International Journal of Behavioral Medicine
2006, Vol. 13, No. 4, 330–339

Effects of Expectation and Caffeine on Arousal, Well-Being, and Reaction Time

Rainer Schneider, Mauritz Grüner, Alexandra Heiland, Martina Keller, Zuzana Kujanová, Martin Peper, Maximilian Riegl, Stefan Schmidt, Petra Volz, and Harald Walach

The objective of this study is to determine the impact of expectation associated with placebo and caffeine ingestion. We used a three-armed, randomized, double-blind design. Two three-armed experiments varying instruction (true, false, control) investigated the role of expectations of changes in arousal (blood pressure, heart rate), subjective well-being, and reaction time (RT). In Experiment 1 (N = 45), decaffeinated coffee was administered, and expectations were produced in one group by making them believe they had ingested caffeinated coffee. In Experiment 2 (N = 45), caffeinated orange juice was given in both experimental groups, but only one was informed about the true content. In Experiment 1, a significant effect for subjective alertness was found in the placebo treatment compared to the control group. However, for RT and well-being no significant effects were found. In Experiment 2, no significant expectancy effects were found. Caffeine produced large effects for blood pressure in both treatments compared to the control group, but the effects were larger for the false information group. For subjective well-being (alertness, calmness), considerable but nonsignificant changes were found for correctly informed participants, indicating possible additivity of pharmacologic effect and expectations. The results tentatively indicate that placebo and expectancy effects primarily show through introspection.

Key words: arousal, blood pressure, expectation, placebo effects, reaction time, well-being

Beecher's sweeping article on the powerfulness of placebo (Beecher, 1955) instigated an ever-growing and passionate controversy. The discussion was related to a variety of issues, for instance, methodological (Hróbjartsson, 2002), etiological (Papakostas & Dras, 2001), definitional (Moerman & Jonas, 2002), or ethical (Temple & Ellenberg, 2000). From the debate on placebo effects, it can be derived that they are best studied by randomizing participants to treatment groups, blinding participants and experimenters, and including zero control groups (Haour, 2005; Walach & Jonas, 2004). This is usually achieved by employing randomized controlled trials (RCT), but the mechanisms of the placebo effect cannot be fully understood when using placebos only as control against a specific treatment (Schneider, in press). Research involving systematic variation of treatment conditions shows that the psychological meaning of placebo administration may change considerably. For example, in a reevaluation of a meta-analysis (Hróbjartsson & Gøtzsche, 2001) conducted by Vase, Riley, and Price (2002), including clinical and experimental studies, an overall effect of $d = .91$ was found when the placebo effect was investigated in terms of the informational context and its psychological meaning. This effect was seven times larger than the one found by Hróbjartsson and Gøtzsche ($d = .13$), who restricted their analysis on placebos as controls for pharmacological effects.

Rainer Schneider and Stefan Schmidt, Institute for Environmental Medicine and Hospital Epidemiology & European Office of the Samueli Institute, University Hospital, Freiburg, Germany; Mauritz Grüner, Alexandra Heiland, Martina Keller, Martin Peper, Maximilian Riegl, and Petra Volz, Department of Psychology, University of Freiburg, Germany; Zuzana Kujanová, Department of Psychology, University of Prešov, Slovakia; Harald Walach, Institute for Environmental Medicine and Hospital Epidemiology & European Office of the Samueli Institute, University Hospital, Freiburg, Germany and University of Northampton, School of Social Sciences, & European Office of the Samueli Institute, Northampton, UK.

This study was funded by the Samueli Institute, Alexandria, Virginia. We are grateful to the participants as well as to W. Keller and A. Schneider for their commitment as experimenters. We also thank R. Trittler for sharing his pharmacologic expertise with us.

Correspondence concerning this article should be addressed to Rainer Schneider, University Hospital Freiburg, Institute for Environmental Medicine and Hospital Epidemiology, Germany, Breisacher Strasse 115 B, 79106 Freiburg. E-mail: rainer.schneider@uniklinik-freiburg.de

One way to bring about the placebo effect is to experimentally elicit expectations (Kirsch, 1999; Stewart-Williams & Podd, 2004). They may be simply produced. For instance, in an experiment investigating placebo analgesia in irritable bowel syndrome patients, verbal suggestion containing coarse information on the agent's pain alleviating properties significantly and equally, the verum reduced pain intensity and pain unpleasantness over time (Vase, Robinson, Verne, & Price, 2003). Most important, pain expectancy explained 77% of pain intensity, an effect amounting to a difference from natural history of $d = 2.0$. In a study by the same authors (Verne, Robinson, Vase, & Price, 2003) mimicking the setting of a clinical trial, that is, administering placebo analgesia without fostering expectations via verbal instructions, this difference was less than half as large ($d = .89$).

The Caffeine Placebo Paradigm

As an experimental model, the caffeine placebo paradigm is particularly beneficial because the effects can be studied in healthy individuals, and there is a prevailing social stereotype for the effects of coffee (even for nonhabitual drinkers) embracing a wide range of bodily reactions. Decaffeinated coffee has repeatedly yielded placebo effects when participants were made to believe that they consumed caffeine (Fillmore, Mulvihill, & Vogel-Sprott, 1994; Fillmore & Vogel-Sprott, 1992; Kirsch & Weixel, 1988; Lotshaw, Bradley, & Brooks, 1996). These effects have shown in various indicators, such as physiological measures (e.g., blood pressure), behavioral variables (e.g., psychomotor performance), and subjective experience (e.g., mood). However, placebo effects show in different measures, and establishing the most sensitive indicators is difficult.

