Published in final edited form as:
Exp Clin Psychopharmacol. 2008 December ; 16(6): 521–531. doi:10.1037/a0014246.

Cigarette Abstinence Impairs Memory and Metacognition Despite Administration of 2 mg Nicotine Gum

William L. Kelemen and **Erika K. Fulton**
California State University, Long Beach

Abstract

We assessed the effects of cigarette abstinence (non-abstinent vs. minimum 8 hours abstinent) and nicotine gum (0 mg vs. 2 mg nicotine) on sustained attention, free recall, and metacognition using a within-subjects design. Moderate smokers (10 women and 22 men) received one training session followed by 4 test sessions on consecutive days. Nicotine gum improved sustained attention in both abstinent and non-abstinent states, but had no significant effect on predicted or actual recall levels. Cigarette abstinence significantly impaired free recall and reduced the magnitude of participants' predictions of their own performance. In addition, participants were significantly more overconfident about their future memory when abstinent. Thus, nicotine gum can improve smokers' performance in basic aspects of cognition (e.g., sustained attention) but may not alleviate the detrimental effects of cigarette abstinence on higher-level processes such memory and metacognition.

Keywords

Cigarette Smoking; Memory; Metacognition; Nicotine; Sustained Attention

Nicotine replacement therapies (NRTs) are clinically effective in increasing rates of smoking cessation (Hughes, Shiffman, Callas, & Zhang, 2003). Beyond cessation, researchers have identified at least eight other potential uses of NRTs, including temporary withdrawal management and smoking reduction (Shiffman, Gitchell, Warner, Slade, Henningfield, & Pinney, 2002). There is a growing controversy about whether smokers who are not actively trying to quit should be encouraged to replace their use of traditional cigarettes with NRTs, including nicotine gum, or other potential reduced-exposure products, such as less carcinogenic cigarettes, “pseudo cigarettes,” low nitrosamine smokeless tobacco, etc. (see special issue of *Nicotine & Tobacco Research*, 2002, Volume 4, Supplement 2). Compared with the large body of research on the effectiveness of NRTs for permanent smoking cessation attempts, little is known about the cognitive effects of nicotine gum used by smokers to support periods of temporary cigarette abstinence. The present study addressed this gap in the literature.

Two reviews of human studies on nicotine and cognition suggest that nicotine intake is associated with reliable cognitive improvements in abstinent smokers and smaller, inconsistent cognitive changes in non-deprived smokers and nonsmokers (Heishman, Taylor, & Henningfield, 1994; Sherwood, 1993). One of the most robust findings in the literature is that

Mailing Address: 1250 Bellflower Blvd. Long Beach, CA 90840-0901 USA Telephone: 562-985-5030 FAX: 562-985-8004 E-mail: wkelemen@csulb.edu.

nicotine improves sustained attention (i.e., vigilance) in cigarette-deprived smokers, especially by using tests of rapid visual information processing (Wesnes & Parrott, 1992). Some researchers (Pritchard & Robinson, 1998; Rezvani & Levin, 2001; Warburton, 1998) have suggested that nicotine provides an absolute increase in performance, rather than mere relief of withdrawal symptoms. The controversy may be due, in part, to the failure to include appropriate groups such as non-abstinent smokers (Hughes, 1991). Therefore, one important component of this research is the repeated testing of smokers in both abstinent and non-abstinent states to distinguish effects due to relief of withdrawal from possible net benefits of nicotine on cognition.

Nicotine's influence on human memory is complex and poorly understood. Some research on delayed memory suggests a benefit of nicotine beyond increased attention during the study phase of an experiment. For example, Colrain, Mangan, Pellett, and Bates (1992) showed that smoking after the study phase improved paired-associate recall. Rusted and Warburton and their colleagues have shown that smoking cigarettes containing nicotine improves associative learning and memory compared with nicotine-free cigarettes in a variety of memory tasks (Rusted, Graupner, Tennant, & Warburton, 1998; Rusted, Graupner, & Warburton, 1995; Warburton, Rusted, & Fowler, 1992). Like other stimulants, nicotine activates adrenergic pathways and would be expected to increase attentional performance; unlike some other stimulants, however, nicotine activates cholinergic pathways that are more specific to learning and memory (Levin, 1992). Thus, it is plausible that nicotine could increase both attention and memory in humans.

The relationship between actual versus expected memory effects is central to the concept of metacognition, which has been described as “what you know about what you know” (p. xi, Metcalfe & Shimamura, 1994; see also Dunlosky & Metcalfe, 2008; Nelson, 1992). In students, for example, poor metacognitive monitoring can contribute to poor performance overall through inappropriate study regulation (Koriat, Ma'ayan, & Ravit, 2006). Abstinent smokers using NRTs might show metacognitive biases (e.g., overconfidence) even if nicotine itself produced no measurable effect on memory. For example, low levels of oxygen experienced at high altitudes have been shown to selectively reduce metacognitive judgments with no concomitant effect on recall (Nelson et al., 1990). The reverse effect also has been shown in pharmacological studies of metacognition: alcohol intoxication can impair memory without the expected increase in overconfidence (Nelson, McSpadden, Fromme, and Marlatt, 1986), and caffeine state-dependency can influence recall with no change in the magnitude of participants' predicted performance (Kelemen & Creeley, 2003). Because Juliano and Brandon (2004) have shown that smokers' expectancies for NRTs can differ substantially from their expectancies regarding cigarettes, it is important to examine both predicted and actual recall in regard to nicotine gum in the present study.

In the present research, we tested the effects of 2 mg nicotine gum compared with a control gum (0 mg nicotine) on moderate smokers in both 8-hour abstinent and non-abstinent states using three cognitive tasks: sustained attention, delayed memory, and metacognition. We hypothesized main effects of gum type and abstinence level for each of the three dependent measures, with the largest effects predicted in the sustained attention task, which has been shown to be maximally sensitive to nicotine (Heishman et al., 1994; Sherwood, 1993). For all tasks, a significant improvement due to nicotine gum combined with no significant interaction between gum type and abstinence interval would be consistent with an absolute benefit of nicotine. In this case, a more stringent test for an absolute benefit would be conducted by comparing performance of non-abstinent smokers in the nicotine gum versus non-nicotine gum conditions. Apart from the cognitive measures, a secondary goal was to examine the effects of gum type and abstinence level on other self-reported levels of arousal, using the same analytical strategy. Finally, we sought to assess the integrity of our double-blind design using multiple

measures, which is not commonly done in cognitive research on NRTs (Mooney, White, & Hatsukami, 2004).

Method

Participants

A total of 41 participants were recruited from California State University, Long Beach, and were paid $250 each for their participation over 5 days. A bonus of $30 was paid to participants who appeared on time with no cancellations for all sessions. Volunteers completed a telephone screening to ensure their eligibility for participation. The inclusion criteria specified that participants must (a) have smoked a minimum of 10 cigarettes/day for at least one year (mean years smoking = 7.1, *SD* = 7.7), (b) not be trying to quit smoking, (c) have reported no major health concerns (from a checklist), (d) have scored a 5 or higher on the Fagerström Test of Nicotine Dependence (FTND, Heatherton, Kozlowski, Frecker, & Fagerström, 1991; mean score in the phone interview = 5.8, *SD* = 0.8), (e) not be pregnant or nursing a baby. Data from nine participants were discarded due to experimenter errors during testing, participants' non-compliance with instructions, or attrition. These participants were replaced to maintain the counterbalancing scheme. Overall, 32 participants (10 women, 22 men) provided usable data for analyses. The age range of the participants was 18 – 54 years old (*M* = 24.5 years, *SD* = 8.6).

