Bi-directional effect of increasing doses of baclofen on reinforcement learning

Jean Terrier[†], Andres Ort[†], Cédric Yvon, Arnaud Saj, Patrik Vuilleumier and Christian Lüscher*

Department of Basic Neurosciences, Medical Faculty, University of Geneva, Geneva, Switzerland

In rodents as well as in humans, efficient reinforcement learning depends on dopamine (DA) released from ventral tegmental area (VTA) neurons. It has been shown that in brain slices of mice, GABA$_B$-receptor agonists at low concentrations increase the firing frequency of VTA–DA neurons, while high concentrations reduce the firing frequency. It remains however elusive whether baclofen can modulate reinforcement learning in humans. Here, in a double-blind study in 34 healthy human volunteers, we tested the effects of a low and a high concentration of oral baclofen, a high affinity GABA$_B$-receptor agonist, in a gambling task associated with monetary reward. A low (20 mg) dose of baclofen increased the efficiency of reward-associated learning but had no effect on the avoidance of monetary loss. A high (50 mg) dose of baclofen on the other hand did not affect the learning curve. At the end of the task, subjects who received 20 mg baclofen p.o. were more accurate in choosing the symbol linked to the highest probability of earning money compared to the control group (89.55 ± 1.39 vs. $81.07 \pm 1.55\%$, $p = 0.002$). Our results support a model where baclofen, at low concentrations, causes a disinhibition of DA neurons, increases DA levels and thus facilitates reinforcement learning.

Keywords: **instrumental learning, mesolimbic dopamine system, reward-prediction error, baclofen, bi-directional effect, ventral tegmental area, anti-craving treatment**

INTRODUCTION

In his paper on "The Law of Effect," *Thorndike* stipulated that: "of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur" (Thorndike, 1898). Since then, it has been suggested that the mesolimbic dopamine (DA) system is involved in this learning by coding for a "reward-prediction error" (Schultz et al., 1997). The mesocorticolimbic DA system originates in the ventral tegmental area (VTA), which projects to the nucleus accumbens (NAc) and the prefrontal cortex. Under physiologic conditions, mesocorticolimbic projections release DA in response to natural rewards such as food and sex, which are critical for the survival of the species. This process reflects the fact that it is important for an organism to learn the circumstances under which rewards are obtained (Balland and Lüscher, 2008). When an external reward is delivered, DA neurons elicit a strong learning signal indicating whether the value of the current state is better or worse than predicted (Schultz et al., 1997), rather than euphoria or pleasure (Balland and Lüscher, 2008). This signal therefore allows rapid acquisition of predictive cues and efficient behaviors that are successful in obtaining rewards (Bechara et al., 1998).

Evidence that this system can be pharmacologically modulated by changes in DA function has been provided by Pessiglione et al. (2006). In their study, human volunteers carried out a learning task that involved money gains and losses while functional magnetic resonance images (fMRI) were collected. When mesocorticolimbic DA was boosted by L-DOPA, the participants learned faster and earned more money. Conversely, when DA signaling was inhibited

by haloperidol, participants learned slower and earned less money compared to the control group. Interestingly, no shift of the learning curves was observed when participants were in the loss condition, which suggests that other processes are involved in aversive learning. In a separate study using the Iowa gambling task, an activation of the ventral striatum has also been shown by fMRI (Li et al., 2010).

The effect of DA on learning can be explained by a modulation of the mesocorticolimbic system of circuits involved in action planning and decision-making. In many mammals, at least two systems exist to predict the value of an action: the planning or explicit system, which takes a given situation, predicts an outcome and evaluates that outcome; and the habit or implicit system, which takes a given situation and identifies the best remembered action to take (Redish et al., 2008). The flexible planning system involves the ventral and dorsomedial striatum, the prelimbic medial prefrontal cortex and the orbitofrontal cortex, as well as the entorhinal cortex and hippocampus, with an involvement of DA inputs from the VTA. The habit system involves the dorsolateral striatum, the infralimbic medial prefrontal cortex as well as the parietal cortex, with an involvement of DA inputs from the pars compacta of the substantia nigra (SNc; Redish et al., 2008). The mesocorticolimbic system thus has a central role in evaluating the value of predicted outcomes during decision-making and planning. An over-evaluation of a predicted value by the DA system might alter the decision system leading to addictive behaviors (Redish et al., 2008). Another mechanism leading to automatic decision-making and even addiction could be the recruitment of the habit system by the NAc via feedback loops to the dorsal striatum (Koob and Volkow, 2010). Understanding how modulation of DA can alter valuation and decision-processing therefore has profound implication for understanding motivated

behaviors and addiction. Here we propose to pharmacologically modulate DA release with the GABA$_B$-receptor agonist baclofen and observe the effect of this modulation on an instrumental learning task.