Two large German studies, conducted in our research group, failed to replicate caffeine placebo effects for physiological, psychological, and cognitive parameters (Walach, Schmidt, Bihr, & Wiesch, 2001; Walach, Schmidt, Dirhold, & Nosch, 2002). Varying expectations, provided by accurate information, deceptive information, or ambiguous information (double-blind condition), did not produce effects distinct from those found in the control group. At least two reasons appeared to account for these findings: First, the induced expectations could have been too weak to produce an effect because the "dose" was suggested to equal one cup of coffee, for which many people do not expect too strong an effect. Second, both studies were the first to involve complete double-blindness in the sense that the experimenters were ignorant to both research hypotheses and the experimental designs. This suggests that past positive findings on caffeine placebo effects could not have been exempt from systematic biases such as experimenter effects.

Purpose of this Study

This study was designed to investigate placebo and expectancy effects in two different ways. Placebo effects were assumed when no active agent was administered, but when participants were made to expect caffeine effects. In accordance with the definition proposed by Moerman and Jonas (2002), we defined the placebo effect as the psychological/physiological effect produced by the meaning of the intervention and the expectation associated with it. Such expectations may bear on conditioning (i.e., learning of physiological effects). For example, on the one hand, individuals who are made to believe they drink caffeinated coffee, when actually drinking decaffeinated coffee, experience changes in their functioning simply by smelling or tasting the beverage (Ader, 1993, 2000; Wickrameskera, 1980). On the other hand, such effects do not exclusively depend on concrete learning episodes, but rather may be brought about by prevailing stereotypes about the agent. To explore this, in the first experiment we administered placebo (decaffeinated coffee) and varied the instruction of participants. Expectations were implemented by brewing a "very strong" cup of decaffeinated coffee and by deceiving one experimental group as to the content of the beverage by leading them to believe they ingested caffeinated coffee. It was expected that these individuals should display a placebo effect such that they actually "responded" to the alleged active agent.

Alternatively, expectancy effects may be brought about by mechanisms other than prior learning history or prevailing stereotypes, for example, by camouflaging the active agent in a substance normally not associated with a stimulating effect (Flaten & Blumenthal, 1999; Mikalsen, Bertelsen, & Flaten, 2001). In this case, expectancy effects are separable from pharmacologic effects and, in conjunction, add up to an effect exceeding that of any single component. Thus, in the second experiment, caffeine was mixed with orange juice, and information was again varied such that one experimental group was falsely informed regarding the content of the beverage.

Therefore this study aimed at disentangling pharmacologic from psychological effects: Experiment 1 investigated social stereotypes of decaffeinated coffee and its concomitants (e.g., conditioned effects of smell and taste), whereas Experiment 2 explored expectations over and above pharmacologic effects. Both experiments were run under double-blind conditions in the sense that neither the falsely informed experimental groups nor the experimenters knew about the content of the substances administered. To assess the placebo effect (rather than the placebo response inflated by confounding factors such as regression, time effects, spontaneous fluctuations, etc.), zero control

groups were included against which the placebo groups were compared.

Methods

Experiment 1

Participants

The sample in Experiment 1 consisted of $N = 45$ German adults (35 women and 10 men) recruited from responses to a local newspaper advertisement addressing a general interest in the investigation of caffeine effects. Participants were included in the study if they met none of the following criteria: pregnancy, breast feeding, consumption of medicine or drugs, heart or circulation disease, continuing psychological or psychiatric treatment, and nonage. The mean age of the sample was 31 years ($SD = 11.8$ years; range, 18 to 62 years). All participants signed informed consent prior to the start of the study. With their signature, participants acknowledged that they were informed about the purpose of the study and the effects of caffeine and that they participated voluntarily.

Measures

Blood pressure (systolic and diastolic) and heart rate were measured with a calibrated digital oscillometric sphygmomanometer, the boso-carat (boso Inc., Germany), which automatically inflates the arm cuff and shows the values on a Liquid Cristal display. Participants were asked to rest for 10 min before having their blood pressure taken. They sat on a chair relaxing with their extended left arm lying on a table. The cuff was wrapped around the upper arm, with the lower edge placed 1–2 cm above the inner side of the elbow joint. The level of the cuff was placed at the same level as the heart during measurement.

Reaction time (RT) was measured with a test module from the interactive Test Battery for Attentional Performance (Zimmerman & Fimm, 1992), which was developed to assess subfunctions of attention. This test measured alertness reactions by means of a simple stimulus-response paradigm with a visual imperative stimulus (cross) presented on a computer screen at a distance of 60 cm. During the test, the participants had their forearm comfortably laid on the table to handle the key panel. Each trial contained 20 stimuli that appeared in the middle of the screen according to an algorithm varying interstimulus intervals. Each cross had to be responded to within a time window of 2 sec. Responses shorter than 100 msec were automatically repeated. This alertness subtest measured the ability to maintain or increase arousal when stimuli are expected and are assessed by averaging the single RTs per trial. Retest reliability coefficient is $r = .81$, and repeated measurements are not susceptible to learning effects.

Subjective well-being was assessed with the Multidimensional Well-Being Questionnaire (Steyer, Schwenkmezger, Notz, & Eid, 1997). It assesses current well-being according to the dimensions positive/negative mood, alertness/weariness, and calmness/disconcertment and has been widely used in German psychopharmacological studies. Each scale consists of eight bipolar items with five anchors from which minima and maxima are labeled (*not at all* to *very much so*). Internal consistency and test–retest reliability of all scales is very good (≥ .87).

