Pharmacology Biochemistry and Behavior, Vol. 63, No. 4, pp. 589–598, 1999

0091-3057/99/$–see front matter

ELSEVIER

PII S0091-3057(99)00022-2

Caffeine and Stroop Interference

J. LEON KENEMANS, JOAN S. T. WIELEMAN, MIJKE ZEEGERS AND MARINUS N. VERBATEN

Department of Psychopharmacology, Faculty of Pharmacy, Rudolf Magnus Institute for Neurosciences, Utrecht University, 3508 TB Utrecht, The Netherlands

Received 23 September 1998; Revised 19 January 1999; Accepted 19 January 1999

KENEMANS, J. L., J. S. T. WIELEMAN, M. ZEEGERS AND M. N. VERBATEN. *Caffeine and stroop interference.* PHARMACOL BIOCHEM BEHAV **63**(4) 589–598, 1999.—In two experiments the hypothesis that caffeine reduces Stroop interference was tested. In the first experiment interference was measured as the reduction in performance when subjects had to indicate the numerosity of strings of incongruent digits, relative to neutral-symbol strings. In the second experiment the incongruent condition consisted of naming the color of words referring to incongruent colors, and was compared to color naming of neutral strings. In the number-digit task 250 mg caffeine reduced interference at the level of error rates, relative to placebo. In the color-word task interference was reduced at the level of reaction times. These results were obtained with blocked presentations of incongruent and neutral conditions. The color-word experiment also contained a condition in which neutral and incongruent trials were mixed within one sequence. In this mixed condition caffeine still reduced overall reaction times, but no longer specifically interference. It is argued that this dissociation reflects a caffeine-induced increase in flexibility. The results are discussed in relation to failed previous attempts to demonstrate increased selectivity under caffeine using non-Stroop tasks, the importance of including pretreatment sessions to detect artificial effects, and the possible contribution of withdrawal effects. © 1999 Elsevier Science Inc.

Caffeine Selective attention Stroop interference Withdrawal

REPORTS on beneficial effects of caffeine in various information-processing tasks are abundant. In vigilance and related tasks in which occasional target stimuli have to be detected, caffeine consistently increases speed and rate of detection [e.g., (4,10,12,27)]. In choice reaction-time tasks, in which overt decisions have to be made for every stimulus, caffeine also generally results in shortening of reaction times [e.g., (1,6,11,14)]. The present article addresses the possibility that at least part of these beneficial effects is mediated by caffeine-induced selectivity of information processing or attention. According to the traditional view, stimulant drugs act to arouse the CNS, which may in turn result in increased selectivity of information processing (1,3,7,10). A classic way to probe such selectivity is to have subjects perform varieties of the so-called "Stroop" task (26).

Generally, the Stroop task consists of two conditions. In an incongruent condition two spatially integrated pieces of information are presented approximately simultaneously—one that the task demands should define the response, and another that evokes a conflicting response tendency that must be suppressed. In a second neutral condition the irrelevant piece of information does not, or to a lesser extent, evoke the conflicting response tendency [for a review, see (16)]. The one famous example is to have subjects name the color in which incongruent color names are written (the neutral condition being one with noncolor-name stimuli). The difference in performance between the two conditions, or the "Stroop interference" score, can be used as a measure of selectivity. According to the dominating view, patterns of interference are determined by the relative strengths of the functional connections between input and output representations (2,21). For example, normally the connection between a color-word stimulus input and the tendency to read the word is stronger than the connection between that same input and the tendency to name the color the word has been written in.

Studies on the effects of stimulants on Stroop interference have yielded somewhat inconsistent results. As to nicotine, both a decreasing effect (9) as well as no (clear) effects have been reported (22,28). The same holds for methamphetamine (9), while the one study on ephedrine found no effect (17). When stimulants are expected to reduce interference then depressants may be expected to increase it. Increasing effects were indeed found for promethazine (20), scopolamine (28), and amobarbital, but not for pentobarbital (9).

A particular inconsistency has arisen with respect to the effects of caffeine. Whereas Foreman and colleagues (5) found

Requests for reprints should be addressed to J. L. Kenemans, Department of Psychopharmacology, University of Utrecht, P.O. Box 80082, 3508 TB Utrecht, The Netherlands.

caffeine to actually increase Stroop interference, Hasenfratz and Bättig (6) reported a reducing effect. Although task parameters were highly comparable across the two studies, there were clear differences with respect to more general aspects of the procedure. First, Foreman and colleagues used a between-subjects design with three groups (placebo, and 125 and 250 mg caffeine) and no pretreatment test sessions. In contrast, Hasenfratz and Bättig tested subjects on four different occasions (placebo, 250 mg caffeine, smoking, and caffeine + smoking); they also included pretreatment sessions, the results of which were accounted for statistically. One possibility is that the results of Foreman and colleagues reflect a priori differences between the three groups that were not corrected for by pretreatment data. This possibility seems somewhat less likely when considering that the increasing effect of caffeine was dose-dependent in a fairly linear way. A second potential factor concerns the amount of practice, which was low in the Foreman and colleagues study, but relatively high in the Hasenfratz and Bättig study (four sessions, each including a pretreatment part); there are suggestions that caffeine is more beneficial in highly practiced tasks (13). A third possibility concerns the effective dosage, as the average weight of the male subjects in the Foreman and colleagues study must have been considerably higher than that of the female subjects in the Hasenfratz and Bättig study.

Uncovering the nature of the effect of caffeine on Stroop interference is of particular interest, as its effects on other measures of selectivity are relatively weak. With respect to selectivity for specific locations in the visual field, as well as to distractibility by information at irrelevant locations, Kenemans and Verbaten (11) found no effect of caffeine, even though performance in general was improved. On the other hand, studies using even-related brain potentials have revealed stronger selective processing under caffeine (10,15), but the functional interpretation of these results is far from clear. In contrast, Stroop interference may be viewed as an index of selection between competing response tendencies. As such, it clearly differs from situations in which selectivity concerns competing stimulus inputs originating from different locations in the visual field (as described above); note that in the traditional Stroop setup all stimulus information is spatially integrated.

In the present study the effects of 250 mg caffeine on Stroop interference were investigated following the experimental setup as described by Hasenfratz and Bättig (6). Specifically, a crossover design was used in which female subjects attended four experimental sessions—two with caffeine and two with placebo. In addition, the possible dependence of posttreatment caffeine effects on short-term practice, i.e., experiencing a pretreatment test session, was explicitly addressed. The dosage of 250 mg is high in light of the typical habitual comsumption pattern, but was nevertheless chosen so as to maintain consistency with the reference studies (5,6).

