Title:

An altered cognitive strategy associated with reduction of synaptic inhibition in the prefrontal cortex after social play deprivation in rats

Authors:

Azar Omrani[1,3,4,*] , Ate Bijlsma[1,2,*] , Marcia Spoelder[2,5,*] , Jeroen P.H. Verharen[2,3,6], Lisa Bauer[1] , Cosette Cornelis[2] , René van Dorland[1] , Louk J.M.J. Vanderschuren[2,*,#], Corette J. Wierenga[1,*,#]

***equal contributions**

[#] corresponding authors

Affiliations:

[1] Department of Biology, Faculty of Science, Utrecht University, Utrecht, the Netherlands

[2] Department of Animals in Science and Society, Division of Behavioural Neuroscience, Faculty of Veterinary Medicine, Utrecht University, Utrecht, the Netherlands

[3] Department of Translational Neuroscience, University Medical Center Utrecht, Utrecht, the Netherlands

[4] current address: Department of CNS Diseases Research, Boehringer Ingelheim Pharma GmbH & Co KG, Biberach, Germany

[5] current address: Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands

[6] current address: Helen Wills Neuroscience Institute, Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, CA, USA

Correspondence to:

Corette J. Wierenga c.j.wierenga@uu.nl

Louk J.M.J. Vanderschuren l.j.m.j.vanderschuren@uu.nl

Acknowledgements: Supported by the Netherlands Organisation for Scientific Research (NWO) ALWOP.2015.105 (LJMJV, CJW, AB). We thank Daniëlle Counotte for preliminary experiments and Ruth Damsteegt for practical assistance.

Author Contribution: The study was conceived by CW and LV. Experiments were designed by AO, MS, LV and CW. Experiments were performed by AO, MS, LB, RD and CC. Analysis was performed by AB, AO, MS, JV, LB and CC. Manuscript was written by AB, LV and CW with critical input from all authors.

Abstract

Experience-dependent organization of neuronal connectivity is a critical component of brain development, but how experience shapes prefrontal cortex (PFC) development is unknown. Here, we assessed how social play behaviour, which is highly abundant during post-weaning development, shapes PFC function and connectivity. We subjected juvenile rats to social play deprivation (SPD), followed by resocialization until adulthood. In a PFC-dependent probabilistic reversal learning task, SPD rats earned a similar number of rewards, but achieved more reversals than control rats. Computational trial-by-trial analysis showed that SPD rats displayed a simplified cognitive strategy. In addition, inhibitory synaptic currents were significantly reduced in layer 5 PFC cells of SPD rats, with specific changes in parvalbumin- and cannabinoid receptor 1-positive perisomatic inhibitory synapses. Thus, SPD has a long-lasting impact on PFC inhibition via synapse-specific alterations, associated with simplified cognitive strategies. We conclude that proper PFC development depends on pertinent social experience during a restricted time period.

Main (Introduction, Results and Discussion)

The developing brain requires proper external input to fine-tune activity and connectivity in neural circuits to ensure optimal functionality throughout life. This process has been extensively studied in the sensory cortices, whereby sensory deprivation during development causes long-lasting deficits in sensory processing resulting from improper synaptic wiring[1,2]. However, how experience-dependent plasticity contributes to the development of other brain structures, such as the prefrontal cortex (PFC)[3,4], is less understood.

The PFC is important for higher cognitive, so-called executive functions[5,6], as well as the neural operations required during social interactions[7,8]. By analogy of sensory cortex development, PFC development may require complex cognitive and social stimuli. Importantly, during the period when the PFC matures, i.e. in between weaning and early adulthood[9], young animals display an abundance of an energetic form of social behaviour known as social play behaviour[10,11]. It is widely held that exploration and experimentation during social play facilitates the development of a rich behavioural repertoire, that allows an individual to quickly adapt in a changeable world. In this way, social play subserves the development of PFC-dependent skills such as flexibility, creativity, and decision-making[10,12,13]. In support of a relationship between PFC maturation and social play behaviour, lesions or inactivation of the PFC have been found to impair social play[14,15].

Lack of social play experience during post-weaning development may cause permanent changes in PFC circuitry[13,16,17]. However, the cellular mechanisms by which social play facilitates PFC development remain elusive. It is well described that sensory deprivation induces specific alterations in inhibitory neurotransmission that affect adult sensory processing[1,18], and early social experiences may therefore shape PFC inhibition. To address the importance of early life social experiences for PFC development, we

tested the hypothesis that social play deprivation (SPD) affects cognitive flexibility and inhibitory signalling in the adult PFC in rats.

First, we assessed the impact of SPD on performance in a PFC-dependent probabilistic reversal learning task (Fig. 1a)[19,20]. In this task, responding on the 'correct' and 'incorrect' levers was rewarded on 80% and 20% of trials, respectively, and position of the 'correct' and 'incorrect' levers switched after 8 consecutive responses on the 'correct' lever. Rats in both the SPD and control (CTL) groups readily acquired the task and achieved the same performance level in terms of rewards obtained and correct lever presses (Fig. 1b-e). However, SPD rats completed more reversals (Fig. 1f,g), suggesting that CTL and SPD rats used a different strategy in this task. To understand this difference in more detail, we assessed win-stay and lose-shift percentages, which provides insight on how animals respond to positive and negative feedback (Fig. 1h,i). In SPD rats, we observed a specific increase in win-stay behaviour, at both the correct and incorrect levers (Fig. 1j,k). SPD and CTL rats responded comparably after nonrewarded trials (Fig. S1a-d). We next performed computational trial-by-trial analysis of the behavioural data[20,21] to reveal possible alterations in the component processes subserving probabilistic reversal learning (see Online Methods). The behaviour of CTL rats was best described by a Rescorla-Wagner Q-learning model throughout testing, but the SPD rats came to display behaviour congruent with a simpler heuristic strategy as reversal learning progressed (Fig. 1l, right panel), showing a tendency to remain at the previously chosen lever (Fig. 1m, S1e)). Thus, SPD rats switched from a learning-based strategy to a more perseverative strategy, whereas CTL rats continued learning throughout training. The increase in win-stay behaviour in SPD rats (Fig. 1j,k) suggests that these rats respond stronger to reward than CTL rats. Indeed, SPD rats ingested more sucrose in a preference test (Fig. S1g) and displayed higher rates of responding for sucrose under a progressive ratio schedule, but not under fixed ratio schedules of

reinforcement (Fig. S1h-i). Together, these results demonstrate that SPD results in simplified cognitive strategies and an increased responsiveness to reward in adulthood.

In order to assess the impact of SPD on PFC circuitry development, we performed voltage-clamp recordings from layer 5 (L5) pyramidal cells of the medial PFC (mPFC) in slices prepared from adult SPD and CTL rats (Fig. 2a-c). We found that the frequency and amplitude of spontaneous inhibitory postsynaptic currents (sIPSCs) was reduced in SPD rats (Fig 2d,e). The frequency of miniature inhibitory currents (mIPSCs) was also reduced in SPD slices (Fig. 2h), while mIPSC amplitudes (Fig. 2i) were not affected. By contrast, excitatory postsynaptic currents (sEPSCs) were unaffected (Fig. 2f,g). The reduction in mIPSC frequency in SPD slices was accompanied by an increase in the average rise time (Fig. 2j), suggesting a decrease in events with fast rise time. Decay kinetics (Fig. 2k) and intrinsic excitability (Fig. S2a-c) were unaffected. Together, these data indicate that SPD leads to selective decrease of GABAergic synaptic inputs onto L5 pyramidal cells in the adult mPFC, in the absence of altered postsynaptic responses to GABA release.

A reduction in mIPSCs with fast rise times suggests that inhibitory synapses at perisomatic locations were specifically affected. Perisomatic synapses are made by parvalbumin (PV) and cholecystokinin (CCK) basket cells[22], of which only the latter express the cannabinoid receptor 1 (CB1-R)[23]. We performed immunohistochemistry on the mPFC of adult SPD and CTL rats (Fig. 3a). The density of GAD67-positive interneurons was reduced in the mPFC of SPD rats (Fig. 3b), which was most pronounced in L5. The density of PV-positive cells was not altered (Fig. S3a). We then quantified inhibitory synaptic markers around the soma of L5 pyramidal neurons (Fig. 3c, S3c-g). The density of VGAT puncta was only slightly reduced, but the density of PV synapses (colocalizing with VGAT) was significantly lower in SPD tissue (Fig. 3d). The density of CB1-R synapses was not altered. Nevertheless, PV and CB1-R puncta intensity was noticeably decreased in SPD tissue (Fig. 3e). This reduction was specific as VGAT puncta intensity was similar in SPD and CTL tissue (Fig. 3e). Puncta size was not affected

(Fig. 3f). We verified that somata of L5 pyramidal cells had similar size in SPD and CTL tissue (Fig. S3b). Together, these data show that inhibitory synaptic input to L5 pyramidal cells is decreased after SPD, with differential changes in perisomatic PV and CB1-R synapses. Our findings therefore demonstrate that the organization of the adult GABAergic system in the mPFC of rats, as well as the cognitive strategy in a PFC-dependent task, is profoundly altered when rats are deprived of social play behaviour during development.

Proper maturation of cortical sensory circuits requires appropriate sensory inputs during development[1,2]. Our findings indicate that the PFC undergoes analogous experience-dependent maturation and that the experience of social play behaviour is particularly important for this process. We show cellular changes underlying altered cognitive strategies in animals deprived of juvenile social play. Importantly, the animals were socially housed after SPD, providing ample opportunity for social interactions for several weeks before testing. However, pronounced changes in PFC organisation and cognition persisted, which emphasizes the importance of social play behaviour for brain development. We observed that adult SPD rats used a simplified behavioural strategy in the probabilistic reversal learning task. SPD rats achieved more reversals, but received the same number of rewards as CTL rats by using a strategy in which they relied less on learning, and more on perseveration-based heuristics. Such heuristics are thought to be cognitively less demanding[24], and may therefore be preferred under certain conditions, especially if this does not lead to a reduction in reward. By contrast, CTL rats used a strategy whereby they integrated component processes such as sensitivity to positive and negative feedback, and weighing the benefits of exploration versus exploitation and response persistence. We and others have previously shown that component processes underlying reversal learning require functional activity in PFC regions[19,20]. Thus, the use of a simplified cognitive strategy in SPD rats indicates altered PFC functionality in these animals. Indeed, we found that perisomatic inhibition onto L5 cells was

reduced in the adult mPFC after SPD. We observed a ~30% reduction in mIPSC frequency, associated with a similar reduction in perisomatic PV synapses. A preferential effect on perisomatic inhibition by PV-expressing cells parallels observations after sensory deprivation during development[25–27]. Low levels of PV and CB1 expression suggest reduced PV cellular function[28,29] and altered endocannabinoid tone[30] , which may interfere with developmental plasticity of PFC circuitry[31] and affect cognitive capacities in adulthood[32]. Importantly and resonating well with our present findings, a reduction in PFC inhibition has previously been linked to impaired cognitive flexibility[33], and PFC PV cell activity has recently been implicated in social behavior[4,34].

