Effects of repeated doses of caffeine on performance and alertness: new data and secondary analyses

Paul Hewlett and Andrew Smith*

Centre for Occupational and Health Psychology
School of Psychology
Cardiff University
63 Park Place, Cardiff CF10 3AS

* Corresponding author

Tel: +44 2920874757
Fax: +44 2920874758

E-mail: SmithAP@cardiff.ac.uk

ABSTRACT

Rationale The effects of caffeine on mood and performance are well established. Some authors suggest that caffeine merely reverses effects of caffeine withdrawal rather than having direct behavioural effects. It has also been suggested that withdrawal may be removed by a first dose of caffeine and further doses have little subsequent effect. These issues were examined here.

Objectives The present study aimed to determine whether caffeine withdrawal influenced mood and performance by comparing regular consumers who had been withdrawn from caffeine overnight with non-consumers. Following this repeated caffeine doses were administered to test the claim that repeated dosing has no extra effect on mood or performance. Secondary analyses of a data collected by Christopher et al. (2003) were also carried out to examine some alternative explanations of their results which showed effects of caffeine after a day of normal caffeine consumption.

Methods One hundred and twenty volunteers participated in the study. Regular caffeine consumption was assessed by questionnaire and this showed that thirty six of the sample did not regularly consume caffeinated beverages. Volunteers were instructed to abstain from caffeine overnight and then completed a baseline session measuring mood and a range of cognitive functions at 08.00 the next day. Following this volunteers were given 0, or 1mg/kg caffeine in a milkshake, glucose solution or water (at 09:00), followed by a second 0 or 1mg/kg caffeine dose (at 09:40) and the test battery repeated at 10:00.

Results The baseline data showed no effect of overnight caffeine withdrawal on mood or performance. In contrast, caffeine challenge improved vigilance performance and prevented decreases in alertness induced by completion of the task battery. The magnitude of these effects increased as a function of the number of doses of caffeine given. Secondary analyses of data from Christopher *et al*. (2003) also confirmed that effects of caffeine did not depend on length of withdrawal.

Conclusions The present findings show no effect of overnight caffeine withdrawal on mood and performance. Caffeine challenge did have the predicted effect on alertness and vigilance, with the size of the effects increasing with caffeine dose. These findings suggest that the effects of caffeine are not due to reversal of effects of withdrawal, a view confirmed by secondary analyses of data collected after a day of normal caffeine consumption.

KEY WORDS - Caffeine; caffeine withdrawal; alertness; cognitive performance

INTRODUCTION

It has been suggested that the behavioural effects of caffeine can be attributed to reversal of effects of caffeine withdrawal (James and Rogers, 2005). Many studies that have interpreted effects of caffeine in terms of reversal of caffeine withdrawal have not actually demonstrated effects of caffeine withdrawal and effects of caffeine in the same study (e.g. Durlach *et al.*, 2002; Smit and Rogers, 2000; Yeomans *et al.*, 2002). This issue has been examined by comparing withdrawn consumers with non-consumers of caffeine and recent research using this approach has largely failed to demonstrate negative effects of caffeine withdrawal (Haskell *et al.*, 2005; Hewlett and Smith, 2006a; Smith *et al.*, 2005). The present study also used this methodology to determine whether effects of caffeine withdrawal could be observed.

One type of evidence put forward to support the withdrawal hypothesis is the finding that repeated doses of caffeine (testing while non-withdrawn or following caffeine pre-load) do not produce greater effects than the initial dose (Robelin and Rogers, 1998; Yeomans *et al.*, 2002). Smith *et al.* (2003) have suggested that caffeine produces different types of effect depending on type of task (processing required) and arousal level of the volunteer. This is consistent with results showing that caffeine influences many different neurotransmitter systems (Fredholm *et al.*, 1999). Caffeine improves performance of tasks involving sustained encoding of new information (e.g. the repeated digits cognitive vigilance task, Smith *et al.*, 2003) even when the alertness level of the person is high. In contrast, the effects of caffeine on lapses of attention and improved performance on simple reaction time tasks are most easily observed when alertness is low. Other tasks, such as those involving episodic memory, are rarely affected by caffeine (see Smith, 2005). It is important, therefore, to consider the impact of caffeine withdrawal and subsequent caffeine challenge on all of these different tasks and this has rarely been attempted in previous studies. The present study had this methodological feature and examined effects of repeated doses of caffeine in alert individuals. It was predicted that caffeine would improve performance on the repeated digits vigilance task but have little effect on memory tasks or tasks that are largely improved by caffeine when alertness is reduced (simple reaction time and lapses of attention in choice reaction time tasks).

Previous studies of effects of repeated doses of caffeine often have deficiencies in design, analysis and interpretation. Yeomans *et al.* (2002) used a cross-over design and showed that order of caffeine treatments had a significant effect, with those who had caffeine first showing better performance. Given this effect of order, the analyses should have been restricted to the first run data, which, at least numerically, would have shown that a second dose of caffeine improves cognitive vigilance. Robelin and Rogers (1998) found that caffeine had no effect on performance of a simple reaction time task in the morning but reduced the post-lunch dip in performance seen in those given placebo. This can be interpreted as caffeine only having an effect on this task when alertness is reduced and that this effect occurs whether or not the person has had one or more caffeinated drinks. Heatherley *et al.* (2005) found that a second dose of caffeine only affected cognitive performance and mood after an 8-hour interval between doses. This is interpreted in terms of withdrawal effects only becoming apparent after 8-hours although no data are shown to indicate whether length of caffeine withdrawal influenced pre-challenge performance. If the groups did not differ pre-challenge it would not be possible to interpret the results in terms of withdrawal. Indeed, it may be the case that those in the 8-hour deprivation group were more fatigued prior to the second drink (due to factors other than withdrawal, e.g. having to get up earlier) and an effect in this group would then be consistent with caffeine having a larger effect in low alertness groups.

