UNIVERSITÀ DEGLI STUDI
DI GENOVA

ISTITUTO ITALIANO
DI TECNOLOGIA

Università degli Studi di Genova

Fondazione Istituto Italiano di Tecnologia

Ph.D course in "Neuroscience", XXXIII Cycle

Curriculum "Neuroscience and Neurotechnology"

Coordinator: Prof. Fabio Benfenati

Brain Circuits of Emotion Discrimination

Supervisor:
Dr. Francesco Papaleo

PhD Candidate:
Federica Maltese

Table of Contents

CHAPTER 1

General Introduction and Outlines of the Thesis

General Introduction

1. Emotion recognition in humans and other species

Emotion recognition refers to the cognitive ability to perceive, process and interpret the emotional value of a social stimulus [1]. This ability is crucial to understand others' emotional state and intentions, and to engage the most appropriate behavioral responses [2, 3]. Indeed, these social processes guide both automatic and volitional behavioral reactions by playing a role in processes such as memory, decision-making, attention and motivation [1].

In humans, this ability is experimentally assessed by the so-called Emotion Recognition Tasks (ERTs). One of the most validated is the 'Ekman's Faces Test', which depicts faces displaying the six basic emotions (disgust, anger, fear, surprise, sadness and happiness) that need to be recognized by the testing subject [4]. In particular, emotions can be presented in graded intensities in order to address diverse severity levels of potential deficits, typical of many psychiatric disorders [5-10].

Emotion recognition is not uniquely present in humans [11]. For example, emotion recognition abilities can be assessed in non-human primates by exposing subjects to drawings [12, 13], static images, movies, or real-life conspecifics [14, 15]. Indeed, non-human primates are able to communicate emotions through facial expressions, with homologies to human expressions [13, 15-23], and they are able to recognize and categorize conspecific's affective states [13]. Moreover, similarly to humans, non-human primates might use the information that derives from processing others' emotions to predict future actions or scenarios and to guide behavioral choices [24-26].

The intrinsic complexity of emotion recognition led for a long time to attribute this process to the most evolved species, such as human and non-human primates [16-18, 27]. However, this is a fundamental ability, which by conveying information about threats or dangers and by mediating affiliative behaviors, is important for survival. Thus, not surprisingly, these processes can be revealed also in other social species including horses, dogs and sheep [28-30]. Communication of emotions strongly depends on the species [31]. Accordingly, the ability to perceive emotions has been investigated with species-specific behavioral settings. In general, in species strongly relying on visual cues, such as primates, horses, sheep and dogs, emotion recognition ability involves the integration of multimodal signals [21, 30, 32]. In particular, by using an adapted version of the human ERTs [28, 33, 34], it has been found that these animals are

able to perceive social signals to infer emotional states [28, 34, 35]. Moreover, they show as well adaptive behavioral responses, such as avoidance for agonistic or stressed expressions compared to neutral ones [28, 29]. Interestingly, horses, dogs and sheep, as well as non-human primates, can discriminate emotions in subjects of other species. Thus, emotion recognition ability is important not only for survival within intragroup relationships, but also to sustain proper evolutionary interspecies interactions [11].

As other mammals, rodents show complex social behaviors as well as a high level of reciprocal social interactions. Moreover, rodents show distinct facial expressions, which have been demonstrated to reliably indicate emotional states [36-38]. In agreement, most recently, it has been shown that rodents are able to perceive and react to the emotional state of conspecifics [39-46], but also to produce complex pro-social behaviors such as helping and consolation [47-49]. In particular, it has been shown that witnessing a conspecific being shocked leads to fear responses in the observer. This process is called 'vicarious freezing', and it is part of the so-called emotional contagion processes that refer to the ability of a subject to match its emotional state to the one of a conspecific in pain or distress [11, 42]. Moreover, it has been shown that rodents can perceive and react to changes in the emotional state of conspecifics even when there is no direct witnessing of the distress [44, 50, 51]. In particular, the exposure to a conspecific that experienced a fear conditioning immediately before, induced affective response such as increased social approach and allogroming [44, 49]. Thus, rodents are sensible to altered affective states in others, reacting to them in different ways. Despite this, none of the currently available paradigms previously addressed what is commonly measured by the emotion recognition tasks in humans, which is namely the basic ability to discriminate emotional states in others.

Rodents are the most widely used laboratory animals that give the possibility to apply innovative technologies to directly investigate the biological substrates of specific behaviors. Thus, developing paradigms able to detect and quantify emotion recognition abilities and approximating different features of equivalent human tasks could provide a powerful tool to investigate the neurobiology of these fundamental social processes.

2. Emotion recognition in psychiatric disorders

Impaired processing of emotion recognition has been reported in a range of disorders, including focal brain lesions [52], Huntington's Disease [53], Parkinson's Disease [54], and neuropsychiatric disorders such as Schizophrenia and Autism Spectrum Disorder (ASD) [55, 56].

Social cognitive impairments such as social perception, emotion recognition and theory of mind are key features in schizophrenia [57], in which these symptoms frequently precedes the onset of psychosis [58]. In particular, patients with schizophrenia show deficits in both emotional and cognitive facets of social processes, with marked impairments in processing non-verbal social affective information while showing normal affect sharing and emotion experiences [59]. It has been reported, especially in patients with marked negative symptoms, impairments in recognizing emotions from facial expressions and in appreciating other people's mental states [58, 60]. Moreover, impaired social cognition in these patients might relate to social anxiety, withdrawal and reduced motivation, leading to poor everyday functioning [57, 61].

