ORIGINAL INVESTIGATION

Peter J. Rogers · Susan V. Heatherley ·
Robert C. Hayward · Helen E. Seers · Joanne Hill ·
Marian Kane

Effects of caffeine and caffeine withdrawal on mood and cognitive performance degraded by sleep restriction

Abstract *Rationale:* It has been suggested that caffeine is most likely to benefit mood and performance when alertness is low. *Objectives:* To measure the effects of caffeine on psychomotor and cognitive performance, mood, blood pressure and heart rate in sleep-restricted participants. To do this in a group of participants who had also been previously deprived of caffeine for 3 weeks, thereby potentially removing the confounding effects of acute caffeine withdrawal. *Methods:* Participants were moderate to moderate–high caffeine consumers who were provided with either decaffeinated tea and/or coffee for 3 weeks (LTW) or regular tea and/or coffee for 3 weeks (overnight caffeine-withdrawn participants, ONW). Then, following overnight caffeine abstinence, they were tested on a battery of tasks assessing mood, cognitive performance, etc. before and after receiving caffeine (1.2 mg/kg) or on another day after receiving placebo. *Results:* Final analyses were based on 17 long-term caffeine-withdrawn participants (LTW) and 17 ONW participants whose salivary caffeine levels on each test day confirmed probable compliance with the instructions concerning restrictions on consumption of caffeine-containing drinks. Acute caffeine withdrawal (ONW) had a number of negative effects, including impairment of cognitive performance, increased headache, and reduced alertness and clear-headedness. Caffeine (versus placebo) did not significantly improve cognitive performance in LTW participants, although it prevented further deterioration of performance in ONW participants. Caffeine increased tapping speed (but tended to impair hand steadiness), increased blood pressure, and had some effects on mood in both groups. *Conclusions:* The findings provide strong support for the withdrawal reversal hypothesis. In particular, cognitive performance was found to be affected adversely by acute caffeine withdrawal and, even in the context of alertness lowered by sleep restriction, cognitive performance was not improved by caffeine in the absence of these withdrawal effects. Different patterns of effects (or lack of effects) of caffeine and caffeine withdrawal were found for other variables, but overall these results also suggest that there is little benefit to be gained from caffeine consumption.

P. J. Rogers (✉) · S. V. Heatherley · R. C. Hayward · H. E. Seers · J. Hill
Department of Experimental Psychology,
University of Bristol,
8 Woodland Road,
Bristol, BS8 1TN, UK
e-mail: Peter.Rogers@bristol.ac.uk
Tel.: +44-117-9288584
Fax: +44-117-9288588

M. Kane
National Diagnostics Centre,
National University of Ireland, Galway,
Galway, Ireland

Introduction

Caffeine is widely regarded as a useful psychostimulant (it is even sold as an over-the-counter medicine to "relieve tiredness and help maintain mental alertness"), although consumers also recognise that caffeine can disrupt sleep and that caffeine withdrawal is associated with various adverse effects, including fatigue and headache. What then are the overall or net benefits (if any) of caffeine consumption for psychological functioning? While the information available from research on the behavioural effects of caffeine is extensive, it does not provide an unequivocal answer to this question (James 1997; Rogers and Dernoncourt 1998). The reasons for this are primarily methodological. While many studies have found that caffeine (versus placebo) increases self-ratings of alertness, and improves mood and cognitive performance, in the vast majority of these studies the participants had a history of regular caffeine consumption, and they were tested after a substantial period of caffeine abstinence (i.e., withdrawal). What this experimental protocol leaves open is the question of whether the results obtained

are due to beneficial effects of caffeine or to deleterious, including fatiguing, effects of caffeine withdrawal. Furthermore, because significant fatigue results from even overnight caffeine withdrawal (e.g., Richardson et al. 1995), it is possible that the psychostimulant effects of caffeine felt by regular caffeine consumers represent only reinstatement of functioning (James 1994). That is, caffeine intake merely restores mood, alertness and performance to baseline levels (i.e., the levels displayed by individuals completely free of the effects of caffeine and acute caffeine withdrawal).

One way to investigate this question is to compare the effects of caffeine in caffeine consumers and non-consumers; although these groups are self-selected and therefore pre-existing differences might account for variation found in responses to caffeine administration (Goldstein et al. 1969; Rogers et al. 2003). Another approach is to test non-caffeine-withdrawn individuals (Smith et al. 1994; Warburton 1995; Warburton and Bersellini 2001), but again this is inconclusive, because the possibility that poorer performance in the placebo condition is due to residual caffeine withdrawal cannot be definitely excluded.

A much better approach to the problem is to measure the psychostimulant effects of an acute caffeine versus placebo challenge in long-term (i.e., 'fully') caffeine-withdrawn participants (LTW) compared with overnight caffeine-withdrawn participants (ONW). A net beneficial effect of caffeine consumption would be demonstrated if caffeine administration led to significant improvements in functioning in LTW as well as in ONW participants, and especially if caffeine brought both groups up to the same level of improved functioning. In contrast, the reinstatement or withdrawal reversal hypothesis predicts that without caffeine ONW participants will perform worse than LTW participants, and that caffeine administration will affect functioning only of ONW participants, bringing their level of functioning up to but not exceeding that of LTW participants. More or less exactly this result was obtained in two independent studies, one using a letter recognition task (James 1998) and the other using a long-duration simple reaction time task (Rogers et al. 1998; see also Bruce et al. 1991).

