Journal of Consulting and Clinical Psychology
2001, Vol. 69, No. 2, 227–239

0022-006X/01/$5.00 DOI: 10.1037//0022-006X.69.2.227

Relaxation and Sleep Compression for Late-Life Insomnia: A Placebo-Controlled Trial

Kenneth L. Lichstein
University of Memphis and Methodist Healthcare of Memphis

Brant W. Riedel and Nancy M. Wilson
University of Memphis

Kristin W. Lester and R. Neal Aguillard
Methodist Healthcare of Memphis

Older adults with insomnia were recruited from the community and randomized to treatments: relaxation, sleep compression, and placebo desensitization. Questionnaire data collected at baseline, posttreatment, and 1-year follow-up and polysomnography data collected at baseline and follow-up yielded the following conclusions: All treatments improved self-reported sleep, but objective sleep was unchanged. Clinical significance analyses yielded the strongest findings supporting the active treatments and suggested that sleep compression was most effective. Results partially supported the conclusion that individuals with high daytime impairment (i.e., fatigue) respond best to treatments that extend sleep, as in relaxation, and individuals with low daytime impairment respond best to treatments that consolidate sleep, as in sleep compression. Strong methodological features including a placebo condition and a treatment implementation scheme elevate the confidence due these findings.

Chronic insomnia, referring to persistent difficulty sleeping, may have a pervasive impact on one's quality of life. A large body of data identifies disturbed mood and anxiety and perceived compromised quality of life as common sequelae of insomnia (Riedel & Lichstein, 2000).

Insomnia in older adults is more common and more troublesome than it is in younger people. Insomnia prevalence in older people often exceeds 25% (e.g., Mellinger, Balter, & Uhlenhuth, 1985), and these same surveys found, in samples of older people, a 30–50% higher rate of insomnia than in samples of younger people. Older adults with insomnia (OAWI) turn to sleep medication at a disproportionately high rate, risking addiction, polypharmacy interactions, exacerbation of sleep apnea (periods of breath cessation lasting 10 s or longer), which is more common in older people, and multiple other side effects (Mellinger & Balter, 1981; Roth, Zorick, Wittig, & Roehrs, 1982).

The combination of high treatment need for OAWI and untoward side effects from hypnotic medications invites great interest in developing psychological interventions, but this has been slow to materialize. Our comprehensive review (Lichstein & Fischer, 1985) in the mid 1980s of psychological interventions for insomnia found 57 studies, but only 3 focused on insomnia in seniors.

In the past decade, interest in treating OAWI has accelerated. Our recent review of psychological interventions for OAWI found about 15 clinical outcome studies (Lichstein, Riedel, & Means, 1999). The main interventions tested were sleep restriction–sleep compression (reducing time in bed to closely match actual time slept), relaxation, stimulus control (eliminating sleep-incompatible behaviors, including wakefulness, from the bedroom), and cognitive therapy, and all claimed at least moderate success.

Unfortunately, there are serious methodological shortcomings with the extant group of studies treating OAWI. First, these treatments rarely convert insomniacs to noninsomniacs. Typically, the treatments improve the insomnia, but there remains a residual less severe problem. There is a clear need for more powerful treatments and for effective tailoring of treatments to individual characteristics. Second, there has been only one placebo-controlled trial in this area, and it used a pill-placebo group (Morin, Colecchi, Stone, Sood, & Brink, 1999). No study has used a psychological treatment placebo. The role of social support, social influence, and expectancy for improvement cannot be ruled out from the majority of clinical trials with OAWI. Third, only three studies of psychological interventions for geriatric insomnia (Engle-Friedman, Bootzin, Hazlewood, & Tsao, 1992; Morin et al., 1999; Morin, Kowatch, Barry, & Walton, 1993) used polysomnography (PSG) to supplement self-reported sleep in assessing treatment effects. In these studies, treatment outcome was weaker by PSG than by self-report, a finding common to younger individuals as well. The present clinical trial attempted to address these three faults.

Kenneth L. Lichstein, Department of Psychology, University of Memphis, and Sleep Disorders Center, Methodist Healthcare of Memphis, Memphis, Tennessee; Brant W. Riedel and Nancy M. Wilson, Department of Psychology, University of Memphis; Kristin W. Lester and R. Neal Aguillard, Sleep Disorders Center, Methodist Healthcare of Memphis.

This research was supported by Grant AG12136 from the National Institute on Aging, by Methodist Healthcare of Memphis, and by the University of Memphis, Department of Psychology's Center for Applied Psychological Research, part of the State of Tennessee's Center of Excellence grant program.

Correspondence concerning this article should be addressed to Kenneth L. Lichstein, Sleep Research Project, Department of Psychology, University of Memphis, 202 Psychology Building, Memphis, Tennessee 38152-3230. Electronic mail may be sent to lichstein@mail.psyc.memphis.edu.

There is a special problem in the diagnosis of insomnia in older adults that merits attention. Age-correlated deterioration in sleep architecture and pattern found in many older adults creates a characteristic sleep pattern among seniors that mimics key characteristics of middle-aged insomnia (Morgan, 1987; Williams, Karacan, & Hursch, 1974). As one's age advances past 60 years, there is ascending likelihood of the following sleep changes: sleep will become lighter, sleep will become more fragmented, and total sleep time will become highly variable, but on the average will decline from 7.5 hr per 24-hr period in middle-aged samples to 6 hr among older adults.

We have advanced a biodevelopmental model of insomnia (Riedel & Lichstein, 2000) to incorporate the sleep changes into an understanding of the insomnia complaint. Biological (constitutional inclination toward short sleep need) and developmental (sleep deterioration associated with the normal aging process) influences abbreviate sleep need and lead to insomnia if there is no corresponding adjustment of sleep goals. Older persons uninformed as to their changing sleep capacity and motivated to preserve a middle-aged sleep pattern may seek treatment to correct perceived pathological sleep. These individuals compose a subgroup of OAWI who present with a disturbed sleep pattern but no excessive daytime impairment (e.g., sleepiness, inattentiveness, or retarded fine motor skills), the presumed legacy of inadequate sleep. This suggests biological sleep needs have been satisfied despite fragmented sleep, and we refer to such individuals as *insomnoids* (Riedel & Lichstein, 2000). Insomnoids can be expected to respond poorly to conventional treatments, such as relaxation, which aim to extend sleep.

The present study factorially contrasted types of psychological treatments (relaxation, sleep compression, and placebo therapy) and levels of daytime impairment (high or low as determined by baseline self-report (SR)) in older adults complaining of insomnia. We investigated which of the active treatments is more effective with elderly adults, which is more effective than placebo treatment and which is more effective with insomnia subtypes. A Treatment × Participant Characteristic interaction was hypothesized whereby the subgroup that failed to satisfy biological sleep needs (as evidenced by high daytime impairment) should benefit the most from the intervention (relaxation) aimed at inducing and extending sleep, and the subgroup not evidencing sleep deprivation effects (low daytime impairment) should benefit the most from the intervention (sleep compression) aimed at eliminating unneeded, prolonged time in bed and resulting perceived insomnia.

Method

Participants

We used public service announcements to recruit volunteers from the community. Participants were compensated with free sleep treatment and $200 paid on an all-or-none basis.

A total of 1,376 people inquired about the study. Forty-five of these people could not be reached for the initial telephone interview despite multiple attempts. A group of 381 declined to participate, and 861 did not satisfy screening criteria. The most common reasons for disqualification were as follows: applicant was too young, sleep apnea was suspected or confirmed by PSG, sleep medication use, or insomnia was secondary to a medical or psychiatric disorder. We found a high rate of apnea that escaped detection by SR measures, especially among men, which emphasizes the need for PSG screening in an older population (Lichstein, Riedel, Lester, & Aguillard, 1999).

A total of 89 participants passed the baseline screening process and were randomly assigned to treatment. Two participants from each treatment condition withdrew from the study during treatment. The following reasons were given for withdrawal: time constraints (2 placebo), family illness (1 sleep compression), depression (1 sleep compression), and no specific reason (2 relaxation). Five participants completed the study through post-treatment but did not return for follow-up. The following reasons were given for missing follow-up: time constraints (1 sleep compression, 1 placebo), personal health problems (1 relaxation, 1 placebo), and family illness and sleep medication use (1 placebo). An additional 4 participants (3 sleep compression, 1 placebo) were eliminated from the study because they were found to have significant sleep apnea at follow-up. This left a sample of 74 participants (27 relaxation, 24 sleep compression, 23 placebo) who completed the study. Two of these participants (1 relaxation, 1 placebo) were missing all self-report data at one phase of the study but completed all other aspects of the study.