Subjective expectations about the effects of caffeinated coffee on blood pressure, heart rate, arousal, and cognitive efficiency were assessed by a self-constructed 5-point Likert scale rating from strongly increase to strongly decrease. Furthermore, drinking habits (frequency, amount) and estimation of actual consumption of the alleged beverage were assessed (*very certain* to *very uncertain*).

Procedure

All participants were asked to fast for a period of 4 hr prior to the experiment. They were asked to refrain from consumption of substances containing stimulants or caffeine (e.g., chocolate, cola) 24 hr prior to participation. They were greeted by a female experimenter in the experimental room of the Department of Psychology, University of Freiburg, and introduced to the study. The experimenter was blind in the sense that she had no knowledge of the substance administered in the experimental condition where participants were falsely informed (i.e., she believed that all participants received a strong cup of coffee). During the whole session of about 1 hr, participants were asked to remain seated in order to reduce artifactual impact of movement on the measurement of blood pressure and heart rate. All dependent measures were taken at two time points, before and after administration of the beverage. To avoid ceiling effects ("white coat hypertension"), the physiological measures were assessed three times within a period of approximately 5 min; baseline measurements were calculated by averaging across the repetitive measurements. Thereafter, baseline values for well-being were taken. Baseline measures for RT were taken after a short trial run of 5 min.

Following the baseline measurements, the experimenter opened an opaque and numbered envelope containing the assignment to the experimental group that had been randomly generated by the first author before the experiment. Group allocation was according to a random sequence generated by the statistical software SPSS using a pseudo-random algorithm. Participants in the treatment group "True information" ($n = 15$) were told that they were to consume a "very strongly"

dosed cup of black and unsweetened decaffeinated coffee. The experimental coffee drink was prepared before the eyes of the participants by taking three heaped scoops (approximately 20 g) out of a can named "Group 1" and brewing it with 125 ml of water in a coffee machine, as witnessed by all participants. Participants of the treatment group "False information" (n = 15) were told they would drink a very strong cup of black and unsweetened regular (i.e., caffeinated) coffee that was taken out of a can named "Group 2." Hence, the instructions in this group aimed at facilitating positive effects associated with learned expectancy of (caffeinated) coffee. Participants of the control group (n = 15) were not exposed to sensory cues of coffee preparation, and they were informed to belong to the group drinking no beverage and receiving no instructions. While the coffee was brewing, participants of the treatment groups were asked to specify their expectations how the beverage would affect them. Participants of the control group were to rate their general expectations of how a strong cup of coffee normally affects them. Then all participants were asked to read a one-page flyer about the "scientifically undisputed" effects of caffeine on the autonomous nervous system, cognitive and bodily efficiency, cardiovascular system, and alertness. All participants were asked to attest their consent by providing their signature. The beverage was to be consumed within 2 min. After that, a waiting period of 15 min followed to increase participants' suggestion in the "false information" group that caffeinated coffee was administered that had to take its effect. During this time, participants were allowed to read magazines.

After the waiting period, posttreatment measures were taken in the same order as before. Blood pressure and heart rate again were also averaged across measurements. Finally, participants of the treatment groups were asked to rate their coffee drinking habits and to estimate whether they actually had consumed the alleged beverage. Subsequently, participants were remunerated with €10 and dismissed. Upon completion of the study, participants of the false information group were informed about the rationale of the study and the fact that they had been administered decaffeinated coffee.

Statistical Analyses

Preplanned analyses were repeated analyses of covariance (ANCOVAs) conducted on physiological, psychological, and reaction measures. In all analyses, baseline values served as the covariate. Effects for group differences were calculated according to the measure d by Cohen (1988). For all measures, a significant difference between the false information group and the control group (placebo effect) was expected. To rule out physiological and psychological effects due to the consumption of decaffeinated coffee, we hypothesized that the true information group and the control group would not differ.

Results

Physiological Measures

All measures were normally distributed and fell within the range of normal physiological values. Retest reliabilities of the three measurements before and after treatment were high and showed no outliers that could have affected measurement validity (systolic blood pressure $r \geq .88$, diastolic blood pressure $r \geq .94$, heart rate $r \geq .78$) and the posttreatment measures ($r \geq .89$, $r \geq .93$, $r \geq .86$, respectively). Therefore, averaging the three measurements at each time point did not appear to have been subject to disproportional biases.

The ANCOVA with systolic blood pressure as dependent variable did not yield any significant difference between the three groups (F[2, 41] = 1.03; p = .37). As can be seen in Table 1, the difference between the experimental group ("false information") and the control group was relatively small (d = .33). Similarly, the ANCOVA with diastolic blood pressure as the de-

Table 1. *Mean, Standard Deviation and p Values for Baseline and Posttreatment Measures of Blood Pressure (mmHg), Heart Rate (Beats per Minute), Reaction Time (ms), Mood, Alertness, and Calmness in Experiment 1*

	True					False					Control				
	Baseline		Post			Baseline		Post			Baseline		Post		
Variable	*M*	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*
SBP	113.2	12.5	111.7	12.3	.315	109.6	13.3	107.4	12.1	.228	110.6	8.7	106.2	11.3	.026
DBP	70.3	9.3	69.3	9.3	.394	66.9	10.9	68.9	11.0	.114	65.1	9.35	64.3	10.2	.471
HR	78.4	7.1	73.6	8.2	.003	69.1	7.1	70.8	4.8	.035	64.3	10.2	70.2	4.8	.004
RT	245.0	57.0	240.0	54.0	.565	237.0	29.0	229.0	23.0	.148	222.0	40.0	223.0	27.0	.882
Mood[a]	31.1	7.4	31.8	5.4	.736	33.9	5.5	34.8	4.0	.225	31.3	6.0	32.8	5.1	.515
Alertness[a]	30.3	7.4	27.5	7.4	.048	26.9	7.6	29.5	6.9	.050	25.9	7.4	25.1	7.4	.593
Calmness[a]	31.4	4.6	31.6	4.0	.853	32.6	5.9	32.6	5.2	1.0	31.8	5.2	33.5	4.9	.150

Note. SBP = Systolic blood pressure; DBP = Diastolic blood pressure; HR = Heart rate; RT = Reaction time. p = p value (two-tailed) for difference baseline–post.