In this first experiment the particular variety of the Stroop task used by Foreman and colleagues and by Hasenfratz and Bättig was employed. In this version subjects have to judge the numerosity of homogenous character strings presented at a high rate, the characters being either neutral or digits incongruent with numerosity. As will be decribed, in the present study the interference effects were rather small, which may have limited the probability of further reductions under caffeine. Therefore, a second experiment was designed using a variety of the classic color-word Stroop task, for which higher interference scores were expected. The second experiment also contained a condition in which neutral and incongruent stimuli were randomly mixed within one sequence, as opposed to the traditional method of blocked presentations of the two conditions. It is possible that subjects' strategies differ between mixed and blocked modes (e.g., more conservative responding in incongruent blocks), and that caffeine effects differ as a function of strategy.

EXPERIMENT 1

Method

Subjects and treatment manipulation. Sixteen reportedly healthy female graduate students served as subjects. Mean age was 21.8 (SD = 2.8), and mean reported daily coffee intake was 3.4 cups (SD = 1.5). Subjects had normal or corrected-to-normal vision. They were paid for their participation. Prior to the experiment they signed written consents that they had been informed about treatment and task conditions and the approximate everyday equivalent of the caffeine dossage to be used. As such, the complete procedure was in compliance with the Declaration of Helsinki on human subjects.

Each subject participated in four sessions, separated by 1 week. Half of the subjects first attended two sessions, including a pretreatment condition, and thereafter two sessions without; for the other half, this order was reversed. Within each half, four treatment orders were evenly distributed across the subjects: Placebo-Caffeine-Placebo-Caffeine, Placebo-Caffeine-Caffeine-Placebo, Caffeine-Placebo-Caffeine-Placebo, and Caffeine-Placebo-Placebo-Caffeine.

Subjects were asked to abstain from substances containing caffeine for at least the 12 h preceding each experimental session. Before each session each subject filled in a questionaire to assess her compliance with the request to abstain; it was emphasized that failure to comply would result in exclusion from the experiment, but that she would still receive a proper financial reward. In two sessions subjects were given 250 mg lactose and in the other two 250 mg caffeine. Each dose was dissolved and administrated in a cup of decaffeinated coffee. Milk powder and sugar were added to the coffee to suit the taste of the subject. In one of the placebo and one of the caffeine sessions, substance administration was preceded by a complete pretreatment session (see Stimuli and Procedure). Forty-five minutes after lactose or caffeine administration subjects (again) performed all tasks.

Stimuli and procedure. The 16 neutral stimuli were strings consisting of either 1, 2, 3, or 4 times a '▲' or a '■,' or a '●' or an "×." These stimuli were randomly mixed in two functionally equivalent neutral-condition blocks of 64 trials each. The 12 incongruent stimuli consisted of either 1, 2, 3, or 4 times a "1," or a "2," or a "3," or a "4," allowing only for incongruent number-digit combinations (four numbers × three noncorresponding digits = 12 stimuli). The incongruent stimuli were also randomly mixed in two functionally equivalent incongruent-condition blocks of 64 trials each.

Stimuli were presented in the center of the visual field on a NEC Multisync monitor (black on white). The strings were 0.3 to 1.9 degrees wide and 0.6 to 0.9 degrees high, depending on the characters used. Stimulus presentation was terminated by the subject's button press, with a maximum duration of 1 s. The interval between two successive stimuli was 80 ms. During the task subjects were seated in a dentist's chair and had to fixate on a cross that was presented in the center of the visual field.

Subjects were instructed to respond according to the number of elements in the string by pressing a button with the middle or index finger of the left or right hand (one element:

left middle; two elements: left index; three: right index; four: right middle). The order of neutral and incongruent blocks was counterbalanced across subjects and treatment orders.

Data analysis. Performance was analyzed in terms of speed and accuracy. Reaction time (RT) was measured as the time between stimulus onset and the button press. Mean RT was based on correct button presses between 350 ms poststimulus and stimulus ending. This rather high lower limit was maintained to exclude delayed responses to preceding stimuli. For each condition in each block outliers (RTs outside a ±2.0 SDs window around the mean) were removed from the data set. Error and omission proportions were also computed per block and per condition. Error proportion was defined as the ratio of the number of incorrect responses vs. the number of incorrect + the number of correct responses. Omission proportions were computed as the ratio of the number of trials without responses vs. the total amount of trials.

Subsequently, RTs, error rates, and omission rates were pooled across target stimuli (fingers) and blocks for each condition of interest, and entered into ANOVAs. A first series of ANOVAs concerned the performance in posttreatment blocks that where preceded by pretreatment blocks. Within-subject factors were treatment (placebo vs. caffeine), time (before vs. after administration), and Stroop condition (neutral vs. incongruent). A between-factor order was also included in the design, subjects being categorized as to whether the two sessions including pretreatment tasks were the first two or the last two of the total four, and as to whether they were administered placebo in the first of these two sessions, or caffeine. This resulted in four order groups, each containing four subjects.

In a second series of ANOVAs the effect of performing in a pretreatment session on posttreatment performance was addressed. Within-subject factors were treatment (placebo vs. caffeine), pretreatment (pretreatment task included or not), and Stroop condition (neutral vs. incongruent); note that the dependent variables concerned posttreatment performance only. As to the between-subject factor order, subjects were categorized as to whether they first attended two sessions including a pretreatment condition and thereafter two sessions without, or the reverse, and according to the order of pharmacological treatment: Placebo-Caffeine-Placebo-Caffeine, Placebo-Caffeine-Caffeine-Placebo, Caffeine-Placebo-Caffeine-Placebo, or Caffeine-Placebo-Placebo-Caffeine (see above, Subjects and Treatment section). This resulted in eight order groups, each containing two subjects.

The order factor was included solely to reduce its contribution to the error terms for the various *F*-tests. However, if a given effect of interest (e.g., treatment) does not significantly depend on order condition, then including the order factor in the model will not increase statistical power but may actually decrease it because of a loss of degrees of freedom (*df*s) for the effect of interest. Therefore, if a given effect did not depend significantly on order (i.e., at $p < 0.25$), the order factor was removed from the model, so as to increase the dfs for the effect of interest. In the first series of ANOVAs (four order groups) these *df*s amounted to (1, 12) with and (1, 15) without order included; in the second series (eight order groups) they amounted to (1, 8) and (1, 15), respectively.

Results

Pre-/postdesign: Reaction times. The left panel of Fig. 1 shows mean RTs as a function of treatment and task condition, before and after administration. RTs were in general shorter posttreatment than pretreatment [time main effect, $F(1, 15) = 38.6$, $p < 0.001$], and on neutral than on incongruent trials [Stroop main effect, $F(1, 12) = 23.0$, $p < 0.001$]. There were no significant effects of treatment, $F(1, 15) = 1.4$, $p > 0.25$, for the treatment × time effect, and $F(1, 15) = 0.9$, for the treatment × time × Stroop effect. Although the left panel of Fig. 1 might suggest otherwise, the relatively short mean RTs in the caffeine/posttreatment condition reflected the contribution of a few subjects with extreme values; in all, only 8 out of the 16 subjects had a larger positive difference between pre- and posttreatment RTs under caffeine than under placebo.

