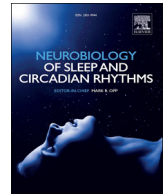




Contents lists available at ScienceDirect

Neurobiology of Sleep and Circadian Rhythms

journal homepage: www.elsevier.com/locate/nbscr

Multiple caffeine doses maintain vigilance, attention, complex motor sequence expression, and manual dexterity during 77 hours of total sleep deprivation

William D.S. Killgore^{a,b,*}, Gary H. Kamimori^a^a Department of Behavioral Biology, Walter Reed Army Institute of Research, USA^b Department of Psychiatry, University of Arizona College of Medicine, USA

ARTICLE INFO

Keywords:

Total sleep deprivation
Caffeine
Boredom/monotony
Multi dose paradigm
PVT
Tower of Hanoi
Grooved pegboard
Vigilance
Attention

ABSTRACT

Sleep deprivation (SD) and fatigue have detrimental effects on performance in operational settings. Few studies have investigated the cumulative effects of SD and fatigue on performance under heavy workload demands. Therefore, we investigated the efficacy of multiple repeated doses of caffeine as a countermeasure to SD and fatigue during 77 h total SD (TSD) during the early morning hours. Twenty-three males and females, 18 – 35 years of age, who identified as moderate caffeine consumers completed the Psychomotor Vigilance Task (PVT) 141 times during the experimental test period. Caffeine was administered in a multi-dose paradigm over three nights without sleep. Participants received either caffeine (200 mg) or placebo at the beginning of each 2-h test block from 0100 – 0900 (800 mg total per night). While PVT speed declined for both groups across all 3 nights, the caffeine group consistently out-performed the placebo group. Caffeine maintained attentiveness (1-5 s lapses) on night 1, but this advantage was lost on nights 2 and 3. Caffeine outperformed placebo for responsive lapses (5-9 s lapses) across all three nights, but caffeine performance was still notably worse than at baseline. Prolonged non-responsive lapses (beyond 10 s) were only reduced by caffeine on night 2. Caffeine was more effective than placebo across all nights at sustaining completion speed of a complex motor sequence task and a manual coordination task. Essentially, caffeine is an effective countermeasure for SD, as it mitigates declines in speed and failures to respond, and sustains motor planning and coordination. However, caffeine does not restore normal functioning during SD and cannot be considered as a replacement for sleep.

1. Introduction

According to the Centers for Disease Control and Prevention (CDC), nearly 30% of American workers routinely obtain less than 6 h of sleep per night (Center for Disease Control and Prevention, 2012). This is concerning, as insufficient sleep is associated with increased lapses of attention (Lim and Dinges, 2008) and can lead to a host of cognitive deficits that can affect performance, decision making, and interpersonal relationships (Banks and Dinges, 2007; Killgore, 2010). Moreover, sleepiness and fatigue resulting from lack of sleep are major contributors to operational errors, traffic accidents, and aviation disasters (Dinges, 1995; Gander et al., 1998a, 1998b; Bendrick et al., 2016; Czeisler et al., 2016; Gottlieb et al., 2018; Matsui et al., 2017; Kalsi et al., 2018; Martiniuk et al., 2013). Several factors can contribute to the neurobehavioral deficits associated with sleep deprivation (SD), including the cumulative build-up of the homeostatic drive for sleep,

which accrues as time awake increases, and the circadian rhythm of alertness, which fluctuates across the 24-h day (Borbely et al., 1989). While these biological variables are the major drivers of the adverse effects of SD, additional factors such as monotony, time-on-task, and cognitive fatigue can also interact to exacerbate these deficits (Lim and Dinges, 2008; Grant et al., 2017; Satterfield et al., 2017; Maire et al., 2014; Stutts et al., 2003; Thiffault and Bergeron, 2003).

The cumulative effects of SD and fatigue on performance are well documented. Vigilance and simple attention, as assessed by the psychomotor vigilance test (PVT) (Dinges and Powell, 1985), are the aspects of performance that are most consistently impaired by SD (Lim and Dinges, 2008). The PVT is a relatively simple test of sustained attention that is highly sensitive to the slowing of response times and attentional lapses that emerge during SD (Balkin et al., 2004), and which shows no demonstrable improvement with practice (Lim and Dinges, 2008, 2010). Part of the sensitivity of the PVT to prolonged SD

* Corresponding author. Department of Psychiatry, University of Arizona, 1501 N Campbell Ave, Room 7303B, Tucson, AZ, 85724, USA.

E-mail address: killgore@psychiatry.arizona.edu (W.D.S. Killgore).

<https://doi.org/10.1016/j.nbscr.2020.100051>

Received 18 January 2020; Received in revised form 24 May 2020; Accepted 27 May 2020

Available online 31 May 2020

2451-9944/© 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

is likely due to its simple and monotonous format, which quickly unmasks waning attentional capacities. As the duration of SD increases, the probability of simple attentional lapses and cognitive slowing on the PVT is heightened, and the risk of these deficits is compounded as the duration of work, or “time-on-task,” is increased (Grant et al., 2017; Satterfield et al., 2017). Most evidence suggests that vigilance decrements from SD are most likely to emerge during long, unengaging, tedious tasks, whereas more complex activities that are more cognitively interesting are less likely to show SD-related impairments (Harrison and Horne, 2000). This is critical, as many industries involving 24-h operations (e.g., medical, military, transportation) often require completion of simple monotonous tasks with low-level demands (e.g., monitoring, surveillance, driving) that induce fatigue and boredom. While a relatively large research literature has addressed the effects of sleep deprivation on relatively short and discrete cognitive and vigilance tasks, little work has examined the ability to sustain performance for prolonged periods of time while engaging in monotonous, yet cognitively fatiguing activity during sleep deprivation.

Because sufficient restorative sleep is not always possible in certain operational contexts, there has been great interest in studying the effects of pharmacological countermeasures to sleepiness and fatigue. In particular, caffeine has been extensively studied for its capacity to sustain alertness and vigilance during periods of SD (Kamimori et al., 1995, 2005; Killgore et al., 2009a; Syed et al., 2005; Paech et al., 2016; Wesensten et al., 2005; McLellan et al., 2004, 2005a, 2005b; Tikuisis et al., 2004; Tharion et al., 2003; Lieberman et al., 2002). Caffeine administered in a multi-dose paradigm is effective at maintaining alertness in sleep deprived and fatigued individuals as well as military populations during short periods of SD (Paech et al., 2016; McLellan et al., 2005a, 2005b; Kamimori et al., 2005). Many studies have examined the effects of caffeine under periods of short- and long-term SD, but few studies have examined the effectiveness of multi-dose caffeine in maintaining prolonged vigilance under total sleep deprivation (TSD), with a repetitive high cognitive workload task during the early morning hours. The need to fill this information gap is critical and has high applicability for many operational settings where individuals must sustain performance on repetitive, uninteresting, and tedious tasks, without sleep, and during an adverse circadian time. Under such circumstances, individuals may not simply show cognitive slowing and brief attentional lapses, but may find it virtually impossible to maintain a continuous stream of conscious wakefulness despite engaging in focused and demanding cognitive activity. Such individuals may find themselves experiencing an uncontrollable “sleep attack” and unable to rouse themselves voluntarily (or even with mild external prodding). These longer bouts of non-responsiveness could be particularly disastrous in certain industries. It is currently unknown whether caffeine is effective at reducing these types of non-responsive periods under extreme sleep deprivation and prolonged monotony. Additionally, there is little evidence regarding the effects of caffeine on repeated higher order planning and sequencing abilities or manual dexterity during sleep deprivation.

The goal of the present study was to determine the effectiveness of multi-dose caffeine on the ability to sustain vigilance, cognitive, and manual dexterity performance under conditions of high monotony/boredom, heavy workload, persistent time-on-task effects, and circadian time-of-day effects. In particular, we focused on longer lapse periods than typically studied in prior work to see whether caffeine would reduce the frequency, duration, and severity of these longer periods of non-responsiveness. We studied the effects of repeated bi-hourly overnight administrations of caffeine (200 mg each administration for a total of 800 mg per night) on performance during 77 h of TSD under heavy workload during the early morning hours. Cognitive workload was maintained by contiguous repeated administrations of the PVT (36 administrations during each 8 h overnight session). We examined vigilance and attention on the PVT, as well as complex motor sequence expression, and manual dexterity/hand-eye coordination at

multiple times each night. Our primary hypothesis was that caffeine would be superior to placebo for sustaining psychomotor vigilance performance (speed and all lapse durations) across all three nights. We also had secondary hypotheses predicting that the ability to maintain a complex overlearned sequence over multiple trials and switch flexibly to a new task set would be sustained by caffeine relative to placebo. Finally, we hypothesized that the ability to maintain manual coordination and speed would also be greater among those receiving caffeine relative to placebo.

2. Methods

2.1. Participants

Twenty-nine healthy, non-smoking, enlisted military personnel volunteered to participate in a three-night sleep deprivation study involving potential administration of caffeine. One participant withdrew on the first day in the middle of the afternoon after complaining of feeling sick. A second participant withdrew due to feeling “tired” at 0430 on the first night of sleep deprivation, and a third participant withdrew at 0830 on the second night of sleep deprivation, complaining of headache. Three others were not included in the analysis due to partial data loss. The final sample reported here consisted of twenty-three (19 male; 4 female) participants ranging in age from 20 to 35 years (M age = 25.3, SD = 4.1). All participants had attained at least 12 years of formal education (M = 14.0, SD = 1.6). Participants reported their racial background as Black (47.8%), White (34.8%), Hispanic (13.0%), or Asian (4.3%). The body mass index (BMI) ranged from 20.8 to 35.1 (M = 26.0, SD = 3.4). Prior to study commencement, participants were required to pass a physical examination. Exclusionary criteria included any past or current physical/mental health problems, sleep problems, drug abuse, regular caffeine consumption ≥ 300 mg of caffeine per day, or a history of caffeine sensitivity. Based on self-report, all participants were required to be on a normal day/night sleep schedule and report sleeping between 7 to 8 h per night. Participants were non-nicotine users validated by nicotine/cotinine testing at the time of the screening physical. All participants completed a caffeine history questionnaire that asked about use of various caffeine containing products and average use of those products per day and per week. Based on this questionnaire, all participants were self-reported moderate to low caffeine users, ranging from 0-104 mg/day (M = 12.3, SD = 24.3). Prior to each study run, participants were assigned to four-person groups and provided with an initial study date (actual group size varied depending on participant availability and scheduling). Each group was randomly assigned to receive caffeine (n = 12; 9 male; M age = 25.6, SD = 4.0 years) or placebo (n = 11; 10 male; M age = 24.9, SD = 4.4 years) in a double-blind fashion. Groups did not differ significantly in terms of age or gender. To minimize variability, female participants were screened for pregnancy and did not use any form of hormonal contraceptives three months prior to the study, as they are known to modify caffeine pharmacokinetics (Teichmann, 1990). All participants abstained from alcohol, stimulants, and other psychoactive drugs for 48 h prior to initiation of the study. Written informed consent was obtained from all participants. This study was approved by the Walter Reed Army Institute of Research Human Use Review Committee and the U. S. Army Human Subjects Research Review Board.