Baclofen (*p*-chlorophenyl-GABA) acts as a high affinity γ-aminobutyric acid type B (GABA$_B$) receptor agonist. Its primary action as spasmolytic agent is mediated by increasing K$^+$ conductance that results in postsynaptic inhibition (Cruz et al., 2004; Katzung, 2009). In addition, baclofen causes presynaptic inhibition by reducing Ca^{2+} influx and the release probability of excitatory transmitters in the brain and spinal cord (Katzung, 2009). Interestingly, baclofen may also modulate DA release in the mesocorticolimbic system by targeting VTA neurons (Lomazzi et al., 2008). A recent model proposed by Cruz et al. (2004) shows a bi-directional control of DA activity by increasing doses of baclofen. In this model, low-dose baclofen preferentially inhibits γ-aminobutyric acid (GABA) neurons, which control in part DA neuron activity, causing a disinhibition of DA neurons. Conversely, high dose baclofen inhibits the firing of DA neurons, resulting in a decrease of transmitter release to the NAc in the ventral striatum. A possible explanation for this phenomenon is based on a different coupling efficiency between GABA$_B$-receptors, G-proteins, regulator of G-proteins signaling (RGS) proteins, and G-protein-gated inwardly rectifying potassium channels (GIRK/Kir3), forming a macromolecular signaling complex (Lomazzi et al., 2008). Indeed, it has been shown that the concentration that produces 50% of the maximal effect (EC$_{50}$) of baclofen is one order of magnitude lower in GABA neurons than in DA neurons. Therefore, low doses of baclofen preferentially inhibit GABA neurons activity (Cruz et al., 2004; Labouèbe et al., 2007).

In this study, the focus was on the question whether the prediction error of reward signals (i.e., DA neurons firing) in healthy human subjects can be modulated by increasing doses of baclofen. We predicted that low-dose baclofen would disinhibit DA neurons, eventually increase DA release, and thus make the behavioral instrumental learning process more efficient. Conversely, high dose would inhibit DA neurons and therefore reduce the learning rate.

MATERIALS AND METHODS
EXPERIMENTAL PROCEDURE
The local ethics committee and Swissmedic approved the study (CER 07-074, NAC 07-029, Swissmed: 2008DR2044). The present experiments constituted a randomized, double-blind, placebo-controlled study using low- and high-dose of baclofen. Informed consent was obtained from all subjects.
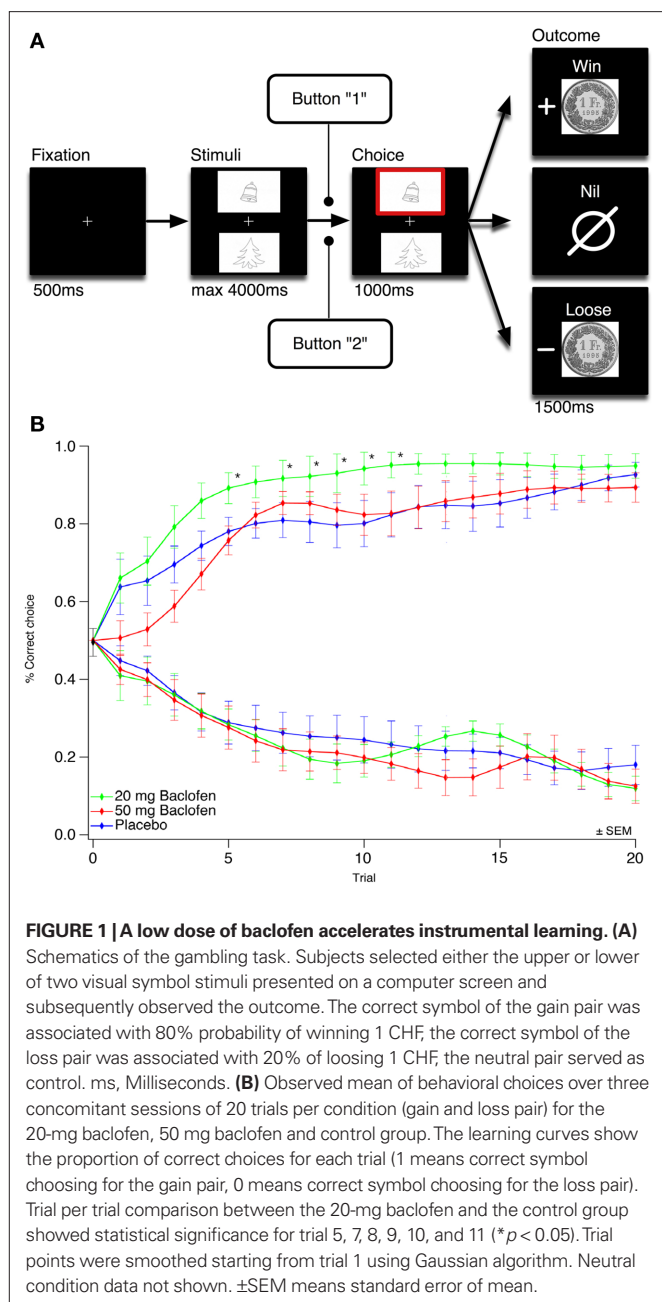
A total of 36 healthy male subjects were recruited at the University of Geneva. Exclusion criteria were an age below 18 or above 35 years, weight below 60 or above 90 kg, regular consumption of drugs or medications, regular gambling (≥1 time/week, i.e., Casino, lottery, poker), and history of psychiatric or neurological disease. These 36 subjects were randomly split into three groups: 12 subjects received 20 mg of baclofen, 12 subjects received 50 mg of baclofen, and 12 subjects received a placebo. The 50-mg group took 10 mg on the first day and progressively increased the dosage by 10 mg every day to reach 50 mg at day 5. The 20-mg group took the placebo during the first 3 days, took the first 10 mg of baclofen