This methodology has therefore already contributed important new information relevant to estimating the impact of caffeine consumption in everyday life. However, its application to date has been restricted to the performance of simple tasks during the morning when alertness is relatively high. Consequently, it remains possible that there are circumstances under which caffeine consumption will bring about a significant net improvement of functioning. In particular, it has been suggested that caffeine is most likely to benefit performance when alertness is low (e.g., Johnson et al. 1990; Lorist et al. 1994; Smith et al. 1994; Horne and Reyner 1996; Lieberman et al. 2002; Reyner and Horne 2002; Wesensten et al. 2002). Accordingly, the present study investigated the effects of caffeine on self-rated mood and alertness, and performance of various cognitive and psychomotor tasks in LTW and ONW participants whose sleep had been restricted to 5 h on the night before testing. Overnight caffeine withdrawal was used because

this is a feature of many previous studies of the psychostimulant effects of caffeine (James 1997; Rogers and Dernoncourt 1998), and it is also part of everyday patterns of caffeine consumption (Smit and Rogers 2002).

Materials and methods

Participants

Forty-eight participants (29 female) aged between 20 and 34 years took part in the study. They were all students at the University of Bristol, and they participated for payment. The study was described as an investigation into the effects of a fruit squash drink on mood and cognitive function in the context of sleep deprivation. The participants were moderate to high caffeine consumers, all consuming at least three cups of tea or coffee a day (this information was gathered by email before the participants had been recruited to the study). They were either non-smokers or light smokers, and reported that they had no food sensitivities and were not pregnant or breastfeeding. All participants read an information sheet and signed a consent form before taking part in the study.

Study design and procedure

Participants were provided with supplies of tea and/or coffee and instructed to avoid all other caffeine-containing drinks during the 3 weeks immediately before testing. Twenty-three of the participants were provided with de-caffeinated tea and/or coffee (LTW) and 25 were provided with regular tea and/or coffee (ONW). They were told that the purpose of this part of the study was to investigate responses to changes in "brands of tea and coffee", and no reference was made to the caffeine content of the supplies. They were monitored closely during this period, which included periodic visits to the laboratory to replenish their supplies of tea and/or coffee, and to download data from electronic diaries used for measuring mood, and tea and coffee consumption patterns (details of this part of the study will be reported elsewhere—in preparation).

On the night before testing, participants' sleep was restricted to 5 h. Information collected on their usual sleeping patterns revealed that this was a reduction of between 1 and 5.5 h (mean 3 h) in a normal night's sleep for these individuals. In order to ensure that they slept for only 5 h, they arrived at the laboratory at 10.30 P.M., where they were supervised while they watched films until 2.30 A.M. After this they were returned home by taxi to get to sleep at around 3 A.M. They were then required to check into the laboratory again at 9 A.M., so it was anticipated they would wake about 8 A.M. in order to arrive on time.

These sleep-restricted participants were tested after overnight caffeine abstinence on a battery of mood and cognitive performance tasks and other measures before (baseline) and beginning 30 min after administration of 1.2 mg per kg body weight of caffeine and on a separate occasion after

placebo (order counterbalanced across participants). This is a dose of 84 mg of caffeine for a person weighing 70 kg, which is similar to the amount of caffeine contained in a serving of instant coffee (James 1997). Testing took place 2 days apart either in the morning or early afternoon (counterbalanced across participants) on Tuesdays and Thursdays. Note that this meant that half the LTW participants tested on Thursday had received caffeine 2 days previously. It was assumed that this single exposure to a modest dose of caffeine would not significantly affect their status as long-term withdrawn. The alternative procedure of withdrawing them perhaps again for a further 3 weeks was felt to be unnecessarily demanding and costly. On both occasions each participant provided a sample of saliva, which was analysed (see below) to check on compliance with the instructions regarding consumption of caffeine-containing drinks. Participants were tested alone, either in a room with an experimenter present (blood pressure, heart rate and hand steadiness), or accommodated in a separate, private booth (cognitive and psychomotor tasks and mood).

These procedures were approved by the University of Bristol, Department of Experimental Psychology Human Research Ethics Committee, and the study was performed in accordance with the ethical standards laid down by the 1964 Declaration of Helsinki.

Drug administration

Caffeine was administered double-blind in a blackcurrant drink (Blackcurrant Squash, Sainsbury's Supermarkets, UK, diluted according to manufacturer's instructions). Food grade caffeine hydrochloride was dissolved in this drink at a concentration 0.336 mg/ml and the volume of drink given to individual participants was varied so that each participant received a dose of 1.2 mg caffeine per kg body weight (e.g., 250-ml drink for a 70-kg person). The placebo drink was the same volume of blackcurrant squash drink, without caffeine.

Test battery

Various tasks were used to measure cognitive and psychomotor performance. Included in this battery were tasks which assessed the ability to sustain attention (e.g., simple reaction time and focus of attention tasks) and the ability to inhibit responses (impulsivity task), tasks which assessed memory, and a cognitively demanding, reasoning task. Two psychomotor tasks assessed tapping speed and hand steadiness. Mood, including alertness, some physical sensations (symptoms), and perceived task demand were measured using self-rating scales. Blood pressure and heart rate were also recorded. The cognitive performance and tapping tasks, and self-report rating scales were programmed and presented using E-Prime 1.0 (Psychology Software Tools, 1996–2001) run on networked PCs with 15-in. colour monitors and standard keyboards. Participants were tested in individual booths to ensure minimal unintended distraction. The tasks are described in detail below in the order that they were completed. The total time taken to complete this test battery was about 50 min.

Self-reported mood and physical sensations Participants were instructed to rate mood and physical sensation states according to how they were feeling "at the moment" using 9-point rating scales presented one at a time on the computer monitor. Ratings were made by keying the appropriate number from 1 to 9. Four aspects of mood were rated on bipolar scales: energetic mood (drowsy/sluggish–energetic/alert), tense mood (tense–relaxed), hedonic tone (sad/gloomy–happy/cheerful), and overall mood (bad mood–good mood). Physical sensations were rated on one bipolar scale and four unipolar scales. The physical sensation descriptors were clear-headed–muzzy/dazed, light-headed/feeling faint, jittery/shaky, and headache. On the latter four unipolar scales 1 represented "not at all" and 9 represented "extremely".