2.2. Caffeine administration

Caffeine was administered in a chewing gum formulation (*Military Energy Gum™*; MarketRight Inc., Plano, IL). Kamimori et al. (2002) demonstrated that caffeine administered in a chewing gum formulation is absorbed at a significantly faster rate, and reaches peak blood concentrations faster than caffeine administered in a capsule form (Kamimori et al., 2002). Each dose (placebo or 200 mg) was divided between two pieces of gum and administered at 0100, 0300, 0500 and

Day 1	Day 2	Day 3	Day 4	Day 5
Sleep at Home	Night 0 (sleep)	Night 1	Night 2	Night 3
		<p>Baseline (PVT): 0000 - 0045 0015 – PVT 0025 – PVT 0035 – PVT</p> <p>Block 1: 0100 - 0300 0100 – Gum #1 0105 – PVT 0115 – PVT^a 0125 – PVT 0135 – PVT^a 0145 – GP 0150 – PVT 0200 – TOH 0215 – PVT^a 0225 – PVT 0235 – PVT^a 0245 – PVT^a</p> <p>Block 2: 0300 – 0500 0300 – Gum #2 0305 – PVT 0315 – PVT^a 0325 – PVT 0335 – PVT^a 0345 – GP 0350 – PVT 0400 – TOH 0415 – PVT^a 0425 – PVT 0435 – PVT^a 0445 – PVT^a</p> <p>Block 3: 0500 – 0700 0500 – Gum #3 0505 – PVT 0515 – PVT^a 0525 – PVT 0535 – PVT^a 0545 – GP 0550 – PVT 0600 – TOH 0615 – PVT^a 0625 – PVT 0635 – PVT^a 0645 – PVT^a</p> <p>Block 4: 0700 – 0900 0700 – Gum #4 0705 – PVT 0715 – PVT^a 0725 – PVT 0735 – PVT^a 0745 – GP 0750 – PVT 0800 – TOH 0815 – PVT^a 0825 – PVT 0835 – PVT^a 0845 – PVT</p>	<p>Baseline (PVT): 0000 - 0045 0015 – PVT 0025 – PVT 0035 – PVT</p> <p>Block 1: 0100 - 0300 0100 – Gum #1 0105 – PVT 0115 – PVT^a 0125 – PVT 0135 – PVT^a 0145 – GP 0150 – PVT 0200 – TOH 0215 – PVT^a 0225 – PVT 0235 – PVT^a 0245 – PVT^a</p> <p>Block 2: 0300 – 0500 0300 – Gum #2 0305 – PVT 0315 – PVT^a 0325 – PVT 0335 – PVT^a 0345 – GP 0350 – PVT 0400 – TOH 0415 – PVT^a 0425 – PVT 0435 – PVT^a 0445 – PVT^a</p> <p>Block 3: 0500 – 0700 0500 – Gum #3 0505 – PVT 0515 – PVT^a 0525 – PVT 0535 – PVT^a 0545 – GP 0550 – PVT 0600 – TOH 0615 – PVT^a 0625 – PVT 0635 – PVT^a 0645 – PVT^a</p> <p>Block 4: 0700 – 0900 0700 – Gum #4 0705 – PVT 0715 – PVT^a 0725 – PVT 0735 – PVT^a 0745 – GP 0750 – PVT 0800 – TOH 0815 – PVT^a 0825 – PVT 0835 – PVT^a 0845 – PVT</p>	<p>Baseline (PVT): 0000 - 0045 0015 – PVT 0025 – PVT 0035 – PVT</p> <p>Block 1: 0100 - 0300 0100 – Gum #1 0105 – PVT 0115 – PVT^a 0125 – PVT 0135 – PVT^a 0145 – GP 0150 – PVT 0200 – TOH 0215 – PVT^a 0225 – PVT 0235 – PVT^a 0245 – PVT^a</p> <p>Block 2: 0300 – 0500 0300 – Gum #2 0305 – PVT 0315 – PVT^a 0325 – PVT 0335 – PVT^a 0345 – GP 0350 – PVT 0400 – TOH 0415 – PVT^a 0425 – PVT 0435 – PVT^a 0445 – PVT^a</p> <p>Block 3: 0500 – 0700 0500 – Gum #3 0505 – PVT 0515 – PVT^a 0525 – PVT 0535 – PVT^a 0545 – GP 0550 – PVT 0600 – TOH 0615 – PVT^a 0625 – PVT 0635 – PVT^a 0645 – PVT^a</p> <p>Block 4: 0700 – 0900 0700 – Gum #4 0705 – PVT 0715 – PVT^a 0725 – PVT 0735 – PVT^a 0745 – GP 0750 – PVT 0800 – TOH 0815 – PVT^a 0825 – PVT 0835 – PVT^a 0845 – PVT</p>
<p>1900—Participants arrive at lab</p> <p>2000—GP, PVT, TOH Training</p> <p>2300—Participants begin 8 hour sleep period</p>	<p>0700—Participants awoken</p> <p>0935—Begin practice testing: GP, PVT, TOH ~every 2 hr</p> <p>Baseline (TOH), GP) 1605 – TOH 1745 – GP 1800 – TOH 1945 – GP 2000 – TOH 2145 – GP</p>			<p>1200 – 2000 – 8 hours Recovery sleep</p>

Fig. 1. Study procedure and timeline. *Day* describes the number of days participants remained in the lab. *Night* describes nights of sleep deprivation and experimental testing. Psychomotor Vigilance Task; TOH – Tower of Hanoi; GP – Grooved Pegboard. Caffeine gum administrations in bold. ^aTime block details: Provides a detailed description of each time block during the experimental test days. PVT^b – PVT, SSS, and SEQ were administered.

0700 for a cumulative dose of 0 mg or 800 mg caffeine each night. We have previously demonstrated that 200 mg of caffeine administered via multiple pieces of caffeinated gum is highly effective at sustaining cognitive and behavioral performance during sleep deprivation (Kamimori et al., 2005, 2015; McLellan et al., 2004; LaJambe et al., 2005). To minimize expectancy effects, all participants in the same 4-person cohort received the same study intervention (i.e., caffeine gum or placebo gum). Gum was administered to each participant in their own private bedroom. Participants chewed the gum for 5 min in the presence of a research technician and then it was expectorated.

2.3. Measures

PVT. The primary outcome variable for this study was a modified version of the psychomotor vigilance test (PVT) (Dinges and Powell, 1985), which measures simple reaction time to a pseudo-randomly appearing visual stimulus over a sustained vigilance period. A brief, 5-min version of the PVT was administered on a Personal Data Assistant (PDA) (Thorne et al., 2005) (see Fig. 1). For each trial, a “bullseye” target appeared in the center of the PDA screen and participants were required to respond as quickly as possible by pressing an assigned button on the PDA. Stimuli were presented with an inter-stimulus interval that varied from 1 to 5 s in 1-s increments randomized without replacement in blocks of 10 to control run lengths, equalize presentation frequencies, and to hold stimulus rates to approximately ten times/min. Dependent measures averaged or summed across the 5-min PVT session, included mean speed (reciprocal of average response latency across trials) and number of three different types of attentional lapses (described below).

During the PVT administration, participants were closely monitored by research staff, and if a participant failed to respond to the stimulus, a verbal prompt was made between five and 8 s after stimulus onset, followed by a second verbal prompt (if necessary), and then finally by a physical intervention (i.e., tap on the shoulder) if no response had been made. For the present analysis, we characterize three types of lapses based on PVT response time and the level of intervention required to arouse the participant to the point of responding: 1) **Attentional Lapse (AL)**: these simple lapses are defined as momentary periods of non-responsiveness with reaction times lasting between 1000 and 4999 ms. ALs are characterized by a brief period of disengagement that was spontaneously self-corrected without external intervention or prompting from study staff; 2) **Responsive Lapse (RL)**: these lapses represented more severe disengagements or momentary sleep periods, and were defined as lasting from 5000–9999 ms. An RL was recorded if the participant responded to the PVT before 10 s had elapsed, typically in response to one of the verbal prompts or the subsequent physical stimulus. Finally, a 3) **Non-Responsive Lapse (NRL)** was recorded if none of these stimuli resulted in a PVT response within 10000 ms.

Participants completed the PVT 141 times throughout the study. Eight practice sessions were completed on Day 2 (0940, 1150, 1350, 1555, 1750, 1950, 2150, 2350 h). Forty-seven tests were completed on days 3 and 4, and 39 tests were completed on Day 5 (see Fig. 1). Baseline data were collected at 0015, 0025, and 0035 on Days 3 (Night 1) through 5 (Night 3). However, due to the effects of SD on performance, only baseline data from Night 1 were used for statistical comparisons. Following the baseline testing period, there were four 2-h test blocks on Days 3 through 5: block 1 (0100 – 0300), block 2 (0300 – 0500), block 3 (0500 – 0700), and block 4 (0700 – 0900). Nine 5-min PVTs were administered in each block.

The Tower of Hanoi (TOH). The TOH is a mathematical puzzle that assesses planning ability. The version used here consists of three vertical pegs and five disks descending in size (see Fig. 2). The TOH has historically been considered as an executive functioning test of planning and sequencing (Killgore et al., 2009a). However, since the algorithm can be easily learned and applied recursively, we sought to test how well caffeine would sustain optimal performance of this algorithm

over three nights of sleep deprivation after it had been well learned. Therefore, participants were taught the algorithm on the evening of Day 1 and practiced it until they were able to solve the puzzle nearly perfectly (i.e., in ≤ 33 moves) on two consecutive administrations. Upon awakening the next day, participants then completed the TOH every 2 h over the course of 73 h of sleep deprivation. The total trial time to solve the puzzle was recorded for each administration. The puzzle was terminated if the correct solution had not been achieved in 100 moves.

Additionally, as a metric of cognitive flexibility, we surprised participants with a similar appearing but slightly different version of the TOH at 0935 on day 5, after the completion of all other tasks in the study (i.e., 75.5 h of total sleep deprivation). Instead of the standard 5-ring version that they had completed repeatedly over the preceding days, they were presented with a novel 6-ring version with three pegs and were informed that the version was different than what they had completed previously. This version looked similar but required a minimum of 63 moves to solve perfectly. The total trial time to solve the puzzle was recorded. The puzzle was terminated if the correct solution had not been achieved in 200 moves.