Participants were asked to abstain from using alcohol for 24 hours before their first appointment and until the end of the 5-day study and to abstain from caffeine for 3 hours before each appointment. Participants were asked to abstain from acidic beverages for 1 hour before each appointment to prevent reduced absorption of nicotine. They were also asked to abstain from nicotine for 8 hours before two of their appointments. The mean number of hours reported since last cigarette in the abstinent condition was 13.8 (*SD* = 6.3) whereas the mean in the non-abstinent condition was 0.4 (*SD* = 0.7). To ensure nicotine deprivation, all participants completed an end tidal carbon monoxide (CO) breath analysis (Vitalograph, Lenexa, KS) upon arrival. Participants were not allowed to complete the day's tasks if their CO level exceeded 15 parts per million (ppm).

Design

A 2 × 2 within-subjects factorial design was used. The two independent variables were type of gum (0 mg vs. 2 mg nicotine) and level of abstinence (non-abstinent vs. 8 hr abstinence). Participants completed one day of testing (2 hours) to familiarize themselves with the tasks, followed by four days of testing in 1 of 4 experimental conditions. The order of conditions was randomly assigned based on participants' order of appearance and partially counterbalanced (i.e., each of the four conditions was applied across the four days equally). Participants completed all sessions at the same time of day plus or minus 1 hour in groups ranging from 1-3; most were tested individually over consecutive days for convenience. All procedures were approved by the California State University, Long Beach Institutional Review Board.

Types of Gum

Two types of gum were used: (a) 2 mg nicotine gum (i.e., *Nicorette® Fruit Chill)*, and (b) nicotine-free control gum, which was manufactured as a “taste sample” by the same company and contained capsaicin to mimic the “hot” flavor of the nicotine gum. Both types of gum were matched for size and taste; nevertheless it was necessary to wrap each in a single piece of confectionary gum to conceal the dosage stamped on the nicotine gum. The dosage of 2 mg was selected because it was recommended by the manufacturer for smokers who smoke less than 25 cigarettes per day, which accounted for 97% of our sample. The nicotine gum was

purchased commercially and the control gum was provided by GlaxoSmithKline at the request of the first author.

Questionnaires

Participants completed a total of eight different questionnaires over the course of the experiment. A 6-item version of the Fagerström Test for Nicotine Dependence (FTND; Heatherton et al., 1991), and a 12-item Cigarette Dependence Scale (CDS; Etter, Le Houezec, & Perneger, 2003) were administered on Day 1. One item was added to the FTND, which asked participants how long they had been smoking. Participants also completed a 20-item version of the Smoking Consequences Questionnaire-Adult (Copeland, Brandon, & Quinn, 1995) on Day 1, which contained the Stimulation/State Enhancement subscale, the Boredom Reduction subscale, and the Negative Reinforcement/ Negative Affect Reduction subscale. These data were not central to the hypotheses investigated and so they are not discussed further. The 20-item version of Activation-Deactivation Adjective Checklist (AD-ACL; Thayer, 1986) was used to measure participants' self-reported levels of arousal, specifically, tension and energy, two times per testing session on Days 2-5. A demographic survey, gum evaluation scale (based on Westman, Levin, & Rose 1992), post-test questionnaire, and questionnaire of adverse events, all created by the first author, also were used. The gum evaluation scale asked participants to rate the gum on 12 dimensions, whether they thought it contained nicotine or not, and how confident they were in their guess. The post-test questionnaire asked participants to rate the amount of nicotine they thought their gum contained, the effect they believed the gum had on their ability to learn the nouns, and the effect they believed the gum had on their performance in the attention task. To help ensure participants' safety, a questionnaire on adverse events was constructed that solicited ratings of possible side effects from nicotine gum on a scale from 1 to 10 (1 = not at all, 10 = severe). The reported levels of side effects were nominal, so this questionnaire is not discussed further.

Cognitive Tasks

Study and prediction task—Participants were asked to study 40 concrete, English nouns (e.g. market), presented on a computer monitor. The order of presentation was randomized for each participant, and all nouns appeared for two study trials of 3 s each. Immediately following the second trial, participants were asked to make item-by-item predictions about how likely they were to recall each noun (also known as judgments of learning; Nelson, 1992), on a scale of 0 (*definitely will not recall)* to 100 (*definitely will recall).* After studying and providing judgments for all 40 nouns, participants were asked to make a prediction of how many total words (0 – 40) they would recall (i.e., make an aggregate prediction). This task took approximately 10 minutes to complete, and the magnitude of each type of predictive judgment was analyzed.

Sustained attention task—After the study and prediction task, participants completed a 12-minute test of sustained attention using a rapid visual information processing task (Bakan, 1963; Wesnes & Warburton, 1983). Participants were shown a series of digits on the screen, one at a time, at the rate of 75 per minute. Participants were asked to press the space bar whenever 3 consecutive odd or even digits appeared, which occurred an average of 6 times per minute. Pilot testing revealed that these parameters yielded intermediate levels of performance (i.e., hit rates near 50%) so that ceiling and floor effects could be avoided. Response latencies, hit rates, and false alarm rates were assessed.

Recall task—Immediately after the sustained attention task, the participants were asked to write down as many of the 40 words they studied that day as they could remember, with no time limitation. Misspellings were counted as correct if the first three letters of the participants' response matched those of the target word. Mean numbers of words recalled were assessed.

Measures of metacognitive accuracy—Metacognitive accuracy involves the relationship between predicted and actual recall. Consistent with recommendations by Weaver and Kelemen (1997), metacognitive accuracy was assessed with two measures: Goodman-Kruskal Gamma correlations and bias indices. Gamma is a measure of relative (item-by-item) metacognitive accuracy, and it is the best measure of metacognitive resolution (Nelson, 1984). Gamma ranges from −1 (completely inaccurate metacognition) to 1 (completely accurate metacognition), with 0 indicating chance performance. Bias scores also were calculated to assess absolute monitoring accuracy (i.e., general under- vs. overconfidence), because Gamma is insensitive to overall levels of predicted and actual performance. Thus, Gamma and Bias scores measure different aspects of metacognitive accuracy. The reliability of recall, predicted recall, and bias scores tend to be higher (Spearman correlations = 0.69, 0.52, and 0.57, respectively) than the reliability of Gamma correlations (Spearman correlations = 0.07; see Kelemen, Frost, & Weaver, 2000 for more details).

Procedure

Day 1—Participants signed an informed consent form upon arrival, which stated that the purpose of the study was to "test how nicotine gum influences mood, attention, memory, and memory monitoring." Participants completed a CO breath analysis to familiarize them with the instrument. They then completed four questionnaires in the following order: demographics, FTND, CDS, and Smoking Consequences Questionnaire. Participants then received instructions for the tasks, which involved the noun study and prediction task, the sustained attention task, and the free recall task; this sequence comprised one "cycle". Participants completed three cycles with novel stimuli in each case; a 3-minute break was provided between each cycle. Data from these sessions were not analyzed; the sessions were conducted to enhance participant' familiarization with the procedures and to maximize stability of their performance on Days 2-5.

Days 2-5—Participants were asked to abstain from smoking for 8 hours before arrival on 2 of the 4 test days. CO breath analyses were conducted on all four days for consistency, but the levels were checked so that they did not exceed 15 ppm on abstinence days. Participants' mean CO level on abstinence days was 7.11 (SD = 3.0) compared with a mean CO on non-abstinence days of 18.7 (SD = 1.7). This difference was statistically significant, $t(31) = 7.11$, $p < .001$, suggesting that participants did comply with the abstinence instructions. The time of their last cigarette also was recorded. Participants who arrived having abstained when they had been instructed to smoke as usual were escorted outside to smoke a cigarette before beginning. Participants who did not abstain when instructed to do so were not allowed to participate that day. After the breath analysis, participants completed a baseline AD-ACL followed by a brief (3-minute) "warm-up" version of the sustained attention task. They then chewed a piece of gum for 20 minutes, as guided by a computer program, which instructed them when to chew and when to "park" the gum between their lower left or right gum and cheek. After 20-minutes, participants disposed of their gum and completed a gum evaluation scale. They then completed the study and prediction task, followed by the sustained attention task and free recall test. They then completed a post-task AD-ACL and another gum evaluation scale. Lastly they completed the post-test questionnaire and questionnaire of adverse events. The procedures for Days 2-5 were identical, only the type of gum and abstinence interval varied systemically on each day and different stimuli were included in the memory and attention tasks.