Hand steadiness Participants held a 1.59-mm-diameter stylus in holes drilled in a 1-mm-thick metal plate, which stood on a table so that it leaned away from the participant at an angle of 50° from the vertical. The holes measured 3.96, 3.18, 2.76, 2.36 and 1.98 mm in diameter (32011 Steadiness tester, Lafayette Instrument, IN, USA). Participants were instructed to hold the stylus in each hole for 30 s, as far as possible avoiding contact with the metal plate. They were permitted to use their preferred hand but not use their other hand or the table top for support. A buzzer indicated if the stylus touched the side of the hole, and chimes indicated the start and finish of the 30 s. The number of contacts and the time spent in contact (ms) were recorded for each hole using an interface and programme designed by a colleague (W.S. Maggs, Department of Electrical and Electronic Engineering.).

Systolic and diastolic blood pressure and heart rate Systolic and diastolic blood pressure and heart rate were measured using the Omron 711 Intellisense automatic inflation monitor (Omron Healthcare UK, West Sussex, UK).

Two-finger tapping task Participants were required to press alternately the n and m keys on the computer keyboard as quickly as possible using their first and index fingers of their preferred hand. They were told that the task would end automatically when they had tapped 300 times, and that this would be signalled on the computer monitor. The time taken to make 300 alternate taps was recorded.

Memory (immediate and delayed recall) Fifteen words were presented one at a time. They were displayed for 1 s with an inter-stimulus interval of 1 s. The words were nouns between four and seven letters long and they were all of similar word frequency and imagery. No plurals, or names were used. Before presentation of the words participants were told that they would be required to remember as many words as possible. Immediately after presentation of the words, then again 25 min later, participants were

given 2 min in which to write down all the words that they could remember. For both immediate and delayed recall the dependent variable was the number of words correctly recalled.

Long duration variable foreperiod Simple Reaction Time (SRT) task For this task participants were instructed to press the space bar as quickly as possible upon detection of a stimulus, a small star, presented in the centre of the computer screen. There was a variable stimulus onset of 1, 3, 7 and 16 s. The trials were divided into six blocks, each consisting of four of each of the variable stimulus onset durations occurring in random order. The dependent variable was mean reaction time.

Focus of attention This was based on the two-choice reaction time task developed by Eriksen and Eriksen (1974) (see also Broadbent et al. 1989). On each trial three warning crosses were presented on the computer monitor for 500 ms and then replaced by a target letter (A or B). This target was either presented alone or accompanied by distracter stimuli on both sides. The distracters were stars, letters the same as the target letter, or letters different from the target letter, that were positioned either near or far from the target. Participants were required to indicate whether the target was an A or B by pressing keys labelled A and B on the computer keyboard (A = J key and B = F key on the keyboard). Twenty-four continuous blocks of 16 trials were completed. The dependent variables were mean reaction time, number of errors, and Eriksen effect scores (Broadbent et al. 1989).

Reasoning task In this task (Baddeley 1999) participants were presented with a series of sentences, each describing the order of presentation of two letters, A and B. Each sentence was either followed by the letters AB, or BA. The participants' task was to decide whether the sentence correctly described the order of the letter pair. Participants were given the following example; "B follows A–AB" which is true and "B precedes A–AB" which is false. The sentences also varied in difficulty, for example, "A does not proceed B" and "B is not proceeded by A". There were 32 different sentences that ran continuously in random order for 3 min. Participants responded by pressing the F key labelled T (true) and J key labelled F (false) on the computer keyboard. The dependent variable was mean number of correct responses.

Impulsivity (inhibition) task The Test of Variables of Attention described by Leark et al. (1996) was adapted as follows. Participants saw different coloured squares appear successively one at a time in the centre of the computer monitor. They were required to press the space bar as quickly as possible on appearance of each coloured square, except for the blue one for which they were to respond to by doing nothing. There was no inter-trial interval (i.e., the next stimulus was presented immediately the space bar

was pressed). For blue squares, when the space bar was not pressed, the stimulus remained displayed for 2 s, after which it disappeared and the next trial followed immediately. There were six blocks of 65 trials (15 red, 15 green, 15 yellow, 15 pink and five blue), and within each block these trials occurred in random order. The dependent variables were the number of correctly withheld responses and mean reaction time of correct responses.

Self reported task demand On completing the test battery, participants rated how difficult, effortful and tiring they found the tasks to be. Responses were made on scales ranging from 1 to 9 where 1 represented "not at all" and 9 represented "extremely".

Habitual caffeine intake

Average daily caffeine consumption was calculated from drink intake questionnaires (Richardson et al. 1995) using the following estimates adapted from James (1997): instant coffee 90 mg/cup, tea 60 mg/cup, cola 30 mg/can, energy drinks such as Red Bull 80 mg/can, and caffeine tablets 50 mg/tablet.

Enzyme immunoassay for caffeine

A competitive enzymeimmunoassay for caffeine was established and validated against an HPLC assay and then used for analysis of caffeine levels in saliva samples. Wells were initially coated with rabbit anti-mouse immunoglobulins (DAKO Cat. no. Z0109) and then with anti-caffeine antibody (Biodesign Cat. no. G45110M AMS Biotechnology, UK) at 1 in 128,000 dilution. Standard solutions (5–500 ng/ml; 0.026–2.6 μM) of caffeine (Sigma Cat. no. C-8960) were prepared in phosphate-buffered saline containing 0.1% bovine serum albumin (PBS–BSA). The caffeine derivative, 7-(5-carboxypentyl)-1,3-dimethylxanthine (Department of Biochemistry, University of Surrey, UK), was conjugated to HRP using the mixed-anhydride method (Erlanger et al. 1957), as modified by Dawson et al. (1978). Salivary samples were thawed overnight before analysis, centrifuged at 10,000 rpm for 15 min using a Heraeus Sepatech Biofuge A centrifuge and supernatants diluted 1 in 100 in PBS–BSA. For the assay, 10 μl standard, control saliva or diluted unknown saliva was added to individual wells in duplicate, followed immediately by 100 μl caffeine–peroxidase conjugate at 1/800 dilution. Plates were shaken briefly and then incubated for 1 h at 37°C. After washing, bound peroxidase activity was determined using TMB as substrate. A standard curve of %B/ Bo (absorbance reading for each well (B) divided by the average absorbance for the zero standard wells (Bo) ×100) versus standard concentration was used to determine the concentration of caffeine in the samples.