Pre-/postdesign: Errors and omissions. The right panel of Fig. 1 shows mean error proportions as a function of treatment and task condition, before and after administration. Stroop interference was reduced or even reversed after caffeine administration, relative to placebo [treatment × time × Stroop, $F(1, 15) = 5.1$, $p < 0.05$]. This was further confirmed in that a significant time × Stroop interaction was found for the caffeine condition, $F(1, 15) = 6.8$, $p < 0.05$, but not for placebo. Furthermore, under caffeine the time effect was significant for the incongruent Stroop condition, $F(1, 15) = 17.5$, $p < 0.001$, but not for the neutral one. Finally, the posttreatment Stroop effect was significant under placebo, $F(1, 15) = 5.6$, $p < 0.05$, not under caffeine, $F(1, 15) = 2.7$, $p > 0.1$.

Omission rates. Mean omission rates were smaller than 0.04 in each condition. There were more omissions pre- than posttreatment [time main effect, $F(1, 12) = 13.8$, $p < 0.005$] and in incongruent, relative to neutral conditions [Stroop main effect, $F(1, 12) = 7.6$, $p < 0.05$], without there being any other effects. Thus, the pattern of significant effects equalled the one found for RT, suggesting that very slow responses made a major contribution to the estimated omission rates.

Effects of the pretreatment session on posttreatment performance: RT. The left panel of Fig. 2 shows mean posttreatment RTs as a function of treatment, task condition, and of whether these posttreatment sessions were preceded by pretreatment sessions. RTs were in general shorter after caffeine than after placebo [treatment main effect, $F(1, 8) = 12.4$, $p < 0.01$], with pretreatment than without [pretreatment main effect, $F(1, 8) = 11.5$, $p < 0.01$], and on neutral than on incongruent trials [Stroop main effect, $F(1, 8) = 23.0$, $p < 0.001$]. There were no significant interaction effects, $F(1, 8) = 0.03$ and 0.21 for the treatment × pretreatment and the treatment × pretreatment × Stroop effect, respectively.

These results present a seeming contradiction in that treatment effects were significant, whereas they were not after correction for pretreatment values (see section on pre-/postdesign: Reaction Times). Looking only at the posttreatment conditions that were preceded by pretreatment, conditions also revealed a main effect of treatment ($p < 0.004$). As apparent from the analysis of the pre-/postdesign, this main effect disappears after correction for the pretreatment values.

Effects of the pretreatment session on posttreatment performance: Errors and omissions. The right panel of Fig. 2 shows mean posttreatment error proportions as a function of treatment and task condition, and of whether these posttreatment sessions were preceded by pretreatment sessions. Stroop interference was reduced or even reversed after caffeine administration, relative to placebo [treatment × Stroop, $F(1, 8) = 16.9$, $p < 0.005$]. This was further confirmed in that the Stroop effect was significant for the caffeine condition, $F(1, 8) = 6.6$, $p < 0.05$, but not for placebo. The effect of treatment also depended on whether the tasks where performed pretreatment as well: there were more errors under caffeine when there was no pretreatment session [pretreatment × treatment effect,

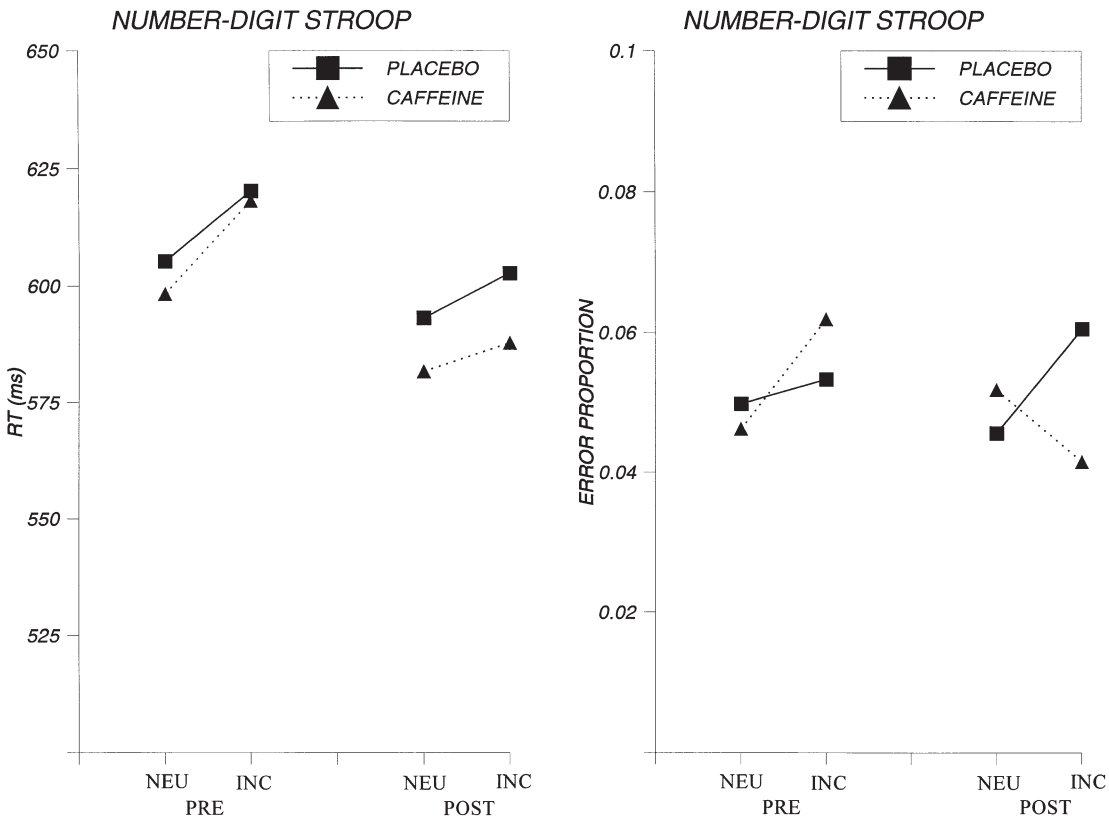

FIG. 1. Mean reaction times (RT, left panel) and error proportions (right panel) pre- and posttreatment (placebo or caffeine), in neutral (NEU) and incongruent (INC) number-digit Stroop conditions.

$F(1, 8) = 13.9$, $p < 0.01$; treatment effect without pretreatment session, $F(1, 8) = 8.7$, $p < 0.05$; NS with pretreatment session]. There was no other significant effect, in particular, no three-way interdependence of treatment, Stroop, and pretreatment effects, $F(1, 15) = 2.7$, $p = 0.125$.