Results

All tests of statistical significance were conducted with an alpha level set at .05, unless otherwise noted. Effect sizes are expressed as partial eta-squared (η^2).

Sustained Attention Task

Sustained attention performance is reported first because it was hypothesized to be most sensitive to the effects of nicotine. Mean proportions of hits and false alarms were computed for each of the four conditions (see top panel of Figure 1). Data from two participants who produced greater numbers of false alarms than hits in all four conditions were omitted. For the analysis of hits, a 2 (abstinence level) × 2 (gum type) within-subjects ANOVA revealed significant main effects of nicotine level and abstinence level, $F(1,29) = 6.00$, $\eta^2 = 0.17$ and $F(1,29) = 9.66$, $\eta^2 = 0.25$, respectively. The interaction term was not statistically significant. For the analysis of false alarms, a 2 × 2 within-subjects ANOVA revealed no statistically significant main effects or interaction. Thus, the proportion of hits in the sustained attention task was significantly higher for nicotine gum compared with non-nicotine gum, and also was higher during non-abstinent conditions compared with abstinent conditions. The absence of a significant interaction implies that the main effect of nicotine represents an absolute benefit in this task rather than relief from withdrawal during 8 hours of cigarette abstinence. A planned comparison was conducted on data from non-abstinent smokers in the nicotine gum versus non-nicotine gum conditions to test for this effect directly. A one-tailed paired *t*-test confirmed a significant increase in hit rate due to nicotine, $t(29) = 1.77$, $\eta^2 = 0.10$.

The pattern of response latencies for hits mirrored that obtained for actual hit rates (see bottom panel of Figure 1). Analysis of response latencies was conducted using a 2 (abstinence level) × 2 (gum type) within-subjects ANOVA. A statistically significant main effect emerged for abstinence level $F(1,29) = 4.56$, $\eta^2 = 0.14$; the effect of nicotine level approached statistical significance, $F(1,29) = 3.22$, $p = .08$, $\eta^2 = 0.10$, and the interaction term was not statistically significant. As with the hit rates, the effect size of abstinence interval exceeded that of nicotine level, and we saw no evidence of an interaction in the response latency data. In this case, however, the planned one-tailed *t*-test on gum type using data from non-abstinent smokers did not reach statistical significance, $t(29) = -1.61$, $p = .059$.

Memory Task

Degrees of freedom vary slightly in the following analyses because some participants failed to provide a prediction of performance during one (or more) test sessions; in these cases, the data for those participants were excluded from that particular analysis only.

Predicted performance—Participants provided two types of predictions regarding their future memory performance while studying the nouns: 40 item-by-item predictions (ranging from 0-100 each) and a single aggregate prediction ranging from 0-40 at the end of the task. Both types of judgments are shown as proportions in Table 1 for consistency. For item-by-item predictions, a 2 (abstinence level) × 2 (gum type) within-subjects ANOVA revealed a significant main effect of abstinence level, $F(1,30) = 5.59$, $\eta^2 = 0.16$. Thus, cigarette abstinence produced a statistically significant reduction in the magnitude of smokers' predictions in their future memory (see top panel of Figure 2). The main effect of nicotine level and the interaction between abstinence and nicotine level were not statistically significant. Aggregate predictions of performance showed a similar pattern of means, but a 2 × 2 within-subjects ANOVA showed no statistically significant effects.

Actual performance—The third row of Table 1 lists mean recall performance across the four conditions. The results of a 2 (abstinence level) × 2 (gum type) ANOVA were the same as those obtained for item-by-item predictions. That is, we observed a significant main effect of abstinence level on memory recall, $F(1,31) = 14.63$, $\eta^2 = 0.32$ (see top panel of Figure 2). Neither the main effect of nicotine level nor the interaction between abstinence and nicotine level were statistically significant.

Metacognitive Performance

Bias scores were calculated by subtracting mean recall from mean predicted accuracy for each participant in each condition. Thus, positive bias scores indicated overconfidence, and negative bias scores indicated underconfidence (see bottom two rows of Table 1). The main effect of abstinence level was statistically significant for both item-by-item bias and aggregate bias, $F(1,30) = 9.57$, $\eta^2 = 0.24$ and $F(1,28) = 9.68$, $\eta^2 = 0.26$, respectively. In both cases, smokers in the abstinence condition were significantly more overconfident in their future memory compared with their same predictions in non-abstinent conditions. No other main effects or interactions were statistically significant.

In addition to showing differences in bias between abstinent versus non-abstinent conditions, we also examined the overall magnitude of bias compared to 0 (which would indicate no bias, or perfect calibration). This analysis tested whether participants were significantly overconfident overall. Mean item-by-item bias scores by abstinence conditions are shown in the bottom panel of Figure 2. Note that the 95% confidence interval for the abstinent condition does not include 0, indicating statistically significant bias (overconfidence) overall. In contrast, mean bias in the non-abstinent condition did not differ significantly from 0. Thus, the level of overconfidence in the abstinent condition was reliably non-zero as well as significantly higher than the bias scores in the non-abstinent condition.

Mean Gamma scores were moderate and consistent (ranging from 0.52 – 0.56) across conditions, as shown in Table 1. Mean Gammas were significantly non-zero in all four conditions using one-sample t-tests (ps < .001); these findings confirm that participants were able to reliably monitor the subsequent performance, although their accuracy was far from perfect (i.e., Gamma = + 1.0). Using a 2 (abstinence level) × 2 (gum type) ANOVA, no statistically significant differences emerged. In sum, absolute monitoring accuracy (i.e., Bias) was significantly influenced by abstinence level, but relative monitoring accuracy (i.e., Gamma) was not affected by gum type nor abstinence level.

Self-Reported Levels of Energy and Tension

Participants' levels of self-reported arousal over time were assessed using the energy and tension subscales of the AD-ACL. A 2 (abstinence level) × 2 (gum type) × 2 (time assessed) within-subjects ANOVA revealed statistically significant main effects of abstinence and time, $F(1,31) = 20.42$, $\eta^2 = 0.40$ and $F(1,31) = 11.35$, $\eta^2 = 0.27$, respectively. Specifically, self-reported energy levels were higher in the non-abstinence conditions ($M = 10.61$, $SD = 2.49$) compared with the abstinence conditions ($M = 8.80$, $SD = 2.61$), and mean energy levels were higher before the cognitive tasks ($M = 10.54$, $SD = 2.97$) compared with afterward ($M = 8.88$, $SD = 2.35$). The same three-way ANOVA was conducted on tension scores, with statistically significant main effects noted only for abstinence level, $F(1,31) = 11.02$, $\eta^2 = 0.26$. In this case, tension levels were higher in the abstinence conditions ($M = 8.30$, $SD = 2.30$ compared with the non-abstinence conditions ($M = 7.00$, $SD = 1.32$). Thus, 8 hours of cigarette abstinence produced significantly decreased levels of energy and increased levels of tension, and these effects were not alleviated by 2 mg nicotine gum.