Grooved Pegboard (GP). The GP (Matthews and Klove, 1964) assesses fine motor coordination and motor speed. The GP is a 5x5 metal board with 25 holes each with randomly positioned keyhole slots (see Fig. 2). One-inch metal pegs with a keylike shape are stored in an attached adjacent dish. Pegs must be picked up from the dish and rotated to match the keyhole slot before they can be inserted. Once timing began, participants were required to pick up and insert the pegs as quickly as possible using their dominant hand, working from left to right (for right handed participants, or from left to right for left handed individuals), without skipping any slots. Total trial time was recorded.

Side Effects Questionnaire (SEQ). The SEQ is a seven item self-report inventory that assesses the extent to which participants experienced symptoms commonly attributed to caffeine. Items include: headache, dizziness, nervousness, lightheadedness, incoordination, nausea, and vomiting. Participants rated each item on a scale from 1 – *not experiencing* to 4 – *severe*. Participants completed the SEQ five times during each block.

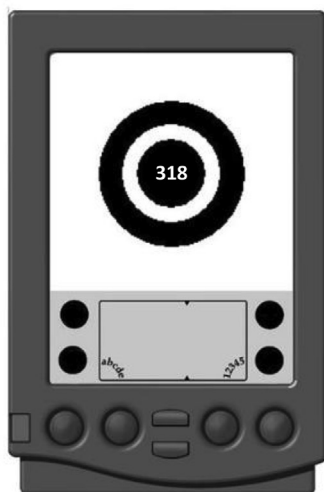
Stanford Sleepiness Scale (SSS). The SSS (Hoddes et al., 1973) is a single-item self-report inventory assessing subjective sleepiness. Participants rated items on a scale ranging from 1 – *feeling active and vital, alert, or wide awake* to 7 – *no longer fighting sleep, sleep onset soon, having dream-like thoughts*. A computerized version of this test was administered near the end of each 2-h test block. Participants completed the SSS five times during each block.

2.4. Design and procedure

The present study used a double-blind design with three continuous nights without sleep and primary testing in the early morning hours. TSD was defined as continuous wakefulness with no naps. Each study run included a maximum of four participants and was 77 h in duration. No more than two study runs in a row were of the same assignment (placebo or caffeine). Participants were randomly assigned to either the placebo or caffeine group according to their individual availability to participate. Monetary compensation was awarded upon completion of all the test trials. The 5-min PVT was the primary performance measure, with TOH and GP as secondary outcomes.

Participants reported to the lab at 1900 on Day 1 and began training on the tasks. From 2300 (Day 1) to 0700 (Day 2) participants completed an 8-h sleep period in separate bedrooms. They were awakened at 0700 h (Day 2) and were kept awake for the next 77 h. Participants completed 8 practice sessions on Day 2. Baseline data for the TOH and GP were recorded from 1605 to 2145 (Day 2). Baseline PVT data for each 24-h period was collected between 0000 and 0045 each night. Due to the effects of SD, baseline TOH and GP data from Day 2, and baseline PVT data from Night 1 were used as a comparison for all subsequent time blocks in all analyses. At 0100 h (Day 3), participants began

A) Psychomotor Vigilance Test (PVT)

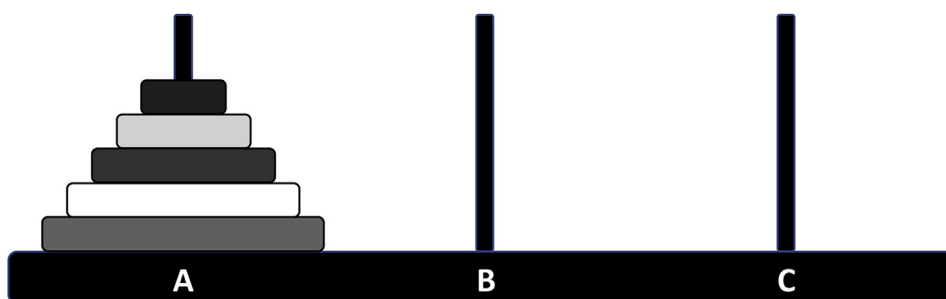


B) Grooved Pegboard Test



Fig. 2. Tasks used to assess the effects of sleep deprivation. A) Palm-PVT: a 5-min version of the psychomotor vigilance test (PVT) administered on a hand-held personal assistant device. Participants were required to press one of the response keys at the bottom of the device as quickly as possible each time the “bullseye” target appeared. Response time was displayed in the center. B) Grooved Pegboard (BP): participants were required to place oddly shaped metal pegs in small keyholes as rapidly as possible. C) Tower of Hanoi (TOH): Participants were required to move a stack of colored rings from peg A to peg C as quickly as possible, while following the rules that only one ring could be moved at a time, and that a larger ring could never be placed on top of a smaller ring.

C) Tower of Hanoi



completing the four 2-h experimental test blocks (0100-0300, 0300-0500, 0500-0700, 0700-0900 h). Participants began each test block by chewing 2 pieces of gum containing either placebo or caffeine (100 mg/piece) for 5 min. After chewing for 5 min participants expectorated the gum and immediately began testing. The PVT, TOH, GP, SSS, and SEQ were completed during each test block (Fig. 1). A variety of neuropsychological tasks and questionnaires were administered during the daytime period and the results are reported elsewhere (Kahn-Greene et al., 2006, 2007; Killgore et al., 2007a, 2007b, 2007c, 2008a, 2008b, 2011; McBride et al., 2006). Following the completion of the third night of testing participants completed an 8-h recovery sleep period (1200-2000 h) prior to release from the laboratory.

2.5. Statistical design

A three-way mixed analysis of variance (ANCOVA) was used to analyze the PVT using drug condition (between subjects factor: caffeine/placebo), block (within subjects factor: Baseline/Blocks 1 – 4), and night (within subjects factor: Nights 1 – 3), correcting for baseline performance. Prior work in our lab has suggested that caffeine effects are quite robust for PVT speed, with effect sizes (η_p^2 (Lim and Dinges, 2008)) ranging from moderate to large. In designing the study, a power analysis assuming at least a moderate effect size (η_p^2 (Lim and Dinges, 2008) $\geq .06$), with 2 groups and 12 repeated measures (i.e., blocks) and a .5 correlation among repeated measures suggested at least $n = 10$ to have .80 power for the interactions. Therefore, assuming moderate to large effect sizes, the planned sample size of $n = 24$ was deemed to be

sufficiently powered to detect effects within that range. For the TOH and GP, a three-way mixed analysis of variance (ANCOVA) was used with drug group (between subjects factor: caffeine/placebo), block (within subjects factor: Blocks 1 – 4), and night (within subjects factor: Nights 1 – 3), correcting for baseline performance. A logistic regression was used to analyze the SEQ and SSS data to determine which group (caffeine or placebo) was more likely to report each symptom throughout the study, with SEQ scores and night as predictors. Descriptive statistics are reported as mean \pm standard deviation. SEQ and SSS results are reporting using odds ratio (OR) and 95% confidence interval.

3. Results

3.1. PVT

Speed. There was no difference between groups in speed at baseline on Night 1, ($t_{16,21} = 1.09, p = .293$). There was a significant condition \times block \times night interaction for speed, ($F_{6,120} = 2.88, p = .012, \eta_p^2 = .126$). There was a significant main effect of Night ($F_{2,40} = 14.36, p < .001, \eta_p^2 = .418$), suggesting that performance declined across nights, but this decline was not moderated by drug condition ($F_{2,40} = 1.45, p = .247, \eta_p^2 = .068$). For caffeine, Night 1 speed was significantly faster than Nights 2 or 3, which did not differ from each other (see Table 1). This was true for placebo as well. We then examined speed for each night separately. On Night 1, there was a main effect of drug condition whereby caffeine outperformed placebo for

Table 1
Within group comparisons between nights for each outcome.

Speed						Attentional Lapses (AL)					
Condition	Night	Mean	SE	Comparison Night	Bonferroni p	Condition	Night	Mean	SE	Comparison Night	Bonferroni p
Caffeine	1	2.537	0.118	> 2	.038	Caffeine	1	0.791	0.321	< 2	< .001
	2	0.741	0.078	< 1	.038		2	5.463	0.636	> 1	< .001
	3	0.716	0.104	< 1	< .001		3	5.366	0.362	> 1	< .001
Placebo	1	2.823	0.131	> 2	< .001	Placebo	1	3.387	0.382	< 2	.027
	2	1.977	0.284	< 1	< .001		2	5.437	0.580	> 1	.027
	3	1.563	0.228	< 1	< .001		3	5.785	0.482	> 1	.017

Responsive Lapses (RL)						Non-Responsive Lapses (NRL)					
Condition	Night	Mean	SE	Comparison Night	Bonferroni p	Condition	Night	Mean	SE	Comparison Night	Bonferroni p
Caffeine	1	0.074	0.24	< 2	.011	Caffeine	1	0.000	0.000	< 2	.039
	2	3.074	0.698	> 1	.011		2	0.420	0.132	> 1	.039
	3	4.581	0.948	> 1	.006		3	1.049	0.364	-	-
Placebo	1	1.794	0.448	< 2	< .001	Placebo	1	0.242	0.127	< 2	.003
	2	6.872	0.534	> 1	< .001		2	2.878	0.589	> 1	.003
	3	7.412	0.574	> 1	< .001		3	2.500	0.569	> 1	.014

