NPS-OR-10-011

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

EFFECTS OF SLEEP ON TRAINING EFFECTIVENESS IN SOLDIERS AT FORT LEONARD WOOD, MISSOURI

by

Nita Lewis Miller, Ph.D.
Lawrence G. Shattuck, Ph.D.
Anthony P. Tvaryanas
Panagiotis Matsangas

November 2010

Approved for public release; distribution is unlimited

Prepared for: Leonard Wood Institute
Fort Leonard Wood, MO 65473

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Daniel T. Oliver
President

Leonard A. Ferrari
Executive Vice President and
Provost

This report was prepared for and funded by the Leonard Wood Institute, 197 Replacement Avenue, Fort Leonard Wood, Missouri 65473.

This report was prepared by:

NITA LEWIS MILLER
Associate Professor of
Operations Research

LAWRENCE G. SHATTUCK
Senior Lecturer of Operations Research

ANTHONY P. TVARYANAS
NPS Ph.D. Candidate

PANAGIOTIS MATSANGAS
NPS Ph.D. Candidate

Reviewed by:

RONALD D. FRICKER
Associate Chairman for Research
Department of Operations Research

ROBERT F. DELL
Chairman
Department of Operations Research

Released by:

KARL VAN BIBBER
Vice President and
Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 11-2010	**2. REPORT TYPE** Technical Report	**3. DATES COVERED** *(From - To)* June 2009-April 2010

4. TITLE AND SUBTITLE
Effects of Sleep on Training Effectiveness in Soldiers at Fort Leonard Wood, Missouri

5a. CONTRACT NUMBER W911NF-07-2-0062
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Nita Lewis Miller, Ph.D.
Lawrence G. Shattuck, Ph.D.
Anthony P. Tvaryanas
Panagiotis Matsangas

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER
NPS-OR-10-011

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Leonard Wood Institute
197 Replacement Avenue
Fort Leonard Wood, MO 65473

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT
This study examined the effect of alterations in the timing of sleep within the circadian cycle on the amount of total nightly sleep and its influence on various indicators of mood and performance of U.S. Army Soldiers attending Basic Combat Training (BCT) at Fort Leonard Wood, Missouri. The quasi-experimental study design compared Soldiers assigned to one of two training companies: a company using the standard BCT sleep regimen (8:30 p.m. to 4:30 a.m.) or a company using a phase-delayed sleep regimen (11:00 p.m. to 7:00 a.m.), the latter being more in line with the biologically driven sleep-wake patterns of adolescents. Demographic and psychophysiological measures were collected at the start of the study using standard survey instruments and methods. A random sample of approximately 24% of Soldiers wore wrist activity monitors to unobtrusively record sleep quantity and quality. Weekly assessments were made of subjective fatigue and mood throughout BCT. Data on physical fitness, marksmanship, and attrition from BCT were extracted from organizational training records.

The study sample was comprised of 392 Soldiers, 209 in the intervention group and 183 in the comparison group. Based on actigraphic data, it was shown that Soldiers on the modified sleep schedule obtained 33 more minutes of total sleep per night than those on the standard sleep schedule. Soldiers in the intervention group reported less total mood disturbance relative to baseline, but the effect size was modest and diminished over the course of BCT. Improvements in Soldier marksmanship performance over a series of record fires was positively correlated with average nightly sleep during the week preceding the record fires, when basic marksmanship tasks were being learned. By the end of BCT, Soldiers in the comparison group were 2.3 times more likely to have occupationally significant fatigue and were 5.5 times more likely to report poor sleep quality, as assessed using validated survey instruments, than those in the comparison group. There was no effect of the sleep scheduling intervention on physical fitness scores or the relative risk for attrition. Overall, increasing sleep and concomitantly decreasing fatigue had a small but measurable influence on various indicators of Soldier functioning even after controlling for a variety of factors that affect performance.

15. SUBJECT TERMS
Biological rhythms, Human Systems Integration, Marksmanship, Military training, Operations research, Psychology, Physiology, Sleep

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UU	141	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This study examined the effect of alterations in the timing of sleep within the circadian cycle on the amount of total nightly sleep and its influence on various indicators of mood and performance of U.S. Army Soldiers attending Basic Combat Training (BCT) at Fort Leonard Wood, Missouri. The quasi-experimental study design compared Soldiers assigned to one of two training companies: a company using the standard BCT sleep regimen (8:30 p.m. to 4:30 a.m.) and a company using a phase-delayed sleep regimen (11:00 p.m. to 7:00 a.m.), the latter being more in line with the biologically driven sleep-wake patterns of adolescents. Demographic and psychophysiological measures were collected at the start of the study using standard survey instruments and methods. A random sample of approximately 24% of Soldiers wore wrist activity monitors to unobtrusively record sleep quantity and quality. Weekly assessments were made of subjective fatigue and mood throughout BCT. Data on physical fitness, marksmanship, and attrition from BCT were extracted from organizational training records.

The study sample was comprised of 392 Soldiers: 209 in the intervention group and 183 in the comparison group. Based on actigraphic data, it was shown that Soldiers on the modified sleep schedule obtained 33 more minutes of total sleep per night than those on the standard sleep schedule. Soldiers in the intervention group reported less total mood disturbance relative to baseline, but the effect size was modest and diminished over the course of BCT. Improvements in Soldier marksmanship performance over a series of record fires was positively correlated to the average nightly sleep during the week preceding the record fires, when basic marksmanship tasks were being learned. By the end of BCT, Soldiers in the comparison group were 2.3 times more likely to have occupationally significant fatigue and were 5.5 times more likely to report poor sleep quality (as assessed using validated survey instruments), than those in the comparison group. Sleep scheduling intervention had no effect on physical fitness scores or the relative risk for attrition. Overall, increased sleep, and its resultant decrease in fatigue,

had a small, but measurable, influence on various indicators of Soldier functioning, even after controlling for a variety of factors that could affect performance.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Recognizing that adolescents comprise the majority of military accessions, this study evaluated the performance impact of accommodating adolescent alterations in sleeping and waking patterns. Specifically, this study examined the effect of alterations in the timing of sleep within the circadian cycle on the amount of total nightly sleep and its influence on various indicators of mood and performance of U.S. Army Soldiers attending Basic Combat Training (BCT) at Fort Leonard Wood, Missouri. The quasi-experimental study design compared Soldiers assigned to one of two training companies: a company using the standard BCT sleep regimen (i.e., sleep period 8:30 p.m. to 4:30 a.m.) or a company using a phase-delayed sleep regimen (i.e., sleep period 11:00 p.m. to 7:00 a.m.), the latter being more in line with the biologically driven sleep-wake patterns of adolescents. Demographic and psychophysiological measures were collected at the start of the study using standard survey instruments and methods. A random sample of approximately 24% of Soldiers wore wrist activity monitors to unobtrusively record sleep quantity and quality. Weekly assessments were made of subjective fatigue and mood throughout the course of BCT. Data on physical fitness, marksmanship, and attrition from BCT were extracted from organizational training records.

The study sample was comprised of 392 Soldiers, 209 in the intervention group and 183 in the comparison group. Based on actigraphic data obtained from a sample of 94 Soldiers, it was shown that Soldiers on the modified sleep schedule obtained 33 more minutes of total sleep per night than those on the standard sleep schedule. The additional sleep obtained as a result of the sleep scheduling intervention was observed to have a modest impact on the mood state of Soldiers. Irrespective of treatment condition, the general trend was for Soldiers to report decreased feelings (relative to baseline) of tension-anxiety, depression-dejection, fatigue-inertia, and confusion-bewilderment over the course of BCT. Soldiers in the intervention group reported less anger-hostility and had lower total mood disturbance scores relative to the comparison group early in training, but these differences declined during BCT. Soldiers in the intervention group

reported significantly greater feelings of vigor, but the effect size was modest. While the effects of chronotype were mixed overall, the preponderance of evidence suggested that the phase-delayed sleep schedule preferentially impacted, in a positive direction, the mood state of evening chronotype Soldiers.

Sleep was also shown to be an important determinant of Soldier basic rifle marksmanship performance. In this study, it was demonstrated that the degree of improvement in marksmanship performance over the serial record fires was positively correlated to average nightly sleep during the week preceding the record fires, which was when basic marksmanship tasks were being learned. Thus, sleep appeared to potentiate the learning and recall of marksmanship skills. What is more, the effect size of sleep was greater than that attributable to prior experience with firearms.

Furthermore, the sleep scheduling intervention had significant safety and health effects. By the end of BCT, Soldiers in the comparison group were 2.3 times more likely to have occupationally significant fatigue and were 5.5 times more likely to report poor sleep quality (as assessed using validated survey instruments), than those in the comparison group. Moreover, the odds of Soldiers reporting poor quality sleep actually decreased for those in the intervention group relative to the start of the study, suggesting that the phase-delayed sleep schedule was an improvement over Soldiers' baseline sleep schedule. In contrast, sleep scheduling intervention had no effect on physical fitness scores or the relative risk for attrition.

In summary, increasing sleep had a small, but measurable, influence on various indicators of Soldier functioning, even after controlling for a variety of factors that affect performance. Although Soldiers' responses to the sleep schedule intervention were modest, it should be appreciated that the majority of outcome measures in BCT were not highly sensitive to the effects of fatigue. Thus, the most important finding of the study may be the impact of the schedule intervention on sleep quality during BCT—that is, Soldiers completing BCT using the phase-delayed sleep schedule had significant improvements in sleep hygiene, such that they graduated from training in a better physiological state than when they started. The significance of this finding may not be fully appreciated until Soldiers' subsequent performance is assessed during the more cognitively demanding secondary military occupational specialty training courses.

THIS PAGE INTENTNIONALLY LEFT BLANK

ACKNOWLEDGEMENTS

The authors wish to thank the U.S. Army Maneuver Support Center and Fort Leonard Wood for their sponsorship and assistance in conducting this study. The authors also wish to specifically thank the personnel of the Army Research Laboratory Human Research and Effectiveness Directorate's Maneuver and Mobility Branch at Fort Leonard Wood, who provided invaluable assistance in administering survey instruments, collecting actigraphic data, and providing local coordination with the participating training companies. Finally, the authors thank the officers, drill sergeants, and Soldiers of Bravo and Charlie Companies, 3rd Battalion, 10th Infantry Regiment, 3rd Chemical Brigade for their support and participation in this study.

I. INTRODUCTION

> The bed is a bundle of paradoxes: we go to it with reluctance, yet we quit it with regret; we make up our minds every night to leave it early, but we make up our bodies every morning to keep it late (Colton, 1837, p. 164).

A. STATEMENT OF THE PROBLEM

Military training regimes often include some degree of sleep deprivation, whether it is by design or unintentional. Several studies have demonstrated that sleep deprivation is prevalent in military training and education programs. For example, Killgore, Estrada, Wildzunas, and Balkin (2008), using actigraphy to assess sleep in Soldiers attending military training at the Noncommissioned Officer Academy and the Warrant Officer Candidate School, reported Soldiers obtained an average of 5.8 hours of sleep per night. Miller, Shattuck and Matsangas (2010), reporting on the preliminary results of a 4-year longitudinal study of sleep in U.S Military Academy (USMA) cadets based on actigraphy data, found that cadets averaged 5.4 hours of sleep per night. This finding is substantially less than the approximately eight hours of sleep per night required by healthy adults to maintain cognitive effectiveness (Anch, Browman, Mitler, & Walsh, 1988). Additionally, this finding is more than two hours less sleep per night than cadets reported receiving prior to arriving at the USMA (Miller & Shattuck, 2005). It is also important to recognize that military recruits are adolescents, or young adults, in their late teens and early twenties. Biologically driven sleep-wake patterns in this age group differ from those of more mature adults, with delayed bedtimes, later awakenings, and longer sleep periods (i.e., on the order of 0.50 to 1.25 more hours of sleep per night) (Carskadon, Acebo, Richardson, Tate, & Seifer, 1997; Carskadon, Wolfson, Acebo, Tzischinsky, & Seifer, 1998; Wolfson & Carskadon, 2003). Thus, the general population of military recruits may actually require from 8.50 to 9.25 hours of sleep per night for optimal performance (Miller & Shattuck, 2005).

Chronic sleep deprivation from multiple nights of less than eight hours of sleep will cause sleep debt and fatigue. A vast body of research has shown that the effects of fatigue include decreased vigilance, adverse mood changes, perceptual and cognitive decrements (Krueger, 1990; Belenky, Wesensten, Thorne, Thomas, Sing, Redmond, et al., 2003; Van Dongen, Maislin, Mullington, & Dinges, 2003), impaired judgment, and increased risk taking (Killgore, Balkin, & Wesensten, 2006), and even decreased marksmanship (Tharion, Shukitt-Hale, & Lieberman, 2003; McLellan, Kamimori, Bell, Smith, Johnson, & Belenky, 2005). Contrary to popular opinion in the military, research has shown that motivation can only partially compensate for the adverse effects of sleep deprivation (Pigeau, Angus, & O'Neil, 1995).

Of particular relevance to military training, the ability of individuals to learn and retain information is reduced by sleep deprivation (literature summarized in Miller, Matsangas, & Shattuck, 2007). For example, Graham (2000) reports that learning curves drop dramatically for adolescents obtaining 4-6 hours of sleep relative to those obtaining eight hours per night. In the military training environment, Andrews (2004) conducted a retrospective comparison of the academic performance of Navy recruits before and after the training command leadership changed the sleep regime from six to eight hours per night. It was observed that recruits who received eight hours of sleep per night scored, on average, 11% higher than their counterparts who received only six hours of sleep, although Andrews was unable to discount the impact of other, concurrent changes at the training command. In contrast, Baldus (2002) collected actigraphic data on 31 Navy recruits at the same training command who were all assigned to two sleep conditions (9:00 p.m. to 5:00 a.m. and 10:00 p.m. to 6:00 a.m.) in a cross-over study design. It was shown that recruits obtained an additional 22 minutes of sleep when on the 1-hour phase-delayed sleep schedule, but no attempt was made to correlate this observation with measures of recruit performance.

Moreover, Killgore and colleagues (2008), evaluating the effectiveness of actigraphy as a predictor of cognitive performance, found significant positive correlations between Soldier academic exam scores in six military education programs (i.e., programs of instruction at the Noncommissioned Officer Academy and Warrant Officer Candidate School at Fort Rucker, Alabama) and the following sleep indices: average hours of sleep per night and hours slept in the 24- and 48-hour periods preceding an exam. They also report that the average amount of sleep obtained by Soldiers accounted for approximately

40% of the variance in exam scores—a finding that underscores the impact of fatigue on learning and memory. A similar result was reported by Trickel, Barnes, and Egget (2000), who found that sleep habits accounted for most of the variance in the academic performance of freshman college students.

Physical health is an equally important concern in military recruit populations, particularly because the close living conditions are conducive to the spread of communicable disease. Individual physical health and, in turn, public health, also depends on individuals receiving adequate amounts of sleep. Research has shown that disturbances of sleep-wake homeostasis are accompanied by alterations in the immunological, neuroendocrine, and thermoregulatory functions of the body and, hence, contribute to pathological processes such as infectious disease (Moldofsky, 1995). Lange, Perras, Fehm, and Born (2003) also report that sleep enhances antibody production and the immune response to vaccination. Besides illness, sleep deprivation threatens health by increasing the risk for injuries resulting from accidents. For example, Thorne, Thomas, Russo, Sing, Balkin, Wesensten, et al. (1999) demonstrated that accidents increase progressively as sleep duration decreases to 7, 5, and 3 hours per night over a period of one week.

Scientific literature suggests there is a high prevalence of fatigue in military recruits, which has important implications for Soldier training, health, and safety. Well-controlled laboratory experiments have demonstrated a convincing dose-response relationship between sleep deprivation and degraded cognitive performance (Belenky et al., 2003; Driskell, Hughes, Willis, Cannon-Bowers, & Salas, 1991; Driskell & Salas, 1996; Hursh & Bell, 2001; Van Dongen et al., 2003) (as discussed in Miller, Matsangas, & Shattuck, 2007). However, the design of prior studies of fatigue in military training environments has been primarily descriptive in nature, limited to correlations between sleep and academic test performance, and many of the recommendations for follow-on research have yet to be followed. The only field study to directly examine the effect of a phase-delayed sleep scheduling intervention in the military training environment (Baldus, 2002) did not include any assessment of performance outcomes. Thus, whether designing schedules to minimize fatigue would have a direct effect on outcomes in the military training environment remains an open question (Miller, Shattuck, Matsangas & Dyche, 2008).

The scarcity of information on the benefit of sleep scheduling interventions for military training is regrettable because it is the sort of evidence that senior decision makers require if they are to support fatigue-sensitive revisions to training regimes. If sleep scheduling is found to have a significant effect on overall training effectiveness and recruit attrition, health, and safety, then two options become available for the military training community:

- Performance thresholds of achievement for basic military training can be increased, while maintaining the present length of training (optimizing training effectiveness), or
- Thresholds of achievement can be maintained and the length of training decreased (optimizing training efficiency).

Preliminary evidence suggests that sleep, and conversely fatigue, may account for nearly half the variability in academic performance during military training (Killgore et al., 2008). Additionally, implementing a phase-delayed sleep scheduling intervention during military training appears to result in measurable increases in total daily sleep (Baldus, 2002). Collectively, these observations suggest that sleep scheduling is a potentially powerful lever for manipulating the performance of military training programs—and one that is immediately within our grasp without making a significant investment in new technologies. Since training is a potential bottleneck in meeting wartime manpower needs as well as a recurring life-cycle cost for all weapon systems, even a more modest 10% improvement in trainee performance, as suggested by Andrews (2004), is significant when one considers the cumulative impact across military training programs.

This study attempts to contribute to the knowledge base by exploring the influence of sleep scheduling in the Basic Combat Training environment on Soldiers' achievement of entry-level standards and combat skills. This study examines the direct effect of sleep scheduling on motivation and mood state and training, health, and safety outcomes, while controlling for such individual differences as sleep habits, personality, and personnel aptitudes.

B. PURPOSE OF THE STUDY

The purpose of this study is to examine the effect of alterations in the timing of sleep within the circadian cycle on the amount of total nightly sleep and its influence on various indicators of mood and performance of U.S. Army Soldiers attending Basic Combat Training at Fort Leonard Wood, Missouri. The study design compares Soldiers assigned to one of two training companies: a company using the standard Basic Combat Training sleep regimen (i.e., sleep period 8:30 p.m. to 4:30 a.m.) or a company using a phase-delayed sleep regimen (i.e., sleep period 11:00 p.m. to 7:00 a.m.), the latter being more in line with the biologically driven sleep-wake patterns of adolescents.

To account for some of the myriad factors that are assumed to play a role in daytime functioning, a number of factors are selected as control variables or covariates (Table 1). These control variables include background information about each Soldier (e.g., age, sex, caffeine and tobacco habits, prior experience with firearms, etc.) and information about their sleep habits, personality, resilience, and personnel aptitudes. The inclusion of these individual characteristics is important to this study because we predict that sleep timing will have a small, but measurable, influence on daytime functioning, even after controlling for the contributions of the usual variables thought to affect mood state and performance.

Table 1. Summary of study variables.

Independent variables	Dependent variables
Age	Attrition
Caffeine and tobacco habits	Basic rifle marksmanship
Personality	Mood state
Personnel aptitude	Physical fitness
Prior experience with firearms	
Resilience	
Sex	
Sleep habits	
Sleep schedule	

Consequently, at weekly intervals, Soldiers are asked to identify their mood state over the prior week of training. Mood state is defined by six general factors identified in the Profile of Mood States (POMS) (McNair, Lorr, & Droppleman, 1981). These six factors are tension-anxiety, depression-dejection, anger-hostility, vigor-activity, fatigue-inertia, and confusion-bewilderment. These six factors can also be aggregated into a total mood disorder score. The study primarily examines three major performance outcomes of concern to the military training organization: attrition, basic rifle marksmanship, and physical fitness.

C. THEORETICAL PERSPECTIVE

In formulating a theoretical perspective for considering the effects of a sleep scheduling intervention on training effectiveness, biomathematical models of sleep and circadian processes provide a useful prototype. The first biomathematical models of sleep and circadian processes were developed more than 20 years ago in an effort to explain the timing of the human sleep-wake activity cycle. In the intervening years, a number of applied biomathematical models of fatigue and performance have been developed from the first generation of models of sleep-wake cycles. These applied biomathematical models typically use information about sleep history, duration of wakefulness, and circadian phase to predict performance capability and risk. They are currently used to assess the potential contribution of fatigue to performance degradation at specific points in time, to develop and evaluate work/rest schedules, to plan work and sleep in operational missions, and to determine the timing of fatigue countermeasures to anticipated performance decrements (Neri, 2004). The March 2004 edition of the journal, *Aviation*, *Space*, *and Environmental Medicine*, provides a comprehensive review and model-to-data comparisons of seven of the current biomathematical models of human fatigue and performance. Those interested in more information on the biomathematical modeling of fatigue and performance should refer to this resource and its bibliographies.

The U.S. Defense Department has long pursued applied research concerning fatigue in military operations and has developed several biomathematical fatigue models. One of these models, known as the Sleep, Activity, Fatigue, and Task Effectiveness (SAFTE) Model, has achieved relatively wide acceptance and seen practical application within the Fatigue Avoidance Scheduling Tool (FAST) (Hursh, Redmond, Johnson, Thorne, Belenky, Balkin, et al., 2004). The SAFTE model is shown in Figure 1 using a system dynamics modeling stock and flow diagram. The conceptual architecture of the SAFTE model centers on a sleep reservoir, representing sleep-dependent processes that govern the capacity to perform cognitive work. Using the language of system dynamics modeling, the stock of this reservoir is cognitive work capacity. Sleep is a replenishing flow into the reservoir, while wakefulness is a depleting flow out of the reservoir. Replenishment, in terms of sleep accumulation, is determined by information about the time-of-day of sleep, reservoir level (i.e., sleep debt), and sleep quality (i.e., sleep fragmentation). The system modeled in Figure 1 provides output in terms of performance effectiveness, which is simultaneously modulated by circadian effects and the level of the reservoir (Hursh et al., 2004).

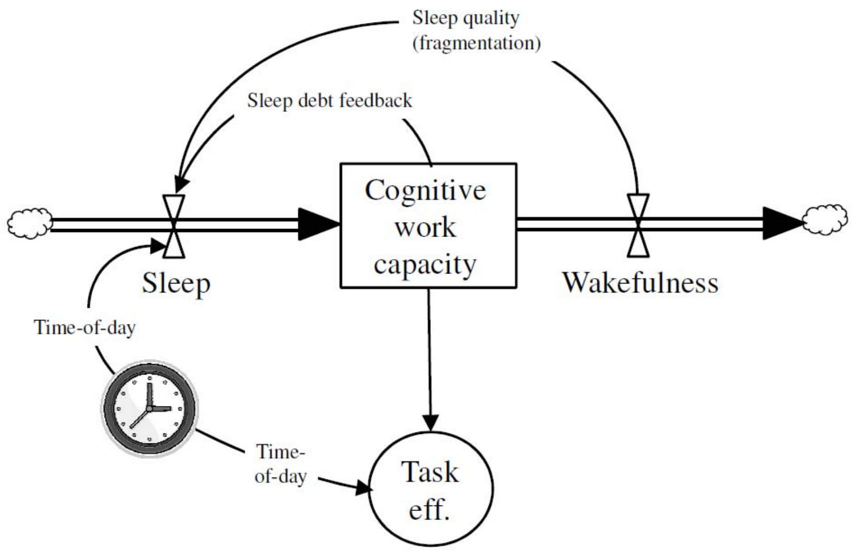


Figure 1. Stock and flow diagram of the SAFTE model.

The SAFTE model has been shown to predict changes in cognitive capacity as measured by standard laboratory tests of cognitive performance with reported coefficients of determination ranging from 89% to 94%. It is presumed that these cognitive tasks measure changes in the fundamental capacity to perform a variety of real-world tasks that rely on such cognitive skills as discrimination, reaction time, mental processing, reasoning, and language comprehension and production. Although specific military tasks may vary in their reliance on these skills, Hursh and colleagues (2004) assert that it is reasonable to assume that changes in military task performance will correlate with changes in the underlying cognitive effectiveness. Hence, one would expect to see a direct relationship between measured changes in cognitive effectiveness and military task performance.

Based on the structure of the SAFTE model, the reservoir or stock of cognitive work capacity, shown in Figure 1, will reach a time-averaged equilibrium state provided an individual remains on a constant schedule (Hursh et al., 2004). Consequently, the following statement represents the underlying logic for designing and conducting this study. If we design a schedule so the timing of sleep-wake periods improves the overall equilibrium state of Soldiers' work capacity reservoirs (and consequently, their cognitive task effectiveness) it follows that (1) individual Soldier task performance should improve, resulting in a greater proportion of recruits who meet specified performance criteria; and (2) this effect should be greater for those Soldiers with lower personnel aptitudes, as their performance margin, relative to the specified performance criteria, is expected to be smaller. The predicted relationship between personnel aptitude, schedule, and their interaction and the outcome, as expressed in terms of the proportion of proficient trainees, is illustrated in Figure 2. As shown, Schedule A results in a more favorable equilibrium state of Soldiers' reservoirs than Schedule B, which is to say that Schedule A is more complementary to Soldiers' natural circadian cycles. Hence, Schedule A is more effective overall, but it is particularly beneficial for recruits on the lower end of the personnel aptitude spectrum.

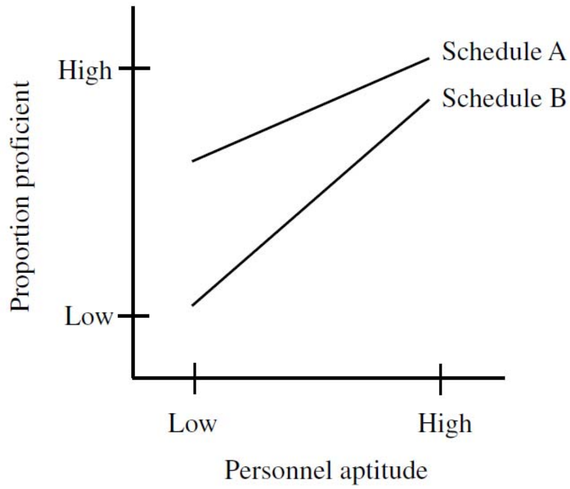


Figure 2. Hypothetical interactive effects of aptitude and two training schedules on learning outcomes (expressed in terms of the proportion of proficient trainees).

It is worth noting that if we replaced the word "schedule" with "treatment" in Figure 2, we would have the depiction of an ordinal aptitude treatment interaction (ATI) as described in ATI theory (Whitener, 1989). The underlying premise of ATI theory is that learning, and subsequent performance, is higher when the learning method, or treatment, capitalizes on an individual's cognitive aptitudes (Snow, 1978). In a twist on ATI theory, this study involves no change in learning methods *per se*, but rather, the treatment changes the relative availability of cognitive resources. Again, the underling logic for this study would suggest that if a schedule enhances cognitive resources, then (1) this change should be manifest by increased performance on learning tasks, and (2) performance enhancements should be greater for those individuals with less aptitude, given their overall higher demand for cognitive resources during training.

D. STUDY HYPOTHESES

The following hypotheses guide this study:

H_1: Participants on the modified, phase-delayed sleep schedule will obtain more daily sleep than participants following the standard Basic Combat Training schedule.

H_2: Participants on the modified sleep schedule will have less decrement in mood state (relative to baseline) than participants following the standard Basic Combat Training sleep schedule.

H_3: Participants on the modified sleep schedule will exhibit greater improvement in basic rifle marksmanship scores than participants following the standard Basic Combat Training sleep schedule.

H_4: Participants on the modified sleep schedule will exhibit greater improvement in physical fitness scores than participants following the standard Basic Combat Training sleep schedule.

H_5: The likelihood of participants on the modified sleep schedule reporting occupationally significant fatigue will be lower than that for participants following the standard Basic Combat Training sleep schedule.

H_6: The likelihood of participants on the modified sleep schedule reporting poor sleep quality will be lower than that for participants following the standard Basic Combat Training sleep schedule.

H_7: The likelihood of participants on the modified sleep schedule attriting from training will be lower than that for participants following the standard Basic Combat Training sleep schedule.

E. DELIMITATIONS AND LIMITATIONS

A Delimitation:

This study is confined to assessing and observing U.S. Army Soldiers assigned to two companies within a combat support training battalion at Fort Leonard Wood, Missouri.

Limitations:

The study sample consists of Soldier accessions into military occupational specialties within the U.S. Army's combat support branch. Since combat support units may differ from combat arms and combat service support units in terms of the distributions of sex and personnel aptitudes, this study may not be generalizable to all Army training programs.

The study sample consists of Soldier accessions into the U.S. Army in the month of August. Since the demographics of Soldiers entering Basic Combat Training exhibit a seasonal variation, the findings of this study may not directly apply to other Basic Combat Training classes at the study location.

II. METHODS

A. RESEARCH DESIGN

The study protocol was approved by the Naval Postgraduate School Institutional Review Board in accordance with 32 Code of Federal Regulations 219 and SECNAV Instruction 3900.39D. The study used a quasi-experimental study design that was embedded within the Army's 63-day Basic Combat Training program of instruction. The intervention and comparison groups were selected without random assignment, although group assignment to the treatment condition was random. Participant assignment to group was made by the U.S. Army based on factors that were unobservable by the research team, but which were not altered for the purpose of this study. That is, the research team took the groups as they were created by the U.S. Army based on their normal mode of operations for managing Basic Combat Training. The study intervention consisted of a modification of the timing of sleep and wake periods; otherwise, no change was made to the content, instructional methods, or sequence of Basic Combat Training events. The intervention group used a phase-delayed (i.e., 11:00 p.m. to 7:00 a.m.) sleep regimen with opportune midday naps, while the comparison group maintained the standard (i.e,. 8:30 p.m. to 4:30 a.m.) sleep regimen. The barracks used by the intervention group were modified with black-out curtains to mitigate the effect of morning light; no modifications were made to the barracks used by the comparison group.

B. PARTICIPANTS

Participants for the comparison group were solicited from among those Soldiers starting Basic Combat Training on August 14, 2009, and assigned to Charlie Company, 3rd Battalion, 10th Infantry Regiment, 3rd Chemical Brigade (C/3-10 IN BN), Fort Leonard Wood, Missouri. Similarly, participants for the intervention group were solicited from among those Soldiers starting Basic Combat Training on August 21, 2009, and assigned to Bravo Company, 3rd Battalion, 10th Infantry Regiment (B/3-10 IN BN). Participants for both groups were solicited during Basic Combat Training inprocessing by a civilian member of the research team to mitigate the potential for implied coercion by rank. Soldiers who chose not to participant in the study (less than 1%) still followed the training company's schedule and accomplished all training events, but they did not complete any of the study-related instruments.

C. DATA COLLECTION INSTRUMENTS AND VARIABLES

1. Actiwatch

The Actiwatch® (Model AW-64, Philips Respironics, Bend, Oregon) is a 16-gram, 28 x 27 x 10-millimeter wristwatch-like device worn on the nondominant wrist that objectively measures activity and rest patterns. With each participant movement, a highly sensitive accelerometer generates a variable voltage that is digitally processed and sampled at a frequency of 32 Hertz. The signal is integrated over a user-selected epoch and a value expressed as activity counts is recorded in the on-board memory. Data are downloaded to a computer and may be expressed graphically as an actogram or reported in American standard code for information interchange (ASCII) format numerically as total activity counts per epoch.

2. Basic Rifle Marksmanship

Objective evaluation of rifle marksmanship skill was made based on "record fire" score. During a Basic Combat Training record fire, Soldiers are given an M16/M4 series rifle and 40 rounds of ammunition and presented with 40 timed target exposures at ranges from 50 to 300 meters. Twenty targets are engaged with 20 rounds from the prone supported position, ten targets are engaged with ten rounds from the prone unsupported position, and ten targets are engaged with ten rounds from the kneeling position—while wearing a helmet and load-bearing equipment. The standard is to obtain at least 23 target hits on the 40 targets exposed. Soldiers complete a practice record fire on days 29 and 30 of Basic Combat Training and an official record fire on day 32 of Basic Combat Training, for a total of three sequential record fires (Directorate Basic Combat Training Doctrine and Training Development, 2008).

3. General Technical Aptitude

Objective evaluation of individual aptitude was made based on General Technical (GT) score as derived from the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB is a 216-item inventory containing nine separately timed subtests: General Science, Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, Auto and Shop, Mathematics Knowledge, Mechanical Comprehension, Electronics Information, and Assembling Objects. The ASVAB is not an intelligence test, but rather, is specifically designed to measure an individual's aptitude to be trained in specific jobs. GT score is a composite of the Arithmetic Reasoning, Word Knowledge, and Paragraph Comprehension subtests, and it is often a major determinant of the occupational specialties for which a person can be considered in the military.

4. Mood State

Subjective evaluation of mood was made with the Profile of Mood States (POMS) (McNair et al., 1981). The POMS is a 65-item questionnaire that measures affect or mood on six scales: (1) tension-anxiety, (2) depression-dejection, (3) anger-hostility, (4) vigor-activity, (5) fatigue-inertia, and (6) confusion-bewilderment. An aggregate mood disturbance score is calculated by summing the scores on the six scales and negatively weighting the vigor-activity score.