Autism spectrum disorders (ASD) have been described as neurodevelopmental disorders with marked impairments in social communication and social interaction, including multiple nonverbal behaviors such as facial expressions and lack of social or emotional sharing [62]. In particular, problems related to emotion processing have been seen as hallmark symptoms of these pathologies. However, findings on emotion recognition deficits in ASD are inconsistent: in most cases, autistic patients have difficulty labeling, or most often, matching emotions, but in some cases, they perform as well as controls [56, 63, 64]. These disparate findings have been attributed to demographic characteristics of participants, task demands and different variables measured [56, 65].

Emotion recognition impairments have a deleterious impact on patients' daily life, more than non-social cognitive deficits [66]. In general, social cognitive abilities enable subjects to interact effectively with their social environment, thus deficits on this domain could lead to social misperceptions, inappropriate interpersonal reactions or social withdrawal [57]. Unfortunately, there are still no effective therapeutic strategies for these social behavior problems, mostly because of the lack of knowledge on the neuronal substrates underlying them. In order to achieve this, it is important to understand with a cell-, circuit- and system-specific way the biology of emotion recognition.

3. Neuronal substrates of emotion recognition

Our current knowledge on neuronal substrates of emotion recognition derives mostly from brain lesions and functional neuroimaging studies in healthy subjects and patients [67-69]. Recognition of socially relevant information requires a network of neural assembles, which process the perception of social signals and bring this information to brain areas involved in motivation, emotion and adaptive behaviors [70-73]. This 'social brain' network encompasses: the sensory and association neocortices for social perceptual processing; a network involving amygdala, prefrontal cortex, cingulate cortex and somatosensory-related cortices for the mediation between perception and cognitive processing; hypothalamus, brainstem nuclei, basal ganglia, and motor cortices for carrying out the actions involved in the social behavior [70]. In particular, it has emerged that interaction between limbic structures and frontal cortical areas is implicated in the social cognitive processes of emotion recognition [74], with the amygdala and the prefrontal cortex (PFC) playing a prominent role [60].

The amygdala is strictly implicated in linking perceptual representations to cognition and behavior, according to the emotional or social value of the stimuli [1]. Indeed, lesion studies in humans have demonstrated that the amygdala is involved in the recognition of others' emotions through their facial expressions [52]. Moreover, the amygdala is crucial for rapid processing of ambiguous, potentially threatening or dangerous stimuli [75]. Nevertheless, recent findings also implicate the amygdala in the evaluation of positively-valenced emotional stimuli such as happy or surprised expressions [76-79].

The PFC, instead, is mainly involved in a general "top-down" control of cognitive processes [61]. Indeed, the PFC has been found to act at higher cognitive levels such as processing representations of emotions and their value [80-82], and understanding others' thoughts and intentions ('theory of mind') [83, 84]. However, it has been shown that the PFC is also engaged already during rapid processing of socially or emotionally salient features shown in facial expressions [85]. In general, PFC lesion studies in humans have shown impairments in recognition of facial and vocal emotional expressions [86], and other non verbal social and emotional cues such as body language [87].

It has been proposed that the amygdala and the PFC might act in synergy, with the PFC tempering the amygdala, in order to exert control over impulsive and aggressive social behaviors,

and the amygdala modulating the PFC responses through inputs regarding vigilance, threat and ambiguity of the stimuli [88].

Altered activation of amygdala and PFC regions during emotion recognition have emerged from neuroimaging studies on patients with ASD and schizophrenia. For example, in ASD it has been reported an altered functional connectivity between the amygdala and other regions when viewing emotional faces: increased connectivity with ventro-medial pre-frontal cortex when viewing happy faces, and decreased connectivity with the inferior frontal gyrus when viewing sad faces [56]. Similarly, in patients with schizophrenia, an altered activation of both amygdala and PFC during the presentation of emotional stimuli has been reported, with the amygdala being impaired during processing of facial expression, while the PFC during theory of mind [89-91].

Despite this knowledge, the specific mechanisms at the circuit- and cell-specific levels that guide the activity of the above mentioned brain regions during emotion recognition is something still unreachable in humans. In this context, rodents' studies could be useful to identify fundamental biological principles of social cognition, which could then be extended to other species. In the last years, emotion based behavioral rodents' studies have confirmed some overlapping substrate with other species, particularly the amygdala, anterior cingulate and insula cortex [43, 51, 92-94]. In particular, in a rat model of vicarious freezing it has been shown that a specific amygdala circuit, from the lateral nucleus to the medial nucleus of the amygdala (LA–MeA), is required for using social cues to learn about environmental cues that signal imminent threats, subsequently guiding appropriate social behaviors [93]. In another study, it has been demonstrated in rats that the anterior cingulate cortex (ACC) is uniquely activated during acquisition of vicarious fear conditioning, and its activation is necessary to express a freezing response in the following days [92]. As mentioned above, cortico-limbic communication is highly involved in social cognitive processes. In a mouse model of observational fear learning, it has been shown that cortical input (ACC) transmits information derived from observation to the amygdala (basolateral amygdala), in order to instruct encoding the aversive value of learned cues [94]. Finally, it has been shown that the insular cortex activity and its modulation by oxytocin (OXT) are necessary for social affective behaviors in rats, modulating the social approach towards a conspecific in distress [51].

Nevertheless, the specific neuronal substrates underlying emotion recognition processes are still unknown. Thus, appropriate behavioral paradigms able to approximate human emotion

recognition tasks are needed in order to be able to achieve mechanistic insights on emotion recognition.

4. Neuropeptides as modulators of emotion recognition

The neuropeptides OXT and arginine vasopressin (AVP) have been described to have a key role in the modulation of social behaviors in both rodents [95, 96], and human patients [97, 98].