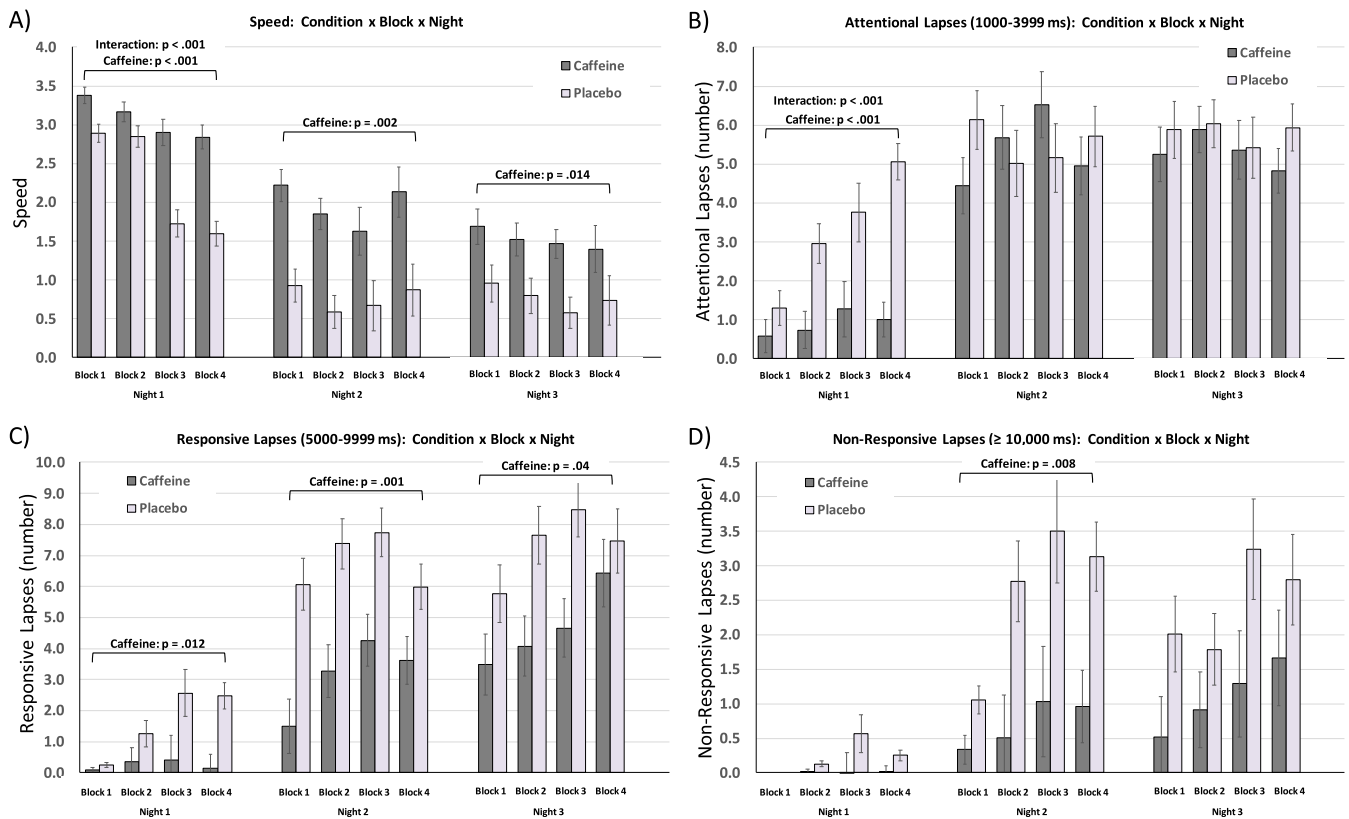


Fig. 3. Caffeine and placebo group performance for a) mean speed (error bars represent ± 1 SE) b) frequency counts of attentional lapses—AL c) frequency counts of responsive lapses—RL and d) frequency counts of non-responsive lapses—NRL on all 3 nights for all 4 blocks including baseline from Night 1.

Night 1 as a whole ($F_{1,20} = 20.26, p < .001, \eta_p^2 = .503$; see Fig. 3A), and a significant condition x block interaction ($F_{1,77,35,30} = 14.79, p < .001, \eta_p^2 = .425$) suggesting a greater decline for placebo across the first night compared to caffeine. For Night 2, there was no significant condition x block interaction ($F_{2,07,41,38} = 0.45, p = .649, \eta_p^2 = .022$), but there was a significant main effect of drug condition ($F_{1,20} = 12.95, p = .002, \eta_p^2 = .393$), suggesting that caffeine outperformed placebo across blocks (see Fig. 3A). For Night 3, there was no significant condition x block interaction ($F_{1,91,38,26} = 0.15, p = .849,$

$\eta_p^2 = .008$), but there was again a significant main effect of drug condition ($F_{1,20} = 7.26, p = .014, \eta_p^2 = .266$), suggesting that caffeine outperformed placebo across blocks (see Fig. 3A).

Attentional Lapses (AL). There was no group difference in ALs at baseline on Night 1 ($t_{21} = 0.58, p = .568$). There was no significant 3-way interaction ($F_{4,23,84,52} = 2.35, p = .058, \eta_p^2 = .105$). There was, however a significant effect of night ($F_{2,40} = 49.18, p < .001, \eta_p^2 = .711$), suggesting a general worsening of ALs over the three nights. For caffeine, Night 1 had fewer ALs than either Night 2 or Night 3, which

did not differ significantly from one another (see Table 1). This same pattern was seen for placebo. There was a significant drug condition x night interaction ($F_{2,40} = 4.16, p = .023, \eta_p^2 = .172$), suggesting that caffeine may have differential effects over the course of three nights. Specifically, on Night 1, there was a significant main effect of drug condition ($F_{1,20} = 19.09, p < .001, \eta_p^2 = .488$), with caffeine significantly outperforming placebo (see Fig. 3B), but this was moderated by a significant condition x drug interaction ($F_{3,60} = 4.69, p = .005, \eta_p^2 = .876$), suggesting that caffeine sustained performance throughout the first night, while placebo showed a progressive increase in lapses throughout the night. On the other hand, on Night 2 ($F_{1,20} = 0.14, p = .906, \eta_p^2 = .001$) and Night 3, the drug conditions did not differ in ALs ($F_{1,20} = 0.66, p = .426, \eta_p^2 = .032$).

Responsive Lapses (RL). There was no group difference in RLs at baseline on Night 1 ($t_{13,17} = 1.17, p = .264$). There was a significant condition x block x night interaction for RLs, ($F_{6,96} = 2.22, p = .048, \eta_p^2 = .122$), and a significant main effect of Night ($F_{2,32} = 51.63, p < .001, \eta_p^2 = .763$), suggesting that RLs increased across nights, although this increase was not moderated by drug condition ($F_{2,32} = 2.44, p = .103, \eta_p^2 = .132$). For caffeine, there were fewer RLs on Night 1 than on Night 2 or 3, which did not differ from one another (see Table 1). This same pattern was evident for placebo. On Night 1, there was no condition x block interaction ($F_{1,56,24,93} = 2.93, p < .083, \eta_p^2 = .155$), but there was a significant main effect of drug condition, with caffeine showing significantly fewer RLs than placebo for Night 1 as a whole ($F_{1,16} = 8.05, p < .012, \eta_p^2 = .335$; see Fig. 3C). Similarly, for Night 2, there was no significant condition x block interaction ($F_{3,48} = 1.11, p = .353, \eta_p^2 = .065$), whereas there was a significant main effect of drug condition ($F_{1,16} = 17.29, p = .001, \eta_p^2 = .519$), suggesting that caffeine was superior to placebo at reducing RLs across blocks (see Fig. 3C). Similarly, for Night 3, there was no significant condition x block interaction ($F_{3,48} = 2.29, p = .090, \eta_p^2 = .125$), but there was a significant main effect of drug condition ($F_{1,16} = 5.02, p = .040, \eta_p^2 = .239$), suggesting that caffeine outperformed placebo across blocks (see Fig. 3C).

Non-responsive Lapses (NRL). There was no group difference in NRLs at baseline on Night 1 ($t_9 = 1.00, p = .343$). There was no significant 3-way interaction ($F_{3,97,64,44} = 1.39, p = .248, \eta_p^2 = .080$). Similar to the findings for other outcome variables, there was a significant main effect of night ($F_{2,32} = 14.04, p < .001, \eta_p^2 = .467$), suggesting a general worsening of NRLs over the three nights. This was qualified by a significant drug condition x night interaction ($F_{2,32} = 3.57, p = .040, \eta_p^2 = .182$). For caffeine, there were fewer NRLs on Night 1 than Night 2, but no other differences, while for placebo, there were fewer NRLs on Night 1 than Night 2 and Night 3, which did not differ from one another (see Table 1). On Night 1, there was no main effect of drug condition ($F_{1,16} = 2.99, p = .103, \eta_p^2 = .158$), with most participants making no NRLs regardless of condition (see Fig. 3D). On the other hand, on Night 2 there was a significant main effect of condition ($F_{1,16} = 9.03, p = .008, \eta_p^2 = .361$), with caffeine producing significantly fewer NRLs than placebo. By Night 3, the drug conditions did not differ significantly in the average number of NRLs ($F_{1,16} = 3.02, p = .102, \eta_p^2 = .159$).

3.2. TOH

Repeated 5-Ring Puzzle. Greenhouse-Geisser corrections were applied due to significant sphericity for the within-subjects analyses. There was no significant condition x night ($F_{1,45, 54,95} = 0.73, p = .45, \eta_p^2 = .037$), condition x block ($F_{2,42, 54,95} = 0.98, p = .396, \eta_p^2 = .049$), or condition x night x block ($F_{2,89, 54,95} = 1.25, p = .301, \eta_p^2 = .062$) interaction for the TOH. There was, however, a main effect of drug condition ($F_{1, 19} = 11.61, p = .003, \eta_p^2 = .379$). This large effect size suggests that the caffeine group generally completed the overlearned TOH puzzle faster than the placebo group across the period of sleep deprivation (see Fig. 4). On average, the caffeine group was able to complete the task in 24.91 s (SE = 2.29), compared to the placebo

group, which completed the task in 35.95 s (SE = 2.29).

Single 6-Ring Puzzle. Only one participant (placebo) was unable to complete the 6-ring puzzle within 200 moves or less. A simple comparison of completion times between the drug conditions, controlling for baseline TOH 5-ring performance, suggested that caffeine ($M = 114.6, SD = 24.1$ s) provided significant advantage over placebo ($M = 165.5, SD = 76.9$ s; $F_{1,18}, p = .048$). The effect size ($\eta_p^2 = 0.20$) for this difference is large, suggesting a potential advantage of caffeine for sustaining this aspect of cognitive flexibility (i.e., the ability to consider and implement a novel solution) under extreme sleep loss.

3.3. Grooved Pegboard (GP)

Greenhouse-Geisser corrections were applied due to significant sphericity for the within-subjects analyses. There was no significant condition x night ($F_{1,54, 47,35} = 0.78, p = .437, \eta_p^2 = .044$), condition x block ($F_{2,11, 47,35} = 0.41, p = .68, \eta_p^2 = .023$), or condition x night x block ($F_{2,79, 47,35} = 1.77, p = .169, \eta_p^2 = .094$) interactions for the GP. There was, however, a main effect of drug condition ($F_{1, 17} = 8.23, p = .011, \eta_p^2 = .326$), suggesting that the caffeine group was generally faster at correctly placing the full complement of pegs than the placebo group across the three nights of sleep deprivation (see Fig. 5). On average, the caffeine group was able to complete the task in 55.04 s (SE = 1.84), compared to the placebo group, which completed the task in 63.38 s (SE = 2.25).