In conclusion, we here shed light on how social play experience in juvenile rats – roughly equivalent to childhood in humans - contributes to PFC maturation. Social play enhances neuronal activity in a broad network of limbic and corticostriatal structures[35–37]. This integrated neuronal activity is likely to induce synaptic plasticity, but specific mechanisms remain to be identified. Our data provide important clues in this regard, showing that juvenile social play is crucial for the development of perisomatic inhibitory synapses onto L5 cells in the PFC to shape adaptive cognitive strategies.

References

1. Hensch, T. K. Critical period plasticity in local cortical circuits. *Nat. Rev. Neurosci.* **6**, 877–888 (2005).
2. Gainey, M. A. & Feldman, D. E. Multiple shared mechanisms for homeostatic plasticity in rodent somatosensory and visual cortex. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160157 (2017).
3. Larsen, B. & Luna, B. Adolescence as a neurobiological critical period for the development of higher-order cognition. *Neurosci. Biobehav. Rev.* **94**, 179–195 (2018).
4. Bicks, L. K. *et al.* Prefrontal parvalbumin interneurons require juvenile social experience to

establish adult social behavior. *Nat. Commun.* **11**, 1003 (2020).

5. Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
6. Floresco, S. B., Block, A. E. & Tse, M. T. L. Inactivation of the medial prefrontal cortex of the rat impairs strategy set-shifting, but not reversal learning, using a novel, automated procedure. *Behav. Brain Res.* **190**, 85–96 (2008).
7. Frith, C. D. & Frith, U. Mechanisms of Social Cognition. *Annu. Rev. Psychol.* **63**, 287–313 (2012).
8. Rilling, J. & Sanfey, A. The neuroscience of social decision-making. *Annu. Rev. Psychol.* **62**, 23–48 (2012).
9. Kolb, B. *et al.* Experience and the developing prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* **109 Suppl2**, 17186–17193 (2012).
10. Pellis, S. M. & Pellis, V. *The playful brain: venturing to the limits of neuroscience*. (Oneworld Publications, 2009).
11. Manduca, A. *et al.* Social play behavior, ultrasonic vocalizations and their modulation by morphine and amphetamine in Wistar and Sprague-Dawley rats. *Psychopharmacology (Berl).* **231**, 1661–1673 (2014).
12. Spinka, M., Newberry, R. C. & Bekoff, M. Mammalian play: training for the unexpected. *Q. Rev. Biol.* **76**, 141–168 (2001).
13. Vanderschuren, L. J. M. J. & Trezza, V. What the laboratory rat has taught us about social play behavior: role in behavioral development and neural mechanisms. *Curr. Top. Behav. Neurosci.* **16**, 189–212 (2014).
14. Bell, H. C., McCaffrey, D. R., Forgie, M. L., Kolb, B. & Pellis, S. M. The role of the medial prefrontal cortex in the play fighting of rats. *Behav. Neurosci.* **123**, 1158–1168 (2009).
15. van Kerkhof, L. W., Damsteegt, R., Trezza, V., Voorn, P. & Vanderschuren, L. J. Social Play

Behavior in Adolescent Rats is Mediated by Functional Activity in Medial Prefrontal Cortex and Striatum. *Neuropsychopharmacology* **38**, 1899–1909 (2013).

16. Baarendse, P. J. J., Counotte, D. S., O'Donnell, P. & Vanderschuren, L. J. M. J. Early social experience is critical for the development of cognitive control and dopamine modulation of prefrontal cortex function. *Neuropsychopharmacology* **38**, 1485–1494 (2013).

17. Bell, H. C., Pellis, S. M. & Kolb, B. Juvenile peer play experience and the development of the orbitofrontal and medial prefrontal cortices. *Behav. Brain Res.* **207**, 7–13 (2010).

18. Turrigiano, G. G. & Nelson, S. B. Homeostatic plasticity in the developing nervous system. *Nat. Rev. Neurosci.* **5**, 97–107 (2004).

19. Dalton, G. L., Wang, N. Y., Phillips, A. G. & Floresco, S. B. Multifaceted Contributions by Different Regions of the Orbitofrontal and Medial Prefrontal Cortex to Probabilistic Reversal Learning. *J. Neurosci.* **36**, 1996–2006 (2016).

20. Verharen, J. P. H., Ouden, H. E. M. den, Adan, R. A. H. & Vanderschuren, L. J. M. J. Modulation of value-based decision making behavior by subregions of the rat prefrontal cortex. *Psychopharmacology (Berl).* doi: 10.1007/s00213-020-05454-7 (2020).

21. Verharen, J. P. H. *et al.* A neuronal mechanism underlying decision-making deficits during hyperdopaminergic states. *Nat. Commun.* **9**, 1–15 (2018).

22. Whissell, P. D., Cajanding, J. D., Fogel, N. & Kim, J. C. Comparative density of CCK- and PV-GABA cells within the cortex and hippocampus. *Front. Neuroanat.* **9**, 1–16 (2015).

23. Katona, I. *et al.* Presynaptically located CB1 cannabinoid receptors regulate GABA release from axon terminals of specific hippocampal interneurons. *J. Neurosci.* **19**, 4544–4558 (1999).

24. Christie, S. T. & Schrater, P. Cognitive cost as dynamic allocation of energetic resources. *Front. Neurosci.* **9**, 1–15 (2015).

25. Jiao, Y., Zhang, C., Yanagawa, Y. & Sun, Q. Q. Major effects of sensory experiences on the

neocortical inhibitory circuits. *J. Neurosci.* **26**, 8691–8701 (2006).

26. Mowery, T. M. *et al.* Preserving inhibition during developmental hearing loss rescues auditory learning and perception. *J. Neurosci.* **39**, 8347–8361 (2019).

27. Hensch, T. *et al.* Local GABA circuit control of experience-dependent plasticity in developing visual cortex. *Science (80-.).* **282**, 1504–1508 (1998).

28. Donato, F., Rompani, S. B. & Caroni, P. Parvalbumin-expressing basket-cell network plasticity induced by experience regulates adult learning. *Nature* **504**, 272–276 (2013).

29. Caballero, A., Flores-Barrera, E., Cass, D. K. & Tseng, K. Y. Differential regulation of parvalbumin and calretinin interneurons in the prefrontal cortex during adolescence. *Brain Struct. Funct.* **219**, 395–406 (2014).

30. Sciolino, N. R. *et al.* Social isolation and chronic handling alter endocannabinoid signaling and behavioral reactivity to context in adult rats. *Neuroscience* **168**, 371–386 (2010).

31. Caballero, A. & Tseng, K. Y. GABAergic Function as a Limiting Factor for Prefrontal Maturation during Adolescence. *Trends Neurosci.* **39**, 441–448 (2016).

32. Donato, F., Chowdhury, A., Lahr, M. & Caroni, P. Early- and Late-Born Parvalbumin Basket Cell Subpopulations Exhibiting Distinct Regulation and Roles in Learning. *Neuron* **85**, 770–786 (2015).

33. Gruber, A. J. *et al.* More is less: a disinhibited prefrontal cortex impairs cognitive flexibility. *J. Neurosci.* **30**, 17102–17110 (2010).

34. Sun, Q. *et al.* Ventral Hippocampal-Prefrontal Interaction Affects Social Behavior via Parvalbumin Positive Neurons in the Medial Prefrontal Cortex. *iScience* **23**, 100894 (2020).

35. van Kerkhof, L. W. M. *et al.* Cellular activation in limbic brain systems during social play behaviour in rats. *Brain Struct. Funct.* **219**, 1181–1211 (2013).

36. Manduca, A. *et al.* Dopaminergic neurotransmission in the nucleus accumbens modulates social play behavior in rats. *Neuropsychopharmacology* **41**, 2215–2223 (2016).

37. Manduca, A. *et al.* Interacting cannabinoid and opioid receptors in the nucleus accumbens core control adolescent social play. *Front. Behav. Neurosci.* **10**, 1–17 (2016).

38. Bari, A. *et al.* Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology* **35**, 1290–301 (2010).

39. Rigoux, L., Stephan, K. E., Friston, K. J. & Daunizeau, J. Bayesian model selection for group studies - Revisited. *Neuroimage* **84**, 971–985 (2014).

40. Penny, W. D. *et al.* Comparing families of dynamic causal models. *PLoS Comput. Biol.* **6**, (2010).

41. Rescorla, R. A. & Wagner, A. R. *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement, Classical Conditioning II*. (Appleton-Century-Crofts., 1972).

42. Boekhoudt, L. *et al.* Enhancing excitability of dopamine neurons promotes motivational behaviour through increased action initiation. *Eur. Neuropsychopharmacol.* **28**, 171–184 (2018).

Methods

Animals and housing conditions

All experiments were approved by the Animal Ethics Committee of the Utrecht University and were conducted in agreement with Dutch laws (Wet op de Dierproeven, 1996) and European regulations (Guideline 86/609/EEC). Male Lister Hooded rats were obtained from Charles River (Germany) on postnatal day (P) 14 in litters with nursing mothers. All rats were subject to a reversed 12:12h light-dark cycle with ad libitum access to water and food. Rats were weaned on P21 and were either subjected to the social play deprivation (SPD) group or the control (CTL) group. CTL rats were housed together in groups of four during P21-42. On P42, CTL rats were housed in pairs for the remainder of the experiment. SPD rats were pair-housed with a rat from a different mother. During P21 to P42 a transparent Plexiglas divider containing small holes was placed in the middle of their home cage creating two separate but identical compartments. SPD rats were therefore able to see, smell and hear one another but they were unable to physically socially engage. On P42, the Plexiglas divider was removed and SPD rats were housed in pairs for the remainder of the experiment. All rats were housed in pairs for at least 4 weeks until early adulthood (10 weeks of age) after which experimentation began. All experiments were conducted during the active phase of the animals (10:00 - 17:00). One week before the start of behavioural testing, the rats were subjected to food-restriction and were maintained at 85% of their free-feeding weight for the duration of the behavioural experiment. Rats were provided with ~ 20 sucrose pellets (45mg, BioServ) in their home cage for two subsequent days before their first exposure to the operant conditioning chamber to reduce potential food neophobia. Rats were weighed and handled at least once a week throughout the course of the experiment.