We conducted a number of analyses to determine if there were significant differences between treatment conditions on demographic and sleep variables collected during the participant screening process. The mean age of participants was 68.03 years (SD = 7.04, range = 59–92), $F(2, 71) = 0.01$, $p = .99$. More women (n = 53) than men (n = 21) participated, but the female-to-male ratio did not differ significantly across treatment conditions, $\chi^2(2, N = 74) = 0.98$, $p = .61$. In addition, no treatment group differences were observed for race (60 White, 12 African American, 2 Hispanic), $\chi^2(4, N = 74) = 1.81$, $p = .77$; marital status (58% married), $\chi^2(2, N = 74) = 4.41$, $p = .11$; or education level (M = 14.70, SD = 2.89), $F(2, 71) = 0.85$, $p = .43$. Also, there were no treatment condition differences on sleep-related variables such as duration of insomnia (M = 8.93 years, SD = 11.54, range = 6 months–51 years), $F(2, 71) = 2.38$, $p = .10$; number of nights per week insomnia was experienced (M = 5.34, SD = 1.72), $F(2, 71) = 0.51$, $p = .60$; and type of insomnia (maintenance: 56, sleep-onset: 3, mixture of maintenance and sleep-onset: 15), $\chi^2(4, N = 74) = 0.19$, $p = .99$. A breakdown of these data is provided in Table 1.

Participant Screening

Participant screening proceeded in four stages.

Stage 1. We used a 20-min structured telephone interview to collect information on the following criteria.

1. The key sleep criteria were difficulty initiating or maintaining sleep for at least 6 months, complaint of impaired daytime functioning, and indications of learned sleep-preventing associations (psychophysiologic insomnia, American Sleep Disorders Association, 1990). We added the following quantitative criteria: sleep onset or awake time during the night must exceed 30 min, and insomnia must be present on the average at least three times per week.
2. The volunteer is 59 or older, is free from other sleep disorders, and is free from medical or psychiatric disorders or sleep-active medications (for the past month) that could affect sleep.
3. We screened out volunteers who used high levels of caffeine (consumes caffeinated beverages past 2:00 PM), nicotine (10 or more cigarettes a day), or alcohol (consumes more than seven alcoholic beverages per week or consumes alcohol at bedtime at least once per week).

Stage 2. We applied additional criteria in a face-to-face interview to rule out medical and psychiatric contributors to insomnia and neurological deficits.

1. We used the Trait scale of the State-Trait Anxiety Inventory (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) to screen for anxiety (cutoff of 43), and the Geriatric Depression Scale (Yesavage et al., 1983) to screen for depression (cutoff of 15).
2. We used the Mini–Mental State Exam (Folstein, Folstein, &

Table 1
Participant Characteristics

Variable	Relaxation	Sleep compression	Placebo
	Parametric variables		
M (and *SD*) for:			
Age	68.11 (8.31)	67.92 (6.77)	68.04 (5.91)
Education (years)	14.96 (3.28)	14.06 (2.12)	15.04 (3.13)
Insomnia duration (years)	12.37 (15.89)	8.42 (9.30)	5.41 (5.04)
No. of insomnia nights per week	5.59 (1.61)	5.27 (1.91)	5.11 (1.69)
	Nonparametric variables		
Gender	21 women	17 women	15 women
	6 men	7 men	8 men
Race	23 White	18 White	19 White
	4 African American	5 African American	3 African American
	0 Hispanic	1 Hispanic	1 Hispanic
Marital status	19 married	10 married	14 married
	8 not married	14 not married	9 not married
Type of insomnia	20 maintenance	18 maintenance	18 maintenance
	1 sleep-onset	1 sleep-onset	1 sleep-onset
	6 mixed	5 mixed	4 mixed

McHugh, 1975) to screen for cognitive impairment. We chose a conservative criterion of 27 to gain entry. For individuals with less than a 9th-grade education, the score had to exceed 20, because their performance on this test is burdened by inadequate educational preparation (Murden, McRae, Kaner, & Bucknam, 1991).

3. We used the Cornell Medical Index (CMI, Brodman, Erdmann, Wolff, & Miskovitz, 1986) to screen for general physical health. The CMI consists of about 220 questions inquiring about the presence of a wide range of symptoms or diseases. We inquired how any "yes" response affected sleep. Individuals were disqualified if their reported history revealed variations in insomnia severity mirrored the changing course of their medical status.

Stage 3. The validity of the insomnia complaint was judged from 2 weeks of self-report data. These data must satisfy the criteria applied in the telephone interview.

Stage 4. In the baseline PSG, individuals exhibiting sleep apnea breathing interruptions (apnea or hypopnea [partial apneas] lasting at least 10 s and occurring at the rate of 15 or more per hr), periodic limb movements (producing myoclonic arousals at the rate of 10 or more per hr), or other sleep disorders beside insomnia were terminated from the study. Also, a urine sample was collected on the second of each pair of PSGs. Volunteers were disqualified if these were positive for sleep medications or other drugs having sleep-active properties.

Research Setting and Apparatus

The study was conducted at two settings, the Psychological Services Center, The University of Memphis, and the Sleep Disorders Center, Methodist Healthcare of Memphis.

A Nihon Koden #4312 polygraph was used for the all-night sleep studies (PSG). Monitoring on the first night of each pair of sleep studies consisted of two electroencephalography (EEG) measures, two electro-oculography (EOG), and chin electromyography (EMG) according to standard placements (Rechtschaffen & Kales, 1968) to score sleep stages. Supplementary channels included oxygen saturation level, bilateral anterior tibialis EMG, heart rate (EKG), thoracic strain gauge, and a nasal-oral thermistor. These are needed to screen out veiled associated sleep disorders that mimic the clinical appearance of insomnia. The second night of each pair of sleep studies consisted of an abbreviated protocol of EEG, EOG, and chin EMG to score sleep stages. In the event the first night yielded ambiguous findings, the full hook-up used the first night was again used.

Drug Screens

Participants were repeatedly warned that a urine screen would be taken on one of the two nights of the baseline and of the follow-up sleep studies and that positive findings for sleep-active medications was a disqualifier. In fact, the urine samples were always taken on the second night of both series of PSGs. This procedure helped guarantee that both PSG nights were drug free. Also, participants were informed that one urine screen would be taken at a random point during the treatment period from a randomly selected subset of the participants. In fact, none were taken, but its anticipation was kept alive throughout the treatment phase.

Dependent Measures

Daytime functioning. The Insomnia Impact Scale (IIS, Hoelscher, Ware, & Bond, 1993) questionnaire contains 40 negative statements about the daytime impact of sleep. These statements sample five areas of impairment: physical, cognitive, emotional, social, and occupational. Respondents rate each item on a 5-point scale and ratings are summed.

The Beliefs and Attitudes about Sleep Scale (BASS, Morin, 1993) contains 30 statements of beliefs about sleep tapping potential distortions and exaggerations. In consultation with Morin, we have made the following small changes. First, a higher score indicated less distortion on just one of the items (#23). For all others, a higher score indicated greater distortion. Item 23 was reworded to agree with all other items. Second, we replaced the continuous analogue scale with a Likert scale to simplify scoring.

The Fatigue Severity Scale (FSS, Krupp, LaRocca, Muir-Nash, & Steinberg, 1989) is composed of 9 items asserting the intrusion of fatigue in different aspects of living. Each item is rated from 1 = *strongly disagree* to 7 = *strongly agree*, and the test score is the average rating.

The Epworth Sleepiness Scale (ESS, Johns, 1991) measures trait daytime sleepiness. Respondents rated how likely they were to doze in eight commonly encountered restful situations. The scale ranges from 0 (*would never doze*) to 3 (*high chance of dozing*).

Nighttime sleep. Participants slept at the Sleep Disorders Center for 2 consecutive nights at baseline and at 1-year follow-up. PSG records were

manually scored in 30-s epochs by a registered PSG technician uninformed as to treatment condition, according to the criteria of Rechtschaffen and Kales (1968). A second technician randomly selected one third of the records and scored them independently. Discrepancies were resolved by a meeting of the technicians.

The records yielded sleep stage percentage (stages awake, light sleep [Stages 1 and 2], deep sleep [Stages 3 and 4], and REM), and absolute values for initial sleep latency, total sleep time (TST, actual time slept), wake time after sleep onset (WASO), number of awakenings during the night, and sleep efficiency percent (SE, the ratio of TST to total time in Bed × 100).

To collect SR sleep data, we gave the participants a sleep questionnaire (given in Lichstein, Riedel, & Means, 1999) each morning for 2 weeks at baseline, posttreatment, and follow-up. The questionnaire yields the same sleep-pattern data calculated from the PSG: initial sleep latency, TST, WASO, number of awakenings during the night, and SE, as well as our sole measure of daytime napping. The sleep questionnaire also contained a rating scale assessing perceived quality of sleep (from 1 = *very poor* to 5 = *excellent*). All SR data were double scored and all discrepancies resolved. For purposes of data analysis, the SR sleep data were collapsed over each 2-week period.

Therapists

Graduate students in clinical psychology served as therapists. Each treated an equal number of participants from each group. Therapists were trained according to the procedures outlined in the Treatment Implementation section (see delivery induction).