Gum Evaluation

Table 2 shows participants' mean ratings of the gum on 12 dimensions immediately after chewing,[1] along with the outcome of inferential statistical tests. Nicotine level had a significant main effect in 7 of 12 cases (throat and chest sensations, strength, nausea, concentration, craving reduction, taste, and irritability reduction). Thus, the presence of nicotine in gum increased participants' self-reported ability to concentrate, which is consistent with the main effect of nicotine in the sustained attention task. However, this subjective report of increased concentration in the presence of nicotine did not increase participants' confidence in their own

future recall, nor the level of recall itself, during the memory task. In addition, the strong effects of nicotine on several variables that participants could potentially use to identify the type of gum (e.g., presence of throat and chest sensations) led us to closely examine the integrity of our double-blind manipulations.

Blindness Checks

Both participants and experimenters were asked to judge twice during each testing session whether the type of gum administered contained nicotine, once immediately after chewing the gum for 20 minutes and again after completing all the cognitive tasks. Participants also provided confidence ratings for their judgments. The accuracy of these judgments was examined by day to control for possible response bias (see Hughes & Krahn, 1985), and the mean accuracy is shown in Figure 3. The mean accuracy for each day was compared to 50% (chance performance) using one-sample t-tests. Participants' judgments tended to be more accurate (i.e., significantly above 50% in 5 of 8 cases) compared with experimenters' guesses (i.e., significantly above 50% in 1 of 8 cases).

The relatively high rate of participants' identification of gum type raised the possibility of a confounding variable of smokers' beliefs or expectancies regarding nicotine and cognition compared with the “pure” effects of nicotine that were the intended focus of this experiment. Although our participants were unfamiliar with nicotine gum at the start of the procedures, they all were at least moderate smokers who likely had established beliefs about the effects of nicotine in cigarettes on their cognition, which could influence their beliefs about the effects of nicotine in gum. To explore the possibility, we divided participants into two groups for a series of exploratory analyses: “accurate” raters (i.e., those participants who correctly identified the type of gum administered on all four test days, $n = 10$) and “inaccurate” raters (i.e., those participants who incorrectly identified the type of gum administered on one or more test days, $n = 22$). The patterns of means for these two groups were similar across the dependent measures (see Appendix). To formally test for differences, all previous analyses were recomputed including type of rater as a between-subjects variable. For example, the two-way (abstinence level × gum type) within subjects ANOVAs were reanalyzed as three-way (abstinence level × gum type × rater type) mixed ANOVAs. Importantly, the main effect of rater type did not approach statistical significance in any of the analyses (all ps > .25).[2] Thus, there was little discernable influence of correct identification even though the rate of participants' correct identification was higher than anticipated.

Post-test Assessment

Near the end of the experiment each day, participants answered three questions. First, they were asked “How would you rate the amount of the nicotine that your gum contained?” with alternatives ranging from −10 (labeled “None”) to +10 (labeled “Very Large”), and 0 labeled “Moderate.” Nicotine gum was rated as approximately “moderate” ($M = 0.77$, $SD = 3.84$),

[1]The gum was rated twice: immediately after chewing and at the end of the experiment. For brevity, Table 2 reports ratings obtained immediately after chewing only; ratings at the end of the experiment were very similar. The seven statistically significant effects of gum type remained so when the gum was rated at the end of the experiment; slightly reduced effect size magnitude was noted in the second ratings. Thus, participants' evaluation of gum remained stable throughout the experiment each day.
Minor differences did emerge on three other dimensions: (a) the significant main effect of abstinence level on dizziness at Time 1 did not emerge at Time 2; (b) the significant interaction between abstinence and gum type at Time 1 did not emerge at Time 2 and was replaced by a significant reduction in hunger due to nicotine gum; and (c) a statistically significant increase in awakeness due to nicotine gum emerged at Time 2.

[2]One statistically significant two-way interaction between abstinence level and rater type emerged in the analysis of hit rates for sustained attention. The pattern of means from the penultimate row of the Appendix illustrates a stronger effect of abstinence on hit rates for participants who inaccurately judged the gum compared with accurate participants. Separate two-way ANOVAs on each group of participants confirmed this effect: the main effect of abstinence was significant for inaccurate participants, $F(1, 21) = 28.31$ $\eta^2 = 0.57$ but not for accurate participants, $F(1, 9) = 0.39$. It is not evident why the effects of abstinence would be lessened in participants who knew the type of gum they received, and so this finding is not explored further.

which was significantly higher compared with the control gum ($M = -4.55$, $SD = 4.10$), $F(1, 31) = 28.24$, $\eta^2 = 0.48$. The second question was "Please rate the effect you believe the gum had on your ability to learn the 40 words" with alternatives ranging from −10 (labeled "Impaired") to +10 (labeled "Improved"), with 0 labeled "No Effect." Here, the influence of gum type was non-significant; in contrast, the mean ratings of participants in the non-abstinent condition were significantly higher ($M = 0.47$ $SD = 1.99$) compared with the abstinent condition ($M = -0.92$, $SD = 2.24$), $F(1, 31) = 9.14$, $\eta^2 = 0.23$. This pattern of data reflects participants' actual memory performance, although in this case the performance was attributed to the effects of gum. Whether participants genuinely attributed their differential memory performance across conditions to the gum or whether this finding reflects demand characteristics of the question is unclear. The final question each day was "Please rate the effect you believe the gum had on your ability to identify 3 consecutive odd or even numbers on the attention test" with alternatives ranging from −10 (labeled "Impaired") to +10 (labeled "Improved"), with 0 labeled "No Effect." Here, no statistically significant differences emerged, despite the finding that the sustained attention task yielded the largest effect sizes for nicotine gum type.

Discussion

The influence of gum type (0 vs. 2 mg nicotine) as well as abstinence level (non-abstinent vs. 8 hours of abstinence) was examined across three cognitive tasks. Overall, abstinence level played a stronger and more consistent role in this research. In the sustained attention task, abstinence reduced hit rates and increased response latencies. In the memory task, participants were less confident and recalled fewer items during conditions of abstinence. From a metacognitive perspective, the reduction in confidence did not fully reflect the magnitude of memory impairment, resulting in a net bias (i.e., overconfidence) in absolute monitoring accuracy. In other words, smokers who abstained from cigarettes for 8 hours remembered fewer items and yet remained significantly overconfident about their subsequent memory performance, which represents a double-threat to abstinent smokers using NRTs to maintain their cognitive performance.

These data are the first to document deficit in metacognitive accuracy related to cigarette abstinence. Because bias scores were computed from two separate dependent measures (i.e., predicted minus actual recall), the underlying source of this deficit is uncertain. For example, overconfidence can emerge when recall declines in the presence of stable predictions, or when recall is stable but prediction magnitude increases (see Connor, Dunlosky, & Hertzog, 1997 for examples and discussion). The former represents a memory-based change in absolute metacognitive accuracy whereas the latter represents a more "pure" metacognitive deficit. Regardless of the source, the magnitude of overconfidence (approximately 10%, see Figure 2), was large enough to warrant concern outside the laboratory. In the classroom, for example, recalling 10% fewer items than anticipated on an exam could be enough to reduce a course grade. In fact, college students with low standardized test scores can show the equivalent magnitude of overconfidence even after 5 exams, whereas students with high standardized test scores adjust their confidence more readily (Kelemen, Winningham, & Weaver, 2007).

The present data suggest that a single 2 mg dose of nicotine gum will not stave off the detrimental memory effects of prolonged cigarette abstinence, and perhaps worse, naïve users of nicotine gum may falsely believe that it will. Our study was limited, however, in that only a single, modest dose (2 mg) of nicotine gum was tested. West, Jarvis, Russell, Carruthers, & Feyerabend (1984) showed that 2 mg of nicotine gum produced only 27% of the nicotine plasma level observed with cigarette smoking. Thus, future research is needed to determine whether or not higher doses of nicotine gum would reduce overconfidence, or if smokers' performance might improve with repeated use of nicotine gum.