5. Personality

A personality assessment was accomplished using the Neuroticism-Extroversion-Openness Five-Factor Inventory (NEO-FFI) (Costa & McCrae, 1992). The NEO-FFI is essentially a short form of the Revised NEO Personality Inventory (NEO-PI-R). It consists of 60 items from the NEO-PI-R that are used to score the five domains: (1) neuroticism, (2) extraversion, (3) openness, (4) agreeableness, and (5) conscientiousness. It does not contain the items for assessing the facets within each domain. The NEO-FFI is designed for use in circumstances in which time is too limited to present the entire NEO-PI-R or when only scores on the five domains are required (Weiner & Greene, 2008).

6. Physical Fitness

Objective evaluation of physical fitness was made based on Army Physical Fitness Test (APFT) scores. Soldiers complete a physical fitness assessment consisting of three measured events: push-ups, sit-ups, and a timed 2-mile run. Raw scores are scaled for both age and sex. Soldiers must earn a score of 150 points or higher on the end-of-training APFT with 50 points or more in each event to graduate from Basic Combat Training (Directorate Basic Combat Training Doctrine and Training Development, 2008). Soldiers complete two diagnostic APFTs during the third and sixth weeks of Basic Combat Training and a final APFT in the eighth week of training.

7. Resilience

Assessment of resilience to stress was accomplished using the Response to Stressful Experiences Scale (RSES) (Johnson, Polusny, Erbes, King, King, Litz, et al., 2008). The RSES was developed by researchers with the National Center for Post Traumatic Stress Disorder to rate psychological traits that promote resilience, which is the ability to undergo stress and still retain mental health and well-being. It consists of 22 items and identifies six factors that are key to psychological resilience: (1) positive outlook, (2) spirituality, (3) active coping, (4) self-confidence, (5) learning and making meaning, and (6) acceptance of limits. The RSES has been tested on more than 1,000 active-duty military personnel (Naval Center for Combat and Operational Stress Control, 2009).

8. Sleep Habits

Subjective assessments of sleep habits were made using three validated survey instruments. The first instrument was the Pittsburgh Sleep Quality Index (PSQI), a self-rated questionnaire designed to measure sleep quality in clinical populations by looking at sleep in the previous month. Nineteen individual items generate the following seven scores: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) habitual sleep efficiency, (5) sleep disturbances, (6) use of sleeping medications, and (7) daytime dysfunction. A review of this survey’s reliability asserts that the PSQI is useful to both psychiatric clinical practice and research activities (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989).

The second instrument was the Epworth Sleepiness Scale (ESS) (Johns, 1991), which is an 8-item scale commonly used to diagnose sleep disorders and considered a valid and reliable self-report of sleepiness. Participants use an integer number from 0 to 3, corresponding to the likelihood (never, slight, moderate, and high, respectively) that they would fall asleep in eight situations such as sitting and reading, watching television, as a passenger in a car for an hour, etc. Ratings above 10 out of a possible 24 are cause for concern with respect to an underlying sleep disorder.

The third instrument was the Morningness-Eveningness Questionnaire (MEQ) published by Horne and Ostberg (1976), which contains 19 questions aimed at determining when, during the daily temporal span, individuals have the maximum propensity to be active. Most questions are preferential, in the sense that the respondent is asked to indicate when they would prefer, rather than when they actually do, wake up or begin sleep. Questions are multiple-choice and each answer is assigned a value such that their sum gives a score ranging from 16 to 86, with lower values corresponding to evening chronotypes and higher values indicating morning chronotypes.

9. Study Questionnaires

The prestudy questionnaire contained ten questions aimed at potential covariates that could influence study outcome measures. Four questions asked participants for their age, sex, height, and weight. One question asked participants to quantify their frequency of exercise during the preceding month, both in terms of the number and duration of exercise sessions. Another question asked whether participants regularly used firearm(s) and, if so, to characterize the type of firearm(s), reason(s) for use, and frequency of use. Three questions addressed use of caffeinated beverages, tobacco, and medications. Lastly, one question asked participants to quantify the amount of sleep per day they required to feel ready to start the day.

The poststudy questionnaire consisted of six questions. Similar to the pretest questionnaire, two questions addressed use of caffeinated beverages and medications, and one question asked participants to quantify the amount of sleep per day they required to feel ready to start the day. One question asked participants about the frequency with which they fell asleep during activities. Another question asked participants to provide an ordinal ranking on a 5-item Likert scale of the adequacy of both their sleep and that of their peers during Basic Combat Training. The final question asked participants' preference for the timing of daily physical training.

D. PROCEDURES

1. General

Prior to beginning the study, each participant received a full briefing on the purposes of the study and assurances about the confidentiality of the data. Once informed consent was obtained, each participant completed the prestudy questionnaire followed by the ESS, PSQI, MEQ, RSES, POMS, and NEO Five Factor Inventory (Table 2). Participants subsequently accomplished the POMS at weekly intervals throughout Basic Combat Training. At the completion of Basic Combat Training, participants received an out-briefing and completed the poststudy questionnaire followed by the ESS, PSQI, and the final POMS. For each participant, data were collected on general technical aptitude, basic rifle marksmanship, and physical fitness scores from preexisting local databases. Attritions were determined from training company graduation rosters.

Table 2. Schedule for data-generating events.

↓Data Event Week→	1	2	3	4	5	6	7	8	9
Actigraphy*	X	X	X	X	X	X	X	X	X
Army Physical Fitness Test			X			X		X	
Basic Rifle Marksmanship					X				
Epworth Sleepiness Scale	X								X
Morningness-Eveningness Questionnaire	X								
NEO Five-Factor Inventory	X								
Pittsburgh Sleep Quality Index	X								X
Profile of Mood States	X	X	X	X	X	X	X	X	X
Response to Stressful Experiences Scale	X								
Study Questionnaires	X								X

*Actigraphy data was collected on a random subsample of the study participants.

2. Actigraphy

A random sample, comprised of approximately 20% of participants in each study group, was selected for actigraphic data collection. Participants agreeing to actigraphic data collection were issued an Actiwatch® on Day 1 to track sleep and activity patterns in a relatively unobtrusive fashion. Participants were asked to wear the Actiwatch® continuously on the wrist of their nondominant hand during all waking and sleeping periods and not to remove it for showering. The Actiwatch® was collected from each participant during Week 4 (intervention group) or Week 5 (comparison group) for downloading of data and reinitialization of the Actiwatch® data collection mode. Once the data collection period was complete, the data were taken back to the laboratory and, using Actiware® version 5.57.0006 software, scored for sleep times.

3. Statistical Analysis

For the prestudy and poststudy questionnaires and the ESS, PSQI, MEQ, and RSES survey instruments, item nonresponse was handled using stochastic regression imputation to reduce the bias that could be caused by ignoring records with missing data (Kim & Curry, 1977; Brick & Kalton, 1996). For the NEO-FFI and the POMS survey instruments, item nonresponse was handled per the guidance in the associated survey technical manuals. In the case of the weekly POMS, which were administered repetitively throughout the course of training, no attempt was made to address unit or partial nonresponses. Microsoft® Office Excel® 2007 was used to develop the study database; histograms of the actigraphy data were created using the Analysis ToolPak add-in. Analyses were undertaken with the Statistical Package for the Social Sciences (SPSS) version 11. All data were assessed for normalcy, and parametric and nonparametric approaches were used accordingly for descriptive statistical analyses. Separate univariate and repeated measures analyses of covariance (ANCOVAs) were used to test major hypotheses involving measures with one dependent variable. Repeated measures were analyzed using a univariate approach with a fixed effect for time when there were a substantial number of unit nonresponses, thereby reducing the danger of biased repeated measures estimates of treatment effects caused by ignoring records with missing responses. ANCOVA results were examined to determine whether there were sphericity violations of sufficient magnitude to warrant the use of Huynh-Feldt adjusted degrees of freedom. Multivariate analysis of covariance (MANCOVA) was used to test hypotheses involving measures with more than one dependent variable. Box and Levene tests were used to assure the multivariate assumptions of equality of covariance matrices and that equality of error variances across groups was not violated. Lastly, logistic regression was used to test major hypotheses involving measures with a binary dependent variable.

III. RESULTS

A. PARTICIPANTS (*N* = 392)

The study sample was comprised of 392 participants, 209 in the intervention group and 183 in the comparison group. Participants' responses on the prestudy questionnaire and survey instruments are summarized in Tables 3 through 5 by treatment condition, that being either assignment to the intervention or comparison group. Figures 3 through 5 display histograms for a select subset of questions from the PSQI asking participants about their baseline sleep schedule. From the outset of the study, the intervention and comparison groups were generally comparable, although they did differ on some of the measured variables:

- Participants in the intervention group tended to have a higher body mass index (i.e., body weight corrected for height) than those in the comparison group.
- A greater proportion of participants in the intervention group were in the National Guard/Reserves as compared to the comparison group.
- Participants in the comparison group reported higher levels of neuroticism on the NEO-FFI, while participants in the intervention group reported higher levels of conscientiousness.
- Participants in the comparison group tended to have higher global scores on the prestudy PSQI, mainly because of increased daytime dysfunction. Also, a greater proportion of participants in the comparison group met the threshold score for being classified as potentially having poor quality sleep.
- Participants in the intervention group had higher levels of spirituality, active coping, and self-efficacy, and hence, overall resilience, as assessed by the RSES at the outset of the study.

Table 3. Summary of intervention and comparison study groups at outset of study.

Variable	Group		*p*-value
	Intervention (*n* = 209)	Comparison (*n* = 183)	
Age (yrs), median (IQR)	20 (18-23)	20 (18-24)	0.762^{M}
Body mass index (kg·m^{-2}), median (IQR)	25.4 (22.9-28.4)	23.6 (21.6-26.8)	0.002^{M}*
Body mass index category, no. (%)			
Underweight	5 (2.4)	6 (3.3)	0.021^{C}*
Normal	87 (41.6)	102 (55.7)	
Overweight	81 (38.8)	57 (31.1)	
Obese	36 (17.2)	18 (9.8)	
Caffeine			
Consume caffeinated beverages, no. (%)	116 (55.5)	110 (60.1)	0.357^{C}
Caffeine use (mg·d^{-1}), median (IQR)	39.0 (0-157.5)	61.0 (0-177.0)	0.248^{M}
Component, no. (%)			
National Guard	72 (34.4)	58 (31.7)	<0.001^{C}*
Regular	82 (39.2)	109 (59.6)	
Reserves	55 (26.3)	16 (8.7)	
Epworth Sleepiness Scale			
Total score, median (IQR)	8 (6-11)	9 (6-11)	0.562^{M}
Excessive fatigue (score > 10), no. (%)	52 (24.9)	52 (28.4)	0.429^{C}
Exercise frequency (hrs·wk^{-1}), median (IQR)	2.5 (1.0-4.5)	3.0 (1.5-5.9)	0.071^{M}
Firearms			
Regularly use firearm, no. (%)	51 (24.4)	39 (21.3)	0.468^{C}
Type of firearm, no. (%)			
Rifle	44 (21.1)	31(16.9)	0.302^{C}
Handgun	28 (13.4)	23 (12.6)	0.808^{C}
Use of firearm, no. (%)			
Hunting	36 (17.2)	28 (15.3)	0.607^{C}
Sport shooting	32 (15.3)	28(15.3)	0.998^{C}
Other	7 (3.8)	4 (1.9)	0.253^{C}
Frequency of use (days·yr^{-1}), median (IQR)	0 (0-0)	0 (0-0)	0.540^{M}

*Significant at ≤ 0.05 level.
^{C}Chi square statistic, ^{M}Mann-Whitney U.
Note: IQR = interquartile range.

Table 4. Summary of intervention and comparison study groups at outset of study (continued).

Variable	Group		*p*-value
	Intervention (n = 209)	Comparison (n = 183)	
GT score, median (IQR)	105 (96-114)	108 (99-116)	0.057^{M}
Morningness-Eveningness Questionnaire			
Total score, median (IQR)	50 (45-55)	49 (42-56)	0.498^{M}
Chronotype, no (%)			
Evening type	39 (18.7)	34 (18.6)	0.291^{C}
Neither type	140 (67.0)	112 (61.2)	
Morning type	30 (14.3)	37 (20.2)	
NEO Five Factor Inventory, median (IQR)			
Neuroticism	52 (45-59)	55 (47-63)	0.012^{M}*
Extraversion	53 (46-61)	53 (46-60)	0.601^{M}
Openness to experience	48 (41-58)	50 (41-57)	0.712^{M}
Agreeableness	46 (36-53)	44 (36-52)	0.224^{M}
Conscientiousness	50 (43-57)	46 (38-53)	0.003^{M}*
Pittsburgh Sleep Quality Index			
Global score, median (IQR)	6 (4-9)	7 (5-10)	0.048^{M}*
Poor sleep quality (score > 5), no. (%)	123 (58.9%)	129 (70.5%)	0.016^{C}*
Component scores, median (IQR)			
Subjective sleep quality	1 (1-1)	1 (1-2)	0.190^{M}
Sleep latency	2 (1-4)	2 (1-4)	0.817^{M}
Sleep duration	0 (0-2)	0 (0-2)	0.430^{M}
Habitual sleep efficiency	0 (0-0)	0 (0-0)	0.203^{M}
Sleep disturbances	1 (1-2)	1 (1-2)	0.399^{M}
Use of sleeping medication	0 (0-0)	0 (0-0)	0.400^{M}
Daytime dysfunction	1 (0-1)	1 (1-1)	0.001^{M}*
Rank, no (%)			
E01	82 (39.2)	62 (33.9)	0.514^{C}
E02	69 (33.0)	58 (31.7)	
E03	43 (20.6)	48 (26.2)	
E04	15 (7.2)	15 (8.2)	

*Significant at ≤ 0.05 level.
^{C}Chi square statistic, ^{M}Mann-Whitney U.
Note: IQR = interquartile range.

Table 5. Summary of intervention and comparison study groups at outset of study (continued).

Variable	Group: Intervention (n = 209)	Group: Comparison (n = 183)	p-value
Response to Stressful Experiences Scale			
Global score, median (IQR)	69 (60-78)	67 (58-75)	0.008[M]*
Factor scores, median (IQR)			
Positive appraisal	7.3 (6.2-8.3)	7.0 (5.9-8.0)	0.141[M]
Spirituality	2.9 (2.9-3.8)	2.9 (2.7-3.8)	0.001[M]*
Active coping	10.8 (8.9-12.2)	10.2 (8.2-11.5)	0.001[M]*
Self-efficacy	3.2 (2.4-3.2)	2.4 (2.4-3.2)	0.029[M]*
Learning and meaning-making	6.6 (5.4-8.0)	6.5 (5.1-7.3)	0.025[M]*
Acceptance of limitations	4.9 (3.5-5.6)	4.3 (3.5-5.0)	0.055[M]*
Sex, no. (%)			
Female	67 (32.1)	52 (28.4)	0.434[C]
Male	142 (67.9)	131 (71.6)	
Tobacco			
Regularly use tobacco, no (%)	81 (38.8)	68 (37.2)	0.745[C]
Frequency of use ($cigs \cdot wk^{-1}$), median (IQR)	0 (0-28)	0 (0-16)	0.519[M]

*Significant at ≤ 0.05 level.
[C]Chi square statistic, [M]Mann-Whitney U.
Note: IQR = interquartile range.

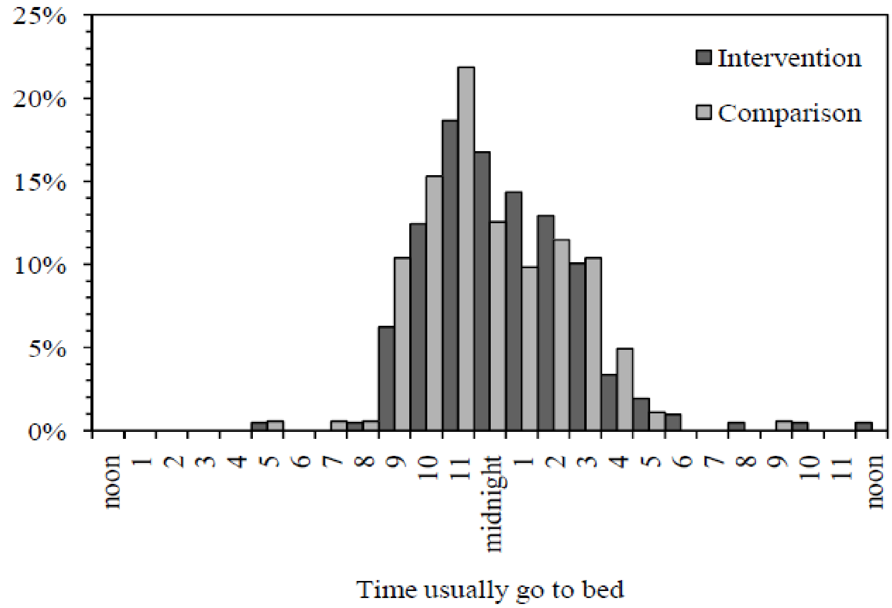


Figure 3. Histogram of participants' reported usual bed time (PSQI question 1).

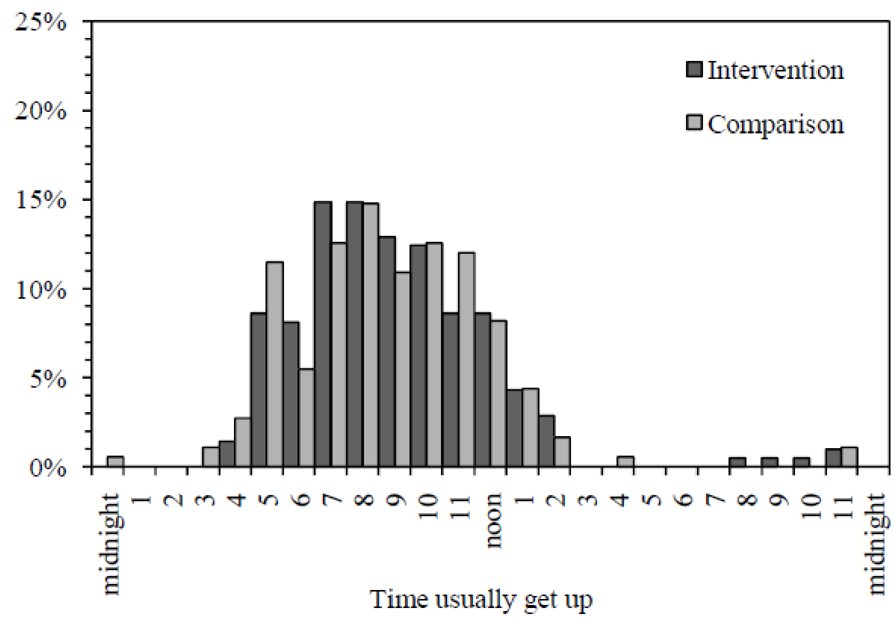


Figure 4. Histogram of participants' reported usual wakeup time (PSQI question 3).

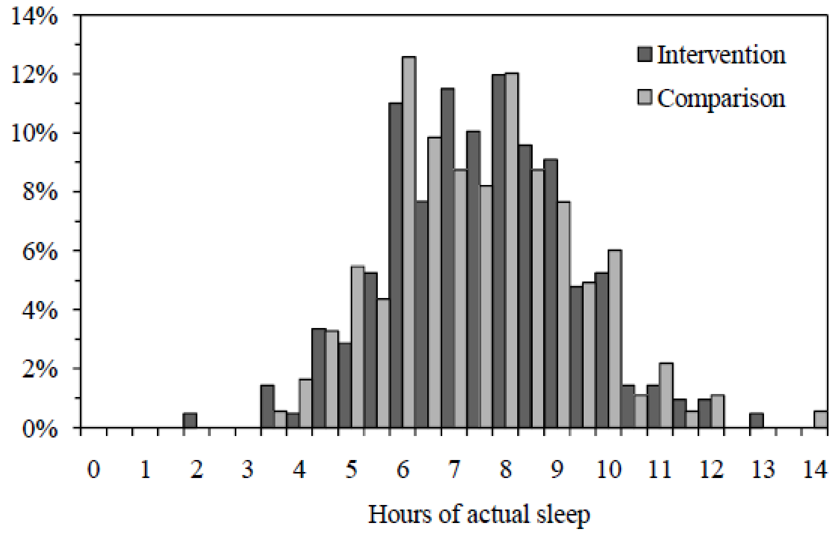


Figure 5. Histogram of participants' reported hours of sleep per night (PSQI question 4).

B. ACTIGRAPHY SUBSAMPLE

1. Participants (*n* = 95)

What follows in this subsection is limited to the subsample of 95 participants (53 in the intervention group and 42 in the comparison group), randomly selected to wear Actiwatches®. Due to unexplained technical difficulties, data were not recorded on Actiwatches® given to one participant in the comparison group. Consequently, this participant's other data were censored in the subsequent analysis, thereby leaving us with a subsample of 94 participants. Across the subsample, on average, 83.8 (standard deviation 9.6; range 36-92) participants had a valid Actiware® score for any given day of Basic Combat Training. A one-way analysis of variance (ANOVA) was used to compare the number of participants per day with a valid Actiware® score by week of training. Overall, there was a significant difference in week ($F_{8,52}$ = 3.205, p = 0.005), but Bonferroni post-hoc tests showed that this difference was only between Week 2 (mean 90.7 participants) and Week 9 (mean 73.4 participants).

Participants' responses on the study questionnaire and survey instruments are summarized in Tables 6 through 8 by treatment condition, that being either assignment to the intervention or comparison group. Figures 6 through 8 display histograms for a select subset of questions from the PSQI asking participants about their baseline sleep schedule. From the outset of the study, the intervention and comparison groups were comparable on practically all measured variables. The only statistically significant difference between groups was the percentage of those handling firearms who reported using a rifle. All the participants in the intervention group who reported handling firearms used a rifle, while slightly more than half of those in the comparison group did so. There was also a tendency for participants in the intervention group to have a higher body mass index than those in the comparison group, but this difference was not statistically significant. Likewise, there was a tendency for a greater proportion of participants in the intervention group to be in the National Guard/Reserves as compared to the comparison group, but this difference was also not statistically significant.

Table 6. Summary of intervention and comparison study groups for actigraphy subsample at outset of study.

Variable	Group		*p*-value
	Intervention (*n* = 53, 25%)	Comparison (*n* = 41, 22%)	
Age (yrs), median (IQR)	19 (18-23)	20 (18-24)	0.320^{M}
Body mass index ($kg \cdot m^{-2}$), median (IQR)	25.1 (22.2-27.8)	23.1 (21.4-26.0)	0.074^{M}
Body mass index category, no. (%)			
Underweight	1 (1.9)	1 (2.4)	0.232^{V}
Normal	24 (45.3)	27 (65.9)	
Overweight	18 (34.0)	9 (22.0)	
Obese	10 (18.9)	4 (9.8)	
Caffeine			
Consume caffeinated beverages, no. (%)	35 (66.0)	20 (48.8)	0.092^{C}
Caffeine use ($mg \cdot d^{-1}$), median (IQR)	164 (108-288)	144 (72-305)	0.327^{M}
Component, no. (%)			
National Guard/Reserve	30 (56.6)	16 (39.0)	0.091^{C}
Regular	23 (43.4)	25 (61.0)	
Epworth Sleepiness Scale			
Total score, mean (SD)	7.9 (3.2)	7.4 (3.5)	0.473^{T}
Excessive fatigue (score > 10), no. (%)	9 (17.0)	7 (17.1)	0.991^{C}
Exercise frequency ($hrs \cdot wk^{-1}$), median (IQR)	2.0 (1.4-4.2)	3.0 (1.4-6)	0.226^{M}

Firearms			
Regularly use firearm, no. (%)	11 (20.8)	7 (17.1)	0.653[C]
Type of firearm, no. (%)			
Rifle	11 (100)	4 (57.1)	0.043[F]*
Handgun	4 (36.4)	4 (57.1)	0.630[F]
Use of firearm, no. (%)			
Hunting	7 (63.6)	3 (42.9)	0.630[F]
Sport shooting	8 (72.7)	4 (57.1)	0.627[F]
Other	0 (0)	2(28.6)	0.137[F]
Frequency of use (days·yr^{-1}), median (IQR)	30 (20-45)	45 (25-50)	0.340[M]

*Significant at ≤ 0.05 level.
[C]Chi square statistic, [F]Fisher's Exact Test, [M]Mann-Whitney U, [T]Student's *t*-test, [V]Cramer's V.
Notes: IQR = interquartile range; SD = standard deviation.

Table 7. Summary of intervention and comparison study groups for actigraphy subsample at outset of study (continued).

Variable	Group		*p*-value
	Intervention (*n* = 53, 25%)	Comparison (*n* = 41, 22%)	
GT score, median (IQR)	108 (96-116)	110 (99-121)	0.354[M]
Morningness-Eveningness Questionnaire			
Total score, mean (SD)	50.6 (8.9)	47.2 (9.7)	0.086[T]
Chronotype, no (%)			
Evening type	11 (20.8)	15 (36.6)	0.226[C]
Neither type	31 (58.5)	20 (48.8)	
Morning type	11 (20.8)	6 (14.6)	
NEO Five Factor Inventory			
Neuroticism, median (IQR)	52 (44-56)	51 (46-63)	0.706[M]
Extraversion, mean (SD)	53.5 (11.5)	54.1 (9.0)	0.786[T]
Openness to experience, mean (SD)	50.7 (12.6)	49.7 (11.1)	0.683[T]
Agreeableness, mean (SD)	45.4 (11.4)	43.7 (11.4)	0.495[T]
Conscientiousness, median (IQR)	46 (42-59)	48 (41-53)	0.359[M]
Pittsburgh Sleep Quality Index			
Global score, mean (SD)	6.3 (2.5)	6.71 (2.8)	0.468[T]
Poor sleep quality (score > 5), no. (%)	32 (60.4)	28 (68.3)	0.428[C]
Component scores, median (IQR)			
Subjective sleep quality	1 (1-1)	1 (1-2)	0.147[M]
Sleep latency	2 (1-4)	2 (1-3)	0.745[M]
Sleep duration	0 (0-1)	0 (0-1)	0.504[M]

Habitual sleep efficiency	0 (0-0)	0 (0-0)	0.211[M]
Sleep disturbances	1 (1-2)	1 (1-2)	0.114[M]
Use of sleeping medication	0 (0-0)	0 (0-0)	0.699[M]
Daytime dysfunction	1 (0-1)	1 (0-1)	0.378[M]
Rank, no (%)			
E01	18 (34.0)	16 (39.0)	0.759[C]
E02	20 (37.7)	12 (29.3)	
E03	12 (22.6)	9 (22.0)	
E04	3 (5.7)	4 (9.8)	

[C]Chi square statistic, [M]Mann-Whitney U, [T]Student's *t*-test.
Notes: IQR = interquartile range; SD = standard deviation.

Table 8. Summary of intervention and comparison study groups for actigraphy subsample at outset of study (continued).

Variable	Group		*p*-value
	Intervention ($n = 53$, 25%)	Comparison ($n = 41$, 22%)	
Response to Stressful Experiences Scale			
Global score, mean (SD)	68.3 (12.0)	65.1 (13.0)	0.233[T]
Factor scores, median (IQR)			
Positive appraisal	7.6 (6.1-8.3)	6.8 (6.2-8.5)	0.819[M]
Spirituality	2.9 (2.9-3.8)	2.9 (2.9-3.8)	0.716[M]
Active coping	8.7 (10.2-11.9)	10.2 (8.4-11.5)	0.778[M]
Self-efficacy	3.2 (2.4-3.2)	2.4 (2.4-3.2)	0.778[M]
Learning and meaning-making	7.2 (5.0-8.0)	6.5 (5.4-8.3)	0.310[M]
Acceptance of limitations	4.3 (3.5-5.6)	4.3 (3.5-5.6)	0.816[M]
Sex, no. (%)			
Female	20 (37.7)	15 (36.6)	0.909[C]
Male	33 (62.3)	26 (63.4)	
Tobacco			
Regularly use tobacco, no (%)	22 (41.5)	15 (36.6)	0.628[C]
Frequency of use ($cigs \cdot wk^{-1}$), median (IQR)	49 (19-101)	35 (8-105)	0.577[M]

[C]Chi square statistic, [M]Mann-Whitney U, [T]Student's *t*-test.
Notes: IQR = interquartile range; SD = standard deviation.

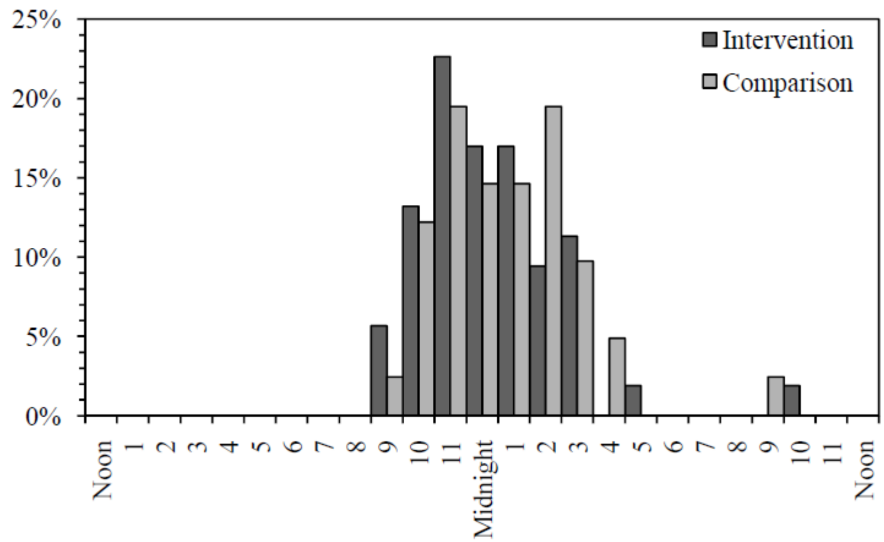


Figure 6. Histogram of participants' reported usual bed time (PSQI question 1) in actigraphy subsample.

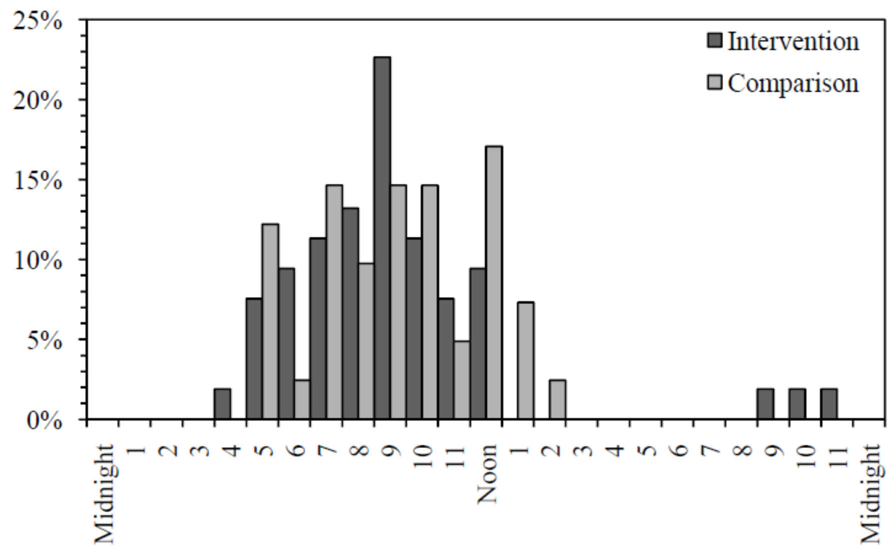


Figure 7. Histogram of participants' reported usual wakeup time (PSQI question 3) in actigraphy subsample.

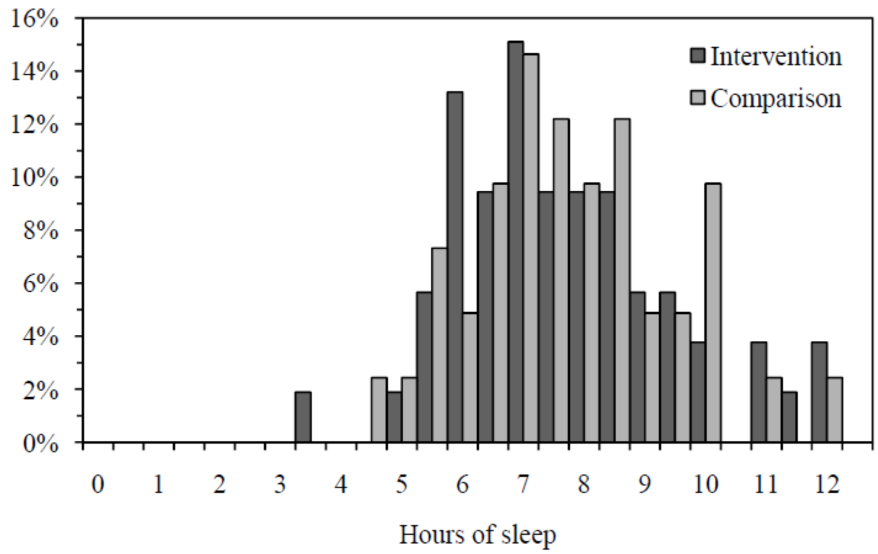


Figure 8. Histogram of participants' reported hours of sleep per night (PSQI question 4) in actigraphy subsample.