Treatment Implementation Variables

Lichstein, Riedel, and Grieve (1994) proposed a treatment implementation model whereby steps must be taken to ensure that the treatment is delivered as intended (*delivery*), is comprehended by the patient as intended (*receipt*), and is practiced out of session as intended (*enactment*). Induction strategies were used (e.g., detailed treatment manuals and mock treatment sessions to induce delivery, tailoring treatments to maximize treatment acceptance by the participant to induce receipt, and reminder sheets to comply with home assignments to induce enactment) to heighten the likelihood that the treatment components were properly implemented. We used assessments (e.g., listening to tapes and rating treatment sessions to assess delivery, using relaxation ratings and a stimulus control quiz to assess receipt, and using compliance logs to assess enactment) to determine the degree of implementation in each component.[1]

Treatment Credibility

Treatment credibility was evaluated by ratings on four 10-point scales (higher ratings reflected higher credibility) with statements tapping the following dimensions: (a) reasonableness of treatment, (b) opinion of therapist, (c) expectation for improvement, and (d) willingness to recommend treatment to a friend.

Procedures

Volunteers received a screening telephone interview, and a follow-up screening interview at the Sleep Disorders Center, when informed consent was obtained. Participants who passed these hurdles were given daily sleep questionnaires to cover 14 nights, an IIS, an BASS, an FSS, and an ESS. Participants returned this material in a franked envelope. This procedure had a threefold purpose: (a) to obtain verification of insomnia, (b) to collect baseline data, and (c) to establish that the volunteer could comply with simple instructions and supply questionnaire information.

Qualifying participants were stratified on gender, SR SE, and IIS scores based on estimated median splits and randomly assigned to the three conditions within strata. All participants were exposed to a common set of procedures as follows.

Baseline. Having already collected SR data, participants submitted to two consecutive PSGs, including urine screens on night two. When the light went out was determined by the participants, and sleep recording continued until the participants signaled that they wanted to arise in the morning.

Treatment. Treatment commenced within 2 weeks following the PSG assessments. Treatment consisted of 6 weekly individual sessions, each lasting about 45 min. All participants received sleep hygiene instructions in the first treatment session. These are composed of seven instructions: avoid caffeine after noon, avoid naps past 2:00 pm, avoid exercise within 2 hr of bedtime, avoid nicotine within 2 hr of bedtime, avoid alcohol within 2 hr of bedtime, avoid heavy meals within 2 hr of bedtime, and arise at a fixed time each morning plus or minus .5 hr.

The treatment credibility questionnaire was administered at the end of the second treatment session. This point was selected because the participant had sufficient exposure to the procedure and to the therapist to make an informed judgment but insufficient treatment exposure for treatment response to bias the questionnaire responses. During the first and fourth treatment sessions, the therapist reminded the participant that a urine screen would be taken from a randomly selected subset of participants during a randomly selected treatment session.

Posttreatment and follow-up. Posttreatment and follow-up assessment repeated the entire baseline assessment, except PSGs were not conducted at posttreatment.

Treatment Conditions

Relaxation therapy (REL). We used a 10-min hybrid REL procedure (given verbatim in Lichstein, 2000) consisting of (a) emphasizing a relaxed attitude, (b) directing five slow, deep breaths including a softly spoken "relax" self-instruction with each exhale, (c) slowly reviewing the body in sequential parts while focusing on relaxed sensations, sometimes termed *passive relaxation,* and (d) directing slow, silent repetition of the autogenic phrase, "I am at peace, my arms and legs are heavy and warm." Participants were introduced to REL during the first session. Remaining sessions were used to refine the participant's technique, tailor the technique to the individual participant, and troubleshoot any problems the participant may have. Participants were strongly encouraged to practice REL at home twice a day: once at any time and once at bedtime.

Sleep compression (SC). This treatment comprises devising idiographic sleep goals and gradually conforming one's sleep schedule to these goals. Mean TST was determined from the 2 weeks of baseline sleep questionnaires, as was total time spent in bed. The difference was divided by 5, and allotted time in bed was compressed by this amount weekly by delaying bedtime or advancing arise time assisted by an alarm clock. Under unusual circumstances, if the participant was accustomed to daytime napping, and the participant was resistant to abandoning the nap, this habit was accommodated in the distribution of time in bed. If napping was included, the nap should be no more than 30 min and should be terminated by 2 PM. Participants in this condition only continued SR sleep recording during the treatment phase to permit monitoring of the sleep schedule needed for the treatment.

Treatment sessions one through five presented a revised sleep schedule each week. By the last session, time in bed had been gradually compressed to match sleep time. Some participants were resistant to this goal and we negotiated within 30 min of baseline TST with these individuals. Also, time in bed was increased by 15 min if SE surpassed 90%. The sixth and

[1] A more complete rendition of the treatment implementation scheme and associated analyses are available from Kenneth L. Lichstein.

final treatment session was used to review progress and encourage future adherence to the new schedule. Weekly discussions included planning activities such as reading, hobbies, and chores to constructively fill the out-of-bed time liberated by SC, so that sleep or time in bed does not become a reaction to boredom.

Placebo desensitization (PL). We used a PL treatment introduced by Steinmark and Borkovec (1974), which they called quasi-desensitization. The procedure mimics the structure of Wolpe's systematic desensitization and is the most frequently used PL treatment in the insomnia literature.

The procedure consisted of constructing a temporal hierarchy of 10 bedtime events, such as a snack at 10 PM, brush teeth, etc. Including this many events in the hierarchy made the length of this procedure comparable to that of our REL condition. The participant alternately imagined a hierarchy event (30 s) and a neutral scene (30 s). This procedure is expected to be therapeutically inert, because if any of the bedtime events do provoke anxiety, this procedure should not diminish the resulting sleep disruption. The neutral paired scene replaces relaxation in systematic desensitization and is not expected to reproduce relaxation's palliative effect on hierarchy events.

Like participants in the REL group, these participants were instructed to practice the procedure at home twice a day, with the second practice to conclude not less than 2 hr from bedtime to minimize the possibility that it would exert adverse or beneficial effects on sleep. Individuals in this condition were offered treatment by either REL or SC at the conclusion of their participation.

Results

First Night Effect and PSG Sleep

The first night effect refers to the tendency for individuals to sleep worse than usual during an initial night of laboratory PSG. To test the presence of a first night effect, we conducted a repeated measures multivariate analysis of variance (MANOVA) across the two baseline PSG nights. PSG-measured sleep latency, number of awakenings, WASO, TST, SE, Stages 1 and 2 sleep percentage, Stages 3 and 4 sleep percentage, and REM percentage were included as the dependent variables. The MANOVA was significant, Roy's $\theta = .76$, $F(8, 65) = 6.14$, $p < .01$, and indicated that sleep was significantly worse on Night 1 than on Night 2. At follow-up, the MANOVA was also significant, Roy's $\theta = .49$, $F(8, 66) = 4.00$, $p < .01$, again indicating that sleep improved significantly from Night 1 to Night 2. Therefore, we treated the first PSG night at baseline and follow-up as an adaptation night and limited PSG analyses to Night 2.

Sleep

We followed these procedures during the statistical analyses of sleep. One of the main hypotheses of the present study was that treatments would vary in effectiveness dependent on level of baseline daytime functioning, as measured by FSS, IIS, ESS, or BASS. Therefore, we conducted four MANOVAs on the home sleep diaries, using a median split on these four measures to create a daytime functioning factor. More specifically, these four tests were 3 (treatments: REL, SC, PL) × 2 (daytime functioning types: high, low) × 3 (times, as repeated measure: baseline, posttreatment, follow-up) MANOVAs that included sleep latency, number of awakenings, WASO, TST, SE, and sleep quality as dependent variables. Each of these four MANOVAs resulted in a significant Treatment × Time interaction ($p < .05$), but only the MANOVA that used FSS produced a significant three-way interaction, Roy's $\theta = .47$, $F(12, 56) = 2.19$, $p < .05$. Our reported results for the home sleep diaries focus on follow-up analyses for this three-way interaction.

We also used the FSS as our daytime functioning measure for the PSG analyses. The initial test for PSG sleep was a 3 (treatments) × 2 (FSS types) × 2 (Time: baseline, follow-up) MANOVA that included sleep latency, number of awakenings, WASO, TST, SE, Stages 1 and 2 sleep percentage, Stages 3 and 4 sleep percentage, and REM sleep percentage as dependent variables.

In univariate follow-up testing, we used a multivariate test of significance for the repeated measures factor and interactions involving it to defend against inflated Type I error rates. We used the Tukey test and Bonferroni-corrected paired t tests to follow up on significant main and simple effects for the independent and repeated measures factors, respectively.

Home sleep diaries. Sleep diary results are summarized in Table 2. For sleep latency, significant results occurred for the Time effect, Roy's $\theta = .60$, $F(2, 65) = 19.57$, $p < .01$, and Treatment × Time interaction, Roy's $\theta = .12$, $F(2, 66) = 3.98$, $p < .05$. Follow-up testing for the significant interaction revealed that for the REL and PL groups, posttreatment sleep latency was significantly shorter than baseline and follow-up sleep latency. In the SC group, posttreatment sleep latency was significantly lower than baseline sleep latency, and no other significant differences were observed across time. There were no significant between-groups effects.

For number of awakenings, the main effect for Time, Roy's $\theta = .14$, $F(2, 65) = 4.56$, $p < .05$, and the Treatment × Time interaction, Roy's $\theta = .23$, $F(2, 66) = 7.51$, $p < .01$, were significant. Analysis of simple effects showed that each treatment group exhibited a different pattern of results across time. The REL group had fewer awakenings at posttreatment than at baseline, but no other comparisons were significant. The SC group had fewer awakenings at posttreatment and follow-up relative to baseline. Number of awakenings in the PL group did not change significantly across time. There was one significant between-groups effect: SC participants had significantly fewer awakenings than the PL group at follow-up.

