The Journal of General Psychology, 2003, *130*(1), 70–86

State-Dependent Memory Effects Using Caffeine and Placebo Do Not Extend to Metamemory

WILLIAM L. KELEMEN
Department of Psychology
California State University, Long Beach

CATHERINE E. CREELEY
Department of Psychology
University of Missouri, St. Louis

ABSTRACT. The authors examined the impact of caffeine on human memory and predictions of memory (i.e., metamemory). On Day 1, 83 college students drank a sweetened beverage containing either caffeine (4 mg/kg body weight) or a placebo before they studied 40 pairs of words. While the participants studied, they predicted their future memory performance for each word pair. On Day 2, the participants again received caffeine or a placebo before the memory test. The participants who drank the same beverage on both days (either caffeine or a placebo) recalled more word pairs than did those who drank different beverages (caffeine on 1 day and a placebo on the other day). In contrast, memory predictions were more accurate when the beverages did not match on both days. These data provide evidence for state-dependent memory when caffeine is used, but not for state-dependent metamemory. People's memory and their predictions of memory can be influenced in different ways if they drink caffeine before they study or take a test.

Key words: caffeine, judgments of learning, metamemory, state-dependent memory

RESEARCH ON CAFFEINE AND HUMAN MEMORY has produced a complex and often inconsistent pattern of results. The putative memory effects of caffeine are not well understood, even though its use is widespread. For example, the literature on word-list recall under the influence of caffeine suggests mixed results. In one study, Warburton (1995) used low doses of caffeine (75 mg and

The authors thank John Dunlosky, George Taylor, and Chuck Weaver for their helpful comments on this article. They also thank Erica Wohldmann for her assistance with data entry.

Address correspondence to William L. Kelemen, Department of Psychology, 1250 Bellflower Boulevard, Long Beach, CA 90840; wkelemen@csulb.edu (e-mail).

150 mg) and found a significant dose-related improvement in delayed recall, but no effect of caffeine on immediate recall. In contrast, Terry and Phifer (1986) also used a low dose of caffeine (100 mg) but found that it impaired both the immediate and the delayed recall of word lists. In another study, caffeine improved recall in women, but only after practice, whereas it impaired recall in men at a dose of 2 mg/kg body weight, but not at 4 mg/kg (Arnold, Petros, Beckwith, Coons, & Gorman, 1987). Arnold and colleagues also reported that caffeine sometimes had no effect on either gender. Other researchers have failed to detect any significant effects of caffeine on either immediate or delayed recall (Loke, 1988; Mitchell & Redman, 1992).

A participant's level of impulsivity is an important variable that can moderate the cognitive effects of caffeine. The results of early research indicated that moderate levels of caffeine impaired the performance of introverts, who were already functioning at a high level of arousal, but that they improved the performance of less intrinsically aroused extroverts (Revelle, Amaral, & Turriff, 1976). Researchers in subsequent studies suggested that the Impulsivity subscale of Eysenck's Introversion–Extroversion factor was more specifically related to the effects of caffeine and that the time of day at which the caffeine was taken also showed an effect. Low-impulsive individuals were impaired by caffeine in the morning, whereas high impulsives were aided by caffeine in the morning. This pattern was reversed in the evening (Revelle, Humphreys, Simon, & Gilliland, 1980). Anderson and Revelle (1994) later examined recognition memory in low impulsives versus high impulsives in the morning and in the evening and they obtained similar results. On the basis of these data, researchers hypothesize that low impulsives are more aroused in the morning and become less aroused as the day passes. High impulsives may be less aroused in the morning but become more aroused during the afternoon and evening. The findings of a large body of additional research support the idea that impulsivity moderates the effects of caffeine (Anderson, Revelle, & Lynch, 1989; Gupta, 1991; Gupta & Gupta, 1990, 1999; Humphreys & Revelle, 1984; MacPherson et al., 1996).

Metamemory

In the present study, we focused on participants' monitoring and control of their own memory process. This process of monitoring and controlling memory is known as *metamemory* and is sometimes described as "what you know about what you know" (Metcalfe & Shimamura, 1994, p. xi). One can measure metamemory if one asks participants to learn novel information and then to make predictions (sometimes called judgments of learning, or JOLs) about the future recall of that information. The predicted recall of the items is then compared with the actual recall of the items to obtain a measure of metamemory. If metamemory accuracy is high, then participants should remember more items that received

high JOLs and they should recall fewer items that received lower JOLs. When memory predictions are made immediately after study (hereinafter, *immediate JOLs*), the correlation between JOL and recall is typically nonzero, but modest. However, if there is a delay of at least 30 s between study and JOL (hereinafter, *delayed JOLs*), predictive accuracy is very high. For example, Nelson and Dunlosky (1991) found that the correlation between predicted and actual recall for immediate JOLs was .38, whereas the correlation for delayed JOL accuracy was nearly perfect (mean correlation = .90). This delayed-JOL effect is robust across a variety of tasks (Dunlosky & Nelson, 1992, 1994, 1997; Kelemen, 2000; Kelemen & Weaver, 1997; Thiede & Dunlosky, 1994; Weaver & Kelemen, 1997).