Data analysis

Saliva samples taken on the two test days showed that a number of participants had clearly failed to comply with the instructions concerning avoidance of caffeine-containing drinks. These participants were excluded from the analyses of the effects of caffeine and caffeine withdrawal presented below. They were six participants from the LTW group who gave one or two saliva samples showing a caffeine concentration >500 ng/ml, and eight participants from the ONW group who gave one or two saliva samples showing a caffeine concentration >2,000 ng/ml. The use of these different 'thresholds' is based on the expectation that LTW participants who had complied would not have detectable levels of caffeine in their saliva (the assay was not reliably sensitive to caffeine concentrations below 500 ng/ml), but that ONW participants might have residual, but low systemic levels of salivary caffeine as a result of caffeine consumed the previous day. The latter assumption is supported by observations from a previous study (Heatherley et al., submitted) in which participants known to have complied with instructions to abstain from caffeine overnight, nevertheless had caffeine concentrations of between 545 and 1,834 ng/ml in saliva sampled during the morning or early afternoon the next day. Note that using the stricter criterion (<500 ng/ml) for the ONW group would have excluded participants with higher intakes of caffeine (throughout the day or in the evening) and/or slower caffeine-elimination rates, which would then have limited the generality of the findings.

The data on cognitive performance and mood were analysed to test the two main predictions of the withdrawal reversal hypothesis (see "Introduction"); namely that, (1) in the absence of caffeine LTW participants will perform better and have better mood than ONW participants, and (2) caffeine administration will affect performance and mood of ONW participants but not of LTW participants. This was done by examining ONW versus LTW differences in pre-treatment (baseline) mood and performance, and examining differences in their subsequent responses to caffeine, respectively (cf. Rogers et al. 2003). Pre-treatment data were analysed using ANOVA with 'withdrawal' (LTW versus ONW) and 'time of testing' (morning versus afternoon) as between-subjects factors. Note that the study design coupled with this method of analysis provides two estimates (one from the occasion when participants subsequently received placebo and one from when they subsequently received caffeine) of pre-treatment dependent variables for each individual for the crucial comparison of LTW versus ONW between-subjects effects. The effects of caffeine administration were analysed by calculating difference scores (post-treatment minus pre-treatment) and subjecting these to ANOVA with 'treatment' (caffeine versus placebo) as a within-subjects factor and 'withdrawal' and 'time of testing' as between-subjects factors. By controlling for 'trait' and 'state' differences in performance measured pre-treatment, this procedure greatly increases sensitivity in relation to detecting effects of caffeine versus placebo. Note that F ratios and P values for main and interaction effects calculated for these difference (change) scores are numerically the same as effects calculated for raw scores with the pre-/post-treatment factor included. Where appropriate, paired comparisons were carried out using t-tests.

Results

Table 1 shows that the two groups of participants (LTW and ONW) did not differ significantly in gender ratio, age, body weight, or habitual level of caffeine intake.

Tables 2, 3 and 4 show the results for cognitive performance, mood and blood pressure. The results in the left-hand panels of these tables show the effects of withdrawal (LTW and ONW) on performance etc., and the results in the right-hand panels show these participants' subsequent responses to caffeine, corresponding respectively to the first and second predictions of the withdrawal reversal hypothesis.

Prior to receiving caffeine the ONW and LTW participants differed on a number of measures of mood and cognitive performance in ways consistent with the withdrawal reversal hypothesis. Specifically, the ONW participants' performance on the focus of attention, reasoning, and impulsivity tasks was significantly worse or marginally significantly worse than the performance of the LTW participants on these tasks. There was also a substantial, though non-significant, difference in reaction time on the SRT task, and ONW participants reported that they found the tasks more tiring, and difficult (marginally significant). Ratings of mood and other feelings also showed marked negative effects of overnight caffeine withdrawal, with the ONW participants reporting heightened feelings of tension, light-headedness, and jitteriness (marginally significant), more headache, and lower clear-headedness, and energy (marginally significant) than the LTW participants. There were no significant pre-treatment differences be-

Dependent variable	Withdrawal		Effect of withdrawal (df=32)
	Long-term (LTW)	Overnight (ONW)	
n	17	17	
Gender (M:F)	7:10	6:11	
Age (years)	25.2±1.0[a]	24.2±0.6	t=0.79, P>0.1
Weight (kg)	66.7±3.1	65.3±2.6	t=0.33, P>0.1
Habitual caffeine intake (mg/day)	404±44	421±45	t=0.27, P>0.1

Table 1 Participant characteristics

[a]Data are means±SE.

Table 2 Results for cognitive and psychomotor performance[a,b]

