Supporting Information

Mizoguchi et al. 10.1073/pnas.1418014112

SI Methods

Gambling Test for Rodents Using a Radial Arm Maze. In other experiments, reward probabilities and reward values of the H-H and L-L arms were varied to test the influence on performance in the gambling test for rodents. Accordingly, reward probability and reward value of the H-H arm were varied as follows: 12.5% probability (2/16 trials) of a large reward of 7 food pellets (315 mg; standard condition), 25% probability (4/16 trials) of a large reward of 3.5 food pellets (158 mg), and 50% probability (8/16 trials) of a large reward of 1.75 food pellets (79 mg). Reward probability and reward value of the L-L arm were varied as follows: 87.5% probability (14/16 trials) of a small reward (1 food pellet, 45 mg), 75% probability (12/16 trials) of a small reward of 1.17 food pellets (52.5 mg), and 50% probability (8/16 trials) of a small reward of 1.75 food pellets (79 mg). Consequently, under these conditions the total reward values of the H-H and L-L arms were always equal (14 food pellets, 630 mg) (Fig. S1 *B* and *C*).

In the double-reward condition, the total reward values of the H-H and L-L arms were both doubled (1,260 mg), but the reward probabilities were the same as those in the standard condition, with a large-reward probability of 12.5% (2/16 trials) in the H-H arm and a small-reward probability of 87.5% (14/16 trials) in the L-L arm. Thus, under the double-reward condition, the large reward in the H-H arm was 14 food pellets (630 mg), and the small reward in the L-L arm was 2 food pellets (90 mg) (Fig. S1 *B* and *C*).

Empty Arm Choice. Ratio of empty arm choice immediately after receiving large reward, small reward, or quinine pellet was expressed as follows: empty arm choice (%) after large reward = [((total number of empty arm choice behaviors immediately after receiving large rewards)/(total number of large rewards acquired)) × 100] (16 trials per day × 14 d = 224 trials), empty arm choice (%) after quinine pellet in the H-H arm = [((total number of empty arm choice behaviors immediately after receiving quinine pellets)/(total number of quinine pellets acquired in the H-H arm at any two trials)) × 100] (16 trials per day × 14 d = 224 trials), empty arm choice (%) after small reward = [((total number of empty arm choice behaviors immediately after receiving small rewards)/(total number of small rewards acquired in the L-L arm at any two trials)) × 100] (16 trials per day × 14 d = 224 trials), and empty arm choice (%) after quinine pellet in the L-L arm = [((total number of empty arm choice behaviors immediately after receiving quinine pellets)/(total number of quinine pellets acquired in the L-L arm)) × 100] (16 trials per day × 14 d = 224 trials).

Gambling Test for Examining the Effect of Posttreatment with METH on Choice Behavior. Rats were subjected to the gambling test for 14 d and then received repeated METH or saline treatment for 14 d. The animals were subjected to the gambling test twice during the METH treatment, on days 7 and 8 and days 14 and 15. A third postgambling test was conducted 1 wk after withdrawal.

Moreover, to determine whether preference for the H-H arm in METH-treated rats is dependent on subsequent experience, the gambling test was performed under a condition of 100% small reward (always one food pellet) in the L-L arm and 0% large reward (always one quinine pellet) in the H-H arm. The animals were first subjected to the gambling test for 14 d under the standard condition and then received repeated METH or saline treatment for 14 d. The animals were subjected to the gambling test twice during saline or METH treatment (i.e., on days 7 and 8 and days 14 and 15) and 1 wk after the withdrawal (on days 21 and 22).

Win–stay or lose–shift behavior was expressed during the postgambling test as well as in Fig. 1.

Food Consumption Test in the Radial Maze Task. Rats were mildly food-restricted. The apparatus contained one start arm and one food arm, which contained a food reward (100 pellets) in a food cup. When the guillotine door was opened in the start arm, rats had no choice but to enter the food arm. Under this condition, rats ate food pellets in the food arm, and once they were satisfied, they came back to the central platform. The following indices of motivation to food were analyzed: approach time (i.e., the time taken by rats to approach the food cup in the food arm after the guillotine door opens), consumption time (i.e., the period during which the rat was eating the food), number of rewards eaten, and number of rewards eaten divided by the consumption time.

Production and Purification of Viral Vectors Expressing Designer Receptors Exclusively Activated by Designer Drug (DREADDs). HEK293 cells were transfected with a pAAV vector plasmid containing the gene of interest, pHelper, and pAAV-RC (serotype 10) (provided by Penn Vector Core) using a standard calcium-phosphate method. Three days later, the transfected cells were collected and suspended in artificial CSF (124 mM NaCl, 3mM KCl, 26 mM $NaHCO_3$, 2 mM $CaCl_2$ 1 mM $MgSO_4$, 1.25 mM KH_2PO_4, 10 mM D-Glucose). After four freeze–thaw cycles and subsequent centrifugation, the supernatant was treated with Benzonase nuclease at 45 °C for 15 min. Purified viruses suspended in artificial CSF were titered by quantitative PCR and stored at −80 °C before use. Plasmids pAAV-CMV-hM3Dq-mCherry and pAAV-CMV-hM4Di-mCherry were constructed from plasmids pAAV-hSyn-FLEX-hM3Dq and pAAV-hSyn-FLEX-hM4Di, respectively, purchased from Addgene (IDs: 44361 and 44362).

c-Fos Immunohistochemistry and Quantitative Analysis. c-Fos immunostaining was performed as described previously (40). Animals were deeply anesthetized with sodium pentobarbital (50 mg/kg) 2 h after the gambling test and then transcardially perfused with ice-cold PBS, followed by 4% paraformaldehyde in PBS. The brains were removed, postfixed in the same fixative, and then cryoprotected in 10–30% sucrose in PBS. Frozen serial coronal slices (40 μm) of the entire brain were made and then incubated with rabbit anti–c-Fos antibody (1:3,000; sc-253; Santa Cruz Biotechnology) for 24 h at 4 °C. We used the EnVision system-HRP (Dako), based on an HRP-labeled polymer conjugated to secondary antibodies. To quantify the number of c-Fos–positive cells in the brain, we used a microscope with a cooled CCD digital camera system (NanoZoomer 2.0; Hamamatsu) to scan the slices and calculated the cell numbers from the digitized images using the Win ROOF image analysis software (ver. 5.6; Mitani Co.). From both the saline- and METH-treated groups, we selected animals that exhibited typical/average responses in the gambling test and then counted c-Fos–positive cells in six to eight different sections from each animal (40). Selected areas were as follows: OFC, INS, anterior cingulated cortex (ACC), prelimbic cortex (PrL), core (NAc) and shell (NAs) of nucleus accumbens, and striatum (St).