The influence of two drugs that are used socially (alcohol and caffeine) on immediate and delayed JOLs has been reported in previous studies. Nelson et al. (1998) used alcohol in their study and they found that it lowered participants' recall of items and that it disrupted one aspect of metamemory: immediate JOLs. However, delayed-JOL accuracy was not affected by alcohol. Because Nelson et al. conducted their study during a single testing session, the encoding and retrieval processes occurred in the same pharmacological state, either drug or placebo.

Kelemen and Creeley (2001) tested the effects of another drug—caffeine—on memory and metamemory with a 2×2 (Drug State on Day $1 \times$ Drug State on Day 2) factorial design. They measured memory when the encoding and retrieval conditions matched (either drug or placebo on both days) and when they did not match (drug on 1 day, placebo on the other). This type of 2×2 design has been used to show state-dependent memory effects in humans for several classes of drugs (see Eich, 1980, for a review). Kelemen and Creeley tested the impact of caffeine in three memory tasks and found a trend toward state-dependent memory. However, the accuracies of both the immediate and the delayed JOLs were unaffected. One criticism of Kelemen and Creeley's study was that the multiple memory tasks they used may have interfered with each other. Interference between the tasks may have reduced the likelihood of observing state-dependent effects in any single task. Contrary to this earlier study, we used only one memory task in the present experiment.

We used a placebo-controlled, 2×2 between-groups design to test the effects of caffeine on a test of memory and metamemory while we controlled for the participants' levels of impulsivity. On Day 1, the participants ingested either 4 mg/kg of caffeine or a placebo before they studied word pairs. They also provided either immediate or delayed JOLs for each item during the period of study. On Day 2, the participants again ingested caffeine or a placebo before they completed a cued-recall test. If caffeine alters memory encoding, then a main effect of the drug should have been seen on Day 1; if caffeine alters memory retrieval, then a main effect should have appeared on Day 2. State-dependent memory would be evident if there was a significant interaction of drug effects on Day 1 and Day 2 in the direction of higher recall when the encoding and retrieval processes matched

(caffeine–caffeine or placebo–placebo) rather than when they did not match (caffeine–placebo or placebo–caffeine).

We also examined the effects of caffeine on metamemory. If caffeine influenced the participants' overall confidence (i.e., JOL magnitude), then a main effect of the drug should have appeared on Day 1, when the JOLs were provided. As noted heretofore, caffeine could influence encoding or retrieval, and the participants' JOLs could be made on the basis of either aspect of memory. According to Koriat (1997), immediate JOLs are made on the basis of information that is present during encoding, whereas delayed JOLs may be made primarily on the basis of the success or failure of retrieval at the time of JOL. Thus, if caffeine alters encoding, then a main effect of the drug should appear for immediate JOLs. However, if caffeine influences retrieval, then delayed JOLs should be altered. Finally, in the present study, we tested the hypothesis that caffeine would produce state-dependent metamemory effects. If so, the correlation between predicted recall and actual recall should have been higher when the drug states during JOL and recall matched, compared with when they did not match.

Method

Participants

A total of 83 undergraduate students (54 women and 29 men) participated in the study. They were paid $15 each. The equipment failed during the testing of 2 of the participants, so we were able to collect usable data from only 81 individuals. All volunteers were screened during a telephone interview so that we could exclude people who had certain health problems (e.g., hypertension, migraine headaches, panic attacks, vertigo, and epilepsy), individuals who consumed more than 500 mg of caffeine daily, and women who were pregnant or nursing a baby.

Materials

On Day 1, the participants completed 3 questionnaires. The first questionnaire comprised a general biographical profile, which included a follow-up screening for health problems; the second dealt with the participants' use of caffeine; and the third was the 30-item Barratt Impulsiveness Scale (BIS-11; Patton, Stanford, & Barratt, 1995). On both days, the participants completed a short, 20-item version of the Activation–Deactivation Adjective Check List (AD-ACL; Thayer, 1978), which is a self-report measure of physiological arousal. We took 40 paired-associate concrete nouns from Paivio, Yuille, and Madigan (1968) for the memory and metamemory test. The stimuli were presented on IBM-compatible PCs that used Micro Experimental Laboratory software (MEL, Version 2.0). The PCs were placed in individual cubicles.

Design and Procedures

We used a placebo-controlled, 2 × 2 (Drug State on Day 1 × Drug State on Day 2) between-groups design. We randomly assigned the participants to one of the four experimental conditions (caffeine–caffeine, caffeine–placebo, placebo–caffeine, or placebo–placebo), with the restriction that each condition was used an equal number of times. On each day, the participants received either 4 mg/kg of powdered caffeine mixed into a 6-oz orange drink (Tang®) or a placebo. We added a quarter teaspoon of salt to all the beverages to match the tastes of the drinks in the caffeine and placebo conditions. The participants were tested in groups of 1–4, at the same time on 2 consecutive days, for approximately 1 hr each day. Because the effects of caffeine can vary with time of day (e.g., Anderson & Revelle, 1994), we conducted all the testing sessions in the afternoon (between 12 p.m. and 6 p.m.). We instructed the participants to abstain from caffeine for 3 hr before they came to the experiment. We also instructed them not to eat anything for 1 hr before they arrived.