		Pre-treatment performance (9-point scale)[a]		Post-treatment change in performance[a,b]				
		Mean±SE	$F_{withdrawal}$ (df=1,30)	Placebo, mean±SE	Caffeine, mean±SE	$F_{withdrawal}$ (df=1,30)	$F_{treatment}$ (df=1,30)	$F_{withdrawal×treat}$ (df=1,30)
Memory: immediate recall (# correct)	LTW	6.79±0.27	F<1	−0.41±0.37	−0.06±0.69	F<1	F<1	F<1
	ONW	6.82±0.39		−0.57±0.52	−0.43±0.55			
Memory: delayed recall (# correct)	LTW	4.97±0.40	F<1	−1.18±0.29	−1.42±0.50	F=1.42 P>0.1	F<1	F<1
	ONW	4.36±0.41		−0.48±0.55	−1.19±0.64			
SRT: response time (ms)	LTW	402±19	F=2.46 P>0.1	−1.4±10.7	2.9±20.3	F<1	F=3.05 P<0.1	F=4.25 P<0.05
	ONW	460±32		46.5±16.1	−7.1±19.8*			
Focus of attention: # errors	LTW	19.3±3.1	F=6.03 P<0.05	2.3±2.2	2.0±2.0	F=1.45 P>0.1	F=3.63 P<0.1	F=3.19 P<0.1
	ONW	32.3±4.3		12.1±4.5	2.5±4.3*			
Focus of attention: response time (ms)	LTW	447±13	F=2.81 P>0.1	−41.5±30.0	−42.2±13.6	F=1.47 P>0.1	F<1	F<1
	ONW	489±21		−21.1±14.5	−23.7±10.6			
Focus of attention: Eriksen effect	LTW	31.4±10.2	F=2.65 P>0.1	−12.3±15.7	−31.7±13.6	F=2.67 P>0.1	F<1	F<1
	ONW	11.9±7.7		−3.4±20.5	4.2±16.7			
Reasoning: # correct	LTW	69.0±4.1	F=5.37 P<0.05	8.69±5.10	2.31±2.99	F<1	F<1	F=2.51 P>0.1
	ONW	54.3±5.1		0.71±2.54	4.82±1.94			
Impulsivity: # of correctly withheld responses	LTW	17.7±1.7	F<1	−4.88±2.34	−3.00±2.21	F<1	F=1.52 P>0.1	F<1
	ONW	17.6±1.8		−5.77±3.06	−3.65±2.28			
Impulsivity: response time (ms)	LTW	294±11	F=3.54 P<0.1	−14.8±10.5	−20.3±9.1	F<1	F=1.94 P>0.1	F<1
	ONW	331±16		−4.3±11.8	−22.8±9.2			
Tapping: time taken (s)	LTW	30.9±2.0	F<1	0.39±0.53	−1.02±0.49	F=3.74 P<0.1	F=5.95 P<0.05	F<1
	ONW	31.4±1.7		1.72±0.97	0.32±0.52			
Hand steadiness: contact time (s)	LTW	13.4±1.1	F<1	−1.49±1.26	0.75±1.13	F=5.03 P<0.05	F=3.54 P<0.1	F<1
	ONW	13.6±1.2		−3.31±0.60	−2.05±0.72			

Pre-treatment data and statistics show the effects of caffeine withdrawal (long-term versus overnight, between-subjects factor), and post-treatment data and statistics show the effects of caffeine withdrawal and treatment (placebo versus caffeine, within-subjects factor) and their interaction.

LTW long-term withdrawn participants (n=17), *ONW* overnight withdrawn participants (n=17).

[a] Morning/afternoon was also included as a factor in these analyses. Effects of caffeine withdrawal and caffeine did not differ significantly (P>0.1) between morning and afternoon (results not shown).

[b] Data are post-treatment scores minus pre-treatment scores.

*P<0.05, versus placebo.

Table 3 Results for mood, physical sensations and task demand

		Pre-treatment mood, etc. (9-point scale)[a]		Post-treatment change in mood, physical sensations and task demand[a,b]				
		Mean±SE	$F_{withdrawal}$ (df=1,30)	Placebo mean±SE	Caffeine mean±SE	$F_{withdrawal}$ (df=1,30)	$F_{treatment}$ (df=1,30)	$F_{withdrawal×treat}$ (df=1,30)
Energetic mood (drowsy–energetic[c])	LTW	3.69±0.39	F=3.84 P<0.1	0.86±0.37	1.47±0.36	F=1.34 P>0.1	F=1.82 P>0.1	F<1
	ONW	2.65±0.35		0.62±0.31	0.73±0.34			
Tense mood (tense–relaxed[c])	LTW	6.01±0.39	F=5.81 P<0.05	-0.18±0.44	-0.38±0.25	F<1	F<1	F<1
	ONW	4.65±0.41		0.18±0.31	-0.03±0.25			
Hedonic tone (sad–happy[c])	LTW	5.53±0.32	F=1.56 P>0.1	0.18±0.27	0.41±0.27	F<1	F<1	F<1
	ONW	5.02±0.25		0.18±0.21	0.21±0.22			
Overall mood (bad–good[c])	LTW	5.70±0.37	F=2.04 P>0.1	0.00±0.26	0.50±0.30	F<1	F=2.29 P>0.1	F<1
	ONW	5.01±0.30		-0.10±0.31	0.23±0.34			
Clear-headed–muzzy/dazed[c]	LTW	5.05±0.53	F=5.15 P<0.05	-0.34±0.36	-1.03±0.42	F<1	F=3.47 P<0.1	F<1
	ONW	6.24±0.32		-0.38±0.32	-0.93±0.28			
Headache[d]	LTW	1.94±0.34	F=6.14 P<0.05	-0.65±0.29	-0.62±0.40	F<1	F<1	F<1
	ONW	3.15±0.35		-0.35±0.29	-0.35±0.28			
Light-headed/feeling faint[d]	LTW	0.86±0.17	F=6.16 P<0.05	-0.15±0.16	-0.68±0.36	F<1	F=4.30 P<0.05	F<1
	ONW	1.41±0.16		0.11±0.31	-0.44±0.21			
Jittery/shaky[d]	LTW	1.77±0.32	F=3.21 P>0.1	-0.29±0.24	0.08±0.38	F<1	F=3.21 P<0.1	F<1
	ONW	2.77±0.45		-0.59±0.32	-0.08±0.12			
Task demand: 'difficult'[d]	LTW	5.37±0.48	F=4.09 P<0.1	-0.62±0.38	-0.24±0.36	F<1	F<1	F=1.21 P>0.1
	ONW	6.73±0.48		-0.27±0.33	-0.42±0.32			
Task demand: 'effort'[d]	LTW	5.82±0.43	F<1	-0.29±0.42	-0.24±0.78	F<1	F<1	F<1
	ONW	6.15±0.38		-0.29±0.60	-0.41±0.68			
Task demand: 'tiring'[d]	LTW	6.09±0.47	F=7.74 P<0.01	-0.71±0.51	-1.06±0.60	F<1	F=2.53 P>0.1	F<1
	ONW	7.65±0.38		-0.06±0.40	-0.94±0.30			

Pre-treatment data and statistics show the effects of caffeine withdrawal (long-term versus overnight, between-subjects factor), and post-treatment data and statistics show the effects of caffeine withdrawal and treatment (placebo versus caffeine, within-subjects factor) and their interaction.