on day 4 and 20 mg on day 5. All groups had to take the last dose 1 h prior to the instrumental learning task. All subjects had to take the same number of compounds over 5 days. Over the 5-days, subjects were asked to report their degree of alertness using the Stanford Sleepiness Scale (SSS) in order to assess for adverse effects at different time points each day (10.00 am–1.00 pm–4.00 pm–7.00 pm) and 30 min before the instrumental test (Hoddes et al., 1973). The SSS is a rating scale measuring current level of subjective sleepiness. It consists of seven statements describing different levels of current sleepiness ranging from "feeling active and vital" (1 point) to "almost in reverie" (7 points). In addition, subjects underwent an auditory digit span test (in order and inversed order), to assess for attention and vigilance just prior to the test.

Baclofen is usually used for its spasmolytic effects at a dosage between 30 and 80 mg/day. Twenty milligram of baclofen therefore represent a low-dose (0.3 mg/kg p.o. respectively for a 70-kg weighted person). At this dose, the predicted plasmatic concentration is about 360 ng/ml after 0.5–1.5 h (=1.60 μM, baclofen MW = 213 g/mol; Compendium Suisse des Medicaments, 2011). In the CSF, the expected value is about 8.5 times lower than in the plasma, which corresponds to 42 ng/ml for the 20-mg dose (nearly 0.20 μM; Compendium Suisse des Medicaments, 2011). These doses theoretically correspond to the concentration activating VTA DA neurons *in vitro* (Cruz et al., 2004). For 50 mg baclofen, the predicted plasmatic concentration is about 900 ng/ml after 0.5–1.5 h (=4.2 μM, baclofen MW = 213 g/mol; Compendium Suisse des Medicaments, 2011). In the CSF, the expected value corresponds to 106 ng/ml for the 50-mg dose (nearly 0.50 μM; Compendium Suisse des Medicaments, 2011). These doses theoretically correspond to the concentration starting to inhibit the VTA DA neurons *in vitro* (Cruz et al., 2004). Plasma elimination half-life of baclofen is situated between 3 and 4 h (Compendium Suisse des Medicaments, 2011).

The subjects had to perform one first practice session in order to become acquainted with the instrumental learning task and three subsequent experimental sessions of the same task adapted from Pessiglione et al. (2006). Each session proposed three new pairs of visual stimuli. Each of the pairs of stimuli (three conditions: "win" to assess the effects of baclofen on the ability of reward learning, "loose" to learn from punishment, and "neutral" as a control) was associated with three pairs of outcomes ("win" +1 CHF/nil, "loose" −1 CHF/nil, "neutral" nil/nil), the two stimuli corresponding to reciprocal probabilities (0.8/0.2 and 0.2/0.8). The neutral pair was nil whatever the stimulus chosen.

The three conditions were randomly presented during each run and the relative position of the visual stimuli was counterbalanced across trials. Each item in the pairs belonged to the same semantic field (e.g., animals, current life objects, transport vehicles etc.), in order not to influence choices by stimulus meaning. In each of the four test sessions, subjects first viewed the cues above and under a central fixation cross (4 s), then indicated their choice by pressing a button on a separate keyboard that led to the appearance of a red frame around the chosen item (1 s), and finally viewed the outcome (1.5 s). A total of 60 trials were administered per session (20 trials per condition, trial 0 was calculated as 0.495 ± 0.036 as the mean and SEM across all starting points of the subjects). One session lasted about 8 min each (**Figure 1**).