Omission rates. Mean omission rates were smaller than 0.04 in each condition. There were fewer omissions under caffeine than under placebo [treatment main effect, $F(1, 8) = 5.4$, $p < 0.05$], with pretreatment session than without [pretreatment main effect, $F(1, 8) = 5.4$, $p < 0.05$] and in neutral, relative to incongruent conditions [Stroop main effect, $F(1, 15) = 9.1$, $p < 0.05$], without there being any other effects. Again, the pattern of significant effects equalled the one found for RT, suggesting that very slow responses made a major contribution to the estimated omission rates.

EXPERIMENT 2

In Experiment 1 Stroop effects on RT were rather small: hardly 20 ms in the present study, as opposed to more than 50 ms in the one by Hasenfratz and Bättig (6). This relatively small effect may have limited the possible extent of caffeine-induced reduction. Therefore, in the second experiment, another variety of the Stroop task was used, for which pilot data indicated that the RT interference effect was considerably larger. This version resembled much more the common modern laboratory procedure, using color-word incongruence and a fixed rate of stimulus presentation.

Method

Subjects and treatment manipulation. Sixteen reportedly healthy female graduate students were recruited as subjects. Mean age was 22.1 (SD = 2.1), and mean reported daily coffe intake was 4.0 cups (SD = 3.2). Subjects had normal or corrected-to-normal vision. They were paid for their participation. Informed consents and compliance with the Declaration of Helsinki were applied as in Experiment 1.

Each subject participated in two sessions, separated by 1 week. Subjects were asked to abstain from substances containing caffeine for at least the 12 h preceding each experimental session. Before each session each subject filled in a questionaire to assess her compliance with the request to abstain; it was emphasized that failure to comply would result in exclusion from the experiment, but that she would still receive a proper financial reward. Both sessions started with a complete series of practice and experimental blocks (see Stimuli and Procedure). Subsequently, in one session subjects were given 250 mg lactose and in the other 250 mg caffeine. Each dose was dissolved and administrated in a cup of decaffeinated coffee. Milk powder and sugar were added to the coffee to suit the taste of the subject. Forty-five minutes after lactose or caffeine administration subjects again performed all tasks. Assignments to treatment conditions were double blind and fixed in advance, the order of treatment conditions being balanced completely across the 16 subjects.

Stimuli and procedure. The 12 neutral stimuli consisted of the strings "????," "****," "00000," and "xxxxx," colored red, blue, yellow, or green (three possible colors for each string).

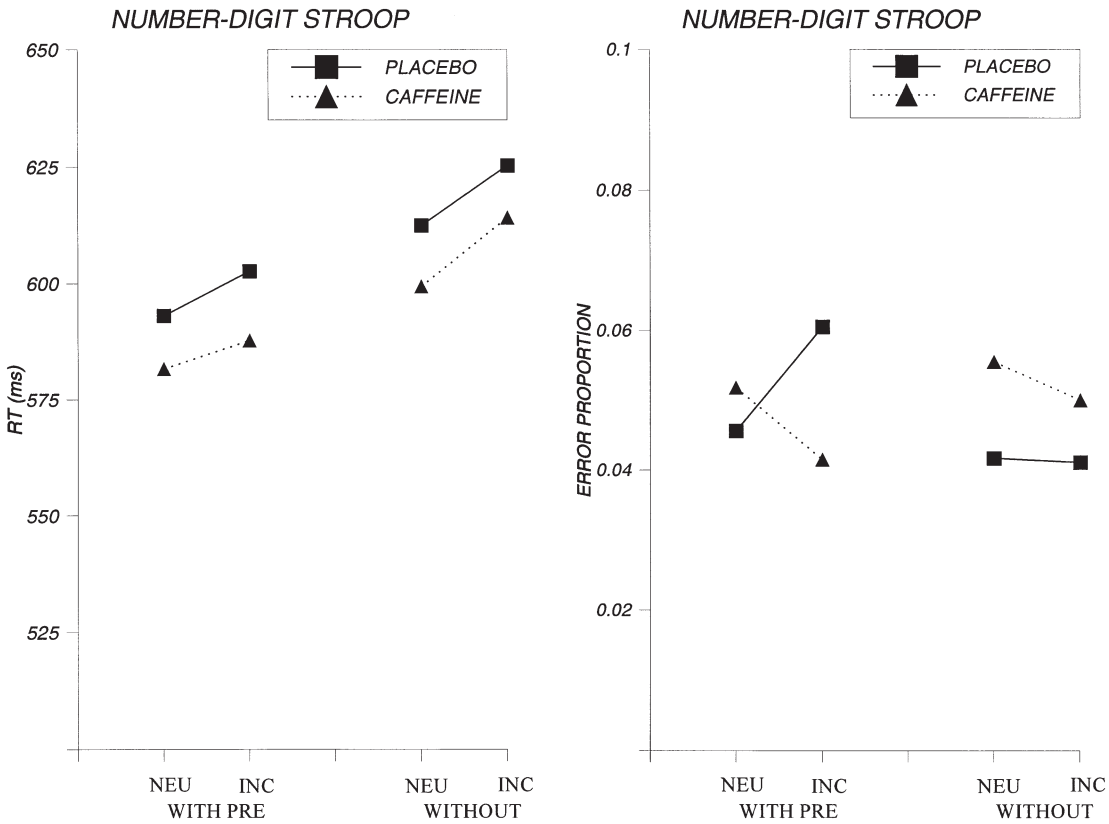

FIG. 2. Mean posttreatment reaction times (RT, left panel) and error proportions (right panel) under placebo and under caffeine, in neutral (NEU) and incongruent (INC) number-digit Stroop conditions, as a function of prior exposure to a pretreatment test.

These stimuli were randomly mixed in two functionally equivalent neutral-condition blocks of 60 trials each. The 12 incongruent stimuli consisted of the words "rood" (red), "blauw" (blue), "geel" (yellow), and "groen" (green). They were printed in these same colors, allowing only for incongruent color-word combinations (four words × three noncorresponding colors = 12 stimuli). The incongruent stimuli were also randomly mixed in two functionally equivalent incongruent-condition blocks of 60 trials each. In addition, the 12 neutral and the 12 incongruent stimuli were randomly mixed in two mixed-condition blocks of 120 trials each.

Stimuli were presented in the center of the visual field on a NEC Multisync monitor (black on white). The strings were 2.6 degrees (four characters) or 3.3 (five characters) wide and 1.7 degrees high. Stimulus duration was 1300 ms, and stimulus onset asynchronies amounted to 1950 ms. During the task subjects were seated in a dentist's chair and had to fixate on a cross that was presented in the centre of the visual field.

Subjects were instructed to respond according to the color in which a string was printed by pressing a button with the middle or index finger of the left or right hand (red: left middle; yellow: left index; green: right index; blue: right middle). Before each series of six experimental blocks (two neutral, two mixed, two incongruent) subjects were trained with three blocks containing 24 congruent color-word combinations only. During these blocks the response was determined by the word, which was, however, congruent with the color. The aim of this task was to strengthen the association between specific color names and specific button presses, which is not a natural one.