For WASO, a significant main effect for Time was observed, Roy's $\theta = .47$, $F(2, 65) = 15.35$, $p < .01$. WASO was significantly lower at posttreatment and follow-up than at baseline.

For TST, the main effects for Time, Roy's $\theta = .64$, $F(2, 65) = 20.67$, $p < .01$, and Treatment, $F(2, 66) = 3.23$, $p < .05$, were significant as were the Treatment × Time, Roy's $\theta = .36$, $F(2, 66) = 11.90$, $p < .01$, and Treatment × FSS Type × Time interactions, Roy's $\theta = .17$, $F(2, 66) = 5.65$, $p < .01$. The three-way interaction is portrayed in Figure 1. Analysis of simple effects for the three-way interaction indicated that high-fatigue REL participants had greater TST at posttreatment ($M = 377.79$, $SD = 88.88$) and follow-up ($M = 404.03$, $SD = 86.46$) than at baseline ($M = 330.77$, $SD = 70.65$). High-fatigue PL participants had more TST at follow-up ($M = 376.36$, $SD = 63.41$) than at baseline ($M = 318.99$, $SD = 54.93$), but posttreatment TST ($M = 361.59$, $SD = 56.17$) did not differ significantly from baseline or follow-up. TST did not change significantly across time in high-fatigue SC participants (baseline $M = 321.42$, $SD = 66.99$; posttreatment $M = 318.51$, $SD = 80.63$; follow-up $M = 350.46$, $SD = 77.44$).

Table 2
Sleep Diary Means and Standard Deviations Across Time

Measure	Baseline		Posttreatment			Follow-up		
	M	*SD*	*M*	*SD*	Effect size	*M*	*SD*	Effect size
Sleep latency (min)								
Relaxation	32.43	19.67	21.58	14.66	.18	27.40	18.50	.40
Sleep compression	32.82	29.80	21.30	16.44	.18	22.58	16.53	.63
Placebo	34.83	21.39	24.15	14.57		36.61	27.04	
No. of awakenings								
Relaxation	2.31	1.17	1.91	1.01	.28	2.00	0.82	.41
Sleep compression	2.39	1.14	1.70	1.08	.44	1.71	0.65	.67
Placebo	2.05	0.77	2.25	1.41		2.50	1.57	
WASO (min)								
Relaxation	66.47	37.05	42.57	26.39	.26	52.00	46.41	.16
Sleep compression	66.60	32.91	42.37	31.96	.24	38.25	27.77	.70
Placebo	72.00	36.07	49.70	28.15		58.19	29.40	
TST (min)								
Relaxation	345.01	78.44	397.79	87.15	.28	404.39	87.93	.43
Sleep compression	327.98	57.57	314.00	81.96	−.89	364.42	69.40	−.14
Placebo	332.25	71.09	376.80	54.92		372.90	53.01	
SE								
Relaxation	71.85	13.82	80.67	12.76	.16	79.62	13.28	.29
Sleep compression	71.10	12.41	78.61	14.83	−.02	81.47	11.83	.48
Placebo	69.22	12.43	78.86	8.76		76.14	10.48	
Sleep quality rating								
Relaxation	2.88	0.55	3.50	0.60	.36	3.38	0.50	.38
Sleep compression	2.80	0.61	3.38	0.57	.16	3.47	0.52	.54
Placebo	2.94	0.47	3.29	0.56		3.18	0.55	
Napping (min/day)								
Relaxation	12.12	13.99	7.12	10.98	−.11	12.49	21.77	−.13
Sleep compression	12.65	13.78	6.54	9.86	−.06	7.49	10.07	.25
Placebo	9.12	13.49	5.93	9.55		10.21	11.38	

Note. The effect sizes compare each treatment group with the placebo group. WASO = wake time after sleep onset; TST = total sleep time; SE = sleep efficiency.

In the low-fatigue group, REL participants reported greater TST at posttreatment (M = 421.13, SD = 82.60) than at baseline (M = 361.63, SD = 86.76) but no other changes across time. The low-fatigue SC group showed greater TST at follow-up (M = 380.92, SD = 57.69) than at posttreatment (M = 308.67, SD = 87.12) or baseline (M = 335.73, SD = 45.98). SC did dip at posttreatment but rebounded nicely at follow-up. The low-fatigue PL group did not report significant changes in TST across time. One between-groups test was significant. In the low-fatigue group at posttreatment, the REL and PL groups (M = 387.32, SD = 53.69) reported greater TST than the SC group did.

For SE, the Time effect, Roy's θ = .76, $F(2, 65) = 24.82$, $p < .01$. Treatment × Time interaction, Roy's θ = .11, $F(2, 66) = 3.58$, $p < .05$, and Treatment × FSS Type × Time interaction, Roy's θ = .17, $F(2, 66) = 5.60$, $p < .01$, were significant. The three-way interaction for SE is shown in Figure 2. Follow-up tests for the three-way interaction indicated that high-fatigue REL and SC participants reported significantly better SE at posttreatment (REL group: M = 79.66, SD = 15.55; SC group: M = 78.75, SD = 10.87) and follow-up (REL group: M = 81.45, SD = 11.34; SC group: M = 78.05, SD = 12.79) than at baseline (REL group: M = 71.11, SD = 12.98; SC group: M = 68.42, SD = 14.71). High-fatigue PL participants reported greater SE at follow-up (M = 76.10, SD = 8.80) than at baseline (M = 68.23, SD = 9.77) but reported no other significant changes across time. In the low fatigue group, the SC participants reported significantly improved SE at follow-up (M = 85.51, SD = 9.62) relative to baseline (M = 74.27, SD = 8.60). Low fatigue PL participants showed greater SE at posttreatment (M = 80.65, SD = 7.51) than at baseline (M = 69.90, SD = 14.34), but SE at follow-up (M = 76.16, SD = 11.86) was not significantly different from baseline. No significant change across time was observed for the low fatigue REL group (baseline M = 72.70, SD = 15.27; posttreatment M = 81.84, SD = 9.02; follow-up M = 77.48, SD = 15.49). No significant between-groups differences were observed.

For the quality of sleep rating, there was a significant main effect for Time, Roy's θ = .81, $F(2, 65) = 26.39$, $p < .01$. Sleep quality was significantly better at posttreatment and follow-up than at baseline.

Napping. A 3 (treatment) × 2 (FSS type) × 3 time ANOVA was conducted on the napping variable. We found a significant Time effect only, Roy's θ = .16, $F(2, 64) = 5.16$, $p < .01$. Napping decreased significantly from baseline to posttreatment, but no other significant changes across time were observed. Napping results are displayed in Table 2.

PSG sleep. We compared the agreement of the two scorers for the eight PSG variables on the PSG records that were double scored. Intraclass correlations ranged from .89 to .98 and all were significant (all $p < .01$).

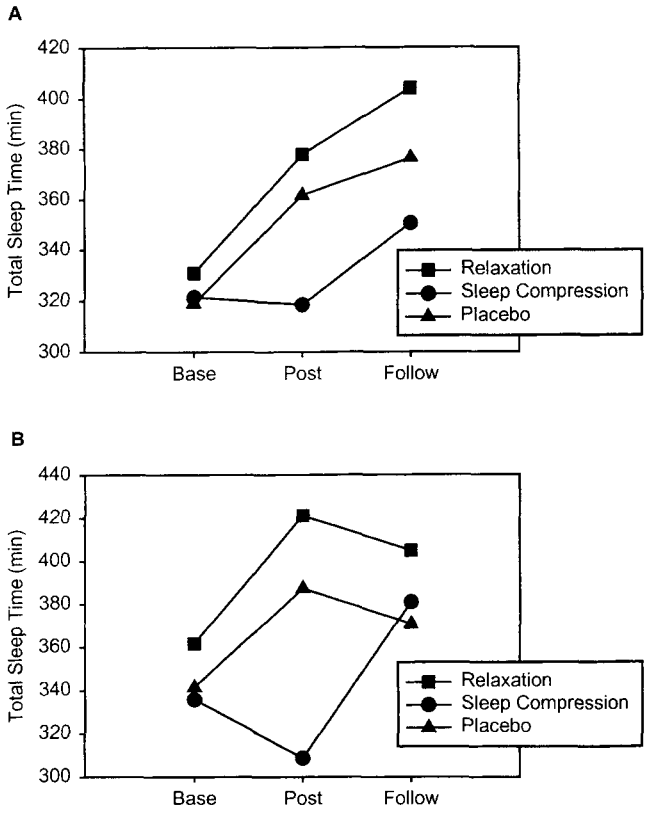

Figure 1. Treatment effects on total sleep time among (A) high-fatigue and (B) low-fatigue individuals. Base = baseline; Post = posttreatment; Follow = follow-up.