LTW long-term withdrawn participants (n=17), *ONW* overnight withdrawn participants (n=17).

[a] Morning/afternoon was also included as a factor in these analyses. Effects of caffeine withdrawal and caffeine did not differ significantly (P>0.1) between morning and afternoon (results not shown).

[b] Data are post-treatment scores minus pre-treatment scores.

[c] Bipolar scales (e.g. drowsy=1, energetic=9).

[d] Unipolar scales (not at all=1, extremely=9).

Table 4 Results for blood pressure and heart rate

	Pre-treatment blood pressure and heart rate (9-point scale)[a,c]		Post-treatment change in blood pressure and heart rate[a,b]				
	Mean±SE	$F_{withdrawal}$ (df=1,29)	Placebo, mean±SE	Caffeine, mean±SE	$F_{withdrawal}$ (df=1,30)	$F_{treatment}$ (df=1,30)	$F_{withdrawal \times treat}$ (df=1,30)
Systolic pressure (mmHg) LTW	119.3±2.5	F<1	−5.03±1.93	0.85±2.64	F<1	F=5.63 P<0.05	F<1
ONW	119.4±2.5		−5.90±2.09	−1.38±2.88			
Diastolic pressure (mmHg) LTW	74.6±2.6	F<1	−2.35±2.03	3.04±1.89	F<1	F=9.24 P<0.01	F<1
ONW	74.4±2.6		−2.88±2.21	2.46±2.06			
Heart rate (bpm) LTW	71.1±3.0	F=1.92 P>0.1	−10.3±2.5	−12.6±2.6	F=1.46 P>0.1	F=1.87 P>0.1	F<1
ONW	76.0±3.1		−13.5±2.7	−16.9±2.8			

Pre-treatment data and statistics show the effects of caffeine withdrawal (long-term versus overnight, between-subjects factor), and post-treatment data and statistics show the effects of caffeine withdrawal and treatment (placebo versus caffeine, within-subjects factor) and their interaction.

LTW long-term withdrawal participants (*n*=17), *ONW* overnight withdrawn participants (*n*=17).

[a]Morning/afternoon was also included as a factor in these analyses. Effects of caffeine withdrawal and caffeine did not differ significantly (*P*>0.1) between morning and afternoon (results not shown).

[b]Body weight was included as a covariate in these analyses.

[c]Data are post-treatment scores minus pre-treatment scores.

tween ONW and LTW participants for psychomotor performance (tapping and hand steadiness tasks), blood pressure or heart rate.

Caffeine (versus placebo) had effects on performance mood and blood pressure. In relation to performance, these effects were broadly as predicted by the withdrawal reversal hypothesis. For two of the cognitive performance tasks, SRT and focus of attention, there was a significant and a marginally significant treatment by caffeine interaction effect, respectively; due to the fact that caffeine significantly affected the performance of the ONW participants but not of the LTW participants. In both cases there was a further deterioration in performance from pre-treatment levels when the ONW participants received placebo, and caffeine prevented this (Table 2). Caffeine did not significantly affect the performance of the LTW participants on any of the cognitive tasks.

In contrast, the (rather weak) effects of caffeine on mood and other feelings occurred in both ONW and LTW participants. These effects were, relative to placebo, a decrease in light-headedness, and an increase in clear-headedness (marginally significant), but also an increase in jitteriness (marginally significant).

Caffeine relative to placebo also improved tapping performance, but tended to increase tremor (impair hand steadiness) (Table 2), and it increased blood pressure (Table 3). The magnitude of each of these effects of caffeine was very similar in the two groups of participants.

None of the effects of caffeine withdrawal and caffeine described above varied systematically with time of testing (morning versus afternoon) (*P*>0.1).

Finally, it is worth noting that analysis of raw post-treatment scores confirmed these various findings. ANOVA, with withdrawal (LTW versus ONW) and time of testing (morning versus afternoon) as factors, showed that in the placebo condition ONW participants performed significantly worse on the SRT, focus of attention (errors), reasoning and impulsivity tasks (response time), and reported significantly worse mood (energetic mood, overall mood, clear-headedness, light-headedness, jitteriness), greater headache, and greater task demand ('difficult' and 'tiring') than LTW participants (*P*<0.05). However, when they had received caffeine, ONW and LTW participants differed significantly (*P*<0.05) on only five variables, namely focus of attention (response time), energetic mood, headache, light-headedness and 'tiring'. ONW and LTW participants did not differ significantly (*P*>0.1) on tapping speed, hand steadiness, blood pressure and heart rate under either the placebo or caffeine conditions.

Discussion

A clear finding from this study was that acute (i.e., overnight) caffeine withdrawal was associated with negative effects, including impaired cognitive performance and the perception that the cognitive tasks were more difficult and tiring to perform, greater headache, reduced alertness and clear-headedness, and an increased feeling of light-

headedness. Feelings of tension and jitteriness were also increased, which might be further symptoms of acute caffeine withdrawal or effects of caffeine consumed the previous day (e.g., Goldstein et al. 1969; Rogers et al. 2003).