FIGURE 1 | A low dose of baclofen accelerates instrumental learning. (A) Schematics of the gambling task. Subjects selected either the upper or lower of two visual symbol stimuli presented on a computer screen and subsequently observed the outcome. The correct symbol of the gain pair was associated with 80% probability of winning 1 CHF, the correct symbol of the loss pair was associated with 20% of loosing 1 CHF, the neutral pair served as control. ms, Milliseconds. **(B)** Observed mean of behavioral choices over three concomitant sessions of 20 trials per condition (gain and loss pair) for the 20-mg baclofen, 50 mg baclofen and control group. The learning curves show the proportion of correct choices for each trial (1 means correct symbol choosing for the gain pair, 0 means correct symbol choosing for the loss pair). Trial per trial comparison between the 20-mg baclofen and the control group showed statistical significance for trial 5, 7, 8, 9, 10, and 11 (*$p < 0.05$). Trial points were smoothed starting from trial 1 using Gaussian algorithm. Neutral condition data not shown. ±SEM means standard error of mean.

To earn money, the subjects had to learn, by trial and error, the stimulus–outcome associations. Subjects were instructed to maximize their earnings. Each subject received the total amount earned during the three experimental sessions. The task was coded using the software e-PRIME.

DATA AND STATISTICAL ANALYSIS

All statistical data were obtained from the three experimental sessions (3×60 trials) and calculated for each condition (gain pair, loss pair, and neutral pair, each 3×20 trials) and group (control, 20 mg baclofen, 50 mg baclofen). The overall mean proportion of correct choices and money gain were calculated for each participant. Learning curves (proportion of correct choices across trials)

in each individual were smoothed with a Gaussian filter (smoothing: 3) using IGOR data analysis software (Wavemetrics, version 6.12A), and an average curve was then obtained for each group in each condition. An ANOVA test and Dunnett's test for comparisons of the drug groups with the placebo group were used to analyze the amount of money gain. For asymmetrically distributed values (overall mean proportion of correct choices, proportion of correct choices trial by trial, SSS scores), we used a Kruskal–Wallis test and *post hoc* comparisons by Mann–Whitney test. All statistical tests were performed with SPSS 17.0.
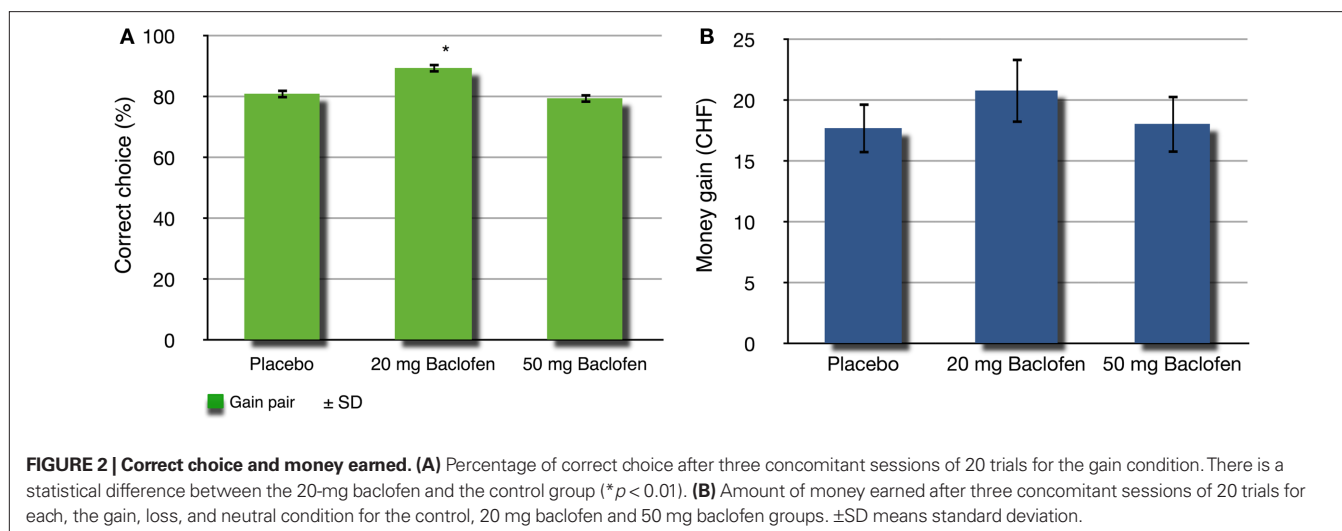
RESULTS

Data from the behavioral learning task was obtained from 34 individuals (11 each in the control and 20 mg baclofen group, 12 in the 50-mg baclofen group). Out of the initial 36 participants, one was excluded in the control group because of discontinuation of substance taking for unknown reasons; and another was excluded in the 20-mg group because of a history of psychiatric disease and addictive behavior not previously detected during the initial examination. All participants had reached a university level and had a mean age between 24 and 27 years. We specifically monitored for the presence of tiredness as a potential side effect. No significant differences were observed on the SSS between the three groups during the whole week and just prior to the task (30 min before the test: placebo group $1.7 \pm 0.0.65$ SD, 20 mg group 1.8 ± 0.61 SD, 50 mg 2.1 ± 0.67 SD points on the SSS). All subjects were able to repeat a five numbers digit span in order and four numbers in reverse order.