The order of neutral, mixed, and incongruent blocks was counterbalanced across subjects and treatment orders.

Data analysis. The initial part of the data analysis, up to the statistical analysis, was identical to that in Experiment 1, with the exception that now mean RT was based on correct button presses between 150 and 1600 ms poststimulus. A first series of ANOVAs evaluated the blocked conditions only, as these were the subject of specific hypotheses, based on the literature and the results of Experiment 1. In a second series of ANOVAs the factor sequence (blocked vs. mixed) was included. The other factors were treatment (placebo vs. caffeine), time (before treatment vs. after administration), and Stroop condition (neutral vs. incongruent), as well as the between-subject factor order of treatment conditions [Placebo first ($n = 8$) or caffeine first ($n = 8$)]. The order factor was included solely to reduce its contribution to the error terms for the various *F*-tests. As in Experiment 1, when the order effect did not interact with the effect of interest, it was removed from the design (resulting in *df*s of 1, 15, instead of 1, 14).

Results

Unless stated otherwise, *F*-values have *df*s of 1 and 14.

Reaction times. Figure 3 shows mean RTs as a function of treatment and task condition, before and after administration. In blocked conditions RTs were shorter after caffeine administration than after placebo administration (treatment × time, $F = 5.7$, $p < 0.05$). RTs were also in general shorter posttreatment than pretreatment (time main effect, $F = 34.8$, $p <$

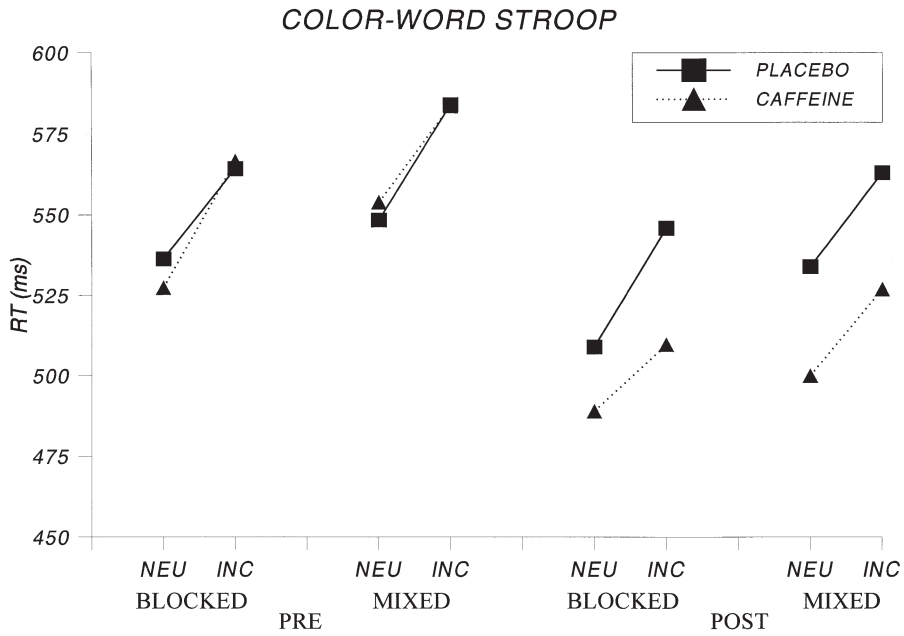

FIG. 3. Mean reaction times (RT) pre- and posttreatment (placebo or caffeine), with color-word Stroop conditions (NEU = neutral, INC = incongruent) either in separate (blocked) or in mixed sequences.

0.001), and on neutral than on incongruent trials (Stroop main effect, $F = 18.1$, $p < 0.001$). In addition, Stroop interference postadministration was smaller under caffeine, relative to placebo [treatment × time × Stroop, $F(1, 15) = 5.1$, $p < 0.05$]. In Fig. 3 (blocked conditions), this is visible as a smaller difference between neutral and incongruent conditions for posttreatment caffeine, relative to posttreatment placebo.

Including the mixed conditions confirmed most of these findings, as is also apparent from Fig. 3 (treatment × time: $F = 8.6$, $p < 0.05$; time: $F = 28.8$, $p < 0.001$; Stroop: $F = 23.4$, $p < 0.001$). A major difference, however, concerns the reduction of Stroop interference under caffeine, which was significant in the blocked conditions, but not that clear in the mixed ones. Consistently, testing the treatment × time × Stroop interaction for mixed conditions only did not yield significance ($F = 0.1$), nor did the same test for mixed and blocked condition pooled ($F = 1.7$, $p > 0.2$). The explicit test for the dependence of the caffeine-induced interference reduction on blocking the Stroop conditions revealed marginal significance [sequence × treatment × time Stroop, $F(1, 15) = 4.1$, $p < 0.062$].

Finally, it may be noted that RTs were in general shorter in blocked than in mixed conditions (sequence main effect, $p < 0.001$). This effect was not modulated by caffeine and did not depend on Stroop condition.

Errors and omissions. Figure 4 shows mean error proportions as a function of treatment and task condition, before and after administration. The analysis over blocked conditions only and the overall one (including mixed conditions as well) yielded essentially the same results. The overall analysis confirmed that error rates were smaller after caffeine administration than after placebo administration [treatment × time, $F(1, 15) = 6.0$, $p < 0.05$; for blocked conditions only $F = 11.1$, $p < 0.005$]. Error rates were also in general higher posttreatment than pretreatment [time main effect, $F = 16.2$, $p < 0.001$; blocked only $F(1, 15) = 14.3$, $p < 0.005$], and lower in blocked than in mixed conditions [sequence main effect, $F(1, 15) = 8.9$, $p < 0.01$].

Stroop interference (the difference between neutral and incongruent) was smaller in the caffeine conditions, but this difference was visible both pre- and posttreatment [treatment × Stroop, $F(1, 15) = 7.5$, $p < 0.05$; blocked only, $F(1, 15) = 7.3$, $p < 0.05$; Treatment × time × Stroop, $F = 0.04$ (overall) and 0.02 (blocked only)]. Thus, the treatment-induced reduction in interference at the level of errors does not reflect a direct effect of caffeine administration. As to what it does reflect, there is probably only one obvious possibility: asymmetric transfer from one treatment session to the other. Further analysis revealed that this was indeed the case. Consider the fact that the order × treatment design as presently used can also be described as an order × session (first vs. second session) design. Note further that the treatment main effect in the first design is perfectly equivalent to the order × session interaction effect in the second design, and that significance of the former and latter are statistically one and the same thing. Figure 5 depicts the order × session interaction. As can be seen, Stroop interference diminishes from a first placebo to a second caffeine session, but increases from caffeine first to placebo second.