OXT and AVP are two closely related neuropeptide sharing a similar structure. OXT and AVP are synthesized primarily in separate neuronal populations in the hypothalamic paraventricular (PVN) and supraoptic (SON) nuclei, and can be released centrally by direct projections into other brain areas where they act as neuromodulators, or peripherally through the posterior pituitary [99]. OXT and AVP exert their action binding a subgroup of receptors: the OXT receptor (OXTR) and the AVP receptor subtypes (V1a, V1b, and V2). Given their similar structure, OXT and AVP can bind, even with a lower efficacy, each other receptors [100].

Particular attention has been directed onto OXT system, which has been reported to play a crucial role in regulation of complex social behaviors, such as attachment, social exploration, recognition and aggression [96, 101-104]. Since neuropeptides cross the blood-brain barrier after intranasal administration [105], a number of studies has been conducted in humans demonstrating that intranasal OXT strengthens social-cognitive processing, by increasing the social salience of stimuli [106-109]. In particular, a popular current hypothesis states that exogenous OXT might strengthen the positive valence of a positive social cue, while what is commonly negative and aversive might be perceived as more negative and aversive [110]. Thus, it has been proposed that during emotion recognition OXT is recruited in decoding the social salience of facial expressions [111].

Several studies reported impaired function of the OXT system in neuropsychiatric disorders associated with social deficits such as autism, schizophrenia and depression [112]. Thus, a number of clinical studies investigated whether intranasal administration of OXT could improve social deficits, particularly emotion recognition, in these disorders [98, 113-116]. However, the results reported so far are mixed and elusive [117]. These controversial effects can be explained by the salience hypothesis of OXT, according to which its action might differ depending on the social context in which it is administered [108]. Alternatively, this could be linked to inter-

individual differences of patients, pointing out the importance of developing personalized therapeutic strategies by considering the specific genetic background and the nature of the social impairments.

Another neuropeptide involved in social cognitive processes is the corticotropin-releasing factor (CRF). The CRF mediates the neuroendocrine and behavioral responses to stressful situations through the regulation of the hypothalamic-pituitary-adrenal axis . In addition to the CRF, the mammalian CRF peptide family contains also the urocortins 1, 2 and 3 (Ucn1, Ucn2 and Ucn3), which were shown to elicit similar biological effects as CRF [118]. The biological activity of CRF is mediated through two receptors, the CRFR type 1 (CRFR1), the most represented in the central nervous system, and type 2 (CRFR2). Initially, the two receptors seemed to have possibly different biological roles, with the CRFR1 showing an anxiogenic effect and the CRFR2 an anxiolytic one. However, recent works have been revealing the complex functional role of both these receptors in behavioral processes [118].

Besides its role as hypothalamic neurohormone, the CRF has also an extensive extrahypothalamic action across the corticostriatal-limbic regions, playing a critical role in modulating behavioral responses to stress, including regulation of emotional and cognitive components of stress responses [119-121]. Moreover, human studies have revealed that the CRF system modulates the effect of stress on empathy. In particular, it was found that specific polymorphisms on CRFR1 result in impaired emotion recognition and empathy [122] and influence brain responses to negative emotional stimuli [123]. More generally, there is a large body of literature indicating an alteration of the CRF system in many psychiatric disorders characterized by social and emotional deficits, such as mood, trauma- and stress-related disorders [124]. Indeed, for adaptive social interactions the ability to cope with the stress that accompanies the social engagement is crucial [125]. In agreement, recent studies in rodents revealed a direct implication of the CRF system in modulating different aspects of social behavior such as general social interaction, social memory, social defeat, sexual behavior, pair bonding, and defensive/aggressive behaviors [126-131]. In particular, the CRF modulation of social behavior is environmentally and developmentally sensitive, species-specific, and often sex-specific [132, 133], and seems to be mediated more by the CRFR2 than CRFR1.

In conclusion, there is a variety of studies, both in humans and rodents, which brings the attention to the essential role of neuropeptides in modulating and fine-tuning social cognitive behaviors, making them potential targets for therapeutic intervention.

References

1. Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. Behav Cogn Neurosci Rev 1, 21-62.

2. Millan, M.J., Agid, Y., Brune, M., Bullmore, E.T., Carter, C.S., Clayton, N.S., Connor, R., Davis, S., Deakin, B., DeRubeis, R.J., et al. (2012). Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat Rev Drug Discov 11, 141-168.

3. Brothers, L., Ring, B., and Kling, A. (1990). Response of neurons in the macaque amygdala to complex social stimuli. Behav Brain Res 41, 199-213.

4. Ekman, P. (1992). Are there basic emotions? Psychol Rev 99, 550-553.

5. Kohler, C.G., Turner, T.H., Bilker, W.B., Brensinger, C.M., Siegel, S.J., Kanes, S.J., Gur, R.E., and Gur, R.C. (2003). Facial emotion recognition in schizophrenia: intensity effects and error pattern. Am J Psychiatry 160, 1768-1774.