To examine the effect of CNO treatment on neural activity in the INS, rats were killed 1.5 h after CNO treatment, and then c-Fos immunohistochemistry (1:1,000; sc-253) was performed as described above. Affinity-purified FITC-conjugated goat anti-rabbit

IgG was used as the secondary antibody. We selected more than ten different sections from each animal, and the images were analyzed with a deconvolution fluorescence microscope system (BZ-9000; Keyence). The average numbers of c-Fos–positive cells were used for statistical analysis.

Statistical Analyses for Reinforcement Learning Model-Based Analysis. We used the maximum likelihood approach to estimate the reinforcement learning model parameter from rats' choice data. To determine which parameter differs between subject groups (saline-treated and METH), we adopted the following procedure. First, we constructed a model set of all combinations for each parameter. The parameters were either allowed to have different values between two subject groups (saline and METH) or shared the same value between the groups. We also examined combinations in which we set either κ_E or κ_Q to 0, both of them to a nonzero value, or both of them to 0. We then used fixed-effect analysis; to obtain a stable estimator, for each model a single parameter set was estimated for all subjects considered as a whole (14). The model parameters were optimized by minimizing the negative log-likelihood using the Matlab function "fmincon." We compared the models based on the Akaike information criterion (AIC). The model that yielded the smallest AIC was deemed the best model. Based on the best model selected by the AIC, we tested whether the differences in parameters between groups were significant by conducting the likelihood ratio test, with the null hypothesis that the improvement in the likelihood of differentiation in the model parameters between groups occurred by chance alone.

SI Results

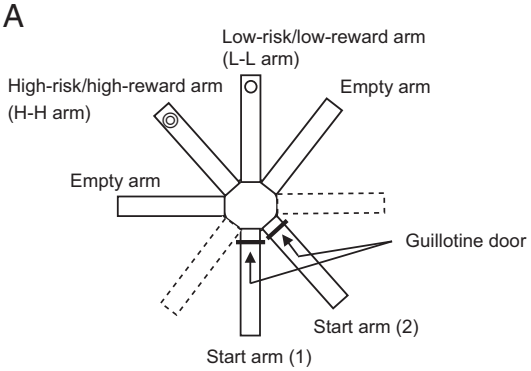
To determine whether posttreatment with METH altered the established preference for L-L arm in the gambling test, rats were subjected to the gambling test for 14 d and then received repeated METH or saline treatment for 14 d. The animals were subjected to the gambling test twice during the METH treatment, on days 7 and 8 and days 14 and 15. A third postgambling test was conducted 1 wk after withdrawal (on days 21 and 22; Fig. S5*A*).

In the saline-treated group, the H-H arm choice ratios in the first, second, and third postgambling tests remained stable at ~10% (Fig. S5*B*). Following METH treatment, the animals chose the H-H arm more than saline-treated control rats did, and this increase was maintained at least 7 d after the withdrawal (Fig. S5*B*). Accordingly, in the second and third postgambling test, H-H arm choice ratios in the METH-treated group were significantly elevated relative to the saline-treated group. There was no difference between saline- and METH-treated rats in the number of entries into empty arms (Fig. S5*C*).

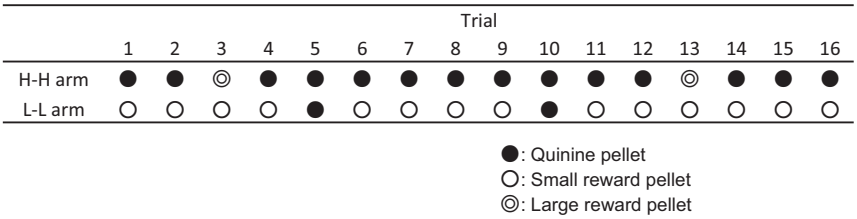
In addition, we analyzed the effect of posttreatment with METH on win–stay/lose–shift behavior in the gambling test (Fig. S5 *D*–*G*). The effects were quite similar to those observed in the pretreatment experiment (Fig. 1 *D*–*G*). When METH-treated rats received a large reward following a choice of the H-H arm, they chose the H-H arm in the next trial more frequently than control rats did (Fig. S5*D*). Alternatively, when METH-treated rats obtained a quinine pellet following a choice of the L-L arm, they chose the H-H arm in the next trial more frequently than did control animals (Fig. S5*G*). There were no differences in H-H arm choice ratios after animals received a quinine pellet in the H-H arm (Fig. S5*E*) or the small reward in the L-L arm (Fig. S5*F*).

Finally, we investigated the effect of repeated METH treatment on arm choice behavior in the gambling test under a condition of 100% small reward (always one food pellet) in the L-L arm and 0% large reward (always one quinine-coated pellet) in the H-H arm. Accordingly, the animals were first subjected to the gambling test for 14 d under the standard condition and then received repeated METH or saline treatment for 14 d. The animals were subjected to the gambling test twice during the METH treatment (i.e., on days 7 and 8 and days 14 and 15) and 1 wk after the withdrawal (on days 21 and 22) (Fig. S6). Under these conditions, both groups of rats preferentially chose the L-L arm, and there were no differences between the saline- and METH-treated groups in H-H arm choice ratio (Fig. S6*B*) or number of entries into empty arms (Fig. S6*C*). Thus, alteration of decision-making induced by METH treatment was manifested under conditions of uncertainty, i.e., when RPEs were inserted during the test.