Other research has examined effects of caffeine after a day of normal consumption. James and Rogers (2005) have criticised these studies for failing to supervise consumption. Indeed, they suggest that the effects of caffeine reported by Christopher *et al.* (2005) could be due to a sub-group of participants who abstained from caffeine for > 6 hours. This issue is considered in a series of secondary analyses in the second section of this paper. Smith *et al.* (2005) examined effects of repeated doses of caffeine following a day of normal consumption. They tested volunteers when they were alert and fatigued, and were able to show effects of caffeine on both cognitive vigilance tasks and those measures that are sensitive to effects of caffeine when the person is fatigued (simple reaction time and lapses of attention in choice reaction time tasks). These effects were found to increase with repeated doses.

The aims of the present study were to examine possible effects of caffeine withdrawal by comparing non-consumers with withdrawn consumers. Following this the study examined effects of repeated doses of caffeine to determine whether effects were restricted to the first dose (Yeomans *et al.*, 2002) or whether a subsequent dose led to greater effects. A range of tasks were used and it was predicted that the repeated digits task would be improved by caffeine whereas simple reaction time, lapses of attention and episodic memory tasks would show no effect. Following this study secondary analyses of the data from Christopher *et al.* (2005) were conducted to determine whether length of caffeine withdrawal influenced the effects of caffeine.

MATERIALS AND METHODS

Design

A between subjects design was used with three factors: regular caffeine consumption dose of caffeine, and nature of the breakfast condition (see below), which determined the vehicle in which the caffeine was given. Regular caffeine consumers and non-consumers were randomly assigned to one of four conditions representing the presence or absence of caffeine in two drinks:

	Drink 1	Drink 2	
Condition 1:	Caffeine	Caffeine	(two doses of caffeine)
Condition 2:	Caffeine	No caffeine	(one dose of caffeine)
Condition 3:	No caffeine	Caffeine	(one dose of caffeine)
Condition 4:	No caffeine	No caffeine	(no caffeine)

Sample size

Studies of the effects of caffeine in alert individuals typically show an effect size of 0.8 SD so at least 24 participants were required per condition. The sample was larger because the caffeine manipulation was part of a larger study involving carbohydrate manipulations.

Participants

120 participants (66% female) took part in the study and these were recruited from the population of students of Cardiff University. Participants were divided into high, low and non-consumers on the basis of levels of regular caffeine consumption. Consumption data was missing for 7 participants and analyses involving this factor were based on an N of 113. Mean (s.d.) caffeine consumption was 212.0 (136.8) mg/day for the higher consumers (over 100mg/day), 53.6 (24.4) mg/ day for the lower consumers (100mg/day or less) and 0.0 mg/day for the non-consumers. Participants were screened for medical conditions/medications. Smokers were excluded to avoid any effects of tobacco or tobacco withdrawal. People with scores of over 55 on the

Spielberger Trait Anxiety Inventory were excluded. Details of the sample, sub-divided by regular consumption, can be seen in Table 1.

Insert Table 1 about here

Procedure

Before the testing day, volunteers were familiarised with the procedure and the test battery. All volunteers were made aware of the nature of the testing and the drinks (water, glucose or milkshake) that would be consumed. Written informed consent was obtained. On the testing day, participants arrived at the unit at 08:00 and carried out their baseline tests having had no breakfast and having had no caffeine or alcohol for at least seven hours and having done no strenuous exercise. This was followed by the appropriate drink type (approx 08:55). They then received the second drink (water) at 09:45 followed by the second test battery at 10:00. The session finished at approximately 10:55. Participants were allowed to drink water freely, but were asked to refrain from any other food or drinks during the testing period. Participants were paid £20.

Nature of breakfasts

Volunteers were randomly assigned to five breakfast conditions which are described in detail in Hewlett (2005):

1. Glucose drink
2. Glucose milk-shake
3. Water
4. Toast (and a drink of water)
5. Sucrose milk-shake

The caffeine was given in these 330 ml drinks.

Dose of Caffeine

The caffeine dose was 1mg/kg body weight.

MOOD AND PERFORMANCE TESTS

The mood and performance test battery (described in detail in Smith et al., 1999) was run on IBM-compatible computers. A response box attached to the computers contained a microchip, which controlled timing of the presentation of stimuli and timed the responses to the nearest millisecond.

The following tasks were chosen as they had previously been used to examine the effects of caffeine. Predictions could therefore be made about the effect of caffeine on these tasks.

Mood

Mood was assessed before and after each battery of tasks using visual analogue rating scales (described in detail in Smith et al., 1999). These have been factor analysed to produce 3 factors: Alertness, Hedonic tone and Anxiety.

Performance Tests (in the order completed)

Free recall (described in detail in Smith et al., 1999)

A list of 20 words was presented on the screen at a rate of one every 2 seconds. At the end of the list volunteers were required to write down as many as possible, in any order, within 2 minutes. Volunteers saw a different list at each test session. The variables analysed were: the number of words correctly recalled and the number of commission errors (words written down that were not in the presented list).

Focused attention choice reaction time task (Broadbent et al., 1986)

Target letters appeared as upper case A's and B's. On each trial three warning crosses appeared on the screen, the outside crosses being separated from the middle one by 1.02 or 2.60 degrees. Volunteers were required to respond to the target presented in the middle of the screen and ignore any distracters presented in the periphery. The crosses were on the screen for 500ms before being replaced by the target letter. The target letter was either accompanied by nothing, asterisks, letters that were the same as the target or letters that differed from the target. The two distracters were identical and the targets and the accompanying the letters were always A or B. The correct response to "A" was to press the A key with the forefinger of the left hand and the correct response to "B" was to press the B key with the forefinger of the right hand. Volunteers had 10 practice trials followed by five blocks of 64 trials. In each block there were equal numbers of near/far conditions, A or B responses and equal number of the four-distracter conditions. The nature of the previous trial was controlled. The scores derived from this task were:

- Mean reaction time.
- The speed of encoding of new information. This is a derived score taken as the difference in reaction time of response between conditions when the target is

alternated from the previous trial and when the target is repeated from the previous trial.

- The number of long responses (>800ms).
- Number of errors.