In the experimental gambling task, we observed a significant difference in choosing the correct stimulus between the groups, specifically for the gain condition (Kruskal–Wallis test, $\chi^2 = 14.56$, df = 2, $p = 0.001$) but not for the loss (Kruskal–Wallis test, $\chi^2 = 5.38$, df = 2, $p = 0.68$) nor the neutral condition (Kruskal–Wallis test, $\chi^2 = 1.57$, df = 2, $p = 0.45$). Comparisons with placebo showed a significant difference for the 20-mg baclofen group for the gain (89.55 ± 1.39 vs. $81.07 \pm 1.55\%$, $p = 0.002$ with Mann–Whitney test) but not for the 50-mg group (79.59 ± 1.63 vs. $81.07 \pm 1.55\%$, $p = 0.734$ with Mann–Whitney test, **Figure 2**). These results indicate that subjects in the 20-mg baclofen group more often chose the correct symbol associated with the highest probability of earning money (gain pair).

Accordingly, the 20-mg baclofen group earned the highest amount of money overall (20.82 ± 2.67 CHF), whereas the 50-mg baclofen group received 18.08 ± 2.39 CHF and the placebo group 17.73 ± 2.08 CHF (**Figure 2**). However, the difference for the monetary gain did not receive significance, in contrast to the proportion of correct choice [ANOVA, $F(2,31) = 0.488$, $p = 0.618$].

Learning curves for the gain, loss, and neutral pairs were obtained and plotted for each group (**Figure 1**). The 20-mg baclofen group showed a faster learning rate over the first 10 trials for the gain pair, with a significant difference in trial per trial comparisons for trials 5 ($p = 0.038$), 7 ($p = 0.046$), 8 ($p = 0.034$), 9 ($p = 0.025$), 10 ($p = 0.032$), and 11 ($p = 0.032$; Mann–Whitney test for each trial after Kruskal–Wallis test showing a significant difference between the three groups). For the loss pair, in contrast, no significant difference was observed between the three groups (Kruskal–Wallis test for each trial, data not shown). This was also the case for the neutral pair (Kruskal–Wallis test for each trial, data not shown).

FIGURE 2 | Correct choice and money earned. (A) Percentage of correct choice after three concomitant sessions of 20 trials for the gain condition. There is a statistical difference between the 20-mg baclofen and the control group (*$p < 0.01$). **(B)** Amount of money earned after three concomitant sessions of 20 trials for each, the gain, loss, and neutral condition for the control, 20 mg baclofen and 50 mg baclofen groups. ±SD means standard deviation.

DISCUSSION

In this study, inspired by a rodent model of the effects of baclofen on the VTA (Compendium Suisse des Medicaments, 2011), we could demonstrate that the GABA$_B$-receptor agonist baclofen causes a significant modulation of reward-driven learning in young, healthy male humans.

REWARD LEARNING

Out of the two dosages used here, enhanced instrumental learning was only observed in the low-dose baclofen group. Participants who received 20 mg of baclofen chose the stimuli linked to the highest probability of earning money significantly more often than the other two groups. This effect is reflected by the greater steepness of the learning curve for this group relative to the placebo group at the first six trials, and a higher plateau thereafter. From this point onwards, all groups reached a relatively stable performance but with generally higher accuracy for the 20-mg group. In addition, participants in the 20-mg baclofen group tended to earn more money after the task completion, as compared to the other groups (although this failed to reach significance). However, the overall amount of money is not a reliable indicator of learning because subjects can also earn money by choosing the wrong symbol in the gain condition (0.2 probability of earning money).

These results are in agreement with those obtained in a recent study using the nearly same learning task showing improvement of learning with L-DOPA in the gain condition but not in the loss condition (Pessiglione et al., 2006). This improvement was correlated with an increase in striatal activity as measured with fMRI. A similar effect was also described in a population with Parkinson's disease with problem gambling and shopping (Voon et al., 2010). The implication of DA for reward processing is now well established. According to the "prediction error hypothesis," most DA neurons encode a "reward-prediction error" (Schultz et al., 1997) indistinctly responding to reward probability, magnitude, and the time when the predicted reward is expected (Schultz, 2007). Moreover, one third of DA neurons show a relatively slow, moderate, but significant activation that increases gradually during the interval between

a reward predicting stimulus and the reward. This activation varies monotonically with risk and could code for the discrepancy between predicted and actual reward (Fiorillo et al., 2003). Such data suggest that DA signals could have an important role in the gain condition of our learning task for evaluating, confirming, and finally learning the risk uncertainties associated with the different reward cues. The effect of 20 mg of baclofen could be explained by an enhancement of this process due to larger release of DA at striatal synapses, acting as potent learning signal and by the involvement of glutamate-dependent forms of plasticity in VTA neurons (Ungless et al., 2001; Saal et al., 2003; Borgland et al., 2004), in the NAc (Kourrich et al., 2007), and prefrontal cortex neurons (Sun et al., 2005). Moreover, we have to remember that there are other brain structures implicated in reward-coding than the mesocorticolimbic DA system. Additional discriminatory information could be provided by the orbitofrontal cortex, striatum, and amygdala (Schultz, 2010).