PSG results are summarized in Table 3. The initial MANOVA conducted for the PSG variables resulted in a significant Treatment effect only, Roy's $\theta = .35$, $F(8, 62) = 2.69$, $p < .05$. This significant effect may reflect a failure of randomization on baseline PSG variables and is of no theoretical or clinical interest. Several PSG sleep variables were at subclinical levels at baseline (latency, TST, Stages 1 and 2 percentage, Stages 3 and 4 percentage, REM percentage), making significant improvement unlikely (i.e., a ceiling or basement effect). We repeated the MANOVA, including only those variables that were clinically deviant at baseline: number of awakenings, WASO, and SE. This MANOVA also produced only a significant Treatment effect, Roy's $\theta = .16$, $F(3, 67) = 3.61$, $p < .05$.

Sleep diaries completed on PSG nights. Because home sleep diary and laboratory PSG measures produced discrepant results, diaries completed at the laboratory were also analyzed. One interpretation of the sleep diary–PSG discrepancy is that participants perceived their sleep as improved, but objectively it was not. However, degree of objectivity was not the only difference between the home sleep diaries and PSG. For example, the setting differed between the two measures, and one measure covered 2 weeks, whereas the other measure was based on one night.

A total of 62 participants completed a sleep diary the morning following each PSG night. If the PSG night diaries showed improvement across time, this would suggest that participants perceived improvement, although there was no objective evidence of progress. However, if PSG night diaries did not indicate improvement across time, this would suggest the possibility that PSG nights were unable to detect the treatment effect measured by the home sleep diaries. For example, home findings might not generalize to the laboratory, or the laboratory assessment period may have been too brief.

A 3 (treatment group) × 2 (FSS type) × 2 (time: baseline, follow-up) MANOVA was conducted on the PSG night sleep diaries. There were no significant main or interaction effects. It is possible that this subsample of 62 participants differed from the sample as a whole. Therefore the 3 × 2 × 2 MANOVA reported for PSG variables was repeated including only these 62 participants, and no significant main or interaction effects were observed. Similarly, a 3 × 2 × 2 MANOVA was used to analyze the home sleep diaries of these 62 participants. This MANOVA resulted in a significant Treatment × Time interaction, Roy's $\theta = .39$, $F(6, 52) = 3.39$, $p < .01$, which suggested sleep improvement across time had differential effects that depended on treatment condition. Therefore, PSG and home sleep diary results for this subsample were similar to those for the total sample, except that the three-way interaction for home diaries was no longer observed. In summary, PSG measures and sleep diaries from the PSG nights suggested no sleep improvement at follow-up, whereas home sleep diaries suggested significant sleep enhancement across time.

Clinical significance. Statistically significant improvement may not be clinically meaningful. Therefore, we also attempted to

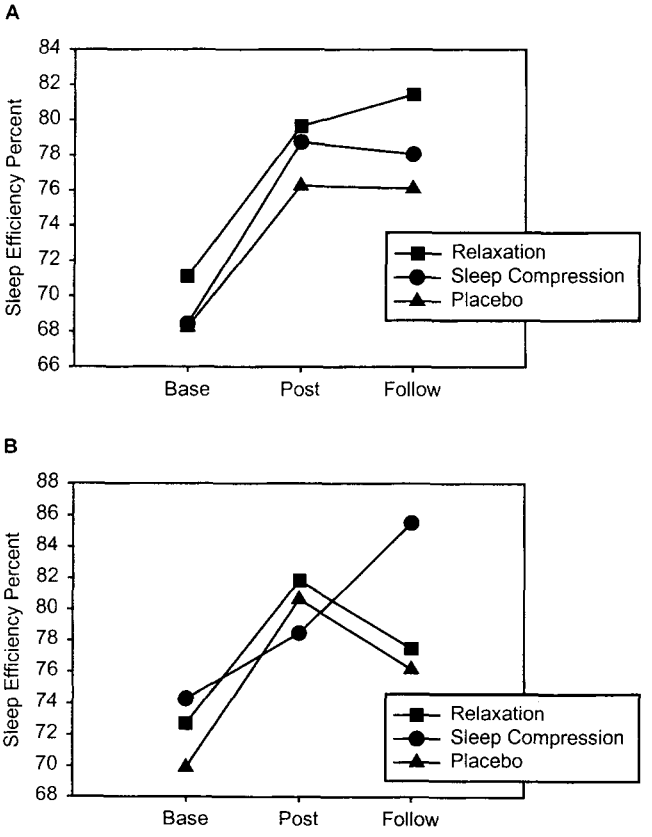

Figure 2. Treatment effects on sleep efficiency percentage among (A) high-fatigue and (B) low-fatigue individuals. Base = baseline; Post = posttreatment; Follow = follow-up.

Table 3
PSG Sleep Means and Standard Deviations Across Time

Measure	Baseline		Follow-up		
	M	*SD*	*M*	*SD*	Effect size
Sleep latency (min)					
Relaxation	25.54	46.71	15.24	16.40	−.32
Sleep compression	11.17	15.11	11.29	15.84	−.03
Placebo	13.30	17.56	10.89	8.92	
No. of awakenings					
Relaxation	20.85	6.59	22.56	11.89	.00
Sleep compression	23.75	11.69	23.29	10.14	−.06
Placebo	26.30	10.24	22.57	13.33	
WASO (min)					
Relaxation	67.56	35.50	74.19	51.71	.01
Sleep compression	58.17	35.27	54.10	41.11	.45
Placebo	68.59	47.70	74.78	49.87	
TST (min)					
Relaxation	364.07	97.70	349.33	83.66	−.09
Sleep compression	383.19	66.28	396.17	77.05	.57
Placebo	377.26	49.70	356.02	62.81	
SE					
Relaxation	77.13	14.88	76.97	12.29	−.14
Sleep compression	83.85	6.11	85.56	7.83	.69
Placebo	81.20	9.35	78.62	11.88	
Stages 1 & 2 (%)					
Relaxation	55.39	11.47	53.50	11.79	.02
Sleep compression	60.35	7.66	59.47	9.73	−.52
Placebo	55.99	11.89	53.76	12.01	
Stages 3 & 4 (%)					
Relaxation	12.96	10.37	14.20	10.16	.28
Sleep compression	8.18	7.43	9.67	7.85	−.21
Placebo	11.85	10.00	11.47	9.30	
REM (%)					
Relaxation	15.81	5.49	15.24	6.30	−.37
Sleep compression	18.73	3.58	19.26	4.62	.38
Placebo	17.34	5.26	17.40	5.29	

Note. The effect sizes compare each treatment group with the placebo group. PSG = polysomnography; WASO = wake time after sleep onset; TST = total sleep time; SE = sleep efficiency.

judge the clinical significance of our treatment outcomes. We chose to focus on sleep latency and WASO because these are the main measures used to assess degree of initiation and maintenance difficulty. A sleep latency or WASO of greater than 30 min is often used as the minimum criterion for diagnosing clinically significant insomnia, and therefore, a clinically significant treatment result would imply a reduction of sleep latency or WASO to a mean of 30 min or less after treatment. In addition to the 30 min criterion, participants were required to reduce mean latency or WASO by at least 15 min from baseline in order for the outcome to be defined as clinically significant. Also, baseline mean latency or WASO had to be greater than 30 min for a participant to be included in the following analyses. Each sleep variable (latency, WASO) was analyzed separately, and both diary and PSG data were examined.

For SR sleep latency, 14 REL, 10 SC, and 10 PL participants had mean sleep latencies greater than 30 min at baseline, $\chi^2(2, N = 72) = 0.78, p = .68$. By posttreatment, 7 REL (50%), 6 SC (60%), and 2 PL (20%) participants showed clinically significant results for latency, but these differences did not reach statistical significance, $\chi^2(2, N = 34) = 3.58, p = .17$. At follow-up, statistically significant differences favoring the SC group emerged, with the following number of participants showing clinically significant results in each treatment group: 2 REL (14%), 6 SC (60%), and 0 PL, $\chi^2(2, N = 34) = 11.13, p < .01$.

For SR WASO, 21 REL, 22 SC, and 20 PL participants had a mean greater than 30 min at baseline, $\chi^2(2, N = 72) = 1.69, p = .43$. At posttreatment, 8 REL (38%), 9 SC (41%), and 4 PL (20%) participants showed clinically significant results for wake time, $\chi^2(2, N = 63) = 2.38, p = .30$. By follow-up, the SC group clearly surpassed the others. Clinically significant progress was observed for 5 REL (24%), 12 SC (55%), and 2 PL participants (10%), $\chi^2(2, N = 63) = 10.47, p < .01$.

Effect size (ES), shown in Table 2, provides another view of clinical impact. At posttreatment and follow-up, we calculated the ES between each treatment and the PL group using pooled error terms. By Cohen's (1987) criteria, REL had a small impact on SR sleep at posttreatment and follow-up (mean ES = 0.20 and 0.28, respectively). SC did poorly at posttreatment (mean ES = 0.01) mainly because of a dropoff in TST, but approached a medium ES by follow-up (mean ES = 0.45).

For PSG, there were too few participants at baseline with clinically significant sleep latency to perform this analysis. For PSG WASO, 23 REL, 19 SC, and 17 PL participants showed wake

time greater than 30 min at baseline, $\chi^2(2, N = 74) = 0.98$, $p = .61$. No significant group differences were observed at follow-up, with 6 REL (26%), 5 SC (26%), and 3 PL (18%) participants showing clinically significant outcomes, $\chi^2(2, N = 59) = 0.49$, $p = .78$. PSG ES for REL (ES: $M = 0.08$) and SC (ES: $M = 0.16$) were undistinguished (see Table 3).