Day 1. When the participants arrived on Day 1, they completed the biographical questionnaire. They were weighed by the experimenter, and they then consumed their beverages. They completed the remaining questionnaires, and they had the option of playing a computer game (Solitaire) until 30 min after ingestion to allow for the absorption of the caffeine. After 30 min, all the participants completed the Thayer AD-ACL to measure their arousal before they began to study the word pairs and follow the JOL procedures.

The experimenter instructed the participants to study the 40 paired associates for a cued-recall test the next day. They studied each item twice on Day 1 so that they would achieve moderate levels of recall on Day 2. During the first study trial, the items were presented in a random order, for 6 s each. The items were then rerandomized and presented again for 6 s each. After the second study trial, the participants provided either an immediate or a delayed JOL for each item. The type of JOL assigned to each item was randomized, with the restriction that 20 immediate and 20 delayed JOLs were included.

In the immediate-JOL condition, immediately after the participant had finished studying, the cue word appeared above the JOL prompt: “How confident are you that you will be able to recall the second word of this word pair on a test tomorrow?” The participants were asked to rate themselves as 0% confident (definitely will not recall), 20%, 40%, 60%, 80%, or 100% confident (definitely will recall). In the delayed-JOL condition, study of the paired associates was followed by a prompt to “press the spacebar to continue.” Delayed JOLs were made in the same manner in which the immediate JOLs were made, but only after all the paired associates had been made. This procedure ensured that at least 19 items were presented between study and JOL. The order of item presentation and the assignment of JOL conditions were randomized for each participant. After the

participants had studied and rated all 40 items, the experimenter administered a posttest questionnaire to determine whether they thought their beverages had contained caffeine and if so, how it may have influenced their performance. In addition, the experimenter encouraged the participants to report any discomfort that they may have experienced during the procedures. Finally, the participants were asked to return at the same time the following day, and they were reminded to refrain from eating and from ingesting caffeine before the procedures on Day 2.

Day 2. When the participants arrived the next day, they immediately ingested their beverage and again played a computer game for 30 min, after which time they completed the arousal questionnaire. They then completed a computerized cued-recall test in the same cubicle as the one they had used on Day 1. On the memory test, the first word was shown to participants, and they were asked to type the associated word. The items were presented in a random order, and the participants were given as much time as they needed to finish the test. After the memory test, the participants completed the same questionnaire that they had completed at the end of the trial on Day 1. They were then debriefed and paid for their participation.

Dependent Measures and Covariates

Three dependent variables were of primary interest: JOL magnitude (confidence); recall (memory); and the relationship between JOL magnitude and recall (metamemory). In regard to metamemory, we analyzed whether the items that received higher JOLs were more likely to be recalled than were the items that received lower JOLs. We computed Goodman–Kruskal gamma correlations (*G*s) between JOLs and recall for each participant so that we could summarize the relationship (for the rationale concerning the use of *G*, see Nelson, 1984). *G* is a measure of association for ordinal data. It ranges from –1 to 1, with zero indicating a complete lack of predictive accuracy. In general, as the participants' relative memory-monitoring accuracy improves, the magnitude of the correlation between predicted and actual recall increases.[1]

The findings of previous research have shown that an individual's level of impulsivity can influence metamemory (Walczyk & Hall, 1989) as well as the cognitive effects of caffeine (Anderson & Revelle, 1994; Anderson, Revelle, & Lynch, 1989; Gupta, 1991; Gupta & Gupta, 1999; Humphreys & Revelle, 1984; MacPherson et al., 1996; Revelle, Humphreys, Simon, & Gilliland, 1980). Therefore, we included the participants' impulsivity scores (BIS-11) as a covariate in all the statistical analyses, except for the analyses of the demographic variables. We included two additional covariates. We instructed the participants to abstain from eating for at least 1 hr before they arrived because food intake can interact with caffeine (e.g., Smith, Kendrick, Maben, & Salmon, 1994; Smith, Rusted, Eaton-Williams, Savory, & Leathwood, 1990). Nevertheless, the amount of time that the participants reported since their last meal varied greatly—from 20 min

to 24 hr. To control for this variation, we included as covariates the amount of time (in hours) that the participants reported since their last meal and the meal serving size (coded as small, medium, or large). These data were collected on the pretest questionnaires that were administered on Day 1, so we did not include them in the analyses that involved Day 2 only.

Results

Three participants showed extreme performance on the memory test: 2 recalled only 1 item, and another recalled all 40 items correctly. These outcomes were not caused by caffeine because all 3 participants received a placebo on both days. Researchers in some studies of metamemory have excluded participants who score less than 5% correct or more than 95% correct to generate more stable outcome measures (e.g., Connor, Dunlosky, & Hertzog, 1997). We adopted these criteria and excluded the 3 participants from the subsequent analyses, which left a total sample size of 78 people. All tests of statistical significance were conducted at $p < .05$.

Demographics

The demographic data for the participants are summarized in Table 1. There were no statistically significant differences between groups for age, weight, daily caffeine use, or impulsivity when we conducted the one-way analyses of variance (ANOVAs).