2. Total Sleep Time

Figure 9 shows the distribution and the parameters obtained from the distribution for daily total sleep obtained per night for all sleep observations gathered during Basic Combat Training according to treatment condition. The spike at three hours in both histograms was believed to be attributable to participants performing night watch duties. The median total sleep obtained per night across all weeks of Basic Combat Training was significantly greater for participants in the intervention versus comparison group (intervention group mean rank = 2,884.0; comparison group mean rank = 2,105.9; $p < 0.001$ based on the Mann-Whitney U test). The National Sleep Foundation (NSF) recommends that adults obtain 7-9 hours of sleep per night. In this study, 15.5% of sleep observations in the intervention group satisfied the NSF recommendation versus only 4.6% in the comparison group—a significant difference ($\chi_1^2 = 152.282$, $p < 0.001$). Restated, the likelihood or odds of an episode of total daily sleep being less than the NSF's recommendation was 3.802 (95% CI: 3.037, 4.761) for the comparison group relative to the intervention group—i.e., the comparison group was nearly four times as likely to be sleep deficient.

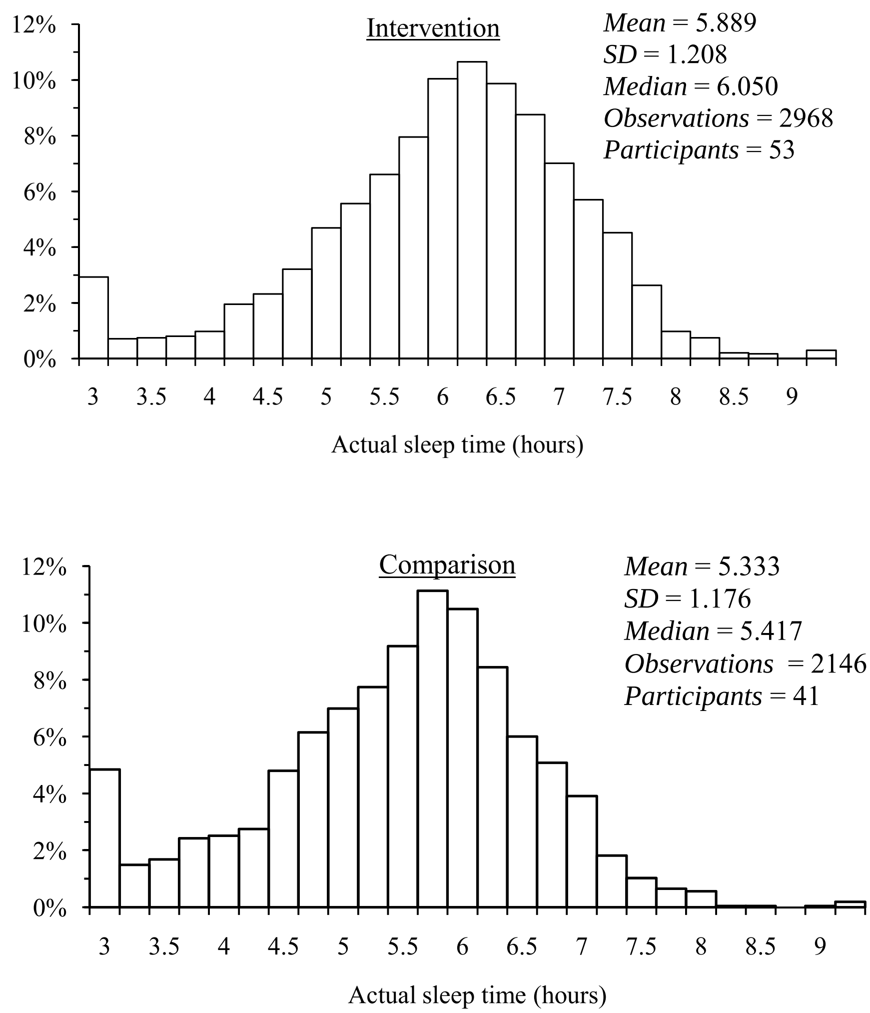


Figure 9. Histograms of total sleep obtained at night for all sleep observations gathered during Basic Combat Training according to treatment condition.

We examined how daily total sleep related to the treatment condition over the course of Basic Combat Training, while accounting for potential covariates and the aforementioned differences between the study groups. However, any approach to analyzing total sleep time needed to address the issue that participants did not necessarily have valid Actiware® scores for every day of Basic Combat Training. This issue was remedied by first computing a weekly average sleep for each participant and then analyzing the dataset as a repeated cross-section design, rather than a within-participant repeated measures design. A 1% significance level (or alpha of 0.01) was also used to counter the resulting increased power of statistical tests. Accordingly, an ANCOVA of weekly average sleep was accomplished using treatment condition, week, and chronotype as fixed effects. Age, caffeine and tobacco use, component, firearm use, fitness factors (body mass index (BMI) and exercise frequency), GT score, personality component scores (NEO-FFI neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness scores), resilience (RSES score), sex, and sleep factors (ESS and PSQI scores) were covariates.

Table 9 provides the results of the univariate analysis of weekly average sleep. There was a significant fixed effect for treatment condition with an estimated marginal mean sleep for the intervention group of 5.876 (99% CI: 5.806, 5.945) versus 5.359 (99% CI: 5.276, 5.442) for the comparison group. That is, controlling for other variables, the intervention group obtained 31 minutes more sleep than the comparison group.

Table 9. Univariate tests for weekly average sleep.

Source	MS	df	*F*	*p*	η^2
Condition	32.384	1	140.162	<0.001*	0.163
Week	15.138	8	65.518	<0.001*	0.422
Chronotype	2.383	2	10.312	<0.001*	0.028
Condition x Week	2.555	8	11.059	<0.001*	0.110
Condition x Chronotype	0.323	2	1.399	0.247	0.004
Chronotype x Week	0.321	16	1.390	0.140	0.030
Condition x Chronotype x Week	0.116	16	0.502	0.947	0.011
Age	2.569	1	11.118	0.001*	0.015
Body mass index	1.476	1	6.390	0.012	0.009
Caffeine use (referent no)	2.490	1	10.779	0.001*	0.015
Component (referent regular)	0.232	1	1.004	0.317	0.001
Epworth Sleepiness Scale	2.491	1	10.781	0.001*	0.015
Exercise frequency	1.860	1	8.052	0.005*	0.011
Firearm use (referent no)	0.301	1	1.301	0.254	0.002
GT score	0.438	1	1.895	0.169	0.003
NEO-FFI					
Neuroticism	0.541	1	2.341	0.126	0.003

Extraversion	0.926	1	4.006	0.046	0.006
Openness to experience	0.090	1	0.387	0.534	0.001
Agreeableness	0.052	1	0.224	0.636	<0.001
Conscientiousness	0.937	1	4.055	0.044	0.006
Pittsburgh Sleep Quality Index	0.357	1	1.545	0.214	0.002
RSES	0.307	1	1.327	0.250	0.002
Sex (referent male)	2.376	1	10.285	0.001*	0.014
Tobacco use (referent no)	0.125	1	0.539	0.463	0.001
Error	0.231	718			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

There was also a significant fixed effect for week (Figure 10), with pairwise differences occurring between week 1 versus weeks 6-9 ($p < 0.001$); week 2 versus weeks 6-9 ($p \leq 0.002$); week 3 versus week 6 and weeks 8-9 ($p < 0.001$); week 4 versus week 6 and weeks 8-9 ($p < 0.001$); week 5 versus weeks 6-9 ($p \leq 0.004$); week 6 versus week 7 ($p < 0.001$); and week 7 versus weeks 8-9 ($p < 0.001$).

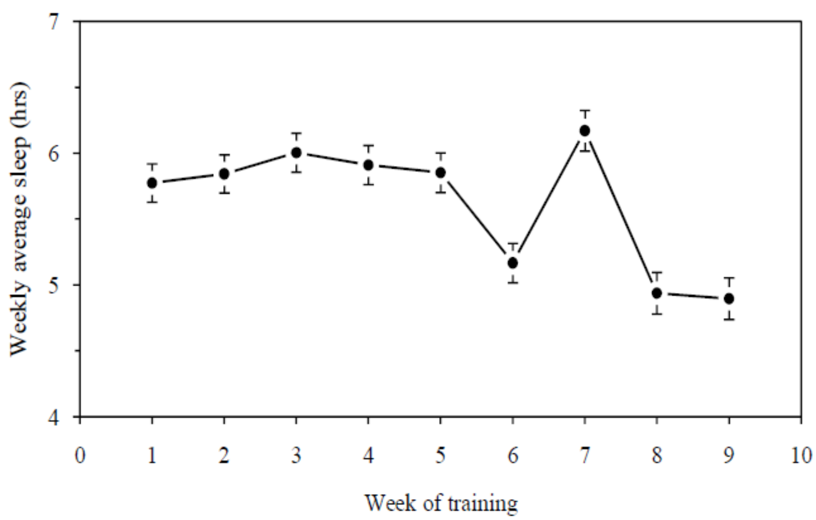


Figure 10. Estimated marginal means for sleep by week of training (error bars are for 99% confidence intervals).

For the significant fixed effect of chronotype (Figure 11), the pairwise differences occurred between morning chronotype versus both evening and indeterminate chronotypes ($p \leq 0.001$).

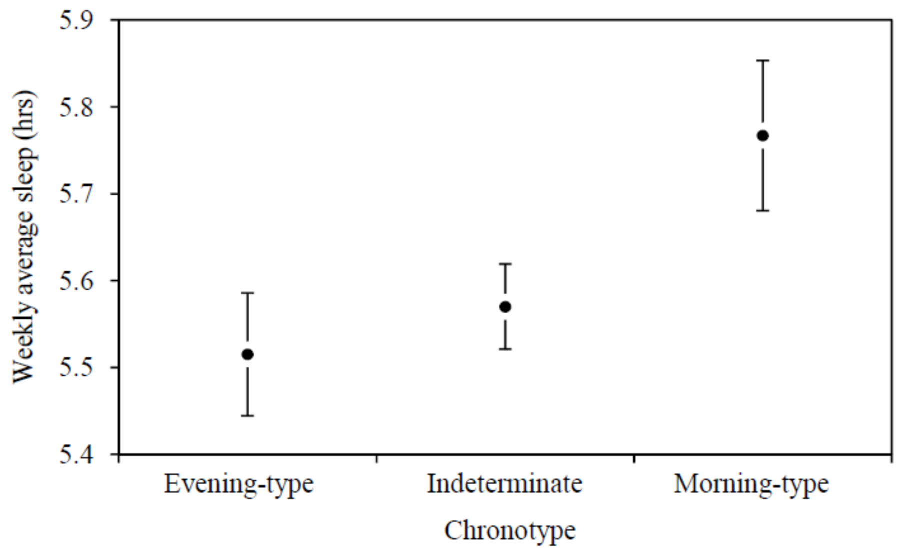


Figure 11. Estimated marginal means for sleep by chronotype (error bars are for 99% confidence intervals).

Additionally, there was a significant interaction effect between treatment condition and week (Figure 12), with participants in the intervention group getting more sleep than those in the comparison group during the first six weeks of training. During the latter three weeks of training, participants in the intervention group got notably less sleep, such that there was no longer a difference between the intervention and comparison groups. This observation was attributed to the field exercises that were conducted throughout the last three weeks of training, during which participants moved from the barracks to an encampment. There was no interaction effect between treatment condition and chronotype or between chronotype and week. Significant covariates included age, caffeine use, ESS score, exercise frequency, and sex.

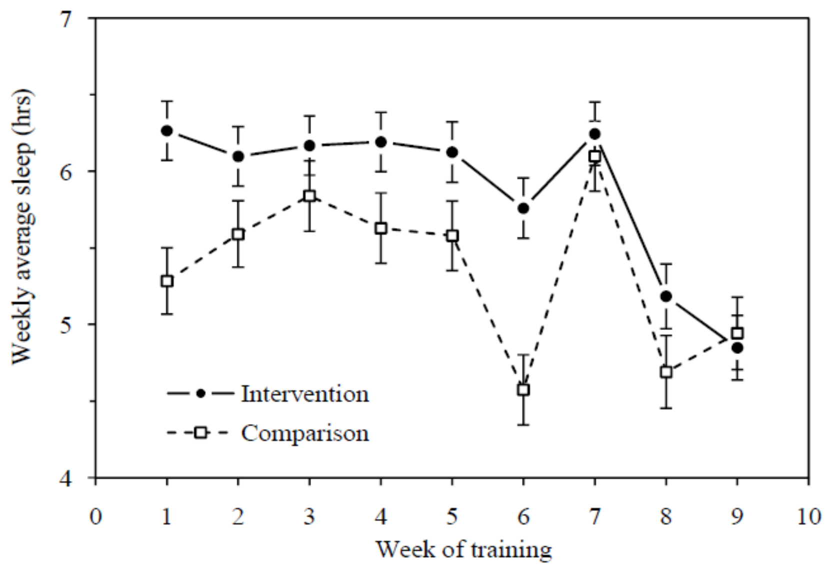
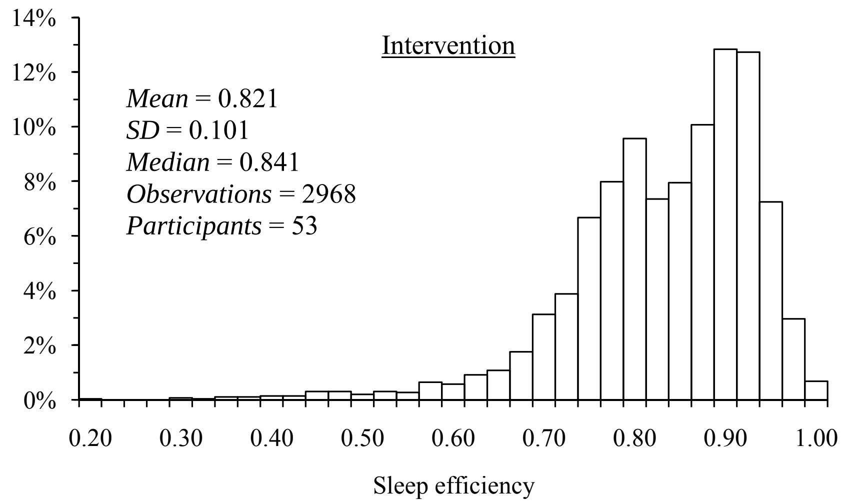


Figure 12. Estimated marginal means for sleep by treatment condition and week of training (error bars are for 99% confidence intervals).

3. Sleep Efficiency

Sleep efficiency was calculated as the ratio of a participant's total sleep time to total time in bed; it represents the proportion of time that a participant was assumed to be "in bed" or attempting sleep that was actually spent asleep (Paquet, Kawinska, & Carrier, 2007). Figure 13 shows the distribution and distributional parameters for sleep efficiency for all sleep observations gathered during Basic Combat Training according to treatment condition.



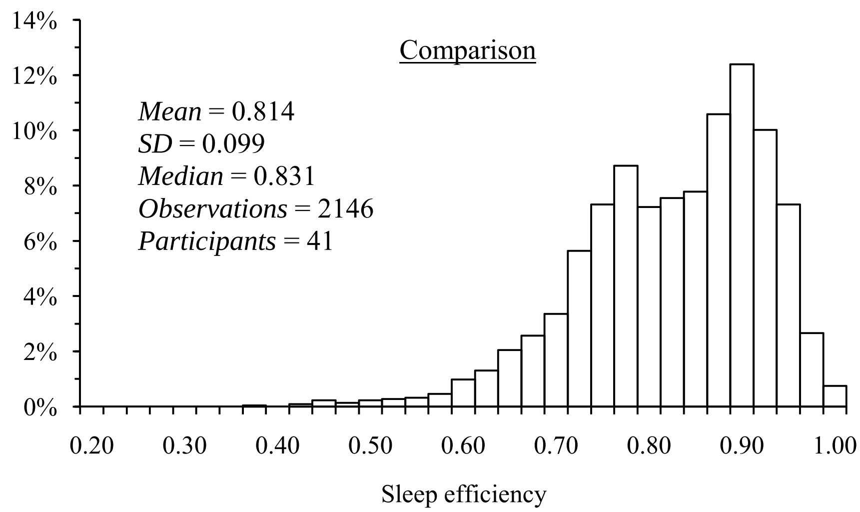


Figure 13. Histograms of sleep efficiency for all sleep observations by treatment condition.

The median sleep efficiency across all weeks of Basic Combat Training was significantly greater for participants in the intervention vice comparison study group (intervention group mean rank = 2,614.3; comparison group mean rank = 2,479.0; $p < 0.001$ based on the Mann-Whitney U test). Nevertheless, the practical significance of a difference in median sleep efficiency of 0.010 is questionable. However, the histograms suggest that the distributions of sleep efficiency for the two groups differed slightly. This impression was investigated further by estimating the population moments using the sample k^{th} moments (Table 10). While the 95% confidence intervals overlapped for the first and second moments, there was a significant difference in the third and fourth moments, which are functions of the distributions' skewness (i.e., symmetry) and kurtosis (i.e., peakedness), respectively.

Table 10. Population moment estimates based on sample k^{th} moments.

	Intervention Group		Comparison Group	
k^{th} Moment	Estimate	95% CI	Estimate	95% CI
First	0.821	(0.817, 0.825)	0.814	(0.810, 0.818)
Second	0.684	(0.678, 0.690)	0.672	(0.666, 0.679)
Third	0.577	(0.571, 0.584)	0.562	(0.555, 0.570)
Fourth	0.492	(0.485, 0.499)	0.476	(0.467, 0.484)

Note: CI = confidence interval.

4. Activity Counts During Sleep

Activity counts reflect movements during sleep and may be a function of the stage of sleep (Monk, Buysse, & Rose, 1999). Figure 14 shows the distribution and distributional parameters for mean activity counts for all sleep observations gathered during Basic Combat Training according to treatment condition. The median activity count during sleep across all weeks of Basic Combat Training was significantly less for participants in the intervention versus comparison study group (intervention group mean rank = 2,504.8; comparison group mean rank = 2,630.4; $p < 0.001$ based the on Mann-Whitney test). However, the histograms appear quite similar; as in the analysis of the sleep efficiency data, population moments were estimated for each distribution using the k^{th} sample moments. It was found that the 95% confidence intervals overlapped for the first four moments of each sample distribution, thereby suggesting that the observed distributions do not significantly differ.

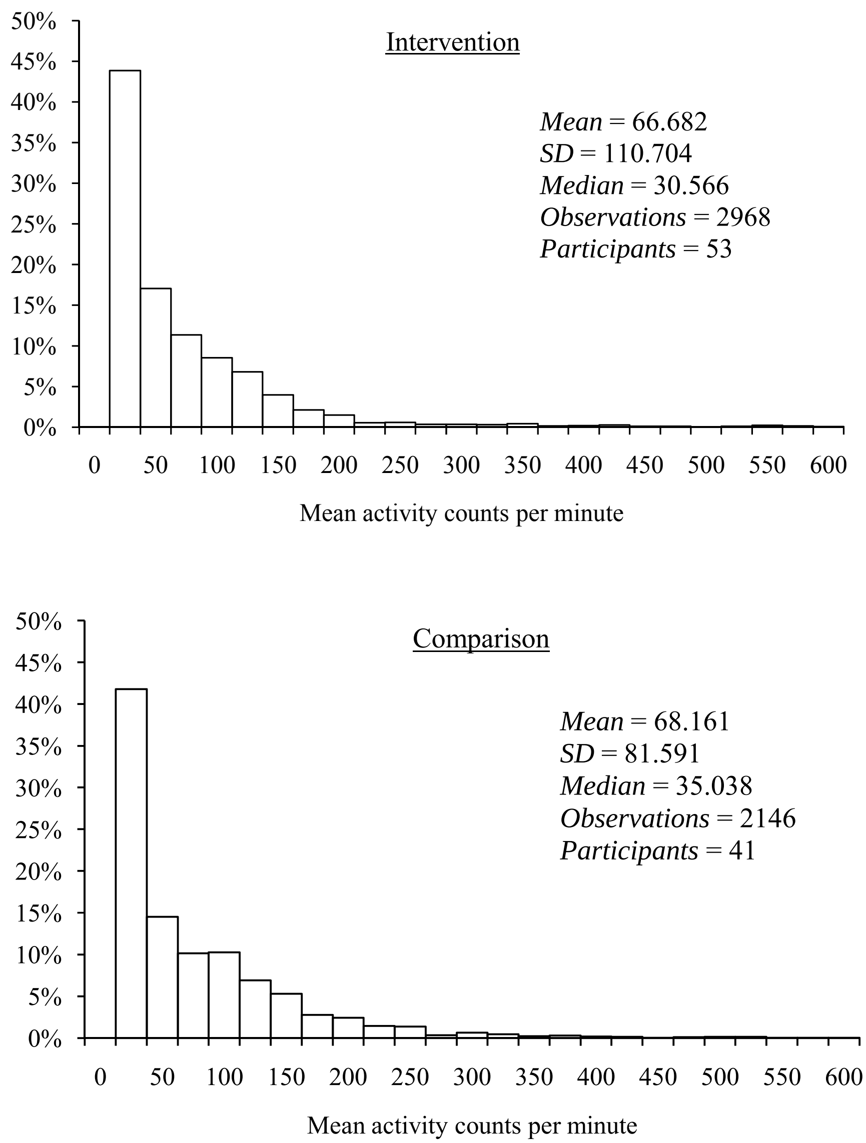


Figure 14. Histograms of mean activity counts for all sleep observations by treatment condition.

C. PROFILE OF MOOD STATES (POMS)

The study examined how POMS factor scores related to the treatment condition over the course of Basic Combat Training, while accounting for potential covariates and the known differences between the study groups. However, any approach to modeling the POMS factor scores needed to address several issues. First, a MANCOVA of the prestudy POMS factor scores with treatment condition as a fixed effect and age, caffeine and tobacco use, component, GT score, firearm use, fitness factors (BMI and exercise frequency), NEO personality component scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates found a significant effect for treatment condition (Wilks' $\lambda = 0.769$, $F_{6,367} = 18.393$, $p < 0.001$). An examination of the univariate ANCOVAs showed that there were significant fixed effects for treatment condition on T-factor (tension-anxiety) scores ($F_{1,372} = 42.094$, $p < 0.001$), D-factor (depression-dejection) scores ($F_{1,372} = 30.305$, $p < 0.001$), A-factor (anger-hostility) scores ($F_{1,372} = 39.278$, $p < 0.001$), V-factor (vigor-activity) scores ($F_{1,372} = 6.961$, $p = 0.009$), F-factor (fatigue-inertia) scores ($F_{1,372} = 100.803$, $p < 0.001$), and C-factor (confusion-bewilderment) scores ($F_{1,372} = 22.397$, $p < 0.001$). It was clearly observed from Figure 15 that the pre-study POMS factor scores, prior to any exposure to the treatment, differed between the study groups.

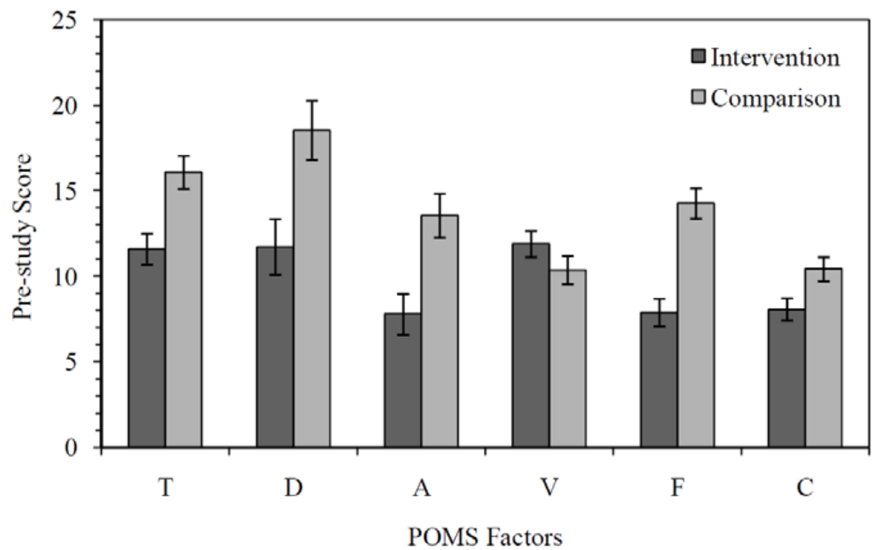


Figure 15. Comparison of estimated marginal means and associated 95% confidence intervals for prestudy POMS factor scores by study group.

These results suggested that the two study groups were not directly comparable at baseline in terms of subjective mood. This issue was remedied by calculating the "delta from baseline" score for each factor—that is, subtracting a participant's prestudy POMS factor score from all their subsequent POMS factor scores. This subtraction had the effect of making all participants' prestudy POMS factor scores zero, while still preserving the magnitude and directionality of variations in their subsequent POMS factor scores. Another issue was the observation that most participants (70.4%) did not have a POMS questionnaire for every week of training. This issue was addressed by analyzing the POMS dataset as a repeated cross-section design rather than a within-participant repeated measures design and using a 1% significance level to counter the resulting increased power of statistical tests.

A MANCOVA of the POMS factor delta from baseline scores was accomplished using treatment condition, week, and chronotype as fixed effects and age, caffeine and tobacco use, component, firearm use, fitness factors (BMI and exercise frequency), GT score, NEO personality component scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates. Table 11 summarizes the results of the multivariate tests. There were significant fixed effects for treatment condition, week, and chronotype as well as significant interaction effects between treatment condition and both week and chronotype. With the exception of exercise frequency, firearm use, NEO extraversion component score, and RSES score, there were significant effects for all the measured covariates.

Table 11. Multivariate tests for POMS delta from baseline scores.

Source	Wilks' λ	*F*	df1	df2	*p*	η^2
Condition	0.992	4.261	6	3037	<0.001*	0.008
Week	0.944	3.694	48	14947	<0.001*	0.010
Chronotype	0.984	4.217	12	6074	<0.001*	0.008
Condition x Week	0.974	1.673	48	14947	0.002*	0.004
Condition x Chronotype	0.990	2.628	12	6074	0.002*	0.005
Chronotype x Week	0.985	0.466	96	17213	1.000	0.002
Condition x Chronotype x Week	0.981	0.617	96	17213	0.999	0.003
Age	0.967	17.008	6	3037	<0.001*	0.033
Body mass index	0.980	10.084	6	3037	<0.001*	0.020
Caffeine use (referent no)	0.981	9.842	6	3037	<0.001*	0.019
Component (referent regular)	0.989	5.812	6	3037	<0.001*	0.011
Epworth Sleepiness Scale	0.956	23.510	6	3037	<0.001*	0.044
Exercise frequency	0.995	2.628	6	3037	0.015	0.005
Firearm use (referent no)	0.996	1.951	6	3037	0.069	0.004
GT score	0.968	16.607	6	3037	<0.001*	0.032

NEO-FFI						
Neuroticism	0.966	17.934	6	3037	<0.001*	0.034
Extraversion	0.995	2.318	6	3037	0.031	0.005
Openness to experience	0.985	7.631	6	3037	<0.001*	0.015
Agreeableness	0.973	14.192	6	3037	<0.001*	0.027
Conscientiousness	0.982	9.075	6	3037	<0.001*	0.018
Pittsburgh Sleep Quality Index	0.984	8.108	6	3037	<0.001*	0.016
RSES	0.995	2.583	6	3037	0.017	0.005
Sex (referent male)	0.973	13.883	6	3037	<0.001*	0.027
Tobacco use (referent no)	0.988	6.158	6	3037	<0.001*	0.012

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

1. Tension-Anxiety (T) Factor

Table 12 provides the results of the relevant univariate tests of between-participant effects for the POMS T-factor delta from baseline scores. There was no significant fixed effect for treatment condition or chronotype.

Table 12. Univariate tests of between-participant effects for POMS T-factor delta from baseline scores.

Source	MS	df	*F*	*p*	η^2
Condition	60.636	1	1.359	0.244	<0.001
Week	335.619	8	7.521	<0.001*	0.019
Chronotype	31.538	2	0.707	0.493	<0.001
Condition x Week	78.945	8	1.769	0.078	0.005
Condition x Chronotype	49.363	2	1.106	0.331	0.001
Age	555.040	1	12.439	<0.001*	0.004
Body mass index	1243.017	1	27.857	<0.001*	0.009
Caffeine use (referent no)	814.800	1	18.260	<0.001*	0.006
Component (referent regular)	219.848	1	4.927	0.027	0.002
Epworth Sleepiness Scale	124.464	1	2.789	0.095	0.001
GT score	1474.994	1	33.055	<0.001*	0.011
NEO-FFI					
Neuroticism	1379.661	1	30.919	<0.001*	0.010

Openness to experience	80.314	1	1.800	0.180	0.001
Agreeableness	14.529	1	0.326	0.568	<0.001
Conscientiousness	20.671	1	0.463	0.496	<0.001
Pittsburgh Sleep Quality Index	762.339	1	17.084	<0.001*	0.006
Sex (referent male)	298.227	1	6.683	0.010*	0.002
Tobacco use (referent no)	706.302	1	15.829	<0.001*	0.005
Error	44.622	3042			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

There was a significant fixed effect for week (Figure 16), with the main pairwise differences occurring between week 1 versus weeks 4, 5, 6,7 and 9 ($p \leq 0.001$) and between week 3 versus week 6 ($p = 0.006$).

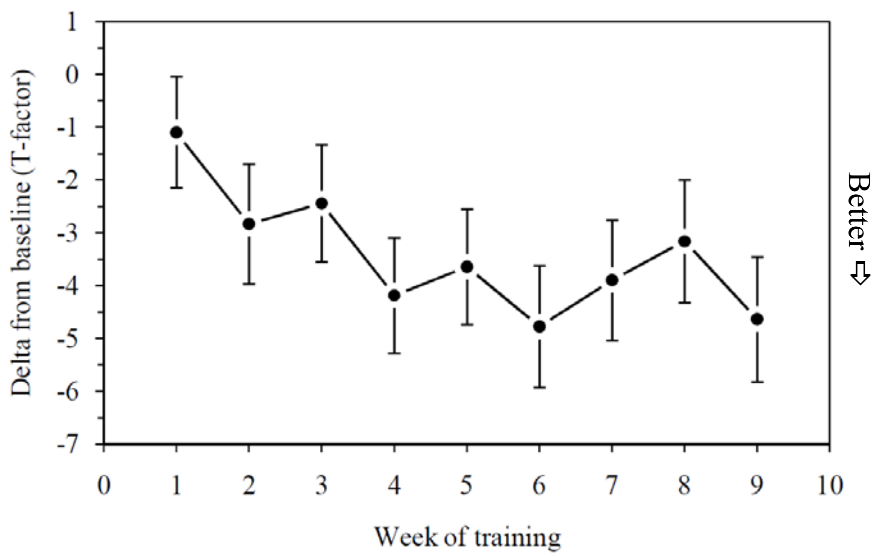


Figure 16. Estimated marginal means for POMS T-factor delta from baseline scores by week of training (error bars are for 99% confidence intervals).

There was no significant interaction effect between treatment condition and either week or chronotype. Thus, the general trend was for T-factor scores to decrease during the first six weeks of training followed by a spike in T-factor scores during weeks 7 and 8. Significant covariates included age, BMI, caffeine and tobacco use, GT score, NEO neuroticism component score, PSQI score, and sex, but only GT score and neuroticism had effect sizes of at least 1% as measured using eta squared.

2. Depression-Dejection (D) Factor

Table 13 provides the results of the univariate tests of between-participant effects for the POMS D-factor delta from baseline scores. Again, there was no significant fixed effect for treatment condition or chronotype.

Table 13. Univariate tests of between-participant effects for POMS D-factor delta from baseline scores.

Source	MS	df	*F*	*p*	η^2
Condition	132.618	1	0.989	0.320	<0.001
Week	1208.472	8	9.015	<0.001*	0.023
Chronotype	299.645	2	2.235	0.107	0.001
Condition x Week	158.458	8	1.182	0.306	0.003
Condition x Chronotype	245.889	2	1.834	0.160	0.001
Age	1014.065	1	7.565	0.006*	0.002
Body mass index	5334.391	1	39.793	<0.001*	0.013
Caffeine use (referent no)	2135.415	1	15.930	<0.001*	0.005
Component (referent regular)	146.044	1	1.089	0.297	<0.001
Epworth Sleepiness Scale	0.044	1	0.000	0.985	<0.001
GT score	856.795	1	6.391	0.012	0.002
NEO-FFI					
Neuroticism	6150.683	1	45.882	<0.001*	0.015
Openness to experience	577.989	1	4.312	0.038	0.001
Agreeableness	2046.344	1	15.265	<0.001*	0.005
Conscientiousness	708.772	1	5.287	0.022	0.002
Pittsburgh Sleep Quality Index	233.218	1	1.740	0.187	0.001
Sex (referent male)	165.777	1	1.237	0.266	<0.001
Tobacco use (referent no)	518.436	1	3.867	0.049	0.001
Error	134.054	3042			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

There was a significant fixed effect for week (Figure 17), with pairwise differences occurring between week 1 versus weeks 4, 5, 6, 7, 8 and 9 ($p \leq 0.002$), week 2 versus week 9 ($p = 0.001$), and week 3 versus week 9 ($p = 0.003$).