Omission rates. Mean omission rates were smaller than 0.01 in each condition. They were lower after caffeine adminstration than after placebo [treatment × time, $F(1, 15) = 6.5$, $p < 0.05$]. No other effect was observed, except for a significant interaction time × sequence × Stroop, $F(1, 15) = 4.7$, $p < 0.05$, for which further testing did not result in an unambiguous interpretation.

GENERAL DISCUSSION

Two experiments on caffeine effects on Stroop interference were described. In the first one, subjects had to judge the

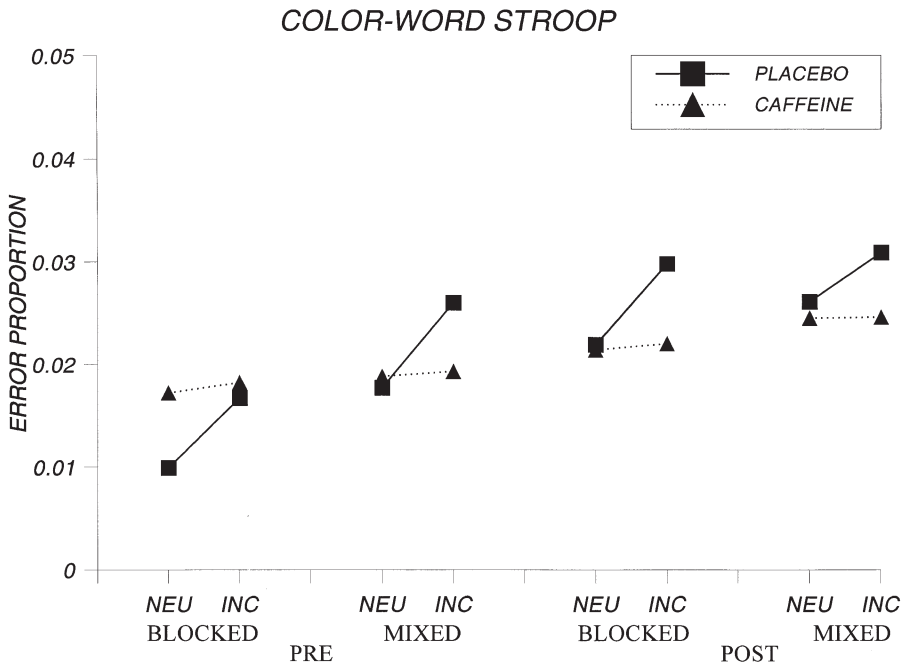

FIG. 4. Mean error proportions pre- and posttreatment (placebo or caffeine), with color-word Stroop conditions (NEU = neutral, INC = incongruent) either in separate (blocked) or in mixed sequences.

numerosity of digit strings consisting of incongruent digits. Relative to a neutral condition, interference was reliably observed with respect to speed of responding (RT). However, there were no consistent effects of caffeine administration on either RT in general or interference as indexed by differential RT. In contrast, error rates were highly dependent on treatment condition. A significant reduction in error rate was found after caffeine administration, specifically in the incongruent condition, not after placebo or in the neutral condition (see Fig. 1). These results suggest that caffeine was beneficial to suppression of the task-irrelevant response tendency, but that this benefit was realized as a reduction in error rate, rather than as increased speed.

A major aim of the first experiment concerned the effects of experiencing a pretreatment session on posttreatment performance. Posttreatment RTs were shorter with than without preceding pretreatment sessions, suggesting that some practice effect did indeed occur (see Fig. 2). Furthermore, the general effect of caffeine on error rates depended on the presence of a pretreatment session: without preceding pretreatment sessions caffeine increased error rates, whereas it did not with preceding pretreatment sessions. However, the effects of caffeine on Stroop interference were not significantly modulated by the experience of a preceding pretreatment session. In all, these results do not support the notion that the use of pretreatment sessions would be of influence on the effects of caffeine on Stroop intereference.

The finding of reduced Stroop interference under caffeine is consistent with the results of Hasenfratz and Bättig (6). However, the latter authors found this reduction at the level of RT (they did not report data on error rates), whereas in the present study it was observed for error rates and not for RT. One factor that may have contributed to this inconsistency concerns the size of the interference effect in terms of RT, as observed under placebo: more than 50 ms in the Hasenfratz and Bättig study, as opposed to hardly 20 ms in the present one. This relatively small effect may have limited the possible extent of caffeine-induced reduction. Therefore, in a second experiment, another variety of the Stroop task was used, involving color-word incongruence and a a fixed rate of stimulus presentation.

In the second experiment the RT interference scores were indeed about twice as large as in the first experiment. Caffeine resulted in a general shortening of RT. Furthermore, when neutral and incongruent trials were presented in separate blocks of trials, caffeine also specifically reduced the interference effect (see Fig. 3). At the same time error rates were also generally smaller, while there was no reduction of the error-rate interference score specifically after caffeine administation (Fig. 4). The combined results of the two experiments may be taken to suggest that caffeine specifically enhances the selectivity for task-appropriate response tendencies, and that this increase in response selectivity is manifest in reduced interference in terms of either speed or accuracy, depending on specific task conditions.

However, at first sight this general conclusion is difficult to reconcile with the results from the mixed conditions in the second experiment. Traditionally [e.g., (26)], and in particular in previous studies on caffeine and Stroop interference, neutral and incongruent conditions are separated in different sequences of trials. This has the disadvantage that a difference in performance between the two conditions may not only reflect the amount of response selectivity, but also the effect of strategic factors. For example, subjects may adopt a much more conservative strategy for incongruent blocks, resulting in relatively long RTs but only small or even reversed differences in error rates. For any given pharmacological manipulation then, the potential effect on response selectivity and that on strategy will be hard to separate. When incongruent and

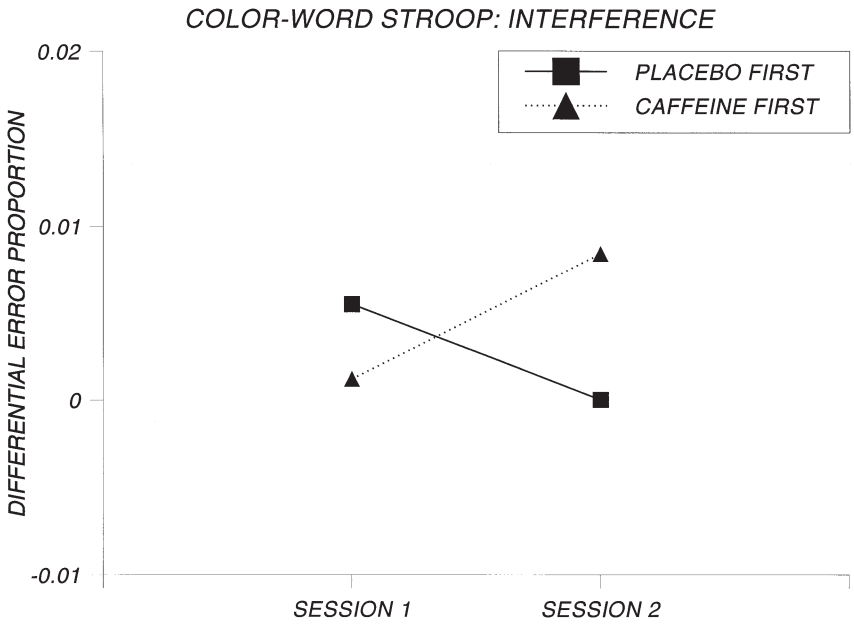

FIG. 5. Mean differences in error proportion between color-word incongruent and neutral conditions, as a function of treatment order and session number. Pre- and posttreatment values, as well as mixed-condition and block-condition data, are pooled.

neutral trials are mixed randomly in one sequence, strategy effects are far less probable, because subjects can no longer predict when either condition will be presented. However, as it turns out, there are other differences between blocked and mixed conditions as well.