TABLE 1
Demographic Characteristics of Participants by Drug State on Day 1 and Day 2

	Drug state (Day 1/Day 2)			
Characteristic	C/C	C/P	P/C	P/P
N	21	21	19	17
Age (years)				
M	23.7	23.3	23.7	21.7
SD	7.7	7.1	6.1	3.8
Weight (lb)				
M	163.2	178.5	149.0	143.5
SD	42.4	73.7	51.7	22.9
Average daily caffeine use (mg)				
M	112.2	86.2	99.4	115.0
SD	76.0	68.3	82.7	91.5
BIS–11 impulsivity scores				
M	61.6	63.3	66.0	62.7
SD	9.3	9.4	9.2	11.4

Note. C = caffeine. P = placebo.

Arousal

We examined the effects of caffeine on self-reported arousal with the use of the short-form AD-ACL (Thayer, 1978). We conducted an analysis of covariance (ANCOVA) for each day using type of substance (caffeine or placebo) as the independent variable. As noted heretofore, we included all three covariates (impulsivity, time since last meal, and meal size) in the Day 1 analysis, and we used impulsivity only as a covariate for Day 2. On Day 1, the participants who received caffeine were more aroused (M = 27.3, SD = 14.0) than were those who received the placebo (M = 20.5, SD = 12.2), $F(1, 73) = 4.74$, MSE = 174.4. On Day 2, caffeine again increased arousal (M = 29.1, SD = 12.2) compared with the placebo (M = 21.1, SD = 11.8), $F(1, 75) = 9.03$, MSE = 143.2. We computed η^2 as a measure of effect size and used guidelines based on Cohen (1988) to interpret η^2: 0.010 = small effect size, 0.059 = medium effect size, and 0.138 = large effect size (see Clark-Carter, 1997, for details). We obtained medium effect sizes, η^2 = .060 for Day 1 and η^2 = .110 for Day 2.

JOL Magnitude

The participants provided JOLs on Day 1 either immediately after they studied the items or after a brief delay. The mean JOL magnitude was similar whether caffeine (M = 0.48, SD = 0.17) or placebo (M = 0.46, SD = 0.15) had been administered. The mean JOL magnitude for immediate JOLs was 0.44 (SD = 0.16) compared with 0.50 (SD = 0.19) for delayed JOLs. We conducted a two-way mixed-design ANCOVA on the magnitude of JOLs with the use of Day 1 substance (caffeine or placebo) and type of judgment (immediate or delayed) as independent variables and all three covariates. There were no statistically significant main effects or interactions.

Recall

The mean performance on the memory test according to Day 1/Day 2 drug states was as follows: caffeine–caffeine (M = 0.37, SD = 0.23); caffeine–placebo (M = 0.31, SD = 0.23); placebo–caffeine (M = 0.33, SD = 0.19); and placebo–placebo (M = 0.37, SD = 0.15). Mean recall across conditions, adjusted for the aforementioned covariates, is shown in Figure 1. For both the adjusted and the unadjusted sets of means, recall was higher when the substance administered was the same on both days (either caffeine or placebo) compared with when different substances were administered. We performed a 2×2 ANCOVA using type of substance administered on Day 1 and Day 2 as independent variables, controlling for the effects of impulsivity, time since last meal, and meal size. Given the pattern of means shown in Figure 1, a significant Day 1 Substance $\times$ Day 2 Substance interaction would be evidence for state-dependent memory. As predicted, a statistically significant interaction emerged, $F(1, 71) = 4.18$, MSE = 0.04. The

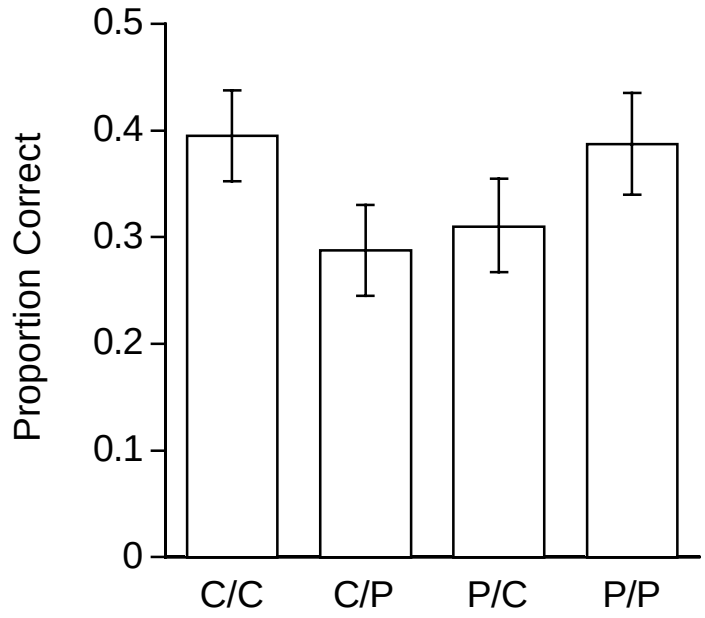

FIGURE 1. Mean recall as a function of drug state on Day 1 and Day 2 (*n* = 78), adjusted for impulsivity, time since last meal, and meal size. C = caffeine. P = placebo. Error bars represent ± 1 standard error of the mean.

observed effect size, $\eta^2 = .056$, was between small and medium. There were no significant main effects for type of substance. Thus, caffeine did not directly influence the encoding or retrieval processes (Day 1 and Day 2, respectively), but we did obtain evidence for state-dependent memory.

