated) subjects who received morning melatonin, only one subject (who shifted the most away from PAD 6) actually worsened on the wrong treatment (and none on the correct treatment). However, adjusting for the 13% placebo response, 5 of 21 subjects worsened and four of these shifted away from PAD 6.

When the depression ratings of the prototypical phase-delayed group taking the treatment of choice (afternoon/evening melatonin) were plotted against PAD, the parabolic minimum occurred at PAD 6. At the vertex, the depression score was 13, which is not too far from the normal range. According to this parabola, in these 11 subjects, 65% of the variance in these inherently noisy depression ratings was explained by the degree of circadian misalignment. Thus, it is possible that this component accounts for most (or perhaps all) of the basis for SAD. However, given that PAD average and range in healthy controls is the same as in SAD patients, at least one other biological or psychological variable must render the SAD patients vulnerable to becoming depressed in the winter when they experience circadian misalignment. By way of analogy, not everyone becomes jet lagged when they travel across time zones.

The “sweet spot” of PAD 6 appears to hold for SAD, in that it was found in the baseline scores when the extant data from an earlier light-treatment study [37] of 49 patients [52,53] were tested a priori. Furthermore, the r-square of the parabola was statistically significant. Moreover, two-thirds of the subjects were phase delayed. Therefore, PAD 6 appears to be heuristically useful, at least for SAD. Perhaps some data sets, particularly for disorders other than SAD, are best fit by a linear regression and not a parabola. But first a review of proposed revisions is in order, along with other criticisms of the PSH.

**Proposed Revisions to the PSH for SAD**

Some investigators have had difficulties with the fact that treatment with evening bright light did not worsen depression in more subjects. In the first morning vs. evening light study, for example, after a baseline week, subjects were randomly assigned to either a week of morning light or evening light, and then crossed-over to the other treatment. During the fourth and final week of the study, subjects received bright light exposure at both times (6–8 a.m.) and (8–10 p.m.). It should be noted that the second bright light pulse is scheduled much later in this study than in the original case report (4–7 p.m.), so as to be more likely to cause a phase delay. Some investigators noted that subjects who received evening light first did not worsen on average. Nor were these investigators impressed by the fact that subjects who received morning light first worsened when switched to evening light [54], because mean ratings in the latter condition were not different from mean ratings at baseline. The explanation of an accompanying placebo
component to light treatment that would counteract evening light’s depressogenic effects (as would be predicted by the PSH) was not persuasive. In our opinion, the placebo component is the same for equal durations of morning and evening light, particularly in the early studies before it became known that morning light was superior (that is, after 1998). In agreement with Eastman 35, we think the placebo component to light is variable between studies, in contrast to what some investigators have opined 55.

While some investigators thought the lack of a placebo control in the first major (cross-over) morning vs. evening light study was problematic, other investigators – ourselves included – thought that the lack of a placebo control in parallel studies that showed no difference between morning and evening light was problematic, rendering interpretation of the results extremely difficult. Furthermore, a parallel-design study is vulnerable to the following confound: raters, and perhaps even subjects, expect the initial phase of the study to precede treatment that is increasingly effective as the study progresses. In one of these studies, for example, treatments were given to separate groups in parallel, following an initial baseline week [32]. Lack of a placebo control in this study would not have been problematic had one treatment proved to work better than another, because raters were unaware of whether the treatment week was testing morning or evening light exposure, as in the cross-over studies. However, since it turned out that there was no difference in depression ratings comparing the two treatments, raters and patients could have been influenced by their knowledge of which weeks were baseline and which were light treatments. Furthermore, without a placebo comparison, it is not possible to distinguish between interpreting the results to mean that both were equally effective or equally ineffective.

Nevertheless, for several years proponents of parallel-design studies [56] were concerned about an order effect in cross-over studies. These concerns arose because some of the earlier cross-over studies showed a greater benefit of morning vs. evening light in the second treatment period compared to the first treatment period. Critics of the PSH favored the importance of the first treatment period (which showed less of a difference between morning and evening light), positing that the second treatment period was confounded by an order effect. Subsequently, however, the antidepressant superiority of morning light was shown to be statistically significant in parallel-design studies, as well as in the first treatment period in cross-over studies, thus rendering this criticism moot; in fact, the three 1998 studies effectively created consensus that morning light was more antidepressant than evening light, at least for most patients with SAD 36–38.