[a]Alertness–calmness range: 8–40; high value indicates good mood, high alertness, high calmness.

pendent variable failed to show a significant effect either (F[2, 41] = 2.05; p = .14). However, the difference between the experimental group and the control group was twice as large as for the systolic pressure (d = 0.65). The ANCOVA for heart rate also failed to show an effect (F[2, 41] = .12; p = .89). The effect for the difference between the experimental group and the control group was very small (d = 0.12). Thus none of the three physiological parameters was indicative of a placebo effect in the sense that participants of the falsely informed treatment group displayed stronger increases in blood pressure or heart rate.

Reaction Time

As depicted in Table 1, participants' reactions to the stimuli decreased in both experimental groups. However, this effect was very small and did not reach statistical significance (F[2, 41] = .25; p = .78). Hence, H1 assuming a placebo effect for the misinformed treatment group could not be confirmed. The difference between the falsely informed group and the control group was small (d = 0.20).

Subjective Well-Being

For both time points, participants generally described themselves as being good tempered, alert, and calm (cf. Table 1). Whereas, for the dimension "mood" and "calmness," no differences were found (F[2, 41] = 2.42; p = .086, and F[2, 41] = .68; p = .52, respectively), the difference between the false information group and the control group was significant for "alertness" (F[2, 41] = 3.49; p = .04), with participants made to believe they drank a strong cup of coffee feeling more alert than the control group. This effect, large in size (d = 0.75), is displayed in Figure 1. The respective effects for mood and calmness were d = 0.12 and d = –0.32.

Discussion

In Experiment 1, double-blind administration of decaffeinated coffee was expected to show a placebo effect in participants if they were made to believe they drank a very strong cup of coffee. In accordance with this assumption, misinformed (false) information about the beverage produced such an effect for one of the subjective well-being factors (alertness). This effect was not due to ingredients of or psychological factors associated with drinking decaffeinated coffee, because informed participants showed no substantial differences from the natural history control group. Also, this effect was not ascribable to an artifact produced by suspicion regarding the experimental design, because the participants expected to respond to the substance in alignment with the instructions: When comparing the two treatment groups regarding their pretreatment expectations, only the falsely informed participants expected the beverage to affect them (t[28]= 4.73; p < .01).

Contrary to our hypotheses, a caffeine placebo effect only showed for subjective alertness. This effect could have been produced by factors other than expectations. Specifically, dietary caffeine use and possible withdrawal symptoms associated with it ("cravings") could have exerted an artifactual impact. For example, Garrett and Griffiths (1998; see also James, Gregg,

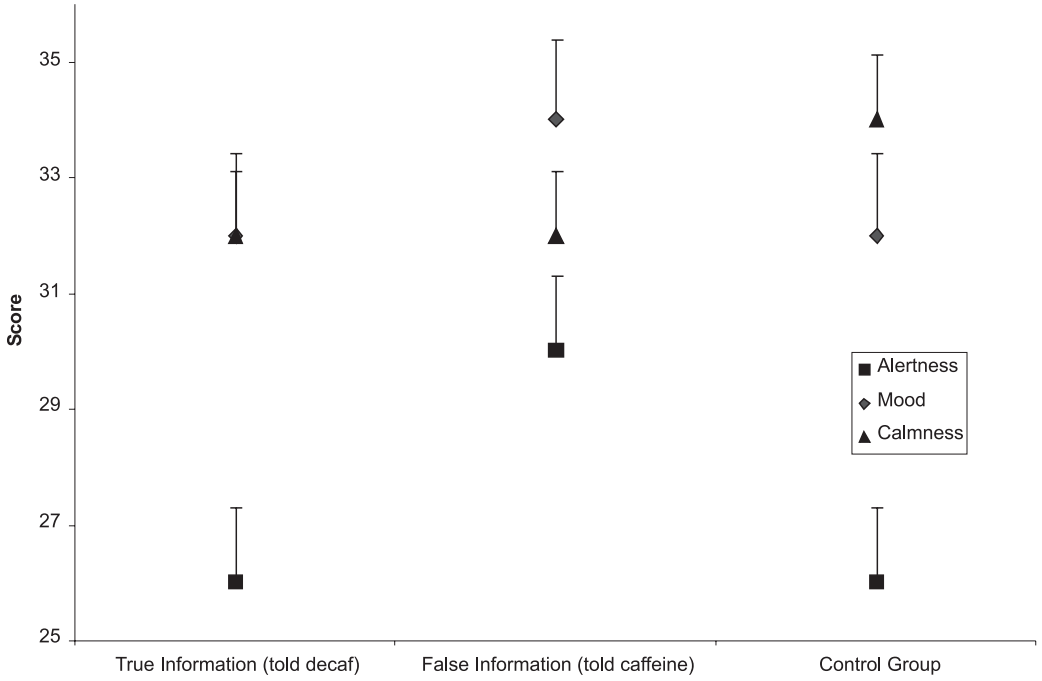

Figure 1. Mean scores (SE) for posttreatment subjective well-being corrected for baseline values in Experiment 1.