Metamemory Accuracy

We assessed the relative accuracy of the participants' metamemory judgments with the computation of the *G* correlations between JOLs and recall described heretofore. Nine participants showed a lack of response variability in either the immediate- or the delayed-JOL condition. Because *G* was indeterminate in these cases, the data from these participants were not included in the analyses hereinafter.[2] We considered two issues concerning the accuracy of memory monitoring. First, we predicted that the timing of JOLs would have a significant impact on metamemory accuracy (e.g., Nelson & Dunlosky, 1991). To test this hypothesis independent of drug effects, we conducted a paired *t* test on immediate- versus delayed-JOL accuracy, for which we included only the participants who had received a placebo on both days. Consistent with past findings, the

delayed JOLs (mean G = .79, SD = .24) were more accurate than were the immediate JOLs (mean G = .43, SD = .25), $t(15) = 4.17$.

A more important question concerned the influence of caffeine on the accuracy of metamemory. One possibility was that state-dependent memory monitoring might emerge and that JOL accuracy would be greater when the substances (caffeine or placebo) active during judgment and the test days were the same. We conducted a three-way mixed design ANCOVA to evaluate the influence of caffeine on both types of JOLs and to test for state-dependent metamemory. The independent variables were type of substance administered on Day 1 (caffeine or placebo), type of substance administered on Day 2 (caffeine or placebo), and type of judgment (immediate JOL or delayed JOL). As in previous analyses, impulsivity, hours since last meal, and meal size were included as covariates. There were no statistically reliable main effects, but two interactions were significant: a two-way Day 1 Substance × Day 2 Substance interaction, $F(1, 62) = 19.16$, $MSE = 0.07$, $\eta^2 = .240$; and a three-way, Day 1 × Day 2 × JOL Type interaction, $F(1, 62) = 5.29$, $MSE = 0.04$, $\eta^2 = .080$. The three-way interaction is shown in Figure 2. Contrary to our hypothesis, metamemory accuracy (G) was not higher when the pharmacological states during judgment and retrieval matched. In fact, we observed the opposite pattern. Mean Gs were higher when the participants ingested caffeine on Day 1 and placebo on Day 2, or

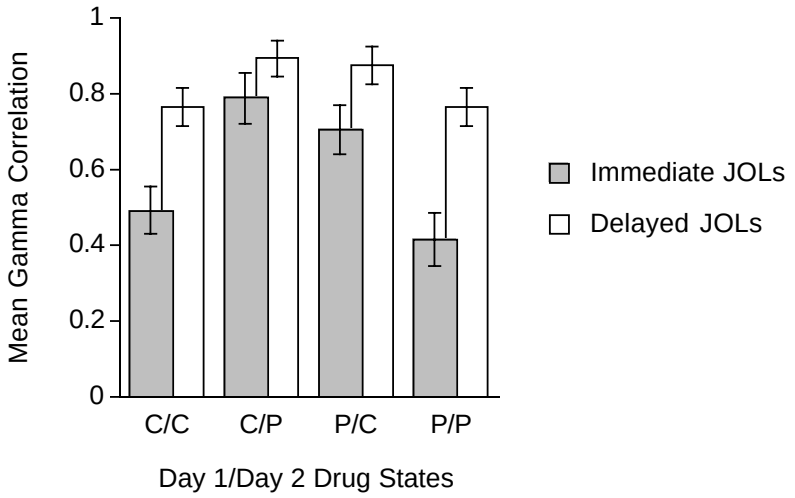

FIGURE 2. Mean gammas for immediate JOLs and delayed JOLs as a function of drug states on Day 1 and Day 2. JOLs = judgments of learning. C = caffeine. P = placebo. Means (n = 69) have been adjusted to control for the effects impulsivity, time since last meal, and meal size. Error bars represent ± 1 standard error of the mean.

vice versa (nonmatching drug states), compared with when the participants ingested either caffeine or placebo on both days (matching drug states).

We conducted two univariate ANCOVAs to explore the three-way interaction: the Day 1 Drug State × Day 2 Drug State interactions were tested separately for immediate JOLs and for delayed JOLs. The interaction was significant for both types of JOLs, $F(1, 63) = 18.99$, $MSE = 0.08$, and $F(1, 70) = 5.69$, $MSE = 0.04$, respectively. Thus, both immediate- and delayed-JOL accuracy was significantly better in nonmatching drug states. However, the magnitude of effect was much larger for immediate JOLs ($\eta^2 = .230$) than it was for delayed JOLs ($\eta^2 = .080$).