A



B



C

	Standard condition			Double-reward condition		
	Reward value	Reward probability	Total reward value	Reward value	Reward probability	Total reward value
H-H arm	315 mg	12.5 %	630 mg	630 mg	12.5 %	1260 mg
L-L arm	45 mg	87.5 %	630 mg	90 mg	87.5 %	1260 mg

Fig. S1. Gambling test for rodents using an eight-arm radial arm maze. (*A*) The apparatus consisted of two start arms and four choice arms [one low-risk/low-return (L-L) arm, one high-risk/high-return (H-H) arm, and two empty arms] on an elevated six-arm Plexiglas radial maze. The other two arms of the original eight-arm maze were not used (removed). The food cups were glued onto the surfaces of the arms. The entrances from the start arms to the central platform could be blocked by black guillotine doors. In the decision-making task, rewards were 45-mg banana-flavored food pellets, and negative outcomes (food that was disappointing in comparison with the reward) were 45-mg quinine-coated food pellets that were unpalatable but not inedible. Rats could obtain either rewards or quinine pellets at the ends of the baited arms, but no food pellets were placed in empty arms. (*B*) Reward probabilities and reward values of the H-H and L-L arms. Under the standard condition, choice of the L-L arm resulted in frequent (14/16 trials, reward probability: 87.5%) small rewards (one 45-mg food pellet; open circle) with infrequent (2/16 trials) negative outcomes (one quinine-coated pellet; closed circle). Choice of the H-H arm resulted in infrequent (2/16 trials, reward probability: 12.5%) large rewards (seven 45-mg food pellets; double open circle) with frequent (14/16 trials) negative outcomes (one quinine-coated pellet; closed circle). Quinine pellets in the L-L arm and large rewards in the H-H arm were provided in a random manner at a rate of 2 out of 16 trials per day. (*C*) Comparison of the standard condition and the double-reward condition.

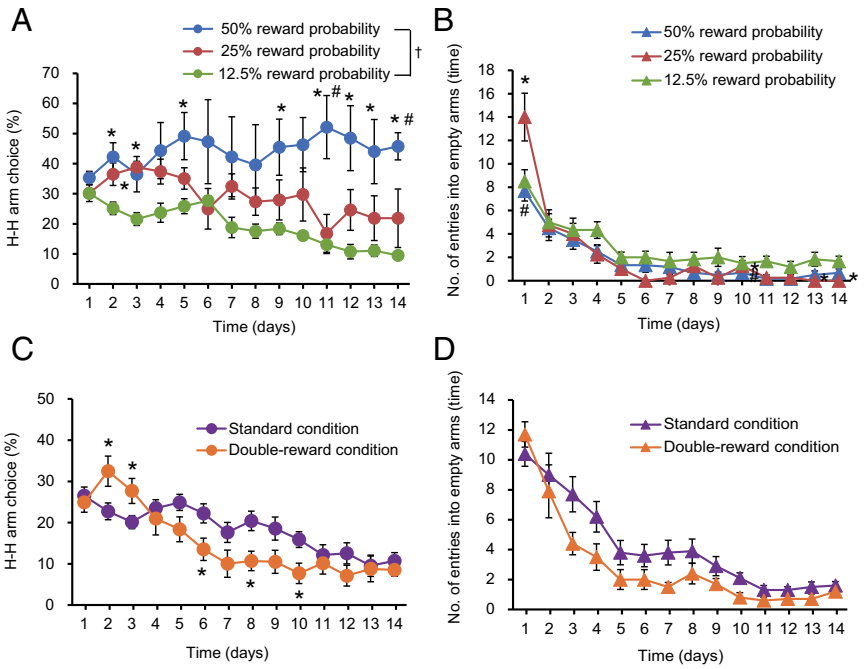


Fig. S2. (*A*) Effect of changing the reward probability on H-H arm choice ratio in the gambling test. Reward probability of H-H arm was changed from 12.5% to 50%, and reward probability of L-L arm was changed from 87.5% to 50%. Values are means ± SE ($n = 4–6$). Repeated ANOVA revealed significant effects on reward probability [$F_{(2,13)} = 5.04$, $P < 0.05$], time [$F_{(13,169)} = 1.81$, $P < 0.05$], and their interaction [$F_{(26,169)} = 1.84$, $P < 0.05$]. $^{\dagger}P < 0.05$. $^{*}P < 0.05$ vs. the 12.5% reward probability group on each day. $^{\#}P < 0.05$ vs. the 25% reward probability group on each day. (*B*) Effect of changing the reward probability on the number of entries into empty arms. Values are means ± SE ($n = 4–6$). There was no effect on reward probability [$F_{(2,13)} = 2.69$, $P > 0.05$], time [$F_{(13,169)} = 59.10$, $P < 0.05$], or their interaction [$F_{(26,169)} = 3.46$, $P < 0.05$]. $^{*}P < 0.05$ vs. the 12.5% reward probability group on each day. $^{\#}P < 0.05$ vs. the 25% reward probability group on each day. (*C*) Effect of changing reward values of the H-H and L-L arms on H-H arm choice ratio. Total reward values of the H-H and L-L arms were doubled from 630 mg to 1260 mg, but the reward probabilities remained the same with the standard condition (12.5% and 87.5%, respectively). Values are means ± SE ($n = 10$). Repeated ANOVA revealed no significant main effect on reward value [$F_{(1,18)} = 2.15$, $P > 0.05$], time [$F_{(13,234)} = 18.18$, $P < 0.05$], or interaction [$F_{(13,234)} = 3.86$, $P < 0.05$]. On days 2, 3, 6, 8, and 10, simple main effects were significant ($P < 0.05$). (*D*) Effect of changing reward values of the H-H and L-L arms on the number of entries in empty arms. Values are means ± SE ($n = 10$). There were no significant main effects [reward value: $F_{(1,18)} = 3.70$, $P = 0.07$; time: $F_{(13,234)} = 49.56$, $P < 0.05$; interaction: $F_{(13,234)} = 1.66$, $P = 0.07$].