6. Pinkham, A.E., Penn, D.L., Green, M.F., and Harvey, P.D. (2016). Social Cognition Psychometric Evaluation: Results of the Initial Psychometric Study. Schizophr Bull 42, 494-504.
7. Blair, R.J., Colledge, E., Murray, L., and Mitchell, D.G. (2001). A selective impairment in the processing of sad and fearful expressions in children with psychopathic tendencies. J Abnorm Child Psychol 29, 491-498.
8. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J Child Psychol Psychiatry 42, 241-251.
9. Cecilione, J.L., Rappaport, L.M., Verhulst, B., Carney, D.M., Blair, R.J.R., Brotman, M.A., Leibenluft, E., Pine, D.S., Roberson-Nay, R., and Hettema, J.M. (2017). Test-retest reliability of the facial expression labeling task. Psychol Assess 29, 1537-1542.
10. Henry, J.D., von Hippel, W., Molenberghs, P., Lee, T., and Sachdev, P.S. (2016). Clinical assessment of social cognitive function in neurological disorders. Nat Rev Neurol 12, 28-39.
11. Ferretti, V., and Papaleo, F. (2019). Understanding others: Emotion recognition in humans and other animals. Genes Brain Behav 18, e12544.
12. Dittrich, W. (1990). Representation of Faces in Longtailed Macaques (Macaca fascicularis). Ethology 85, 265.
13. Parr, L.A. (2003). The discrimination of faces and their emotional content by chimpanzees (Pan troglodytes). Ann N Y Acad Sci 1000, 56-78.
14. Nahm, F.K., Perret, A., Amaral, D.G., and Albright, T.D. (1997). How do monkeys look at faces? J Cogn Neurosci 9, 611-623.
15. Parr, L.A., Dove, T., and Hopkins, W.D. (1998). Why faces may be special: evidence of the inversion effect in chimpanzees. J Cogn Neurosci 10, 615-622.
16. Chevalier-Skolnikoff, S. (1973). Visual and tactile communication in Macaca arctoides and its ontogenetic development. Am J Phys Anthropol 38, 515-518.
17. Sherwood, C.C., Holloway, R.L., Gannon, P.J., Semendeferi, K., Erwin, J.M., Zilles, K., and Hof, P.R. (2003). Neuroanatomical basis of facial expression in monkeys, apes, and humans. Ann N Y Acad Sci 1000, 99-103.
18. van Hooff, J.A.R.A.M. (1962). Facial expressions in higher primates. Symposia of the Zoological Society of London 8.
19. Leopold, D.A., and Rhodes, G. (2010). A comparative view of face perception. J Comp Psychol 124, 233-251.

20. Parr, L.A., Waller, B.M., and Vick, S.J. (2007). New Developments in Understanding Emotional Facial Signals in Chimpanzees. Curr Dir Psychol Sci 16, 117-122.
21. Ghazanfar, A.A., and Logothetis, N.K. (2003). Neuroperception: facial expressions linked to monkey calls. Nature 423, 937-938.
22. Partan, S.R. (2002). Single and Multichannel Signal Composition: Facial Expressions and Vocalizations of Rhesus Macaques (Macaca mulatta). Behaviour 139, 993-1027.
23. Waller, B.M., Whitehouse, J., and Micheletta, J. (2017). Rethinking primate facial expression: A predictive framework. Neurosci Biobehav Rev 82, 13-21.
24. Morimoto, Y., and Fujita, K. (2012). Capuchin monkeys (Cebus apella) use conspecifics' emotional expressions to evaluate emotional valence of objects. Anim Cogn 15, 341-347.
25. Buttelmann, D., Call, J., and Tomasello, M. (2009). Do great apes use emotional expressions to infer desires? Dev Sci 12, 688-698.
26. Waller, B.M., Whitehouse, J., and Micheletta, J. (2016). Macaques can predict social outcomes from facial expressions. Anim Cogn 19, 1031-1036.
27. Andrew, R.J. (1963). Evolution of Facial Expression. Science 142, 1034-1041.
28. Tate, A.J., Fischer, H., Leigh, A.E., and Kendrick, K.M. (2006). Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. Philos Trans R Soc Lond B Biol Sci 361, 2155-2172.
29. Wathan, J., Proops, L., Grounds, K., and McComb, K. (2016). Horses discriminate between facial expressions of conspecifics. Sci Rep 6, 38322.
30. Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., and Mills, D. (2016). Dogs recognize dog and human emotions. Biol Lett 12, 20150883.
31. Hauser, M.D. (1996). The evolution of communication. Cambridge, MA, US: The MIT Press.
32. Barraclough, N.E., Xiao, D., Baker, C.I., Oram, M.W., and Perrett, D.I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci 17, 377-391.
33. Ekman, P.F., W.V. (1978). The Facial Action Coding System: A Tecnique for Measurement of Facial Movement. San Francisco: Consulting Psychologists Press.
34. Wathan, J., Burrows, A.M., Waller, B.M., and McComb, K. (2015). EquiFACS: The Equine Facial Action Coding System. PLoS One 10, e0131738.
35. Wathan, J., and McComb, K. (2014). The eyes and ears are visual indicators of attention in domestic horses. Curr Biol 24, R677-679.
36. Pecina, S., Cagniard, B., Berridge, K.C., Aldridge, J.W., and Zhuang, X. (2003). Hyperdopaminergic mutant mice have higher "wanting" but not "liking" for sweet rewards. J Neurosci 23, 9395-9402.
37. Berridge, K., and Winkielman, P. (2003). What is an unconscious emotion?(The case for unconscious "liking"). Cogn Emot 17, 181-211.
38. Dolensek, N., Gehrlach, D.A., Klein, A.S., and Gogolla, N. (2020). Facial expressions of emotion states and their neuronal correlates in mice. Science 368, 89-94.
39. Panksepp, J.B., and Lahvis, G.P. (2011). Rodent empathy and affective neuroscience. Neurosci Biobehav Rev 35, 1864-1875.
40. Atsak, P., Orre, M., Bakker, P., Cerliani, L., Roozendaal, B., Gazzola, V., Moita, M., and Keysers, C. (2011). Experience modulates vicarious freezing in rats: a model for empathy. PLoS One 6, e21855.
41. Panksepp, J., and Panksepp, J.B. (2013). Toward a cross-species understanding of empathy. Trends Neurosci 36, 489-496.
42. Meyza, K.Z., Bartal, I.B., Monfils, M.H., Panksepp, J.B., and Knapska, E. (2017). The roots of empathy: Through the lens of rodent models. Neurosci Biobehav Rev 76, 216-234.
43. Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H.E., Lin, S.Y., Rabah, D., Kinet, J.P., and Shin, H.S. (2010). Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. Nat Neurosci 13, 482-488.