Some aspects of mood were weakly improved by caffeine and certain other feelings, such as light-headedness, were more clearly improved. Although these effects occurred in both the acutely (ONW) and long-term (LTW) caffeine withdrawn participants, they were not accompanied by improved cognitive performance in the LTW participants. That is, caffeine consumption failed to benefit the performance of participants who were free of the negative effects of acute caffeine withdrawal, even in the context of low alertness induced by sleep restriction. This is a key result, because it contradicts suggestions that caffeine is especially beneficial for performance when alertness is low (e.g., Johnson et al. 1990; Lorist et al. 1994; Smith et al. 1994; Horne and Reyner 1996; Lieberman et al. 2002; Reyner and Horne 2002; Wesensten et al. 2002). Caffeine did affect the cognitive performance of the ONW participants, but merely to prevent yet further deterioration in their poorer pre-treatment performance. In the study by Lieberman et al. (2002) participants were withdrawn from caffeine and 'almost totally sleep-deprived' for 72 h before caffeine administration. The authors state in relation to their findings that the 'typical dietary levels of caffeine intake by the subjects were not high' and therefore 'caffeine withdrawal would have been modest in these volunteers' (p. 254). The evidence provided for this, however, is unconvincing. No data on the usual caffeine intakes of the volunteers are presented and, although pre-study salivary levels of caffeine appeared to be low, information on what time of day the saliva samples were taken is not given. The latter omission is crucial, as the levels reported (mean $\cong$ 600 ng/ml) are similar to those found in the present study after overnight caffeine withdrawal.

The finding that the SRT and focus of attention tasks revealed effects of caffeine and caffeine withdrawal is consistent with the general finding that performance on such vigilance and continuous performance tasks is particularly sensitive to caffeine administration (e.g., James 1997; Rogers and Dernoncourt 1998). Memory performance, on the other hand, appears to be much less reliably affected by caffeine. In the present study the SRT and focus of attention tasks took up respectively 12 and 8 min of the 50 min taken to complete the entire task battery. It is unlikely that more or clearer performance effects would have been found had a higher dose of caffeine been used. The average dose received by participants in this study was 79 mg, and various previous studies show that the dose–response relationship for caffeine and cognitive performance effects is very 'flat' in the range above about 30 mg (Lieberman et al. 1987; Rogers and Dernoncourt 1998; Smit and Rogers 2000).

It is worth noting that sleep restriction did successfully lower alertness. Many of the participants complained of feeling tired during the testing sessions, and their mean pre-treatment alertness ratings (LTW = 3.69, ONW = 2.65) were markedly lower than those of caffeine non-consumers and caffeine consumers (NC = 5.20, C = 4.27) who participated in a similar study but who were not sleep-restricted (Rogers et al. 2003). Sleep restriction also appears to have degraded performance. Parallel data for the SRT task are follows: LTW = 402 ms and ONW = 460 ms (present study), and NC = 371 ms and C = 376 ms (Rogers et al. 2003).

One aspect of psychomotor performance, tapping speed, was improved by caffeine in both ONW and LTW participants; however, at the same time caffeine tended to impair hand steadiness. These observations confirm previous findings (e.g., Richardson et al. 1995). Indeed, increased tremor is a well-known effect of caffeine, and especially of moderate to high doses of the drug (e.g., James 1990, 1997; Arnold et al. 1993; Bovim et al. 1995; Miller et al. 1998). Furthermore, and again consistent with many previous reports (e.g., Goldstein et al. 1969; James 1990, 1997), there were negative effects of caffeine for both groups in relation to feelings of jitteriness (marginally significant) and also blood pressure. Pre-treatment blood pressure did not differ between LTW and ONW participants, and it was increased by caffeine quite substantially and to the same extent in both groups.

The approach of comparing the performance of LTW and ONW participants before and after an acute caffeine challenge has been used previously in only a very few studies (Bruce et al. 1991; James 1998; Rogers et al. 1998). The present results extend the results of those studies by measuring effects on a much wider range of performance and other variables and, crucially, by showing that caffeine did not benefit the LTW participants' performance even though their alertness had been lowered by restricting their sleep. James (1998) and Rogers et al. (1998) both found that caffeine failed to improve LTW participants' performance on sustained attention tasks, and indeed in the latter study, if anything, caffeine tended to impair performance (not significant). However, self-rated alertness of LTW participants was increased by caffeine in James's study, and to a lesser extent in the study by Bruce et al. (1991). The explanation for this mismatch between apparent alerting and performance effects of caffeine is uncertain, although one possibility, as noted previously (Rogers et al. 2003), is that the increase in alertness is a misinterpretation of other subjective effects of caffeine, including an increase in 'jitteriness' (e.g., Goldstein et al. 1969; and the present study).

Bruce et al. (1991) also included a tapping task in their study, and found that there was a trend for caffeine to improve performance in the ONW participants (actually 24 h caffeine withdrawn), but not in the LTW participants. Nevertheless, the small effect of the lower (250 mg) of the two doses of caffeine administered in this study was identical for LTW and ONW participants, which agrees with the present results (Table 2). Five hundred milligrams of caffeine failed to further improve tapping performance, or subject state, in LTW participants, whereas it did so for ONW participants. This perhaps suggests the presence of tolerance to certain performance-disrupting, adverse effects of caffeine in the ONW participants, which is lost with long-term withdrawal.

A surprising finding from the present study was the failure of one third of the participants to comply with the instructions concerning restrictions on consumption of caffeine-containing drinks. It may be that this happened in part because some participants were attempting (mistakenly) to ameliorate the impact of sleep restriction. Nonetheless, such poor compliance is a cause for concern, as inclusion of the non-compliers would have caused the study to underestimate the effects of caffeine consumption and caffeine withdrawal (data not shown). This may well be a general problem in this area of research, as measurement of systemic caffeine concentration has to date been used only rarely to check participant compliance in such studies. In the present study the use of different criteria for including LTW and ONW participants is justified by the fact that systemic caffeine concentration might well remain moderately elevated after overnight caffeine abstinence, especially in slow eliminators and/or in individuals who have consumed caffeine during the evening prior to testing. These are simply variations in everyday patterns of caffeine consumption. If anything, the criteria used would be expected to have reduced the likelihood of finding differences in responses of the LTW and ONW participants, because for the ONW participants presumably the effects of caffeine withdrawal and subsequent caffeine administration would be greatest when systemic caffeine levels are at their lowest.