Kane, & Harte, 2005) showed that caffeine had reinforcing effects on mood disturbance and fatigue in caffeine deprived individuals. To test for "caffeine addiction," we computed a post hoc ANCOVA for alertness including coffee drinking habits (number of cups of coffee per day) and time of experiment (morning, afternoon, late afternoon) as additional covariates. The results showed that the three groups still differed with regard to perceived alertness (F[2, 39] = 3.27, p = .049). Hence, caffeine cravings as a function of coffee consumption did not qualify as an alternative explanation for the observed effect.

Despite the lack of a significant placebo effect in the cardiovascular parameters, there was a (nonsignificant) effect for diastolic blood pressure amounting to a difference of d = 0.65, which, with a larger sample size of about N = 25, would have been significant. Likewise, although subjectively more alert, falsely informed participants did not show improved cognitive performance (RT). This group, however, showed nonsignificantly faster RT after the treatment (d = –0.30). This difference was three times larger than the one found for the informed group, which showed no improvement (d = 0.09).

Experiment 2

Participants

Analogous to the sample in Experiment 1, Experiment 2 also consisted of N = 45 German adults (25 women and 20 men). Recruitment and exclusion criteria were analogous to that of Sample 1. The mean age of the sample was 26.6 years (SD = 8.4 years; range, 20 to 62 years). All participants signed informed consent prior to the start of the study.

Measures

All measures corresponded to the ones used in Experiment 1. In addition, caffeine, serving as the independent variable, was triturated (1g of quinine, 8 g of caffeine, ad 200 g lactose) and administered at a dose of 2 mg caffeine per kg body weight. This dose of caffeine has been shown to produce significant pharmacologic and psychological effects in similar studies (Flaten et al., 1999). Because caffeine has a neutral taste, the slightly bitter tasting quinine was added to reinforce the impression that the orange juice actually contained an active pharmacologic agent.

Procedure

Procedure concurred with that of Experiment 1. The experiments were conducted by a blinded female experimenter who was different from the one in Experiment 1. Also, participants were weighed to assess the amount of caffeine per body weight to be administered. Participants of the true information group (n = 15) were told that they were to consume a glass of caffeinated orange juice that was mixed by the experimenter before the eyes of the participants. They were told the dose and its approximate equivalent of one cup of coffee. The beverage was to be consumed within 1 min. Participants of the false information group (n = 15) were told that they had been assigned to the control group that was to consume orange juice, yet they ingested caffeine. Thus participants in this group were assumed to only show pharmacologic effects and no (learned) expectancy effects. In order to both avoid suspicion and enhance credibility on behalf of the experimenter, the experimenter was told that the substance she blindly mixed with the orange juice was a placebo bitter substance. In so doing, standardization of treatments was ensured, as this was general practice in pharmacologic trials. After a waiting period of 30 min necessary for caffeine to show its effect (Quinlan, Lane, & Aspinall, 1997), posttreatment measures were taken.

Statistical Analyses

All analyses conformed to the ones applied in Experiment 1. However, different hypotheses were formulated. Because pharmacologic and physiological effects were compared to test expectancy effects, the group informed about the caffeinated orange juice was assumed to show stronger effects on blood pressure, heart rate, reaction time, and subjective well-being than the group consuming blinded orange juice (Effect $_{\text{True information}}$ > Effect $_{\text{False information}}$). Furthermore, to assess the true pharmacologic effect, it was hypothesized that the group blindly administered caffeine would differ from the natural history control group, which should not show any effects.

Results

Physiological Measures

All measures were normally distributed and fell within the normal range of physiological values. Inspection of the data, however, yielded more heterogeneity for the three blood pressure measurements before and after the treatment. To correct for measurement bias, outliers (values larger than 15 mm Hg) were replaced by means. In so doing, acceptable retest reliabilities were obtained for both baseline measures (systolic blood pressure $r \geq .88$, diastolic blood pressure $r \geq .88$, heart rate $r \geq .83$) and posttreatment measures ($r \geq .86$, $r \geq .84$, $r \geq .83$). Because the analyses for both data sets did not substantially differ, we report the results for the corrected, more conservative data set.

The ANCOVA with the systolic blood pressure as the dependent variable yielded a significant effect (F[2, 41] = 5.03; p = .011). However, whereas both treatment

groups differed from the control group, the informed treatment group, which was assumed to show an expectancy effect, had lower systolic blood pressure than the falsely informed treatment group (d = –0.58). Conversely, Hypothesis 2 could be confirmed because there was a large pharmacologic effect in the falsely informed participants compared to the control group (d = 1.16). Similarly, the ANCOVA with diastolic blood pressure as the dependent variable yielded a highly significant effect (F[2, 41] = 14.42; p < .001). Again, both treatments produced significant effects compared to the control groups. Contrary to H1, however, there was no difference between the two treatment groups (d = –0.30). The pharmacologic effect found for the falsely informed treatment, alternatively, was very large (d = 1.85), confirming H2. The ANCOVA for heart rate failed to show an effect (F[2, 41] = .84; p = .44). The results for the physiological measures are depicted in Figure 2.

Reaction Time

Falsely informed participants had the slowest posttreatment reaction times (cf. Table 2). However, this effect was very small and did not reach significance (F[2, 41] = 1.69; p = .20). However, the effect for H1 was medium in size (d = –0.47), indicating that participants who were informed about the content of the orange juice tended to show an enhanced reaction time.