44. Knapska, E., Mikosz, M., Werka, T., and Maren, S. (2010). Social modulation of learning in rats. Learn Mem 17, 35-42.
45. Chen, Q., Panksepp, J.B., and Lahvis, G.P. (2009). Empathy is moderated by genetic background in mice. PLoS One 4, e4387.
46. Carrillo, M., Migliorati, F., Bruls, R., Han, Y., Heinemans, M., Pruis, I., Gazzola, V., and Keysers, C. (2015). Repeated Witnessing of Conspecifics in Pain: Effects on Emotional Contagion. PLoS One 10, e0136979.
47. Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and pro-social behavior in rats. Science 334, 1427-1430.
48. Ben-Ami Bartal, I., Rodgers, D.A., Bernardez Sarria, M.S., Decety, J., and Mason, P. (2014). Pro-social behavior in rats is modulated by social experience. Elife 3, e01385.
49. Burkett, J.P., Andari, E., Johnson, Z.V., Curry, D.C., de Waal, F.B., and Young, L.J. (2016). Oxytocin-dependent consolation behavior in rodents. Science 351, 375-378.
50. Knapska, E., Nikolaev, E., Boguszewski, P., Walasek, G., Blaszczyk, J., Kaczmarek, L., and Werka, T. (2006). Between-subject transfer of emotional information evokes specific pattern of amygdala activation. Proc Natl Acad Sci U S A 103, 3858-3862.
51. Rogers-Carter, M.M., Varela, J.A., Gribbons, K.B., Pierce, A.F., McGoey, M.T., Ritchey, M., and Christianson, J.P. (2018). Insular cortex mediates approach and avoidance responses to social affective stimuli. Nat Neurosci 21, 404-414.
52. Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. Nature 372, 669-672.
53. Jacobs, D.H., Shuren, J., Bowers, D., and Heilman, K.M. (1995). Emotional facial imagery, perception, and expression in Parkinson's disease. Neurology 45, 1696-1702.
54. Adolphs, R., Schul, R., and Tranel, D. (1998). Intact recognition of facial emotion in Parkinson's disease. Neuropsychology 12, 253-258.
55. Gur, R.E., McGrath, C., Chan, R.M., Schroeder, L., Turner, T., Turetsky, B.I., Kohler, C., Alsop, D., Maldjian, J., Ragland, J.D., et al. (2002). An fMRI study of facial emotion processing in patients with schizophrenia. Am J Psychiatry 159, 1992-1999.
56. Harms, M.B., Martin, A., and Wallace, G.L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol Rev 20, 290-322.
57. Green, M.F., Penn, D.L., Bentall, R., Carpenter, W.T., Gaebel, W., Gur, R.C., Kring, A.M., Park, S., Silverstein, S.M., and Heinssen, R. (2008). Social cognition in schizophrenia: an NIMH workshop on definitions, assessment, and research opportunities. Schizophr Bull 34, 1211-1220.
58. Brune, M. (2005). Emotion recognition, 'theory of mind,' and social behavior in schizophrenia. Psychiatry Res 133, 135-147.
59. Green, M.F., Horan, W.P., and Lee, J. (2015). Social cognition in schizophrenia. Nat Rev Neurosci 16, 620-631.
60. Maat, A., van Montfort, S.J., de Nijs, J., Derks, E.M., Kahn, R.S., Linszen, D.H., van Os, J., Wiersma, D., Bruggeman, R., Cahn, W., et al. (2015). Emotion processing in schizophrenia is state and trait dependent. Schizophr Res 161, 392-398.
61. Millan, M.J., and Bales, K.L. (2013). Towards improved animal models for evaluating social cognition and its disruption in schizophrenia: the CNTRICS initiative. Neurosci Biobehav Rev 37, 2166-2180.
62. Kanner, L. (1965). Infantile autism and the schizophrenias. Behav Sci 10, 412-420.
63. Black, M.H., Chen, N.T.M., Iyer, K.K., Lipp, O.V., Bolte, S., Falkmer, M., Tan, T., and Girdler, S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. Neurosci Biobehav Rev 80, 488-515.
64. Xavier, J., Vignaud, V., Ruggiero, R., Bodeau, N., Cohen, D., and Chaby, L. (2015). A Multidimensional Approach to the Study of Emotion Recognition in Autism Spectrum Disorders. Front Psychol 6, 1954.