Probabilistic reversal learning

Apparatus: Behavioural testing was conducted in operant conditioning chambers (Med Associates, USA) enclosed in sound-attenuating cubicles equipped with a ventilation fan. Two retractable levers were located on either side of a central food magazine into which sugar pellets could be delivered via a dispenser. A LED cue light was located above each retractable lever. A white house light was mounted in the top-centre of the wall opposite the levers. Online control of the apparatus and data collection was performed using MED-PC (Med Associates) software.

Pre-training: Rats were first habituated once to the operant chamber for 30 min in which the house light was illuminated and 50 sucrose rewards were randomly delivered into the magazine with an average inter-trial interval of 15 s between reward deliveries. On the subsequent days, the rats were trained for 30 min under a Fixed-Ratio 1 (FR1) schedule of reinforcement for a minimum of three consecutive sessions. The session started with the illumination of the house light and the insertion of both levers, which remained inserted for the remainder of the session. One of the two levers was the 'correct' lever rendering a reward when pressed, whereas pressing the other lever had no consequences. There was no limit other than time on the amount of times a rat could press the 'correct' lever. If the rat obtained 50 or more rewards in a session it was required to press the other lever the following day. If it obtained less than 50 rewards the rat was tested on the same schedule the next day. To proceed to (probabilistic) reversal learning pre-training, the rat had to obtain an average of at least 50 rewards over three completed sessions. In case a rat obtained a lower number of rewards during the first three sessions, it was further trained on subsequent days until the criterion was met. A trial started with an inter-trial-interval (ITI) of 5 s with the chamber in darkness, followed by the illumination of the house-light and the insertion of one of the two levers into the chamber. A response within 30 s on the inserted lever resulted in the delivery of a reward. If the rat failed to respond on the lever within 30 s, the lever retracted and the trial was scored as an omission. Rats were trained for $\sim$ 2-3 days to a criterion of at least 50 rewards and had to perform a lever press in more than 80% of the trials. In the subsequent

sessions, the rats were familiarized with the probabilistic nature of the probabilistic reversal learning task, in which a lever press resulted in 80% reward delivery instead of 100%. Rats were trained for ~ 3-4 d to a criterion of at least 50 rewards and had to perform a lever press in more than 80% of the trials before progressing to the probabilistic reversal learning.

Probabilistic reversal learning: The protocol used for this task was modified from those of previous studies [19,20,38]. At the start of each session one of the two levers was randomly selected to be 'correct' and the other 'incorrect'. A response on the 'correct' lever resulted in the delivery of a reward on 80% of the trials, whereas a response on the 'incorrect' lever was reinforced on 20% of trials. Each trial started with a 5 s ITI, followed by the illumination of the house-light and the insertion of both levers into the chamber. After a 'correct' response, both levers retracted but the house light remained illuminated. In case the rat was rewarded, the house light remained illuminated, whereas the house light extinguished in case the rat was not rewarded on the 'correct' lever. An 'incorrect' response or a failure to respond within 30 s after lever insertion (i.e. omission) lead to the retraction of both levers, extinction of the house light so that the chamber returned to its ITI state. When the rat made a string of 8 consecutive trials on the 'correct' lever (regardless of whether they were rewarded or not), contingencies were reversed, meaning the 'correct' lever became the 'incorrect' lever and the previously 'incorrect' lever became the 'correct' lever. This pattern repeated over the course of a daily session. Daily sessions were completed upon performing 200 trials or after 60 minutes have passed, whichever occurred first. Daily sessions were completed upon performing 200 trials or after 60 minutes have passed, whichever occurred first. The number of reversals made during a session was solely limited by the number of trials (200) in a session. The total time to finish a session was not different for SPD and CTL rats (CTL: 29.4 ± 0.8 minutes vs SPD:30.5 ± 0.7 minutes; $p = 0.15$, T test).

Trial-by-trial analysis: This analysis was performed to assess the shifts in choice behaviour between subsequent trials, in order to investigate the sensitivity to positive and negative feedback. Depending on

whether the rat received a reward or not, it can press the same lever on the subsequent trial or shift towards the other lever, resulting in 4 different possibilities (i.e. win-stay, win-shift, lose-stay, lose-shift) for both the 'correct' and 'incorrect' lever. First, the number of times that each possibility occurred was summed per session and was subsequently calculated in percentages. For example, lose-shift behaviour after a press on the incorrect lever was calculated by dividing the number of shifts towards the other lever upon a loss on the incorrect lever by the total number of losses on the incorrect lever, multiplied by 100.

Computational Modelling

Eight different computational models were fit to the trial-by-trial data to assess differences in task strategy between both groups of animals. Best-fit model parameters were estimated using maximum likelihood estimation, using Matlab (version 2018b; The MathWorks Inc.) function 'fmincon'[21]. These maximum likelihood estimates were corrected for model complexity (i.e., the number of free parameters (n_f)) by calculating the Akaike information criterion (AIC) for each session:

$$AIC = 2*(n_f) - 2*\log(likelihood)$$

In which a lower AIC indicates more evidence in favor of the model. These log-model evidence estimates were subsequently used to perform Bayesian model selection[39] using the Matlab package SPM12 (The Wellcome Centre for Human Neuroimaging), taking into account the family to which each model belonged[40]. This yielded the protected exceedance probability for each family of models (Fig. 1L) and for each individual model (Fig. S1E), indicating the probability that each of the (family of) models was more prevalent among the group of rats than the others.

Table 1 contains an overview of the eight computational models. The random choice model is the null model, which assumes that animals choose randomly (i.e., p = 0.5 for each choice, so that the log likelihood is given by [number of trials]*log(0.5)). The second family of models contained strategies based on 'heuristics'; simple strategies to complete the task. The third family contained Q learning models, consisting of four derivatives of the Rescorla-Wagner model[41].

Regardless of value function, all models (except the null model) included a Softmax function to compute the probability of choosing left or right, scaled by the Softmax' inverse temperature β (which was a free parameter in the models). For example, the chance of choosing the right lever in trial t is given by:

$$p_{right,t} = \frac{e^{\beta \times Q_{right,t}}}{e^{\beta \times Q_{left,t}} + e^{\beta \times Q_{right,t}}}$$

In which $Q_{s,t}$ is the value of lever s in trial t. A high value of β indicates consistent choice for the highest valued option, while a low β indicates random choice behavior. Values of β were not different between CTL and SPD rats (See Suppl Table 2).

Table 1 – Overview of computational models

Model family	**#**	**Model name**	**n_f**	**Description**
Random choice	1	Random choice model	0	Animal choses randomly
Heuristics family	2	Win-Stay, Lose-Switch	1	Animal stays on the same lever after winning, moves away from the lever after a loss ;

				i.e., after a 'win' on the right lever, Q_{left} is set to 1, and Q_{right} is set to 1.
	3	Win-Stay, Lose-Random	1	Animal stays on the same lever after winning, randomly picks a lever after a loss; after a loss, both Q values are set to 0, but after a win, Q value of the chosen side becomes 1.
	4	Random choice + stickiness	2	Animal choses randomly but attributes value to the lastly chosen lever [20].
Q learning family	5	Q learning, single learning rate for reward and punishment learning	2	Animal learns from previous decisions. Learning rate for positive and negative feedback are the same [20].
	6	Q learning, separate learning rates for reward and punishment learning	3	Animal learns from previous decisions. Learning rate for positive and negative feedback may differ [20].
	7	Model 6 + stickiness parameter	4	Animal learns from previous decisions. Learning rate for positive and negative feedback are separately calculated. Additionally, the animal attributes value to the lastly chosen lever [20].
	8	Rescorla-Wagner Pearce Hall	4	Animal learns from previous decisions. Learning rate for positive and negative feedback are the

				same. Additionally, the animal may learn better when task volatility is higher (e.g., after a reversal).

Open field, sucrose preference and reinforcement learning tasks

After completion of testing for probabilistic reversal learning, the animals were tested in an open field to investigate locomotion in a novel environment, and for differences in their relative preference for sucrose over water. The open field and sucrose preference tests were replicated in a separate set of 12 CTL and 12 SPD rats, since the data from these two experiments did not significantly differ, the data were collapsed for analysis. Responding for sucrose was tested in a separate set of 12 CTL and 12 SPD rats. *Open field:* The open field consisted of a rectangular grey PVC box (50 x 50 cm; 100 cm height). Each rat was placed in the open field for a duration of 10 minutes. Behaviour was recorded using a video camera and analysed using the video tracking software (Ethovision, Noldus, Wageningen). *Sucrose preference:* The rats were habituated to the testing cages for 2-3 h and received access to two water bottles for a period of 4 consecutive days, with one bottle containing increasing concentrations of sucrose over days (0.0%, 0.1%, 0.5% and 1%). The bottles were weighed prior to and after 2h and 6h of access in each session that started at 9:00 AM. Each concentration was offered once and bottle positions were switched between sessions to avoid development of side-bias. Daily consumption (ml) and the rat's body weight was used to calculate sucrose intake (ml/kg), water intake (ml/kg) and total consumption (ml/kg). Sucrose preference was calculated as the percentage of sucrose consumption of the total fluid intake ((sucrose intake/total intake)*100). *Responding for sucrose:* The rats were habituated to the operant conditioning chambers (Med Associated) after which lever pressing sessions started under a fixed ratio (FR) schedule of reinforcement. During the first three consecutive days, rats were tested under a FR-1 schedule, in which each active lever press resulted in one reward. In the following days the rats were tested on a FR-2, FR-5 and FR-10 schedule of reinforcement (with 2, 5 or 10

lever presses needed for one reward). After acquisition of the FR tasks, the rats were tested in three sessions under a progressive ratio (PR) schedule of reinforcement. Under this schedule, the animals had to meet a response requirement on the active lever that progressively increased with three after every earned reward (1, 2, 4, 6, 9, 12, 15, 20, 25, etc.)[42].

Electrophysiological analysis

Slice preparation: CTL and SPD rats (12-15 weeks of age) were anesthetized by intraperitoneal injection of sodium pentobarbital and then transcardially perfused with ice-cold modified artificial cerebrospinal fluid (ACSF) containing (in mM): 92 N-methyl-D-glucamine (NMDG), 2.5 KCl, 1.25 NaH_2PO_4, 30 $NaHCO_3$, 20 HEPES, 25 glucose, 2 thiourea, 5 Na-ascorbate, 3 Na-pyruvate, 0.5 $CaCl_2.4H_2O$, and 10 $MgSO_4.7H_2O$, bubbled with 95% O_2 and 5% CO_2 (pH 7.3–7.4). Coronal slices of the medial PFC (300 µm) were prepared using a vibratome (Leica VT1200S, Leica Microsystems) in ice-cold modified ACSF. Slices were initially incubated in the carbogenated modified ACSF for 10 min at 34 °C and then transferred into a holding chamber containing standard ACSF containing (in mM): 126 NaCl, 3 KCl, 2 $MgSO_4$, 2 $CaCl_2$, 10 glucose, 1.25 NaH_2PO_4 and 26 $NaHCO_3$ bubbled with 95% O_2 and 5% CO_2 (pH 7.3) at room temperature for at least 1 hour. They were subsequently transferred to the recording chamber, perfused with standard ACSF that is continuously bubbled with 95% O_2 and 5% CO_2 at 30–32°C.