Nicotine gum did increase the hit rate in a sustained attention task relative to control gum in non-abstinent smokers. Moreover, the proportion of hits was significantly higher when nicotine gum was administered with no change in the rate of false alarms. Thus, nicotine did not simply cause a criterion shift in responding; rather, the increased hit rate was likely a genuine increase in sensitivity. These findings are consistent with the view that nicotine does not merely provide relief of withdrawal symptoms in sustained attention, particularly in tests of rapid visual information processing (Sherwood, 1993; see also Foulds et al., 1996; Hindmarch, Kerr, & Sherwood, 1990; Pritchard, Robinson, & Guy, 1992). It is important to note, however, that conflicting findings have emerged using different attention-based tasks (e.g., Park, Knopcik, MacGurk, & Meltzer, 2000). Moreover, the increases due to nicotine did not generalize to "higher" cognitive functions in this study, such as memory or metacognition. Overall, these data support the view that cognitive improvements due to nicotine are likely to be quite constrained (Heishman et al., 1994) and, especially in the case of smoking, outweighed in many orders of magnitude by the unequivocal negative health outcomes.

An important methodological limitation in the present experiment concerned the difficulty of maintaining a double-blind procedure using a within-subjects design over four test sessions. Figure 3 shows that participants' accuracy in guessing the type of gum they received exceeded chance in only 1 of 4 cases on the first two days of testing; by Days 3-4, however, participants' mean levels of performance were significantly above chance in 4 of 4 cases. Other experimenters have documented decreasing levels of blindness in NRT studies across multiple test sessions (e.g., from Session 1 to Session 2 in Dawkins, Powell, West, Powell, & Pickering, 2007), though relatively few studies using NRTs have actually assessed the validity of their purported double-blind procedures. Mooney et al. (2004) analyzed the reports of 73 clinical trials using double-blind studies of NRTs, and only 17 studies (i.e., less than 24% of the cases) reported the results of blindness assessments in participants. Only four studies (i.e., less than 6%) reported attempts to assess blindness in experimenters. Worse still, most reported attempts to assess blindness reported failures (i.e., 12 of 17 cases, or approximately 70%).

Why is it so difficult to maintain double-blind procedures in NRT studies? Perkins and his colleagues have shown that both smokers and non-smokers can be trained to discriminate nicotine from placebo NRTs (Perkins, DiMarco, Grobe, Scierka, & Stiller, 1994; Perkins et al., 1996; Perkins, Sanders, D'Amico, & Wilson, 1997). In the present study, nicotine gum produced large and reliable changes in 7 different dimensions that could be used to distinguish the active gum (Table 2). Thus, simply matching for taste does not guarantee blind procedures in NRT research. We used a variant of Hughes and Krahn's (1985) solution for addressing blindness failures in nicotine research and found minimal differences between accurate and inaccurate participants' performance on our major dependent measures. In the future, however, researchers may consider using an "active placebo" that produces detectable sensations (e.g., a very low-dose NRT) compared with higher doses of a drug to bolster the strength of double-blind manipulations (Mooney et al., 1984).

In conclusion, Henningfield, Fant, Buchhatter, and Stitzer (2005) have argued that NRTs are useful in treating cigarette dependence in part because they provide some positive effects previously supplied by cigarettes, including "desirable mood and attention states, making it easier to handle stressful or boring situations…" (p. 283). Such cognitive changes due to NRTs were the focus of this study, specifically in regard to sustained attention, memory, and metacognition. The results suggest that participants judged that nicotine gum improves concentration and they did show improvements in sustained attention. However, the improvements due to the nicotine gum did not generalize to higher cognitive functions such as memory and metacognition, and importantly, abstinent smokers were overconfident in their performance even after ingesting nicotine gum. Thus, acute administration of a low dose (2

mg) of nicotine gum is not a panacea for cognitive deficits due to nicotine withdrawal in abstinent smokers.

Acknowledgements

This research was funded by the National Institutes of Health, grant S06 GM063119-06S1 awarded to the first author. Portions of these data were presented in November 2007 at the 48th annual meeting of the Psychonomic Society, Long Beach, CA. The authors wish to thank Carlos Salas and Kelsey White for their assistance with data collection.

Appendix

Mean Predicted Recall, Actual Recall, Bias, and Sustained Attention Performance by Condition to Assess the Impact of Gum Identification.

	Nicotine Gum		Non-Nicotine Gum	
Dependent Measure	**Abstinent**	**Non-Abstinent**	**Abstinent**	**Non-Abstinent**
Predicted Recall				
Item-by-Item Prediction	0.43 / 0.43	0.44 / 0.44	0.41 / 0.42	0.43 / 0.44
Aggregate Prediction	0.38 / 0.34	0.41 / 0.36	0.35 / 0.36	0.38 / 0.34
Actual Memory Score	0.34 / 0.32	0.40 / 0.37	0.31 / 0.31	0.42 / 0.37
Metacognitive Accuracy				
Item-by-Item Bias	0.09 / 0.10	0.04 / 0.07	0.11 / 0.09	0.01 / 0.60
Aggregate Bias	0.04 / 0.01	0.03 / 0.00	0.04 / 0.03	−0.04 / −0.03
Sustained Attention				
Hit Rate	0.52 / 0.47	0.51 / 0.54	0.50 / 0.46	0.49 / 0.51
Response Latency (ms)	465 / 486	458 / 478	481 / 489	466 / 486

Note. Entries to the left of each slash are means for 10 participants who correctly identified the gum condition during all four test sessions; Entries to the right of each slash are means for 22 participants who incorrectly identified the type of gum during one or more test sessions.

References

Bakan, P. An analysis of retrospective reports following an auditory vigilance task. In: Buckner, D.; McGrath, JJ., editors. Vigilance: A symposium. McGraw-Hill; New York: 1963.

Colrain IM, Mangan GL, Pellett OL, Bates TC. Effects of post-learning smoking on memory consolidation. Psychopharmacology 1992;108:448–451. [PubMed: 1410157]

Copeland AL, Brandon TH, Quinn EP. The smoking consequences questionnaire-adult: Measurement of smoking outcome expectancies of experienced smokers. Psychological Assessment 1995;7:484–494.

Connor LT, Dunlosky J, Hertzog C. Age-related differences in absolute but not relative metamemory accuracy. Psychology and Aging 1997;12:50–71. [PubMed: 9100268]

Dawkins L, Powell JH, West R, Powell J, Pickering A. A double-blind placebo-controlled experimental study of nicotine: II—Effects on response inhibition and executive functioning. Psychopharmacology 2007;190:457–467. [PubMed: 17205318]

Dunlosky, J.; Metcalfe, J. Metacognition: A Textbook for Cognitive, Educational, Life Span, and Applied Psychology. Sage Publications; Thousand Oaks, CA: 2008.

Etter JF, Le Houezec J, Perneger TV. A self-administered questionnaire to measure addiction to cigarettes: The Cigarette Dependence Scale. Neuropsychopharmacology 2003;28:359–370. [PubMed: 12589389]

Foulds J, Stapleton J, Swettenham J, Bell N, McSorley K, Russell MAH. Cognitive performance effects of subcutaneous nicotine in smokers and never-smokers. Psychopharmacology 1996;127:31–38. [PubMed: 8880941]

Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for nicotine dependence: A revision of the Fagerström tolerance questionnaire. British Journal of Addiction 1991;86:1119–1127. [PubMed: 1932883]

Heishman SJ, Taylor RC, Henningfield JE. Nicotine and smoking: Review of effects on human performance. Experimental and Clinical Psychopharmacology 1994;2:345–395.