65. Nuske, H.J., Vivanti, G., and Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. Cogn Emot 27, 1042-1061.
66. Fett, A.K., Viechtbauer, W., Dominguez, M.D., Penn, D.L., van Os, J., and Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 35, 573-588.
67. Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., et al. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. J Psychiatry Neurosci 34, 418-432.
68. Kret, M.E., and Ploeger, A. (2015). Emotion processing deficits: a liability spectrum providing insight into comorbidity of mental disorders. Neurosci Biobehav Rev 52, 153-171.
69. Adolphs, R., Baron-Cohen, S., and Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. J Cogn Neurosci 14, 1264-1274.
70. Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annu Rev Psychol 60, 693-716.
71. de Waal, F.B.M., and Preston, S.D. (2017). Mammalian empathy: behavioural manifestations and neural basis. Nat Rev Neurosci 18, 498-509.
72. Frith, C.D., and Frith, U. (2012). Mechanisms of social cognition. Annu Rev Psychol 63, 287-313.
73. Vuilleumier, P., and Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia 45, 174-194.
74. Gur, R.C., and Gur, R.E. (2016). Social cognition as an RDoC domain. Am J Med Genet B Neuropsychiatr Genet 171B, 132-141.
75. Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. Nat Rev Neurosci 11, 773-783.
76. Killgore, W.D., and Yurgelun-Todd, D.A. (2004). Activation of the amygdala and anterior cingulate during nonconscious processing of sad versus happy faces. Neuroimage 21, 1215-1223.
77. Sergerie, K., Chochol, C., and Armony, J.L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. Neurosci Biobehav Rev 32, 811-830.
78. Somerville, L.H., Kim, H., Johnstone, T., Alexander, A.L., and Whalen, P.J. (2004). Human amygdala responses during presentation of happy and neutral faces: correlations with state anxiety. Biol Psychiatry 55, 897-903.
79. Fitzgerald, D.A., Angstadt, M., Jelsone, L.M., Nathan, P.J., and Phan, K.L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. Neuroimage 30, 1441-1448.
80. Fellows, L.K. (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. Ann N Y Acad Sci 1239, 51-58.
81. Myers-Schulz, B., and Koenigs, M. (2012). Functional anatomy of ventromedial prefrontal cortex: implications for mood and anxiety disorders. Mol Psychiatry 17, 132-141.
82. Roy, M., Shohamy, D., and Wager, T.D. (2012). Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn Sci 16, 147-156.
83. Gallagher, H.L., and Frith, C.D. (2003). Functional imaging of 'theory of mind'. Trends Cogn Sci 7, 77-83.
84. Lieberman, M.D. (2005). Principles, processes, and puzzles of social cognition: an introduction for the special issue on social cognitive neuroscience. Neuroimage 28, 745-756.
85. Wolf, R.C., Philippi, C.L., Motzkin, J.C., Baskaya, M.K., and Koenigs, M. (2014). Ventromedial prefrontal cortex mediates visual attention during facial emotion recognition. Brain 137, 1772-1780.
86. Hornak, J., Bramham, J., Rolls, E.T., Morris, R.G., O'Doherty, J., Bullock, P.R., and Polkey, C.E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. Brain 126, 1691-1712.

87. Mah, L., Arnold, M.C., and Grafman, J. (2004). Impairment of social perception associated with lesions of the prefrontal cortex. Am J Psychiatry 161, 1247-1255.
88. Adolphs, R. (2002). Neural systems for recognizing emotion. Curr Opin Neurobiol 12, 169-177.
89. Maat, A., van Haren, N.E.M., Bartholomeusz, C.F., Kahn, R.S., and Cahn, W. (2016). Emotion recognition and theory of mind are related to gray matter volume of the prefrontal cortex in schizophrenia. Eur Neuropsychopharmacol 26, 255-264.
90. Pinkham, A.E., Penn, D.L., Perkins, D.O., and Lieberman, J. (2003). Implications for the neural basis of social cognition for the study of schizophrenia. Am J Psychiatry 160, 815-824.
91. Habel, U., Chechko, N., Pauly, K., Koch, K., Backes, V., Seiferth, N., Shah, N.J., Stocker, T., Schneider, F., and Kellermann, T. (2010). Neural correlates of emotion recognition in schizophrenia. Schizophr Res 122, 113-123.
92. Jones, C.E., Agee, L., and Monfils, M.H. (2018). Fear Conditioning by Proxy: Social Transmission of Fear Between Interacting Conspecifics. Curr Protoc Neurosci 83, e43.
93. Twining, R.C., Vantrease, J.E., Love, S., Padival, M., and Rosenkranz, J.A. (2017). An intra-amygdala circuit specifically regulates social fear learning. Nat Neurosci 20, 459-469.
94. Allsop, S.A., Wichmann, R., Mills, F., Burgos-Robles, A., Chang, C.J., Felix-Ortiz, A.C., Vienne, A., Beyeler, A., Izadmehr, E.M., Glober, G., et al. (2018). Corticoamygdala Transfer of Socially Derived Information Gates Observational Learning. Cell 173, 1329-1342 e1318.
95. Bielsky, I.F., and Young, L.J. (2004). Oxytocin, vasopressin, and social recognition in mammals. Peptides 25, 1565-1574.
96. Veenema, A.H., and Neumann, I.D. (2008). Central vasopressin and oxytocin release: regulation of complex social behaviours. Prog Brain Res 170, 261-276.
97. Pedersen, C.A., Gibson, C.M., Rau, S.W., Salimi, K., Smedley, K.L., Casey, R.L., Leserman, J., Jarskog, L.F., and Penn, D.L. (2011). Intranasal oxytocin reduces psychotic symptoms and improves Theory of Mind and social perception in schizophrenia. Schizophr Res 132, 50-53.
98. Feifel, D., Macdonald, K., Nguyen, A., Cobb, P., Warlan, H., Galangue, B., Minassian, A., Becker, O., Cooper, J., Perry, W., et al. (2010). Adjunctive intranasal oxytocin reduces symptoms in schizophrenia patients. Biol Psychiatry 68, 678-680.
99. Johnson, Z.V., and Young, L.J. (2017). Oxytocin and vasopressin neural networks: Implications for social behavioral diversity and translational neuroscience. Neurosci Biobehav Rev 76, 87-98.
100. Grinevich, V., Knobloch-Bollmann, H.S., Eliava, M., Busnelli, M., and Chini, B. (2016). Assembling the Puzzle: Pathways of Oxytocin Signaling in the Brain. Biol Psychiatry 79, 155-164.
101. Meyer-Lindenberg, A., Domes, G., Kirsch, P., and Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. Nat Rev Neurosci 12, 524-538.
102. Stoop, R. (2012). Neuromodulation by oxytocin and vasopressin. Neuron 76, 142-159.
103. Heinrichs, M., and Domes, G. (2008). Neuropeptides and social behaviour: effects of oxytocin and vasopressin in humans. Prog Brain Res 170, 337-350.
104. Bosch, O.J., Meddle, S.L., Beiderbeck, D.I., Douglas, A.J., and Neumann, I.D. (2005). Brain oxytocin correlates with maternal aggression: link to anxiety. J Neurosci 25, 6807-6815.
105. Meredith, M.E., Salameh, T.S., and Banks, W.A. (2015). Intranasal Delivery of Proteins and Peptides in the Treatment of Neurodegenerative Diseases. AAPS J 17, 780-787.
106. Olff, M., Frijling, J.L., Kubzansky, L.D., Bradley, B., Ellenbogen, M.A., Cardoso, C., Bartz, J.A., Yee, J.R., and van Zuiden, M. (2013). The role of oxytocin in social bonding, stress regulation and mental health: an update on the moderating effects of context and interindividual differences. Psychoneuroendocrinology 38, 1883-1894.
107. Bartz, J.A., Zaki, J., Bolger, N., and Ochsner, K.N. (2011). Social effects of oxytocin in humans: context and person matter. Trends Cogn Sci 15, 301-309.
108. Shamay-Tsoory, S.G., and Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. Biol Psychiatry 79, 194-202.
109. Domes, G., Sibold, M., Schulze, L., Lischke, A., Herpertz, S.C., and Heinrichs, M. (2013). Intranasal oxytocin increases covert attention to positive social cues. Psychol Med 43, 1747-1753.