Whole-cell recordings and analysis: Whole-cell patch-clamp recordings were performed from layer 5 pyramidal neurons in the medial PFC. These neurons were visualized with an Olympus BX61W1 microscope using infrared video microscopy and differential interference contrast (DIC) optics. Patch electrodes were pulled from borosilicate glass capillaries and had a resistance of 3-5 MΩ when filled with intracellular solutions. Excitatory postsynaptic currents (EPSCs) were recorded in the presence of bicuculline (10 µM) and with internal solution containing (in mM): 140 K-gluconate, 1 KCl, 10 HEPES, 0.5 EGTA, 4 MgATP, 0.4 Na_2GTP, 4 phosphocreatine (pH 7.3 with KOH). Inhibitory postsynaptic currents

(IPSCs) were recorded in the presence of 6-cyano-7- nitroquinoxaline-2,3-dione (CNQX) (10 µM) and D,L-2-amino-5-phosphopentanoic acid (D,L-AP5) (20 µM), with internal solution containing (in mM): 70 K-gluconate, 70 KCl, 10 HEPES, 0.5 EGTA, 4 MgATP, 0.4 Na_2GTP, 4 phosphocreatine (pH 7.3 with KOH) or 125 CsCl, 2 $MgCl_2$, 5 NaCl, 10 HEPES, 0.2 EGTA, 4 MgATP, 0.4 Na_2GTP (pH 7.3 with CsOH). Action-potential independent miniature IPSCs (mIPSCs) were recorded under the same conditions, but in the presence of 1 µM tetrodotoxin (TTX; Sigma) to block sodium channels. The membrane potential was held at -70 mV for voltage-clamp experiments. Signals were amplified, filtered at 3 kHz and digitized at 10 kHz using an EPC-10 patch-clamp amplifier and PatchMaster v2x73 software. Series resistance was constantly monitored, and the cells were rejected from analysis if the resistance changed by >20%. No series resistance compensation was used. Resting membrane potential was measured in bridge mode (I=0) immediately after obtaining whole-cell access. The basic electrophysiological properties of the cells were determined from the voltage responses to a series of hyperpolarizing and depolarizing square current pulses. Input resistance was determined by the slope of the linear regression line through the voltage-current curve. Passive and active membrane properties were analysed with Clampfit 10 (Axon Instrument) and synaptic currents were analysed with Mini Analysis (Synaptosoft) software. Miniature and spontaneous synaptic currents (IPSCs and EPSCs) data were analyzed with Mini Analysis (Synaptosoft Inc., Decatur, GA). All events were detected with a criterion of a threshold >3× root-mean-square (RMS) of baseline noise. The detected currents were manually inspected to exclude false events.

Immunohistochemistry

Tissue preparation: Rats were anesthetized with Nembutal (i.p. 240mg/kg) and transcardially perfused with 0.1 M phosphatebuffered saline (PBS, pH 7.3-7.4) followed by 4% paraformaldehyde in 0.01 M PBS. The brains were removed from the skull and post-fixed overnight in the same paraformaldehyde solution at 4◦C and subsequently cryoprotected in 30% sucrose for three days at 4◦C. Thereafter, the

brains were rapidly frozen in aluminium foil on dry ice and stored at −80◦C until further use. Brain sections (20 µm thick) from the prefrontal cortex (PFC) between Bregma levels of 4.2 - 2.2mm were made with a Cryostat Leica CM 3050 S. Sections were stored at -80°C until immunohistochemistry was performed. Brain slices were thawed and let dry for 1h at room temperature (RT). Slices were washed in PBS three times for 15 min (3x15min) at RT. Sections were cooked in sodium citric acid buffer (SCAB, 10mM sodium citric acid in demi water, pH6) for 10 min at 97°C in a temperature controlled microwave, cooled for 30 min at 4°C and washed again (3x15min in PBS). Slices were blocked with 400 µl of blocking buffer (10% normal goat-serum, 0,2% triton-X 100 in PBS) for 2h in a wet chamber at RT. Slices were incubated overnight at 4°C in the wet chamber with 250 µl of primary antibodies in blocking buffer. Primary antibodies and their concentrations used can be found in Table 2 (for the interneuron density analysis) and Table 3 (for the synapse analysis). Sections were washed (3x15min in PBS), followed by incubation with the secondary antibodies in blocking buffer for 2h at RT in a wet chamber. Secondary antibodies and their concentrations used can be found in Table 4 (for the interneuron density analysis) and Table 5 (for the synapse analysis). After another wash step (3x15min in PBS), slides were mounted and stored at 4°C until image acquisition.

Table 2 - Primary antibodies for interneuron analysis

Host	**Epitope**	**Concentration**	**Company**	**Order number**
Rat	Ctip2	1:1000	Abcam	ab18465
Rabbit	PV	1:1000	Life technologies	PA1933
Mouse IgG1	NeuN	1:500	Millipore	MAB377
Mouse IgG2a	GAD67	1:500	Millipore	MAB5406

Table 3 - Secondary antibodies for interneuron analysis

Host	Epitope	Fluorophore	Concentration	Company	Order number
Goat	Anti-rat	Alexa Fluor 568	1:500	Life technologies	A11077
Goat	Anti-rabbit	Alexa Fluor 405	1:500	Life technologies	A31556
Goat	Anti-mouse IgG1	Alexa Fluor 647	1:500	Life technologies	A21240
Goat	Anti-mouse IgG2a	Alexa Fluor 488	1:500	Life technologies	A21131

Table 4 - Primary antibodies for synapse analysis

Host	Epitope	Concentration	Company	Order number
Chicken	VGAT	1:1000	Synaptic systems	131006
Rabbit	PV	1:1000	Life technologies	PA1933
Guinea pig	NeuN	1:500	Millipore	ABN90
Mouse	CB1-R	1:1000	Synaptic systems	258011

Table 5 - Secondary antibodies for synapse analysis

Host	Epitope	Fluorophore	Concentration	Company	Order number
Goat	Anti-chicken	Alexa Fluor 647	1:500	Life technologies	A21449
Goat	Anti-rabbit	Alexa Fluor 405	1:500	Life technologies	A31556
Goat	Anti-guinea pig	Alexa Fluor 568	1:500	Life technologies	A11075
Goat	Anti-mouse	Alexa Fluor 488	1:500	Life technologies	A11029

Image acquisition and analysis: For both interneuron density and synapse analysis, images of the brain slices were taken with a Zeiss Confocal microscope (type LSM700). The investigator was blinded to the groups of the sections when acquiring the images and performing the quantifications. Image analysis was performed in ImageJ (National Institutes of Health USA).

Inhibitory cell density analysis: z-stacks were acquired at 20x of all layers of the mPFC. Tile scan z-stacks (1600 x 1280 μm^2, 2 µm steps, total of 10 µm) were acquired of the mPFC in both hemispheres of control (n=6) and SPD (n=6) rats. Antibodies staining for NeuN, Ctip2, GAD67 and PV were used. NeuN (neuronal nuclei) is a nuclear protein specific for neurons and was used as a marker to distinguish neurons specifically. Ctip2 (CtBP (C-terminal binding protein) interacting protein) shows mPFC layer specific expression, which allowed for the identification of the cortical layers (Figure 3A). GAD67 (glutamate decarboxylase), a GABA synthesis enzyme expressed in all inhibitory interneurons, was used to select and determine the inhibitory interneuron density. PV+ inhibitory interneurons were counted to determine the density of the PV+ subtype. Density was determined based on the number of neurons within a certain area (μm^2).

Synapse analysis: z-stacks were acquired at 63x in layer 5 of the mPFC. For each of the rat brains (Control (n=6), SPD (n=6)) z-stacks (102 x 102 μm^2, 0.4 µm steps, total of 12 µm) were acquired in both hemispheres. Antibodies are listed in tables 4&5.Image analysis was performed semi-automatically using custom-written ImageJ macros and MATLAB scripts. NeuN was used to determine the perisomatic outline of the L5 somata. VGAT (vesicular GABA transporter) is expressed in all inhibitory synapses and was used as a general inhibitory synapse marker. PV and CB1-R antibodies were used to select the synapses of their respective subtypes. For each L5 cell, a maximum intensity image was constructed from 4 z-stack slices and everything outside of a 1.5 µm band around the NeuN outline (Fig. S3E) was cleared. The image was then median filtered (Median filter 2.0), and thresholded (Fig. S3G),. Particles larger than 0.2µm and circularity of 0.6-1.0 were selected. PV and CB1-R puncta were only included when col-localized with VGAT.

Data processing and statistical analyses

Statistical analyses were performed with GraphPad Prism (Software Inc.) and RStudio 1_2_5019 (R version 3.6.1, R Foundation for Statistical Computing). Normality of the data was tested with a Shapiro-Wilk test. Differences between two groups were then tested with a nonparametric Mann-Whitney-Wilcoxon test (MW), or a parametric Welch t test (T). Two-way repeated ANOVA (for sessions and condition) was used for multiple comparisons followed by T-tests (with Bonferroni correction). Detailed statistical information of the figures is listed in Supplementary Table 2. All graphs represent the mean ± standard error of the mean (SEM) with individual data points shown in colored circles. Statistical range in all figures: *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$.

Data and code availability

The datasets generated during the current study and custom-written MATLAB scripts are available from the corresponding authors on reasonable request.