Daytime Functioning

Means and standard deviations for daytime functioning variables are given in Table 4. A 3 (treatment condition) × 3 (time) MANOVA was performed on the ESS, FSS, IIS, and BASS. The MANOVA produced a significant Time effect, Roy's $\theta = .38$, $F(8, 61) = 2.93$, $p < .01$, and a significant Treatment × Time interaction, Roy's $\theta = .34$, $F(8, 62) = 2.65$, $p < .05$. During univariate follow-up tests, the only measure that produced a significant main or interaction effect was the BASS, for which a significant Time effect, Roy's $\theta = .27$, $F(2, 67) = 8.92$, $p < .01$, and Treatment × Time interaction, Roy's $\theta = .10$, $F(2, 68) = 3.21$, $p < .05$, were observed. Analysis of simple effects revealed that the BASS changed significantly across time only in the REL group. REL participants had significantly lower BASS scores at follow-up than at baseline, with no other significant comparisons across time. No significant between-groups differences were observed.

The ES data yield a somewhat different picture (see Table 4). On the average, REL consistently produced a small ES in contrast with the PL group; posttreatment $M = 0.27$, and follow-up $M = 0.24$. Daytime improvement in the REL group surpassed that of the SC group, wherein little change occurred; posttreatment $M = 0.08$, and follow-up $M = -0.01$.

Treatment Implementation

Delivery. Audiotapes of therapy sessions were rated for proper delivery and one fourth were scored by a second rater. Percentage agreement between raters was 84%. The mean score for the sessions was 98.82% ($SD = 3.72$), suggesting that treatment had been delivered as intended.

Receipt. Relaxation rating increased significantly (M change = 2.44, $SD = 1.39$), $t(25) = 8.96$, $p < .01$, and pulse rate decreased significantly (M change = − 4.39, $SD = 4.14$), $t(25) = 5.41$, $p < .01$, after treatment within sessions. In the SC group, receipt was measured through the sleep quiz, and 4.33 of 5 ($SD = .70$) questions were answered correctly. In the PL group, participants gave clarity ratings on a 1 (*low*) to 10 (*high*) scale for bedtime ($M = 8.08$, $SD = 1.50$) and neutral ($M = 7.89$, $SD = 1.19$) events in treatment sessions. Therefore, satisfactory receipt occurred in each treatment condition. (See Footnote 1 for information on obtaining a more complete analysis).

Enactment. All participants were given the seven sleep hygiene instructions, and they reported enactment of nearly all of these recommendations at posttreatment ($M = 6.61$, $SD = 0.47$) and follow-up ($M = 6.39$, $SD = 0.84$).

REL participants reported home practice of relaxation 12.72 ($SD = 1.57$) times per week during treatment, 13.02 ($SD = 2.18$) times during posttreatment, and 9.15 ($SD = 5.17$) times at follow-up. They also reported a significant increase in relaxation rating and a significant decrease in pulse rate after home relaxation practice during treatment, posttreatment, and follow-up (all $ps < .05$).

In the SC group, we used a one-way repeated measures ANOVA to examine SR time in bed across study phases, Roy's $\theta = 3.38$, $F(2, 22) = 37.15$, $p < .01$. The participants spent significantly less time in bed at posttreatment ($M = 397.60$, $SD = 67.49$) and follow-up ($M = 445.70$, $SD = 57.76$) than at baseline ($M = 465.02$, $SD = 64.60$). Reported time spent in bed was significantly higher at follow-up than at posttreatment. Time in bed on baseline and follow-up PSG Night 2 was not significantly different, $t(23) = 0.38$, $p = .71$.

Table 4
Daytime Functioning Means and Standard Deviations Across Time

	Baseline		Posttreatment			Follow-up		
Measure	*M*	*SD*	*M*	*SD*	Effect size	*M*	*SD*	Effect size
FSS								
Relaxation	3.72	1.31	3.42	1.27	.38	3.39	1.33	−.14
Sleep compression	3.55	1.21	3.75	1.26	.11	3.48	1.32	−.22
Placebo	3.37	1.26	3.89	1.22		3.22	1.03	
IIS								
Relaxation	99.58	23.09	90.35	20.20	.40	94.54	22.19	.13
Sleep compression	98.08	21.24	96.50	21.83	.14	101.06	27.30	−.17
Placebo	103.50	22.06	99.86	27.44		97.17	18.19	
ESS								
Relaxation	9.12	5.19	8.54	3.97	.13	8.15	3.53	.41
Sleep compression	9.70	5.09	9.22	5.22	−.04	9.17	5.10	.11
Placebo	9.95	4.86	9.05	3.92		9.68	3.99	
BASS								
Relaxation	41.17	9.65	36.49	12.61	.17	33.40	12.09	.57
Sleep compression	40.26	11.37	37.07	12.79	.12	36.65	13.21	.25
Placebo	41.10	11.30	38.36	8.42		39.50	8.77	

Note. The effect sizes compare each treatment group with the placebo group. FSS = Fatigue Severity Scale; IIS = Insomnia Impact Scale; ESS = Epworth Sleepiness Scale; BASS = Beliefs and Attitudes about Sleep Scale.

PL participants reported home practice of imagery 12.51 (SD = 2.23) times per week during treatment, 13.07 (SD = 2.90) times during posttreatment, and 8.24 (SD = 5.18) times at follow-up. Mean bedtime and neutral event clarity ratings for home practice at each study phase were generally high, with ratings ranging from a mean of 6.92 (SD = 1.50) for bedtime clarity at follow-up to a mean of 8.27 (SD = 1.00) for neutral event clarity at posttreatment.

Enactment and treatment outcome. To determine if enactment was correlated with treatment outcome, we calculated partial correlations between the enactment variables listed earlier and posttreatment and follow-up sleep variables, controlling for baseline sleep. We report significant correlations only.

For the sleep hygiene component, greater posttreatment adherence predicted less posttreatment SR WASO (r = −.27), and better follow-up adherence was associated with shorter SR sleep latencies (r = −.40). Adherence during the follow-up period was also related to three follow-up PSG measures: latency (r = −.28), WASO (r = −.41), and SE (r = .38).

In the REL group, a higher number of relaxation practices per week during treatment predicted a lower posttreatment latency (r = −.39) and a higher posttreatment sleep efficiency (r = .44). Also, those participants who spent more time practicing relaxation at follow-up reported lower follow-up sleep latencies (r = −.40).

In the SC group, there was a positive correlation between SR TST and time spent in bed at posttreatment (r = .52) and at follow-up (r = .48). At PSG follow-up, less time in bed was associated with lower sleep latency (r = .55), WASO (r = .50), and TST, (r = .88). Also, a smaller difference between SR posttreatment time in bed and baseline TST predicted lower posttreatment TST (r = .52) but a higher posttreatment sleep quality rating (r = −.45).

For the PL group, no treatment-specific enactment variable was related to outcome. Earlier, we reported the association between the sleep hygiene component and sleep outcome for all groups combined. Because sleep hygiene was an active treatment component inserted in the PL group, we now report this relationship within the PL group only. At posttreatment, sleep hygiene was unrelated to outcome. At follow-up, sleep hygiene adherence was related to SR latency (r = −.56) and three PSG variables: latency (r = −.50), WASO (r = −.67), and SE (r = .49).

Treatment Credibility

Treatment credibility ratings were compared across treatments. Ratings for two of the four credibility questions differed significantly across treatment groups: Question 1 ("reasonableness"), $F(2, 71) = 8.58$, $p < .01$, and Question 4 ("recommend to friend"), $F(2, 71) = 3.27$, $p < .05$. For these two questions, the PL group had lower credibility than the other two treatment conditions, but PL group ratings still fell in the upper end of the 10-point scale (Question 1, M = 6.61, SD = 1.47; Question 4, M = 6.70, SD = 1.92). For the remaining two credibility questions, there was no significant difference between groups: Question 2 ("like my therapist," M = 9.20, SD = 1.05), and Question 3 ("expectation for improvement," M = 6.84, SD = 1.75).

We computed partial correlations between the sum credibility scores and sleep outcome controlling for baseline sleep to estimate the role of expectancy. For all groups combined, credibility was significantly related to SR at posttreatment, number of awakenings (r = −.29), WASO (r = −.24), and quality of sleep rating (r = .31), and at follow-up, number of awakenings (r = −.32) and quality of sleep rating (r = .39). Credibility scores were not related to follow-up PSG data. Considering the PL group only, we observed a stronger immediate credibility effect, but this dissipated by follow-up. At posttreatment, credibility was significantly correlated with number of awakenings (r = −.38), WASO (r = −.53), TST (r = .51), SE (r = .44), and rated quality (r = .49). By follow-up, credibility in the PL group was unrelated to any self-report or PSG sleep variable.