Henningfield JE, Fant RV, Buchhatter AR, Stitzer ML. Pharmacotherapy for nicotine dependence. CA: A Cancer Journal for Clinicians 2005;55:281–299. [PubMed: 16166074]

Hindmarch I, Kerr JS, Sherwood N. Effects of nicotine gum on psychomotor performance in smokers and non-smokers. Psychopharmacology 1990;100:535–541. [PubMed: 2320715]

Hughes JR. Distinguishing withdrawal relief and direct effects of smoking. Psychopharmacology 1991;104:409–410. [PubMed: 1924648]

Hughes JR, Krahn D. Blindness and the validity of the double-blind procedure. Journal of Clinical Psychopharmacology 1985;5:138–142. [PubMed: 3998203]

Hughes JR, Shiffman S, Callas P, Zhang J. A meta-analysis of the efficacy of over-the-counter nicotine replacement. Tobacco Control 2003;12:21–27. [PubMed: 12612357]

Juliano LM, Brandon TH. Smokers' expectancies for nicotine replacement therapy vs. cigarettes. Nicotine & Tobacco Research 2004;6:569–574. [PubMed: 15203790]

Kelemen WL, Creeley CE. State-dependent memory effects using caffeine and placebo do not extend to metamemory. Journal of General Psychology 2003;130:70–86. [PubMed: 12635857]

Kelemen WL, Frost PJ, Weaver CA III. Individual differences in metacognition: Evidence against a general metacognitive ability. Memory & Cognition 2000;28:92–107.

Kelemen WL, Winningham RG, Weaver CA III. Repeated testing sessions and scholastic aptitude in college students' metacognitive accuracy. European Journal of Cognitive Psychology 2007;19:689–717.

Koriat A, Ma'ayan H, Ravit N. The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. Journal of Experimental Psychology: General 2006;135:36–69. [PubMed: 16478315]

Levin ED. Nicotinic systems and cognitive functions. Psychopharmacology 1992;108:417–531. [PubMed: 1357713]

Masson MEJ, Loftus GR. Using confidence intervals for graphically based data interpretation. Canadian Journal of Experimental Psychology 2003;57:203–220. [PubMed: 14596478]

Metcalfe, J.; Shimamura, AP. Metacognition: Knowing about knowing. MIT Press; Cambridge, MA: 1994.

Mooney M, White T, Hatsukami D. The blind spot in the nicotine replacement therapy literature: Assessment of the double-blind in clinical trials. Addictive Behaviors 2004;29:673–684. [PubMed: 15135549]

Nelson TO. A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychological Bulletin 1984;95:109–133. [PubMed: 6544431]

Nelson, TO. Metacognition: Core readings. Allyn & Bacon; Boston: 1992.

Nelson TO, Dunlosky J, White DM, Steinberg J, Townes BD, Anderson D. Cognition and metacognition at extreme altitudes on Mount Everest. Journal of Experimental Psychology: General 1990;119:367–374. [PubMed: 2148573]

Nelson TO, McSpadden M, Fromme K, Marlatt GA. Effects of alcohol intoxication on metamemory and on retrieval from long-term memory. Journal of Experimental Psychology: General 1986;115:247–254. [PubMed: 2944987]

Park S, Knopcik C, MacGurk S, Meltzer H. Nicotine impairs spatial working memory while leaving spatial attention intact. Neuropsychopharmacology 2000;22:200–209. [PubMed: 10649832]

Perkins KA, D'Amico D, Sanders M, Grobe JE, Wilson A, Stiller RL. Influence of training dose on nicotine discrimination in humans. Psychopharmacology 1996;126:132–139. [PubMed: 8856832]

Perkins KA, DiMarco A, Grobe JE, Scierka A, Stiller RL. Nicotine discrimination in male and female smokers. Psychopharmacology 1994;116:407–413. [PubMed: 7701041]

Perkins KA, Sanders M, D'Amico D, Wilson A. Nicotine discrimination and self-administration in humans as a function of smoking status. Psychopharmacology 1997;131:361–370. [PubMed: 9226738]

Pritchard, WS.; Robinson, JH. Effects of nicotine on human performance. In: Snel, J.; Lorist, MM., editors. Nicotine, Caffeine, and Social Drinking: Behaviour and Brain Function. Harwood Academic Press; Amsterdam: 1998. p. 21-82.

Pritchard WS, Robinson JH, Guy TD. Enhancement of continuous performance task reaction time by smoking in non-deprived smokers. Psychopharmacology 1992;108:437–442. [PubMed: 1410155]

Rezwani AH, Levin ED. Cognitive effects of nicotine. Biological Psychiatry 2001;49:258–267. [PubMed: 11230877]

Rusted J, Graupner L, Tennant A, Warburton D. Effortful processing is a requirement for nicotine-induced improvements in memory. Psychopharmacology 1998;138:362–368. [PubMed: 9725759]

Rusted J, Graupner L, Warburton D. Effects of post-trial administration of nicotine on human memory: Evaluating the conditions for improving learning. Psychopharmacology 1995;119:405–413. [PubMed: 7480520]

Sherwood N. Effects of nicotine on human psychomotor performance. Human Psychopharmacology: Clinical and Experimental 1993;8:155–184.

Shiffman S, Gitchell JG, Warner KE, Slade J, Henningfield JE, Pinney JM. Tobacco harm reduction: Conceptual structure and nomenclature for analysis and research. Nicotine & Tobacco Research 2002;4:S113–S129. [PubMed: 12573173]

Thayer RE. Activation-Deactivation Adjective Check List (AD ACL): Current overview and structural analysis. Psychological Reports 1986;58:607–614.

Warburton, DM. Caffeine, nicotine, and attentional performance. In: Snel, J.; Lorist, MM., editors. Nicotine, Caffeine, and Social Drinking: Behaviour and Brain Function. Harwood Academic Press; Amsterdam: 1998. p. 245-256.

Warburton DM, Rusted JM, Fowler J. A comparison of the attentional and consolidation hypotheses for the facilitation of memory by nicotine. Psychopharmacology 1992;108:443–447. [PubMed: 1410156]

Weaver CA III, Kelemen WL. Judgments of learning at delays: Shifts in response patterns or increased metamemory accuracy? Psychological Science 1997;8:318–321.

Wesnes, KA.; Parrott, AC. Smoking, nicotine, and human performance. In: Smith, AP.; Jones, DM., editors. Handbook of Human Performance: Health and Performance (Volume 2). Academic Press; San Diego: 1992. p. 127-168.

Wesnes KA, Warburton DM. Effects of smoking on rapid information processing performance. Neuropsychobiology 1983;9:223–229. [PubMed: 6646394]

West RJ, Jarvis M, Russell MAH, Carruthers M, Feyerabend C. Effect of nicotine replacement on the cigarette withdrawal syndrome. British Journal of Addiction 1984;79:215–219. [PubMed: 6589006]

Westman E, Levin E, Rose J. Smoking while wearing the nicotine patch: Is smoking satisfying or harmful? Clinical Research 1992;40:871A.