Besides enhanced learning by DA agonist, Pessiglione et al. (2006) also reported a significant decrease in the learning curve with the DA receptor antagonist haloperidol, which is known for its strong depressant action on the VTA system (Pessiglione et al., 2006). In a same manner, we expected a decrease in the learning rate with a dosage of 50 mg of baclofen compared to 20 mg and placebo. However, we did not observe such effect. This negative result is most probably due to the fact that the concentration of baclofen in the CSF may be too low (0.5 μM) to sufficiently inhibit DA neurons. For complete abolition of firing, a concentration of 100 μM must be reached *in vitro* (Cruz et al., 2004). However, this concentration virtually corresponds to nearly 10 g of baclofen p.o., a dose that is two orders of magnitude higher than the usual maximum dose (80 mg/day). Furthermore, these dosage-effect relationships may also be strongly influenced by each individual's pharmacokinetics. To inhibit efficiently the VTA system, concentrations close to the maximum dose or even higher concentrations may be necessary, which however can be confounded by the occurrence of adverse effects such as tiredness, muscle weakness, and headache.

AVERSION LEARNING

We did not observe any differences between the three groups in the loss condition, which is consistent with previous data (Pessiglione et al., 2006). Dopaminergic neurons respond mostly with depressed firing rates to aversive stimuli (Ungless et al., 2004; Schultz, 2007). Recent studies however identified different subpopulations of DA neurons that respond to aversive stimuli in being either excited or inhibited (Brischoux et al., 2009; Matsumoto and Hikosaka, 2009). Thus, the inhibited responding subpopulations might encode a prediction error for aversive stimuli (Matsumoto and Hikosaka, 2009). Such neurons are situated in the ventromedial SNc and VTA, projecting mainly to the ventral striatum, which is thought to process reward values as classically assumed (Matsumoto and Hikosaka, 2009). In addition, however, other structures like the lateral habenula (Matsumoto and Hikosaka, 2008) and amygdala (Paton et al., 2006) have neurons responding both to reward and aversive stimuli. These other structures might subserve learning of the loss condition without any impact of DA manipulation used in the study of Pessiglione et al. (2006) and ours.

The importance of DA neurons in the aversive condition needs to be considered and clarified in future studies. In humans, data from fMRI in healthy subjects and patients with Parkinson's disease during a similar instrumental learning task, point to the implication of a distinct brain network including the anterior insula, dorsal striatum, and orbitofrontal cortex, that influences the learning from negative outcomes (Pessiglione et al., 2006; Voon et al., 2010). Although some DA neurons in VTA are activated by aversive events, the largest DA activation is related to reward (Ungless et al., 2004). Alternatively and more specifically, the addictive behaviors in Parkinson's disease may be associated with a shift of the response to reinforcing cues from ventral (impaired) to dorsal striatum, so that the response itself becomes dominated from stimulus–response rather than action–outcome representations (Everitt and Robbins, 2005).

IMPLICATION FOR ADDICTION

The key role of the mesocorticolimbic system in the neurocircuitry of addiction is generally accepted. (Koob and Volkow, 2010), and these pathways could be implicated in addictive behaviors even long after drug exposure (Lüscher and Bellone, 2008). Although addictive drugs have very distinct molecular targets, they all cause an increase in DA concentration in the mesocorticolimbic projection target structures (Lüscher and Ungless, 2006). Moreover, there is strong evidence that drugs of abuse cause a potentiation of excitatory synapses on the VTA DA neurons (Kauer and Malenka, 2007). Synaptic plasticity might therefore represent the cellular mechanism underlying instrumental learning, which is pathologic in addicts (Balland and Lüscher, 2008). Drugs that bind to the G-protein coupled receptors (GPCR) belong to a first group of addictive drugs, which includes morphine, delta-9-tetrahydrocannabinol (THC), and the $GABA_B$-receptor agonist γ-hydroxy-butyric acid (GHB; Lüscher and Ungless, 2006). The action of these drugs is preferentially on GABA interneurons, which normally inhibit DA neurons. Thus, inhibition of GABA neurons leads to a net activation of DA neurons and an increase of DA release, a mechanism referred to as disinhibition.