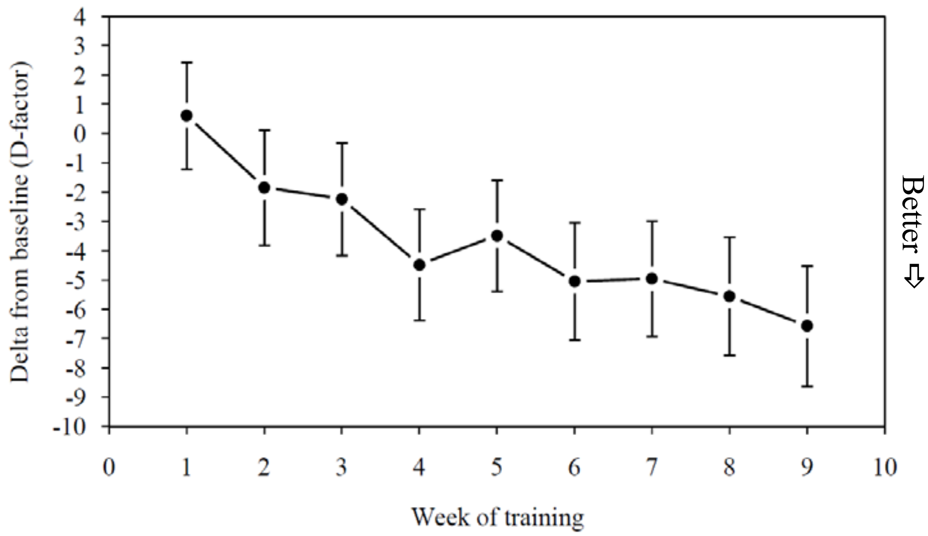


Figure 17. Estimated marginal means for POMS D-factor delta from baseline scores by week of training (error bars are for 99% confidence intervals).

There was no significant interaction effect between treatment condition and either week or chronotype. Thus, the general trend was for D-factor scores to decrease during the course of training, with lower scores meaning less of a depressed mood. Significant covariates included age, BMI, caffeine use, and NEO neuroticism and agreeableness component scores, but only BMI and neuroticism had effect sizes of at least 1%.

3. Anger-Hostility (A) Factor

Table 14 provides the results of the univariate tests of between-participant effects for the POMS A-factor delta from baseline scores. There was no significant fixed effect for treatment condition or chronotype.

Table 14. Univariate tests of between-participant effects for POMS A-factor delta from baseline scores.

Source	MS	df	*F*	*p*	η^2
Condition	5.447	1	0.062	0.803	<0.001
Week	718.227	8	8.172	<0.001*	0.021
Chronotype	118.510	2	1.348	0.260	0.001
Condition x Week	235.186	8	2.676	0.006*	0.007
Condition x Chronotype	200.591	2	2.282	0.102	0.001
Age	1553.745	1	17.679	<0.001*	0.006

Body mass index	1822.769	1	20.740	<0.001*	0.007
Caffeine use (referent no)	538.882	1	6.131	0.013	0.002
Component (referent regular)	38.695	1	0.440	0.507	<0.001
Epworth Sleepiness Scale	34.238	1	0.390	0.533	<0.001
GT score	1301.170	1	14.805	<0.001*	0.005
NEO-FFI					
Neuroticism	176.461	1	2.008	0.157	0.001
Openness to experience	1270.906	1	14.461	<0.001*	0.005
Agreeableness	7.572	1	0.086	0.769	<0.001
Conscientiousness	252.873	1	2.877	0.090	0.001
Pittsburgh Sleep Quality Index	158.508	1	1.804	0.179	0.001
Sex (referent male)	3035.072	1	34.533	<0.001*	0.011
Tobacco use (referent no)	963.306	1	10.961	0.001*	0.004
Error	87.888	3042			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

There was a significant fixed effect for week (Figure 18), with the pairwise differences occurring between week 1 versus week 4, 6, 7, 8 and 9 ($p \leq 0.005$), week 2 versus week 9 ($p = 0.001$), week 3 versus week 9 ($p < 0.001$), and week 5 versus week 9 ($p = 0.002$).

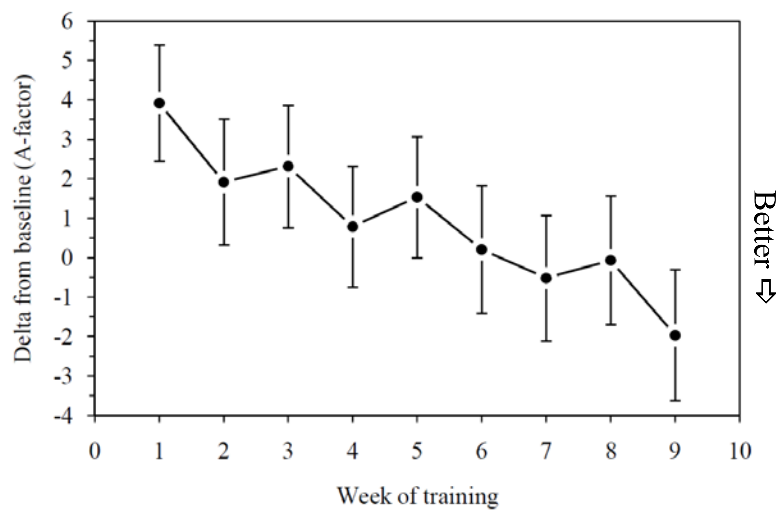


Figure 18. Estimated marginal means for POMS A-factor delta from baseline scores by week of training (error bars are for 99% confidence intervals).

There was a significant interaction effect between treatment condition and week (Figure 19), but not between treatment condition and chronotype.

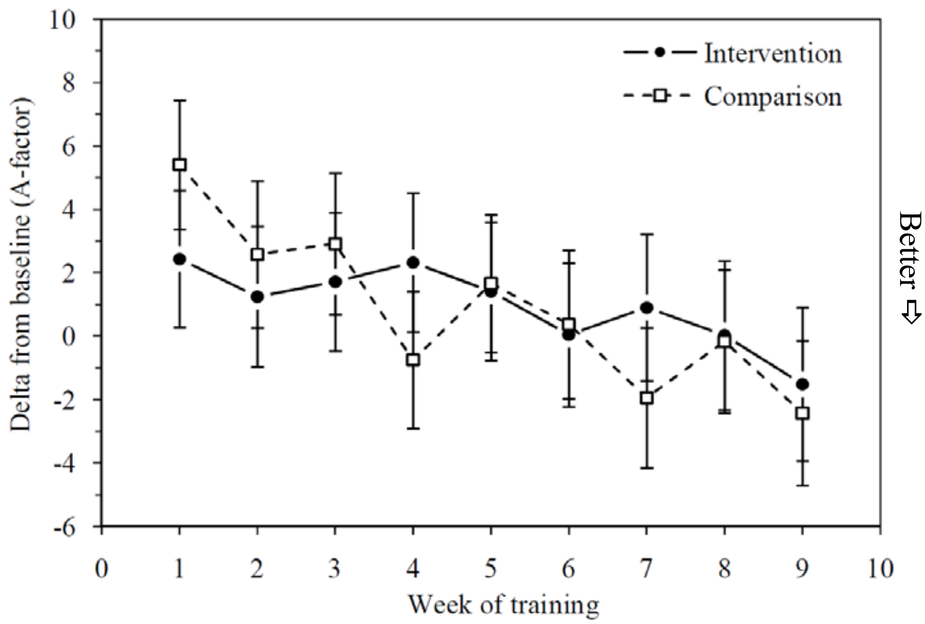


Figure 19. Estimated marginal means for POMS A-factor delta from baseline scores by treatment condition and week of training (error bars are for 99% confidence intervals).

Thus, the comparison group started out with higher A-factor delta from baseline scores, but had a greater rate of decrease in scores over training compared to the intervention group. Significant covariates included age, BMI, GT score, NEO openness to experience component score, sex, and tobacco use, but only sex had an effect size of at least 1%.

4. Vigor-Activity (V) Factor

Table 15 provides the results of the univariate tests of between-participant effects for the POMS V-factor delta from baseline scores. There was a significant fixed effect for treatment condition with an estimated marginal mean score for the intervention group of 1.229 (99% CI: 0.830, 1.628) versus 0.098 (99% CI: –0.347, 0.543) for the comparison group. That is, controlling for other variables, the intervention group exhibited a mood of greater vigorousness and ebullience and higher energy than the comparison group.

Table 15. Univariate tests of between-participant effects for POMS V-factor delta from baseline scores.

Source	MS	df	F	p	η^2
Condition	394.489	1	10.232	0.001*	0.003
Week	17.975	8	0.466	0.881	0.001
Chronotype	574.906	2	14.911	<0.001*	0.010
Condition x Week	78.426	8	2.034	0.039	0.005
Condition x Chronotype	94.740	2	2.457	0.086	0.002
Age	3039.636	1	78.838	<0.001*	0.025
Body mass index	571.114	1	14.813	<0.001*	0.005
Caffeine use (referent no)	377.387	1	9.788	0.002*	0.003
Component (referent regular)	494.366	1	12.822	<0.001*	0.004
Epworth Sleepiness Scale	2844.343	1	73.773	<0.001*	0.024
GT score	1283.601	1	33.292	<0.001*	0.011
NEO-FFI					
Neuroticism	1037.429	1	26.907	<0.001*	0.009
Openness to experience	479.607	1	12.439	<0.001*	0.004
Agreeableness	224.950	1	5.834	0.016	0.002
Conscientiousness	378.944	1	9.829	0.002*	0.003
Pittsburgh Sleep Quality Index	395.210	1	10.250	0.001*	0.003
Sex (referent male)	561.431	1	14.562	<0.001*	0.005
Tobacco use (referent no)	40.373	1	1.047	0.306	<0.001
Error	38.555	3042			

*Significant at ≤ 0.01 level.

Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

There was no significant fixed effect for week, but there was a significant effect for chronotype (Figure 20), with the main pairwise difference occurring between evening and indeterminate chronotypes ($p < 0.001$).

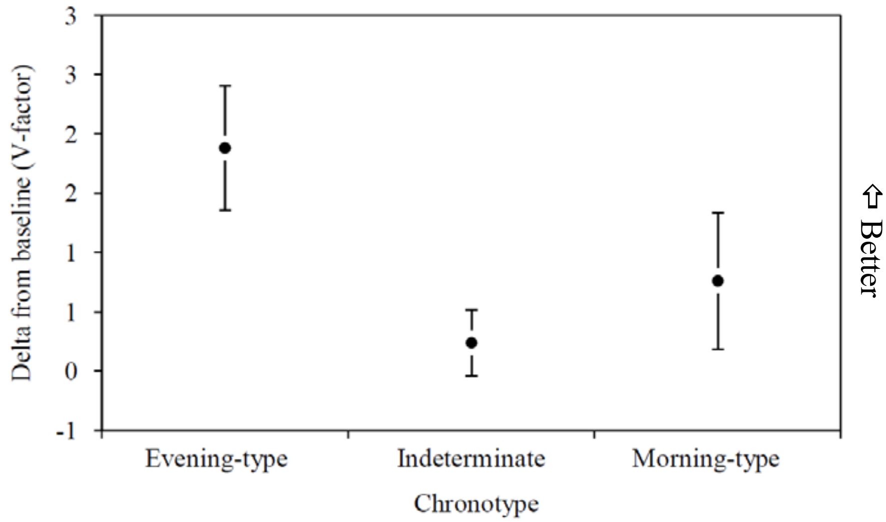


Figure 20. Estimated marginal means for POMS V-factor delta from baseline scores by chronotype (error bars are for 99% confidence intervals).

Significant covariates included age, BMI, caffeine use, component, ESS score, GT score, NEO (neuroticism, openness to experience, and agreeableness component scores), PSQI score, and sex. Only age, ESS score, and GT score had effect sizes of at least 1%.

5. Fatigue-Inertia (F) Factor

Table 16 provides the results of the univariate tests of between-participant effects for the POMS F-factor delta from baseline scores. There were no significant fixed effects of either treatment condition or chronotype. However, there was a significant fixed effect of week as well as a significant interaction effect between treatment condition and week.

Table 16. Univariate tests of between-participant effects for POMS F-factor delta from baseline scores.

Source	MS	df	*F*	*p*	η^2
Condition	119.754	1	3.092	0.079	0.001
Week	401.350	8	10.362	<0.001*	0.027
Chronotype	23.846	2	0.616	0.540	<0.001
Condition x Week	163.341	8	4.217	<0.001*	0.011
Condition x Chronotype	111.529	2	2.880	0.056	0.002
Age	1100.898	1	28.424	<0.001*	0.009
Body mass index	1451.967	1	37.488	<0.001*	0.012

Caffeine use (referent no)	112.819	1	2.913	0.088	0.001
Component (referent regular)	16.907	1	0.437	0.509	<0.001
Epworth Sleepiness Scale	2118.381	1	54.694	<0.001*	0.018
GT score	753.970	1	19.467	<0.001*	0.006
NEO-FFI					
Neuroticism	627.055	1	16.190	<0.001*	0.005
Openness to experience	8.629	1	0.223	0.637	<0.001
Agreeableness	1108.981	1	28.633	<0.001*	0.009
Conscientiousness	899.462	1	23.223	<0.001*	0.008
Pittsburgh Sleep Quality Index	33.364	1	0.861	0.353	<0.001
Sex (referent male)	472.120	1	12.190	<0.001*	0.004
Tobacco use (referent no)	33.269	1	0.859	0.354	<0.001
Error	38.731	3042			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

For the fixed effect of week (Figure 21), the pairwise differences occurred between week 1 versus week 4, 6, 7, 8 and 9 ($p < 0.001$); week 2 versus week 7 ($p = 0.009$); week 3 versus weeks 6, 7, and 9 ($p \leq 0.009$); and week 5 versus weeks 4, 6–7, and 9 ($p \leq 0.005$).

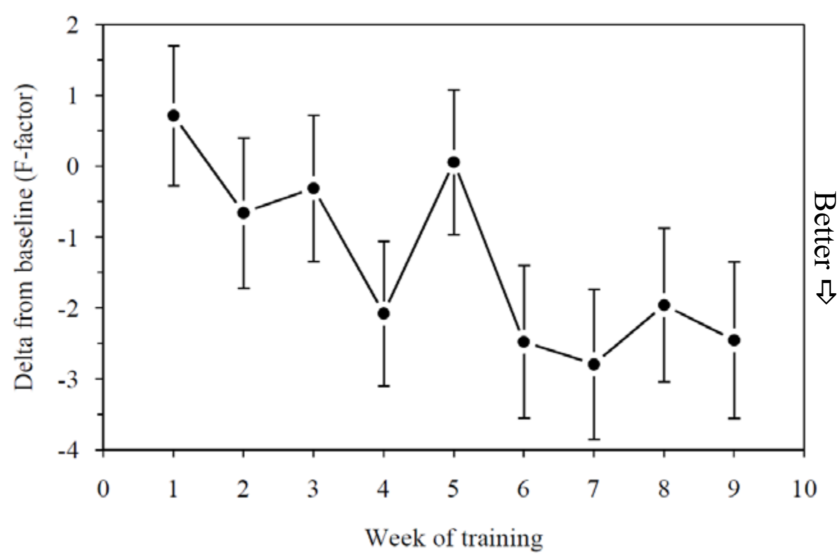


Figure 21. Estimated marginal means for POMS F-factor delta from baseline scores by week of training (error bars are for 99% confidence intervals).

In terms of the significant interaction effect (Figure 22), the comparison group started out with a higher mean F-factor score, but had a greater rate of decrease in scores over training as compared to the intervention group.

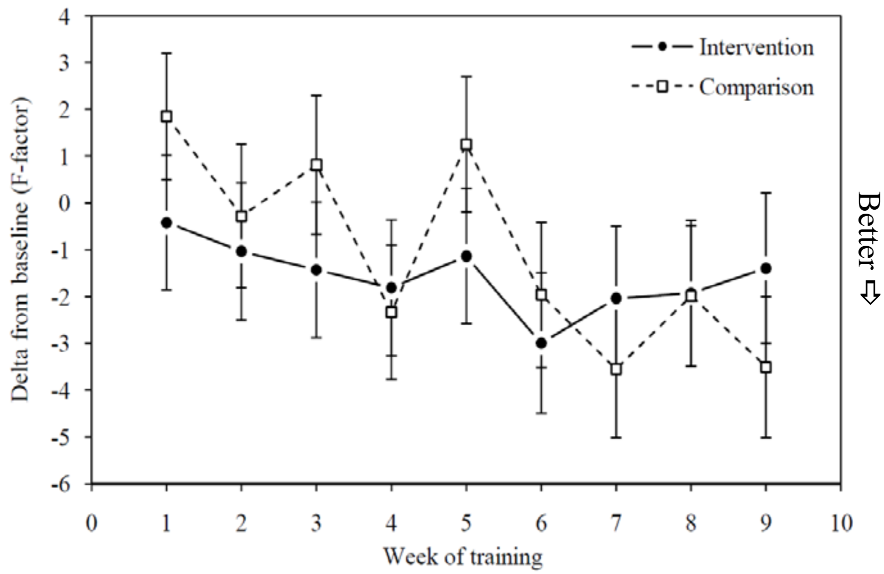


Figure 22. Estimated marginal means for POMS F-factor delta from baseline scores by treatment condition and week of training (error bars are for 99% confidence intervals).

Significant covariates included age, BMI, ESS score, GT score, NEO (neuroticism, agreeableness, and conscientiousness component scores), and sex. Only BMI and ESS score had effect sizes of at least 1%.

6. Confusion-Bewilderment (C) Factor

Table 17 provides the results of the univariate tests of between-participant effects for the POMS C-factor delta from baseline scores. There was no significant fixed effect for treatment condition or chronotype.

Table 17. Univariate tests of between-participant effects for POMS C-factor delta from baseline scores.

Source	MS	df	F	p	η^2
Condition	26.964	1	1.117	0.291	<0.001
Week	274.662	8	11.383	<0.001*	0.029
Chronotype	5.940	2	0.246	0.782	<0.001
Condition x Week	27.565	8	1.142	0.331	0.003

Condition x Chronotype	27.612	2	1.144	0.319	0.001
Age	30.062	1	1.246	0.264	<0.001
Body mass index	790.474	1	32.760	<0.001*	0.011
Caffeine use (referent no)	38.958	1	1.615	0.204	0.001
Component (referent regular)	274.152	1	11.362	0.001*	0.004
Epworth Sleepiness Scale	248.181	1	10.286	0.001*	0.003
GT score	72.149	1	2.990	0.084	0.001
NEO-FFI					
Neuroticism	181.822	1	7.535	0.006*	0.002
Openness to experience	2.737	1	0.113	0.736	<0.001
Agreeableness	92.860	1	3.848	0.050	0.001
Conscientiousness	286.123	1	11.858	0.001*	0.004
Pittsburgh Sleep Quality Index	449.225	1	18.618	<0.001*	0.006
Sex (referent male)	57.315	1	2.375	0.123	0.001
Tobacco use (referent no)	446.382	1	18.500	<0.001*	0.006
Error	24.129	3042			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

There was a significant fixed effect for week (Figure 23), with pairwise differences occurring between week 1 versus weeks 3 through 9 ($p < 0.006$) and week 2 versus weeks 6 through 9 ($p \leq 0.005$).

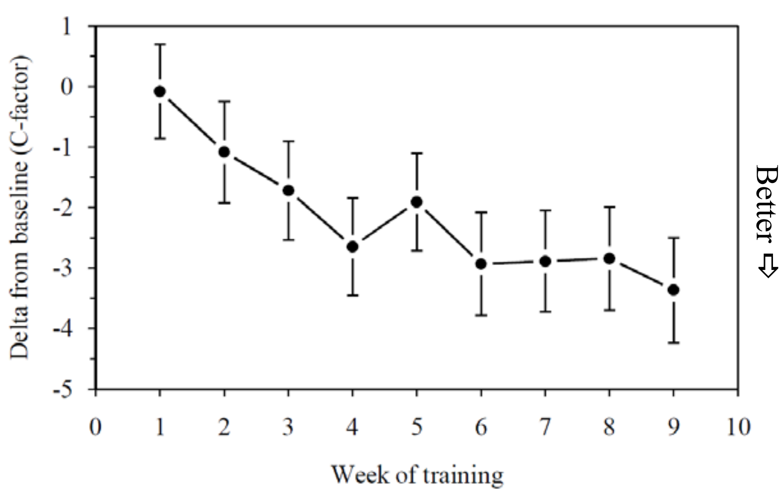


Figure 23. Estimated marginal means for POMS C-factor delta from baseline scores by week of training (error bars are for 99% confidence intervals).

There was no significant interaction effect between treatment condition and either week or chronotype. Thus, the trend was for C-factor scores to decrease during the course of training. Significant covariates included BMI, component, ESS score, NEO neuroticism and conscientiousness component scores, PSQI score, and tobacco use, but only BMI had an effect size of at least 1%.

7. Total Mood Disturbance (TMD) Score

A TMD score was obtained from the POMS by simply summing the scores across all six factors, while negatively weighting vigor. Accordingly, the TMD score provides a single global estimate of affective state (McNair et al., 1981). An ANCOVA of TMD delta from baseline scores was accomplished using treatment condition, week, and chronotype as fixed effects and age, caffeine and tobacco use, component, firearm use, fitness factors (BMI and exercise frequency), GT score, personality component scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates (Table 18).

Table 18. Univariate tests for Total Mood Disturbance delta from baseline scores.

Source	MS	df	*F*	*p*	η^2
Condition	253.538	1	0.221	0.638	<0.001
Week	12915.545	8	11.276	<0.001*	0.029
Chronotype	1400.551	2	1.223	0.295	0.001
Condition x Week	3306.386	8	2.887	0.003*	0.008
Condition x Chronotype	2040.045	2	1.781	0.169	0.001
Chronotype x Week	839.027	16	0.733	0.763	0.004
Condition x Chronotype x Week	1137.775	16	0.993	0.461	0.005
Age	36498.019	1	31.865	<0.001*	0.010
Body mass index	58619.151	1	51.178	<0.001*	0.017
Caffeine use (referent no)	5566.435	1	4.860	0.028	0.002
Component (referent regular)	153.641	1	0.134	0.714	<0.001
Epworth Sleepiness Scale	17536.589	1	15.311	<0.001*	0.005
Exercise frequency	2809.579	1	2.453	0.117	0.001
Firearm use (referent no)	557.135	1	0.486	0.486	<0.001
GT score	10973.626	1	9.581	0.002*	0.003
NEO-FFI					

Neuroticism	40202.835	1	35.100	<0.001*	0.011
Extraversion	2535.015	1	2.213	0.137	0.001
Openness to experience	5919.692	1	5.168	0.023	0.002
Agreeableness	9377.554	1	8.187	0.004*	0.003
Conscientiousness	10897.472	1	9.514	0.002*	0.003
Pittsburgh Sleep Quality Index	2656.198	1	2.319	0.128	0.001
RSES	7.987	1	0.007	0.933	<0.001
Sex (referent male)	8096.891	1	7.069	0.008*	0.002
Tobacco use (referent no)	12866.257	1	11.233	0.001*	0.004
Error	1145.388	3039			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

There were no significant fixed effects for treatment condition or chronotype. However, there was a significant fixed effect for week as well as a significant interaction effect between treatment condition and week. For the fixed effect of week (Figure 24), pairwise differences occurred between week 1 versus weeks 4 through 9 ($p \leq 0.004$), week 2 versus week 9 ($p < 0.001$), week 3 versus week 9 ($p < 0.001$), and week 5 versus week 9 ($p = 0.007$).

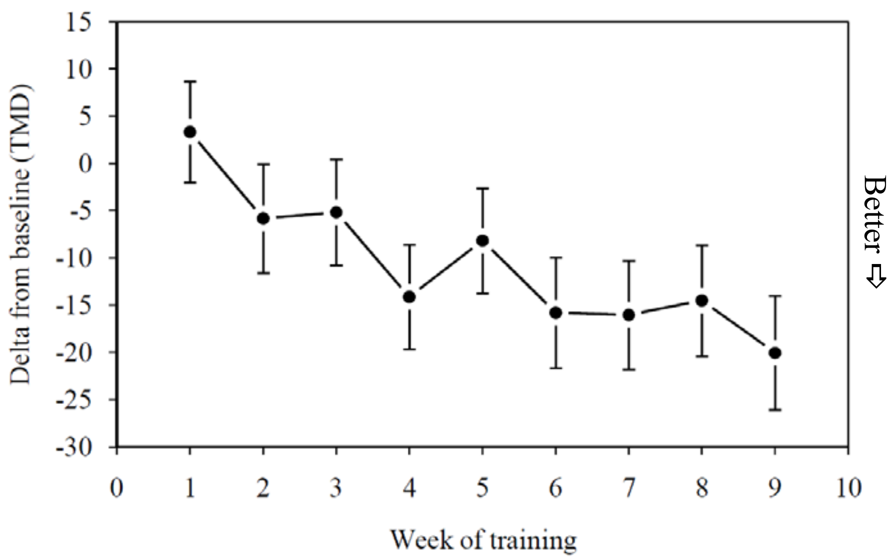


Figure 24. Estimated marginal means for POMS Total Mood Disturbance (TMD) delta from baseline scores by week of training (error bars are for 99% confidence intervals).

As shown in Figure 25, the comparison group started out with a higher mean TMD score, but had a greater rate of decrease in scores over the course of training relative to the intervention group.

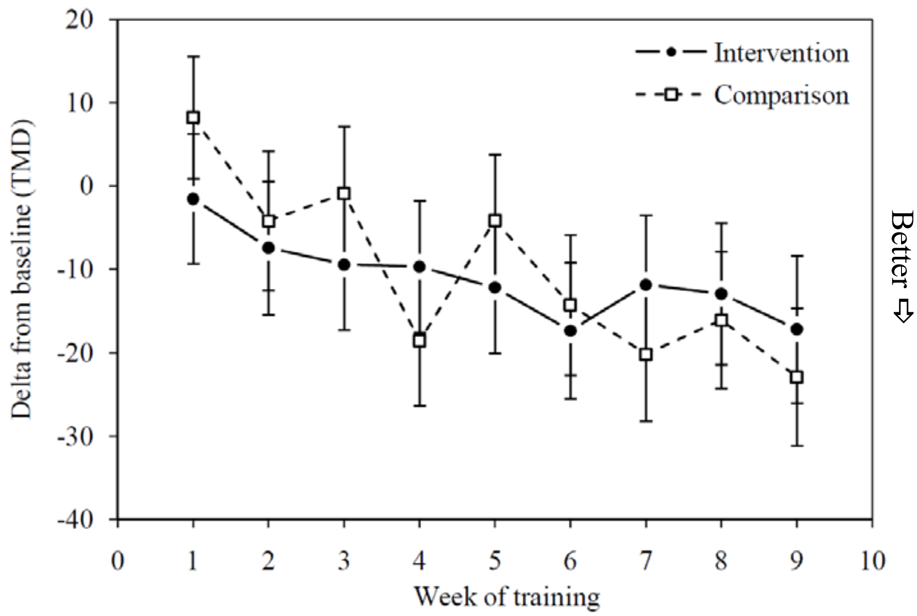


Figure 25. Estimated marginal means for POMS Total Mood Disturbance (TMD) delta from baseline scores by treatment condition and week of training (error bars are for 99% confidence intervals).

Significant covariates included age, BMI, ESS score, GT score, NEO (neuroticism, agreeableness, and conscientiousness component scores), sex, and tobacco use. Only age, BMI, and neuroticism had effect sizes of at least 1%.

8. Actigraphy Subsample

The analysis of the POMS data was repeated for the subsample of participants for which actigraphy data was available. The same analytic approach was used with the exception that weekly average hours slept was used as the covariate. Table 19 summarizes the results of the multivariate tests. There was no significant fixed effect of treatment condition or week, but there was a significant fixed effect of chronotype as well as a significant interaction effect between treatment condition and chronotype. There was also a significant multivariate effect of the covariate, weekly average hours slept, but the covariate was not significant in any of the subsequent univariate tests.

Table 19. Multivariate tests for POMS delta from baseline scores for actigraphy subsample.

Source	Wilks' λ	*F*	df1	df2	*P*	η^2
Condition	0.989	1.258	6	686	0.275	0.011
Week	0.907	1.415	48	3379	0.032	0.016
Chronotype	0.863	8.749	12	1372	<0.001*	0.071
Condition x Week	0.960	0.584	48	3379	0.990	0.007
Condition x Chronotype	0.874	7.945	12	1372	<0.001*	0.065
Chronotype x Week	0.942	0.429	96	3893	1.000	0.010
Condition x Chronotype x Week	0.947	0.394	96	3893	1.000	0.009
Average weekly sleep	0.971	3.458	6	686	0.002*	0.029

*Significant at ≤ 0.01 level.
Note: MS = Mean square.

The analysis of the respective univariate tests revealed significant fixed effects of chronotype for T-factor ($F_{2,691} = 15.888$, $p < 0.001$, $\eta^2 = 0.044$), D-factor ($F_{2,691} = 14.710$, $p < 0.001$, $\eta^2 = 0.041$), A-factor ($F_{2,691} = 9.508$, $p < 0.001$, $\eta^2 = 0.027$), V-factor ($F_{2,691} = 7.730$, $p < 0.001$, $\eta^2 = 0.022$), F-factor ($F_{2,691} = 16.262$, $p < 0.001$, $\eta^2 = 0.045$), and C-factor ($F_{2,691} = 21.489$, $p < 0.001$, $\eta^2 = 0.059$). In the case of T-factor, D-factor, and F-factor, pairwise differences occurred between indeterminate versus both evening and morning chronotypes; the basic pattern was as shown in Figure 26 for T-factor.

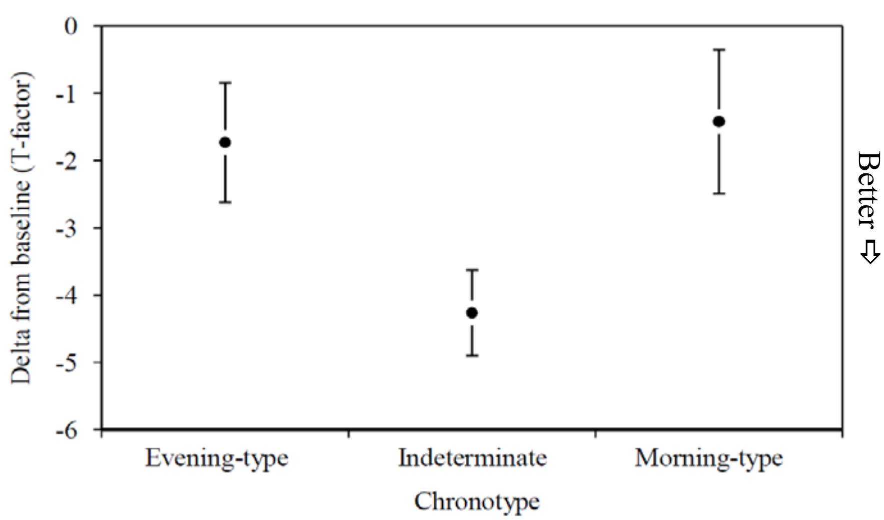


Figure 26. Estimated marginal means for POMS T-factor delta from baseline scores by chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

For A-factor, the pairwise difference occurred between indeterminate and morning chronotypes (Figure 27), whereas the pairwise difference occurred between evening versus morning chronotypes for V-factor (Figure 28).