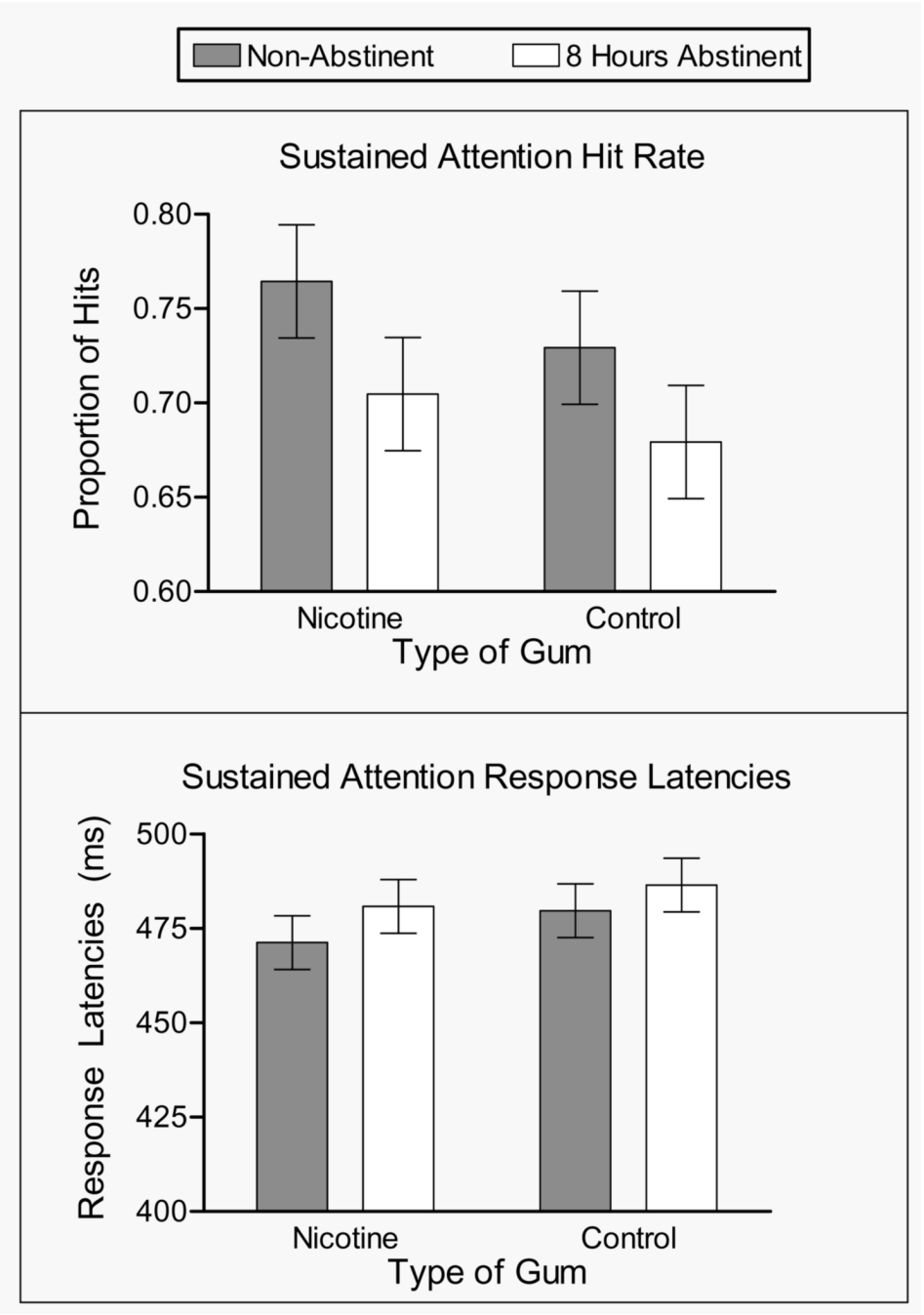

Figure 1.
Top panel: Mean proportion of hits by condition in the sustained attention task by condition. Bottom panel: Mean response latencies (in milliseconds) by condition in the sustained attention task. For both panels, error bars are 95% confidence intervals using a pooled *MS* error term for within-subjects designs (see pp. 211-212 of Masson & Loftus, 2003).

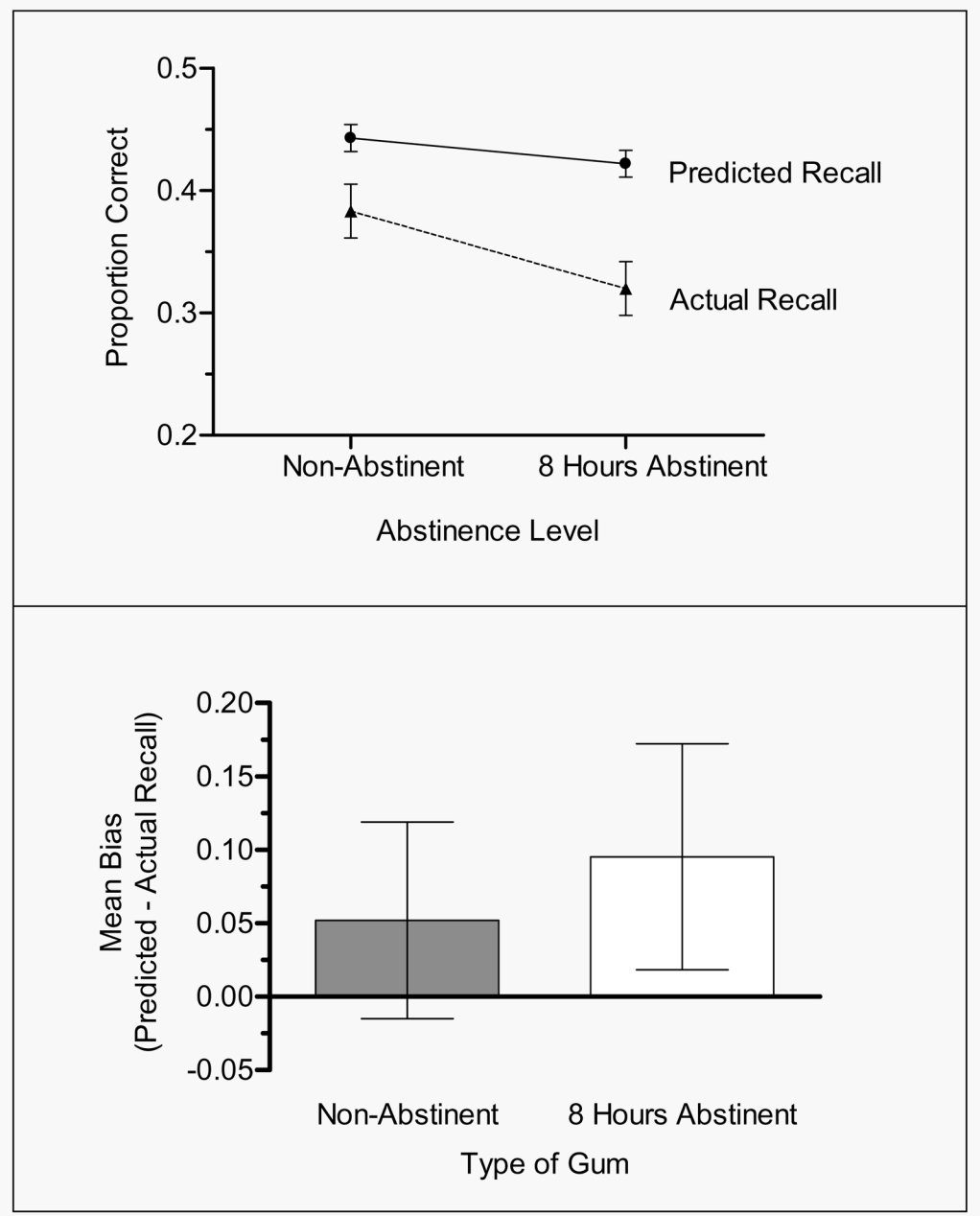

Figure 2.
Top panel: Mean predicted and actual recall performance by abstinence condition. Error bars are 95% confidence intervals using a pooled *MS* error term for within-subjects designs. Bottom panel: Bias scores (i.e., predicted minus actual recall) by abstinence condition. Error bars are standard (between subjects) 95% confidence intervals to facilitate comparison with zero.