Categoric search choice reaction time task (Broadbent et al., 1986)

Each trial started with the appearance of two crosses in the positions occupied by the non-targets in the focused attention task (i.e. 2.04 or 5.20 degrees apart). Volunteers did not know which of the crosses the target would follow. The letter A or B was presented alone in half the trials and was accompanied by a digit (1-7) in the other half. The number of near/far stimuli, A verses B responses and digit/blank conditions was controlled. Half the trials led to compatible responses (i.e. the letter A on the left side of the screen or letter B on the right) and the other half were incompatible. The nature of the previous trial was also controlled. In respect to practice, number of trials etc., this task was identical to the focused attention. The variables measured were:

- Mean reaction time.
- The speed of encoding of new information. This is a derived score taken as the difference in reaction time of response between conditions when the target is alternated from the previous trial and when the target is repeated from the previous trial.
- The number of long responses (>1000ms).
- Number of errors.

Variable fore-period simple reaction time (described in detail in Smith et al., 1999)

A square was displayed on the screen and at varying intervals (between 1 and 8 seconds) a small filled white square appeared in the centre of the larger square. Volunteers were required to press a response key as soon as they detected the small square. The task lasted for 5 minutes.

Verbal reasoning (Baddeley, 1968)

Volunteers were shown a statement such as “A follows B: BA” and they had to decide whether the statement was a true description of the order of the letters and press the appropriate response key. Sentences ranged in syntactic complexity from simple active to passive negative (e.g. A is not followed by B). The task lasted for 5 minutes. The variables measured on this task were:

- Speed of performance - the number of trials completed within the 5 minutes.

- The percentage of correct responses.

Repeated digit detection task (described in detail in Smith et al., 1999)

Volunteers were shown three digit numbers on the screen at a rate of 100 per minute. Each number differed from the preceding one by one digit. Occasionally (8 times/minute) the same number was presented twice in succession. Volunteers had to detect these repetitions as quickly as possible. The task lasted for 5 minutes. Variables recorded were:

- Number of hits (correct detection of targets).
- Mean reaction time for hits.
- Number of false alarms (a response when no target was presented).

Because the literature sometimes shows an effect of caffeine on hit rate and sometimes RT, the RT and hit scores were combined to form a general efficiency score (hits /RT). This provides a measure whereby better performance (higher hit rate or faster RT) gives a higher score.

Semantic Memory (Baddeley, 1981)

This was a sentence verification task in which a series of sentences (e.g. "Crocodiles attend religious services") were presented. Participants were required to indicate as quickly as possible whether the statements presented were true or false. The variables measured were:

- The number of trials completed in the 5 minutes.
- The percentage of correct responses.

Delayed recognition memory (described in detail in Smith et al., 1999)

At the end of the testing session volunteers were shown a list of 40 words that included the 20 shown at the beginning of the testing session and 20 new distracters. Volunteers had to decide as quickly as possible which words were in the original list of twenty. The lists differed at each session. The variables measured were:

- Mean reaction time for the different response categories (hits, correct rejections, false alarms and misses).
- Number of correct responses (i.e. correctly accepting target words and correctly rejecting distracter words).
- The number of false alarms (i.e. wrongly accepting a distracter as being in the original list).

RESULTS

Excluded participants and missing data

Nine volunteers performed at chance level on the logical reasoning test and their data were excluded from the analyses of that task. Four volunteers were also excluded from the analysis of the semantic memory task because of their poor performance. There were missing data for 9 volunteers in the free recall task.

Withdrawal effects

Initial analyses were carried out on the baseline data to determine whether regular consumers who had abstained from consuming caffeine differed from non-consumers. These results are shown in Table 2. The only significant difference between the consumer groups was found in the simple reaction time task ($F(2,112) = 3.0$, $p<0.05$) where the reaction time of the lower consumers was faster than that of the higher consumers ($p<0.05$) and non-consumers ($p<0.05$).

Insert Table 2 about here

Post-drink effects

Table 3 shows the characteristics of the participants in the three different dose conditions. There were no significant differences between the dose conditions for these variables.

Analyses of covariance were carried out with the pre-drink measures as the covariate. The independent variables were type of breakfast (vehicle in which caffeine given) and dose of caffeine. The mood and performance measures from the second session were the dependent variables

Effects of breakfast

The effects of the type of breakfast on these tests are described in detail elsewhere (Hewlett, 2005) and can be briefly summarised as follows. Breakfast altered pre-session alertness, with those in the toast condition reporting highest alertness ($F(4,119) = 3.6$, $p<0.01$). Breakfast also had a significant effect on speed of encoding in the categoric search task ($F(4,119) = 3.5$, $p=0.01$). Those in the water condition were slower to encode new information than those in the toast and glucose milkshake conditions (both p's<0.05).

There were no significant interactions between the caffeine and breakfast conditions showing that any effects of caffeine did not vary as a function of the drink in which it was given.

Effects of repeated doses of caffeine on alertness

As expected, post-task alertness was lower than pre-task (F (1,119) = 30.6, p<0.001). Mean pre-task alertness was 240.3 (s.e. 4.2) and post-task it was 218.4 (s.e. 5.2). There was no difference between caffeine conditions pre-task (p>0.05). There was a statistically significant effect of caffeine on post-task alertness (F (2,119) = 8.9, p<0.001) where alertness was related to the number of caffeine doses. It can be seen in Figure 1 that the more caffeine consumed, the higher the alertness. Planned comparisons showed that alertness was approximately 14% higher in the one dose than no dose condition (p<0.01), 26% higher in the two dose than no dose condition (p<0.001) and 11% higher in the two dose than one dose condition (p<0.05 - see Figure 1).

Insert Figure 1 about here

Effects of repeated doses of caffeine on performance

There was a statistically significant effect of caffeine on performance of the vigilance task (F (2,119) = 9.4, p<0.001). Planned comparisons showed that performance scores were approximately 12% higher in the one dose than no dose condition (p<0.005), 19% higher in the two doses than no dose condition (p<0.001) and 6% higher in the two doses than one dose condition (see Figure 2).

Insert Figure 2 about here

None of the other tasks showed significant effects of caffeine (see Table 4).

Insert Table 4 about here

Caffeine dose and level of regular consumption

These analyses included dose of caffeine and regular caffeine consumption as independent variables. There was only one significant interaction between these variables (speed of encoding of new information, categoric search task – see Table 5) showing that the effects of different doses of caffeine were consistent across different consumer groups.