Recall of Easy Versus Difficult Items

Because of the surprising nature of these findings, we conducted an additional set of analyses to examine more closely the state-dependent interaction for recall. We hypothesized that the impact of nonmatching drug states on recall might have been unequally distributed across items that received different JOL ratings. If a change in drug state selectively impaired memory for items that were rated low at the time of JOL (presumably the more difficult items), but had less impact on items that were rated high (easy items), then the mean *G* correlation between JOLs and recall might increase in nonmatching drug states because the participants remembered fewer of the items that were rated low. To test this idea, we examined recall of difficult items (those receiving the three lowest JOL ratings) separately from recall of easy items (those receiving the three highest JOL ratings). Items were classified as easy or difficult on the basis of all overall JOL ratings (i.e., both the immediate and the delayed JOLs). We conducted ANCOVAs on easy and difficult items using drug state (matching vs. nonmatching) as the independent variable and including impulsivity, time since last meal, and meal size as covariates. We found no significant difference in recall for easy items in matching drug states ($M = 0.57$, $SD = 0.24$) compared with nonmatching drug states ($M = 0.56$, $SD = 0.24$). However, difficult items were significantly less likely to be recalled in nonmatching drug states ($M = 0.13$, $SD = 0.17$) compared with matching drug states ($M = 0.20$, $SD = 0.18$), $F(1, 73) = 6.21$, $MSE = 0.03$. This difference represents a medium effect size, $\eta^2 = .080$. Thus, to the extent that nonmatching drug states interfered with recall, the impact was statistically significant only for items that were previously rated as less likely to be recalled. The importance of this observation is considered in the next section.

Discussion

We examined three main issues in regard to caffeine, memory, and metamemory. First, we obtained reliable state-dependent memory effects, although we observed no main effects of caffeine for encoding processes (Day 1) or for retrieval processes (Day 2). Recall was higher when the participants were in the

same drug state (caffeine or placebo) for both days than when they were in different drug states on the 2 days. Second, caffeine did not affect the magnitude of the metamemory judgments that the participants provided on Day 1—we found no evidence that caffeine influenced the participants' confidence in their future memory performances. Third, contrary to our predictions, the participants' metamemory judgments were more accurate when the drug states on Day 1 and Day 2 did not match. We found no evidence to support the hypothesis of state-dependent metamemory, whereas we did observe state-dependency effects for recall. Thus, caffeine may influence memory and metamemory in different ways.

Caffeine and Memory

In the present study, we obtained evidence for state-dependent memory effects in college students. Unadjusted mean recall differed by about 5% across conditions, and the effect size was modest. Although drug–state dependency has been demonstrated for several classes of drugs in humans, the effects of caffeine are often variable. The results of previous studies may have been inconsistent because important moderating variables were not adequately controlled. Revelle and colleagues (Anderson & Revelle, 1994; Revelle et al., 1980) have shown that several factors, including impulsivity and time of day, can moderate the effects of caffeine (see van der Stelt & Snel, 1998, for a list of other proposed moderators). In this study, we tested all the participants in the afternoon, and the effects of impulsivity and food intake were statistically controlled.

As with any pharmacological study, one might wonder whether the memory effects would change at different doses of caffeine. We administered a moderate dose of caffeine (4 mg/kg)—equivalent to about 2 cups of coffee in a 150-lb (68-kg) person—which many other researchers have used in previous studies of caffeine and cognition (e.g., Anderson & Revelle, 1994; Arnold et al., 1987; Gupta, 1991; Liguori, Grass, & Hughes, 1999; MacPherson et al., 1996; Mitchell & Redman, 1992). An average adult in the United States consumes this amount of caffeine per day (Barone & Roberts, 1996). It is possible that significant main effects of caffeine on encoding or retrieval might emerge at higher doses, such as those common in animal studies (Gauvin & Holloway, 1999). However, our dose had the advantage of being large enough to be noticed by the participants (e.g., in self-reported arousal) and to still be relevant to the level of caffeine use that is often reported by adults.

Caffeine and Metamemory

Another goal of this study was to explore how caffeine influences two aspects of metamemory: JOL magnitude and JOL accuracy. In regard to the former, researchers have shown that caffeine influences mood and arousal (Rusted, 1999), and we wondered whether these changes would be reflected in the magnitude of the participants' memory judgments. Although caffeine increased self-reported

arousal, we observed no reliable differences in JOL magnitude between people who received caffeine and people who received a placebo. These findings replicate the results reported by Kelemen and Creeley (2001). Although null results are difficult to interpret, we have tested a total of 225 students in both studies with the use of four different tasks, and we have seen no evidence that caffeine influences JOL magnitude. Thus, caffeine does not seem to alter participants' overall confidence in their future memory performance.

Contrary to our expectations, metamemory accuracy (*G*) was higher when the drug states on Days 1 and 2 did not match, which is the opposite of state-dependent metamemory. The magnitude of this effect was larger for immediate JOLs than it was for delayed JOLs, perhaps because delayed JOL accuracy was already near ceiling. Another possibility is that immediate JOLs are more sensitive to changes in drug states than are delayed JOLs. Only one other study has shown a significant drug effect on JOLs (Nelson et al., 1998), and in that case, alcohol altered immediate JOL accuracy but it did not alter delayed JOL accuracy. It seems that immediate and delayed JOLs are based on different mnemonic cues (e.g., Koriat, 1997), so future research should perhaps examine whether these two judgments also differ in their sensitivity to other pharmacological manipulations.