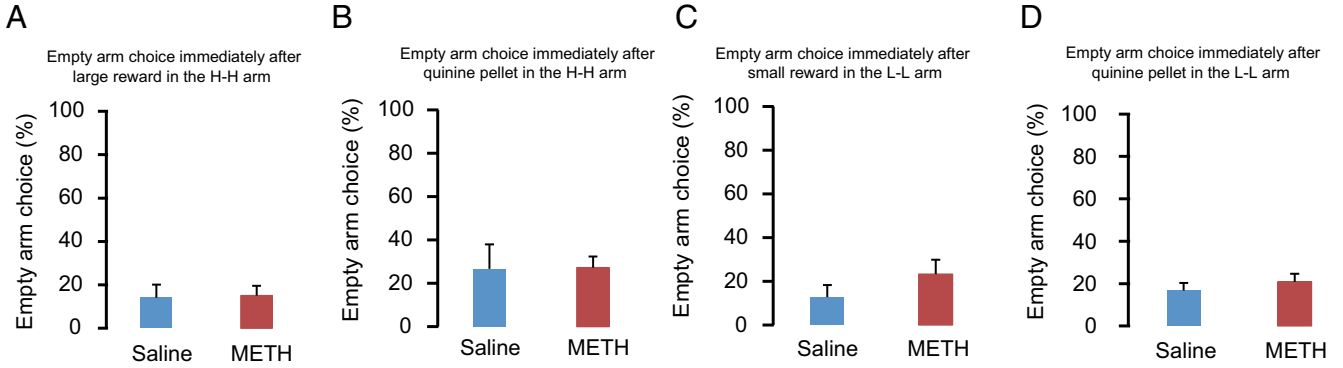


Fig. S3. (*A* and *C*) Win–empty/(*B* and *D*) lose–empty behavior of METH-treated rats in the gambling test. Values are means ± SE ($n = 5–6$). There were no significant effects ($P > 0.05$ by *u* test).