3.4. Symptom effects questionnaire (SEQ)

Participants in the caffeine group were more likely to report experiencing lightheadedness, nausea, and nervousness compared to the placebo group on all three nights. Caffeine was also associated with greater dizziness on night 1 and 3, and greater incoordination on night 2. Participants in the placebo group were more likely to report experiencing headaches on night 2 and 3, and less incoordination on night 2 compared to the caffeine group. Vomiting was not likely to be reported by either group (see Table 2).

3.5. SSS

Participants in the placebo group were more likely to report being sleepy on all nights compared to the caffeine group (see Table 2).

4. Discussion

On the whole, caffeine was superior to placebo for all tasks over the course of sleep deprivation, but the magnitude of observed benefits was often task specific and was moderated by the duration of sleep loss. We discuss the specific findings and their implications in the following sections.

4.1. Psychomotor vigilance under prolonged monotony

Prior sleep deprivation studies have generally focused on periodic assessments requiring only relatively brief periods (e.g., 5 to 10 min) of sustained vigilance collected every few hours, typically interspersed with other relatively undemanding tasks (e.g., mood or sleepiness scales) or generous breaks often involving relaxation, social contact, mild entertainment, or refreshments. Most relevant to this paper, each night between 0100 and 0900, when circadian pressure was accumulating, 36 PVT bouts were carried out in succession (i.e., approximately one bout every 10 to 15 min over an 8-h period). This schedule was designed to push the limits of human vigilance performance to allow a full evaluation of caffeine's effectiveness under conditions of accumulating homeostatic sleep drive, circadian pressure, mental fatigue, and prolonged tedium.

PVT Speed. Despite the high workload and extreme monotony,

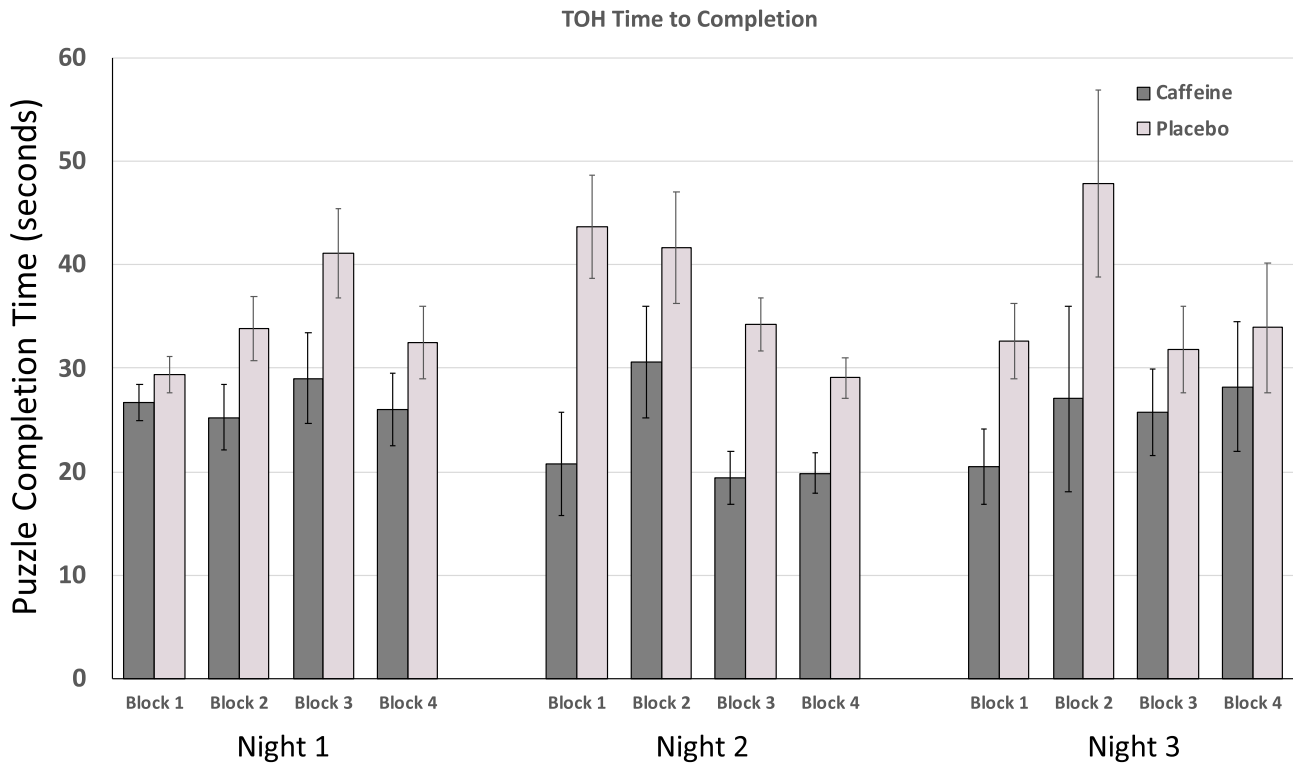


Fig. 4. Mean completion time for the Tower of Hanoi (TOH) puzzle at each block.

repeated overnight administration of caffeine maintained faster speed of PVT performance relative to placebo during all 3 nights. However, while caffeine consistently maintained faster responses than placebo throughout all sessions, there was still a significant progressive decline in performance speed each night for the caffeine group, which was particularly notable between the first and second nights. Previous work

has suggested that vigilance performance declines by approximately 25% for every 24-h of TSD (Belenky et al., 1994). However, in the present study both the caffeine and placebo groups demonstrated greater than a 50% decrease in speed from nights 1 to 2, suggesting that cognitive fatigue and time-on-task effects, above and beyond the effects of SD, may have impacted performance. These findings suggest that

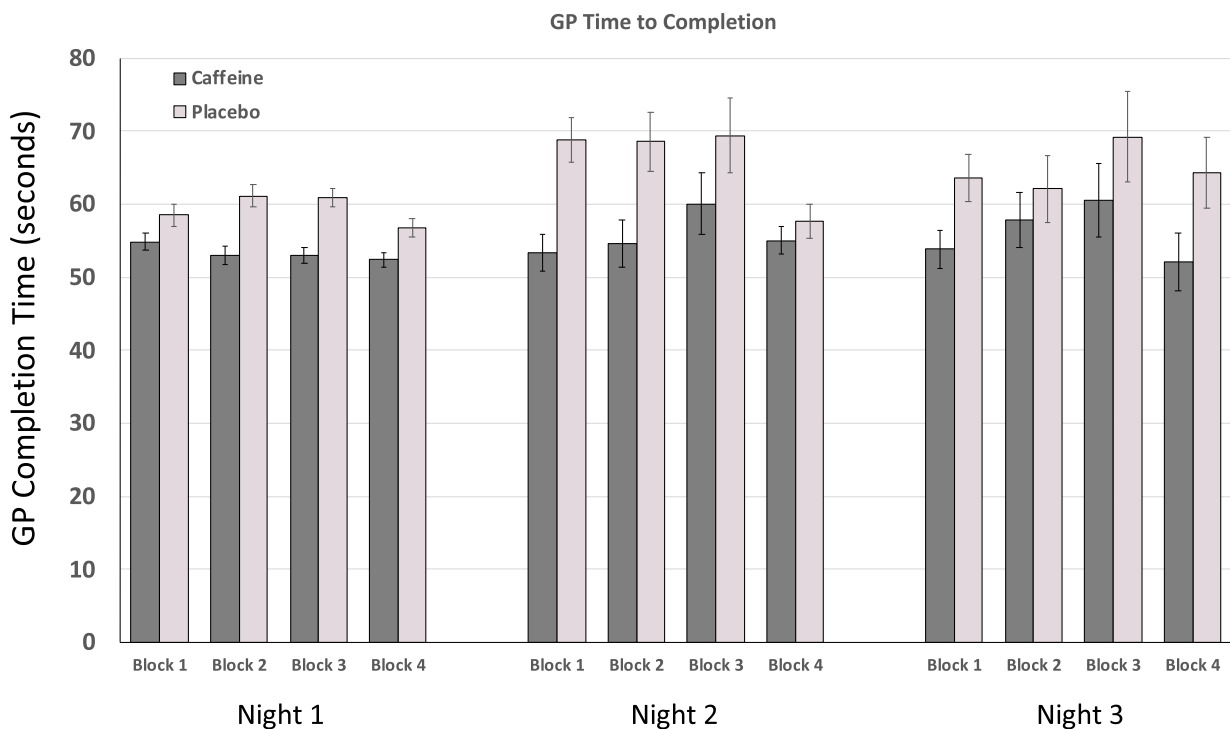


Fig. 5. Mean completion time for the Grooved Pegboard (GP) task at each block.

Table 2
Symptom group means and Odds Ratio (OR) estimates with 95% confidence intervals.

Symptom	Night	Caffeine M (SD)	Placebo M (SD)	OR	95% Confidence Interval	
					Lower	Upper
Lightheadedness	1	1.15 (0.27)	1.01 (0.06)	10.889	3.277	36.182
	2	1.15 (0.28)	1.03 (0.08)	7.739	2.836	19.197
	3	1.21 (0.57)	1.00 (0.02)	30.474	4.104	226.3
Nausea	1	1.10 (0.20)	1.02 (0.06)	4.607	1.712	12.401
	2	1.15 (0.32)	1.07 (0.12)	2.175	1.118	4.229
	3	1.11 (0.36)	1.02 (0.05)	6.650	1.943	22.757
Nervousness	1	1.07 (0.15)	1.00 (0.02)	14.203	1.858	108.547
	2	1.05 (0.12)	1.00 (0.00)	9.503	1.21	74.615
	3	1.10 (0.28)	1.00 (0.00)	24.233	3.263	179.959
Headache	1	1.10 (0.29)	1.12 (0.30)	0.883	0.622	1.253
	2	1.08 (0.17)	1.17 (0.38)	0.442	0.241	0.809
	3	1.04 (0.09)	1.04 (0.09)	0.210	0.098	0.450
Dizziness	1	1.13 (0.28)	1.00 (0.03)	14.65	3.439	62.407
	2	1.17 (0.35)	1.09 (0.19)	1.691	0.926	3.088
	3	1.13 (0.39)	1.00 (0.02)	20.903	2.798	156.145
Incoordination	1	1.08 (0.17)	1.03 (0.13)	2.416	1.035	5.641
	2	1.21 (0.41)	1.07 (0.20)	2.119	1.128	3.981
	3	1.09 (0.22)	1.09 (0.30)	0.72	0.356	1.457
Vomiting	1	1.00 (0.00)	1.00 (0.02)	-	-	-
	2	1.03 (0.10)	1.00 (0.02)	2.119	1.128	3.981
	3	1.00 (0.00)	1.00 (0.02)	-	-	-
Sleepiness	1	2.49 (0.68)	3.56 (0.88)	5.201	3.609	7.497
	2	3.78 (0.96)	4.93 (1.05)	4.032	2.825	5.754
	3	3.38 (1.31)	4.20 (1.24)	2.295	1.636	3.218

Bolded figures indicate significant difference between caffeine and placebo group on likelihood of reporting.
- indicates parameters cannot be calculated because all participants had the same response.

while caffeine is highly effective at maintaining psychomotor vigilance speed relative to placebo, it does not sustain this performance at baseline levels for more than a few hours during the first night of prolonged wakefulness. The reasons for this cannot be determined from the present data. A likely possibility is that by the second night, the buildup of adenosine is so great that the dose of caffeine used here was simply insufficient to provide adequate blockade of adenosine receptors. Another possibility is that the repeated dosing of caffeine on the first night resulted in upregulation of adenosine receptors and the development of tolerance to caffeine. Further research will be necessary to establish this. Nonetheless, from a practical standpoint, we find that after the first night, additional dosing of caffeine was not able to fully counter the cumulative effects of continued SD. Over severely prolonged wakefulness, the effects of repeated dosing of caffeine, while clearly better than placebo, are insufficient to restore vigilance response times normal rested levels, a finding that consistent with prior research (Paech et al., 2016).