Supplementary Table 1. Model parameters for the two most prevalent models

Day	Winning model CTL	Winning model SPD	Best-fit parameters model 4 (random + stickiness)			Best-fit parameters model 7 (Q learning dual + stickiness)		
			CTL	SPD	T-test	CTL	SPD	T-test
1	5	5	π = 0.140 β = 4.763	π = 0.256 β = 4.842	P = 0.027 P = 0.651	α+ = 0.111 α- = 0.120 π = -0.012 β = 6.021	α+ = 0.050 α- = 0.054 π = 0.111 β = 4.939	P = 0.472 P = 0.448 P = 0.085 P = 0.510
2	5	6	π = 0.151 β =4.426	π = 0.315 β = 4.435	P = 0.035 P = 0.981	α+ = 0.122 α- = 0.129 π = -0.033 β = 5.943	α+ = 0.056 α- = 0.061 π = 0.119 β = 5.667	P = 0.438 P = 0.475 P = 0.072 P = 0.867
3	5	5	π = 0.244 β = 4.925	π = 0.400 β = 4.327	P = 0.013 P = 0.087	α+ = 0.077 α- = 0.120 π = 0.046 β = 5.273	α+ = 0.139 α- = 0.243 π = 0.215 β = 5.523	P = 0.500 P = 0.341 P = 0.011 P = 0.869
4	7	7	π = 0.343 β = 4.431	π = 0.564 β = 4.901	P = 0.014 P = 0.011	α+ = 0.057 α- = 0.087 π = 0.158 β = 4.498	α+ = 0.268 α- = 0.419 π = 0.391 β = 4.152	P = 0.071 P = 0.009 P = 0.015 P = 0.784
5	5	7	π = 0.360 β = 4.754	π = 0.519 β = 4.991	P = 0.089 P = 0.080	α+ = 0.091 α- = 0.107 π = 0.164 β = 4.656	α+ = 0.216 α- = 0.355 π = 0.327 β = 4.918	P = 0.224 P = 0.054 P = 0.073 P = 0.860
6	5	7	π = 0.308 β = 4.733	π = 0.581 β = 4.827	P = 0.007 P = 0.695	α+ = 0.069 α- = 0.193 π = 0.087 β = 5.033	α+ = 0.286 α- = 0.473 π = 0.409 β = 4.749	P = 0.047 P = 0.060 P = 0.002 P = 0.852
7	6	4	π = 0.348 β = 4.755	π = 0.647 β = 4.349	P = 0.001 P = 0.261	α+ = 0.099 α- = 0.207 π = 0.140 β = 5.098	α+ = 0.471 α- = 0.497 π = 0.396 β = 4.247	P = 0.007 P = 0.052 P = 0.009 P = 0.551
8	7	7	π = 0.523 β = 4.765	π = 0.814 β = 4.584	P = 0.018 P = 0.358	α+ = 0.247 α- = 0.378 π = 0.298 β = 2.999	α+ = 0.363 α- = 0.578 π = 0.530 β = 3.273	P = 0.350 P = 0.201 P = 0.064 P = 0.799
9	7	4	π = 0.614 β = 4.850	π = 0.801 β = 4.757	P = 0.072 P = 0.590	α+ = 0.318 α- = 0.303 π = 0.394 β = 4.100	α+ = 0.503 α- = 0.689 π = 0.449 β = 2.311	P = 0.163 P = 0.007 P = 0.175 P = 0.593
10	7	4	π = 0.586 β = 4.838	π = 0.852 β = 4.681	P = 0.015 P = 0.386	α+ = 0.317 α- = 0.337 π = 0.330 β = 3.116	α+ = 0.380 α- = 0.595 π = 0.543 β = 3.978	P = 0.642 P = 0.042 P = 0.057 P = 0.490
average			**π = 0.36 β = 4.7**	**π = 0.58 β = 4.7**	**P < 0.001 P = 0.475**	**α+ = 0.15 α- = 0.20 π = 0.16 β = 4.7**	**α+ = 0.27 α- = 0.40 π = 0.35 β = 4.4**	**P = 0.013 P = 0.002 P < 0.001 P = 0.496**

Best-fit model parameters (maximum likelihood estimates) for the two models that are most prevalent among the rats (models 4 and 7; model 5 is nested in model 7 and was thus not given). Estimate of stickiness parameter π is consistently higher in SPD rats in all sessions.

Supplementary Table 2. Statistical parameters

	DESCRIPTION	N		NORMALITY TEST	P-VALUE		STATISTICAL TEST	DF/T/F/W/ETC…	P-VALUE	
FIG 1		CTL	SPD		CTL	SPD				
	B	12	12	Shapiro–Wilk test	0.6871	0.09222	Welch's T-Test	t(21.974) = -1.434	0.1656	
	C	12	12				Two way repeated Anova	F(9,220) = 7.091 (Session)	<0.0005	***
								F(1,220) = 3.476 (Condition)	0.0636	
								F(9,220) = 0.667 (Interaction)	0.7380	
	D COR: CTL vs SPD	12	12	Shapiro–Wilk test	0.07551	0.9179	Welch's T-Test	t(21.784) = -1.3728	0.1838	
	D INC: CTL vs SPD	12	12	Shapiro–Wilk test	0.1164	0.05508	Welch's T-Test	t(21.648) = 2.063	0.05131	
	D CTL: COR vs INC	12	12	Shapiro–Wilk test	0.07551	0.1164	Welch's T-Test	t(21.971) = 25.258	<0.0005	***
	D SPD: COR vs INC	12	12	Shapiro–Wilk test	0.05508	0.9179	Welch's T-Test	t(21.239) = 28.809	<0.0005	***
	E COR: CTL vs SPD	12	12				Two way repeated Anova	F(9,220) = 9.374 (Session)	<0.0005	***
								F(1,220) = 4.900 (Condition)	0.0279	*
								F(9,220) = 0.802 (Interaction)	0.6146	
	E INC: CTL vs SPD	12	12				Two way repeated Anova	F(9,220) = 13.882 (Session)	<0.0005	***
								F(1,220) = 8.439 (Condition)	0.00405	**
								F(9,220) = 0.921 (Interaction)	0.50781	
	F	12	12	Shapiro–Wilk test	0.6932	0.1202	Welch's T-Test	t(19.661) = -3.8016	0.001146	**
	G	12	12				Two way repeated Anova	F(9,220) = 16.30 (Session)	<0.0005	***
								F(1,220) = 31.76 (Condition)	<0.0005	***
								F(9,220) = 0.47 (Interaction)	0.894	
	H	12	12	Shapiro–Wilk test	0.7974	0.008435	Mann-Whitney-Wilcoxon Test	W = 9	<0.0005	***
	I	12	12	Shapiro–Wilk test	0.7941	0.5805	Welch's T-Test	t(20.39) = 0.16529	0.8703	
	J	12	12				Two way repeated Anova	F(9,220) = 22.109 (Session)	<0.0005	***
								F(1,220) = 83.739 (Condition)	<0.0005	***
								F(9,220) = 1.136 (Interaction)	0.338	
	K	12	12				Two way repeated Anova	F(9,220) = 1.903 (Session)	0.052663	
								F(1,220) = 13.792 (Condition)	0.000259	***
								F(9,220) = 1.444 (Interaction)	0.170808	

	Description	N		Normality test	P-Value		Statistical test	df/t/F/W/etc…	P-Value	
Fig 2		CTL	SPD		CTL	SPD				
	D	15	12	Shapiro–Wilk test	0.1843	0.8677	Welch's T-Test	t(15.835) = 3.6075	0.002396	**
	E	15	12	Shapiro–Wilk test	0.1787	0.4352	Welch's T-Test	t(22.266) = 2.317	0.03008	*
	F	14	13	Shapiro–Wilk test	0.6275	0.4761	Welch's T-Test	t(24.983) = 1.1342	0.2675	
	G	14	13	Shapiro–Wilk test	0.978	0.261	Welch's T-Test	t(18.064) = -0.95029	0.3545	
	H	13	15	Shapiro–Wilk test	0.8291	0.4944	Welch's T-Test	t(25.561) = 3.9494	0.0005464	**
	I	13	15	Shapiro–Wilk test	0.003542	0.4568	Welch's T-Test	t(21.061) = 0.68133	0.5031	
	J	13	15	Shapiro–Wilk test	0.9864	0.3444	Welch's T-Test	t(22.061) = -3.8678	0.0008288	**
	K	13	15	Shapiro–Wilk test	0.1929	0.3649	Welch's T-Test	t(25.46) = -0.85983	0.3979	

	Description	N		Normality test	P-Value		Statistical test	df/t/F/W/etc…	P-Value	
Fig 3		CTL	SPD		CTL	SPD				
	B overall	12	12	Shapiro–Wilk test	0.7391	0.1259	Welch's T-Test	t(21.793) = -2.4967	0.02059	*
	B L1	12	12	Shapiro–Wilk test	0.5039	0.9563	Welch's T-Test	t(20.028) = -2.0206	0.0569	
	B L2	12	12	Shapiro–Wilk test	0.9914	0.03668	Mann-Whitney-Wilcoxon Test	W = 44	0.1135	
	B L3	12	12	Shapiro–Wilk test	0.9367	0.07767	Welch's T-Test	t(20.977) = -0.81354	0.425	
	B L5	12	12	Shapiro–Wilk test	0.3553	0.7425	Welch's T-Test	t(19.999) = -3.3557	0.003147	**
	B L6	12	12	Shapiro–Wilk test	0.9757	0.2855	Welch's T-Test	t(21.838) = 0.28213	0.7805	
	D VGAT	49	34	Shapiro–Wilk test	0.1081	0.9815	Welch's T-Test	t(66.199) = 1.3243	0.19	
	D PV	44	31	Shapiro–Wilk test	0.12553	0.254	Welch's T-Test	t(72.209) = 2.4129	0.01837	*
	D CB1R	46	31	Shapiro–Wilk test	0.0004908	0.003577	Mann-Whitney-Wilcoxon Test	W = 624	0.3601	
	E VGAT	49	34	Shapiro–Wilk test	0.6378	0.01052	Mann-Whitney-Wilcoxon Test	W = 904	0.5139	
	E PV	44	31	Shapiro–Wilk test	0.04778	0.169	Mann-Whitney-Wilcoxon Test	W = 1032	<0.0005	***
	E CB1R	46	31	Shapiro–Wilk test	0.01128	0.03466	Mann-Whitney-Wilcoxon Test	W = 940	0.01864	*
	F VGAT	49	34	Shapiro–Wilk test	0.07667	0.01047	Mann-Whitney-Wilcoxon Test	W = 822.5	0.9262	
	F PV	44	31	Shapiro–Wilk test	0.8942	0.3963	Welch's T-Test	t(62.624) = -1.621	0.11	
	F CB1R	46	31	Shapiro–Wilk test	0.3503	0.0005745	Mann-Whitney-Wilcoxon Test	W = 523	0.04902	*