We cannot conclusively explain the finding of better metamemory in nonmatching drug states. One possible explanation is that nonmatching conditions impair the recall of difficult items more than they impair the recall of easy items without altering JOLs. If so, *G* correlations between JOLs and recall would be higher in nonmatching conditions. Consider two students with identical sets of JOLs (mean JOL for difficult items = 0.25 and mean JOL for easy items = 0.75) but different memory performance. Student A (in a matching drug–state condition) recalls 35% of the difficult items and 75% of the easy items, whereas Student B (in a nonmatching condition) recalls only 25% of the difficult items and 75% of the easy items. Student A shows better overall recall than does Student B, but *G* would tend to be higher for Student B because that person recalled fewer of the items that were rated lower. In this case, the higher *G* for Student B is a result of poorer memory for difficult items.

Consistent with this interpretation, caffeine had no significant effect on JOL magnitude. However, nonmatching drug states significantly impaired the recall of difficult items, but the recall of easy items did not vary between conditions. In addition, the effect size of the state-dependent interaction in recall was between small and medium. It seems plausible that a variable with a modest effect on recall might exert its strongest influence on items that were originally learned less well. However, our explanation is post hoc, because the finding was unexpected. We did not manipulate item difficulty directly, but instead relied on the participants' JOL ratings to determine which items were perceived to be the most difficult for each person. If we were to include items that had been normed for difficulty, then it would have allowed for a more precise assessment of the relationship between state dependency and item difficulty.

The results of the present study complement earlier findings in two ways. First, they show that caffeine can produce state-dependent memory in cued recall. Kelemen and Creeley (2001) found a significant drug–state interaction when they used free recall but not cued recall or recognition, which was consistent with the proposal by Eich (1980) that free-recall tests are maximally sensitive to state-dependent effects. However, the use of multiple tasks in that study may have reduced the likelihood of observing state-dependent effects in less sensitive tests because state-dependent memory did emerge in the present study when we used a cued-recall test alone. Second, the present findings are the first to show differences in metamemory accuracy (*G*) according to caffeine state. However, we have argued that these differences may be owing to the influence of caffeine on memory rather than its influence on JOLs. That Kelemen and Creeley (2001) did not observe the memory effects of caffeine when they used cued recall may explain why no changes in *G* appeared in that study.

The accurate assessment of one's own state of knowledge can be important in numerous situations. Consider the case of a student who is preparing for an examination. During study, students must select effective study strategies, judge how likely they are to remember target information on a future exam, and allocate their study time accordingly. These metamemory decisions can affect future test performance. Our data suggest that students' memory and metamemory accuracy can be altered by the interaction of caffeine drug states during study and test. In a broader sense, these data highlight the value of including measures of metamemory in pharmacological studies of cognition so that the more subtle aspects of the effects of a drug on different memory systems can be detected. Variation in caffeine states from day to day can produce modest changes in different aspects of cognition.

NOTES

1. We also examined bias scores as a second measure of metamemory accuracy. Bias is the signed difference between a participant's JOLs and recall (positive bias scores indicate overconfidence and negative bias scores indicate underconfidence). However, there were no significant differences for bias scores in any condition, and we have not reported those null results.

2. *G* correlations are undefined when there is a lack of variability in predictions (providing the same JOL rating for all items in a given condition) or recall (remembering all or none of the items in a given condition). One participant correctly recalled all 20 delayed JOL items; 8 participants did not recall any of the immediate JOL items. Thus, *G* correlations could not be computed in these instances.

REFERENCES

Anderson, K. J., & Revelle, W. R. (1994). Impulsivity and time of day: Is rate of change in arousal a function of impulsivity? *Journal of Personality and Social Psychology, 67,* 334–344.

Anderson, K. J., Revelle, W., & Lynch, M. J. (1989). Caffeine, impulsivity, and memory scanning: A comparison of two explanations for the Yerkes–Dodson effect. *Motivation and Emotion, 13,* 1–20.

Arnold, M. E., Petros, T. V., Beckwith, B. E., Coons, G., & Gorman, N. (1987). The effects of caffeine, impulsivity, and sex on memory for word lists. *Physiology & Behavior, 41,* 25–30.

Barone, J., & Roberts, H. (1996). Caffeine consumption. *Food & Chemical Toxicology, 34,* 119–129.

Clark-Carter, D. (1997). *Doing quantitative psychological research: From design to report.* East Sussex, UK: Psychology Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Connor, L. T., Dunlosky, J., & Hertzog, C. (1997). Age-related differences in absolute but not relative metamemory accuracy. *Psychology and Aging, 12,* 50–71.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments for learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20,* 374–380.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language, 33,* 545–565.

Dunlosky, J., & Nelson, T. O. (1997). Similarity between the cue for judgments of learning (JOL) and the cue for test is not the primary determinant of JOL accuracy. *Journal of Memory and Language, 36,* 34–49.

Eich, J. E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition, 8,* 157–173.