PVT Lapses. Not only does SD slow response times, it also leads to instability of wakefulness and unpredictable lapses of attention, which are characterized by momentary failures to respond to a stimulus (Doran et al., 2001; Durmer and Dinges, 2005). Lapses of attention are of particular concern for many occupations. Failure to detect and respond appropriately to a stimulus at a critical moment can have disastrous consequences when high speeds, heavy equipment, or short response-windows are involved. Prolonged SD leads to increased frequency and duration of attentional lapses (Doran et al., 2001), which are most commonly defined as a response time on the PVT lasting longer than 500 ms. The 500-ms criterion has proven to be a useful metric and is widely used due to its sensitivity to sleep loss and ecological validity for many transportation-related accidents. Nonetheless, the 500-ms criterion is arbitrary, and it is possible that other lapse durations may be of particular relevance for understanding the effects of sleep loss and the usefulness of countermeasures such as caffeine. In fact, it is well known that as the period of sleep loss is extended, the duration of lapses will lengthen and eventually progress to uncontrollable “sleep attacks” in which wakefulness ceases (Goel et al., 2009). We therefore extended our assessment of lapses far beyond the

brief 500 ms lapses used in most studies. Here, we defined three forms of lapses of increasing severity, with potential ecological validity to many occupations. With increasing levels of severity, these lapses included 1) Attentional Lapses (ALs), which were defined as periods of disengagement lasting from one to 5 s during which time the individual spontaneously aroused and responded without prompting, 2) Responsive Lapses (RLs) lasting from five to 10 s and requiring some form of verbal or even mild physical intervention (i.e., shoulder tap) to restore arousal, and 3) Non-Responsive Lapses (NRLs) which were defined as lapses where arousal was not restored within 10 s despite two verbal prompts and one mild physical intervention.

Caffeine was more effective than placebo at reducing the number of ALs, but only for the first night of SD. During night 1, the placebo group showed a clear monotonic increase in the frequency of ALs as the night progressed, consistent with the expected effects of SD, monotony, and prolonged time-on-task effects (Lim and Dinges, 2008; Grant et al., 2017; Satterfield et al., 2017; Maire et al., 2014). In contrast, the caffeine group sustained performance throughout the night, with very few ALs observed. However, by the second night, ALs were common for both groups, and there was no significant difference between the caffeine and placebo groups at night 2 or night 3. This is consistent with findings from other studies using the standard definition of lapses (i.e., < 500 ms), which also show that the effects of caffeine are most apparent during the first night of SD but its advantage is significantly attenuated when the period of sleep loss is extended (Paech et al., 2016). Similar declines in the prolonged effectiveness of caffeine have been reported during multiple days of partial sleep restriction as well (Doty et al., 2017).

With increasing sleep pressure, it becomes more difficult to spontaneously arouse from an attentional lapse and the individual may need external stimulation restore attention or wakefulness. Compared to ALs, the longer RLs represent more severe disengagement that continues beyond 5 s and requires a verbal or physical prompt to restore attention or wakefulness. It is likely that RLs reflect the initial impingement of sleep phenomena onto wakefulness and that the period of disengagement would have persisted longer, perhaps leading to the full onset of sleep, without the intervention by the technician. Overall, caffeine

appeared to be protective against RLs. Across all three nights, caffeine was more effective than placebo at reducing the frequency of these more severe periods of attentional disengagement. Nonetheless, even with caffeine, RLs were still common during the second and third nights without sleep, reinforcing the notion that caffeine, in the doses used here, does not restore normal levels of wakefulness and vigilance during prolonged SD under highly monotonous and mentally fatiguing conditions.

Finally, mounting sleep pressure can become so overpowering that individuals are not easily aroused from a lapse even by multiple verbal and physical prompts. When such prompts did not rouse the participant within 10 s, a NRL was scored. The severity of these lapses is highly suggestive of initial entry into sleep. Overall, these NRLs were rare throughout the three nights of TSD. In fact, NRLs were rarely recorded on night 1 regardless of whether caffeine or placebo was administered. During subsequent nights, the placebo group demonstrated numerous NRLs, whereas the caffeine group showed fewer of these lapses, although the group difference was only statistically significant for night 2. These findings suggest that caffeine was modestly effective at protecting against NRLs, but this protection was far from complete, as these severe non-responsive episodes were occasionally observed in the caffeine group and became more frequent with each passing night of SD.

4.2. Execution of an overlearned sequence

Many occupations require individuals to carry out complex but well-learned sequences of activities at various times. To assess this capacity, we used a computerized version of the TOH (5-ring version). Participants were provided with a detailed step-by-step verbal and visual tutorial on how to implement a simple algorithm to solve the problem in the minimum number of moves, and were given ample practice opportunities to learn the skill to near-perfect criterion on day 1. In other words, by the time of overnight testing on SD night 1, the task was highly familiar, well-practiced, and consolidated during the baseline night of sleep (Nagai et al., 2017; Nettersheim et al., 2015).

Overall, participants remained effective at solving the 5-ring TOH puzzles across all three nights and performances did not change dramatically across the course of each night. This is likely due to the brief and engaging nature of the TOH (Harrison and Horne, 2000), which was administered half-way through each block of PVT assessments, and was generally a welcome break from the monotony of the repeated administrations of the PVT. Nonetheless, caffeine was associated with consistently faster TOH completion times than placebo, a finding that was particularly notable during the second night of SD. Interestingly, most performances during the overnight period were faster than baseline performances, particularly for the caffeine group, suggesting that participants continued to learn the task sequence between baseline and SD testing, but that the ability to continuously implement that cognitive-motor sequence during SD was better sustained by caffeine. Such findings are consistent with prior field research with sleep deprived Soldiers showing that caffeine is more effective than placebo at sustaining complex marksmanship activities such as target engagement time (i.e., time from target appearance to first shot fired), number of shots fired at enemy targets, correct identification of friend and foe (Tikuissis et al., 2004), and sighting time (Tharion et al., 2003), but had no appreciable effect on shooting precision (Tikuissis et al., 2004; Tharion et al., 2003). The present findings suggest that repeated 200 mg doses of caffeine bihourly during the overnight period are effective at sustaining performance on a well-learned complex sequence during periods of extended SD.

4.3. Mental flexibility and set shifting

It has been suggested that one higher-order cognitive effect of SD is a decrement in the ability to think flexibly and shift mental set when confronted with a novel problem (Honn et al., 2019; Whitney et al.,

2015; Couyoumdjian et al., 2010). To assess whether caffeine might prove useful in ameliorating this effect, we surprised participants with a novel version of the TOH on the final morning of the study, after 74.5 h of TSD (2.5 h after the last dose of caffeine). This required individuals to inhibit the previously overlearned solution and think flexibly to solve the new but similar appearing task. We found that after three nights of TSD, those who had received repeated dosing of caffeine were significantly faster at correctly solving the novel 6-ring version of the task relative to those who had received placebo. This finding has practical implications, suggesting that caffeine may sustain some aspects of cognitive flexibility relative to placebo despite repeated dosing and extreme sleep loss.

The present finding that caffeine enhanced the ability to solve a novel planning problem is consistent with our prior work showing that a single 600 mg dose of caffeine administered at 44 h of SD led to significantly fewer moves than placebo to solve a single novel administration of the 5-ring TOH administered 3.5 h later at 0630 (Killgore et al., 2009b). However, contrary to the present findings, caffeine in that study did not enhance the speed of responses. It is possible that the differences between findings could be accounted for by the differences in dosing (i.e., one single large bolus versus bi-hourly smaller doses), circadian timing of testing (pre-nadir versus post-nadir), or the complexity of the task (i.e., 5- versus 6-ring). However, in a prior report, we also showed that the same dosing of caffeine used here was more effective than placebo at sustaining sleep deprived performance across three administrations of the Tower of London (TOL), a similar planning and sequencing task (Killgore et al., 2014), suggesting that caffeine does appear to sustain this aspect of executive functioning. It is not clear from these data, however, whether this effect was due to greater immediate enhancement of TOH performance by caffeine (e.g., greater alertness), or whether the repeated practice solving earlier TOH puzzles while under the influence of caffeine was advantageous at helping those individuals learn general conceptual strategies about the task, which were subsequently implemented on the final 6-ring trial. The answer to this this interesting question will require additional research.

4.4. Manual dexterity and coordination

Many occupations that require occasional prolonged wakefulness also often necessitate the ability to maintain hand-eye motor coordination and manual dexterity (e.g., surgeons, military drone operators, automobile/machine operators). Despite their occupational importance, these capacities have been much less extensively studied in the context of SD than vigilance and cognitive performance. Limited evidence suggests that manual dexterity and surgical performance among physicians can be adversely affected by even a single night of SD (Hirkani and Yogi, 2017; Banfi et al., 2019). Similarly, SD affects several aspects of marksmanship involving psychomotor coordination, including target detection time, sighting time, and shooting accuracy (McLellan et al., 2005b; Tikuissis et al., 2004; Tharion et al., 2003). While caffeine is an effective countermeasure for sustaining vigilance and attention, some have raised concerns over hand tremors and possible effects on fine motor movements and dexterity (Humayun et al., 1997; Fargen et al., 2016). Therefore, as a measure of manual dexterity, hand-eye coordination, and psychomotor speed, the Grooved Pegboard (GP) was administered every 2 h throughout the course of the present study, making it an overlearned well-practiced task by the time of the assessment sessions. Similar to the procedure for the TOH, we assessed GP performance four times during each overnight sleep deprivation period between 0100 and 0900. Although we did not find any differences in performance within or across nights for either group, we found a clear advantage of the caffeine group relative to the placebo group overall. On the whole, participants who received caffeine completed the GP task significantly faster than those who did not. This argues in favor of repeated dosing of caffeine to sustain accurate hand-eye coordination and manual dexterity during periods of prolonged SD.