As mentioned above, low-dose baclofen may have addictive properties since it preferentially disinhibits DA neurons, which increases the learning of reward signals. However, in contrast to GHB, addictive behaviors are not widely observed for baclofen, which is also a $GABA_B$-receptor agonist (European Monitoring Centre for Drugs and Drug Addiction, 2010). This apparent contradiction can be explained by their difference in affinity for the $GABA_B$-receptor (high-affinity for baclofen, low-affinity for GHB; Cruz et al., 2004). Thus, typical therapeutic doses of baclofen, particularly when given repetitively are most likely sufficient to suppress physiological DA firing and explain why baclofen is normally not abused (Labouèbe et al., 2007), while concentrations obtained with typical recreational use of GHB will preferentially affect VTA GABA neurons.

In line with this interpretation, rodent studies show that baclofen reduces self-administration of a number of drugs (Brebner et al., 2002) and is considered a putative anti-craving compound in humans (Cousins et al., 2002). A double-blind controlled study with a relatively low dosage (30 mg/day) of baclofen has shown its efficacy vs. placebo on sobriety and dropouts in alcohol-dependant patients (Addolorato et al., 2007), while in most case-reports up to 120 mg/day was used in order to obtain the same effects (Ameisen, 2005; Agabio et al., 2007; Bucknam, 2007). Also well documented is the reduction in cigaret consumption with a relatively high dosage of 80 mg/day (Franklin et al., 2009). However, the efficacy of these regimens remains controversial, as other studies reported only modest relief of symptoms (Garbutt et al., 2010). Patient adherence (low half-life of baclofen) and disease heterogeneity (for example anxious vs. non-anxious populations) may limit those studies. The potential of baclofen as a putative anti-craving compound in aiding the initiation, alleviation, and maintenance of drug abstinence is therefore still a highly discussed topic and certainly needs further clinical research.

CONCLUSION

Our randomized, double-blind, and placebo-controlled study revealed a positive reinforcement in healthy subjects taking a single dose of 20 mg of baclofen during an instrumental learning task involving monetary reward. At this dosage, subjects were more efficient in choosing the stimulus linked to the highest probability of earning money, as compared with the placebo group. These results suggest a reinforcement of prediction error learning signals by baclofen for reward stimuli, and thus corroborate with *in vitro* studies showing an enhanced activation of DA neurons with low-dose baclofen. However, these mechanisms must be confirmed by using fMRI or labeled baclofen with carbon-11, which will eventually correlate our findings with increased activity in the mesocorticolimbic DA system and associated areas. In contrast, learning was not affected by a higher dosage of 50 mg of baclofen. Such a finding suggests that even higher dosages are needed to efficiently inhibit the VTA reward system *in vivo* and to eventually serve as an anti-craving treatment.

ACKNOWLEDGMENTS

REFERENCES

Addolorato, G., Leggio, L., Ferrulli, A., Cardone, S., Vonghia, L., Mirijello, A., Abenavoli, L., D'Angelo, C., Caputo, F., Zambon, A., Haber, P. S., and Gasbarrini, G. (2007). Effectiveness and safety of baclofen for maintenance of alcohol abstinence in alcohol dependent patients with liver cirrhosis: randomized, doubleblind controlled study. *Lancet* 370, 1915–1922.

Agabio, R., Marras, P., Addolorato, G., Carpiniello, B., and Gessa, G. L. (2007). Baclofen suppresses alcohol intake and craving for alcohol in a schizophrenic alcohol dependent patient: a case report. *J. Clin. Psychopharmacol.* 27, 319–320.

Ameisen, O. (2005). Complete and prolonged suppression of symptoms and consequences of alcohol-dependence using high-dose baclofen: a self-case report of a physician. *Alcohol Alcohol.* 40, 147–150.

Balland, B., and Lüscher, C. (2008). Addiction: the dark side of learning. *Pediatr. Res.* 63, 1.

Bechara, A., Nader, K., and van der Kooy, D. (1998). A two-separate-motivational-systems hypothesis of opioid addiction. *Pharmacol. Biochem. Behav.* 59, 1–17.

Borgland, S. L., Malenka, R. C., and Bonci, A. (2004). Acute and chronic cocaine-induced potentiation of synaptic strength in the ventral tegmental area: electrophysiological and behavioral correlates in individual rats. *J. Neurosci.* 24, 7482–7490.

Brebner, K., Childress, A. R., and Roberts, D. C. (2002). A potential role for $GABA_B$ agonists in the treatment of psychostimulant addiction. *Alcohol Alcohol.* 37, 478–484.

Brischoux, F., Chakraborty, S., Brierley, D. I., and Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4894–4899.

Bucknam, W. (2007). Suppression of symptoms of alcohol dependence and craving using high-dose baclofen. *Alcohol Alcohol.* 42, 158–160.