DISCUSSION

According to the withdrawal hypothesis, caffeine restores mood and performance degraded by the prior withdrawal of caffeine. The present results show similar scores for regular consumers after overnight withdrawal and non-consumers. James and Rogers (2005) suggest that non-consumers are a self-selected minority who may be atypically sensitive to caffeine or differ from consumers in other ways. In contrast to this view, Hewlett and Smith (2006b) found that the major difference between non-consumers and regular consumers was that non-consumers disliked the beverage in which caffeine was normally given. Smith *et al.* (2005) found no differences between consumers and non-consumers in terms of speed of metabolism of caffeine.

Some studies of the effects of repeated caffeine doses have reported an effect of the first caffeine dose but not an effect of a second and this has been taken as evidence for the withdrawal hypothesis. These studies can be criticised for assuming the presence of withdrawal effects in order to explain the findings. The present study removes the possible interpretation of reversal of withdrawal by demonstrating no withdrawal effect to reverse. James and Rogers (2005) were also critical of studies that allowed participants to provide their own caffeine pre-load. The present study controlled this in the laboratory. The suggestion that a pre-load may not remove withdrawal equally was shown not to be relevant as baseline comparisons showed no negative effect of caffeine withdrawal. Any effects of caffeine subsequently found must be explained some other way.

As expected, pre-task mood was unaffected in the present experiment. This is important as it is consistent with the idea that caffeine increases alertness when arousal is low and is consistent with other studies showing no effect of caffeine on alertness when arousal is high (see Lieberman, 1992). Also as predicted, caffeine challenge did increase alertness measured after performing the tasks. This supports other results showing beneficial effects of caffeine on alertness in low alertness situations (e.g. Smith *et al*, 1999; Lorist *et al*, 1994). Furthermore, this study showed that a second dose of caffeine produced a greater increase in alertness than a single dose. This was also observed for performance of the repeated digits vigilance task. Other tasks showed no effect of caffeine and this is consistent with the literature on caffeine in general (e.g. absence of effects on episodic memory tasks) or effects of caffeine in alert individuals (absence of effects of caffeine on simple reaction time of alert volunteers).Effects of caffeine on the speed of encoding have been found in alert participants (Smith *et al.*, 2003) and while the numerical trend was in the predicted direction, the effect was not significant in the present study. This could possibly reflect the nature of the drink masking the effect of caffeine.

The next section of the paper considers effects of caffeine following a day of normal consumption. A secondary analysis of data from Christopher *et al.* (2005) is presented to resolve some issues raised in a review of this study (James and Rogers, 2005).

Secondary analysis of Christopher et *al.* (2005)

Christopher *et al.* (2005) and Smith *et al.* (2005) have shown that caffeine can improve mood and performance in individuals who have been allowed to consume caffeinated beverages during the day prior to evening testing. In the early evening, caffeine was associated with greater alertness, faster speed of encoding of new information and better vigilance performance. As volunteers became fatigued during the evening (due to declining circadian alertness and prolonged task performance), simple reaction time was improved by caffeine and the number of long responses reduced by caffeine. These effects of caffeine increased with dose.

James and Rogers (2005) have criticised the Christopher *et al.* study because caffeine levels in saliva were high at the start of the day. This probably reflects the acute effects of caffeine ingestion on saliva levels (Walther *et al.*, 1983) as the volunteers may not have rinsed their mouths prior to providing the sample (a procedure that was applied when samples were collected in the laboratory later in the day). James and Rogers also suggested that the effects obtained by Christopher *et al.* may have been due to a small proportion of volunteers (> 6 hours deprivation) responding to caffeine with improved mood and performance, with the remainder (< 6 hours deprived) showing no effects. This was examined in the following set of secondary analyses.

Results

Fifteen (7 in the caffeine condition, 8 in the placebo condition) of the 68 volunteers did not consume any caffeinated beverages for more than six hours before the evening tests. Analyses of covariance were carried out distinguishing those who had not consumed caffeine for > 6 hours before testing and those who had consumed caffeine < 6 hours before testing. The results (see Table 6) showed that caffeine improved alertness and reduced reaction times in the repeated digits vigilance task for both groups (> 6 hours and < 6 hours withdrawal).

Insert Table 6 about here

CONCLUSION

In conclusion, the present study was unable to demonstrate differences in mood and performance between withdrawn consumers and non-consumers of caffeine. Caffeine challenge produced the predicted profile of behavioural changes and these effects increased with caffeine dose. Secondary analyses of data from a study of effects of caffeine following a day of normal consumption showed that caffeine had similar effects in those who had abstained from caffeine for > 6 hours and those who had consumed caffeine up to the time of testing. Overall, these results provide little support for the view that the behavioural effects of caffeine reflect the reversal of effects of caffeine withdrawal. Alternative explanations must be considered (see Smith, 2005) and these must be able to account for the task specificity of the effects of caffeine and the fact that effects will vary depending on the alertness level of the person.

ACKNOWLEDGEMENT

The research described in this paper formed part of the PhD thesis of the first author (Hewlett, 2005). Some of Professor Smith's caffeine research has been supported by the Institute for the Scientific Investigation of Coffee

REFERENCES

Baddeley A. 1968. A three minute reasoning test based on grammatical transformation. *Psychonomic Science* **10:** 341-342.

Baddeley A. 1981. The cognitive psychology of everyday life. *British Journal of Psychology* **72**: 257-269.

Broadbent DE, Broadbent MHP, Jones J L. 1986. Performance correlates of self-reported cognitive failure and obsessionality. *British Journal of Clinical Psychology* 25: 285-299.

Durlach PJ, Edmunds R, Howard L, Tipper SP. 2002. A rapid effect of caffeine on two choice reaction time tasks. *Nutritional Neuroscience* **5**: 433-442.

Fredholm B, Battig K, Holmen J, Nehlig A, Zvartau, E. 1999. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacological Reviews* **91:** 83-133.

Haskell CF, Kennedy DO, Wesnes KA , Scholey AB. 2005. Cognitive and mood improvements of caffeine in habitual consumers and habitual non-consumers of caffeine. *Psychopharmacology* **179**: 813-825.