Taken together, the findings from this study provide further strong support for the withdrawal reversal hypothesis. In particular, cognitive performance was found to be affected adversely by acute caffeine withdrawal, and cognitive performance was not improved by caffeine in the absence of these adverse effects. Different patterns of effects (or lack of effects) of caffeine and caffeine withdrawal were found for other variables, but overall these results also suggest that there is little benefit to be gained from caffeine consumption.

References

Arnold RW, Springer DT, Engel WK, Helveston EM (1993) The effect of wrist rest, caffeine, and oral timolol on the hand steadiness of ophthalmologists. Ann Ophthalmol 25:250–253

Baddeley AD (1999) Essentials of human memory. Psychology Press, Hove

Bovim G, Naess P, Helle J, Sand T (1995) Caffeine influence on the motor steadiness battery in neuropsychological tests. J Clin Exp Neuropsychol 17:472–476

Broadbent DE, Broadbent HP, Jones JJ (1989) Time of day as an instrument for the analysis of attention. Eur J Cogn Psychol 1:69–94

Bruce M, Scott N, Shine P, Lader M (1991) Caffeine withdrawal: a contrast of withdrawal symptoms in normal subjects who have abstained from caffeine for 24 hours and for 7 days. J Psychopharmacol 5:129–134

Dawson EL, Dennison EHL, van Weeman BK (1978) A simple and efficient method for raising steroid antibodies in rabbits. Steroids 31:357–366

Eriksen BA, Eriksen CW (1974) Effects of noise letters upon identification of target in a non-search task. Percept Psychophys 16:143–149

Erlanger B, Borek F, Beiser SM, Lieberman S (1957) Preparation and characterisation of conjugates of bovine serum albumin with testosterone and with cortisone. J Biol Chem 228:713–727

Goldstein A, Kaizer S, Whitby O (1969) Psychotropic effects of caffeine in man. IV. Quantitative and qualitative differences associated with habituation to coffee. Clin Pharmacol Ther 10:489–497

Horne JA, Reyner LA (1996) Counteracting driver sleepiness: effects of napping, caffeine and placebo. Psychophysiology 33:306–309

James JE (1990) The influence of user status and anxious disposition on the hypertensive effects of caffeine. Int J Psychophysiol 10:171–179

James JE (1994) Does caffeine enhance or merely restore degraded performance? Neuropsychobiology 30:124–125

James JE (1997) Understanding caffeine: a biobehavioral analysis. Sage, Thousand Oaks, CA

James JE (1998) Acute and chronic effects of caffeine on performance, mood, and sleep. Neuropsychobiology 38:32–41

Johnson LC, Spinweber CL, Gomez SA, Matteson LT (1990) Daytime sleepiness, performance, mood, nocturnal sleep: the effect of benzodiazepine and caffeine on their relationship. Sleep 13:121–135

Leark RA, Dupuy T, Greenberg L, Corman C, Kindschi C (1996) Test of variables of attention: professional manual. Universal Attention Disorders, Los Alamitos, CA

Lieberman HR, Wurtman RJ, Emde GG, Roberts C, Coviella ILG (1987) The effects of low doses of caffeine on human performance and mood. Psychopharmacology 92:308–312

Lieberman HR, Tharion WJ, Shukitt-Hale B, Speckman KL, Tulley R (2002) Effects of caffeine, sleep loss, and stress on cognitive performance and mood during U.S. Navy SEAL training. Psychopharmacology 164:250–261

Lorist MM, Snel J, Kok A (1994) Influence of caffeine on information processing in well rested and fatigued subjects. Psychopharmacology 113:411–421

Miller LS, Lombardo TW, Fowler SC (1998) Caffeine, but not time of day, increases whole-arm physiological tremor in non-smoking moderate users. Clin Exp Pharmacol Physiol 25:131–133

Reyner LA, Horne JA (2002) Efficacy of a 'functional energy drink' in counteracting driver sleepiness. Physiol Behav 75:331–335

Richardson NJ, Rogers PJ, Elliman NA, O'Dell RJ (1995) Mood and performance effects of caffeine in relation to acute and chronic caffeine deprivation. Pharmacol Biochem Behav 52:313–320

Rogers PJ, Dernoncourt C (1998) Regular caffeine consumption: a balance of adverse and beneficial effects for mood and psychomotor performance. Pharmacol Biochem Behav 59:1039–1045

Rogers PJ, Stephens S, Day JEL (1998) Contrasting performance effects of caffeine after overnight and chronic caffeine withdrawal. J Psychopharmacol 12:A13

Rogers PJ, Martin J, Smith C, Heatherley SV, Smit HJ (2003) Absence of reinforcing, mood and psychomotor performance effects of caffeine in habitual non-consumers of caffeine. Psychopharmacology 167:54–62

Smit HJ, Rogers PJ (2000) Effects of low doses of caffeine on cognitive performance, mood and thirst in low and higher caffeine consumers. Psychopharmacology 152:167–173

Smit HJ, Rogers PJ (2002) Effects of caffeine on mood. Pharma-copsychoecologia 15:231–257

Smith AP, Maben A, Brockman P (1994) Effects of evening meals and caffeine on cognitive performance, mood and cardiovas-cular functioning. Appetite 22:57–65

Warburton DM (1995) Effects of caffeine on cognition and mood without caffeine abstinence. Psychopharmacology 119:66–70

Warburton DM, Bersellini E (2001) An evaluation of a caffeinated taurine drink on mood, memory and information processing in healthy volunteers without caffeine abstinence. Psychopharma-cology 158:322–328

Wesensten NJ, Belenky G, Kautz MA, Thorne DR, Reichardt RM, Balkin TJ (2002) Maintaining alertness and performance during sleep deprivation: modafinil versus caffeine. Psychopharma-cology 159:238–247