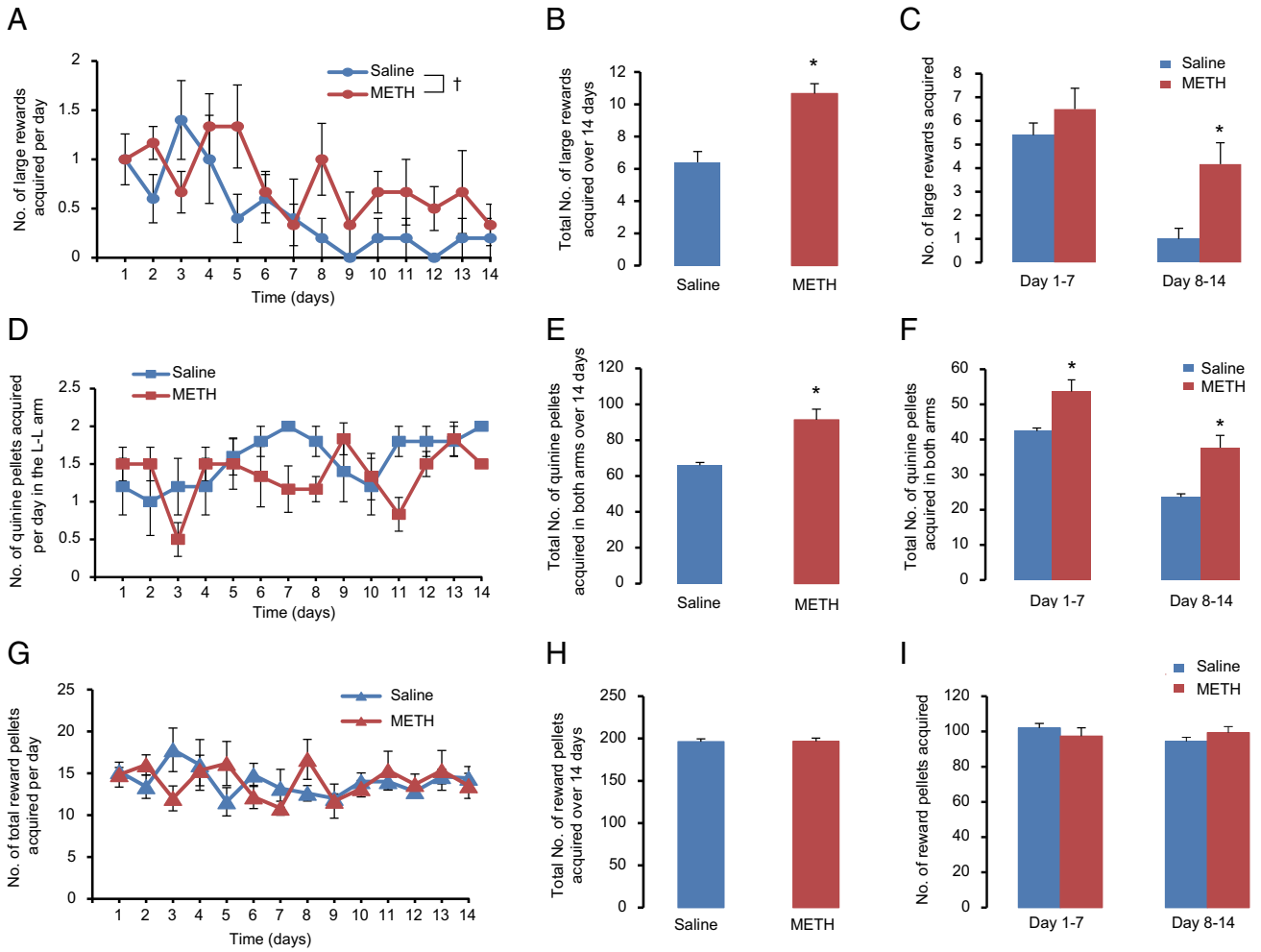


Fig. S4. Difference between saline- and METH-treated rats in the number of reward pellets acquired. (*A–C*) Number of large rewards acquired in the H-H arm. Values are means ± SE ($n = 5–6$). $^{\dagger}P < 0.05$. $^{*}P < 0.05$ vs. saline-treated group. (*A*) Number of large rewards acquired per day. (*B*) Total number of large rewards acquired over 14 d. (*C*) Number of large rewards acquired in the first (days 1–7) and second (days 8–14) halves of the gambling test. There were significant effects in the number of large rewards acquired per day [*A*; treatment: $F_{(1,9)} = 21.76$, $P < 0.05$; time: $F_{(13,117)} = 2.59$, $P < 0.05$; interaction: $F_{(13,117)} = 1.04$, $P > 0.05$] and in the total number of large rewards acquired over 14 d (*B*; $P < 0.05$). There was no difference between the two groups in the first block (*C*; days 1–7, $P > 0.05$; days 8–14, $P < 0.05$). (*D–F*) Number of quinine pellets acquired. Values are means ± SE ($n = 5–6$). (*D*) Number of quinine pellets acquired per day in the L-L arm. There was no significant main effect on treatment [$F_{(1,9)} = 1.99$, $P > 0.05$], time [$F_{(13,117)} = 1.72$, $P > 0.05$], or their interaction [$F_{(13,117)} = 1.69$, $P > 0.05$]. (*E*) Total number of quinine pellets acquired in both arms over 14 d. $^{*}P < 0.05$ vs. saline-treated group. (*F*) Total number of quinine pellets acquired in both arms in the first (days 1–7) and second (days 8–14) halves of the gambling test. $^{*}P < 0.05$ vs. the saline-treated group. (*G–I*) Number of reward pellets acquired. Values are means ± SE ($n = 5–6$). (*G*) Number of reward pellets acquired per day. There was no significant effect on treatment [$F_{(1,9)} = 0.003$, $P > 0.05$], time [$F_{(13,117)} = 0.83$, $P > 0.05$], or their interaction [$F_{(13,117)} = 1.19$, $P > 0.05$]. (*H*) Total number of reward pellets acquired over 14 d ($P > 0.05$). (*I*) Number of reward pellets acquired in the first half (days 1–7) and second half (days 8–14) of the gambling test (days 1–7, $P > 0.05$; days 8–14, $P > 0.05$).