Based on what was thought to be an order effect (that later was not replicated) [57], the first revision to the PSH was posited by the Terman research group, to wit, that bright light is antidepressant in SAD except when it causes phase delays [58]. In other words, the increased efficacy of evening light when it is scheduled as a first treatment is because it causes smaller phase delays than when it is scheduled as a second treatment (in the latter situation the light PRC has been advanced by morning light given in the first treatment period exposing more of the delay zone to be stimulated by evening light). In any event, in their 2001 study the Terman group retracted this revision in favor of a second revision (see below) [57]. Our group, incidentally, interpreted any reduction, statistically significant or not, in the superiority of morning light in the first treatment period to the placebo component of light [37]. As a first treatment, any light treatment is expected to be somewhat antidepressant. As a second treatment, however, evening light suffers in comparison to the subject’s prior benefit with morning light.

Before moving on to the discussion of the revision of the PSH, a brief review of the use of the DLMO may be helpful. The DLMO was assessed in many of the earliest studies of the PSH [26,59,60]. Compared to the average of normal, healthy controls, the average time of the
DLMO was slightly delayed in SAD patients, but in some studies this finding was not statistically significant. The latter case was of concern to some critics of the PSH, despite the fact that the PSH posited that the DLMO was reflecting an ipsative (intra-individual, probably state-dependent) and not necessarily normative difference, that is, most patients with SAD became depressed in the fall/winter at least in part due to a phase delay compared to when they were euthymic in the spring/summer [25]. Therefore, in our opinion, a DLMO in SAD patients that is not delayed compared to controls does not invalidate the PSH. It would not be surprising if the DLMOs of most controls delayed in the winter as well. In fact, it would not be surprising if the DLMOs of most controls also delayed in the winter compared to the summer. There are very few studies of circadian phase across the seasons. In a study that compared SAD patients to controls, the DLMOff (dim light melatonin offset) and SynOff (melatonin synthesis offset) but not the DLMO advanced in the summer in the patients compared to controls [34]. Clearly, more studies across the seasons are needed.

In addition to dichotomous comparisons of the mean DLMOs between SAD patients and controls, correlational analyses were also undertaken, even in the earliest studies. The first such correlation showed a statistically significant correlation between depression ratings and DLMO clock time of the group means for each treatment condition: first baseline, two hours of morning light, 0.5 hour of morning light and second baseline [54]. Of note, in order to keep the raters blind in this parallel study, half of the subjects began with a light-treatment week followed by a baseline week. Nevertheless, even though raters could not know which weeks were treatment weeks, subjects knew. The second such correlation comparing morning and evening light following a baseline week in a cross-over study was also statistically significant [29].

The more meaningful correlational analysis using a separate data point for each individual was first provided by the Terman group [58], utilizing the data from our first two major morning vs. evening light studies [26,29]. A similar analysis in a larger number of their own subjects published in 2001 [57] helped refocus attention on the importance of correlational, rather than dichotomous, analyses, even though the Termans were not able to show antidepressant superiority of morning vs. evening light. In this study, the second revision of the PSH was posited: the Termans proposed that patients with SAD improved depending on the magnitude of the phase advance produced by morning light. In other words, patients who advance their DLMO three hours will do better than those who advance two hours and patients who advance their DLMOs one hour will not do as well as either of the other two groups. The PAD between the DLMO and sleep was not thought to be important, and waketime often has to be scheduled earlier in order to accommodate an early clock time of morning light exposure to provide a sufficient phase advance. According to the original PSH, however, an advance in either bedtime or waketime should be minimized, because this would work against increasing PAD to 6.