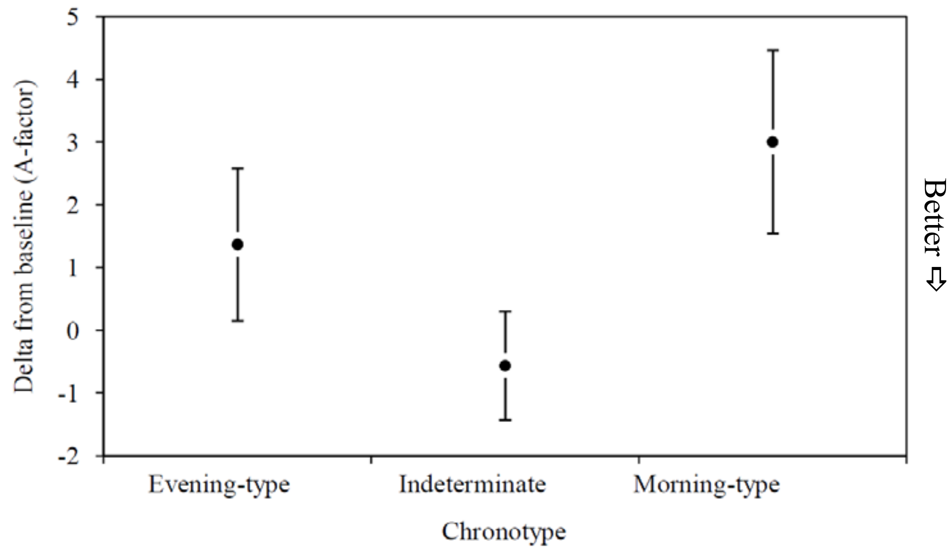


Figure 27. Estimated marginal means for POMS A-factor delta from baseline scores by chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

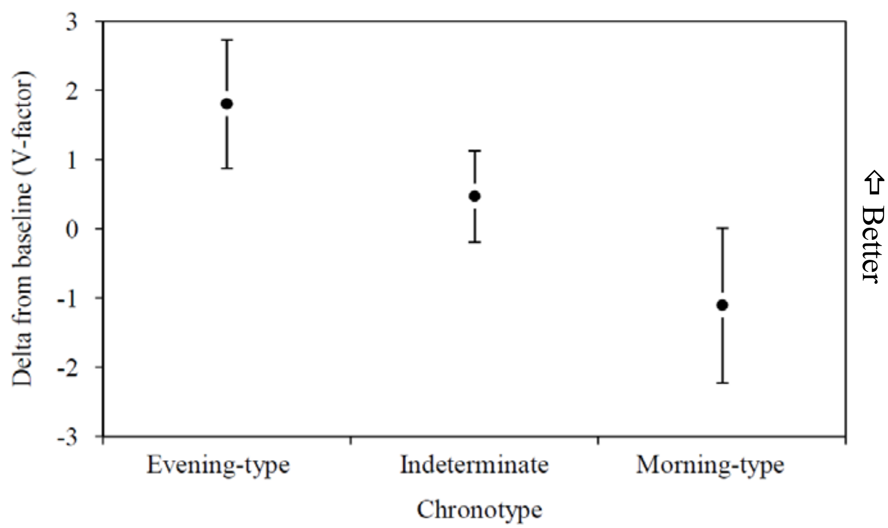


Figure 28. Estimated marginal means for POMS V-factor delta from baseline scores by chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

In the case of C-factor, the pairwise differences occurred between evening and both indeterminate and morning chronotypes (Figure 29).

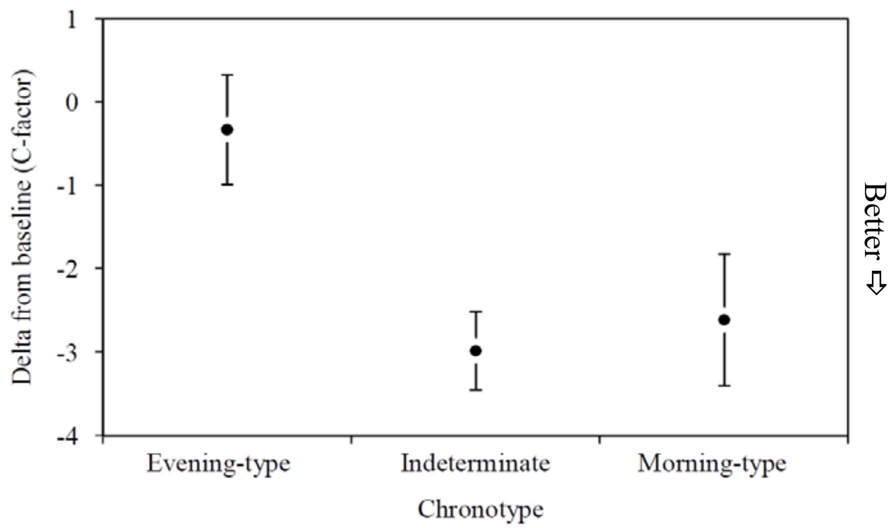


Figure 29. Estimated marginal means for POMS C-factor delta from baseline scores by chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

The univariate tests also revealed significant interaction effects between treatment condition and chronotype for T-factor ($F_{2,691}$ = 14.882, p < 0.001, η^2 = 0.041), D-factor ($F_{2,691}$ = 18.472, p < 0.001, η^2 = 0.051), A-factor ($F_{2,691}$ = 6.264, p = 0.002, η^2 = 0.018), V-factor ($F_{2,691}$ = 9.716, p < 0.001, η^2 = 0.027), and C-factor ($F_{2,691}$ = 19.404, p < 0.001, η^2 = 0.053). Figure 30 illustrates the interaction effect for D-factor; T-factor, A-factor, and C-factor followed similar patterns with evening and indeterminate chronotype participants having lower scores in the intervention group versus the comparison group, while the opposite was true for morning chronotype participants.

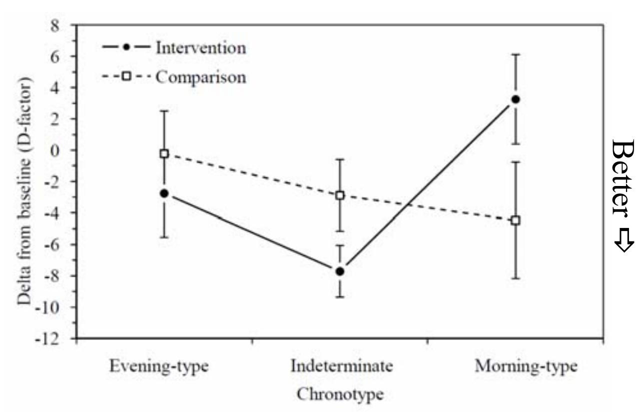


Figure 30. Estimated marginal means for POMS D-factor delta from baseline scores by treatment condition and chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

Figure 31 illustrates the interaction effect for V-factor, with evening chronotype participants having lower scores in the intervention group versus the comparison group, while the opposite was true for intermediate and morning chronotype participants.

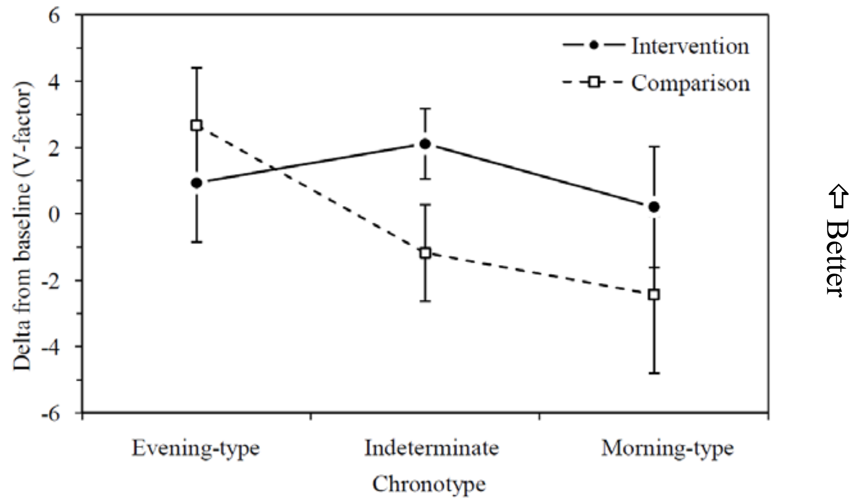


Figure 31. Estimated marginal means for POMS V-factor delta from baseline scores by treatment condition and chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

The univariate analysis of TMD delta from baseline scores for the subsample of participants with actigraphy data (Table 20) showed significant fixed effects for week and chronotype, but not for treatment condition.

Table 20. Univariate tests of between-participant effects for Total Mood Disturbance delta from baseline scores for actigraphy subsample.

Source	MS	df	*F*	*p*	η^2
Condition	2.322	1	0.003	0.960	<0.001
Week	2623.315	8	2.889	0.004*	0.032
Chronotype	16401.755	2	18.060	<0.001*	0.050
Condition x Week	1065.655	8	1.173	0.313	0.013
Condition x Chronotype	11831.703	2	13.028	<0.001*	0.036
Chronotype x Week	305.067	16	0.336	0.993	0.008
Condition x Chronotype x Week	387.332	16	0.426	0.976	0.010
Average weekly sleep	35.315	1	0.039	0.844	<0.001
Error	908.191	690			

*Significant at ≤ 0.01 level.
Note: MS = Mean square.

For the fixed effect of week (Figure 32), a pairwise difference occurred between week 1 versus week 9 ($p = 0.009$).

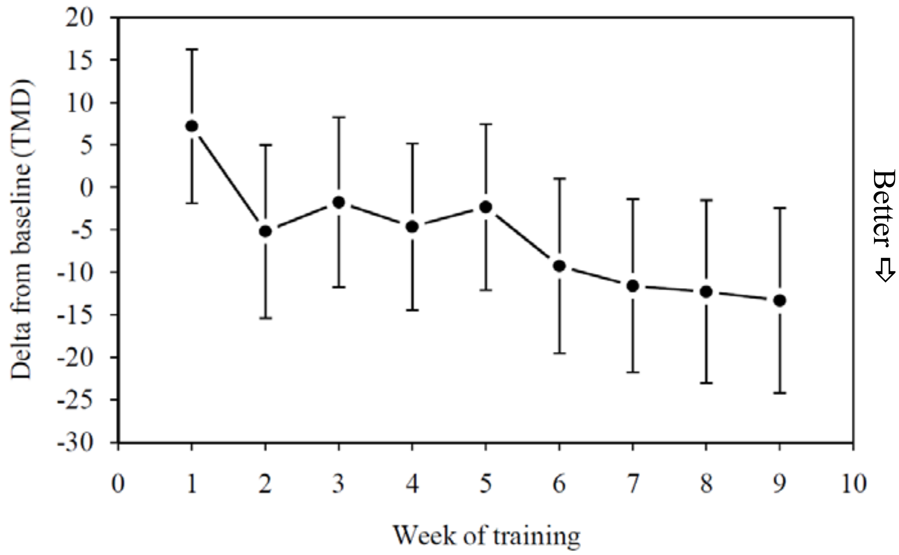


Figure 32. Estimated marginal means for POMS Total Mood Disturbance (TMD) delta from baseline scores by week of training for the actigraphy subsample (error bars are for 99% confidence intervals).

For the fixed effect of chronotype (Figure 33), pairwise differences occurred between indeterminate chronotype versus both evening and morning chronotypes ($p < 0.001$).

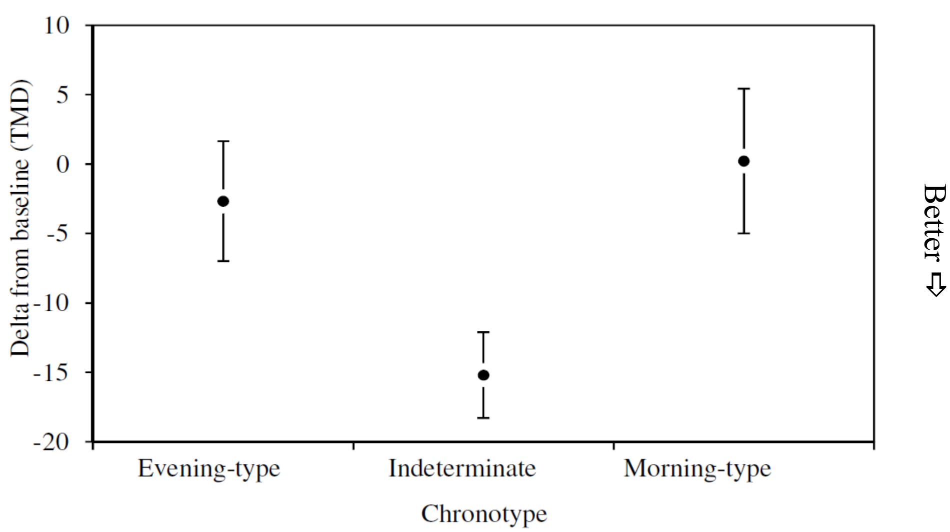


Figure 33. Estimated marginal means for POMS Total Mood Disturbance (TMD) delta from baseline scores by chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

There was also a significant interaction effect between treatment condition and chronotype (Figure 34), with evening and indeterminate chronotype participants having lower scores in the intervention group versus the comparison group, while the opposite was true for morning chronotype participants. There was no significant effect of the covariate, weekly average hours slept.

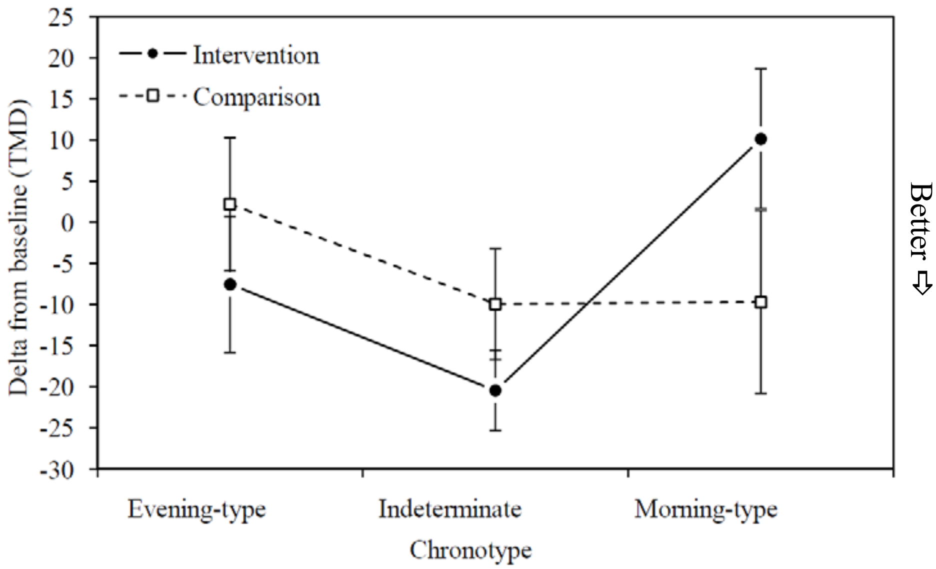


Figure 34. Estimated marginal means for POMS Total Mood Disturbance (TMD) delta from baseline scores by study condition and chronotype for actigraphy subsample (error bars are for 99% confidence intervals).

D. BASIC RIFLE MARKSMANSHIP

We assessed how participants' basic rifle marksmanship performance (on record fires) was related to treatment condition and chronotype, while accounting for potential covariates. However, when the marksmanship database was received from each company, several issues needed to be addressed prior to choosing an analytical approach. First, although both companies were issued the same number of rounds per participant for basic rifle marksmanship training, each company fired those rounds at a different rate. The intervention group accomplished record fires on four separate days, while the comparison group did so on three separate days. Accordingly, there were a maximum of four scores for each participant in the database for the intervention group and three scores per participant in the database for the comparison group. Additionally, not every participant accomplished the available maximum number of record fires.

These issues were addressed by analyzing the marksmanship scores using a simple pre/post repeated measures design in which the first recorded marksmanship score for each participant was denoted as the pre score and the last score was denoted as the post score. A repeated measures ANCOVA of marksmanship score was accomplished using practice as a within-participant effect; study condition and chronotype as fixed between-participant effects; and age, caffeine and tobacco use, component, firearm use, GT score, personality component scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates. In addition, given that marksmanship fundamentals were taught during the week prior to the record fires, POMS measurements from the week prior to ($t^* - 1$) and the week of (t^*) the record fires were also included as covariates.

A total of 372 participants, 201 in the intervention group (90% of the initial cohort) and 171 in the comparison group (87% of the initial cohort), had at least two observations recorded in the marksmanship databases. Tables 21 and 22 display the results for the within-participant model. Based on a 5% significance level, there was no significant within-participant effect of practice.

Table 21. Within-participant effects for marksmanship score.

Source	MS	df	F	p	η^2
Practice	53.799	1	2.662	0.104	0.008
Practice x Condition	196.757	1	9.737	0.002*	0.030
Practice x Chronotype	5.314	2	0.263	0.769	0.002
Practice x Condition x Chronotype	1.235	2	0.061	0.941	<0.001
Practice x Age	0.777	1	0.038	0.845	<0.001
Practice x Body mass index	14.825	1	0.734	0.392	0.002
Practice x Caffeine use (referent no)	25.043	1	1.239	0.266	0.004
Practice x Component (referent regular)	2.255	1	0.112	0.739	<0.001
Practice x Epworth Sleepiness Scale	11.565	1	0.572	0.450	0.002
Practice x Firearm use (referent no)	2.682	1	0.133	0.716	<0.001
Practice x GT score	45.644	1	2.259	0.134	0.007
Practice x NEO neuroticism	50.031	1	2.476	0.117	0.008
Practice x NEO extraversion	74.857	1	3.705	0.055	0.012
Practice x NEO openness to experience	8.837	1	0.437	0.509	0.001

Practice x NEO agreeableness	7.876	1	0.390	0.533	0.001
Practice x NEO conscientiousness	0.163	1	0.008	0.928	<0.001
Practice x PSQI	6.056	1	0.300	0.584	0.001
Practice x POMS week *t** – 1 T-factor	3.562	1	0.176	0.675	0.001
Practice x POMS week *t** – 1 D-factor	0.810	1	0.040	0.841	<0.001
Practice x POMS week *t** – 1 A-factor	27.994	1	1.385	0.240	0.004
Practice x POMS week *t** – 1 V-factor	0.865	1	0.043	0.836	<0.001
Practice x POMS week *t** – 1 F-factor	18.454	1	0.913	0.340	0.003
Practice x POMS week *t** – 1 C-factor	20.848	1	1.032	0.311	0.003

*Significant at ≤ 0.05 level.
Notes: GT score = General technical aptitude score; MS = Mean square; PSQI = Pittsburgh Sleep Quality Index; POMS = Profile of Mood States.

Table 22. Within-participant effects for marksmanship score (continued).

Source	MS	df	*F*	*p*	η^2
Practice x POMS week *t** T-factor	6.477	1	0.321	0.572	0.001
Practice x POMS week *t** D-factor	0.014	1	0.001	0.979	<0.001
Practice x POMS week *t** A-factor	16.824	1	0.833	0.362	0.003
Practice x POMS week *t** V-factor	8.999	1	0.445	0.505	0.001
Practice x POMS week *t** F-factor	17.276	1	0.855	0.356	0.003
Practice x POMS week *t** C-factor	83.390	1	4.127	0.043*	0.013
Practice x RSES	0.680	1	0.034	0.855	<0.001
Practice x Sex (referent male)	10.100	1	0.500	0.480	0.002
Practice x Tobacco use (referent no)	0.740	1	0.037	0.848	<0.001
Error	20.206	313			

*Significant at ≤ 0.05 level.
Notes: MS = Mean square; POMS = Profile of Mood States; RSES = Response to Stressful Experiences Scale.

There was a significant interaction effect between practice and treatment condition, but there was no interaction effect between practice and chronotype. Participants in the intervention group had significantly lower initial scores than participants in the comparison group, but participants in the intervention group had greater improvement in scores with practice such that their final scores were equivalent to those of participants in the comparison group (Figure 35). There was also a significant within-participant interaction between practice and *t** week POMS C-factor score.

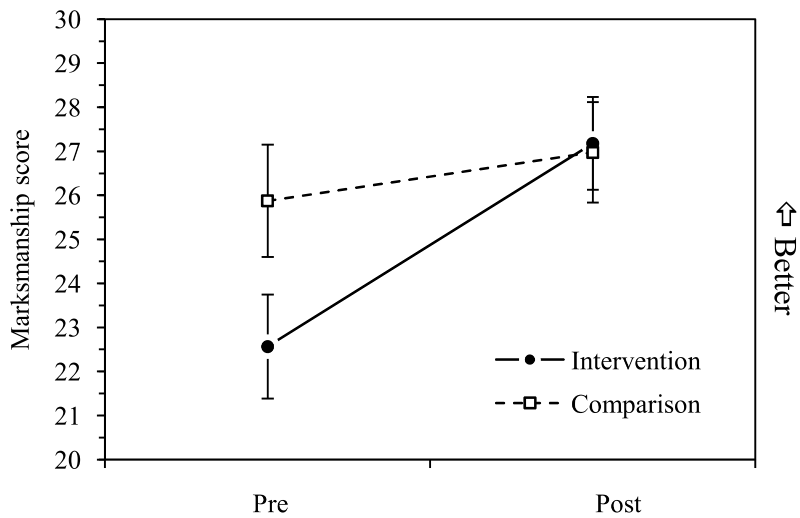


Figure 35. Estimated marginal means for first and last marksmanship scores by treatment condition (error bars are for 95% confidence intervals).

In terms of the between-participant model (Tables 23 and 24), there was a significant fixed effect for treatment condition, with an estimated marginal mean score for the intervention group of 24.872 (95% CI: 23.973, 25.453) versus 26.425 (95% CI: 25.772, 27.397) for the comparison group. Fixed effect of chronotype was not significant, nor was there an interaction effect between treatment condition and chronotype. The only significant covariates were prior use of firearms and sex.

Table 23. Between-participant effects for marksmanship score.

Source	MS	df	*F*	*p*	η^2
Condition	153.391	1	4.183	0.042*	0.013
Chronotype	5.402	2	0.147	0.863	0.001
Condition x Chronotype	43.510	2	1.186	0.307	0.008
Age	0.078	1	0.002	0.963	0.000
Body mass index	30.719	1	0.838	0.361	0.003
Caffeine use (referent no)	55.449	1	1.512	0.220	0.005
Component (referent regular)	23.717	1	0.647	0.422	0.002
Epworth Sleepiness Scale	74.759	1	2.039	0.154	0.006
Firearm use (referent no)	173.043	1	4.719	0.031*	0.015
GT score	84.001	1	2.291	0.131	0.007
NEO-FFI					

Neuroticism	11.672	1	0.318	0.573	0.001
Extraversion	5.767	1	0.157	0.692	0.001
Openness to experience	77.751	1	2.120	0.146	0.007
Agreeableness	41.375	1	1.128	0.289	0.004
Conscientiousness	16.079	1	0.438	0.508	0.001
Pittsburgh Sleep Quality Index	38.364	1	1.046	0.307	0.003

*Significant at ≤ 0.05 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory.

Table 24. Between-participant effects for marksmanship score (continued).

Source	MS	df	F	p	η^2
Week t^* – 1 Profile of Mood States					
T-factor	86.493	1	2.359	0.126	0.007
D-factor	27.612	1	0.753	0.386	0.002
A-factor	0.089	1	0.002	0.961	0.000
V-factor	0.902	1	0.025	0.876	0.000
F-factor	129.144	1	3.522	0.062	0.011
C-factor	22.697	1	0.619	0.432	0.002
Week t^* Profile of Mood States					
T-factor	0.415	1	0.011	0.915	0.000
D-factor	57.535	1	1.569	0.211	0.005
A-factor	5.613	1	0.153	0.696	0.000
V-factor	46.526	1	1.269	0.261	0.004
F-factor	15.798	1	0.431	0.512	0.001
C-factor	0.325	1	0.009	0.925	0.000
RSES	10.603	1	0.289	0.591	0.001
Sex (referent male)	434.120	1	11.838	0.001*	0.036
Tobacco use (referent no)	7.273	1	0.198	0.656	0.001
Error	36.673	313			

*Significant at ≤ 0.05 level.
Notes: MS = Mean square; RSES = Response to Stressful Experiences Scale.

The analysis was repeated for the subsample of participants for which actigraphy data was available. The same general analytic approach was used except that the average hours slept during the week prior to (t^* – 1) and the week of (t^*) the record fires were used as the covariates. A total of 90 participants, 52 (98% of the initial subcohort) in the intervention group and 38 (93% of the initial subcohort) in the comparison group, had at least two observations recorded in the marksmanship databases. Table 25 displays the results for the within-participant model. Again, using a 5% significance level, there was no significant within-participant effect of practice, but there was a significant interaction effect between practice and treatment condition.

Table 25. Within-participant effects for marksmanship score for the actigraphy subsample.

Source	MS	df	*F*	*p*	η^2
Practice	5.079	1	0.289	0.593	0.004
Practice x Condition	105.668	1	6.003	0.017*	0.071
Practice x Chronotype	1.681	2	0.095	0.909	0.002
Practice x Condition x Chronotype	3.893	2	0.221	0.802	0.006
Practice x Week t^* – 1 average sleep	65.360	1	3.713	0.058	0.045
Practice x Week t^* average sleep	21.476	1	1.220	0.273	0.015
Error	17.602	78			

*Significant at ≤ 0.05 level.
Note: MS = Mean square.

Although the intervention and comparison groups did not differ in terms of mean initial and final scores, there was a trend for participants in the intervention group to have a greater improvement in scores with practice than participants in the comparison group (Figure 36). There was no interaction effect between practice and chronotype, nor were there any interaction effects between practice and the covariates.

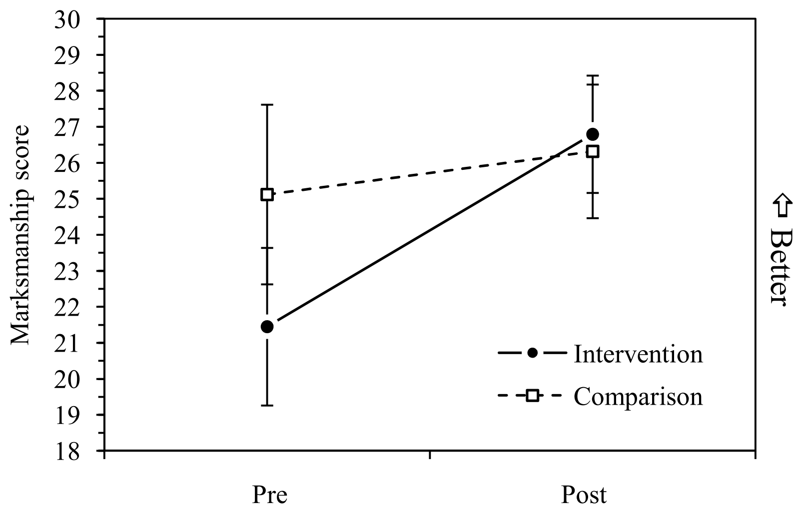


Figure 36. Estimated marginal means for first and last marksmanship scores by treatment condition for the actigraphy subsample (error bars are for 95% confidence intervals).

In terms of the between-participant model (Table 26), there was no significant fixed effect of treatment condition in the presence of the sleep covariates. Additionally, there was no significant fixed effect for chronotype, nor was there an interaction effect between treatment condition and chronotype. There was, however, a significant effect for the covariate, week $t^* - 1$ average sleep, but not week t^* average sleep.

Table 26. Between-participant effects for marksmanship score for the actigraphy subsample.

Source	MS	df	*F*	*p*	η^2
Condition	62.723	1	1.439	0.234	0.018
Chronotype	5.237	2	0.120	0.887	0.003
Condition x Chronotype	56.897	2	1.305	0.277	0.032
Week $t^* - 1$ average sleep	177.670	1	4.076	0.047*	0.050
Week t^* average sleep	48.316	1	1.108	0.296	0.014
Error	43.589	78			

*Significant at ≤ 0.05 level.
Note: MS = Mean square.

E. PHYSICAL FITNESS

It was of interest to determine how participants' performance on the Army Physical Fitness Test related to treatment condition and chronotype, while accounting for potential covariates. However, an issue was identified upon receipt of the physical fitness database from each company that needed to be addressed prior to choosing an analytic approach. Forty-nine (12.5%) participants had no scores reported for any of the three physical fitness tests, 10.2% of the remaining 343 participants had no scores reported for either one or two of the physical fitness tests. This issue was addressed by analyzing the physical fitness dataset as a repeated cross-section design, rather than a within-participant repeated measures design, and using a 1% significance level to counter the resulting increased power of statistical tests.

A MANCOVA of the component physical fitness scores (push-ups, sit-ups, and run) was accomplished using treatment condition, week, and chronotype as fixed effects and age, caffeine and tobacco use, component, fitness factors (BMI and exercise frequency), GT score, personality component scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates. In addition, POMS measurements from the week of the corresponding physical fitness test were also included as covariates. Tables 27 and 28 summarize the results of the multivariate tests. There were significant fixed effects for treatment condition, week, and chronotype as well as a significant interaction effect between treatment condition and week. There were also significant effects for the covariates age, BMI, exercise frequency, GT score, NEO neuroticism component score, POMS A-factor score, and sex.

Table 27. Multivariate tests for physical fitness component scores.

Source	Wilks' λ	*F*	df1	df2	*p*	η^2
Condition	0.964	11.037	3	884	<0.001*	0.036
Week	0.955	6.868	6	1768	<0.001*	0.023
Chronotype	0.963	5.676	6	1768	<0.001*	0.019
Condition x Week	0.978	3.317	6	1768	0.003*	0.011
Condition x Chronotype	0.994	0.838	6	1768	0.540	0.003

Chronotype x Week	0.995	0.396	12	2339	0.966	0.002
Condition x Chronotype x Week	0.994	0.425	12	2339	0.954	0.002
Age	0.952	14.765	3	884	<0.001*	0.048
Body mass index	0.887	37.504	3	884	<0.001*	0.113
Caffeine use (referent no)	1.000	0.045	3	884	0.987	<0.001
Component (referent regular)	0.996	1.201	3	884	0.308	0.004
Epworth Sleepiness Scale	0.997	0.919	3	884	0.431	0.003
Exercise frequency	0.981	5.601	3	884	0.001*	0.019
GT score	0.976	7.391	3	884	<0.001*	0.024
NEO-FFI						
Neuroticism	0.975	7.442	3	884	<0.001*	0.025
Extraversion	0.990	3.011	3	884	0.029	0.010
Openness to experience	0.999	0.196	3	884	0.899	0.001
Agreeableness	0.999	0.376	3	884	0.770	0.001
Conscientiousness	0.990	2.840	3	884	0.037	0.010

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score.

Table 28. Multivariate tests for physical fitness component scores (continued).

Source	Wilks' λ	*F*	df1	df2	*p*	η^2
Pittsburgh Sleep Quality Index	0.991	2.758	3	884	0.041	0.009
Profile of Mood States						
T-factor	0.994	1.645	3	884	0.177	0.006
D-factor	0.996	1.120	3	884	0.340	0.004
A-factor	0.976	7.167	3	884	<0.001*	0.024
V-factor	0.995	1.465	3	884	0.223	0.005
F-factor	0.993	2.177	3	884	0.089	0.007
C-factor	0.997	0.918	3	884	0.432	0.003
RSES	0.993	1.948	3	884	0.120	0.007
Sex (referent male)	0.944	17.607	3	884	<0.001*	0.056
Tobacco use (referent no)	0.999	0.242	3	884	0.867	0.001

*Significant at ≤ 0.01 level.
Notes: NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

Table 29 provides the results of the relevant univariate tests of between-participant effects for push-up score. There were significant fixed effects for treatment condition and week as well as a significant interaction effect between condition and week. The estimated marginal mean push-up score for the intervention group was 76.404 (99% CI: 73.992, 78.816) versus 70.475 (99% CI: 67.921, 73.028) for the comparison group.

Table 29. Univariate tests of between-participant effects for push-up score.

Source	MS	df	*F*	*p*	η^2
Condition	3727.319	1	16.107	<0.001*	0.018
Week	3250.914	2	14.048	<0.001*	0.031
Chronotype	333.852	2	1.443	0.237	0.003
Condition x Week	1588.026	2	6.862	0.001*	0.015
Age	920.453	1	3.978	0.046	0.004
Body mass index	6508.729	1	28.126	<0.001*	0.031
Exercise frequency	3338.788	1	14.428	<0.001*	0.016
GT score	1573.779	1	6.801	0.009*	0.008
NEO-FFI neuroticism	994.902	1	4.299	0.038	0.005
POMS A-factor	842.023	1	3.639	0.057	0.004
Sex (referent male)	1622.487	1	7.011	0.008*	0.008
Error	231.413	886			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; POMS = Profile of Mood States.

For the fixed effect of week (Figure 37), the pairwise difference occurred between week 3 versus week 8 ($p < 0.001$). Note that physical fitness assessments were only accomplished in weeks 3, 6, and 9.

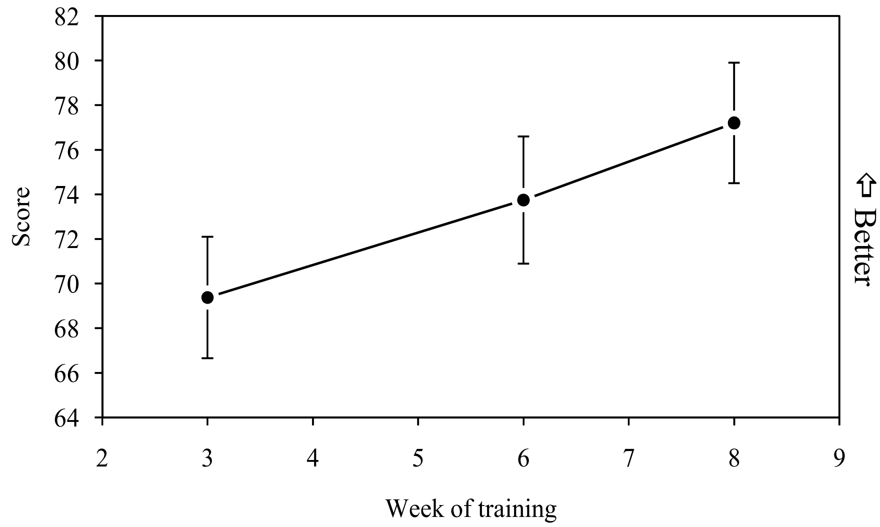


Figure 37. Estimated marginal means for push-up score by week of training (error bars are for 99% confidence intervals).