Heatherley SV, Hayward RC, Seers HE, Rogers PJ. 2005. Cognitive and psychomotor performance, mood and pressor effects of caffeine after 4, 6 and 8 h caffeine abstinence. *Psychopharmacology* **178:** 461-470.

Herbert M., Johns M, Dore C. 1976. Factor analysis of analogue scales measuring subjective feelings before and after sleep. *British Journal of Medical Psychology* **49:** 373-379.

Hewlett PJ. 2005. Effects of caffeine and macronutrient variation on mood and cognitive performance. PhD Thesis, University of Wales Cardiff.

Hewlett P, Smith A. 2006a. Acute effects of caffeine in volunteers with differing patterns of regular consumption. *Human Psychopharmacology Clinical and Experimental* 21: 167-180.

Hewlett PJ, Smith AP. 2006b. Correlates of daily caffeine consumption. *Appetite* **46:** 97-99.

James JE, Rogers PJ. 2005. Effects of caffeine on performance and mood: withdrawal reversal is the most plausible explanation. *Psychopharmacology* **182:** 1-8.

Lieberman H. 1992. Caffeine. In Smith, A. and Jones, D. (Eds). *Handbook of human performance*. San Diego, Academic Press.

Lorist M, Snel J, Kok A. 1994. Influence of caffeine on information processing stages in well rested and fatigued subjects. *Psychopharmacology* **113:** 411-421.

Robelin M, Rogers P. 1998. Mood and psychomotor performance effects of the first, but not of subsequent, cup-of-coffee equivalent doses of caffeine consumed after overnight caffeine abstinence. *Behavioural Pharmacology* **9:** 611-618.

Smit HJ, Rogers PJ. 2000. effects of low doses of caffeine on cognitive performance, mood and thirst in low and high caffeine consumers. *Psychopharmacology* **152**: 167-173.

Smith A, Brice C, Nash J, Rich N, Nutt D. 2003. Caffeine and central noradrenaline: effects on mood, cognitive performance, eye movements and cardiovascular function. *Journal of Psychopharmacology* **17:** 283-292.

Smith A, Sturgess W, Gallagher J. 1999. Effects of a low dose of caffeine given in different drinks on mood and performance. *Human Psychopharmacology Clinical and Experimental* **14:** 473-482.

Smith A, Sutherland D, Christopher G. 2005. Effects of repeated doses of caffeine on mood and performance of alert and fatigued volunteers. *Journal of Psychopharmacology* **19**: 620-626.

Smith AP. 2005. Caffeine. In Lieberman H, Kanarek, R, Prasad C. (eds): *Nutritional Neuroscience.* Taylor & Francis. pp.335-359.

Walther H, Banditt P, Kohler E. 1983. Significance of caffeine values in serum, saliva and urine – determination of pharmacokinetic data by non-invasive methods in psychopharmacologic studies. *Pharmacopsychiatria* **16**: 166-170.

Yeomans M, Ripley T, Davies L, Rusted J, Rogers P. 2002. Effects of caffeine on performance and mood depend on the level of caffeine abstinence. *Psychopharmacology* **164:** 241-249.

Table 1. Age and personality scores of the different consumer groups (means and s.e.s)

Consumer status	Age	Imp	TA
Non (n=36)	22.8 (0.7)	4.4 (0.3)	39.5 (1.3)
Lower (n=39)	21.5 (0.7)	4.6 (0.3)	39.0 (1.3)
Higher (n=38)	21.2 (0.7)	4.2 (0.3)	39.5 (1.2)

TA = trait anxiety. Imp = impulsivity

Table 2. Mean (s.e.) baseline scores in the different consumer groups.

Variable	Consumer status (n)	Mean (s.e.)	Consumer group effect
Pre-session alertness (high scores=greater alertness)	Non (38) Lower (39) Higher (36)	220.7 (9.7) 214.9 (9.3) 203.3 (9.5)	F (2,112)=0.9, p>0.05
Pre-session Hedonic tone (high scores=more positive mood)	Non (38) Lower (39) Higher (36)	187.0 (6.4) 181.1 (6.1) 181.3 (6.2)	F (2,112)=0.3, p>0.05
Pre-session anxiety (high scores=greater calm)	Non (38) Lower (39) Higher (36)	93.7 (2.9) 88.4 (2.7) 86.9 (2.8)	F (2,112)=1.7, p>0.05
Post-session alertness (high scores=greater alertness)	Non (38) Lower (39) Higher (36)	189.9 (8.7) 180.1 (8.6) 195.0 (8.4)	F (2,112)=0.7, p>0.05
Post-session Hedonic tone (high scores=more positive mood)	Non (38) Lower (39) Higher (36)	172.3 (6.2) 164.0 (6.0) 166.9 (6.1)	F (2,112)=0.5, p>0.05
Post-session anxiety (high scores=greater calm)	Non (38) Lower (39) Higher (36)	90.5 (2.5) 88.1 (2.4) 86.2 (2.4)	F (2,112)=0.8, p>0.05
Free recall: Number recalled	Non (34) Lower (38) Higher (35)	10.6 (0.6) 10.6 (0.6) 10.6 (0.6)	F (2,106)=0.1, p>0.05
Free recall: Commission errors	Non (38) Lower (39) Higher (36)	0.3 (0.2) 0.5 (0.1) 0.7 (0.2)	F (2,106)=1.7, p>0.05
Focused attention Mean RT (ms)	Non (38) Lower (39) Higher (36)	383.2 (8.6) 384.4 (8.2) 372.3 (8.3)	F (2,112)=0.6, p>0.05
Focused attention Speed of encoding (ms)	Non (38) Lower (39) Higher (36)	22.7 (3.6) 21.2 (3.5) 24.4 (3.5)	F (2,112)=0.2, p>0.05
Focused attention Long responses	Non (38) Lower (39) Higher (36)	1.2 (0.4) 0.9 (0.4) 1.6 (0.4)	F (2,112)=0.6, p>0.05
Focused attention Errors	Non (38) Lower (39) Higher (36)	14.6 (1.7) 12.1 (1.7) 12.6 (1.7)	F (2,112)=0.6, p>0.05
Categoric search Mean RT (ms)	Non (38) Lower (39) Higher (36)	484.7 (10.2) 484.4 (9.8) 476.2 (10.0)	F (2,112)=0.2, p>0.05
Categoric search Speed of encoding (ms)	Non (38) Lower (39) Higher (36)	13.2 (3.5) 17.6 (3.3) 17.2 (3.4)	F (2,112)=0.5, p>0.05