This revision differs from the PSH in three ways: one, it does not take into account a phase-advanced subgroup of SAD patients that require a corrective phase delay, which is part of the original PSH; two, the PSH is based on the PAD between the DLMO (and its related rhythms) and the sleep/wake cycle(and its related rhythms), whereas this revision is concerned only with the clock time of the DLMO; three, the PSH envisions a “sweet spot” for the time of the DLMO relative to sleep (allowing for the possibility of over-shifting past the sweet spot), whereas this revision explicitly states that the greater the phase advance, the greater the antidepressant response (in plots of clinical response vs. phase, over-shifting past the sweet spot with very large phase shifts would be described by a parabolic fit of the data and not a linear regression).

There are several important clinical implications related to the differences between the original PSH and this revision. The latter recommends only morning light -- for all patients. The original PSH recommends evening light for the phase-advanced type and provides a way to identify them (a baseline DLMO that is > 6 hours before mid-sleep), whereas work-ups based on this
revision do not include a way to phase type. Nor does this revision allow for the possibility that a patient could become overly phase shifted. According to this revision, patients who have the earliest DLMOs, should receive morning light at the earliest times. Evening light would probably be the treatment of choice for these patients, because according to the PSH they would likely be phase typed as phase advanced based on a DLMO/mid-sleep interval >6 hours. For patients that would qualify for morning light according to both the PSH and this revision, according to the PSH sleep times should not be shifted in the same direction as the DLMO, whereas in order to accommodate for relatively early light morning light exposures, proponents of this revision would have patients awaken earlier than usual in order to achieve the maximal possible phase advance in the clock time of the DLMO.

While sleep times are the way to phase type patients with ASPS and DSPS in order to inform the correct times to administer phase-resetting agents, phase typing of SAD patients should be based on whether or not PAD is greater or less than six. Sleep times alone are not a reliable way to do phase typing. In fact, if SAD patients have delayed sleep, they may be even more likely to have a DLMO that is relatively advanced with respect to sleep. The Terman group continues to recommend sleep times and morningness/eveningness ratings (that to some extent correlate with sleep time) to specify how early a patient should be awakened in order to accommodate relatively early light exposures, although recently some exceptions for recommending evening light exposure have recently been considered [61].

Morningness/evening ratings are influenced by sleep times. As predicted by the PSH and recently shown [62], the typical patient with SAD has circadian rhythms that are delayed (perhaps the result of a long intrinsic circadian period), but is required to sleep at an earlier time in the winter than desired. Interestingly, all three individuals with long intrinsic circadian periods who were morning types were retrospectively determined to have SAD [63]. These three individuals would be expected to have a DLMO that is phase delayed with respect to mid-sleep, because people with longer intrinsic circadian periods should have a delayed phase angle of entrainment. In any event, morningness/eveningness ratings (which correlate with both DLMO clock times and sleep times) may be of questionable benefit in predicting whether or not a SAD patient will preferentially benefit from morning vs. evening light.

As mentioned above, the sweet spot of PAD 6 has been found now in two different groups of SAD patients. However, it may turn out that other patients should be phase typed on a more individual basis. So far, this seems to be the case with patients who have unipolar non-seasonal depression, in that their data is best fit linearly and not parabolically [64]. Since it has only been seven years since the Terman linear correlational study was published [57] and only two years since our parabolic correlation [7], the jury is out as to how many SAD patients are of the phase-advanced type and how likely is over-shifting of the phase-delayed types.