We conducted a telephone interview after study completion with PL participants reporting sleep improvement. Two participants could not be reached, and 3 participants denied their sleep had improved. Implementation of sleep hygiene procedures was the most common reason given for improvement (n = 5). Three participants indicated that they were highly motivated to improve their sleep and that this was the main reason for their success. Other reported reasons for improvement were a relaxation effect from the imagery procedure (n = 2) and a significant decrease in life stress (n = 1). Each participant indicated that the $200 payment had no influence.

Discussion

In this clinical trial, we tested psychological treatments for insomnia in older adults against a placebo treatment and explored the interaction between individuals' level of daytime impairment and type of intervention. Our main findings are that psychological treatments for insomnia in older adults are effective, but this conclusion does not stand without qualification. These results were obtained in the sleep diaries but not in the PSG data. Further, we found evidence supporting the conclusion that the psychological treatments outperformed the PL treatment, but this finding too is not without ambiguity. Last, insomnia subtypes, defined by level of daytime fatigue, did to some extent differentially respond to treatments geared to extending or compressing sleep.

The present study is the most methodologically sophisticated clinical trial to come from our laboratory. The main features were extensive screening to establish that we were testing primary insomnia; drug screens to guard against sleep-active substances; a treatment implementation scheme that assured the intended treatment was being tested, the treatment was mastered by the participant, and the participant used the treatment at home; and a PL control. We therefore conclude that our results can be viewed with a high degree of confidence.

The main sleep diary findings are as follows, ignoring the PL group for the moment. Both REL and SC produced significant improvement in sleep latency at posttreatment, but these gains did not hold up at follow-up. A similar pattern obtained for number of awakenings during the night and WASO. For these variables, SC better maintained its gains at follow-up. Both groups produced significant increments in TST, SE, and sleep quality ratings. REL outperformed SC in TST, SC held a small edge in sleep quality ratings at follow-up, and the groups were comparable in SE. In sum, REL and SC produced significant improvement over time in every sleep variable. At 1-year follow-up, REL did not fare as well as SC. The performance of SC in the present study was comparable to the similar method of sleep restriction in the few studies with

older adults when it was a unitary intervention (Bliwise, Friedman, Nekich, & Yesavage, 1995; Brooks, Friedman, Bliwise, & Yesavage, 1993).

Although PL participants rated their treatment as less credible than the other groups on half the credibility measures, their sleep improvement was good. Sleep data for this group did not significantly differ from the treatment groups for latency to sleep, WASO, TST (in the high-fatigue group), SE, and rated sleep quality, and were weaker than the treatment groups for number of awakenings during the night and TST (in the low-fatigue group). In almost all cases, PL group posttreatment and follow-up scores were weaker than those for the treatment groups, but these differences did not usually reach statistical significance.

When symptomatic individuals seeking treatment are returned to nonsymptomatic levels following treatment, this reflects clinical significance, in contrast to statistical significance. Clinical significance speaks to the magnitude or meaningfulness of change, is considered a primary criterion for judging the merits of interventions, and can be evaluated by a number of approaches (Jacobson, Roberts, Berns, & McGlinchey, 1999; Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999).

Tests of clinical significance more clearly separated the PL group. By follow-up, there was little clinically significant change in the PL group for SR sleep and its separation from the treatment groups was statistically significant. The standard of clinical significance clearly showed SC was the strongest treatment. These results mirror those of Friedman, Bliwise, Yesavage, and Salom (1991), who compared progressive relaxation and sleep restriction. ES data partially corroborated these findings but differed to some degree because the clinical significance analyses were based on two measures and the ES analyses combined all measures. REL consistently produced small average ES, whereas SC showed no net improvement at posttreatment climbing to an average moderate ES at follow-up.

This PL has been shown to have credibility equivalent to those of the active treatments in insomnia studies with middle-aged participants (e.g., Lacks, Bertelson, Gans, & Kunkel, 1983; Steinmark & Borkovec, 1974). We piloted variations of the rationale for the PL group, but ratings still slumped on half the scales. This PL may be less suited to older adults. Our analyses suggest that credibility was related to short-term but not to long-term outcome.

Sleep hygiene instructions cannot be construed as a placebo, and therefore, our PL group was corrupted with an active therapeutic component. In retrospect, this was clearly an error. These sleep hygiene instructions were particularly comprehensive and robust, and this likely increased their impact. For example, they included arising at a fixed time each morning, which is an ingredient of both stimulus control and SC treatments. In the present study, sleep hygiene was perceived as influential in debriefing, and sleep hygiene enactment was correlated with several SR and PSG outcome variables for all groups considered jointly, and for the PL group analyzed separately. Although sleep hygiene has generally been shown to be a weak to moderately strong treatment (Morin, Culbert, & Schwartz, 1994), there is little prior data on this approach with OAWI.

SC differs from sleep restriction treatment of insomnia (Spielman, Saskin, & Thorpy, 1987) and is modeled after our prior trials (Lichstein, 1988; Riedel, Lichstein, & Dwyer, 1995). Spielman et al. (1987) eliminated the entire excess time in bed immediately and fine-tuned the allotted time in bed in subsequent weeks depending on the participant's satisfying sleep efficiency ratio criteria. This approach can produce a sudden change in lifestyle for elderly adults to which they have difficulty adjusting and may induce some measure of sleep deprivation resulting in increased daytime sleepiness (Glovinsky & Spielman, 1991). Our approach of incrementally reducing time in bed seemed to be well tolerated by our participants.

We had two reasons for being optimistic about the present REL treatment. First, Friedman et al. (1991) used four treatment sessions, and in the present study we used six. Relaxation effectiveness is correlated with the number of sessions (Carlson & Hoyle, 1993). Our future research will use eight sessions. Second, we anticipated that our method of REL was more suited to older adults than progressive relaxation (PR). PR may be less suitable for older adults because it is procedurally complex and physically demanding. The present method avoids these potential pitfalls and also possesses the advantage of comprising discreet components. The participant is given a menu of techniques and can emphasize those parts found most appealing. However, REL achieved only moderate success.

In sharp contrast to the SR results, the PSG data showed no significant change for any variable. However, negative PSG results do not automatically discount positive SR findings. Clearly, PSG change is more difficult to obtain than SR change and, when it does occur, PSG change is usually of smaller magnitude (e.g., Engle-Friedman et al., 1992; Morin et al., 1993). One factor contributing to weaker PSG results is a function of the number of data points sampled for analysis. Fourteen data points were gathered to yield each SR variable, but we should expect much more noise in the PSG data because variables were derived from only one night of recording, and sleep in OAWI has been shown to be highly unstable when sampling fewer than about 7 days of data (Wohlgemuth, Edinger, Fins, & Sullivan, 1999). Further, our omission of posttreatment PSG, for cost-saving purposes, precludes our concluding that short-term PSG improvement failed to occur.

Alternatively, the PSG results should not necessarily be dismissed. Analyses comparing SR data collected on PSG nights with PSG data were in agreement that sleep change did not occur during laboratory assessments. This reciprocal confirmation bolsters the validity of both the PSG data and the SR data taken on PSG nights and, by extrapolation, enhances the confidence accorded the home SR data as well.

We found a significant interaction with reported daytime fatigue in two important sleep measures, SE and TST. There is accumulating evidence that fatigue is a common characteristic of insomnia. A survey of patients at a sleep disorders center found that individuals diagnosed with primary insomnia reported clinically significant levels of fatigue that were significantly higher than even those patients with sleep apnea (Lichstein, Means, Noe, & Aguillard, 1997). Among older adults complaining of high-distress insomnia, reported fatigue was four times that of noncomplaining sleepers and double that of low-distress OAWI (Fichten et al., 1995).

Factoring in fatigue with SE, we reasoned that SC would be most helpful with low fatigue individuals because they exhibited minimal daytime impairment. This is consistent with our theory of insomnoid states discussed in the introduction to this article. Pre-

sumably, these individuals satisfied their biological sleep need, albeit with a fragmented, unsatisfying sleep pattern. Consistent with this hypothesis, low-fatigue individuals receiving SC attained an SE of 85.5 at follow-up compared with an SE of 78.1 among high-fatigue participants. Further, among low-fatigue individuals, the SE of 85.5 was the highest score at follow-up of the three groups. Similarly, we reasoned that individuals reporting high fatigue were needing more sleep and that REL was better suited to extending sleep and would be more effective with this subgroup. As hypothesized, at follow-up in the high-fatigue group, participants treated with REL attained the highest SE of any group (81.5), whereas low-fatigue participants receiving REL showed virtually no change from baseline to follow-up.

Considering fatigue and TST, we similarly reasoned that REL would work best for high fatigue individuals and SC would work best for low-fatigue individuals, and these expectations were partially confirmed. SC worked well with low-fatigue individuals and was ineffective with high-fatigue participants. REL performed well with both low- and high-fatigue participants, but results extending out to follow-up were strongest when high fatigue was present.