Compendium Suisse des Medicaments. (2011). Basle: Documed.

Cousins, M. S., Roberts, D. C., and de Wit, H. (2002). $GABA_B$ receptor agonists for the treatment of drug addiction: a review of recent findings. *Drug Alcohol Depend.* 65, 209–220.

Cruz, H. G., Ivanova, T., Lunn, M. L., Stoffel, M., Slesinger, P. A., and Lüscher, C. (2004). Bi-directional effects of $GABA_B$ receptor agonists on the mesolimbic dopamine system. *Nat. Neurosci.* 7, 153–159.

European Monitoring Centre for Drugs and Drug Addiction (EMCDDA). (2010). *Statistical Bulletin*. Available at: www.emcdda.europa.eu/stats10

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902.

Franklin, T. R., Harper, D., Kampman, K., Kildea-McCrea, S., Jens, W., Lynch, K. G., O'Brien, C. P., and Childress, A. R. (2009). The $GABA_B$ agonist baclofen reduces cigarette consumption in a preliminary double-blind placebo-controlled smoking reduction study. *Drug Alcohol Depend.* 103, 30–36.

Garbutt, J. C., Kampov-Polevoy, A. B., Gallop, R., Kalka-Juhl, L., and Flannery, B. A. (2010). Efficacy and safety of baclofen for alcohol dependence: a randomized, double-blind, placebo-controlled trial. *Alcohol. Clin. Exp. Res.* 11, 1849–1857.

Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., and Dement, W. C. (1973). Quantification of sleepiness: a new approach. *Psychophysiology* 14, 431–436.

Katzung, B. (2009). *Basic and Clinical Pharmacology*, 11th Edn. New York, NY: McGraw Hill Companies.

Kauer, J. A., and Malenka, R. C. (2007). Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8, 844–858.

Koob, G. F., and Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology* 35, 217–238.

Kourrich, S., Rothwell, P. E., Klug, J. R., and Thomas, M. J. (2007). Cocaine experience controls bidirectional synaptic plasticity in the nucleus accumbens. *J. Neurosci.* 27, 7921–7928.

Labouèbe, G., Lomazzi, M., Cruz, H. G., Creton, C., Luján, R., Li, M., Yanagawa, Y., Obata, K., Watanabe, M., Wickman, K., Boyer, S. B., Slesinger, P. A, and Lüscher, C. (2007). RGS2 modulates coupling between $GABA_B$ receptors and GIRK channels in dopamine neurons of the ventral tegmental area. *Nat. Neurosci.* 10, 1559–1568.

Li, X., Lu, Z. L., D'Argembeau, A., Ng, M., and Bechara, A. (2010). The Iowa gambling task in fMRI images. *Hum. Brain Mapp.* 31, 410–423.

Lomazzi, M., Slesinger, P. A., and Lüscher, C. (2008). Addictive drugs modulate GIRK-channel signaling by regulating RGS proteins. *Trends Pharmacol. Sci.* 29, 544–549.

Lüscher, C., and Bellone, C. (2008). Cocaine-evoked synaptic plasticity: a key to addiction? *Nat. Neurosci.* 11, 737–738.

Lüscher, C., and Ungless, M. A. (2006). The mechanistic classification of addictive drugs. *PLoS Med.* 3, e437. doi: 10.1371/journal.pmed.0030437

Matsumoto, M., and Hikosaka, O. (2008). Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci.* 12, 77–84.

Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841.

Paton, J. J., Belova, M. A., Morrison, S. E., and Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., and Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.

Redish, A. D., Jensen, S., and Johnson, A. (2008). A unified framework for addiction: vulnerabilities in the decision process. *Behav. Brain Sci.* 31, 415–437.

Saal, D., Dong, Y., Bonci, A., and Malenka, R. C. (2003). Drugs of abuse and stress trigger a common synaptic adaptation in dopamine neurons. *Neuron* 37, 577–582.

Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci.* 30, 203–210.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6, 24.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Sun, X., Zhao, Y., and Wolf, M. E. (2005). Dopamine receptor stimulation modulates AMPA receptor synaptic insertion in prefrontal cortex neurons. *J. Neurosci.* 25, 7342–7351.

Thorndike, E. (1898). Some experiments on animal intelligence. *Science* 7, 818–824.

Ungless, M. A., Magill, P. J., and Bolam, J. P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303, 2040–2042.

Ungless, M. A., Whistler, J. L., Malenka, R. C., and Bonci, A. (2001). Single cocaine exposure in vivo induces long-term potentiation in dopamine neurons. *Nature* 411, 583–587.

Voon, V., Pessiglione, M., Brezing, C., Gallea, C., Fernandez, H. H., Dolan, R. J., and Hallett, M. (2010). Mechanisms underlying dopamine mediated reward bias in compulsive behaviors. *Neuron* 65, 135–142.