110. De Dreu, C.K., Baas, M., and Boot, N.C. (2015). Oxytocin enables novelty seeking and creative performance through upregulated approach: evidence and avenues for future research. Wiley Interdiscip Rev Cogn Sci 6, 409-417.
111. Prehn, K., Kazzer, P., Lischke, A., Heinrichs, M., Herpertz, S.C., and Domes, G. (2013). Effects of intranasal oxytocin on pupil dilation indicate increased salience of socioaffective stimuli. Psychophysiology 50, 528-537.
112. Iovino, M., Messana, T., De Pergola, G., Iovino, E., Dicuonzo, F., Guastamacchia, E., Giagulli, V.A., and Triggiani, V. (2018). The Role of Neurohypophyseal Hormones Vasopressin and Oxytocin in Neuropsychiatric Disorders. Endocr Metab Immune Disord Drug Targets 18, 341-347.
113. Young, L.J., and Barrett, C.E. (2015). Neuroscience. Can oxytocin treat autism? Science 347, 825-826.
114. Domes, G., Normann, C., and Heinrichs, M. (2016). The effect of oxytocin on attention to angry and happy faces in chronic depression. BMC Psychiatry 16, 92.
115. Domes, G., Kumbier, E., Heinrichs, M., and Herpertz, S.C. (2014). Oxytocin promotes facial emotion recognition and amygdala reactivity in adults with asperger syndrome. Neuropsychopharmacology 39, 698-706.
116. Guastella, A.J., Einfeld, S.L., Gray, K.M., Rinehart, N.J., Tonge, B.J., Lambert, T.J., and Hickie, I.B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. Biol Psychiatry 67, 692-694.
117. Quintana, D.S., and Guastella, A.J. (2020). An Allostatic Theory of Oxytocin. Trends Cogn Sci 24, 515-528.
118. Deussing, J.M., and Chen, A. (2018). The Corticotropin-Releasing Factor Family: Physiology of the Stress Response. Physiol Rev 98, 2225-2286.
119. Dedic, N., Chen, A., and Deussing, J.M. (2018). The CRF Family of Neuropeptides and their Receptors - Mediators of the Central Stress Response. Curr Mol Pharmacol 11, 4-31.
120. Sanford, C.A., Soden, M.E., Baird, M.A., Miller, S.M., Schulkin, J., Palmiter, R.D., Clark, M., and Zweifel, L.S. (2017). A Central Amygdala CRF Circuit Facilitates Learning about Weak Threats. Neuron 93, 164-178.
121. Sterley, T.L., Baimoukhametova, D., Fuzesi, T., Zurek, A.A., Daviu, N., Rasiah, N.P., Rosenegger, D., and Bains, J.S. (2018). Social transmission and buffering of synaptic changes after stress. Nat Neurosci 21, 393-403.
122. Grimm, S., Wirth, K., Fan, Y., Weigand, A., Gartner, M., Feeser, M., Dziobek, I., Bajbouj, M., and Aust, S. (2017). The interaction of corticotropin-releasing hormone receptor gene and early life stress on emotional empathy. Behav Brain Res 329, 180-185.
123. Hsu, D.T., Mickey, B.J., Langenecker, S.A., Heitzeg, M.M., Love, T.M., Wang, H., Kennedy, S.E., Pecina, M., Shafir, T., Hodgkinson, C.A., et al. (2012). Variation in the corticotropin-releasing hormone receptor 1 (CRHR1) gene influences fMRI signal responses during emotional stimulus processing. J Neurosci 32, 3253-3260.
124. Sanders, J., and Nemeroff, C. (2016). The CRF System as a Therapeutic Target for Neuropsychiatric Disorders. Trends Pharmacol Sci 37, 1045-1054.
125. Iarocci, G., Yager, J., and Elfers, T. (2007). What gene-environment interactions can tell us about social competence in typical and atypical populations. Brain Cogn 65, 112-127.
126. Hostetler, C.M., and Ryabinin, A.E. (2013). The CRF system and social behavior: a review. Front Neurosci 7, 92.
127. Gehlert, D.R., Shekhar, A., Morin, S.M., Hipskind, P.A., Zink, C., Gackenheimer, S.L., Shaw, J., Fitz, S.D., and Sajdyk, T.J. (2005). Stress and central Urocortin increase anxiety-like behavior in the social interaction test via the CRF1 receptor. Eur J Pharmacol 509, 145-153.
128. Heinrichs, S.C. (2003). Modulation of social learning in rats by brain corticotropin-releasing factor. Brain Res 994, 107-114.
129. Lim, M.M., Liu, Y., Ryabinin, A.E., Bai, Y., Wang, Z., and Young, L.J. (2007). CRF receptors in the nucleus accumbens modulate partner preference in prairie voles. Horm Behav 51, 508-515.