4.5. Side effect profile

Compared to the caffeine group, those in the placebo group were more likely to report sleepiness and headache, while those receiving caffeine were more likely to report lightheadedness, nausea, and nervousness on all 3 nights. Hence, among low to moderate habitual users of caffeine, repeated overnight administration appears to reduce negative symptoms caused by SD such as sleepiness and headache, but may be associated with increased physical sensations associated with nervousness, which are often attributed to caffeine.

4.6. Limitations

While the present data provide compelling evidence regarding the extent to which repeated doses of caffeine can and cannot sustain various aspects of sleep deprived performance during prolonged periods of monotony and high work load, these findings need to be interpreted in light of several caveats. First, the number of total PVT trials was not always consistent across test bouts. This occurred because the version of PVT used here was timed to last exactly 5 min, with randomized presentations of stimuli occurring from 1 to 5 s after each preceding response. Typically, there are 85 stimulus presentations. However, on those rare occasions where participants experienced NRLs (responses ≥ 10 s), the number of possible stimulus presentations was necessarily reduced. Second, our implementation of NRLs in this protocol may have also contributed to the potential for a floor effect, as anecdotal reports from lab technicians suggests that some NRLs may have reflected full onset of sleep. Consequently, performance in some sessions may have hit a near floor level without opportunity to show further declines. Third, evidence suggests that prior sleep history can affect subsequent performance during sleep deprivation (Rupp et al., 2009). Although participants were asked to maintain a regular sleep schedule prior to the study, and all participants completed a baseline night of sleep in the lab before the SD phase, we cannot rule out the potential contribution of prior sleep history on the present findings. Future studies would benefit from the use of wrist actigraphy to verify sleep history. Fourth, while the effects of SD and caffeine are relatively robust, our sample sizes were modest and may have obscured more subtle findings. Finally, our sample was composed of enlisted military personnel, who may have self-selected for service based on their trait-like capacity to sustain performance during prolonged periods of wakefulness (Van Dongen et al., 2004). Further work in civilian samples will be necessary before these findings can be validly generalized to the broader population.

5. Summary and conclusions

Under the tedious and mentally fatiguing conditions of this study, caffeine was more effective than placebo at sustaining several aspects of performance, including psychomotor vigilance speed and attention, complex motor sequence expression, cognitive flexibility, and manual dexterity/hand-eye coordination across three nights of TSD. Caffeine was particularly effective at sustaining PVT speed preventing simple attentional lapses on the first night, and minimizing more severe responsive lapses on the second night of TSD. Nonetheless, while clearly superior to placebo, caffeine failed to sustain psychomotor vigilance performance at normal rested baseline levels beyond the first few hours of SD. Even with repeated doses of caffeine, psychomotor vigilance speed and the frequency of attentional lapses continued to worsen with each passing night, with decrements exceeding more than 50% within the first 24-h. Caffeine is an effective countermeasure to vigilance and attention decrements during periods of TSD; however, caffeine does not maintain performance at baseline levels as sleep drive increases and should not be considered as a replacement for sleep.

Statement of significance

Optimal sleep cannot always be obtained in some operational/professional settings. It is, therefore, critical to understand the effectiveness and limitations of stimulant countermeasures such as caffeine, on alertness, vigilance, and higher cognitive processing. The present study is the first to investigate the efficacy of multiple repeated overnight dosing of caffeine during 3 nights of total sleep deprivation under highly monotonous and mentally fatiguing conditions. While significantly improving performance over placebo, caffeine was ineffective at restoring baseline levels of psychomotor vigilance beyond the first few hours of total sleep deprivation. Moreover, dangerous prolonged attentional lapses were prevalent after the first night even with 800 mg of total nightly caffeine. Caffeine is more effective than placebo but does not replace sleep.

Disclaimer

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its publication. The opinions or assertions contained herein are the private views of the authors, and are not to be construed as official, or as reflecting true views of the Department of the Army or Department of Defense. The investigators have adhered to the policies for protection of human subjects as prescribed in AR 70-25.

Financial disclosure

None.

Non-financial disclosure

None.

CRedit authorship contribution statement

William D.S. Killgore: Investigation, Formal analysis, Writing - original draft, Supervision, Project administration. **Gary H. Kamimori:** Conceptualization, Investigation, Formal analysis, Writing - review & editing, Supervision, Project administration.

Declaration of competing interest

The authors have no conflicts of interest to declare.

Acknowledgments

The authors wish to acknowledge the contributions of Christina LaValle, M.S. and Michael J. Egnoto, Ph.D. for their comments and suggestions regarding this manuscript, as well as the contributions of the staff of the WRAIR Department of Behavioral Biology, and the research volunteers who participated in this project.

References

- Balkin, T.J., Bliese, P.D., Belenky, G., et al., 2004. Comparative utility of instruments for monitoring sleepiness-related performance decrements in the operational environment. *J. Sleep Res.* 13 (3), 219–227.
- Banfi, T., Coletto, E., d'Ascanio, P., et al., 2019. Effects of sleep deprivation on surgeons dexterity. *Front. Neurol.* 10, 595.
- Banks, S., Dinges, D.F., 2007. Behavioral and physiological consequences of sleep restriction. *J. Clin. Sleep Med. : JCSM* 3 (5), 519–528 official publication of the American Academy of Sleep Medicine.
- Belenky, G., Penetar, D., Thorne, D., et al., 1994. The effects of sleep deprivation on performance during continuous combat operations. In: Marriott, B.M. (Ed.), *Food Components to Enhance Performance*. National Academy Press, Washington, D.C., pp. 127–135.
- Bendrick, G.A., Beckett, S.A., Klerman, E.B., 2016. Human fatigue and the crash of the