It may turn out that the Terman revision of the PSH for SAD may be more applicable to other groups of patients, at least with respect to a linear vs. a parabolic plot of the data. However, even with linear analyses, we recommend that future studies plot data differently than the Terman group [57]. Although their post-treatment change scores in depression ratings were based on the change from the initial baseline condition for all subjects; these were apparently plotted against change in DLMO clock time from the initial baseline condition of only those subjects who received morning light first; for those subjects who received evening light first, this condition (that is evening light), instead of the baseline condition, was apparently used in the change scores for DLMO clock times, even though the baseline condition was used in the change scores for depression ratings in the same individuals. Therefore, if the discussion above is not mistaken, the change scores are not consistent, and for half of the subjects the change scores for the DLMO clock times are not based on the same initial condition and time point as the change scores for the depression ratings.
It may also turn out that DLMO clock time confers a greater level of statistical significance than PAD when plotted against depression ratings, particularly in linear plots. Nevertheless, we continue to think that PAD is to be preferred over DLMO clock time for several reasons, even though we and others have found DLMO clock times to lead to significant analyses when sleep times are held constant, rendering the difference between DLMO circadian and clock times moot, at least for changes during the course of the study. However, the same cannot be said for initial phase typing. PAD also takes into account different preferred sleep times in individuals in which sleep may have nothing to do with their disorder. PAD also takes into account various influences on sleep times that may or may not have to do with their disorder or clinical state and are more likely to be causal of sleep and psychiatric symptoms, whereas changes in DLMO clock time could easily be the result of changes in sleep times vis-à-vis the perceived light/dark cycle.

In any event, the recommendations we made two decades ago [65] appear to be current with one modification (see table 1), first suggested for SAD by the Terman group [30]. The first SAD studies used light of 2000–2,500 lux [2,26,66]. The Terman group suggested using 10,000 lux in the treatment of SAD [30]. There is probably intensity/duration reciprocity, so that a shorter duration of 10,000 lux than originally recommended for 2,000 lux, at least for maintenance after induction of the treatment response (that can take up to two weeks to be complete).

Research continues to uphold the original treatment recommendations for SAD [65], which considered the possibility that patients could be overly phase shifted with too much light. Most (at least 2/3rds) of SAD patients are phase delayed (that is, have DLMO that are delayed with respect to their sleep/wake cycle). Therefore, bright light should be tried first in the morning, even in patients with early sleep times. In fact, these patients likely have a DLMO that is delayed with respect to the sleep/wake cycle, in that they likely have a sleep/wake cycle that is advanced with respect to the DLMO. In fact, the modal patient is prevented from indulging their inclination to sleep later in the winter because of work or family obligations. Years ago, we found preliminary evidence for the therapeutic efficacy of delaying sleep in SAD [67], which is another way of accomplishing the goal of having waketime coincide with bright exposure (even on a cloudy, winter day, sunlight exposure an hour or more after dawn is at least 10,000 lux). These recommendations will be refined as salivary DLMOs become commercially available. Samples can be collected at home. In almost all cases, the DLMO occurs before sleep onset, so collections can be completed before bedtime. Another very important feature that will make this test much more convenient is the use of orange goggles that will obviate the need for dim light, or at least light that is so dim that reading is uncomfortable, which is the standard recommendation.

The PSH as Applied to Other Disorders that May Have a Circadian Misalignment Component

As with the study of blind people (see chapter by Uchiyama and Lockley this volume), SAD is a useful model for studying the effects of light deprivation on the circadian system of humans. Totally blind people who completely lack light perception provide an unfortunate but useful experiment of nature which is the only way to study human circadian rhythms in the absence of the confounding effects of light [48,68–72]. Of course, light deprivation in the winter compared to the summer is not comparable to what a totally blind person experiences. However, SAD is an excellent model for a circadian rhythm affective disorder and perhaps for a certain type of circadian rhythm sleep disorder, non-restorative sleep due to internal circadian misalignment.
People with circadian misalignment, even when they have little difficulty going to sleep and waking up at conventional times, may have non-restorative sleep, which appears to be part of the dysphoric mood constellation of symptoms in patients with SAD and in people with unipolar non-seasonal depression [64]. In an item analysis of the SIGH-SAD baseline ratings of the subjects in the melatonin-administration study [7], three items accounted for the results found for the 29 items, even though the former had a range that was 1/10th of the latter [53]. These three items were: depression severity, psychic anxiety severity and severity of agitation observed by the interviewer. These three items might constitute the nub of an endotype that corresponds to the circadian misalignment component for SAD; of note, these three items resulted in a statistically significant parabolic correlation with PAD in the baseline data in our latest light-treatment SAD study extant data set [37]. Perhaps a subgroup of people have non-restorative sleep when their circadian rhythms become misaligned, even in the absence of any other symptoms. It would not be surprising, therefore, if the salivary DLMO was incorporated into the standard PSG test routinely done in clinical sleep labs, particularly if there is a low level of suspicion that the PSG will reveal any abnormalities. The PSG, following a few days of documenting sleep and wake times, will allow calculation of the PAD. The sleep lab PAD could be followed up with additional assessments to provide even more information about the relationship between the patient’s sleep and psychiatric symptoms and circadian misalignment.