Regarding the interaction effect (Figure 38), the intervention and comparison groups differed in mean push-up score at week 3, but participants in the comparison group improved at a faster rate than those in the intervention group, such that there were no differences in mean score by weeks 6 and 8.

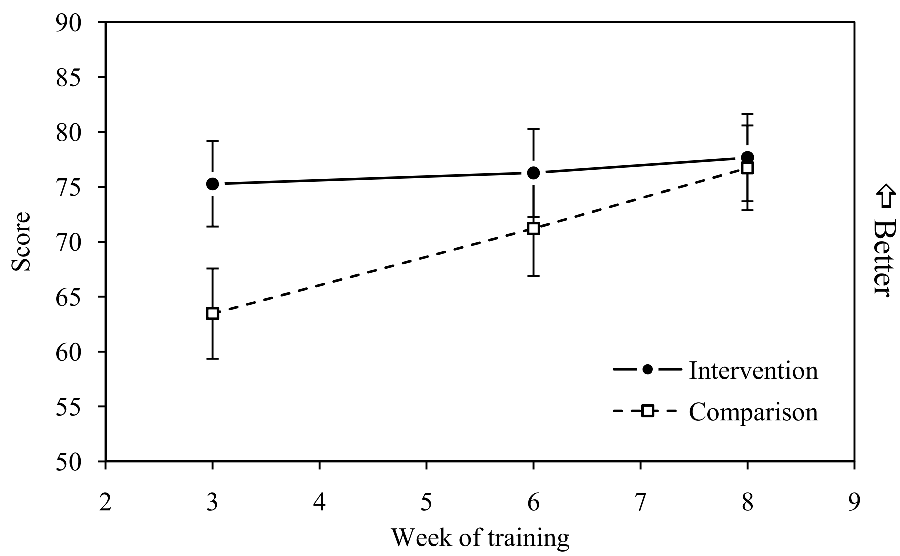


Figure 38. Estimated marginal means for push-up score by treatment condition and week of training (error bars are for 99% confidence intervals).

Significant covariates included age, BMI, exercise frequency, GT score, and sex, although BMI and exercise frequency had effect sizes that were two to four times greater than the effect sizes of GT score and sex.

Table 30 provides the results of the relevant univariate tests of between-participant effects for sit-up score. There were significant fixed effects for treatment condition, week, and chronotype as well as a significant interaction effect between treatment condition and week. The estimated marginal mean push-up score for the intervention group was 73.128 (99% CI: 70.840, 75.416) versus 68.353 (99% CI: 65.930, 70.775) for the comparison group.

Table 30. Univariate tests of between-participant effects for sit-up score.

Source	MS	df	*F*	*p*	η^2
Condition	2417.448	1	11.610	0.001*	0.013
Week	2642.599	2	12.691	<0.001*	0.028
Chronotype	1071.267	2	5.145	0.006*	0.011
Condition x Week	1196.870	2	5.748	0.003*	0.013
Age	159.669	1	0.767	0.381	0.001
Body mass index	9580.624	1	46.010	<0.001*	0.049
Exercise frequency	1782.953	1	8.563	0.004*	0.010
GT score	4162.000	1	19.988	<0.001*	0.022
NEO-FFI neuroticism	2535.853	1	12.178	0.001*	0.014
POMS A-factor	236.754	1	1.137	0.287	0.001
Sex (referent male)	4519.173	1	21.703	<0.001*	0.024
Error	208.227	886			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; POMS = Profile of Mood States.

For the fixed effect of week (Figure 39), the pairwise difference occurred between week 3 versus week 8 ($p < 0.001$).

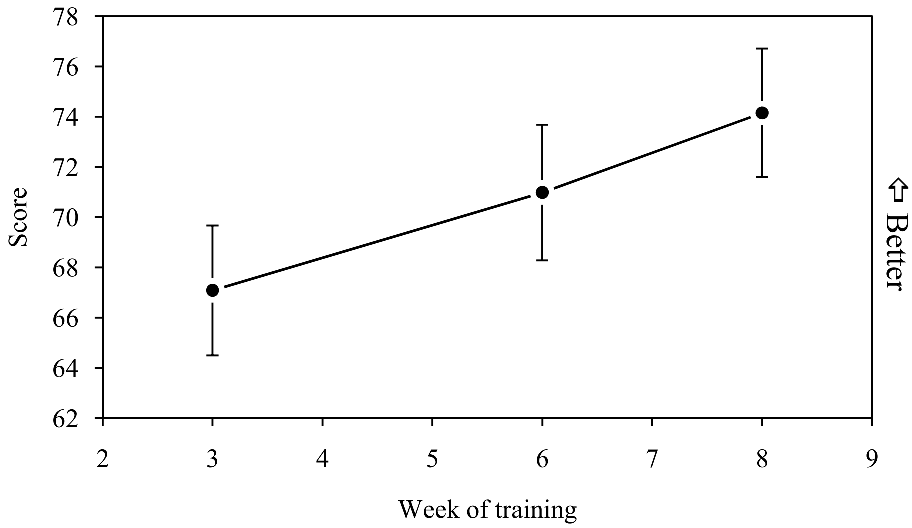


Figure 39. Estimated marginal means for sit-up score by week of training (error bars are for 99% confidence intervals).

For the fixed effect of chronotype (Figure 40), the pairwise difference occurred between evening versus indeterminate chronotypes ($p = 0.004$).

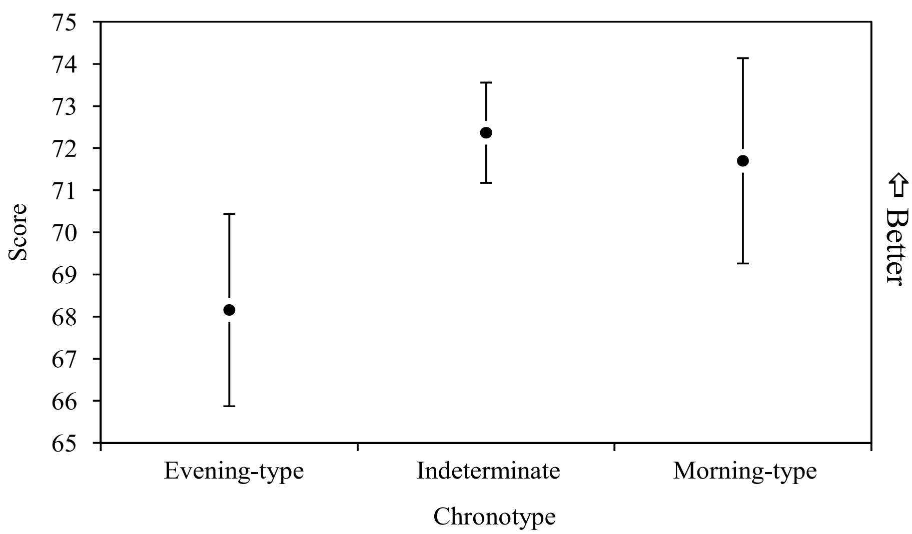


Figure 40. Estimated marginal means for sit-up score by chronotype (error bars are for 99% confidence intervals).

Regarding the interaction effect (Figure 41), the intervention and comparison groups differed in mean sit-up score at week 3, but participants in the comparison group improved at a faster rate than those in the intervention group, such that there were no differences in mean score by weeks 6 and 8.

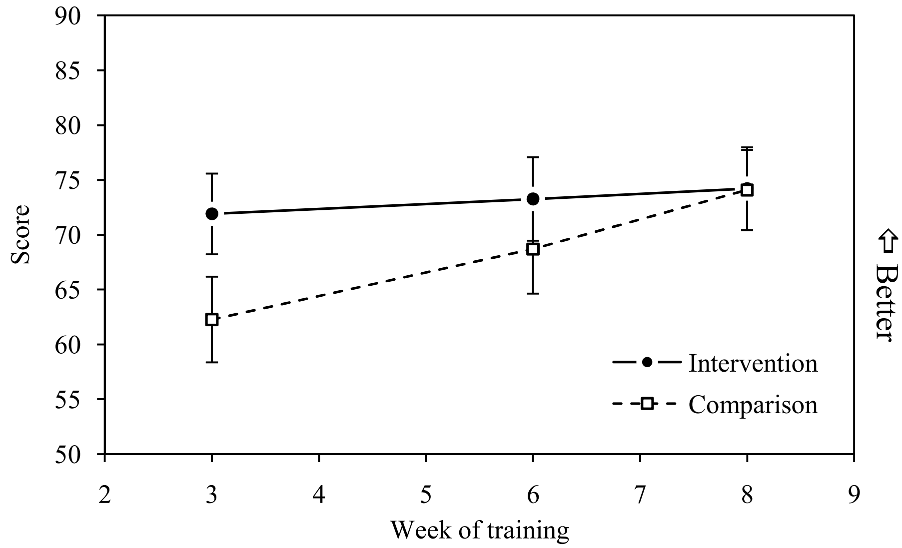


Figure 41. Estimated marginal means for sit-up score by treatment condition and week of training (error bars are for 99% confidence intervals).

Significant covariates included BMI, exercise frequency, GT score, NEO neuroticism score, and sex. Body mass index was the most important covariate in terms of effect size.

Table 31 provides the results of the relevant univariate tests of between-participant effects for the physical fitness test run score. There was no significant fixed effect for treatment condition, but there were significant fixed effects for week and chronotype.

Table 31. Univariate tests of between-participant effects for run score.

Source	MS	df	*F*	*p*	η^2
Condition	435.740	1	1.680	0.195	0.002
Week	2423.699	2	9.346	<0.001*	0.021
Chronotype	3811.444	2	14.697	<0.001*	0.032
Condition x Week	740.598	2	2.856	0.058	0.006
Age	10994.891	1	42.395	<0.001*	0.046
Body mass index	25556.018	1	98.541	<0.001*	0.100
Exercise frequency	354.690	1	1.368	0.243	0.002
GT score	2126.456	1	8.199	0.004*	0.009
NEO-FFI neuroticism	565.387	1	2.180	0.140	0.002

POMS A-factor	5532.681	1	21.333	<0.001*	0.024
Sex (referent male)	367.816	1	1.418	0.234	0.002
Error	259.343	886			

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; POMS = Profile of Mood States.

For the fixed effect of week (Figure 42), pairwise differences occurred between week 3 versus both week 6 and week 8 ($p < 0.002$).

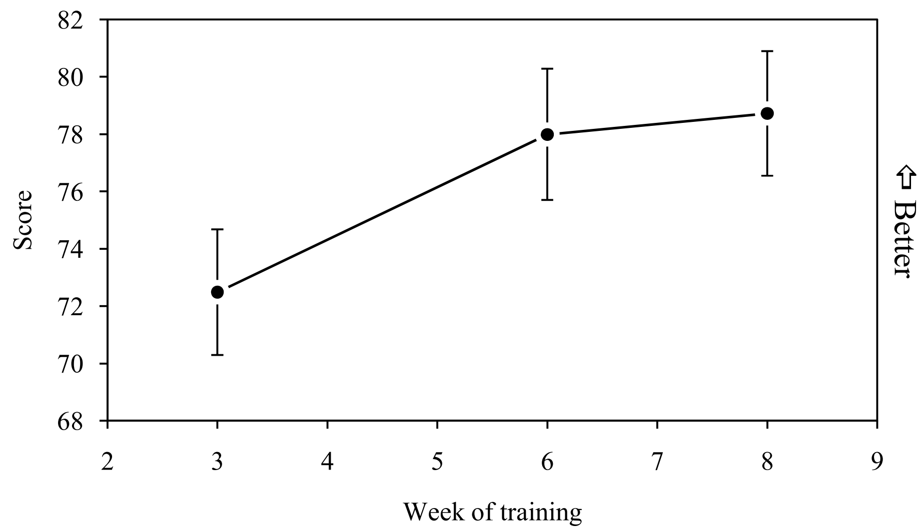


Figure 42. Estimated marginal means for run score by week of training (error bars are for 99% confidence intervals).

For the fixed effect of chronotype (Figure 43), pairwise differences occurred between evening versus both indeterminate and morning chronotypes ($p < 0.009$). Thus, evening chronotypes were slower than indeterminate and morning chronotypes.

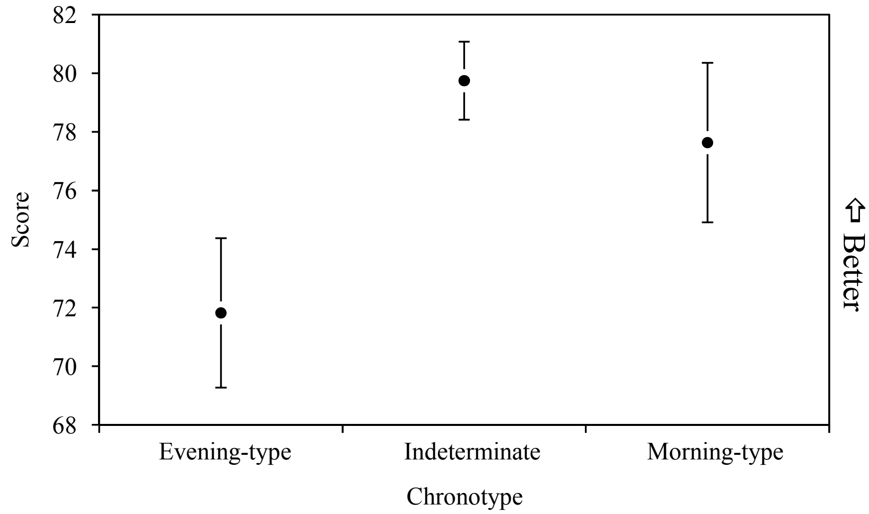


Figure 43. Estimated marginal means for run score by chronotype (error bars are for 99% confidence intervals).

There was no significant interaction effect between study condition and week. Significant covariates included age, BMI, GT score, and POMS A-factor score, although BMI was the most important covariate in terms of effect size.

The APFT score provides a single global estimate of physical fitness and is obtained by summing the scores across the three component fitness assessment activities. An ANCOVA of APFT score was accomplished using treatment condition, week, and chronotype as fixed effects and age, caffeine and tobacco use, component, fitness factors (BMI and exercise frequency), GT score, personality component scores, POMS factor scores, RSES score, sex, and sleep factors (ESS and PSQI scores) as covariates (Tables 32 and 33). There was no significant fixed effect for treatment condition, but there were significant fixed effects for week and chronotype as well as a significant interaction effect between treatment condition and week.

Table 32. Univariate tests of between-participant effects for Army Physical Fitness Test score.

Source	MS	df	*F*	*p*	η^2
Condition	7867.295	1	6.214	0.013	0.007
Week	24182.956	2	19.102	<0.001*	0.041
Chronotype	12473.396	2	9.853	<0.001*	0.022
Condition x Week	9496.913	2	7.501	0.001*	0.017

Condition x Chronotype	453.751	2	0.358	0.699	0.001
Chronotype x Week	779.760	4	0.616	0.651	0.003
Condition x Chronotype x Week	752.311	4	0.594	0.667	0.003
Age	21989.056	1	17.369	<0.001*	0.019
Body mass index	114926.602	1	90.779	<0.001*	0.093
Caffeine use (referent no)	20.595	1	0.016	0.899	<0.001
Component (referent regular)	32.099	1	0.025	0.874	<0.001
Epworth Sleepiness Scale	3086.853	1	2.438	0.119	0.003
Exercise frequency	14194.166	1	11.212	0.001*	0.012
GT score	23105.988	1	18.251	<0.001*	0.020

*Significant at ≤ 0.01 level.
Notes: GT score = General technical aptitude score; MS = Mean square.

Table 33. Univariate tests of between-participant effects for Army Physical Fitness Test score (continued).

Source	MS	df	*F*	*p*	η^2
NEO-FFI					
Neuroticism	3257.315	1	2.573	0.109	0.003
Extraversion	2419.963	1	1.911	0.167	0.002
Openness to experience	335.026	1	0.265	0.607	<0.001
Agreeableness	949.270	1	0.750	0.387	0.001
Conscientiousness	192.961	1	0.152	0.696	<0.001
Profile of Mood States					
T-factor	5577.076	1	4.405	0.036	0.005
D-factor	81.731	1	0.065	0.799	<0.001
A-factor	14252.349	1	11.258	0.001*	0.013
V-factor	5049.279	1	3.988	0.046	0.004
F-factor	3378.278	1	2.668	0.103	0.003
C-factor	280.387	1	0.221	0.638	<0.001
RSES	3535.514	1	2.793	0.095	0.003
Sex (referent male)	2184.334	1	1.725	0.189	0.002
Tobacco use	179.788	1	0.142	0.706	<0.001
Error	1266.002	886			

*Significant at ≤ 0.01 level.
Notes: MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

For the fixed effect of week (Figure 44), pairwise differences in APFT scores occurred between week 3 versus both week 6 and week 8 ($p < 0.001$).

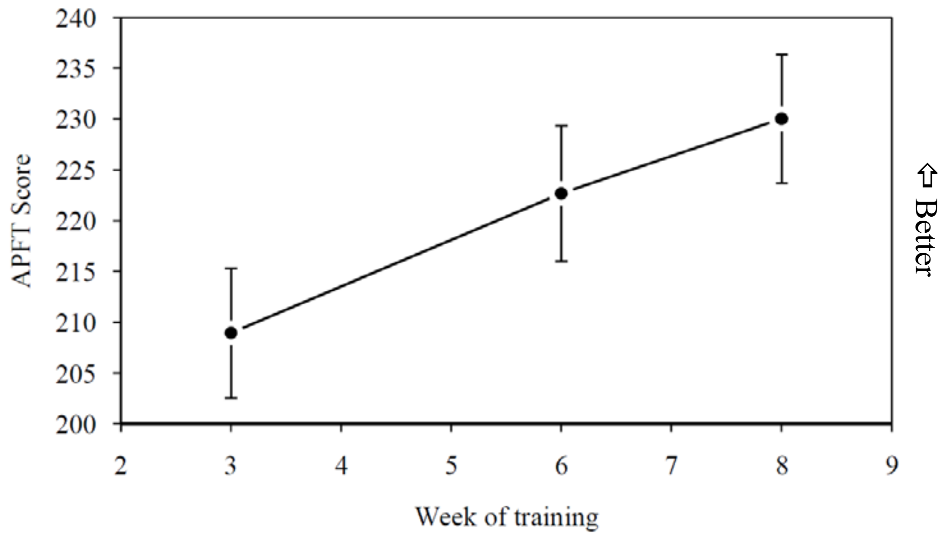


Figure 44. Estimated marginal means for Army Physical Fitness Test (APFT) score by week of training (error bars are for 99% confidence intervals).

For the fixed effect of chronotype (Figure 45), the pairwise difference in APFT scores occurred between evening versus indeterminate chronotypes ($p < 0.001$).

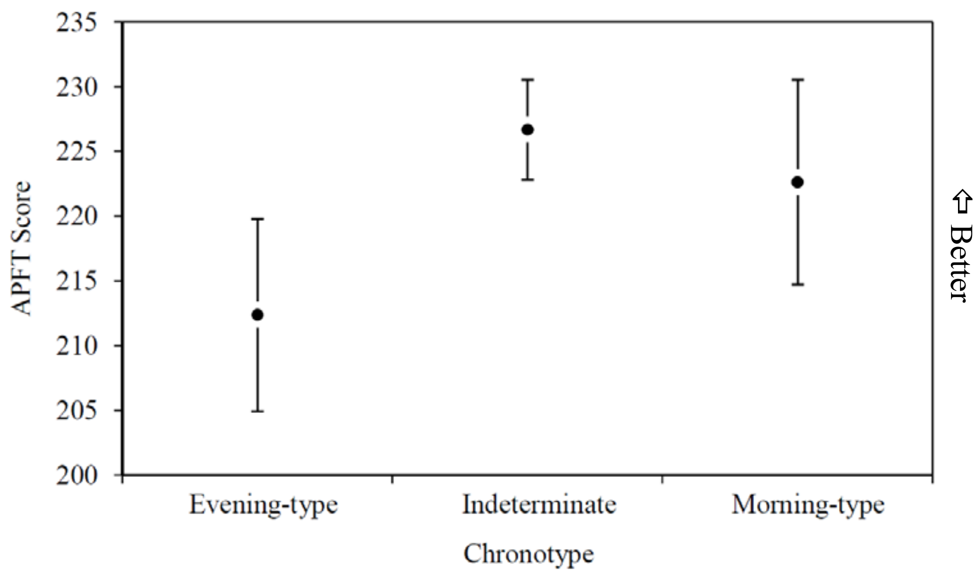


Figure 45. Estimated marginal means for Army Physical Fitness Test (APFT) score by chronotype (error bars are for 99% confidence intervals).

Regarding the interaction effect (Figure 46), the intervention and comparison groups differed in mean APFT score at week 3, but participants in the comparison group improved at a faster rate than those in the intervention group, such that there were no differences in mean score by weeks 6 and 8.

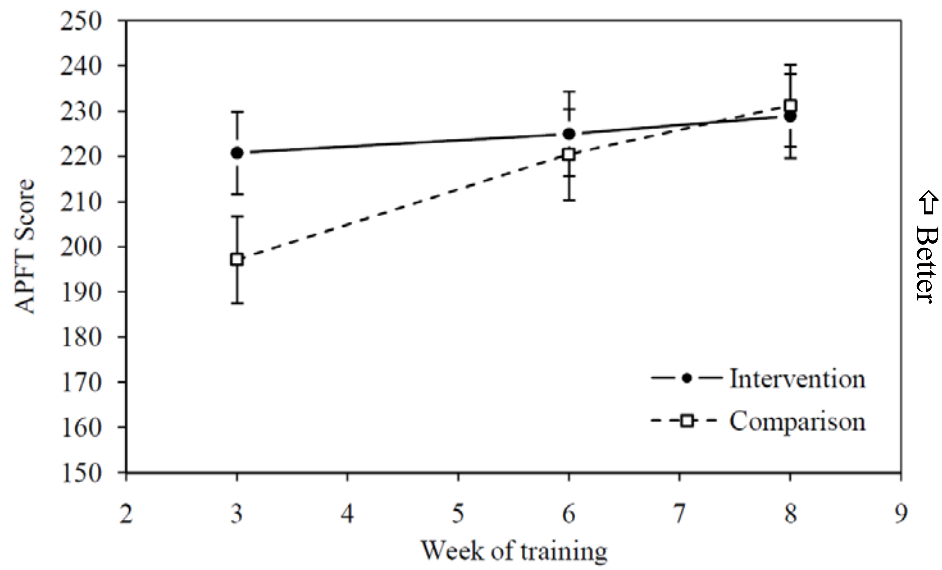


Figure 46. Estimated marginal means for Army Physical Fitness Test (APFT) score by treatment condition and week of training (error bars are for 99% confidence intervals).

Significant covariates included age, BMI, exercise frequency, GT score, and POMS A-factor score, but BMI was clearly the most important covariate based on effect size. The analysis of the fitness data was repeated for the subsample of participants for which actigraphy data was available. The same analytic approach was used with the exception that average hours slept per week was used as the covariate. Multivariate tests showed that there was not a significant overall effect of average hours slept per week. Similarly, the univariate analysis of APFT scores for the subsample of participants with actigraphy data showed no significant effect for the covariate, average hours slept per week.

F. POSTSTUDY QUESTIONNAIRE

Both the prestudy and poststudy questionnaires assessed participant sleep using two standardized survey instruments: the ESS and the PSQI. The effect of the treatment intervention on ESS and PSQI scores was assessed using a pre/poststudy design. A repeated measures ANCOVA of ESS and PSQI scores was accomplished using time as a within-participant effect; treatment condition and chronotype as fixed between-participant effects; and age, caffeine and tobacco use, component, firearm use, fitness factors (BMI and exercise frequency), GT score, personality component scores, RSES score, and sex as covariates. Because of participant attrition, there were missing poststudy questionnaires for 44 participants (21%) in the intervention group and 31 participants (17%) in the comparison group. This difference was not statistically significant.

1. Epworth Sleepiness Scale (ESS)

Based on a 5% significance level, in terms of within-participant effects of ESS score (Table 34), there was no significant within-participant effect of time, nor was there a significant interaction effect between time and chronotype. There were significant interaction effects between time and the fixed effect, treatment condition, as well as the covariate GT score.

Table 34. Within-participant effects for Epworth Sleepiness Scale score.

Source	MS	df	*F*	*p*	η^2
Time	3.157	1	0.231	0.631	0.001
Time x Condition	259.141	1	18.943	<0.001*	0.060
Time x Chronotype	7.304	2	0.534	0.587	0.004
Time x Condition x Chronotype	14.891	2	1.089	0.338	0.007
Time x Age	2.853	1	0.209	0.648	0.001
Time x Body mass index	7.710	1	0.564	0.453	0.002
Time x Caffeine use (referent no)	1.979	1	0.145	0.704	<0.001
Time x Component (referent regular)	0.406	1	0.030	0.863	<0.001
Time x Exercise frequency	4.765	1	0.348	0.556	0.001
Time x Firearm use (referent no)	13.056	1	0.954	0.329	0.003
Time x GT score	111.942	1	8.183	0.005*	0.027
Time x NEO neuroticism	0.476	1	0.035	0.852	<0.001
Time x NEO extraversion	0.261	1	0.019	0.890	<0.001
Time x NEO openness to experience	4.235	1	0.310	0.578	0.001
Time x NEO agreeableness	44.847	1	3.278	0.071	0.011

Time x NEO conscientiousness	4.997	1	0.365	0.546	0.001
Time x RSES	0.091	1	0.007	0.935	<0.001
Time x Sex (referent male)	38.794	1	2.836	0.093	0.009
Time x Tobacco (referent no)	3.389	1	0.248	0.619	0.001
Error	13.680	296			

*Significant at ≤ 0.05 level.
Notes: MS = Mean square; RSES = Response to Stressful Experiences Scale.

The interaction effect between time and treatment condition is shown in Figure 47. ESS scores increased significantly for participants in the comparison group over the course of training, but remained unchanged for those in the intervention group. Consequently, the groups' mean scores differed significantly at the poststudy assessment, with the comparison group reporting greater sleepiness.

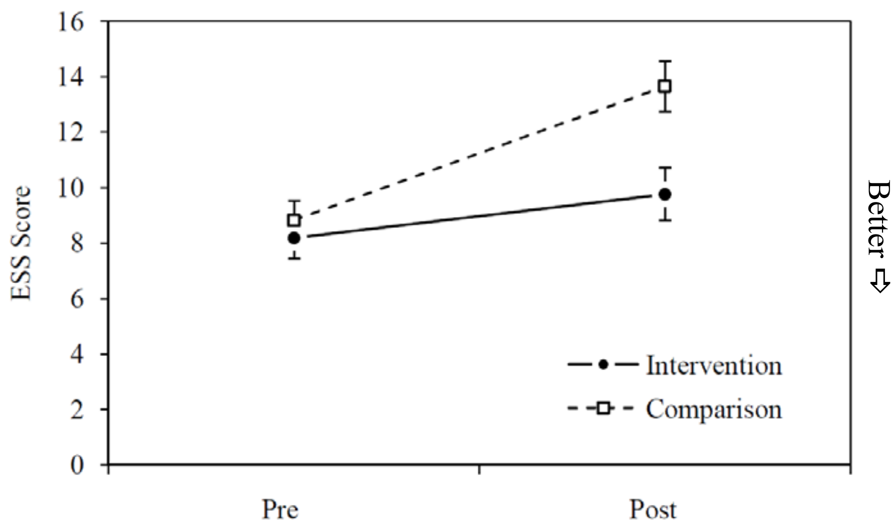


Figure 47. Estimated marginal means for ESS score by treatment condition and week of training (error bars are for 95% confidence intervals).

In terms of between-participant effects for ESS score (Table 35), there was a significant fixed effect of treatment condition, with an estimated marginal mean ESS score of 8.978 (95% CI: 8.297, 9.659) in the intervention group versus 11.242 (95% CI: 10.595, 11.888) in the comparison group.

Table 35. Between-participant effects for Epworth Sleepiness Scale score.

Source	MS	df	*F*	*p*	η^2
Condition	503.762	1	21.635	<0.001*	0.068
Chronotype	104.965	2	4.508	0.012*	0.030
Condition x Chronotype	3.886	2	0.167	0.846	0.001
Age	0.156	1	0.007	0.935	<0.001
Body mass index	4.916	1	0.211	0.646	0.001
Caffeine use (referent no)	5.897	1	0.253	0.615	0.001
Component (referent regular)	20.799	1	0.893	0.345	0.003
Exercise frequency	14.138	1	0.607	0.436	0.002
Firearm use (referent no)	17.778	1	0.764	0.383	0.003
GT score	70.499	1	3.028	0.083	0.010
NEO-FFI					
Neuroticism	27.178	1	1.167	0.281	0.004
Extraversion	34.900	1	1.499	0.222	0.005
Openness to experience	29.898	1	1.284	0.258	0.004
Agreeableness	13.613	1	0.585	0.445	0.002
Conscientiousness	12.016	1	0.516	0.473	0.002
RSES	49.023	1	2.105	0.148	0.007
Sex (referent male)	345.942	1	14.857	<0.001*	0.048
Tobacco use	96.270	1	4.135	0.043*	0.014
Error	23.285	296			

*Significant at ≤ 0.05 level.
Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

There was also a significant fixed effect of chronotype (Figure 48), with the pairwise difference in ESS score occurring between evening and morning chronotypes ($p = 0.009$). There was no significant interaction effect for ESS score between treatment condition and chronotype. Significant covariates included sex and tobacco use, with females and smokers reporting greater sleepiness.

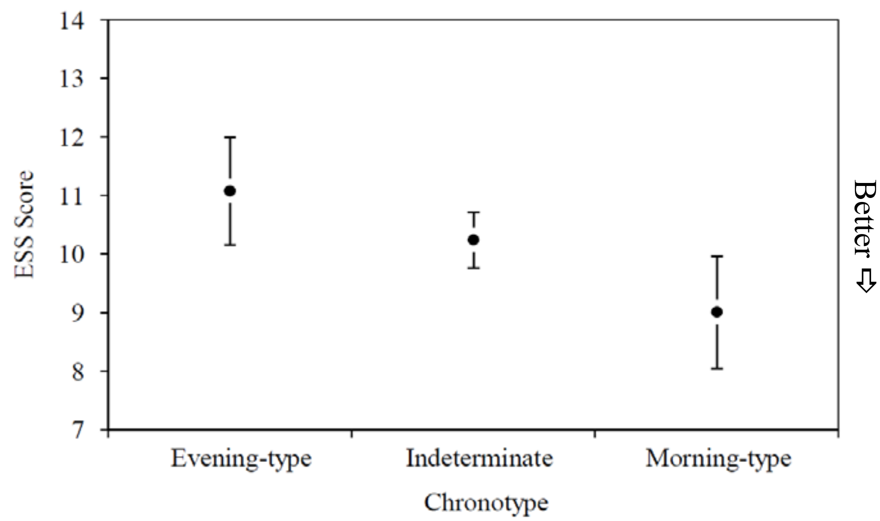


Figure 48. Estimated marginal means for ESS score by chronotype (error bars are for 95% confidence intervals).

Scores above ten on the ESS are indicative of excessive sleepiness and are a cause for concern with respect to performance (Miller, 2006). Applying this standard to our study sample, the odds ratio for a participant reporting excessive sleepiness being in the comparison relative to the intervention group was 1.198 (95% CI: 0.765, 1.874) prior to training and 2.331 (95% CI: 1.478, 3.679) at the completion of training. There was no difference in the odds of participants in the intervention and comparison groups being excessively sleepy at the start of training. However, participants in the comparison group were approximately 1.5 to 3.5 times more likely to be excessively sleepy by the conclusion of training, which was indicative of their sleep debt accrual throughout the course of Basic Combat Training.

2. Pittsburgh Sleep Quality Index (PSQI)

In terms of within-participant effects of PSQI score (Table 36), there was no significant fixed effect of time, nor was there a significant interaction effect between time and chronotype.

Table 36. Within-participant effects for Pittsburgh Sleep Quality Index score.

Source	MS	df	*F*	*p*	η^2
Condition	503.762	1	21.635	<0.001*	0.068
Chronotype	104.965	2	4.508	0.012*	0.030

Condition x Chronotype	3.886	2	0.167	0.846	0.001
Age	0.156	1	0.007	0.935	<0.001
Body mass index	4.916	1	0.211	0.646	0.001
Caffeine use (referent no)	5.897	1	0.253	0.615	0.001
Component (referent regular)	20.799	1	0.893	0.345	0.003
Exercise frequency	14.138	1	0.607	0.436	0.002
Firearm use (referent no)	17.778	1	0.764	0.383	0.003
GT score	70.499	1	3.028	0.083	0.010
NEO-FFI					
Neuroticism	27.178	1	1.167	0.281	0.004
Extraversion	34.900	1	1.499	0.222	0.005
Openness to experience	29.898	1	1.284	0.258	0.004
Agreeableness	13.613	1	0.585	0.445	0.002
Conscientiousness	12.016	1	0.516	0.473	0.002
RSES	49.023	1	2.105	0.148	0.007
Sex (referent male)	345.942	1	14.857	<0.001*	0.048
Tobacco use	96.270	1	4.135	0.043*	0.014
Error	23.285	296			

*Significant at ≤ 0.05 level.
Notes: GT score = General technical aptitude score; MS = Mean square; RSES = Response to Stressful Experiences Scale.