- airship Italia. *Polar Res.* 35 (1).
- Borbely, A.A., Achermann, P., Trachsel, L., Tobler, I., 1989. Sleep initiation and initial sleep intensity: interactions of homeostatic and circadian mechanisms. *J. Biol. Rhythm.* 4 (2), 149–160.
- Center for Disease Control and Prevention, 2012. Short sleep duration among workers - United States, 2010. *MMWR (Morb. Mortal. Wkly. Rep.)* 61 (16), 281–285.
- Couyoumdjian, A., Sdoia, S., Tempesta, D., et al., 2010. The effects of sleep and sleep deprivation on task-switching performance. *J. Sleep Res.* 19, 64–70.
- Czeisler, C.A., Wickwire, E.M., Barger, L.K., et al., 2016. Sleep-deprived motor vehicle operators are unfit to drive: a multidisciplinary expert consensus statement on drowsy driving. *Sleep Health* 2 (2), 94–99.
- Dinges, D.F., 1995. An overview of sleepiness and accidents. *J. Sleep Res.* 4 (S2), 4–14.
- Dinges, D.F., Powell, J.W., 1985. Microcomputer analyses of performance on a portable, simple, visual RT task during sustained operations. *Behav. Res. Methods Instrum. Comput.* 17, 652–655.
- Doran, S.M., Van Dongen, H.P., Dinges, D.F., 2001. Sustained attention performance during sleep deprivation: evidence of state instability. *Arch. Ital. Biol.* 139 (3), 253–267.
- Doty, T.J., So, C.J., Bergman, E.M., et al., 2017. Limited efficacy of caffeine and recovery costs during and following 5 Days of chronic sleep restriction. *Sleep* 40 (12).
- Durmer, J.S., Dinges, D.F., 2005. Neurocognitive consequences of sleep deprivation. *Semin. Neurol.* 25 (1), 117–129.
- Fargen, K.M., Turner, R.D., Spiotta, A.M., 2016. Factors that affect physiologic tremor and dexterity during surgery: a primer for neurosurgeons. *World Neurosurg.* 86, 384–389.
- Gander, P.H., Gregory, K.B., Miller, D.L., Graeber, R.C., Connell, L.J., Rosekind, M.R., 1998a. Flight crew fatigue V: long-haul air transport operations. *Aviat Space Environ. Med.* 69 (9 Suppl. 1), B37–B48.
- Gander, P.H., Gregory, K.B., Graeber, R.C., Connell, L.J., Miller, D.L., Rosekind, M.R., 1998b. Flight crew fatigue II: short-haul fixed-wing air transport operations. *Aviat Space Environ. Med.* 69 (9 Suppl. 1), B8–B15.
- Goel, N., Rao, H., Durmer, J.S., Dinges, D.F., 2009. Neurocognitive consequences of sleep deprivation. *Semin. Neurol.* 29 (4), 320–339.
- Gottlieb, D.J., Ellenbogen, J.M., Bianchi, M.T., Czeisler, C.A., 2018. Sleep deficiency and motor vehicle crash risk in the general population: a prospective cohort study. *BMC Med.* 16 (1), 44.
- Grant, D.A., Honn, K.A., Layton, M.E., Riedy, S.M., Van Dongen, H.P.A., 2017. 3-minute smartphone-based and tablet-based psychomotor vigilance tests for the assessment of reduced alertness due to sleep deprivation. *Behav. Res. Methods* 49 (3), 1020–1029.
- Harrison, Y., Horne, J.A., 2000. The impact of sleep deprivation on decision making: a review. *J. Exp. Psychol. Appl.* 6 (3), 236–249.
- Hirkani, M., Yogi, J., 2017. Effect of sleep deprivation on finger dexterity in resident doctors. *Natl. J. Physiol. Pharm. Pharmacol.* 7 (6).
- Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., Dement, W., 1973. Quantification of sleepiness: a new approach. *Psychophysiology* 10, 431–436.
- Honn, K.A., Hinson, J.M., Whitney, P., Van Dongen, H.P.A., 2019. Cognitive flexibility: a distinct element of performance impairment due to sleep deprivation. *Accid. Anal. Prev.* 126, 191–197.
- Humayun, M.U., Rader, R.S., Pieramici, D.J., Awh, C.C., de Juan Jr., E., 1997. Quantitative measurement of the effects of caffeine and propranolol on surgeon hand tremor. *Arch. Ophthalmol.* 115 (3), 371–374.
- Kahn-Greene, E.T., Lipizzi, E.L., Conrad, A.K., Kamimori, G.H., Killgore, W.D.S., 2006. Sleep deprivation adversely affects interpersonal responses to frustration. *Pers. Individ. Differ.* 41, 1433–1443.
- Kahn-Greene, E.T., Killgore, D.B., Kamimori, G.H., Balkin, T.J., Killgore, W.D.S., 2007. The effects of sleep deprivation on symptoms of psychopathology in healthy adults. *Sleep Med.* 8 (3), 215–221.
- Kalsi, J., Tervo, T., Bachour, A., Partinen, M., 2018. Sleep versus non-sleep-related fatal road accidents. *Sleep Med.* 51, 148–152.
- Kamimori, G.H., Lugo, S.I., Penetar, D.M., et al., 1995. Dose-dependent caffeine pharmacokinetics during severe sleep deprivation in humans. *Int. J. Clin. Pharm. Ther.* 33 (3), 182–186.
- Kamimori, G.H., Karyekar, C.S., Otterstetter, R., et al., 2002. The rate of absorption and relative bioavailability of caffeine administered in chewing gum versus capsules to normal healthy volunteers. *Int. J. Pharm.* 234 (1–2), 159–167.
- Kamimori, G.H., Johnson, D., Thorne, D., Belenky, G., 2005. Multiple caffeine doses maintain vigilance during early morning operations. *Aviat Space Environ. Med.* 76 (11), 1046–1050.
- Kamimori, G.H., McLellan, T.M., Tate, C.M., Voss, D.M., Niro, P., Lieberman, H.R., 2015. Caffeine improves reaction time, vigilance and logical reasoning during extended periods with restricted opportunities for sleep. *Psychopharmacology (Berlin)* 232 (12), 2031–2042.
- Killgore, W.D.S., 2010. Effects of sleep deprivation on cognition. *Prog. Brain Res.* 185, 105–129.
- Killgore, W.D.S., Kahn-Greene, E.T., Killgore, D.B., Kamimori, G.H., Balkin, T.J., 2007a. Effects of acute caffeine withdrawal on Short Category Test performance in sleep-deprived individuals. *Percept. Mot. Skills* 105 (3 Pt 2), 1265–1274.
- Killgore, W.D.S., Killgore, D.B., Day, L.M., Li, C., Kamimori, G.H., Balkin, T.J., 2007b. The effects of 53 hours of sleep deprivation on moral judgment. *Sleep* 30 (3), 345–352.
- Killgore, W.D.S., Lipizzi, E.L., Kamimori, G.H., Balkin, T.J., 2007c. Caffeine effects on risky decision-making after 75 hours of sleep deprivation. *Aviat Space Environ. Med.* 78 (10), 957–962.
- Killgore, W.D.S., Kahn-Greene, E.T., Lipizzi, E.L., Newman, R.A., Kamimori, G.H., Balkin, T.J., 2008a. Sleep deprivation reduces perceived emotional intelligence and constructive thinking skills. *Sleep Med.* 9 (5), 517–526.
- Killgore, W.D.S., McBride, S.A., Killgore, D.B., Balkin, T.J., Kamimori, G.H., 2008b. Baseline odor identification ability predicts degradation of psychomotor vigilance during 77 hours of sleep deprivation. *Int. J. Neurosci.* 118 (9), 1207–1225.
- Killgore, W.D.S., Kahn-Greene, E.T., Grugle, N.L., Killgore, D.B., Balkin, T.J., 2009a. Sustaining executive functions during sleep deprivation: a comparison of caffeine, dextroamphetamine, and modafinil. *Sleep* 32 (2), 205–216.
- Killgore, W.D., Kahn-Greene, E.T., Grugle, N.L., Killgore, D.B., Balkin, T.J., 2009b. Sustaining executive functions during sleep deprivation: a comparison of caffeine, dextroamphetamine, and modafinil. *Sleep* 32 (2), 205–216.
- Killgore, W.D.S., Kamimori, G.H., Balkin, T.J., 2011. Caffeine protects against increased risk-taking propensity during severe sleep deprivation. *J. Sleep Res.* 20 (3), 395–403.
- Killgore, W.D., Kamimori, G.H., Balkin, T.J., 2014. Caffeine improves the efficiency of planning and sequencing abilities during sleep deprivation. *J. Clin. Psychopharmacol.* 34 (5), 660–662.
- LaJambe, C.M., Kamimori, G.H., Belenky, G., Balkin, T.J., 2005. Caffeine effects on recovery sleep following 27 h total sleep deprivation. *Aviat Space Environ. Med.* 76 (2), 108–113.
- Lieberman, H.R., Tharion, W.J., Shukitt-Hale, B., Speckman, K.L., Tulley, R., 2002. Effects of caffeine, sleep loss, and stress on cognitive performance and mood during U.S. Navy SEAL training. *Psychopharmacology (Berlin)* 164 (3), 250–261.
- Lim, J., Dinges, D.F., 2008. Sleep deprivation and vigilant attention. *Ann. N. Y. Acad. Sci.* 1129, 305–322.
- Lim, J., Dinges, D.F., 2010. A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychol. Bull.* 136 (3), 375–389.
- Maire, M., Reichert, C.F., Gabel, V., et al., 2014. Time-on-task decrement in vigilance is modulated by inter-individual vulnerability to homeostatic sleep pressure manipulation. *Front. Behav. Neurosci.* 8, 59.
- Martiniuk, A.L., Senserrick, T., Lo, S., et al., 2013. Sleep-deprived young drivers and the risk for crash: the DRIVE prospective cohort study. *JAMA Pediatr.* 167 (7), 647–655.
- Matsui, K., Sasai-Sakuma, T., Ishigoaka, J., Inoue, Y., 2017. Insufficient sleep rather than the apnea-hypopnea index can be associated with sleepiness-related driving problems of Japanese obstructive sleep apnea syndrome patients residing in metropolitan areas. *Sleep Med.* 33, 19–22.
- Matthews, C.G., Klove, K., 1964. Instruction Manual for the Adult Neuropsychology Test Battery. University of Wisconsin Medical School, Madison, WI.
- McBride, S.A., Balkin, T.J., Kamimori, G.H., Killgore, W.D.S., 2006. Olfactory decrements as a function of two nights of sleep deprivation. *J. Sensory Stud.* 21, 456–463.
- McLellan, T.M., Bell, D.G., Kamimori, G.H., 2004. Caffeine improves physical performance during 24 h of active wakefulness. *Aviat Space Environ. Med.* 75 (8), 666–672.
- McLellan, T.M., Kamimori, G.H., Voss, D.M., Bell, D.G., Cole, K.G., Johnson, D., 2005a. Caffeine maintains vigilance and improves run times during night operations for Special Forces. *Aviat Space Environ. Med.* 76 (7), 647–654.
- McLellan, T.M., Kamimori, G.H., Bell, D.G., Smith, I.F., Johnson, D., Belenky, G., 2005b. Caffeine maintains vigilance and marksmanship in simulated urban operations with sleep deprivation. *Aviat Space Environ. Med.* 76 (1), 39–45.
- Nagai, H., de Vivo, L., Bellesi, M., Ghilardi, M.F., Tononi, G., Cirelli, C., 2017. Sleep consolidates motor learning of complex movement sequences in mice. *Sleep* 40 (2).
- Nettersheim, A., Hallschmid, M., Born, J., Diekelmann, S., 2015. The role of sleep in motor sequence consolidation: stabilization rather than enhancement. *J. Neurosci.* 35 (17), 6696–6702.
- Paech, G.M., Banks, S., Pajcin, M., et al., 2016. Caffeine administration at night during extended wakefulness effectively mitigates performance impairment but not subjective assessments of fatigue and sleepiness. *Pharmacol. Biochem. Behav.* 145, 27–32.
- Rupp, T.L., Wesensten, N.J., Bliese, P.D., Balkin, T.J., 2009. Banking sleep: realization of benefits during subsequent sleep restriction and recovery. *Sleep* 32 (3), 311–321.
- Satterfield, B.C., Wisor, J.P., Schmidt, M.A., Van Dongen, H.P.A., 2017. Time-on-Task effect during sleep deprivation in healthy young adults is modulated by dopamine transporter genotype. *Sleep* 40 (12).
- Stutts, J.C., Wilkins, J.W., Scott Osberg, J., Vaughn, B.V., 2003. Driver risk factors for sleep-related crashes. *Accid. Anal. Prev.* 35 (3), 321–331.
- Syed, S.A., Kamimori, G.H., Kelly, W., Eddington, N.D., 2005. Multiple dose pharmacokinetics of caffeine administered in chewing gum to normal healthy volunteers. *Biopharm Drug Dispos.* 26 (9), 403–409.
- Teichmann, A.T., 1990. Influence of oral contraceptives on drug therapy. *Am. J. Obstet. Gynecol.* 163 (6 Pt 2), 2208–2213.
- Tharion, W.J., Shukitt-Hale, B., Lieberman, H.R., 2003. Caffeine effects on marksmanship during high-stress military training with 72 hour sleep deprivation. *Aviat Space Environ. Med.* 74 (4), 309–314.
- Thiffault, P., Bergeron, J., 2003. Monotony of road environment and driver fatigue: a simulator study. *Accid. Anal. Prev.* 35 (3), 381–391.
- Thorne, D.R., Johnson, D.E., Redmond, D.P., Sing, H.C., Belenky, G., Shapiro, J.M., 2005. The Walter Reed palm-held psychomotor vigilance test. *Behav. Res. Methods* 37 (1), 111–118.
- Tikusis, P., Keefe, A.A., McLellan, T.M., Kamimori, G., 2004. Caffeine restores engagement speed but not shooting precision following 22 h of active wakefulness. *Aviat Space Environ. Med.* 75 (9), 771–776.
- Van Dongen, H.P., Baynard, M.D., Maislin, G., Dinges, D.F., 2004. Systematic inter-individual differences in neurobehavioral impairment from sleep loss: evidence of trait-like differential vulnerability. *Sleep* 27 (3), 423–433.
- Wesensten, N.J., Killgore, W.D.S., Balkin, T.J., 2005. Performance and alertness effects of caffeine, dextroamphetamine, and modafinil during sleep deprivation. *J. Sleep Res.* 14 (3), 255–266.
- Whitney, P., Hinson, J.M., Jackson, M.L., Van Dongen, H.P., 2015. Feedback blunting: total sleep deprivation impairs decision making that requires updating based on feedback. *Sleep* 38 (5), 745–754.