Categoric search Long responses	Non (38) Lower (39) Higher (36)	12.2 (2.0) 11.4 (1.9) 14.2 (1.9)	F (2,112)=0.6, p>0.05
Categoric search Errors	Non (38) Lower (39) Higher (36)	17.7 (2.1) 16.5 (2.0) 16.7 (2.0)	F (2,112)=0.1, p>0.05
Simple reaction time (ms)	Non (38) Lower (39) Higher (36)	340.0 (9.6) 312.5 (9.2) 345.8 (9.3)	**F (2,112)=3.7, p<0.05**
Verbal reasoning Number of trials	Non (33) Lower (38) Higher (35)	86.9 (5.3) 86.6 (5.1) 93.3 (5.4)	F (2,105)=0.5, p>0.05
Verbal reasoning % correct	Non (33) Lower (38) Higher (35)	89.6 (1.4) 90.5 (1.4) 90.1 (1.5)	F (2,105)=0.1, p>0.05
Repeated digits Hits/RT x100	Non (38) Lower (39) Higher (36)	4.0 (0.2) 4.1 (0.2) 3.9 (0.2)	F (2,112)=0.4, p>0.05
Repeated digits False alarms	Non (38) Lower (39) Higher (36)	19.5 (2.0) 21.5 (1.9) 21.9 (1.9)	F (2,112)=0.5, p>0.05
Semantic memory Number of trials	Non (37) Lower (38) Higher (34)	110.3 (3.6) 111.6 (3.4) 109.6 (3.5)	F (2,108)=0.1, p>0.05
Semantic memory % correct	Non (37) Lower (38) Higher (34)	94.8 (0.5) 95.8 (0.5) 94.9 (0.5)	F (2,108)=1.2, p>0.05
Recognition memory Number correct	Non (38) Lower (39) Higher (36)	33.0 (0.7) 32.6 (0.7) 33.0 (0.7)	F (2,112)=0.1, p>0.05
Recognition memory Mean RT (ms)	Non (34) Lower (38) Higher (35)	820.7 (32.6) 865.7 (31.6) 896.9 (31.8)	F (2,112)=1.2, p>0.05

Table 3. Age, personality scores and daily caffeine consumption for the different caffeine dose groups (means and s.e.s)

Dose (mg/kg)	Age	Imp	TA	Caffeine/day (mg)
0	22.7 (0.8)	4.5 (0.3)	39.6 (1.5)	60.7 (21.2)
1	22.0 (0.5)	4.4 (0.2)	38.0 (1.0)	93.0 (15.2)
2	20.6 (0.8)	4.3 (0.3)	41.0 (1.5)	87.4 (21.6)

TA = trait anxiety. Imp = impulsivity

Table 4. Mean (s.e.) scores for mood and performance in the different caffeine dose conditions. Scores are adjusted means from ANCOVA.

Variable	Caffeine dose mg/kg (n)	Mean (s.e.)	Significance level
Pre-session Alertness	0 (30)	243.3 (6.6)	F(2,119)=1.5 p>0.05
	1 (60)	234.9 (4.7)	
	2 (30)	247.9 (6.6)	
Pre-session Hedonic tone	0 (30)	191.0 (4.2)	F(2,119)=1.9 p>0.05
	1 (60)	185.6 (3.0)	
	2 (30)	195.4 (4.2)	
Pre-session Anxiety	0 (30)	88.3 (2.3)	F(2,119)=0.8 p>0.05
	1 (60)	85.1 (1.6)	
	2 (30)	87.6 (2.3)	
Post-session Hedonic tone	0 (30)	181.8 (4.5)	F(2,119)=0.7 p>0.05
	1 (60)	183.3 (3.2)	
	2 (30)	189.0 (4.5)	
Post-session Anxiety	0 (30)	90.0 (2.1)	F(2,119)=0.7 p>0.05
	1 (60)	87.4 (1.5)	
	2 (30)	86.9 (2.1)	
Free recall Number recalled	0 (29)	9.7 (0.5)	F(2,110)=0.2 p>0.05
	1 (52)	9.7 (0.4)	
	2 (30)	9.4 (0.5)	
Free recall Commission errors	0 (29)	0.8 (0.2)	F(2,110)=0.1 p>0.05
	1 (52)	0.9 (0.1)	
	2 (30)	0.9 (0.2)	
Focused attention Mean RT (ms)	0 (30)	365.7 (4.3)	F(2,119)=1.5 p>0.05
	1 (60)	362.9 (3.0)	
	2 (30)	372.0 (4.3)	
Focused attention Speed of encoding (ms)	0 (30)	19.0 (2.6)	F(2,119)=1.1 p>0.05
	1 (60)	15.0 (1.8)	
	2 (30)	14.2 (2.6)	
Focused attention Long responses	0 (30)	0.6 (0.3)	F(2,119)=0.8 p>0.05
	1 (60)	1.0 (0.2)	
	2 (30)	0.9 (0.3)	
Focused attention Errors	0 (30)	17.6 (1.4)	F(2,119)=0.1 p>0.05
	1 (60)	17.5 (1.0)	
	2 (30)	16.9 (1.4)	
Categoric search Mean RT (ms)	0 (30)	462.3 (5.7)	F(2,119)=0.7 p>0.05
	1 (60)	470.5 (4.0)	
	2 (30)	468.7 (5.7)	
Categoric search Speed of encoding (ms)	0 (30)	13.2 (2.6)	F(2,119)=0.6 p>0.05
	1 (60)	10.9 (1.9)	
	2 (30)	9.2 (2.6)	
Categoric search Long responses	0 (30)	11.1 (1.2)	F(2,119)=1.9 p>0.05
	1 (60)	10.2 (0.8)	
	2 (30)	7.9 (1.2)	