Subjective Well-Being

Similar to Experiment 1, participants were generally good tempered, alert, and calm (cf. Table 2) at the beginning of the experiment. The ANCOVAs revealed that neither for the dimension "mood" (F[2, 41] = .97; p = .39) nor for "alertness" (F[2, 41] = 1.44; p = .25) or "calmness" (F[2, 41] = 2.03; p = .14) were significant effects found. However, participants who drank blinded orange juice as opposed to control participants reported more positive mood (d = 0.51), increased alertness (d = 0.37), and less calmness (d = –0.42). Conversely, informed participants, who were assumed to show an expectancy effect, tended to be more alert (d = 0.25) and less calm (d = –0.32) than the falsely informed participants, yet they reported a less positive mood (d = –0.23).

Discussion

In Experiment 2, participants informed about the stimulating content of the orange juice were assumed to feel better, produce faster reaction times and show larger cardiovascular responses than those surreptitiously administered caffeine. Expectations about the drug's effect were thought to be additive to the pharmacologic effect, thereby exerting a stronger effect. Furthermore, intake of caffeine alone was expected to affect functioning and thus be different from natural history (control group).

The results failed to show additivity of pharmacologic and expectancy effect for all dependent variables measured. Even more so, for neither of the comparisons did the experimental group informed about the true content of the beverage show significantly larger effects. However, caffeine produced

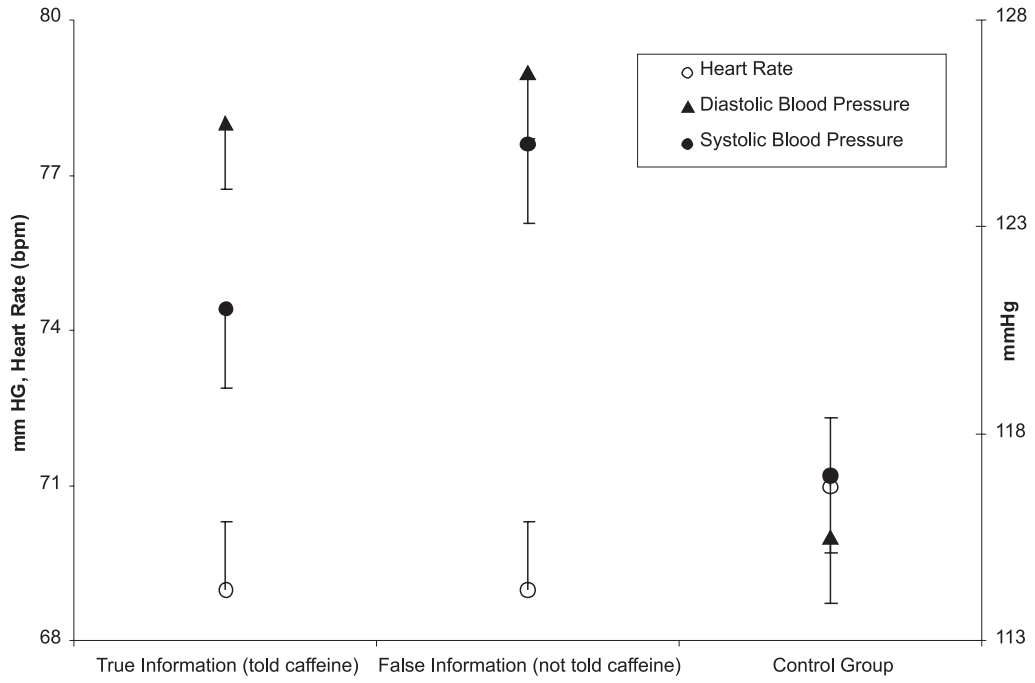

Figure 2. Mean scores (SE) for posttreatment physiological measures corrected for baseline values in Experiment 2.

Table 2. *Mean and Standard Deviation and p Values for Baseline and Posttreatment Measures of Blood Pressure (mmHg), Heart Rate (Beats per Minute), Reaction Time (ms), Mood, Alertness, and Calmness in Experiment 2*

	True					False					Control				
	Baseline		Post			Baseline		Post			Baseline		Post		
	M	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*
SBP	119.4	9.8	117.8	11.1	.502	122.1	15.8	124.7	16.1	.275	125.7	13.5	118.7	9.5	.005
DBP	73.5	8.0	75.8	9.3	.069	74.3	14.1	78.4	12.7	.015	78.1	9.7	71.9	6.4	.001
HR	80.4	9.0	70.8	8.4	< .001	74.5	14.1	66.9	11.0	.001	78.4	12.0	71.6	11.5	.001
RT	233.0	37.0	226.0	38.0	.442	228.0	31.0	235.0	37.0	.014	228.0	25.0	218.0	23.0	.050
Mood	34.1	3.9	34.5	3.3	.147	34.3	4.1	34.9	2.9	.553	35.0	2.5	34.3	4.2	.739
Alertness	27.5	6.9	29.9	7.1	.251	28.2	6.2	29.1	5.7	.879	29.5	6.6	28.1	7.6	.270
Calmness	33.2	3.9	32.1	5.4	.392	34.3	4.1	34.2	5.7	.417	31.4	4.7	34.5	3.2	.086

Note. SBP = Systolic blood pressure; DBP = Diastolic blood pressure; HR = Heart rate; RT = Reaction time. Mood, alertness and calmness have a range of 8–40; high value indicates good mood, high alertness, high calmness. $p = p$ value (two-tailed) for difference baseline–post.

large effects for blood pressure. The effects of d = 1.16 for systolic blood pressure and d = 1.85 for diastolic blood pressure in the falsely informed treatment group and d = 0.57 and d = 1.5 in the informed treatment group were larger than those found in similar studies testing caffeine with comparable doses (Quinlan et al., 1997; Zwyghuizen-Doorenbos, Roehrs, Lipschutz, Timms, & Roth, 1990). Given the rather small dose administered (approximately 150 mg on average), it may be concluded that administration of pure caffeine, dissolved in orange juice, may possess a higher bioavailability than chemically bound caffeine. Because the size of this effect was mainly attributable to the decreases of blood pressure in the control group, caffeine seemed to primarily retard relaxation effects associated with inactivity during a certain time.