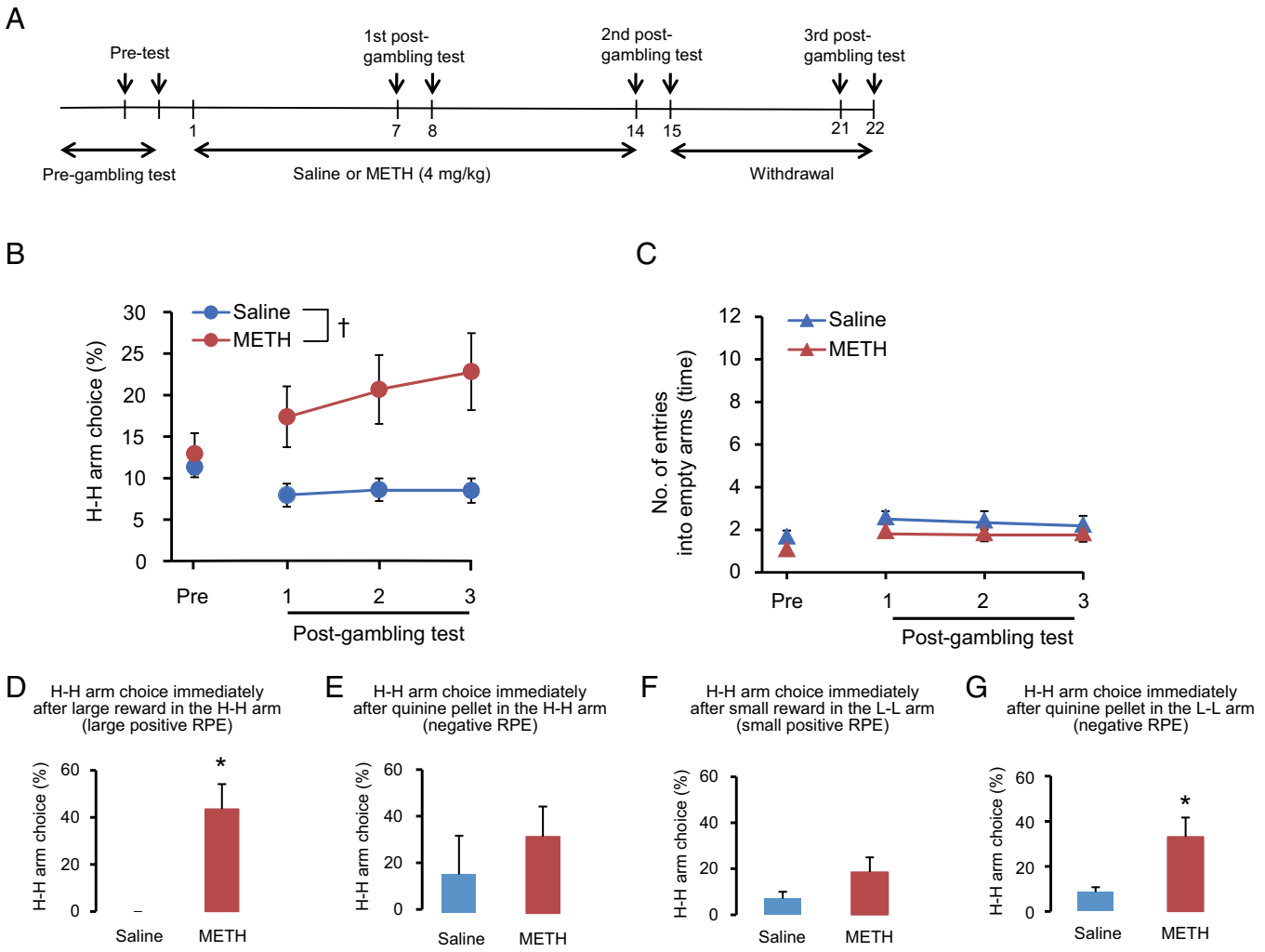


Fig. S5. Posttreatment with METH altered the established preference for the L-L arm in the gambling test. (*A*) Experimental scheme of the behavioral assessment. Animals that had been previously subjected to the gambling test were given METH (4 mg/kg, s.c.) once a day for 14 consecutive days. The effect of posttreatment with METH on choice behavior in the gambling test was investigated three times (the first gambling test on days 7 and 8, the second on days 14 and 15, and the third on days 21 and 22) after initiation of METH administration. Data from the first, second, and third gambling tests were expressed as mean performance over 2 d. (*B*) Effect on H-H arm choice ratio in the gambling test. Values are means ± SE ($n = 10$). $^{\dagger}P < 0.05$. Repeated ANOVA revealed significant effects [treatment: $F_{(1,18)} = 8.90$, $P < 0.05$; time: $F_{(2,36)} = 1.89$, $P > 0.05$; interaction: $F_{(2,36)} = 1.23$, $P > 0.05$]. (*C*) Effect on the number of entries into empty arms in the gambling test. Values are means ± SE ($n = 10$). There were no significant effects [treatment: $F_{(1,18)} = 2.35$, $P > 0.05$; time: $F_{(2,36)} = 0.24$, $P > 0.05$; interaction: $F_{(2,36)} = 0.08$, $P > 0.05$]. (*D*–*G*) Win–stay/lose–shift behavior in the three postgambling tests (first through third tests). Values are means ± SE [(*D*) $n = 7$–9, (*E*) $n = 6$–10, (*F*) $n = 10$, and (*G*) $n = 10$]. $^{*}P < 0.05$ vs. saline-treated group. There were significant effects in *D* ($P < 0.05$ by *u* test) and *G* ($P < 0.05$ by *u* test) but not in *E* ($P > 0.05$ by *u* test) or *F* ($P > 0.05$ by *u* test).