There were significant interaction effects of PSQI score between time and the fixed effect, treatment condition, as well as the covariate age. The interaction effect with treatment condition is shown in Figure 49. PSQI scores increased for participants in the comparison group and decreased for participants in the intervention group over the course of training, such that the groups mean scores differed significantly at the poststudy assessment.

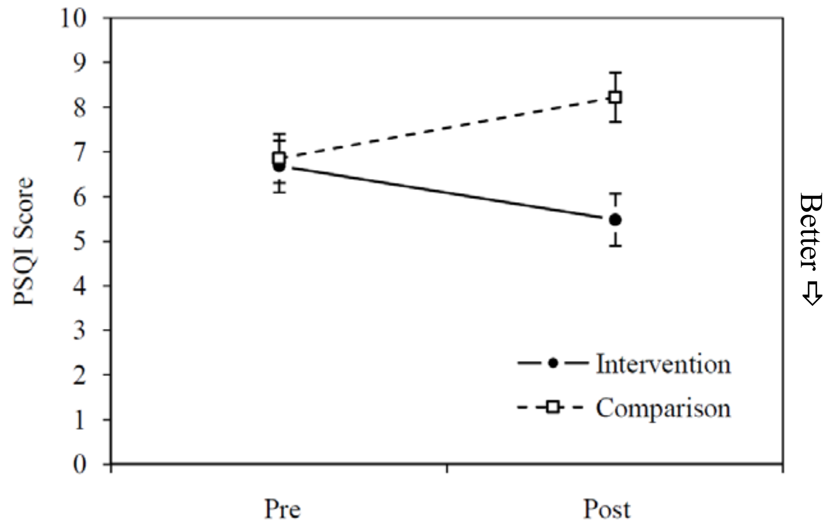


Figure 49. Estimated marginal means for PSQI score by treatment condition and pre/posttraining (error bars are for 95% confidence intervals).

In terms of between-participant effects of PSQI score (Table 37), there was a significant fixed effect of treatment condition, with an estimated marginal mean PSQI score of 6.082 (95% CI: 5.629, 6.536) in the intervention group versus 7.539 (95% CI: 7.109, 7.970) in the comparison group. There was no significant fixed effect of chronotype, nor was there a significant interaction effect between treatment condition and chronotype. Significant covariates included age and the NEO personality components of neuroticism, openness to experience, agreeableness, and conscientiousness scores.

Table 37. Between-participant effects for Pittsburgh Sleep Quality Index score.

Source	MS	df	F	p	η^2
Condition	208.769	1	20.244	<0.001*	0.064
Chronotype	9.839	2	0.954	0.386	0.006
Condition x Chronotype	9.636	2	0.934	0.394	0.006
Age	185.963	1	18.033	<0.001*	0.057
Body mass index	17.835	1	1.729	0.189	0.006
Caffeine use (referent no)	8.543	1	0.828	0.363	0.003
Component (referent regular)	1.432	1	0.139	0.710	<0.001
Exercise frequency	19.454	1	1.886	0.171	0.006
Firearm use (referent no)	30.064	1	2.915	0.089	0.010
GT score	33.465	1	3.245	0.073	0.011

NEO-FFI					
Neuroticism	97.425	1	9.447	0.002*	0.031
Extraversion	2.788	1	0.270	0.603	0.001
Openness to experience	89.635	1	8.692	0.003*	0.029
Agreeableness	180.261	1	17.480	<0.001*	0.056
Conscientiousness	47.638	1	4.619	0.032*	0.015
RSES	5.616	1	0.545	0.461	0.002
Sex (referent male)	17.329	1	1.680	0.196	0.006
Tobacco use	3.049	1	0.296	0.587	0.001
Error	10.312	296			

*Significant at ≤ 0.05 level.

Notes: GT score = General technical aptitude score; MS = Mean square; NEO-FFI = NEO Five-Factor Inventory; RSES = Response to Stressful Experiences Scale.

Scores above five on the PSQI are indicative of poor sleep quality. Applying this standard to our study sample, the odds ratio for a participant having poor quality sleep being in the comparison relative to the intervention group was 1.684 (95% CI: 1.106, 2.565) prior to training and 5.477 (95% CI: 3.343, 8.972) at the completion of training. Moreover, the odds of a participant having poor sleep quality decreased in the intervention group from pretraining (odds = 0.791; 95% CI: 0.659, 0.950) to posttraining (odds = 0.470; 95% CI: 0.377, 0.586). In contrast, the odds of a participant having poor sleep quality increased in the comparison group from pretraining (odds = 1.332; 95% CI: 1.047, 1.696) to posttraining (odds = 2.574; 95% CI: 1.889, 2.509).

3. Ordinal Sleep Ratings

Participants provided ordinal ratings of the adequacy of the sleep obtained by themselves and peers using a 5-item Likert scale. Figure 50 provides histograms of the participants’ ratings by treatment condition.

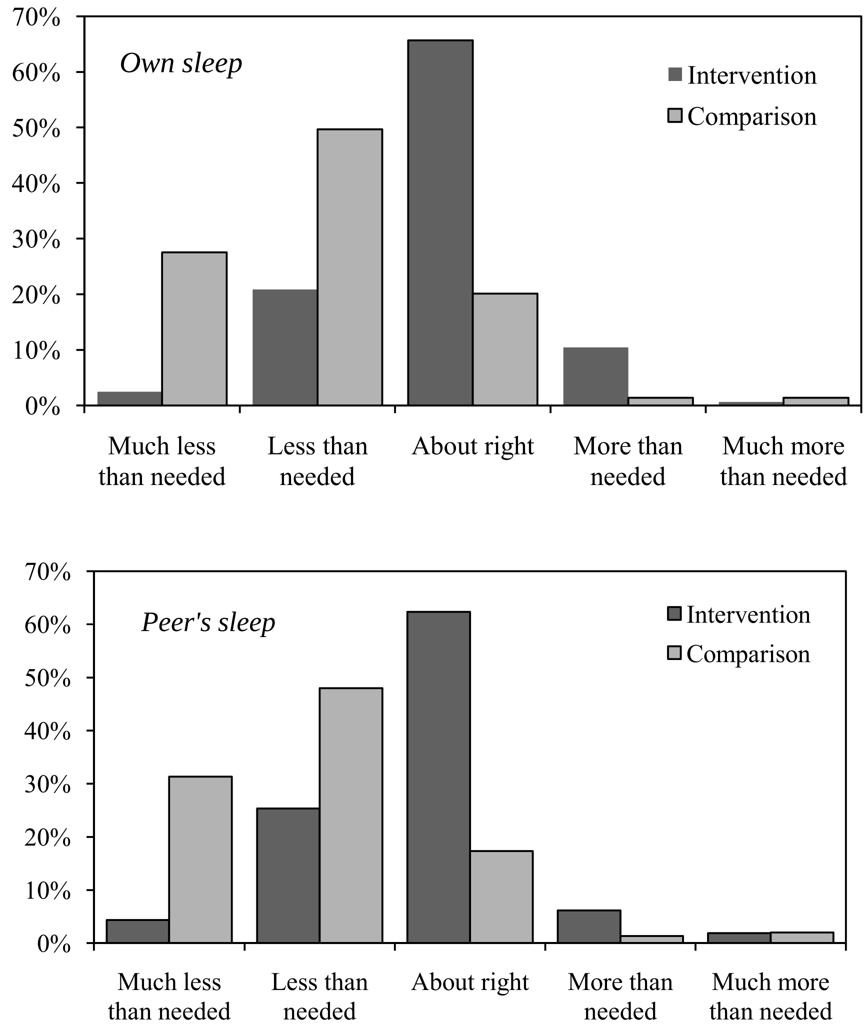


Figure 50. Histogram of participants' ratings of their own sleep (top) and their peers' sleep (bottom) by treatment condition.

The distributions of ratings by the comparison group were positively skewed versus those of the intervention group, which were symmetric unimodal. The mean rank for both ratings was higher for the intervention group than the comparison group: own sleep (intervention mean rank = 203.0, comparison mean rank = 110.5, Mann-Whitney U = 5164.5, $p < 0.001$) and peers' sleep (intervention mean rank = 198.6, comparison mean rank = 112.4, U = 5495.0, $p < 0.001$). There were small to moderate negative correlations between participants' ordinal ratings of the adequacy of their own sleep and their posttraining ESS ($\rho = -0.351$, $p < 0.001$) and PSQI scores ($\rho = -0.505$, $p < 0.001$).

Similarly, there was a negative correlation between participants' own sleep ratings and posttraining POMS total mood disturbance scores ($\rho = -0.370$, $p < 0.001$).

4. Frequency of Sleep During Activities

Participants were asked to report, on average, how often they fell asleep during activities such as classes, training, or lectures. Figure 51 provides a histogram of the participants' responses by treatment condition.

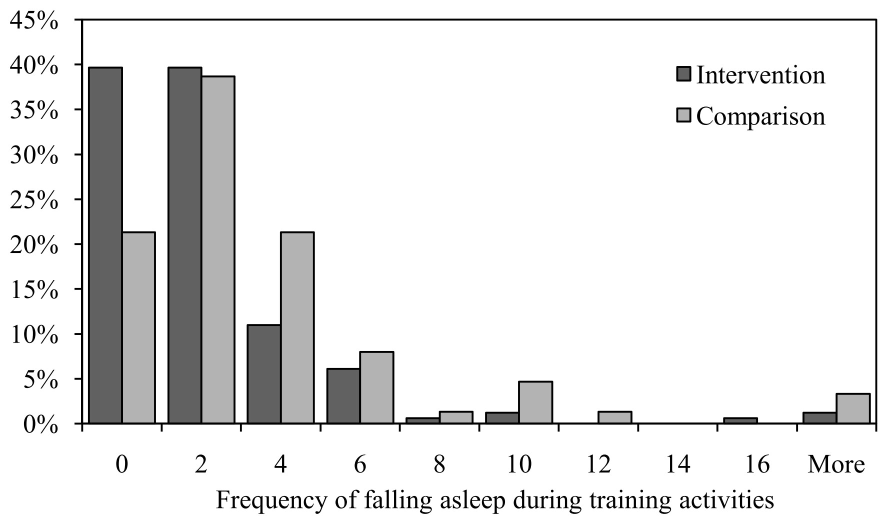


Figure 51. Histogram of daily frequency that participants report falling asleep during activities by treatment condition.

The distribution of responses for both groups was positively skewed, but the distribution for the intervention group was platykurtic at 0-2, while that of the comparison group was mesokurtic between 0-4. A comparison of mean ranks confirmed that participants in the intervention group reported significantly fewer episodes of falling asleep than those in the comparison group (intervention mean rank = 137.5, comparison mean rank = 179.4, Mann-Whitney U = 9011.0, $p < 0.001$). There was a small positive correlation between the frequency that participants fell asleep during activities and their post-training ESS ($\rho = 0.365$, $p < 0.001$) and PSQI scores ($\rho = 0.291$, $p < 0.001$). There was also a positive correlation between the frequency that participants fell asleep during activities the posttraining POMS total mood disturbance score ($\rho = 0.206$, $p < 0.001$). Additionally, there was a small negative correlation between a participant's ordinal rating of their sleep and the frequency with which they reported falling asleep during activities ($\rho = -0.250$, $p < 0.001$).

5. Preference in Timing of Physical Fitness Training

Participants were asked to indicate their preference for the best time of day for physical fitness training. Figure 52 provides a histogram of the participants' responses by treatment condition.

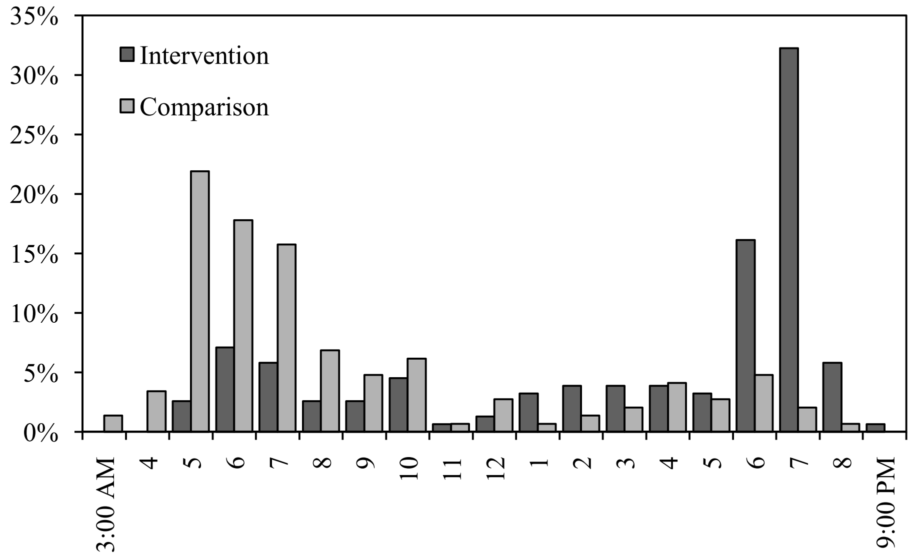


Figure 52. Histogram of participants' preferred time of the day for physical fitness training by treatment condition.

The distribution of responses for both groups was bimodal, with the primary peak occurring near the respectively scheduled company physical fitness training times. Hence, participants in the intervention group indicated they generally preferred to conduct physical fitness training in the evenings as per their training schedule. Similarly, participants in the comparison group preferred to conduct physical fitness training in the mornings as per their training schedule. There was a small, negative correlation between a participant's Morningness-Eveningness Questionnaire score and their time preference for physical fitness training ($\rho = -0.272$, $p < 0.001$). Thus, evening chronotype participants preferred physical fitness training in the evening and morning chronotype participants preferred training in the morning.

G. ATTRITION

It was of interest to determine how participants' likelihood of completing training related to treatment condition and other potential measured covariates. The databases submitted by each of the training companies indicated whether each participant successfully completed training. However, for those participants who did not complete training, the databases did not uniformly indicate when an attrition occurred and for what reason. Moreover, the final disposition of participants who did not graduate was not always determined, with some being separated from the Army, others on convalescent leave pending recovery from an injury or awaiting a physical evaluation board, and still others washing back to reaccomplish either portions of or the entire course of training. Additionally, participants who did not meet physical fitness standards could also be sent to a special training company to focus on further physical conditioning. Thus, a participant being classified as an attrite does not necessarily equate with them being lost to the Army. Accordingly, it was decided to analyze the likelihood of a participant not graduating with their initial training cohort using a simple binary logistic regression model and limiting the covariates to those measured during the initial study enrollment.

Overall, 35 (16.7%) participants in the intervention group failed to graduate with their cohort as compared to 33 (18.1%) participants in the comparison group, a nonsignificant difference $\left(\chi_1^2 = 0.130, p = 0.718\right)$. Table 38 shows the results for the fitted binary logistic regression model for failure to graduate. Accordingly, the odds ratios (ORs), calculated from the exponential of the estimated regression coefficients, should be interpreted in terms of the likelihood of failing to graduate with one's initial training cohort. The classification accuracy of the model was 83.9% using a cutoff of 0.5.

Table 38. Results for the fitted binary logistic regression model for failure to graduate with initial training cohort.

Analysis variables	Estimate	Standard error	df	Wald	*p*
Intercept	–7.387	1.317	1	31.450	<0.001
Body mass index	0.104	0.033	1	9.882	0.002

NEO-FFI neuroticism	0.039	0.017	1	5.507	0.019
POMS depression-dejection factor	0.024	0.011	1	4.651	0.031
Sex (referent male)	1.514	0.314	1	23.236	<0.001

There was no significant effect of treatment condition on the likelihood of failure to graduate. However, being female (OR = 4.545; 95% CI: 2.456, 8.411), increased body mass index (OR = 1.110; 95% CI: 1.1040, 1.184), higher scores of neuroticism as assessed using the NEO-FFI (OR = 1.040; 95% CI: 1.006, 1.074), and depressed mood or sense of inadequacy as measured on the POMS (OR = 1.024; 95% CI: 1.002, 1.046) were all associated with an increased likelihood of failure to graduate.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. DISCUSSION

Most studies of training effectiveness in military environments have concerned themselves primarily with activities that occur during the waking hours. They tend to examine the relationship between time expenditures in training using various modalities and measures of individual or system performance—the archetype being the classic transfer of training study. This study took a decidedly different approach, instead concerning itself primarily with the importance of the hours spent sleeping and their relation to measures of Soldier performance and other indicators of individual functioning during Basic Combat Training. Recognizing that adolescents comprise the majority of military accessions, this study evaluated the impact of accommodating adolescent alterations in sleeping and waking patterns. In particular, the scheduled timing of sleep during training was adjusted to account for the developmental phase delay of the circadian cycle in adolescents. The results of this study indicate that, even after controlling for factors contributing to individual differences, adjusting the scheduled sleep period in a phase-delayed direction was associated with increased daily total sleep and modest improvements in some indicators of daytime functioning. These findings suggest several operationally-relevant effects of accommodating adolescent sleep physiology that military planners may wish to consider in developing future training programs of instruction and associated training schedules.

A. ACTIGRAPHIC MEASURES OF SLEEP

Hypothesis 1 predicted that participants on the modified, phase-delayed sleep schedule would obtain more daily sleep than participants following the standard Basic Combat Training schedule. This hypothesis was supported with participants on the modified sleep schedule obtaining approximately 33 more minutes of total sleep per night than those on the standard sleep schedule. This finding is consistent with that of other studies, such as the School Transition Study (Carskadon, 2001), which have found that early start times are associated with truncated sleep in adolescents. The observed reduction in sleep with early start times is attributed to the developmental phase delay of the circadian cycle in adolescents, which makes it particularly difficult for adolescents to advance the evening retiring time in order to obtain an adequate amount of sleep. Additionally, Carskadon and colleagues (1998) have demonstrated that adolescents do not readily adapt or habituate their circadian cycle to early rising times, although the mechanism underlying this observation is not well understood. It is also interesting to note that a similar phenomenon has been described in adult shift workers with very early morning starts who tend to experience long sleep latencies when attempting to get compensatory sleep in the early evening (Rosa, 2001).

Thus, this study demonstrates that scheduling the sleep period for adolescents and young adults to better align with the phase delay in their circadian cycle results in a significant improvement in total daily sleep without any concomitant adjustment to the quantity of time scheduled for sleep. Regardless of differences in the timing of sleep between the two schedules, morning chronotype participants averaged approximately 15 minutes more sleep than those participants who were evening chronotype. This pattern is consistent with that described by Wolfson (2001) for adolescent students transitioning to a school with an earlier start time: evening chronotype students had more difficulty adjusting to the earlier start time and had less total sleep than did morning chronotype students. The implication is that even with the phase-delayed schedule used in this study, evening chronotype participants experienced greater difficulty adjusting to their new start time. This result is not surprising given the histograms of participants' self-reported wake times prior to Basic Combat Training, which suggest that the transition to military life necessitated earlier start times for the majority of participants. It is also worth noting that the average quantity of sleep obtained by participants was only approximately 60% of the 9.2 hours of daily sleep reportedly needed by adolescents (Mercer, Merritt, & Cowell, 1998; Wolfson, 2001). Lastly, the observation that sleep was reduced for participants using the modified schedule after the sixth week of training is an artifact caused by the commencement of the field exercise portion of Basic Combat Training.

B. MOOD STATES

Hypothesis 2 predicted that participants on the modified sleep schedule would have less decrement in mood state than participants following the standard Basic Combat

Training sleep schedule. There was weak support for this hypothesis based upon the analysis of the entire study sample, which necessarily excluded consideration of a total daily sleep variable in the models. Irrespective of treatment condition, the general trend was for participants to report decreased feelings of tension-anxiety, depression-dejection, fatigue-inertia, and confusion-bewilderment over the course of Basic Combat Training. Participants in the intervention group reported more stable feelings of anger-hostility and exhibited steadier total mood disturbance scores than participants in the comparison group. Participants in the intervention group also tended towards less anger-hostility and lower total mood disturbance scores relative to the comparison group early in training, although these differences declined during Basic Combat Training. Participants in the intervention group reported significantly greater feelings of vigor than those in the comparison group throughout training, but the effect size of treatment condition was very modest in this case. Overall, there was no evidence that characteristics of chronotype significantly affected participants' mood states.

There was partial support for Hypothesis 2, particularly with regards to the effects for the characteristics of chronotype on mood, when the analysis was restricted to the actigraphy subsample and a variable for total daily sleep was included in the models. Irrespective of treatment condition, evening chronotype participants reported more vigor throughout training than morning chronotype participants. However, evening chronotype participants in the intervention group exhibited less self-reported feelings of tension-anxiety, depression-dejection, anger-hostility, and confusion-bewilderment than their morning chronotype counterparts. The opposite pattern occurred in the comparison group, with evening chronotype participants reporting greater feelings of tension-anxiety, depression-dejection, anger-hostility, and confusion-bewilderment than their evening chronotype counterparts. In terms of total mood disturbance score, evening chronotype participants in the intervention group had lower scores than their morning chronotype counterparts, while a trend in the opposing direction was observed for participants in the comparison group. Taken together, these findings suggest that the phase-delayed sleep schedule preferentially impacted, in a positive direction, the mood state of evening chronotype participants. The operational significance of this finding is evident when one appreciates that the majority of military accessions are adolescents who, as a demographic group, tend to exhibit a biological predisposition for eveningness (Carskadon, 2001).

The rather modest impact of the sleep schedule intervention on subjective mood in this study contrasts with other research that has shown that manipulations of the duration and timing of sleep episodes can have marked impacts on mood (Birchler-Pedross, Schröder, Münch, Knoblauch, Blatter, Schnitzler-Sack, et al., 2009; Boivin, Czeisler, Dijk, Duffy, Folkard, Minors, et al., 1997; Danilenko, Cajochen, & Wirz-Justice, 2003; Monk, Buysse, Reynolds, Jarrett, & Kupfer, 1992; Selvi, Gulec, Agargun, & Besiroglu, 2007; Taub & Berger, 1973; Wood & Magnello, 1992). For example, Boivin and colleagues (1997) demonstrated that even moderate changes in the timing of the sleep-wake cycle led to profound effects on mood. Similarly, Danilenko, Cajochen and Wirz-Justice (2003) showed that advancing the sleep-wake cycle daily by just 20 minutes for a week led to significant decrements in subjective mood ratings relative to a control group with stable sleep. Interestingly, Selvi and colleagues (2007) showed that phase preference modified the effect of partial sleep deprivation on mood, with morning chronotypes exhibiting less sensitivity of mood. A pattern similar to that described by Selvi and colleagues was observed, at least for the subsample of the study population who had actigraphy data.

Several hypotheses are suggested to explain the small observed effect of the schedule intervention on subjective mood in this study. Mood is largely a function of situational factors (Chamorro-Premuzic, 2007) and the Basic Combat Training environment represents a complex milieu of such factors. Throughout Basic Combat Training, the military instructor cadre is working to actively shape and influence the mood state of their Soldiers as a means of achieving organizational training objectives. Many factors, such as leader-subordinate and peer-to-peer dynamics, unit morale, and individual perceptions of acute physical and mental stressors, likely contributed to differences in subjective mood among participants. Given the aggregate of observed and unobserved factors in this study, the relationship between sleep and subjective mood was most likely reduced to having a small, but still measurable, effect size. Additionally, while the phase-delayed sleep schedule resulted in increased total daily sleep for participants in the intervention group, the shortfall in daily sleep relative to known adolescent sleep needs for both groups was still large (i.e., on the order of 3-4 hours). Consequently, participants in both groups may have had a significant partial sleep deprivation that then blunted the observed effect of the schedule intervention. Finally, the phase-delayed sleep schedule, while a marked improvement over the standard Basic Combat Training sleep schedule in terms of accommodating adolescent sleep-wake patterns, was still significantly out of phase with participants' baseline patterns as inferred from participant responses on the pretraining PSQI. Such an assertion is supported by Carskadon's (2001) study of adolescent students, which found that school start times around 7:00 a.m. were difficult for adolescent students, and students tended to do better when start times were delayed until 8:00 a.m. or later.

C. BASIC RIFLE MARKSMANSHIP

Hypothesis 3 predicted that participants on the modified sleep schedule would exhibit greater improvement in basic rifle marksmanship scores than those following the standard Basic Combat Training sleep schedule. This hypothesis was supported by the study results, although the analysis of marksmanship performance turned out to be far from straightforward given differences between training companies in initial performance on the first record fire and variability in the number of record fires accomplished by each participant. Despite all this variability, however, it was possible to demonstrate that the degree of improvement in marksmanship performance over the serial record fires was significantly predicted, in part, by a sleep-related variable. Moreover, the effect size of sleep, while relatively small, was still greater than that attributable to prior experience with firearms.

It is noteworthy that sleep during the week preceding the record fires, when basic marksmanship tasks and subtasks were being learned, was more strongly correlated with subsequent performance than sleep during the week of the record fires. This suggests the possibility that sleep was acting as a modifier of training effectiveness. Such an assertion is consistent with research showing that procedural memories improve with subsequent early slow wave sleep (SWS) and late rapid eye movement (REM) sleep, although there is some debate regarding the relative importance of the various stages of sleep. Nevertheless, increasing evidence supports the role of sleep in memory consolidation and latent learning (Fenn, Nusbaum, & Margoliash, 2003; Gais, Plihal, Wagner, & Born, 2000; Karni, Tanne, Rubenstein, Askenasy, & Sagi, 1994; Stickgold, James, & Hobson, 2000; Walker, Brakefield, Hobson, & Stickgold, 2003; Wilson & McNaughton, 1994). For example, Gais and colleagues (2000) observed that memories are, on average, more than three times improved after sleep containing both SWS and REM sleep than after a period of early sleep alone. Thus, the phase-delayed schedule, which was associated with increased total daily sleep, likely increased the opportunity for late REM sleep and thereby potentiated the learning and recall of marksmanship skills.

D. PHYSICAL FITNESS

Hypothesis 4 predicted that participants on the modified sleep schedule would exhibit greater improvement in physical fitness scores than participants following the standard Basic Combat Training sleep schedule. This hypothesis was not supported by the study results. As in the case of the marksmanship data, the use of nonrandomized groups led to significant baseline differences between the intervention and comparison groups, with the intervention group exhibiting higher physical fitness scores early in training. However, these differences diminished over the course of training such that the groups were equivalent on the final physical fitness assessment. Thus, the overall pattern suggested a regression to the mean phenomenon—an assertion that is supported by the absence of any correlation between fitness scores and average total daily sleep for participants in the actigraphy subsample. On the flip side, altering the timing of physical fitness training to accommodate the change in timing of sleep did not appear to harm the performance of participants in the intervention group. Additionally, participants in the intervention group generally expressed a preference for the later timing of their physical fitness training, while participants in the comparison group, on average, preferred the earlier timing of their physical fitness training.

These findings are consistent with that reported in the scientific literature examining the effect of sleep deprivation on exercise performance. Studies of exercise performance after periods of sleep deprivation of up to 72 hours have consistently demonstrated that muscle strength and exercise performance are not affected (Martin, 1981; Martin & Gaddis, 1981; Reilly & Deykin, 1983; Van Helder & Radomski, 1989). While Martin (1981) was able to show that sleep loss reduced work time to exhaustion by an average of 11 percent, this change was attributed to the psychological effects of acute sleep debt because subjects' ratings of exertion were dissociated from any cardiovascular changes. A smaller body of research has also examined the influence of chronotype on diurnal changes in muscle strength. For example, Tamm, Lagerquist, Ley, and Collins (2009) found that evening chronotype individuals could produce a stronger maximum voluntary muscle contraction in the evening, while morning chronotype individuals exhibited no significant change in strength throughout the day. However, the results of this study failed to show any significant effect of chronotype for the strength-based fitness assessments.

E. SLEEP HYGIENE

Hypothesis 5 predicted that for participants whose sleep schedules were modified, the odds of reporting occupationally significant fatigue (defined as an Epworth Sleepiness Scale (ESS) score greater than ten) would be lower than that for participants following the standard Basic Combat Training sleep schedule. This hypothesis was supported by the study results, with participants in the comparison group being 2.3 times more likely to have occupationally significant fatigue at the end of training—a finding with important safety and health implications. At the beginning of the study, participants in the intervention and comparison groups had comparable subjective sleepiness as assessed based on ESS scores. Over the course of training, participants in the comparison group exhibited a significant increase in reported sleepiness, while those in the intervention group reported no change in subjective sleepiness. Overall, evening chronotype participants reported greater sleepiness than morning chronotype participants. This result suggests that the modified sleep schedule, while an improvement over the standard schedule, still did not fully accommodate the developmental phase-delay of the adolescent and young adult circadian cycle.

Hypothesis 6 predicted that for participants whose sleep schedules were modified, the odds of reporting poor sleep quality (defined as Pittsburgh Sleep Quality Index

(PSQI) score greater than five) would be lower than that for participants following the standard Basic Combat Training sleep schedule. This hypothesis was supported by the study results, with participants in the comparison group being 5.5 times more likely to report poor sleep quality at the end of training. Participants in the intervention and comparison groups had comparable sleep quality as assessed based on PSQI score at the start of the study. Over the course of training, participants in the comparison group exhibited a significant degradation in sleep quality, while those in the intervention group exhibited a trend towards improved sleep quality. Additionally, the odds of participants reporting poor quality sleep actually decreased for those in the intervention group relative to the start of the study. This finding suggests that the phase-delayed sleep schedule was an improvement over participants' baseline sleep schedule—or in other words, Basic Combat Training actually improved the sleep hygiene of participants in the intervention group.

To summarize, participants in the intervention group graduating from Basic Combat Training did so in a better physiological state than their counterparts in the comparison group. The operational significance of this finding can be inferred from research on school age adolescents linking sleep patterns and academic performance (Acebo & Carskadon, 2001; Wolfson & Carskadon, 2003). Thus, participants in the intervention group, by way of having improved wake-sleep patterns and increased total daily sleep, were better prepared to undertake the more academically rigorous secondary military occupation-specific training that follows Basic Combat Training. Additionally, they can be expected to be at lower risk for future lost training days or injuries (Acebo, Wolfson, & Carskadon, 1997).

F. ATTRITION

Hypothesis 7 predicted that for participants on the modified sleep schedule, the odds of attriting from training would be lower than that for participants following the standard Basic Combat Training sleep schedule. This hypothesis was not supported by the study results as evidenced by the absence of treatment condition in the final logistic model for attrition. The single largest risk factor for attrition was sex with females more likely to attrite, followed by body mass index (i.e., fitness), neurotic personality characteristics, and depressed subjective mood. Given that the frequency of attrition relative to time was positively skewed—that is, most attrition tends to occur earlier rather than later in training—it is more likely that preexisting conditions or vulnerabilities were the predominant determinant of attrition.

THIS PAGE INTENTIONALLY LEFT BLANK

V. SELECT HUMAN SYSTEMS INTEGRATION ANALYSES

A. THE HUMAN SYSTEMS INTEGRATION PARADIGM

Up to this point, we have described a research study that was conducted from the behavioral sciences paradigm utilizing an experimental methodology and multivariable statistical techniques drawn from experimental psychology. We proposed a series of research hypotheses and developed corresponding statistical models to aid decision making with regards to our accepting or rejecting those research hypotheses (and conversely their corollary null hypotheses). However, if we are to transition from the behavioral sciences to the HSI paradigm, we need to take a subset of research hypotheses that were accepted based upon the statistical models and reformulate those that are of most interest to us in terms of trade-off functions, thereby making possible their direct incorporation in the "system analytic thinking process" (Weisz, 1967, p. 3). The latter is involved whenever there is a choice between various alternative system mixes to meet a particular requirement or threat. Historically, systems analysis has been dominated by mathematically based operations research techniques developed to facilitate the decision making of organizational planners and systems developers (Hughes, 1998). Consequently, the objective of our forthcoming HSI analyses is the development of mathematical trade-off functions that can then be used by decision makers to predict the optimum mix of human performance determinants, whether in terms of cost, effectiveness, or technical feasibility (Weisz, 1967, 1968). This objective will be accomplished using the isoperformance methodology proposed by Jones and Kennedy (Jones & Kennedy, 1996; Jones, 2000). In so doing, we establish the pattern by which human factors research and human considerations can be appropriately represented in systems analyses.

B. BASIC RIFLE MARKSMANSHIP MODEL

The purpose of this section is to develop, in a step-by-step fashion, an isoperformance curve for basic rifle marksmanship. We start with a model, a criterion level, and a confidence level. The model states the functional dependence of marksmanship performance on aptitude and average daily sleep. The criterion indicates the minimal level of performance that one is willing to regard as adequate. The confidence level is the probability of adequate performance, by which we mean that performance will equal or exceed the criterion. What results is essentially a trade-off function for marksmanship in terms of the personnel (i.e., aptitude) and survivability (i.e., fatigue) domains of HSI.