	Description	N		Normality test	P-Value		Statistical test	df/t/F/W/etc…	P-Value	
Fig S1		CTL	SPD		CTL	SPD				
	A	12	12	Shapiro–Wilk test	0.3044	0.8545	Welch's T-Test	t(18.382) = -1.2638	0.2221	
	B	12	12				Two way repeated Anova	F(9,220) = 0.678 (Sessions)	0.7283	
								F(1,220) = 3.457 (Condition)	0.0643	
								F(9,220) = 0.358 (Interaction)	0.9536	
	C	12	12	Shapiro–Wilk test	0.5497	0.8043	Welch's T-Test	t(19.86) = 2.0235	0.05669	
	D	12	12				Two way repeated Anova	F(9,220) = 1.548 (Session)	0.1326	
								F(1,22) = 5.837 (Condition)	0.0165	*
								F(9,220) = 0.491 (Interaction)	0.8801	
	F	20	22	Shapiro–Wilk test	0.5897	0.3375	Welch's T-Test	t(33.224) = -2.1078	0.04267	*
	G Preference	24	24				Two way repeated Anova	F(3,184) = 86.247 (Percentage)	<0.0005	***
								F(1,184) = 3.534 (Condition)	0.0617	
								F(3,184) = 0.121 (Interaction)	0.9474	
	G Water intake	24	24				Two way repeated Anova	F(3,184) = 28.359 (Percentage)	<0.0005	***
								F(1,184) = 1.893 (Condition)	0.171	
								F(3,184) = 0.101 (Interaction)	0.959	
	G Sucrose intake	24	24				Two way repeated Anova	F(3,184) = 80.719 (Percentage)	<0.0005	***
								F(1,184) = 10.030 (Condition)	0.0018	**
								F(3,184) = 5.303 (Interaction)	0.00158	**
	H	12	12				Two way repeated Anova	F(3,88) = 131.735 (FR)	<0.0005	***
								F(1,88) = 0.205 (Condition)	0.652	
								F(3,88) = 0.323 (Interaction)	0.809	
	I	12	12				Two way repeated Anova	F(2,66) = 0.162 (PR)	0.851	
								F(1,66) = 4.583 (Condition)	0.036	*
								F(2,66) = 0.210 (Interaction)	0.811	

	Description	N		Normality test	P-Value		Statistical test	df/t/F/W/etc…	P-Value	
Fig S2		CTL	SPD		CTL	SPD				
	A	20	22	Shapiro–Wilk test	0.1496	0.2064	Welch's T-Test	t(40) = -0.034434	0.9727	
	B	14	12	Shapiro–Wilk test	0.008746	0.4817	Mann-Whitney-Wilcoxon Test	W = 73.5	0.6069	
	C	19	22				Two way repeated Anova	F(3,156) = 44.200 (Current)	<0.0005	***
								F(1,156) = 0.314 (Condition)	0.576	
								F(9,156) = 0.237 (Interaction)	0.871	

	Description	N		Normality test	P-Value		Statistical test	df/t/F/W/etc…	P-Value	
Fig S3		CTL	SPD		CTL	SPD				
	A overall	12	12	Shapiro–Wilk test	0.565	0.1079	Welch's T-Test	t(16.641) = -1.429	0.1715	
	A L1	12	12	Shapiro–Wilk test	<0.0005	<0.0005	Mann-Whitney-Wilcoxon Test	W = 78	0.5807	
	A L2	12	12	Shapiro–Wilk test	0.2388	0.1067	Welch's T-Test	t(21.997) = -2.2509	0.03472	*
	A L3	12	12	Shapiro–Wilk test	0.8051	0.02834	Mann-Whitney-Wilcoxon Test	W = 60	0.5137	
	A L5	12	12	Shapiro–Wilk test	0.3472	0.3823	Welch's T-Test	t(20.132) = -0.71114	0.4852	
	A L6	12	12	Shapiro–Wilk test	0.1141	0.07567	Welch's T-Test	t(21.983) = -0.26971	0.7899	
	B	49	34	Shapiro–Wilk test	0.9587	0.05358	Welch's T-Test	t(67.544) = -1.2866	0.2026	

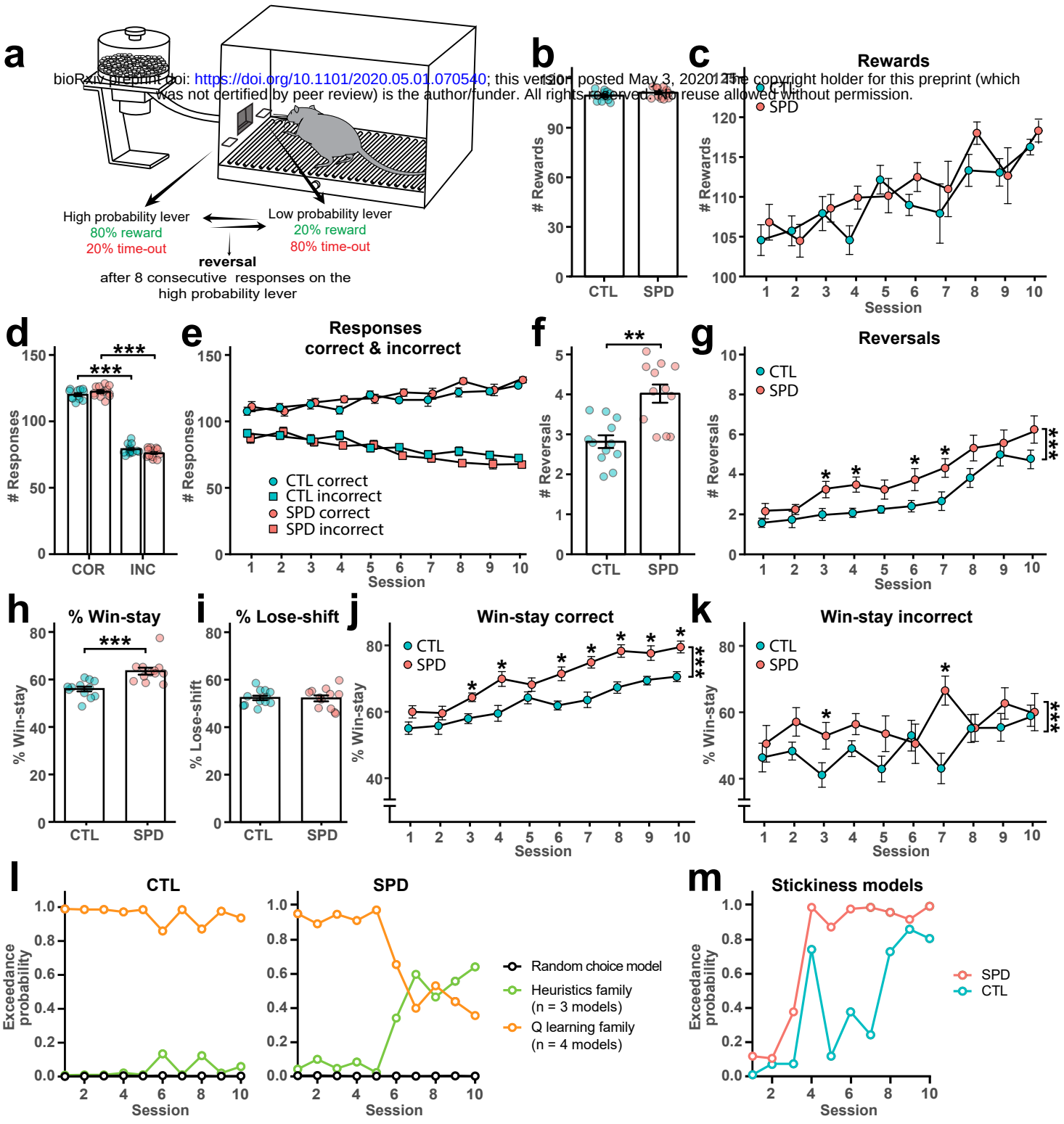


Figure 1

Omrani et al.

Figure 1. *Altered cognitive strategy after social play deprivation*

(**a**) Probabilistic reversal learning task design. (**b**) Average number of rewards for CTL and SPD rats ($p = 0.16$; T test). (c) Number of rewards for each session ($p = 0.06$; 2w ANOVA, condition). (**d**) Average number of correct and incorrect lever presses ($p < 0.0005$ corr vs incorr; T test). (**e**) Number of correct and incorrect lever presses per session ($p = 0.03$ correct; $p = 0.004$ incorrect; 2w ANOVA, condition). (**f**) Average number of reversals ($p = 0.001$; T test). (**g**) Number of reversals per session ($p < 0.0005$; 2w ANOVA, condition). (**h**) Average win-stay behaviour after reinforcement on the correct lever ($p < 0.0005$; MW test). (**i**) Average lose-stay behaviour after reinforcement on the correct lever ($p = 0.87$; T test). (**j**) Correct win-stay behaviour per session ($p < 0.0005$; 2w ANOVA, condition). (**k**) Incorrect win-stay behaviour per session ($p = 0.0003$; 2w ANOVA, condition). (l) Exceedance probability for different families of computational models based on Bayesian model selection. M, Combined exceedance probability for the two families with stickiness (see online methods, model 4 and model 7). Data from 12 CTL and 12 SPD rats. Difference between CTL and SPD: *$p < 0.05$, **$p < 0.005$, ***$p < 0.0005$