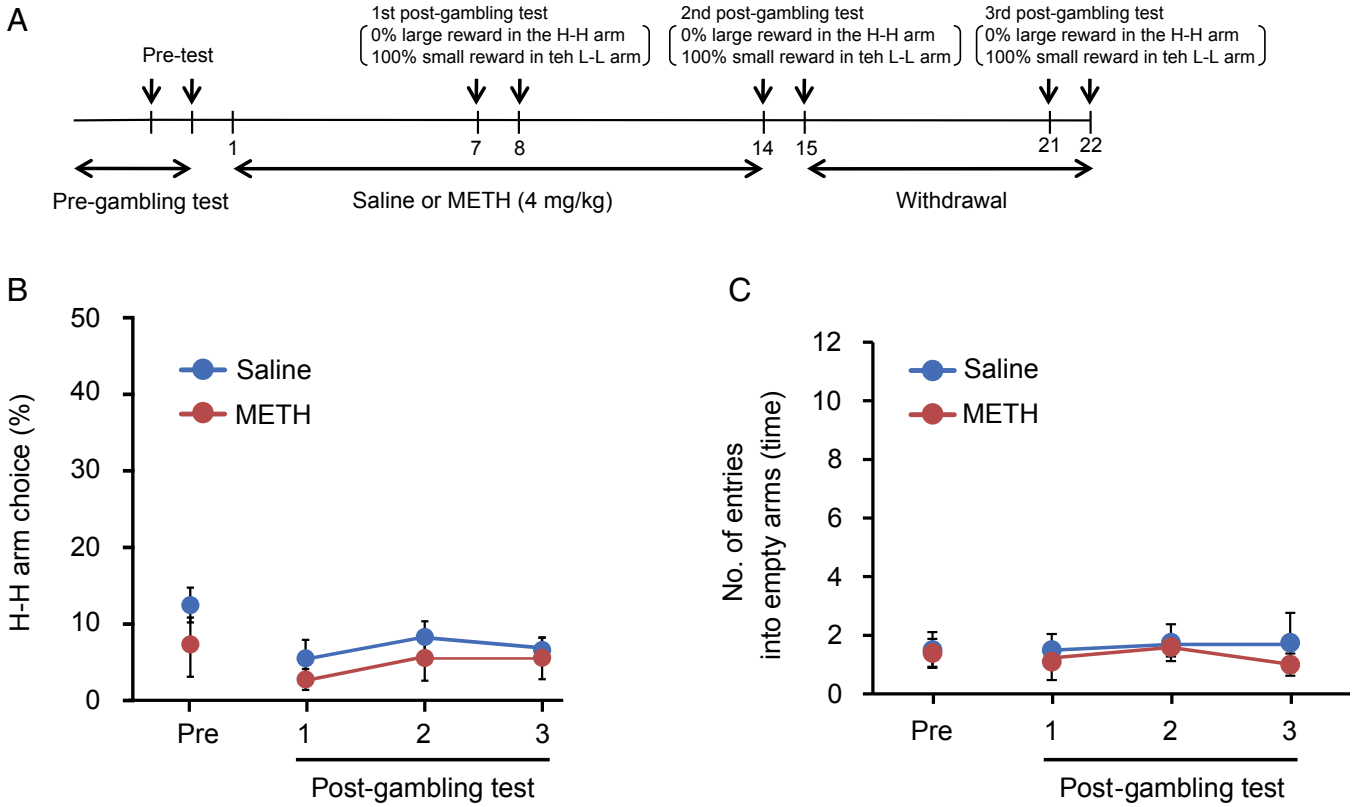


Fig. S6. Preference for the H-H arm in METH-treated rats is dependent on subsequent experience. (*A*) Experimental scheme of behavioral assessment. Animals that had been previously subjected to the gambling test were given METH (4 mg/kg, s.c.) once a day for 14 consecutive days. The effect of posttreatment with METH on choice behavior in the gambling test was investigated three times (first gambling test on days 7 and 8, second on days 14 and 15, and third on days 21 and 22) after initiation of METH administration. Data from the first, second, and third gambling tests are expressed as the mean performance over 2 d. METH or saline-treated rats were trained under a 0% ratio of quinine-coated pellets in the L-L arm and of large reward in the H-H arm. The effect on decision-making of the unexpected experience of a large reward or quinine pellet after posttreatment with METH was investigated 7, 8, 14, and 15 d after the first administration, as well as 21 and 22 d afterward (during withdrawal). (*B*) Effect on H-H arm choice. Values are means ± SE ($n = 4$–5). Repeated ANOVA revealed no significant effects [treatment: $F_{(1,7)} = 0.12$, $P > 0.05$; time: $F_{(2,14)} = 0.33$, $P > 0.05$; interaction: $F_{(2,14)} = 0.07$, $P > 0.05$]. (*C*) Effect on the number of entry into empty arms in the gambling test. Values are means ± SE ($n = 4$–5). Repeated ANOVA revealed no significant effects [treatment: $F_{(1,7)} = 0.36$, $P > 0.05$; time: $F_{(2,14)} = 0.78$, $P > 0.05$; interaction: $F_{(2,14)} = 0.45$, $P > 0.05$].