Categoric search Errors	0 (30)	16.3 (1.4)	F(2,119)=0.3 p>0.05
	1 (60)	17.4 (1.0)	
	2 (30)	16.6 (1.4)	
Simple reaction time (ms)	0 (30)	328.2 (6.0)	F(2,119)=0.4 p>0.05
	1 (60)	322.7 (4.3)	
	2 (30)	321.4 (6.0)	
Verbal reasoning Number completed	0 (29)	101.5 (4.5)	F(2,110)=0.5 p>0.05
	1 (52)	100.2 (3.4)	
	2 (30)	99.4 (4.5)	
Verbal reasoning % correct	0 (29)	90.0 (1.0)	F(2,110)=1.9 p>0.05
	1 (52)	91.5 (0.8)	
	2 (30)	92.7 (1.0)	
Vigilance False alarms	0 (30)	17.8 (1.2)	F(2,119)=0.5 p>0.05
	1 (60)	18.4 (0.9)	
	2 (30)	19.5 (1.2)	
Semantic memory Number completed	0 (29)	116.0 (1.2)	F(2,115)=1.1 p>0.05
	1 (57)	117.5 (0.9)	
	2 (30)	118.4 (1.2)	
Semantic memory % correct	0 (29)	94.4 (0.5)	F(2,115)=1.0 p>0.05
	1 (57)	94.3 (0.3)	
	2 (30)	93.6 (0.5)	
Recognition memory Number correct	0 (30)	30.2 (0.7)	F(2,119)=1.8 p>0.05
	1 (60)	28.9 (0.5)	
	2 (30)	30.0 (0.7)	
Recognition memory Mean RT (ms)	0 (30)	899.5 (24.6)	F(2,119)=1.8 p>0.05
	1 (60)	865.3 (17.4)	
	2 (30)	919.3 (24.5)	

Table 5. Interactions between dose of caffeine and regular consumption.

Variable	Consumer	0mg/kg	1mg/kg	2mg/kg	Interaction
Pre-session Alertness	Higher	234.4 (13.0)	237.8 (8.2)	249.3 (15.9)	F(4,112)=0.5 p>0.05
	Lower	240.8 (11.4)	239.2 (9.8)	249.8 (11.8)	
	Non	255.7 (13.0)	229.6 (10.1)	246.2 (11.2)	
Pre-session Hedonic tone	Higher	194.7 (8.1)	187.7 (5.1)	199.6 (9.9)	F(4,112)=0.3 p>0.05
	Lower	184.9 (7.0)	186.0 (6.1)	197.2 (7.3)	
	Non	195.2 (8.1)	184.7 (6.3)	193.1 (7.0)	
Pre-session Anxiety	Higher	90.8 (4.3)	84.5 (2.7)	82.5 (5.2)	F(4,112)=0.7 p>0.05
	Lower	86.2 (3.7)	87.8 (3.2)	87.7 (3.9)	
	Non	88.3 (4.3)	84.2 (3.3)	90.2 (3.7)	
Post-session alertness	Higher	181.1 (16.3)	217.5 (10.2)	239.2 (20.0)	F(4,112)=0.4 p>0.05
	Lower	194.1 (14.2)	218.3 (12.2)	254.8 (14.8)	
	Non	200.1 (16.4)	224.2 (12.6)	234.9 (14.1)	
Post-session Hedonic tone	Higher	182.2 (8.6)	182.7 (5.3)	185.9 (10.5)	F(4,112)=0.2 p>0.05
	Lower	179.3 (7.4)	186.0 (6.4)	193.4 (7.7)	
	Non	183.2 (8.5)	184.5 (6.6)	186.7 (7.4)	
Post-session Anxiety	Higher	92.4 (4.1)	86.5 (2.5)	82.0 (5.0)	F(4,112)=0.6 p>0.05
	Lower	88.0 (3.5)	89.0 (3.0)	89.8 (3.7)	
	Non	91.2 (4.1)	88.6 (3.1)	87.8 (3.5)	
Free recall Correct	Higher	10.0 (0.9)	9.9 (0.6)	10.8 (1.1)	F(4,105)=1.1 p>0.05
	Lower	10.1 (0.8)	8.8 (0.7)	9.1 (0.8)	
	Non	9.0 (0.9)	10.2 (0.7)	9.0 (0.8)	
Free recall Commission errors	Higher	0.8 (0.4)	1.0 (0.2)	1.4 (0.4)	F(4,105)=0.9 p>0.05
	Lower	0.5 (0.3)	0.7 (0.3)	0.9 (0.3)	
	Non	1.2 (0.3)	0.8 (0.3)	0.6 (0.3)	
Focused attention Mean RT (ms)	Higher	367.3 (7.6)	365.8 (4.8)	366.4 (9.3)	F(4,112)=1.2 p>0.05
	Lower	365.1 (6.6)	357.3 (5.7)	382.3 (6.9)	
	Non	364.3 (7.6)	364.0 (5.9)	365.1 (6.6)	
Focused attention Speed of encoding (ms)	Higher	20.7 (4.8)	14.4 (3.0)	18.6 (5.7)	F(4,112)=0.7 p>0.05
	Lower	20.4 (4.2)	21.3 (3.6)	12.3 (4.4)	
	Non	16.2 (4.8)	12.1 (3.7)	11.6 (4.2)	
Focused attention Long responses	Higher	0.9 (0.5)	1.0 (0.3)	1.4 (0.6)	F(4,112)=0.4 p>0.05
	Lower	0.3 (0.4)	1.2 (0.3)	0.8 (0.4)	
	Non	0.6 (0.5)	1.0 (0.4)	0.9 (0.4)	
Focused attention Errors	Higher	18.9 (2.6)	17.1 (1.6)	18.2 (3.1)	F(4,112)=0.6 p>0.05
	Lower	16.2 (2.2)	14.5 (1.9)	15.3 (2.3)	
	Non	18.2 (2.6)	21.9 (2.0)	18.3 (2.2)	
Categoric search Mean RT (ms)	Higher	466.8 (10.2)	482.5 (6.4)	467.8 (12.5)	F(4,112)=0.8 p>0.05
	Lower	455.0 (8.8)	468.9 (7.6)	470.3 (9.2)	
	Non	467.2 (10.2)	459.3 (7.9)	467.1 (8.8)	
Categoric search Speed of encoding (ms)	Higher	20.0 (4.6)	12.4 (2.9)	2.9 (5.7)	**F(4,112)=3.0 p<0.05**
	Lower	13.4 (4.0)	13.8 (3.5)	1.4 (4.2)	
	Non	5.8 (4.7)	11.4 (3.6)	16.4 (4.0)	
Categoric search Long responses	Higher	13.9 (2.2)	9.4 (1.4)	8.3 (2.7)	F(4,112)=1.1 p>0.05
	Lower	11.3 (1.9)	11.4 (1.7)	7.2 (2.0)	
	Non	8.4 (2.2)	10.4 (1.7)	9.4 (1.9)	