Gauvin, D. V., & Holloway, F. A. (1999). The subjective effects of caffeine: Bridging the gap between animal and human research. In B. S. Gupta & U. Gupta (Eds.), *Caffeine and behavior: Current views and research trends* (pp. 257–281). Washington, DC: CRC Press.

Gupta, U. (1991). Differential effects of caffeine on free recall after semantic and rhyming tests in high and low impulsives. *Psychopharmacology, 105,* 137–140.

Gupta, U., & Gupta, B. S. (1990). Caffeine differentially effects kinesthetic aftereffect in high and low impulsives. *Psychopharmacology, 102,* 102–105.

Gupta, U., & Gupta, B. S. (1999). Caffeine, impulsivity, and performance. In B. S. Gupta & U. Gupta (Eds.), *Caffeine and behavior: Current views and research trends* (pp. 191–205). Washington, DC: CRC Press.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91,* 153–184.

Kelemen, W. L. (2000). Metamemory cues and monitoring accuracy: Judging what you know and what you will know. *Journal of Educational Psychology, 92,* 800–810.

Kelemen, W. L., & Creeley, C. E. (2001). Caffeine (4 mg/kg) influences sustained attention and delayed free recall but not memory predictions. *Human Psychopharmacology: Clinical and Experimental, 16,* 309–319.

Kelemen, W. L., & Weaver, C. A., III. (1997). Enhanced metamemory at delays: Why do judgments of learning improve over time? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1394–1409.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126,* 349–370.

Liguori, A., Grass, J. A., & Hughes, J. R. (1999). Subjective effects of caffeine among introverts and extraverts in the morning and evening. *Experimental and Clinical Psychopharmacology, 7,* 244–249.

Loke, W. H. (1988). Caffeine consumption by college undergraduates. *Psychology: A Journal of Human Behavior, 25,* 8–11.

MacPherson, J., Sternhagen, S., Miller, T., Devitt, M., Petros, T. V., & Beckwith, B. (1996). Effect of caffeine, impulsivity, and gender on the components of text processing and recall. *Experimental and Clinical Psychopharmacology, 4,* 438–446.

Mitchell, P. J., & Redman, J. R. (1992). Effects of caffeine, time of day, and user history on study-related performance. *Psychopharmacology, 109,* 121–126.

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: Knowing about knowing.* Cambridge, MA: MIT Press.

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin, 95,* 109–133.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science, 2,* 267–270.

Nelson, T. O., Graf, A., Dunlosky, J., Marlatt, A., Walker, D., & Luce, K. (1998). Effects of acute alcohol intoxication on recall and on judgments of learning during the acquisition of new information. In G. Mazzoni & T. O. Nelson (Eds.), *Metacognition and cognitive neuropsychology: Monitoring and control processes* (pp. 161–180). Mahwah, NJ: Erlbaum.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs, 76*(1, Pt. 2).

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51,* 768–774.

Revelle, W., Amaral, P., & Turriff, S. (1976). Introversion/extroversion, time, stress, and caffeine: The effect on verbal performance. *Science, 192,* 149–150.

Revelle, W., Humphreys, M. S., Simon, L., & Gilliland, K. (1980). The interactive effect of personality, time of day, and caffeine: A test of the arousal model. *Journal of Experimental Psychology: General, 109,* 1–31.

Rusted, J. (1999). Caffeine and cognitive performance: Effects on mood or mental processing? In B. S. Gupta & U. Gupta (Eds.), *Caffeine and behavior: Current views and research trends* (pp. 221–229). Washington, DC: CRC Press.

Smith, A. P., Kendrick, A., Maben, A., & Salmon, J. (1994). Effects of breakfast and caffeine on cognitive performance, mood, and cardiovascular functioning. *Appetite, 22,* 39–55.

Smith, A. P., Rusted, J. M., Eaton-Williams, P., Savory, M., & Leathwood, P. (1990). Effects of caffeine given before and after lunch on sustained attention. *Neuropsychobiology, 23,* 160–163.

Terry, W. S., & Phifer, B. (1986). Caffeine and memory performance on the AVLT. *Journal of Clinical Psychology, 42,* 860–863.

Thayer, R. E. (1978). Factor analytic and reliability studies on the Activation–Deactivation Adjective Check List. *Psychological Reports, 42,* 747–756.

Thiede, K. W., & Dunlosky, J. (1994). Delaying students' metacognitive monitoring improves their accuracy in predicting their recognition performance. *Journal of Educational Psychology, 86,* 290–302.

van der Stelt, O., & Snel, J. (1998). Caffeine and human performance. In J. Snel & M. M. Lorist (Eds.), *Nicotine, caffeine, and social drinking: Behaviour and brain function* (pp. 167–183). Amsterdam: Harwood Academic.

Walczyk, J. J., & Hall, V. C. (1989). Is the failure to monitor comprehension an instance of cognitive impulsivity? *Journal of Educational Psychology, 81,* 294–298.

Warburton, D. M. (1995). Effects of caffeine on mood without caffeine abstinence. *Psychopharmacology, 119,* 66–70.

Weaver, C. A., III, & Kelemen, W. L. (1997). Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? *Psychological Science, 8,* 318–321.

Manuscript received October 31, 2001
Revision accepted for publication May 16, 2002