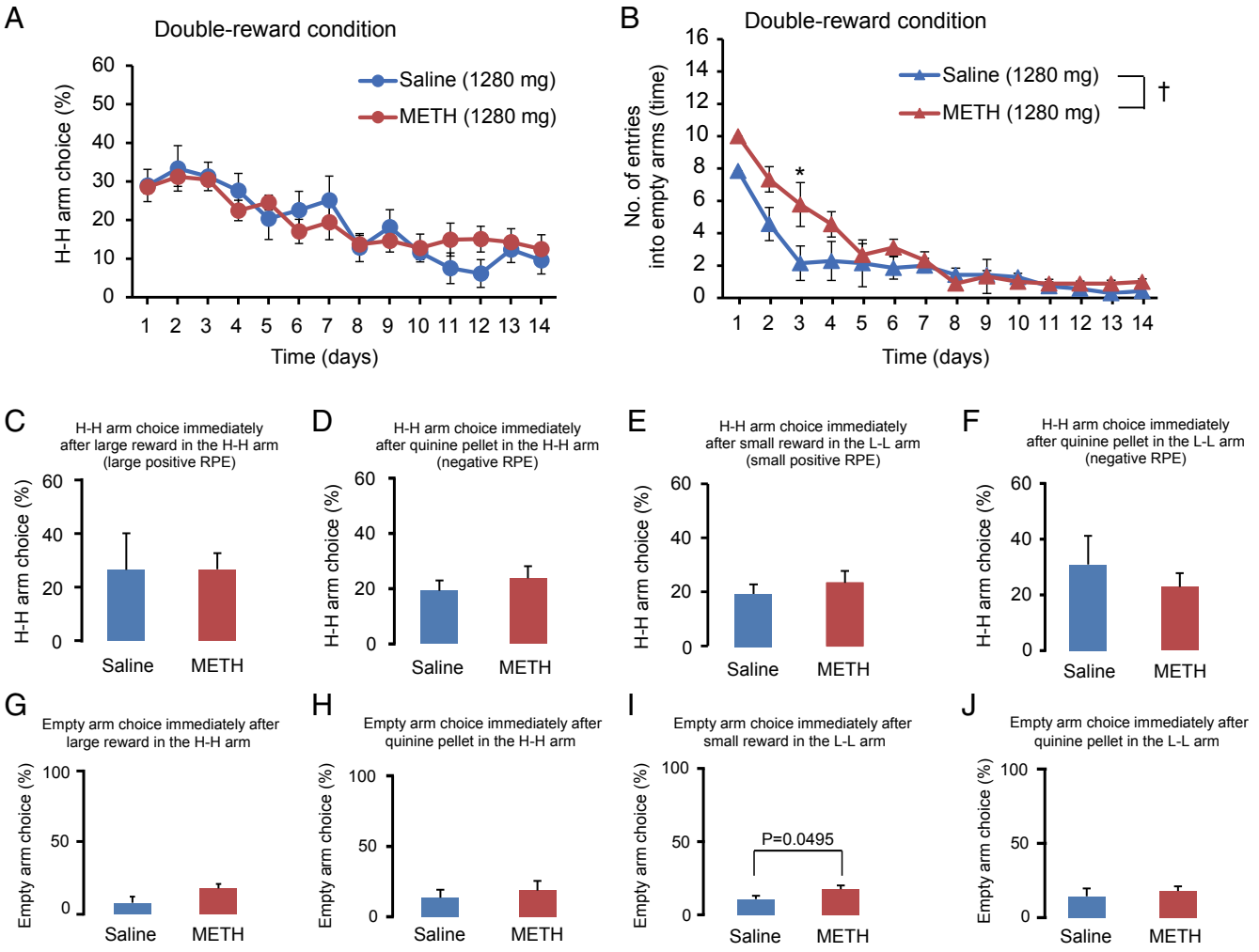


Fig. S7. Performance of chronic METH-treated rats in the gambling test under the double-reward condition. Rats were given METH at a dose of 4 mg/kg once a day for 30 d. The gambling test was initiated after 2 wk of METH withdrawal, as shown in Fig. 1. (*A*) Effect on H-H arm choice ratio in the gambling test. Values are means ± SE (n = 7–9). Repeated ANOVA revealed no significant effect on H-H arm choice ratio [treatment: $F_{(1,14)} = 0.01$, $P > 0.05$; time: $F_{(13,182)} = 12.76$, $P < 0.05$; interaction: $F_{(13,182)} = 1.10$, $P > 0.05$]. (*B*) Number of entries into empty arms. Values are means ± SE (n = 7–9). Repeated ANOVA revealed significant effect on the number of entries into empty arms [treatment: $F_{(1,14)} = 4.95$, $P < 0.05$; time: $F_{(13,182)} = 32.70$, $P < 0.05$; interaction: $F_{(13,182)} = 2.25$, $P < 0.05$]. $^{\dagger}P < 0.05$. $^{*}P < 0.05$ vs. the saline-treated group on each day. (*C*–*F*) Win–stay/lose–shift behavior. Values are means ± SE (n = 7–9). There were no significant effects in any of these figures ($P > 0.05$ by u test). (*G*–*J*) Win–empty/lose–empty behavior. Values are means ± SE (n = 7–9). A significant effect was observed in *I* ($P = 0.0495$ by u test) but not *G*, *H*, or *J* ($P > 0.05$ by u test).

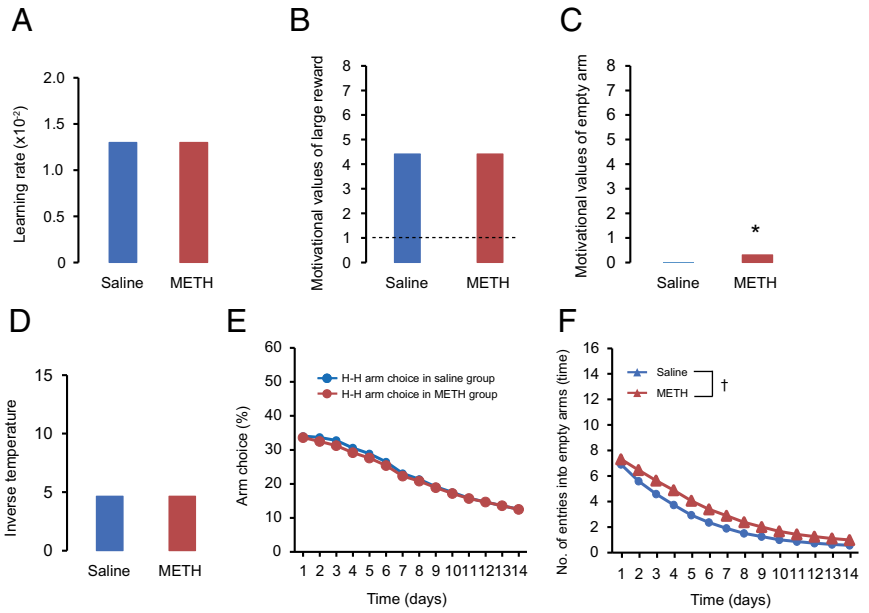


Fig. S8. Reinforcement learning model-based analysis of rats' performance under the double-reward condition. The conventions are the same as in Fig. 2. (*A–D*) The model in which the learning rate differs between METH-treated rats and control rats and the motivational value for empty arm has nonzero value whereas the value for a quinine pellet is set to zero were selected based on the AIC (AIC = 6,595.24). A dotted line indicates one motivational value for small reward in each group. The difference in learning rate did not reach the level of significance [$\chi^2(1) = 2.83$, $P = 0.092$, likelihood ratio test], whereas the difference in the motivational value of the empty arm yielded a significant effect [$\chi^2(1) = 11.20$, $P < 0.001$]. Thus, we used the model in which all of the parameters, except for the motivational value of the empty arm, were shared by the two groups (AIC = 6,596.07). (*E* and *F*) Accordingly, the simulated learning curves (the development of H-H arm choice ratio) and the number of entry in empty arms were almost identical.

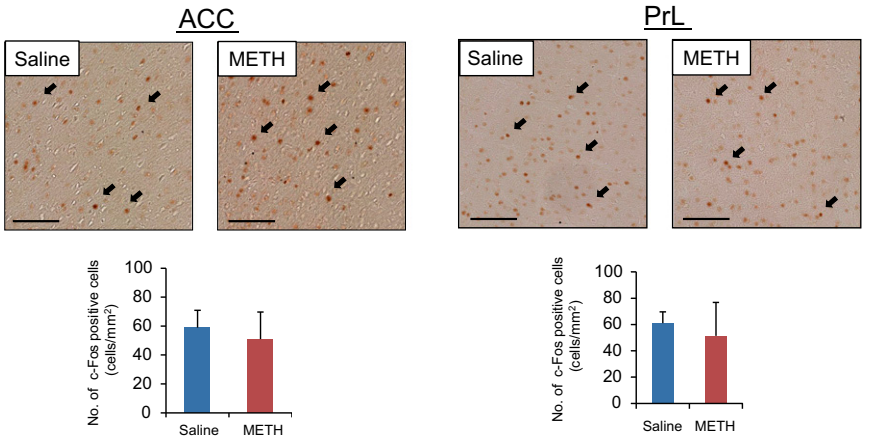


Fig. S9. No changes in c-Fos expression in the ACC and PrL of METH-treated rats were evoked by the gambling test under the standard condition. c-Fos expression was analyzed immunohistochemically in saline- and METH-treated rats 2 h after the gambling test as in Fig. 3. Photographs show typical examples of c-Fos expression in various brain areas of saline- or METH-treated rats after the gambling test under the standard condition. Values are means ± SE (n = 3–4). ACC, anterior cingulate cortex; PrL, prelimbic cortex. (Scale bar: 100 μm.)