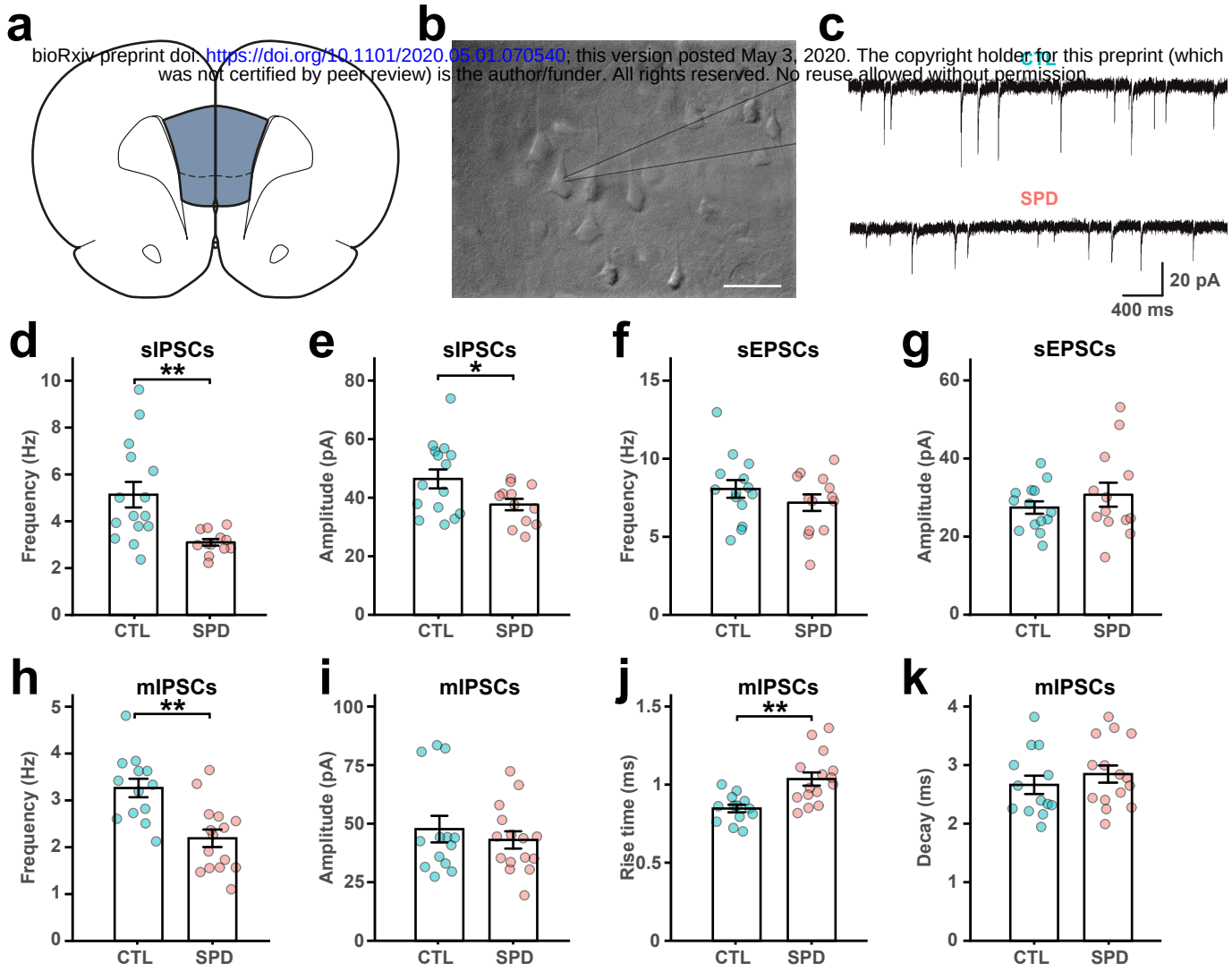


Figure 2

Omrani et al.

Figure 2. *Reduced prefrontal inhibition in L5 pyramidal cells after social play deprivation*

(**a**) Schematic diagram depicting the recording site in the mFPC. (**b**) DIC image of mPFC with the recording electrode (grey lines). Scale bar is 20 µm. (**c**) Example traces of sIPSCs in L5 pyramidal cells. (**d**, **e**) Frequency (**d**) and amplitude (**e**) of sIPSCs in CTL and SPD slices ($p = 0.002$ and $p = 0.03$; T test). (**f**, **g**) Frequency (**f**) and amplitude (**g**) of sEPSCs ($p = 0.27$ and $p = 0.35$; T test). (**h**, **i**) Frequency (**h**) and amplitude (**i**) of mIPSCs ($p = 0.0005$ and $p = 0.50$; T test). (**j**) Rise time of mIPSCs ($p = 0.0008$; T test). (**k**) Decay time of mIPSCs ($p = 0.40$; T test). Data from 15 CTL and 13 SPD brain slices (6 rats per group). Difference between CTL and SPD: $*p < 0.05$, $**p < 0.005$.

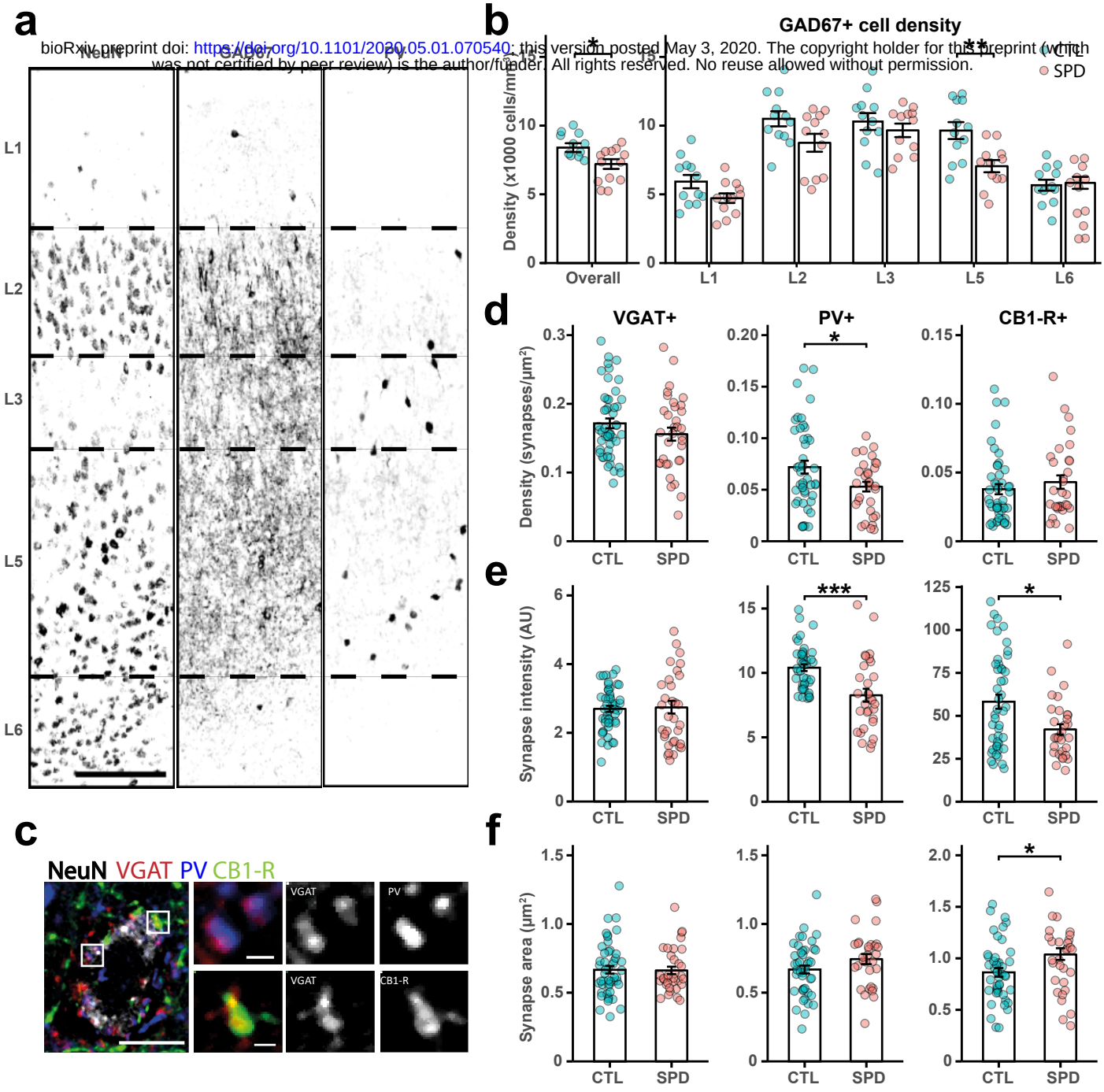
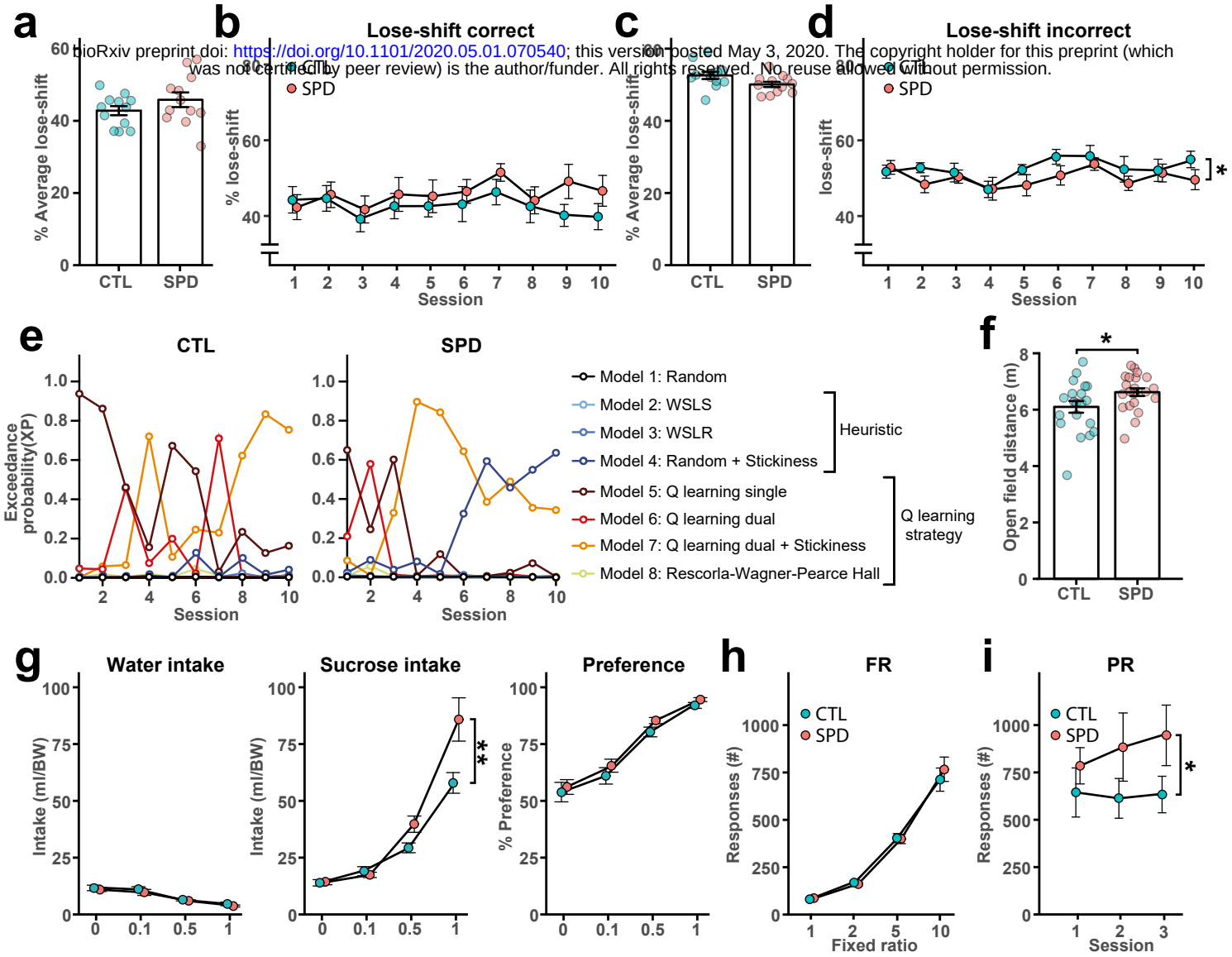


Figure 3
Omrani et al.