Categoric search Errors	Higher	14.3 (2.5)	16.2 (1.6)	19.2 (3.1)	F(4,112)=0.8 p>0.05
	Lower	16.1 (2.2)	16.2 (1.9)	14.9 (2.3)	
	Non	18.9 (2.5)	21.2 (2.0)	16.6 (2.2)	
Simple reaction time (ms)	Higher	329.9 (11.1)	328.5 (6.8)	323.8 (13.4)	F(4,112)=0.1 p>0.05
	Lower	319.8 (9.5)	309.9 (8.3)	306.1 (10.1)	
	Non	336.3 (10.9)	327.8 (8.5)	326.9 (9.5)	
Verbal reasoning Number of trials	Higher	91.1 (8.8)	96.6 (5.8)	96.5 (10.2)	F(4,105)=0.8 p>0.05
	Lower	115.2 (7.2)	102.1 (6.4)	98.9 (7.6)	
	Non	95.5 (8.3)	102.1 (6.7)	101.3 (7.2)	
Verbal reasoning % correct	Higher	90.5 (1.8)	89.8 (1.2)	91.0 (2.1)	F(4,105)=1.0 p>0.05
	Lower	87.8 (1.5)	90.5 (1..3)	93.7 (1.6)	
	Non	92.0 (1.7)	93.4 (1.4)	93.0 (1.5)	
Vigilance Hits/RTx100	Higher	3.7 (0.2)	4.1 (0.2)	3.9 (0.2)	F(4,112)=0.9 p>0.05
	Lower	4.3 (0.1)	4.4 (0.2)	4.4 (0.2)	
	Non	4.6 (0.3)	5.0 (0.2)	4.3 (0.2)	
Vigilance False alarms	Higher	17.9 (2.3)	19.2 (2.0)	16.6 (2.3)	F(4,112)=0.7 p>0.05
	Lower	19.3 (1.4)	17.4 (1.7)	18.7 (1.8)	
	Non	24.0 (2.8)	18.7 (2.1)	18.6 (2.0)	
Semantic memory Number of trials	Higher	115.6 (2.2)	118.1 (1.4)	119.7 (2.7)	F(4,108)=0.5 p>0.05
	Lower	115.4 (1.9)	116.9 (1.7)	115.8 (2.0)	
	Non	118.4 (2.3)	117.6 (1.8)	121.0 (1.9)	
Semantic memory % correct	Higher	94.6 (0.9)	94.4 (0.6)	93.4 (1.1)	F(4,108)=1.1 p>0.05
	Lower	93.8 (0.7)	94.8 (07)	94.6 (0.8)	
	Non	95.3 (0.9)	94.0 (0.7)	92.9 (0.7)	
Recognition memory Number correct	Higher	29.7 (1.3)	28.1 (0.8)	29.1 (1.5)	F(4,112)=0.1 p>0.05
	Lower	29.6 (1.1)	28.8 (0.9)	19.7 (1.1)	
	Non	31.5 (1.2)	29.8 (1.0)	30.9 (1.1)	
Recognition memory Mean RT (ms)	Higher	871.4 (49.4)	849.9 (31.4)	938.6 (56.9)	F(4,106)=1.1 p>0.05
	Lower	959.2 (40.2)	860.5 (36.1)	893.9 (42.1)	
	Non	855.2 (46.5)	896.2 (37.2)	926.3 (42.5)	

Table 6: Effects of caffeine on alertness and repeated digits vigilance RT for volunteers who had abstained from caffeine for > 6 hours and < 6 hours (scores are the adjusted means, s.e.s in parentheses)

	Caffeine (abstained > 6hr)	Caffeine (< 6 hr)	Placebo (> 6hr)	Placebo (< 6hr)
Variable				
Alertness (high scores = greater alertness)	261.4 (13.4)	253.1 (6.5)	227.7 (11.9)	239.6 (6.7)
Repeated digits RT (ms)	671 (17)	666 (9)	691 (17)	692 (9)

Figure 1. Post-task mood: Mean alertness scores in the caffeine conditions. Higher score = greater alertness. Scores are adjusted from ANCOVA and s.e. are shown as error bars.

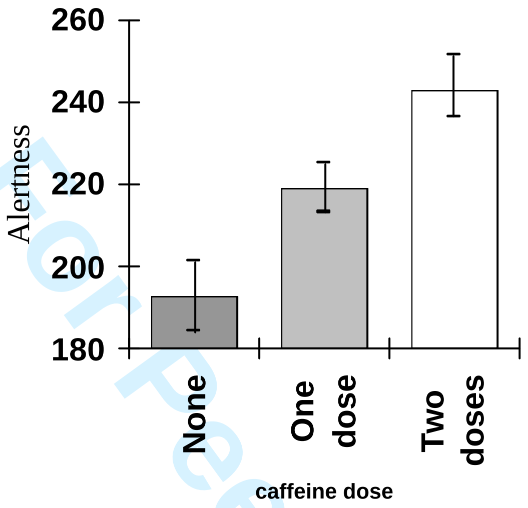

Figure 2. Vigilance task: Mean performance scores in the caffeine conditions. Higher scores = better performance. Scores are adjusted means from the ANCOVA and s.e.s are shown as error bars.