In the mixed condition of the second experiment the reducing effect of caffeine on interference was no longer seen. If the caffeine-induced reduction of the RT interference score in blocked conditions rested on an effect on speed–accuracy tradeoff, then cancelling the contribution of the latter effect (in the mixed conditions) should also cause a change in the effect of caffeine on the error-rate interference score. However, error rates were actually lower in blocked than in mixed conditions, and they were not influenced ($p < 0.087$) by Stroop condition. Furthermore, caffeine reduced the error rate in general, independent of the Stroop condition and of whether these conditions were mixed or blocked. Thus, there is no indication that the caffeine-induced reduction on interference in blocked conditions is mediated by an effect on strategy.

What, then, causes the differences in caffeine effects between blocked and mixed conditions? One relevant consideration is that both speed and accuracy were lower in the mixed conditions. This suggests a difference not so much in strategies, but rather in general task difficulty. It could be that an active inihibition process, directed at the irrelevant-word representations (16), is invoked strongly in blocked incongruent conditions, far less in blocked neutral ones, and to an intermediate extent for both incongruent and neutral mixed conditions. As in the latter case the inhibition is optimal for neither condition, a general decrease in performance, relative to blocked conditions, is observed. It is possible that caffeine reduces the active-inhibition response to mixed neutral stimuli, in addition to enhancing it in response to incongruent stimuli; thus, there would be an additional benefit of caffeine in the mixed-neutral condition, which would weaken the specifity of the benefit in the incongruent condition.

The conclusion that caffeine did not reduce error-rate interference scores in the second experiment was drawn despite the clear-cut reduction that was, in fact, visible in the posttreatment data (Fig. 4). As noted, this reduction was not specific to caffeine administration, as it was equally visible before treatment. This pattern of results may be expected when the transfer (e.g., of learning experience) between placebo and caffeine sessions is asymmetric. This was illustrated in Fig. 5 by plotting the error-rate interference scores as a function of session number and order group. When subjects received placebo in the first session interference was less in the second session for both pre- and posttreament performance. When subjects received caffeine in the first session, interference scores became higher in the second session, again for both pre- and posttreament performance. Thus, the experience of interference after placebo transfers more beneficially to performance of the same task 1 week later, compared with the same experience after caffeine. Whatever the mechanism underlying this asymmetry, it has important methodological implications. First, asymmetrical transfer may result in clear posttreatment drug effects that are equally present at pretreatment. Second, such posttreatment effects may lead to incorrect conclusions if the pretreatment effects are not taken into account (e.g., because they are not measured).

It could be argued that any failure to find a caffeine effect on whatever parameter might be due to a particular aspect of our procedure, viz., the opportunity to subjects for ad lib addition of sugar to the vehicle, which could mask the effects of caffeine. This would imply that caffeine effects are dependent on sugar intake in an interactive way, because it may be assumed that in the present procedure sugar intake was constant across treatment conditions for each subject. Such dependence seems unlikely, because the effects of caffeine on a number of parameters were, in fact, significant as expected, both in the present study and in previous studies that used the same administration procedure (10,11,14,15). The same argu-

ment applies to the possibility that the lack of standardization arising from ad lib addition of milk powder and sugar would reduce the power for detecting caffeine effects. In the studies in which this procedure was maintained, the global effects of caffeine were very much in line with those obtained with other procedures (see, e.g., the discussion of caffeine effects in vigilance and related tasks, in the section on withdrawal effects below).

To conclude, it may be suggested that caffeine enhances response selectivity by enhancing active inhibition of task-irrelevant response tendencies. If conditions in which active inhibition is beneficial alternate unpredictably with ones in which it merely induces costs, then the caffeine effect may be masked. The effects of caffeine on response-selection processes may contribute to its beneficial effects in conditions of rapid and sustained information processing (1,12,24,27), perhaps adding to its beneficial effects on perceptual sensitivity (10,14). The presently found response-selection effects may be contrasted with the lack of caffeine effects on other kinds of response selection, as reported by Kenemans and Verbaten (11). The latter concerned competing response tendencies activated by information at different locations in the visual field, or irrelevant tendencies activated because of their spatial compatibility with the visual location of the stimulus. In contrast, response selection in the present study concerned spatially integrated sources of information while spatial compatibility was not at stake. However, there are other differences as well between the two studies. One of them concerned the amount of response alternatives (four here, two in the Kenemans andVerbaten study), a factor bound to influence response selection and perhaps interacting with caffeine. Another is the predictability of task conditions, which was virtually zero in the Kenemans and Verbaten study; as argued above, such lack of predictability may mask caffeine effects. Future research has to reveal more precisiely the effects of specific task variables, and their interaction with those of caffeine.

The Contribution of Withdrawal Effects

In both experiments the common procedure to have subjects abstain from caffeine for 12 h preceding the sessions was maintained. It is possible then, that differences between caffeine and placebo conditions merely reflect withdrawal effects in the latter condition, which are canceled by caffeine administration in the former. The results of several studies are relevant to this issue. First, some studies have used no or very short (1–3 h) abstinence periods and still found beneficial effects of caffeine on performance [e.g., (25,27)]. Although it might by argued that in such conditions some withdrawal effect may still be present (8), these abstinence periods are much shorter than the generally estimated period necessary for caffeine-withdrawal symptoms to develop (19). Second, several studies [reviewed in (23)] indicate a significant increase in subjectively reported negative symptoms (e.g., tiredness, drowsiness, headache) after about 12 h abstinence. At the same time it turns out that parallel effects of such abstinence on objective performance measures are hard to establish (23).