Figure 3. *Reduction in perisomatic inhibition after social play deprivation*

(**a**) Representative confocal image of NeuN, GAD67 and PV positive neurons in prelimbic cortex layers. Borders between layers are denoted by the dashed lines. Scale bar is 10 µm. (**b**) The density of GAD67-positive cells in the mPFC ($p = 0.02$; T test). (**c**) Confocal images of mPFC layer 5 inhibitory synapses showing NeuN, VGAT, PV and CB1-R immunostaining. Scale bar is 1 Mm. (**d**, **e**, **f**) The density (**d**), intensity (**e**) and area (**f**) of VGAT, PV and CB1-R positive synapses (co-localizations with VGAT) (density: VGAT $p = 0.19$; PV $p = 0.02$; CB1-R $p = 0.36$; intensity: VGAT $p = 0.51$; PV $p < 0.0005$; CB1-R $p = 0.02$; area: VGAT $p = 0.93$; PV $p = 0.11$; CB1-R $p = 0.05$; MW test). Data in **b** from both hemispheres (12 rats per group); in **d-f**: data from 48 CTL cells and 34 SPD cells (6 rats per group). Difference between CTL and SPD: $*p < 0.05$, $**p < 0.005$, $***p < 0.0005$.

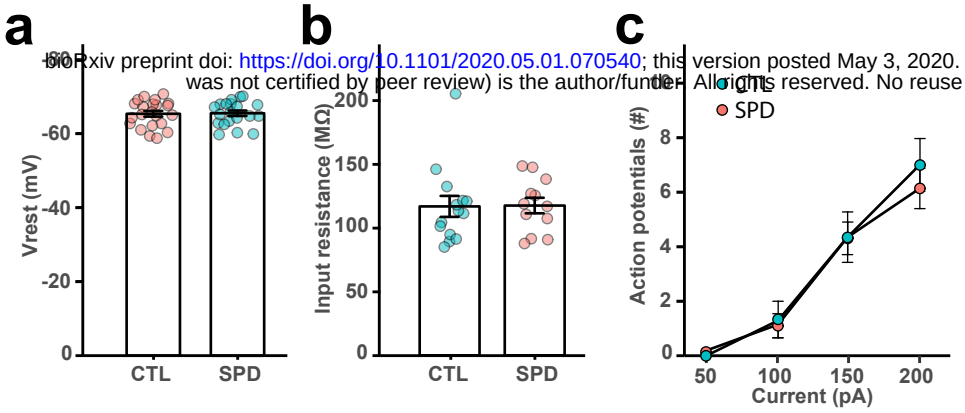


Suppl Figure 1

Omrani et al.

Supplementary Figure 1. *Altered cognitive strategy after social play deprivation*

(**a**) Average lose-shift behaviour after a loss on the correct lever (p = 0.22; T test). (**b**) Lose-shift behaviour on the correct lever per session (p = 0.06; 2w ANOVA, condition). (**c**) Lose-shift behaviour on the incorrect lever after a loss on the incorrect lever (p = 0.06; T test). (**d**) Lose-shift behaviour per session (p = 0.02; 2w ANOVA, condition). (**e**) Exceedance probability for random and specific Q-learning and heuristic family models. (**f**) Distance travelled in open field for CTL and SPD rats (p= 0.04; T test). (**g**) Water and sucrose intake and % sucrose preference (p = 0.17; p = 0.002; p = 0.06; 2w ANOVA, condition). (**h**) Number of rewards during fixed ratio reinforcement schedule (p = 0.65; 2w ANOVA, condition). (**i**) Number of rewards during progressive ratio reinforcement schedule (p = 0.036; 2w ANOVA, condition). Data from 12 CTL and 12 SPD rats. Difference between CTL and SPD: *p < 0.05, **p < 0.005.

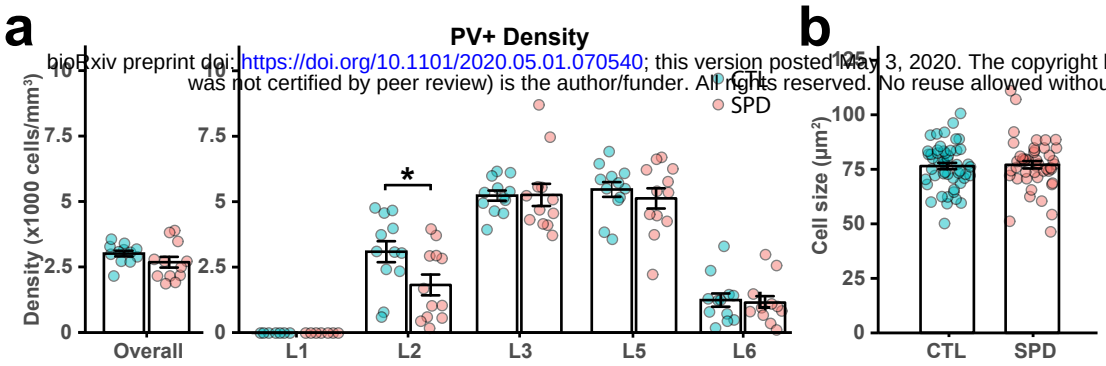


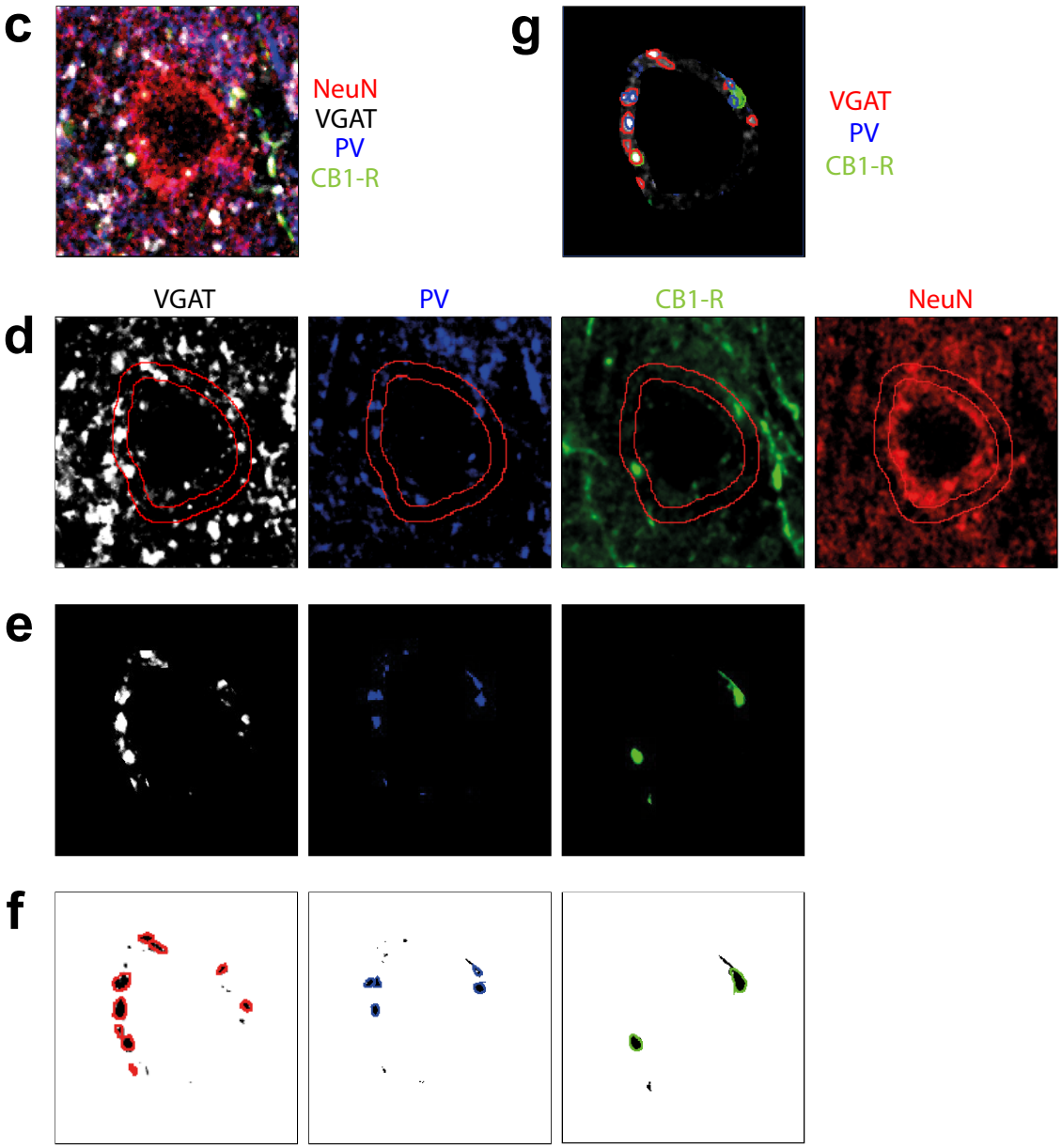
Suppl Figure 2

Omrani et al.

Supplementary Figure 2. *Passive membrane properties of L5 cells are not different*

(**a**) Resting potential of L5 pyramidal neurons in CTL and SPD slices (p = 0.97; T test). (**b**) Input resistance (p = 0.61; MW test). (**c**) Number of action potentials after current injection in CTL and SPD neurons (p = 0.58; 2w ANOVA, condition). Data in **a** from 20 CTL and 22 SPD cells; in **b** from 14 CTL and 12 SPD cells; in C from 19 CTL and 22 SPD cells.





Suppl Figure 3
Omrani et al.

Supplementary Figure 3. *Reduction in perisomatic inhibition after social play deprivation*

(**a**) The overall density of PV-positive cells in the mPFC in CTL and SPD rats ($p = 0.17$; T test). (**b**) Soma size of L5 pyramidal cells ($p = 0.20$; T test). (**c-g**) Workflow of the synaptic analysis. (**c**, **d**) Representative confocal images for VGAT, PV, CB1-R and NeuN immunostaining. (**e**) A 1.5 µm band around the soma was drawn based on the NeuN staining. (**f**) Individual puncta were selected after thresholding based on their size, circularity and co-localization with VGAT. Selected puncta are indicated with a coloured outline. G, Summary of the selected VGAT, PV and CB1-R puncta. Data in **a** from 12 CTL and 12 SPD rats; in **b-c** from 51 CTL and 47 SPD cells. Difference between CTL and SPD: $^*p < 0.05$.