The observed smaller pharmacologic effects in informed participants might have been the results of a counterregulation. For example, addictive individuals have shown to compensate for unsolicited side effects of alcohol when given information about its impairments (Fillmore, Roach, & Rice, 2002). When testing for drinking habits and withdrawal, we were, however, unable to find a significant impact on systolic blood pressure (F[2, 38] = 4.34, p = .02) and diastolic blood pressure (F[2, 38] = 12.72, p < .001). Presumably, knowledge about the stimulant may have been associated with a counterregulation of caffeine effects, as shown in the findings for RT that deteriorated after caffeine intake (cf. Table 2). Whereas intake of caffeine in informed participants showed no impairment of RT from baseline to posttreatment testing (d = –0.19) hidden intake of caffeine did (d = 0.22). Given the large pharmacologic effects in both treatment groups, it may be conjectured whether expectancy could realistically have amplified the pharmacologic effect. Conversely, expectancy appeared to have exerted a nonsignificant impact for alertness (d = 0.62) and calmness (d = –0.74) when the informed and the control groups were compared.

General Discussion

The results tentatively indicate that learning and stereotypes may amplify expectancy effects (Hirt, Lynn, Payne, Krackow, & McCrea, 1999). Individuals with a preconception about a drug respond even when the drug is depleted of the active agent. In Experiment 1, such a preconception of coffee consisted of a subjectively alerting effect. Based on the (nonsignificant) effect size found for diastolic blood pressure (d = 0.65) and the pre- to posttreatment RT improvements found only in the falsely informed individuals (d = 0.30), expectations also affected objective measures at least descriptively. Placebo effects associated with caffeine may thus be best brought about when individuals avail of some knowledge, based on learning and/or stereotypes, about the substance and its effect. The fact that this effect holds under double-blind conditions also abandons alternative explanations for caffeine placebo effects found in past studies. However, placebo effects associated with intake of coffee are varied, as indicated by our failure to find effects for an array of different parameters.

The difficulty to show expectancy effects also showed in Experiment 2 where none of the parameters confirmed our hypothesis of an amplification of pharmacologic effects. If at all, expectancy effects appear not to be entirely contingent on cognitive contents. Information about a drug's stimulating effects may enhance subjective evaluation of arousal over and above pharmacologic effects to a certain degree. The effects of expectations may not necessarily bear on a physiological basis and may even be counterdirectional. Knowledge about caffeine intake may reduce physiological reactions, although caffeine exerts distinct ef-

fects on, for example, cell metabolism, endocrinological feedback systems (insulin, adenosine, catecholamines), and cardiovascular functioning (Barone & Roberts, 1996; Keijzers, De Galan, Tack, & Smits, 2002). This may in part be due to the fact that the nature of the beverage is important in order to stimulate expectations, which, at least in part, are retrieved from past experiences. Expectancy effects (whether associated with a real stimulant or not) seem to be most directly mapped by subjective reports, as they reflect the primary conviction about drug-related changes. This should, of course, not belie the fact that none of the measures produced significant results. This may in part be due to measures applied. For example, we deployed a relatively short task to assess RT, which might not have been sensitive enough to fully map a placebo or expectancy effect. Conversely, the relatively small dose of caffeine was able to (nonsignificantly) deteriorate this rather elementary cognitive task performance. Given the shorter RTs observed in caffeine placebo participants, it may be anticipated that longer runs implying mental fatigue effects could turn out to be more appropriate in future studies. As outlined in the introduction, there is no coherent array of parameters reliably mapping expectancy effects. This, in turn, raises the question as to how future studies should be sufficiently powered. Unless this question is sufficiently answered empirically, we suggest carrying on simultaneously using several measures that allow for a broad assessment.

Conclusions

Effects produced by the caffeine placebo paradigm draws on learning history and/or prevalent stereotypes from which individuals derive expectation. Measures bearing on introspection are best suited to map placebo effects. This also holds for effects associated with expectations of pharmacologic agents. Psychological and pharmacologic effects, however, may be counterdirectional in measures that cannot be introspectively evaluated. Knowledge of the agent's effects affects its subjective experience. These findings point to the significance that psychological factors exert over and above pharmacologic effects.

References

Ader, R. (1993). Conditioned responses in pharmacotherapy research. *Psychological Medicine, 23,* 297–299.

Ader, R. (2000). The placebo effect: If it's all in your head, does that mean you only think you feel better? *Advances in Mind Body Medicine, 16,* 7–11.

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

Beecher, H. K. (1955). The powerful placebo. *Journal of the American Medical Association, 159,* 1602–1606.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Fillmore, M., Mulvihill, L. E., & Vogel-Sprott, M. (1994). The expected drug and its expected effect interact to determine placebo responses to alcohol and caffeine. *Psychopharmacology, 115,* 383–388.