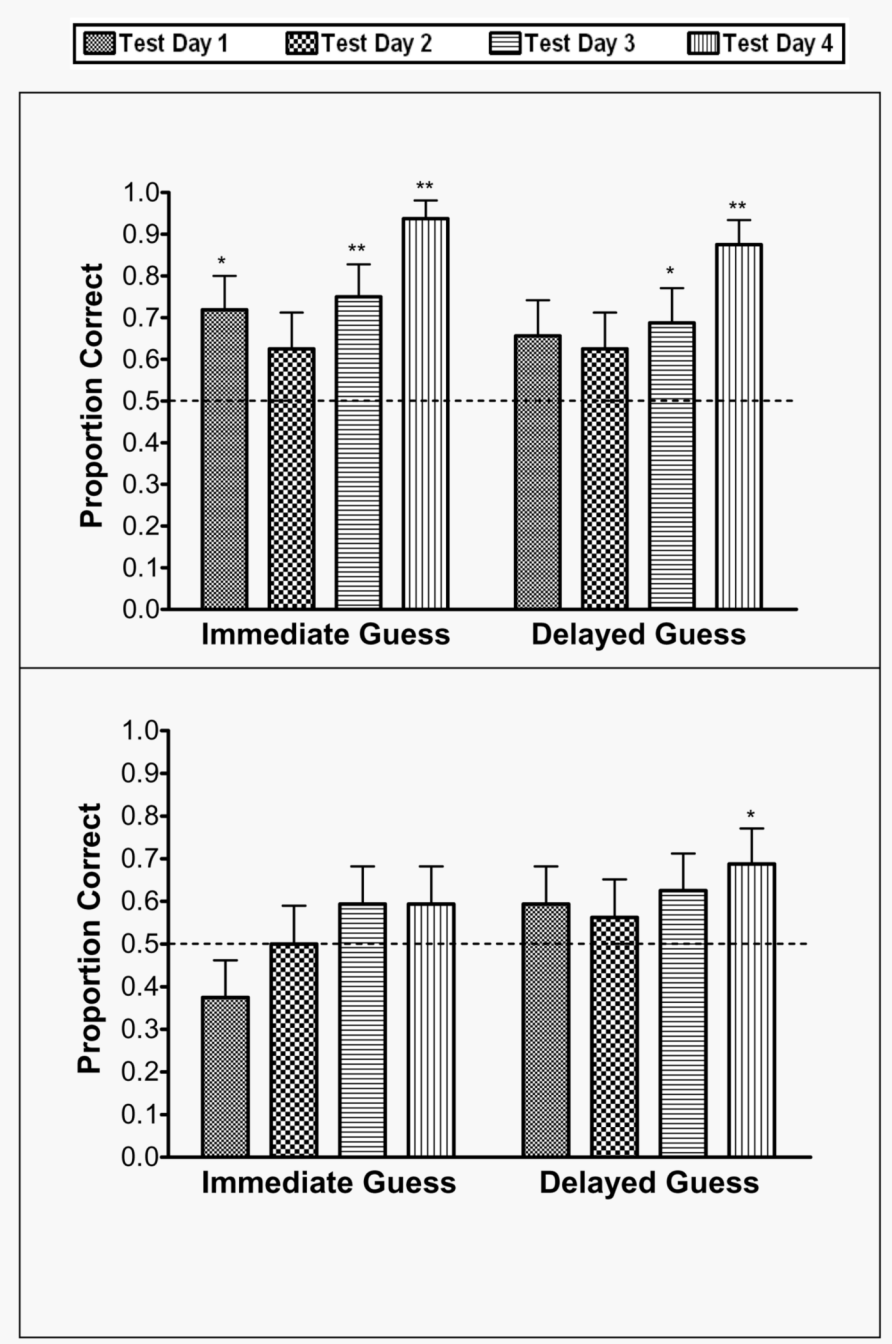

Figure 3.

Top panel: Participants' mean proportion of responses correct in regard to the type of gum administered by test day. Bottom panel: Experimenters' mean proportion of responses correct in regard to the type of gum administered by test day. Error bars are between-subjects standard errors of the mean in both panels. Asterisks indicate a statistically significant difference from 50% using a one-sample *t* test.

Table 1

Mean Predicted Recall, Actual Recall, and Metacognitive Accuracy.

	Nicotine Gum		Non-Nicotine Gum	
Dependent Measure	**Abstinent**	**Non-Abstinent**	**Abstinent**	**Non-Abstinent**
Predicted Recall				
Item-by-Item Prediction*	0.43 (0.03)	0.45 (0.03)	0.41 (0.03)	0.44 (0.03)
Aggregate Prediction	0.36 (0.02)	0.38 (0.03)	0.35 (0.02)	0.36 (0.02)
Actual Memory Score*	0.33 (0.02)	0.38 (0.03)	0.31 (0.02)	0.39 (0.02)
Metacognitive Accuracy				
Item-by-Item Bias*	0.10 (0.04)	0.06 (0.04)	0.09 (0.04)	0.04 (0.03)
Aggregate Bias*	0.02 (0.02)	0.00 (0.03)	0.04 (0.03)	−0.03 (0.02)
Gamma	0.52 (0.05)	0.53 (0.04)	0.56 (0.04)	0.56 (0.04)

Note. Main entries are mean proportions; entries in parentheses are standard errors of the mean.

* indicates a significant main effect of cigarette abstinence level on that variable using a 2 (gum type) × 2 (abstinence level) within-subjects ANOVA, $p < .05$.

Exp Clin Psychopharmacol. Author manuscript; available in PMC 2009 December 1.

Table 2

Mean Ratings and F Ratios by Condition for the Gum Evaluation Scale Immediately after Chewing.

	Nicotine Gum		Non-Nicotine Gum		Effect		
Dimension	**Abstinent**	**Non-Abstinen**	**t Abstinent**	**Non-Abstinent**	**Gum Type**	**Abstinence**	**Gum Type × Abstinence**
Gave me throat and chest sensations.	4.78 (0.34)	4.41 (0.36)	1.55 (0.18)	1.63 (0.22)	107.60***	-----	-----
Was very strong	4.41 (0.25)	4.00 (0.28)	2.44 (0.29)	2.34 (0.29)	53.10***	-----	-----
Made me feel nauseous	2.69 (0.32)	2.19 (0.28)	1.44 (0.17)	1.47 (0.20)	19.32***	-----	-----
Helped me concentrate	3.48 (0.26)	3.63 (0.23)	2.31 (0.26)	2.91 (0.29)	16.50***	-----	-----
Reduced my cravings	3.56 (0.31)	3.66 (0.31)	2.84 (0.33)	2.81 (0.29)	12.00**	-----	-----
Tasted good	3.81 (029)	4.19 (0.27)	4.66 (0.28)	4.91 (0.26)	10.92**	-----	-----
Reduced my irritability	3.44 (0.30)	3.53 (0.31)	2.56 (0.30)	3.16 (0.29)	5.13*	-----	-----
Made me dizzy	2.47 (0.31)	1.88 (0.20)	1.94 (0.28)	1.53 (0.21)	-----	6.44*	-----
Reduced my hunger	3.16 (0.29)	3.91 (0.30)	3.41 (0.35)	3.03 (0.31)	-----	-----	8.18**
Made me feel awake	3.25 (0.33)	3.22 (0.27)	2.50 (0.27)	2.97 (0.28)	-----	-----	-----
Was calming	3.81 (0.31)	4.09 (0.27)	3.22 (0.31)	3.72 (0.31)	-----	-----	-----
Was satisfying	3.78 (0.29)	3.81 (0.30)	3.91 (0.29)	4.25 (0.28)	-----	-----	-----

Note. Main entries under "Condition" are means; entries in parentheses under "Condition" are standard errors of the mean. Main entries under "Effect" reflect the outcomes of statistical comparisons using two-way (Nicotine × Abstinence) ANOVAs; only significant *F* ratios are provided (all effects sizes were $\eta^2 > 0.14$).

* $p < .05$.

** $p < .01$.

*** $p < .001$.