Our first step is to obtain an expression for a model for the expected marksmanship performance for an individual Soldier, *i*. As will be recalled from our earlier analysis of the basic rifle marksmanship data for the actigraphy subsample, participants in the intervention group tended to have lower initial marksmanship scores relative to participants in the comparison group, but they also exhibited a greater improvement in marksmanship performance over serial firings. Additionally, the magnitude of this change was positively correlated with average daily sleep during the week prior to the serial firings ($\rho = 0.341$, $p = 0.001$), which was when they received instruction in rifle marksmanship fundamentals. Moreover, there was no effect of group when sleep was included in the analysis, implying that differences in instructor cadre were not a likely explanation for the observed difference in basic rifle marksmanship. Consequently, we propose the following model for the basic rifle marksmanship data:

$$\Delta \mathrm{S}_i = a + b\left(\mathrm{SLP}_i\right) + \varepsilon_i \qquad (1)$$

where $\Delta \mathrm{S}_i$ is the difference between first and last serial marksmanship scores for the i^{th} Soldier, and SLP_i is the i^{th} Soldier's average daily sleep during the week prior to the serial firings. The constants, *a* and *b*, are parameters estimated during the model fitting and ε_i is a normally distributed error term, with mean equal to zero and variance equal to σ_ε^2.

Table 39 presents a conventional readout for the model in terms of expected mean squares, F ratio, significance level, and effect size. The result is that average daily sleep is a significant determinant of $\Delta \mathrm{S}$, explaining nearly 11% of the variance in the change in marksmanship scores. While average daily sleep has a relatively modest effect on marksmanship performance, it is a determinant that is, at least in Basic Combat Training, controllable by the Army.

Table 39. Expected mean squares, F ratio, significance level, and effect size for the basic rifle marksmanship data from the actigraphy subsample.

Source	MS	df	F	p	η^2
Sleep	412.190	1	11.329	0.001	0.116
Error	36.382	86	—	—	—

We can rewrite Equation 1 as follows:

$$E\left[\Delta S_i\right] = a + b\left(\text{SLP}_i\right) \quad (2)$$

The only difference between the right side of this equation and that of the full model is the absence of the error term. Hence, the expected change in marksmanship performance for the i^{th} Soldier depends only on the determinant SLP_i. The next step is to modify the model so that the left-hand side of Equation 2 is in terms of the excepted final marksmanship score. We begin by noting that $\Delta S_i = S_{2i} - S_{1i}$, where S_{1i} is a Soldier's initial marksmanship score and S_{2i} is their final marksmanship score. Accordingly, we rewrite Equation 2:

$$E\left[S_{2i} - S_{1i}\right] = a + b\left(\text{SLP}_i\right). \quad (3)$$

Since expectation of a difference is simply the difference of expectations:

$$E\left[S_{2i}\right] - E\left[S_{1i}\right] = a + b\left(\text{SLP}_i\right). \quad (4)$$

Rearranging terms:

$$E\left[S_{2i}\right] = E\left[S_{1i}\right] + a + b\left(\text{SLP}_i\right). \quad (5)$$

We next propose replacing the $E\left[S_{1i}\right]$ term with $E\left[S_{1j}\right]$, which is the expectation of the initial marksmanship score for a Soldier in the j^{th} quintile for initial marksmanship performance. Consequently, Equation 5 becomes

$$E\left[S_{2ij}\right] = E\left[S_{1j}\right] + a + b\left(\text{SLP}_{ij}\right), \quad (6)$$

which requires that we recalculate σ_ε^2. It is observed that the penalty for this change is small, with σ_ε^2 now equal to 37.049 as compared to 36.382 previously.

We explain further, since it may not be intuitive why we have proceeded through the following model development steps rather than simply fitting a model directly using S_{1ij}, SLP_{ij}, and S_{2ij}. To start, it was observed that there was a strong multicollinearity between S_{1ij} and SLP_{ij}, which complicates attempts at regression analysis. Another nontrivial problem encountered in this study was the finding that the intervention and comparison groups differed in terms of initial marksmanship performance and, hence, aptitude—an observation that can be attributed to the use of nonrandomly formed groups in the study design. Since the intervention group, which obtained more sleep by study design, had worse initial marksmanship performance, sleep is negatively correlated with initial marksmanship scores (i.e., the effect of sleep was confounded by group differences in aptitude). However, as we showed earlier in this section, sleep is also positively correlated with improvement in serial marksmanship scores irrespective of group. These are contradictory findings. If sleep did indeed have a negative effect on initial marksmanship performance, it would be expected to have a negative effect on serial marksmanship performance as well—but exactly the opposite was observed. Thus, we focused on fitting the latter relationship to minimize potential confounding by the former. In the end, however, we still need to express the model dependent variable in terms of final marksmanship scores, as this is the performance criterion used by the Army.

The second step in developing the isoperformance curve is to determine what expected performance for the i^{th} Soldier in the j^{th} quintile must be if the probability of adequate performance is to equal a specified confidence interval. In our case, the Army has specified a final marksmanship score of 23 as the criterion, and we will presuppose 0.80 is the desired confidence level. These specifications are met if the expected performance for the i^{th} Soldier in the j^{th} quintile is

$$E\left[S_{2ij}\right] = 23 + z\sigma_{\varepsilon}, \tag{7}$$

where z equals 0.84 from tables of the normal curve and $\sigma_{\varepsilon} = \sqrt{37.049}$ (see prior paragraphs). Hence,

$$E\left[S_{2ij}\right] = 23 + 0.84\sqrt{37.049} = 28.11. \tag{8}$$

If the final marksmanship score for the i^{th} Soldier in the j^{th} quintile is to equal or exceed 23 with a probability of 0.80, then the expected final marksmanship score for the Soldier must equal 28.11.

The third and last step is to put Equations 6 and 8 together. Doing so produces

$$28.11 = E\left[\mathrm{S}_{1j}\right] + a + b\left(\mathrm{SLP}_{ij}\right). \tag{9}$$

Equation 9 involves two model parameters (a and b), five sample statistics $\left(E\left[\mathrm{S}_{1j}\right]\right)$, and the determinants SLP_i and quintile j, the latter corresponding to a choice of aptitude level. Rearranging terms so that SLP_{ij} is on the left-hand side, one obtains

$$\mathrm{SLP}_{ij} = \frac{28.11 - E\left[\mathrm{S}_{1j}\right] - a}{b}. \tag{10}$$

The estimated values for the model parameters and sample statistics in Equation 10 are given below:

$$\begin{aligned} a &= -19.052 \quad & E\left[S_{1,1}\right] &= 12.250 \\ b &= 3.861 \quad & E\left[S_{1,2}\right] &= 18.000 \\ & & E\left[S_{1,3}\right] &= 23.579 \\ & & E\left[S_{1,4}\right] &= 27.056 \\ & & E\left[S_{1,5}\right] &= 31.053 \end{aligned}$$

This is the basic rifle marksmanship isoperformance curve. For any given choice of aptitude quintile, j, one can now calculate a value of SLP_{ij} such that the two together produce adequate performance with the specified level of confidence.

Figure 53 presents three isoperformance curves that trade off aptitude, as assessed based upon initial marksmanship score, and average daily sleep. The criterion is set at 23 (i.e., the minimum marksmanship qualification threshold), 27, and 30 (i.e., the sharp shooter qualification threshold). Each isoperformance curve traces combinations of aptitude and average daily sleep that yield equivalent performance in terms of the criterion, which in this case is final marksmanship score. Thus, these isoperformance curves can be read as trade-off functions. For example, Soldiers sleeping 7.55 hours per day will meet the basic rifle marksmanship qualification threshold of a final score of 23 if their initial marksmanship score is at least 18. Alternatively, if Soldiers are allowed to sleep for only 6.77 hours per day, then their initial marksmanship score will need to be at least 21 if they are to achieve the basic rifle marksmanship criterion on their final record fire. In other words, it takes one point in marksmanship aptitude to make up for each 16 minute reduction in Soldiers' average daily sleep during marksmanship instruction.

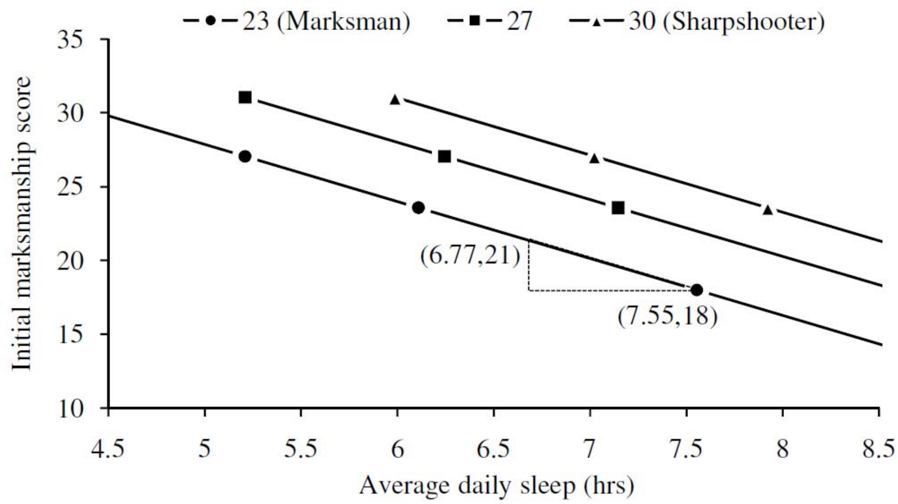


Figure 53. Isoperformance curves trading off aptitude, expressed as initial marksmanship score, and average daily sleep, setting the final marksmanship score criterion levels at 23, 27, and 30 and percentage proficient at 80%.

C. SLEEP QUALITY MODEL

Repeating the process used to create the marksmanship isoperformance model, we next develop an isoperformance curve for posttraining sleep quality as assessed using the PSQI. Again, we start with a model, a criterion level, and a confidence level. The model states the functional dependence of posttraining sleep quality on pretraining sleep quality and average daily sleep during training. Since sleep quality is an important clinical construct and poor sleep quality is a significant symptom of many medical, psychiatric, and sleep disorders (Buysse et al., 1989), we utilize PSQI scores as a metric for the occupational health domain of HSI. In terms of a criterion, a global PSQI score of greater than 5 was shown by Buysse and colleagues (1989) to have a 90% diagnostic sensitivity in distinguishing good sleepers (i.e., healthy individuals) from poor sleepers (i.e., individuals with mood or sleep disorders). What results is essentially a trade-off function in terms of the personnel (i.e., individuals' baseline sleep quality) and survivability (i.e., fatigue) domains of HSI.

Our first step is to obtain an expression for a model of the expected posttraining sleep quality of an individual Soldier, *i*. We propose the following model for the posttraining PSQI data:

$$\text{PSQI}_{2i} = a + b\left(\text{PSQI}_{1i}\right) + c\left(\text{SLP}_i\right) + \varepsilon_i, \tag{11}$$

where PSQI_{2i} is the posttraining PSQI score for the i^{th} Soldier, PSQI_{1i} is the i^{th} Soldier's baseline PSQI score prior to starting training, and SLP_i is the i^{th} Soldier's average daily sleep during training. The constants, *a*, *b*, and *c*, are parameters estimated during the model fitting and ε_i is a normally distributed error term with mean equal to zero and variance equal to σ_ε^2. Table 40 presents a conventional readout for the model in terms of expected mean squares, F ratio, significance level, and effect size. The result is that both baseline sleep quality and average daily sleep are significant determinants of PSQI_2.

Table 40. Expected mean squares, F ratio, significance level, and effect size for the post-training PSQI score data from the actigraphy subsample.

Source	MS	df	*F*	*p*	η^2
Sleep	412.190	1	11.329	0.001	0.116
Error	36.382	86	—	—	—

We can rewrite Equation 11 as follows:

$$E\left[\text{PSQI}_{2i}\right] = a + b\left(\text{PSQI}_{1i}\right) + c\left(\text{SLP}_i\right) \tag{12}$$

The only difference between the right side of this equation and that of the full model is the absence of the error term. Hence, the expected posttraining PSQI score for the i^{th} Soldier depends only on the determinants PSQI_{1i} and SLP_i.

The second step in developing the isoperformance curve is to determine what the expected posttraining PSQI score for the i^{th} Soldier must be if the probability of adequate sleep quality is to equal a specified confidence interval. In this case, we use the cutoff global PSQI score of 5 suggested by Buysse and colleagues (1989) as the criterion, and we will presuppose 0.80 is the desired confidence level. These specifications are met if the expected posttraining PSQI score for the i^{th} Soldier is

$$E\left[\text{PSQI}_{2i}\right] = 5 - z_{0.80}\sigma_{\varepsilon}, \tag{13}$$

where z equals 0.84 from tables of the normal curve and $\sigma_{\varepsilon} = \sqrt{9.183}$. Hence,

$$E\left[\text{PSQI}_{2i}\right] = 5 - 0.84\sqrt{9.183} = 2.455. \tag{14}$$

If the posttraining PSQI score for the i^{th} Soldier is less than or equal to 5 with a probability of 0.80, then the expected posttraining PSQI score for the Soldier must equal 2.455.

The third and last step is to put Equations 12 and 14 together. Doing so produces

$$2.455 = a + b\left(\text{PSQI}_{1i}\right) + c\left(\text{SLP}_{i}\right). \tag{15}$$

Rearranging terms so that SLP_i is on the left-hand side, one obtains

$$\text{SLP}_i = \frac{2.455 - a - b\left(\text{PSQI}_{1i}\right)}{c}. \tag{16}$$

The estimated values for the model parameters in Equation 16 are given below:

$$a = 16.129$$
$$b = 0.296$$
$$c = -2.053$$

This is the posttraining sleep quality isoperformance curve. For any given choice of baseline sleep quality, PSQI_{1i}, one can now calculate a value of SLP_i such that the two together produce adequate posttraining sleep quality (i.e., occupational health) with the specified level of confidence.

Figure 54 presents two isoperformance curves that trade off baseline sleep quality and average daily sleep during training. The criterion is set at 5, the clinical threshold for healthy individuals, and 6.5, the average baseline PSQI score in the study sample. The latter criterion setting represents the option of "doing no harm"—that is, not further exacerbating the sleep quality of already poor sleepers. Each isoperformance curve traces combinations of baseline sleep quality and average daily sleep that yield equivalent performance in terms of the criterion, which in this case is posttraining sleep quality. Consequently, these isoperformance curves can be read as trade-off functions. For example, Soldiers with poor baseline sleep quality (e.g., PSQI = 9.1) can obtain good sleep quality if they are provided 7.98 hours of sleep per night during training. Alternatively, if Soldiers are allowed to sleep for only 7.22 hours per day, then their baseline sleep quality will need to be fairly good (e.g., PSQI = 3.9) if they are to achieve the posttraining sleep quality criterion. In other words, it takes one point in baseline PSQI score to make up for each 9 minutes reduction in Soldiers' average daily sleep during training.

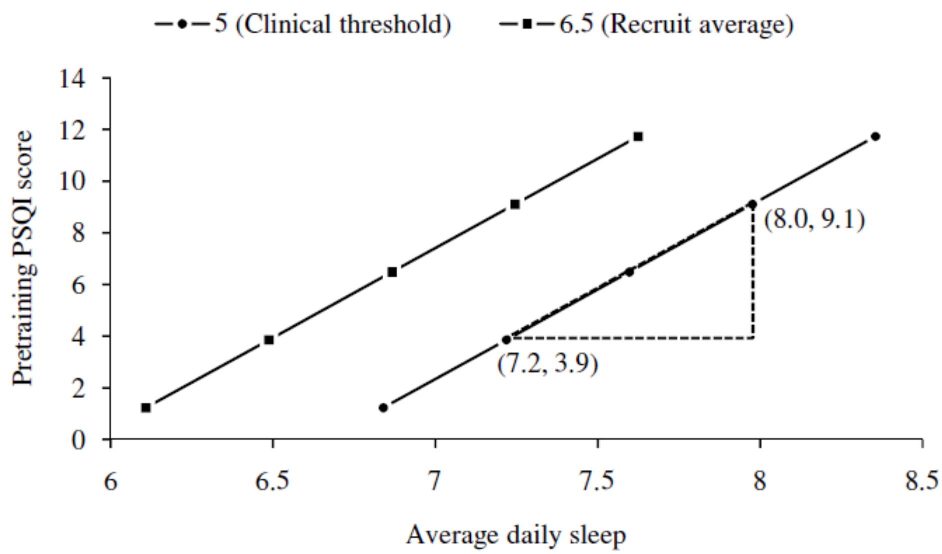


Figure 54. Isoperformance curves trading off baseline sleep quality, expressed as pretraining PSQI score, and average daily sleep, setting the final PSQI score criterion levels at 5 and 6.5 and the assurance level at 80%.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

In summary, increasing sleep and concomitantly decreasing fatigue had a small, but measurable, influence on various indicators of Soldier functioning, even after controlling for a variety of factors that affect performance. Although Soldiers' responses to the phase-delayed schedule intervention were relatively modest, it should be appreciated that the majority of outcome measures in Basic Combat Training are not highly sensitive to the effects of fatigue. Thus, the most important finding of the study may be the impact of the schedule intervention on sleep quality during Basic Combat Training—that is, Soldiers completing Basic Combat Training using the phase-delayed sleep schedule had significant improvements in sleep hygiene, such that they graduated from training in a better physiological state than when they started. Or, in other words, the phase-delayed sleep schedule allowed Soldiers to accomplish the training objectives of Basic Combat Training at a lower cost in terms of their sleep reservoir, thereby leaving them with a greater available cognitive work capacity going forward for subsequent training. The significance of this finding may not be fully appreciated until Soldiers' subsequent performance is assessed during the more cognitively demanding secondary military occupational specialty training courses—a recommendation for follow-up research related to this work.

While insufficient sleep and the consequent fatigue is a recognized problem in our society, concern has mainly been voiced around well-publicized, high-cost disasters resulting from the degraded occupational performance of sleep-deprived adults. The role of sleep in less dramatic circumstances seems to be underappreciated, particularly in the military environment where inadequate sleep is considered part and parcel of the routine starting in basic military training and onward. To the extent that adolescents and young adults entering the Army are unable to obtain sufficient sleep at the appropriate time to facilitate their primary developmental task—that being to master core Soldiering skills and incorporate Army values within their evolving self-identity—there are potentially significant hidden lost opportunity costs being borne by the Army. Our HSI trade-off analyses, derived from the results of a behavioral sciences experiment involving a simple sleep schedule intervention, provide an empirical foundation to begin quantitatively assessing the contribution of sleep to Soldier well-being and performance. What should then emerge is a world view that considers the human sleep reservoir in terms of its contribution to the performance of the human component of weapon systems or the human as a weapon system. Accordingly, the quantity and quality of sleep become limited resource variables that can and must be considered as part of the human factors contribution to systems analyses.

LIST OF REFERENCES

Acebo, C., & Carskadon, M. A. (2001). Influence of irregular sleep patterns on waking behavior. In M. A. Carskadon (Ed.), *Adolescent sleep patterns: Biological*, *social*, *and psychological influences* (chap. 13, pp. 220-235). New York, NY: Cambridge University Press.

Acebo, C., Wolfson, A. R., & Carskadon, M. A. (1997). Relations among self-reported sleep patterns, health, and injuries in adolescents. *Sleep Research*, 26, 149.

Anch, A. M., Browman, C. P., Mitler, M., & Walsh, J. K. (1988). *Sleep: A scientific perspective*. Englewood Cliffs, NJ: Pearson Prentice-Hall, Inc.

Andrews, C. H. (2004). *The relationship between sleep regimen and performance in United States Navy recruits*. Monterey, CA: Naval Postgraduate School.

Baldus, B. R. (2002). *Sleep patterns in U.S. Navy recruits: An assessment of the impact of changing sleep regimens*. Monterey, CA: Maters Thesis at Naval Postgraduate School.

Belenky, G., Wesensten, N. J., Thorne, D. R., Thomas, M. L., Sing, H. C., Redmond, D. P., et al. (2003). Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study. *Journal of Sleep Research*, 12, 1-12.

Birchler-Pedross, A., Schröder, C. M., Münch, M., Knoblauch, V., Blatter, K., Schnitzler-Sack, C., Wirz-Justice, A., & Cajochen, C. (2009). Subjective well-being is modulated by circadian phase, sleep pressure, age, and gender. *Journal of Biological Rhythms*, *24*(3), 232-242.

Boivin, D. B., Czeisler, C. A., Dijk, D. J., Duffy, J. F., Folkard, S., Minors, D. S., Totterdell, P., & Waterhouse, J. M. (1997). Complex interaction of the sleep-wake cycle and circadian phase modulates mood in healthy subjects. *Archives of General Psychiatry*, 54, 145-152

Brick, J. M., & Kalton, G. (1996). Handling missing data in survey research. *Statistical Methods in Medical Research*, 5, 215-238.

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193-213.

Carskadon, M. A. (2001). Factors influencing sleep patterns of adolescents. In M. A. Carskadon (Ed.), *Adolescent sleep patterns: Biological*, *social*, *and psychological influences* (chap. 2, pp. 4-26). New York, NY: Cambridge University Press.

Carskadon, M. A., Acebo, C., Richardson, G. S., Tate, B. A., & Seifer, R. (1997). An approach to studying circadian rhythms of adolescent humans. *Journal of Biological Rhythms*, 12, 278-289.

Carskadon, M. A., Wolfson, A. R., Acebo, C., Tzischinsky, O., & Seifer, R. (1998). Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep*, *21*(8), 871-881.

Chamorro-Premuzic, T. (2007). *Personality and individual differences*. Malden, MA: Blackwell Publishing.

Colton, C. C. (1837). *Lacon, or many things in few words*. London, UK: Longman, Orme, Brown, Green, and Longmans.

Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) professional manual.* Odessa, FL: Psychological Assessment Resources.

Danilenko, K. V., Cajochen, C., & Wirz-Justice, A. (2003). Is sleep per se a zeitgeber in humans? *Journal of Biological Rhythms*, 18, 170-178.

Directorate Basic Combat Training Doctrine and Training Development (2008, March). *Program of instruction for basic combat training* (Version 5). Fort Jackson, SC: Author.

Driskell, J. E., Hughes, S. C., Willis, R. C., Cannon-Bowers, J., & Salas, E. (1991). *Stress*, *stressor*, *and decision-making* (Technical Report for the Naval Training Systems Center). Orlando, FL: Naval Training Systems Center.

Driskell, J.E., & Salas, E. (Eds.) (1996). *Stress and human performance*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature*, 425, 614-616.

Gais, S., Plihal, W., Wagner, U., & Born, J. (2000). Early sleep triggers memory for early visual discrimination skills. *Nature Neuroscience*, *3*(12), 1335-1339.

Graham, M. G. (Ed.). (2000). *Sleep needs*, *patterns*, *and difficulties of adolescents.* National Research Council and Institute of Medicine. Washington, D.C.: National Academy Press.

Horne, J. A., & Ostberg, O. (1976). A self assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, 4, 97-110.

Hughes, T. P. (1998). *Rescuing Prometheus*. New York, NY: Vintage Books.

Hursh, S. R., & Bell, G. B. (2001). *Human performance cognitive model for Air Force information operations* (Technical Report for the Air Force Research Laboratory). Brooks Air Force Base, TX: Air Force Research Laboratory.

Hursh, S. R., Redmond, D. P., Johnson, M. L., Thorne, D. R., Belenky, G., Balkin, T. J., Storm, W. F., Miller, J. C., & Eddy, D. R. (2004). Fatigue models for applied research in warfighting. *Aviation*, *Space*, *and Environmental Medicine*, *75*(3 Suppl), A44-A53.

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

Johnson, D. C., Polusny, M. A., Erbes, C., King, D., King, L., Litz, B., Schnurr, P., Friedman, M., & Southwick, S. (2008). *The response to stressful experiences scale (RSES)*. U.S. Department of Veterans Affairs National Center for PTSD.

Jones, M. B. (2000). Isoperformance and personnel decisions. *Human Factors*, *42*(2), 299-317.

Jones, M. B., & Kennedy, R. S. (1996). Isoperformance curves in applied physiology. *Human Factors*, *38*(1), 167-182.

Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J. M., & Sagi, D. (1994). Dependence on REM sleep of overnight improvement of perceptual skill. *Science*, 265, 679-682.

Killgore, W. D., Estrada, A., Wildzunas, R. M., & Balkin, T. J. (2008). *Sleep and performance measures in soldiers undergoing military relevant training.* Presented at the 26th Army Science Conference, Orlando, FL.

Killgore, W. D., Balkin, T. J., & Wesensten, N. J. (2006). Impaired decision making following 49 h of sleep deprivation. *Journal of Sleep Research*, *15*(1), 7-13.

Kim, J. O., & Curry, J. (1977). The treatment of missing data in multivariate analysis. *Sociological Methods and Research*, 6, 216-240

Krueger, G. P. (1990). Sustained military performance in continuous operations: combatant fatigue, rest, and sleep needs. In R. Gal & A. D. Mangelsdorff (Eds.), *Handbook of military psychology* (pp. 244-277). Chichester, UK: John Wiley and Sons.

Lange, T., Perras, B., Fehm, H. L., & Born, J. (2003). Sleep enhances the human antibody response to Hepatitis A vaccination. *Psychosomatic Medicine*, 65, 831-835.

Martin, B. J. (1981). Effect of sleep deprivation on tolerance of prolonged exercise. *European Journal of Applied Physiology*, 47, 345-354.

Martin, B. J., & Gaddis, G. M. (1981). Exercise after sleep deprivation. *Medicine and Science in Sports and Exercise*, *13*(4), 220-223.

McLellan, T. M., Kamimori, G. H., Bell, D. G., Smith, I. F., Johnson, D., & Belenky, G. (2005). Caffeine maintains vigilance and marksmanship in simulated urban operations with sleep deprivation. *Aviation, Space, and Environmental Medicine*, *76*(1), 39-45.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1981). *Manual for the profile of mood states.* San Diego, CA: Educational and Industrial Testing Services.

Mercer, P. W., Merritt, S. L., & Cowell, J. M. (1998). Differences in reported sleep need among adolescents. *Journal of Adolescent Health*, *23*(5), 259-263.

Miller, D. B. (2005). *Sleep and predicted cognitive performance of new cadets during cadet basic training at the United States Military Academy.* Monterey, CA: Masters Thesis at Naval Postgraduate School.

Miller, J. C. (2006). *Fundamentals of shiftwork scheduling* (Report No. AFRL-HE-BR-2006-0011). Brooks City-Base, TX: Human Effectiveness Directorate, Air Force Research Laboratory.

Miller, N. L., Matsangas, P., & Shattuck, L. G. (2007). Fatigue and its effect on performance in military environments. In P. A. Hancock & J. L. Szalma (Eds.), *Performance under stress* (chap. 12, pp. 231-249). Burlington, VT: Ashgate Publishing Company.

Miller, N. L., & Shattuck, L. G. (2005). Sleep patterns of young men and women enrolled at the United States Military Academy: Results from year one of a four year longitudinal study. *Sleep*, *28*(7), 837-841.

Miller, N.L., Shattuck, L.G. and Matsangas, P (2010). Sleep patterns of College Students: Longitudinal Study of United States Military Academy Cadets. *Sleep*, in press.

Miller, N. L., Shattuck, L. G., Matsangas, P., & Dyche, J. (2008). Sleep and academic performance in U.S. military training and education programs. *Mind*, *Brain*, *and Education*, *2*(1), 29-33.

Moldofsky, H. (1995). Sleep and the immune system. *International Journal of Immunopharmacology*, *17*(8), 649-654.

Monk, T. H., Buysse, D. J., Reynolds, C. F., Jarrett, D. B., & Kupfer, D. J. (1992). Rhythmic vs. homeostatic influences on mood, activation, and performance in young and old men. *Journal of Gerontology*, *47*(4), 221-227.

Monk, T. H., Buysse, D. J., & Rose, L. R. (1999). Wrist actigraphic measures of sleep in space. *Sleep*, *22*(7), 948-954.

Naval Center for Combat and Operational Stress Control (2009). *Resilience: What is it?* Retrieved July 16, 2009, from http://www.med.navy.mil/sites/nmcsd/nccosc/Resilience/Documents/Resilience_What_Is_It.pdf

Neri, D. F. (2004). Fatigue and performance modeling workshop, June 13-14, 2002. *Aviation, Space, and Environmental Medicine*, *75*(3 Suppl), A1-A3.

Paquet, J., Kawinska, A., & Carrier, J. (2007). Wake detection capacity of actigraphy during sleep. *Sleep*, *30*(10), 1362-1369.

Pigeau, R., Angus, B., & O'Neil, P. (1995). Vigilance latencies to aircraft detection among NORAD surveillance operators. *Human Factors*, *37*(3), 622-634.

Reilly, T., & Deykin, T. (1983). Effects of partial sleep loss on subjective states, psychomotor and physical performance tests. *Journal of Human Movement Studies*, 9, 157-170.

Rosa, R. H. (2001). What can the study of work scheduling tell us about adolescent sleep? In M. A. Carskadon (Ed.), *Adolescent sleep patterns: Biological*, *social*, *and psychological influences* (chap. 10, pp. 159-171). New York, NY: Cambridge University Press.

Selvi, Y., Gulec, M., Agargun, M. Y., & Besiroglu, L. (2007). Mood changes after sleep deprivation in morningness-eveningness chronotypes in healthy individuals. *Journal of Sleep Research*, 16, 241-244.

Snow, R. E. (1978). Theory and method for research on aptitude processes. *Intelligence*, 2, 225-278.

Stickgold, R., James, L., & Hobson, A. (2000). Visual discrimination learning requires sleep after training. *Nature Neuroscience*, *3*(12), 1237-1238.

Tamm, A. S., Lagerquist, O., Ley, A. L., & Collins, D. F. (2009). Chronotype influences diurnal variations in the excitability of the human motor cortex and the ability to generate torque during a maximum voluntary contraction. *Journal of Biological Rhythms*, *24*(3), 211-224.

Taub, J. M., & Berger, R. J. (1973). Performance and mood following variations in the length and timing of sleep. *Psychophysiology*, *10*(6), 559-570.

Tharion, W. J., Shukitt-Hale, B., & Lieberman, H. R. (2003). Caffeine effects on marksmanship during high-stress military training with 72 hour sleep deprivation. *Aviation, Space, and Environmental Medicine*, *74*(4), 309-314.

Thorne, D. R., Thomas, M. L., Russo, M.B., Sing, H. C., Balkin T. J., Wesensten, N. J., et al. (1999). Performance on a driving-simulator divided attention task during one week of restricted nightly sleep. *Sleep*, 22(Suppl 1), 301.

Trickel, M. T., Barnes, M. D., & Egget, D. L. (2000). Health-related variables and academic performance among first-year college students: Implications for sleep and other behaviors. *Journal of American College Health*, *49*(3), 125-131.

Van Dongen, H. P. A., Maislin, G., Mullington, J. M., & Dinges, D. F. (2003). The cumulative cost of additional wakefulness: Dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. *Sleep*, 26, 2.

Van Helder, T., & Radomski, M. W. (1989). Sleep deprivation and the effect of exercise performance. *Sports Medicine*, 7, 235-247.

Walker, M. P., Brakefield, T., Hobson, J. A., & Stickgold, R. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature*, 425, 616-620.

Weiner, I. B., & Greene, R. L. (2008). *Handbook of personality assessment*. Hoboken, NJ: John Wiley & Sons.

Weisz, J. D. (1967). *Systems analysis: Human factors research and development contributions* (Technical Note 6-67). Aberdeen Proving Ground, MD: Human Engineering Laboratory. (DTIC Accession No. AD818566)

Weisz, J. D. (1968). *Systems analysis: Manpower resource/system design integration methodology* (Technical Note 9-68). Aberdeen Proving Ground, MD: Human Engineering Laboratory. (DTIC Accession No. AD675481)

Whitener, E. M. (1989). A meta-analytic review of the effect on learning of the interaction between prior achievement and instructional support. *Review of Educational Research*, *59*(1), 65-86.

Wilson, M. A., & McNaughton, B. L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science*, 265, 676-679.

Wolfson, A. R. (2001). Bridging the gap between research and practice: What will adolescents' sleep-wake patterns look like in the 21st century? In M. A. Carskadon (Ed.), *Adolescent sleep patterns: Biological*, *social*, *and psychological influences* (chap. 12, pp. 198-219). New York, NY: Cambridge University Press.

Wolfson, A. R., & Carskadon, M. A. (2003). Understanding adolescents' sleep patterns and school performance: A critical appraisal. *Sleep Medicine Reviews*, *7*(6), 491-506.

Wood, C., & Magnello, M. E. (1992). Diurnal changes in perceptions of energy and mood. *Journal of the Royal Society of Medicine*, 85, 191-194.

INITIAL DISTRIBUTION LIST

1. Research Office (Code 09)..1
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Dudley Knox Library (Code 013)...2
   Naval Postgraduate School
   Monterey, CA 93943-5002

3. Defense Technical Information Center...2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, VA 22060-6218

4. Richard Mastowski (Technical Editor)..2
   Graduate School of Operational and Information Sciences (GSOIS)
   Naval Postgraduate School
   Monterey, CA 93943-5219

5. Dr. Nita Lewis Miller...1
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5219

6. Dr. Lawrence Shattuck...1
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5219

7. Major General Gregg F. Martin, USA..1
   Commandant, U.S. Army War College
   Carlisle Barracks, PA 17013

8. Mr. Joseph Driskill..1
   Leonard Wood Institute
   Fort Leonard Wood, MO 65583