Third, in some studies low and high habitual consumers have been compared after placebo and after caffeine administration. If withdrawal effects are indeed larger in the latter group, the assumption that caffeine benefits to performance are due to cancellation of withdrawal effects leads to the prediction that the caffeine benefit should be larger for high than for low habitual consumers. Mitchell and Redman (18) could not confirm this prediction in any of a variety of psychomotor and cognitive tasks. Such a confirmation was neither found by Fine and colleagues (4) in a prototype vigilance task, even though, after placebo, high habitual consumers did perform worse after 10 h deprivation than low habitual consumers did. The degraded performance after abstinence in high habitual consumers could not be explained by a priori differences between the two groups. This indicates that abstinence may indeed affect performance; at the same time, caffeine benefits were equal for both groups, indicating that they do not merely reflect cancellation of withdrawal effects.

The caffeine benefits of Fine et al. were comparable to those observed by Kenemans and Lorist (10), as well as those in many of the studies reviewed by Koelega (12), in that they concerned reaction times and hit rates, but not false-alarm rates. In the Kenemans and Lorist study a comparison was made between 12 h-deprived subjects receiving caffeine, 12 h-deprived subjects receiving placebo, and nondeprived subjects receiving nothing. The difference between the deprivation/ placebo and the nondeprivation groups was assumed to indicate the extent of withdrawal effects, and the difference between the deprivation/ caffeine and the nondeprivation groups the true, "net" benefit of caffeine. On this assumption it was concluded that withdrawal and net effects were both present, that they were of comparable size for hit rates, and that the net effect was considerably larger than the withdrawal effect for reaction time.

In all, these findings suggest that deprivation periods of at least 10 h may impair performance, but that even in such conditions observed caffeine effects cannot be attributed solely to cancellation of withdrawal effects. It seems reasonable to generalize this conclusion to the present findings: Caffeine augments task performance, both in a general way, and specifically with respect to selection between competing response tendencies.

REFERENCES

1. Anderson, K. J.; Revelle, W.; Lynch, M. J.: Caffeine, impulsivity, and memory scanning: A comparison of two explanations for the Yerkes-Dodson effect. Motivat. Emot. 13:1–20; 1989.
2. Cohen, J. D.; Dunbar, K.; McClelland, J. L.: On the control of automatic processes: A parallel distributed processing account of the Stroop effect. Psychol. Rev. 97:332–361; 1990.
3. Easterbrook, J. A.: The effect of emotion on cue utilization and organization of behavior. Psychol. Rev. 66:183–201; 1959.
4. Fine, B. J.; Kobrick, J. L.; Lieberman, H. R.; Marlowe, B.; Riley, R. H.; Tharion, W. J.: Effects of caffeine or diphenhydramine on visual vigilance. Psychopharmacology (Berlin) 114:233–238; 1994.
5. Foreman, N.; Barraclough, S.; Moore, C.; Mehta, A.; Madon, M.: High doses of caffeine impair performance of a numerical version of the Stroop task in men. Pharmacol. Biochem. Behav. 32:399–403; 1989.
6. Hasenfratz M.; Battig, K.: Action profiles of smoking and caffeine: Stroop effect, EEG, and peripheral physiology. Pharmacol. Biochem. Behav. 42:155–161; 1992.
7. Hockey, G. R. J.: Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. In: Boff, K. R.; Kaufman, L.; Thomas, J. P., eds. Handbook of perception and human performance, vol. 2. New York: Wiley; 1984:44/1–49.
8. James, J. E.: Understanding caffeine: A biobehavioural analysis. Thousand Oaks, CA: Sage; 1997.

9. Jensen, A. R.; Rohwer, W. D.: The Stroop color word task: A review. Acta Psychol. 25:36–93; 1966.
10. Kenemans, J. L.; Lorist, M. M.: Caffeine and selective visual processing. Pharmacol. Biochem. Behav. 52:461–471; 1995.
11. Kenemans, J. L.; Verbaten, M. N.: Caffeine and visuo-spatial attenttion. Psychopharmacology 135: 353–360; 1998.
12. Koelega, H. S.: Stimulant drugs and vigilance performance: A review. Psychopharmacology (Berlin) 111:1–16; 1993.
13. Loke, W. H.: The effects of caffeine and automaticity on a visual information processing task. Hum. Psychopharmacol. 7:379–388; 1992.
14. Lorist, M. M.; Snel, J.; Kok, A.: Influence of caffeine on information processing stages in well rested and fatigued subjects. Psychopharmacology (Berlin) 113:411–421; 1994.
15. Lorist, M. M.; Snel, J.; Kok, A.; Mulder, G.: The influence of caffeine on selective attention in well rested and fatigued subjects. Psychophysiology 31:525–534; 1994.
16. MacLeod, C. M.: Half a century of research on the Stroop effect: An integrative review. Psychol. Bull. 109:163–203; 1991.
17. Mair, R. G.; McEntee, W. J.: Cognitive enhancement in Korsakoff's psychosis by clonidine: A comparison with L-dopa and ephedrine. Psychopharmacology (Berlin) 88:374–380; 1986.
18. Mitchell, P. J.; Redman, J. R.: Effects of caffeine, time of day and user history on study-related performance. Psychopharmacology (Berlin) 109:121–126; 1992.
19. Nehlig, A.; Daval, J. L.; Debry, G.: Caffeine and the central nervous system: Mechanisms of action, biochemical, metabolic, and psychostimulant effects. Brain Res. Rev. 17:139–169; 1992.
20. Parrott, A. C.; Wesnes, K.: Promethazine, scopolamine and cinnarizine: Comparative time course of psychological performance effects. Psychopharmacology (Berlin) 92:513–519; 1987.
21. Phaf, R. H.; Van der Heijden, A. H. C.; Hudson, P. T. W.: SLAM: A connectionist model for attention in visual selection tasks. Cogn. Psychol. 22:273–341; 1990.
22. Provost, S. C.; Woodward, R.: Effects of nicotine gum on repeated administration of the Stroop test. Psychopharmacology (Berlin) 104:536–540; 1991.
23. Rogers, P. J.; Dernoncourt, C.: Regular caffeine consumption: A balance of adverse and beneficial effects for mood and psychomotor performance. Pharmacol. Biochem. Behav. 59:1039–1045; 1998.
24. Smith A. P.; Brockman, P.; Flynn, R.; Maben, A.; Thomas, M.: Investigation of the effects of coffee on alertness and performance during the day and night. Neuropsychobiology 27:217–223; 1993.
25. Smith, A. P.; Maben, A.; Brockman, P.: Effects of evening meals and caffeine on cognitive performance, mood and cardiovascular functioning. Appetite 22:57–65; 1994.
26. Stroop, J. R.: Studies of interference in serial verbal reactions. J. Exp. Psychol. 18:643–663; 1935.
27. Warburton, D. M.: Effects of caffeine on cognition and mood without caffeine abstinence. Psychopharmacology (Berlin) 119: 66–70; 1995.
28. Wesnes, K.; Revell, A.: The separate and combined effects of scopolamine and nicotine on human information processing. Psychopharmacology (Berlin) 84:5–11; 1984.