The results relating level of fatigue to treatment outcome, and their implications for validating our biodevelopmental model of insomnia, should be viewed with caution for two reasons. First, in testing four measures of daytime dysfunction, only one proved significant. This then should be considered an exploratory analysis, and we cannot rule out the explanation that inflated Type I error rates subverted the integrity of the results. Second, the aforementioned summary of findings focused on treatment effects that were confirmatory of our hypotheses relating to the presence or absence of insomnoid states. For both SE and TST, the results included outcomes contrary to predictions from this model. For SE, the SC group showed significant improvement over time among high-fatigue individuals. For TST, the REL group showed significant improvement among low-fatigue individuals. Also, the PL group did register gains in three of four fatigue levels for SE and TST, further muddying the interpretation. Accordingly, doubt will infuse the reliability and generalizability of the interaction findings until it is dispelled by replication.

For the most part, sleep improvement was not accompanied by comparable gains in daytime functioning, although ES data show stronger change in the REL than the SC group. This result is disappointing, but has become a reliable finding in our experience. This disengagement between sleep change and daytime functioning occurred in recent studies of young adults with insomnia (Means, Lichstein, Epperson, & Johnson, 2000) and of older adults with secondary insomnia (Lichstein, Wilson, & Johnson, 2000). The present study had a much longer follow-up period than the two studies just cited, challenging our prior hypothesis that daytime change would come but would lag behind nighttime change.

SC proved effective, and it along with its close kin, sleep restriction, is emerging as one of the strongest treatments for OAWI. REL performed less well, and ambiguity persists in identifying a method of relaxation that is reliably effective for OAWI. Lastly, we would like to encourage further exploration of the role of fatigue and other dimensions of daytime impairment in isolating insomnia subtypes and articulating a model of matching types of insomnia with types of treatment.

References

American Sleep Disorders Association. (1990). *International classification of sleep disorders: Diagnostic and coding manual.* Rochester, MN: Author.

Bliwise, D. L., Friedman, L., Nekich, J. C., & Yesavage, J. A. (1995). Prediction of outcome in behaviorally based insomnia treatments. *Journal of Behavior Therapy and Experimental Psychiatry, 26,* 17–23.

Brodman, K., Erdmann, A. J., Jr., Wolff, H. G., & Miskovitz, P. F. (1986). *Manual: Cornell medical index.* New York: Cornell University Medical College.

Brooks, J. O., III, Friedman, L., Bliwise, D. L., & Yesavage, J. A. (1993). Use of the wrist actigraph to study insomnia in older adults. *Sleep, 16,* 151–155.

Carlson, C. R., & Hoyle, R. H. (1993). Efficacy of abbreviated progressive muscle relaxation training: A quantitative review of behavioral medicine research. *Journal of Consulting and Clinical Psychology, 61,* 1059–1067.

Cohen, J. (1987). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Engle-Friedman, M., Bootzin, R. R., Hazlewood, L., & Tsao, C. (1992). An evaluation of behavioral treatments for insomnia in the older adult. *Journal of Clinical Psychology, 48,* 77–90.

Fichten, C. S., Creti, L., Amsel, R., Brender, W., Weinstein, N., & Libman, E. (1995). Poor sleepers who do not complain of insomnia: Myths and realities about psychological and lifestyle characteristics of older good and poor sleepers. *Journal of Behavioral Medicine, 18,* 189–223.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Friedman, L., Bliwise, D. L., Yesavage, J. A., & Salom, S. R. (1991). A preliminary study comparing sleep restriction and relaxation treatments for insomnia in older adults. *Journal of Gerontology: Psychological Sciences, 46,* P1–P8.

Glovinsky, P. B., & Spielman, A. J. (1991). Sleep restriction therapy. In P. J. Hauri (Ed.), *Case studies in insomnia* (pp. 49–63). New York: Plenum.

Hoelscher, T. J., Ware, J. C., & Bond, T. (1993). Initial validation of the insomnia impact scale [Abstract]. *Sleep Research, 22,* 149.

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology, 67,* 300–307.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth Sleepiness Scale. *Sleep, 14,* 540–545.

Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology, 67,* 285–299.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The Fatigue Severity Scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology, 46,* 1121–1123.

Lacks, P., Bertelson, A. D., Gans, L., & Kunkel, J. (1983). The effectiveness of three behavioral treatments for different degrees of sleep onset insomnia. *Behavior Therapy, 14,* 593–605.

Lichstein, K. L. (1988). Sleep compression treatment of an insomnoid. *Behavior Therapy, 19,* 625–632.

Lichstein, K. L. (2000). Relaxation. In K. L. Lichstein & C. M. Morin (Eds.), *Treatment of late-life insomnia* (pp. 185–206). Thousand Oaks, CA: Sage.

Lichstein, K. L., & Fischer, S. M. (1985). Insomnia. In M. Hersen & A. S.

Bellack (Eds.), *Handbook of clinical behavior therapy with adults* (pp. 319–352). New York: Plenum Press.

Lichstein, K. L., Means, M. K., Noe, S. L., & Aguillard, R. N. (1997). Fatigue and sleep disorders. *Behaviour Research and Therapy, 35,* 733–740.

Lichstein, K. L., Riedel, B. W., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research and Therapy, 16,* 1–29.

Lichstein, K. L., Riedel, B. W., Lester, K. W., & Aguillard, R. N. (1999). Occult sleep apnea in a recruited sample of older adults with insomnia. *Journal of Consulting and Clinical Psychology, 67,* 405–410.

Lichstein, K. L., Riedel, B. W., & Means, M. K. (1999). Psychological treatment of late-life insomnia. In R. Schulz, G. Maddox, & M. P. Lawton (Eds.), *Annual review of gerontology and geriatrics: Vol. 18. Focus on interventions research with older adults* (pp. 74–110). New York: Springer.

Lichstein, K. L., Wilson, N. M., & Johnson, C. T. (2000). Psychological treatment of secondary insomnia. *Psychology and Aging, 15,* 232–240.

Means, M. K., Lichstein, K. L., Epperson, M. T., & Johnson, C. T. (2000). Relaxation therapy for insomnia: Nighttime and day time effects. *Behaviour Research and Therapy, 38,* 665–678.

Mellinger, G. D., & Balter, M. B. (1981). Prevalence and patterns of use of psychotherapeutic drugs: Results from a 1979 national survey of American adults. In G. Tognoni, C. Bellantuono, & M. Lader (Eds.), *Epidemiological impact of psychotropic drugs* (pp. 117–135). Amsterdam: Elsevier.

Mellinger, G. D., Balter, M. B., & Uhlenhuth, E. H. (1985). Insomnia and its treatment. *Archives of General Psychiatry, 42,* 225–232.

Morgan, K. (1987). *Sleep and aging: A research-based guide to sleep in later life.* Baltimore: Johns Hopkins University Press.

Morin, C. M. (1993). *Insomnia: Psychological assessment and management.* New York: Guilford.

Morin, C. M., Colecchi, C., Stone, J., Sood, R., & Brink, D. (1999). Behavioral and pharmacological therapies for late-life insomnia: A randomized controlled trial. *Journal of the American Medical Association, 281,* 991–999.

Morin, C. M., Culbert, J. P., & Schwartz, S. M. (1994). Nonpharmacological interventions for insomnia: A meta-analysis of treatment efficacy. *American Journal of Psychiatry, 151,* 1172–1180.

Morin, C. M., Kowatch, R. A., Barry, T., & Walton, E. (1993). Cognitive–behavior therapy for late-life insomnia. *Journal of Consulting and Clinical Psychology, 61,* 137–146.

Murden, R. A., McRae, T. D., Kaner, S., & Bucknam, M. E. (1991). Mini-Mental State exam scores vary with education in Blacks and Whites. *Journal of the American Geriatrics Society, 39,* 149–155.

Rechtschaffen, A., & Kales, A. (Eds.). (1968). *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects.* (Public Health Service, NIH Publication No. 204). Washington, DC: U.S. Government Printing Office.

Riedel, B. W., & Lichstein, K. L. (2000). Insomnia and daytime functioning. *Sleep Medicine Reviews, 4,* 277–298.

Riedel, B. W., Lichstein, K. L., & Dwyer, W. O. (1995). Sleep compression and sleep education for older insomniacs: Self-help versus therapist guidance. *Psychology and Aging, 10,* 54–63.

Roth, T., Zorick, F., Wittig, R., & Roehrs, T. (1982). Pharmacological and medical considerations in hypnotic use. *Sleep, 5*(Suppl. 1), 46–52.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *State-Trait Anxiety Inventory (Form Y).* Palo Alto, CA: Consulting Psychologists Press.

Spielman, A. J., Saskin, P., & Thorpy, M. J. (1987). Treatment of chronic insomnia by restriction of time in bed. *Sleep, 10,* 45–56.

Steinmark, S. W., & Borkovec, T. D. (1974). Active and placebo treatment effects on moderate insomnia under counterdemand and positive demand instructions. *Journal of Abnormal Psychology, 83,* 157–163.

Williams, R. L., Karacan, I., & Hursch, C. J. (1974). *Electroencephalography (EEG) of human sleep: Clinical applications.* New York: Wiley.

Wohlgemuth, W. K., Edinger, J. D., Fins, A. I., & Sullivan, R. J., Jr. (1999). How many nights are enough? The short-term stability of sleep parameters in elderly insomniacs and normal sleepers. *Psychophysiology, 36,* 233–244.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a Geriatric Depression Screening Scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37–49.

Received August 30, 1999
Revision received August 7, 2000
Accepted August 17, 2000 ■