Fillmore, M., Roach, E., & Rice, J. (2002). Does caffeine counteract alcohol-induced impairment? The ironic effects of expectancy. *Journal of Studies on Alcohol, 63,* 745–754.

Fillmore, M. A., & Vogel-Sprott, M. (1992). Expected effect of caffeine on motor performance predicts the type of response to placebo. *Psychopharmacology, 106,* 209–214.

Flaten, M. A., & Blumenthal, T. D. (1999). Caffeine-associated stimuli elicit conditioned responses: An experimental model of the placebo effect. *Psychopharmacology, 145,* 105–112.

Garrett, B. E., & Griffiths, R. R. (1998). Physical dependence increases the relative reinforcing effects of caffeine versus placebo. *Psychopharmacology, 139,* 195–202.

Haour, F. (2005). Mechanisms of the placebo effect and of conditioning. *Neuroimmunomodulation, 12,* 195–200.

Hirt, E. R., Lynn, S. J., Payne, D. G., Krackow, E., & McCrea, S. M. (1999). Expectancies and memory: Inferring the past from what must have been. In I. Kirsch (Ed.), *How expectancies shape experiences,* pp. 93–125. Washington, DC: American Psychological Association.

Hróbjartsson, A. (2002). What are the main methodological problems in the estimation of placebo effects? *Journal of Clinical Epistemiology, 55,* 430–435.

Hróbjartsson, A., & Gøtzsche, P. C. (2001). Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *New England Journal of Medicine, 344,* 1594–1602.

James, J. E., Gregg, M. E., Kane, M., & Harte, F. (2005). Dietary caffeine, performance and mood: Enhancing and restorative effects after controlling for withdrawal reversal. *Neuropsychology, 52,* 1–10.

Keijzers, G. B., De Galan, B. E., Tack, C. J., & Smits, P. (2002). Caffeine can decrease insulin sensitivity in humans. *Diabetes Care, 25,* 364–369.

Kirsch, I. (1999). *How expectancies shape experience.* Washington, DC: American Psychological Association.

Kirsch, I., & Weixel, L. J. (1988). Double-blind versus deceptive administration of a placebo. *Behavioral Neuroscience, 102,* 319–323.

Lotshaw, S. C., Bradley, J. R., & Brooks, L. R. (1996). Illustrating caffeine's pharmacological and expectancy effects utilizing a balanced placebo design. *Journal of Drug Education, 26,* 13–24.

Mikalsen, A., Bertelsen, B., & Flaten, M. A. (2001). Effects of caffeine, caffeine-associated stimuli, and caffeine-related information on physiological and psychological arousal. *Psychopharmacology, 157,* 373–380.

Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding meaning response. *Annals of Internal Medicine, 136,* 471–476.

Papakostas, Y. G., & Dras, M. D. (2001). Placebos, placebo effect, and the response to the healing situation: The evolution of a concept. *Epilepsia, 42,* 1614–1625.

Quinlan, P., Lane, J., & Aspinall, L. (1997). Effects of hot tea, coffee and water ingestion on physiological responses and mood: The role of caffeine, water and beverage type. *Psychopharmacology, 134,* 164–173.

Schneider, R. (in press). The psychology of the placebo effect: Exploring meaning from a functional account. *Journal of Mind and Behavior.* Manuscript accepted for publication.

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: Dissolving the expectancy versus conditioning debate. *Psychological Bulletin, 130,* 324–340.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Mehrdimensionaler Befindlichkeitsfragebogen* [Multidimensional well-being questionnaire] Göttingen: Hogrefe.

Temple, R., & Ellenberg, D. M. (2000). Placebo-controlled trials and active-control trials in the evaluation of new treatments. Part 1: Ethical and scientific issues. *Annals of Internal Medicine, 133,* 455–463.

Vase, L., Riley, J. L., & Price, D. D. (2002). A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain, 99,* 443–452.

Vase, L., Robinson, M. E., Verne, G. N., & Price, D. D. (2003). The contributions of suggestion, desire, and expectation to placebo effects in irritable bowel syndrome patients. An empirical investigation. *Pain, 105,* 17–25.

Verne, G. N., Robinson, M. E., Vase, L., & Price, D. D. (2003). Reversal of visceral and cutaneous hyperalgesia by local rectal anesthesia in irritable bowel syndrome (IBS) patients. *Pain, 105,* 223–230.

Walach, H., & Jonas, W. B. (2004). Placebo research: The evidence base for harnessing self-healing capacities. *Journal of Alternative and Complementary Medicine, 10,* 103–112.

Walach, H., Schmidt, S., Bihr, Y.-M., & Wiesch, S. (2001). The effects of a caffeine placebo and experimenter expectation on blood pressure, heart rate, well-being, and cognitive performance. *European Psychologist, 6,* 15–25.

Walach, H., Schmidt, S., Dirhold, T., & Nosch, S. (2002). The effects of a caffeine placebo and suggestion on blood pressure, heart rate, well-being and cognitive performance. *International Journal of Psychophysiology, 43,* 247–260.

Wickrameskera, I. (1980). A conditioned response model of the placebo effect: Predictions from the model. *Biofeedback and Self Regulation, 5,* 5–18.

Zimmerman, P. & Fimm, B. (1992). Testbatterie zur Aufmerksamkeitsprüfung (TAP) [Test battery for attentional performance]. Freiburg, Germany: Psychologische Testsysteme.

Zwyghuizen-Doorenbos, A., Roehrs, T. A., Lipschutz, L., Timms, V., & Roth, T. (1990). Effects of caffeine on alertness. *Psychopharmacology, 100,* 36–39.