Conclusion

Clearly, the DLMO/mid-sleep PAD needs to be assessed in a variety of sleep and psychiatric disorders. Thus far, we have found that symptom severity -- as measured by the Connor’s parent rating cognitive problems/inattention subscale and ADHD index subscale in attention deficit hyperactivity disorder [73], the Ham-21 depression scores in non-seasonal depression [64] and the Profile of Mood States Brief form ratings in healthy medical students -- correlates with PAD, in that the more the DLMO is phase delayed with respect to mid-sleep, the greater the symptom severity [74]. Correlations, either linear or parabolic, with symptom severity should lead to the safe and effective use of bright light and/or melatonin at least as add-on treatments, assuming that the circadian misalignment is causal, as it is in SAD, even if it accounts for a small component of the disorder. Patients will probably have to be phase typed on an individual basis, in which PAD 6 may or may not be as useful as it is in SAD. An alternative way to phase type on an individual basis would be to determine whether there is a positive or negative slope, even if not statistically significant, on at least 3–4 data points in which symptom severity is plotted against PAD: individuals who tend to be more symptomatic when more delayed could be phase typed as delayed (and vice versa). Low-dose melatonin (or bright light) could then be scheduled at the correct time (to cause an increase or a decrease in PAD, respectively) to validate phase type and treatment parameters, as well as to determine causality between PAD and symptom severity. Also, mid-sleep may not turn out to be the best marker for sleep phase when calculating PAD, although it does take into account inter-individual (and intra-individual) differences in sleep duration. For continued monitoring of therapeutic efficacy, DLMO and PAD will likely be very important, in addition to clinical assessment of improvement or relapse. The DLMO is also the very best way to identify the phase of the light and melatonin PRCs and thus will help optimize when to schedule treatment times more precisely. In conclusion, the above work may lead towards an extension of a heuristically useful model in medicine and psychiatry. The role of light, melatonin and biological rhythms in a biopsychosocial-environmental model has yet to be fully understood and appreciated.

Acknowledgments

This work was supported by Grant Nos. 5R01 HD042125, 5-R01-AG021826-02, and 5R01EY018312 from the National Institutes of Health. A.J.L. was also supported by a National Alliance for Research on Schizophrenia and Depression Distinguished Investigator Award, and J.S.E. was supported by a National Alliance for Research on Schizophrenia and Depression Junior Investigator Award.
References


32. Wirz-Justice A. Light therapy in seasonal affective disorder is independent of time of day or circadian phase. Arch Gen Psychiatry 1993;50


Figure 1.
Schematic diagram of normal phase relationships (rounded to the nearest integer) between sleep phase markers, the 10 pg/ml plasma dim light melatonin onset (DLMO) derived from historical controls. The present study used the melatonin/mid-sleep interval (phase angle difference, or PAD) of 6 hours as the hypothesized therapeutic window for optimal circadian alignment. Sleep times were determined actigraphically. Plasma melatonin levels were obtained under dim light every 30 minutes in the evening. The operational definition of the melatonin onset is the interpolated time of continuous rise above the threshold of 10 pg/ml; for example, if the melatonin level at 8 p.m. was 5 pg/ml and at 8:30 p.m. was 15 pg /ml, the melatonin onset would be 8:15 p.m. Adapted from[7], with permission.
Figure 2. **Figure 2 (left).** Effect of light on melatonin secretion. Each point represents the mean concentration of melatonin (+/− standard error) for six subjects. **Figure 2 (right).** Effect of different light intensities on melatonin secretion. The averaged values for two subjects are shown. Symbols: (O) 500 lux; (X) 2500 lux; (●) 1500 lux; and (□) asleep in the dark. Melatonin levels were measured by mass spectrometry [13]. These early studies were responsible for an increased awareness of the importance of the light/dark cycle as a zeitgeber (time cue) for human circadian rhythms and for the use of the dim light melatonin onset (DLMO) as a circadian phase marker and of bright light as a circadian phase resetting agent in the treatment of circadian phase disorders, including winter depression.
(SAD) and the circadian disorders experienced by totally blind people. From [21], with permission.
Figure 3.
The optimal times to schedule bright light exposure and low-dose melatonin administration to cause circadian phase shifts are based on their respective phase response curves (PRCs) which are about 12 hours out of phase with each other. The 10 pg/ml plasma (3 pg/ml saliva) melatonin onset marking circadian time (CT) 14, can be used to indicate when advance and delay responses occur, in order to maximize phase shifts. The crossover times are eight hours before (circadian time 6), and four hours after (circadian time 18), the melatonin onset. Also indicated are clock times typical for individuals who awaken at 6 a.m. (0600). Optimally, exogenous melatonin should overlap with either the onset or the offset of the endogenous melatonin profile. High doses (greater than about 5 mg) may be less effective than lower doses, because of spillover onto the wrong zone of the melatonin phase response curve. The crossover times for the light PRC are based on the one published by Czeisler and co-workers [75] in the Johnson Atlas of PRCs [76]; the optimal light times for scheduling light are based on earlier work [26] and the melatonin crossover and optimal scheduling times are based on the melatonin PRC [4939]. Adapted from [62] with permission.
Figure 4.
Percent change in SIGH-SAD score as a function of net change in absolute deviation toward and away from PAD 6 in PM-melatonin treated advanced and delayed subjects. 13% has been added to the change in SIGH-SAD score to remove the average placebo response. Pretreatment vs. post-treatment shifts with respect to PAD 6 account for 35% of the variance. Adapted from [7], with permission.
TABLE 1

Treatment Guidelines for Patients with Seasonal Affective Disorder [Adapted from 65]

<table>
<thead>
<tr>
<th>Treatment Guidelines for Patients with Seasonal Affective Disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>• If patients do not have early morning awakening, schedule 1–2 hours of 2500–10,000 lux exposure immediately upon awakening.</td>
</tr>
<tr>
<td>• If patients begin treatment on the weekend, they may not have to arise earlier to accommodate the morning light exposure; early rising may retard the response for a few days.</td>
</tr>
<tr>
<td>• The response begins 2 to 4 days after beginning light therapy and is usually complete within 2 weeks.</td>
</tr>
<tr>
<td>• These patients should minimize any advance in their sleep time and should avoid bright light in the evening.</td>
</tr>
<tr>
<td>• If patients do not respond to treatment, they may need a longer duration of morning light.</td>
</tr>
<tr>
<td>• If patients respond only transiently or begin to complain of early morning awakening or severe fatigue in the evening, they may be becoming overly phase advanced due to too much morning light. The duration of morning light should be reduced but still begun immediately upon awakening or some late evening light exposure could be added.</td>
</tr>
<tr>
<td>• Some patients may respond to an immediate “energizing” effect of bright light exposure (this may be a placebo effect), which if not administered too late in the evening might be helpful.</td>
</tr>
<tr>
<td>• Once a response has been achieved, the duration and frequency of light exposures can be reduced. Always begin light exposure immediately upon awakening or a little later if patients become overly phase advanced.</td>
</tr>
<tr>
<td>• If there is still no response, a trial of evening bright light (7–9 pm) may be necessary. These patients should minimize any delay in their sleep time and should avoid bright light in the morning.</td>
</tr>
<tr>
<td>• Appropriate precautions should be taken to avoid any possibility of eye discomfort or injury (e.g., an eye history and exam if indicated, instructions never to stare at the sun, use of safe artificial light sources, and recommendation of follow-up check ups).</td>
</tr>
</tbody>
</table>