130. Cooper, M.A., and Huhman, K.L. (2005). Corticotropin-releasing factor type II (CRF-sub-2) receptors in the bed nucleus of the stria terminalis modulate conditioned defeat in Syrian hamsters (Mesocricetus auratus). Behav Neurosci 119, 1042-1051.
131. Kobayashi, T., Kiyokawa, Y., Arata, S., Takeuchi, Y., and Mori, Y. (2013). c-Fos expression during the modulation of sexual behavior by an alarm pheromone. Behav Brain Res 237, 230-237.
132. Lee, Y., Fitz, S., Johnson, P.L., and Shekhar, A. (2008). Repeated stimulation of CRF receptors in the BNST of rats selectively induces social but not panic-like anxiety. Neuropsychopharmacology 33, 2586-2594.
133. Deussing, J.M., Breu, J., Kuhne, C., Kallnik, M., Bunck, M., Glasl, L., Yen, Y.C., Schmidt, M.V., Zurmuhlen, R., Vogl, A.M., et al. (2010). Urocortin 3 modulates social discrimination abilities via corticotropin-releasing hormone receptor type 2. J Neurosci 30, 9103-9116.

Scope and Outlines of the thesis

The overreaching goal of my thesis was to investigate the ability of mice to detect, process and react to altered affective states in conspecifics, and to dissect the neuronal substrates underlying these functions.

1. Assessing "emotion recognition" in mice

My first effort has been to develop and validate a reliable paradigm, which allows to measure mice ability to discriminate different affective states in conspecifics: the *Emotion Discrimination Task (EDT)*. In **Chapters 2** and **4** are reported the development and validation of this test, proving that mice can detect and differently react to either negative (fear or stress) or positive (relief) affective states in conspecifics. In **Chapter 5**, I report recent investigations in which I am extending the EDT in order to explore how a previous "affective" experience (i.e. stress) could modulate mice reaction to the same altered affective state in others. These extensive behavioral validations are providing an effective platform to explore in mice the neurobiology of emotion recognition and empathy.

2. Neuronal substrates of "emotion recognition" in mice

Taking advantage of our new behavioral tools, in **Chapter 2**, I report our study revealing how specific endogenous OXT signaling mediates emotion discrimination in mice. In particular, OXT projections from the paraventricular nucleus (PVN) of the hypothalamus to the Central Amygdala are critical neuronal substrates mediating this ability with important implications for social dysfunctions in psychiatric and neurodevelopmental disorders.

In **Chapter 3**, I expand this topic by presenting a commentary in which we discussed the general implication of OXT in different social contexts and potential future perspectives.

In **Chapter 4**, I report our findings on how the PFC modulates emotion discrimination abilities. In particular, the somatostatin-positive interneurons of the PFC represent a major neuronal substrate mediating these processes.

Finally, in **Chapter 5**, I show our recent investigation on the CRF system involvement in emotion discrimination and self-experience of an aversive affective state.

CHAPTER 2

Oxytocin signaling in the central amygdala modulates emotion discrimination in mice

Valentina Ferretti[1#], **Federica Maltese**[1#], Gabriella Contarini[1], Marco Nigro[1], Alessandra Bonavia[1], Huiping Huang[1], Valentina Gigliucci[2], Giovanni Morelli[1], Diego Scheggia[1], Francesca Managò[1], Giulia Castellani[1], Arthur Lefevre[3], Laura Cancedda[1], Bice Chini[2], Valery Grinevich[3,4,5], Francesco Papaleo[1]

[1]*Department of Neuroscience and Brain Technologies, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy.*

[2]*CNR, Institute of Neuroscience, Milan, Italy.*

[3]*Schaller Research Group on Neuropeptides, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany.*

[4]*CellNetwork Cluster of Excellence, University of Heidelberg, 69120 Heidelberg, Germany.*

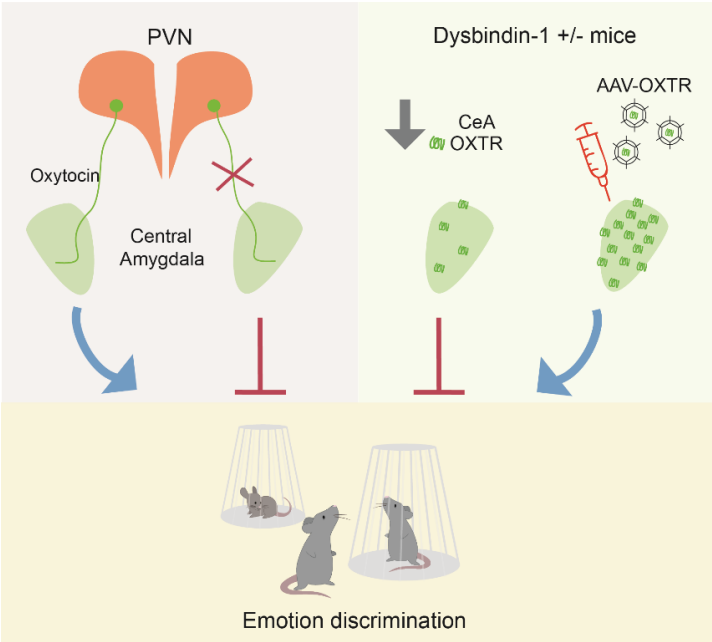
[5]*Central Institute of Mental Health, 68159 Mannheim, Germany.*

[#]co-first authors.

Published in *Current Biology*. 2019 Jun 17; 29(12):1938-1953

Abstract

Recognition of other's emotions influences the way social animals interact and adapt to the environment. The neuropeptide oxytocin (OXT) has been implicated in different aspects of emotion processing. However, the role of endogenous OXT brain pathways in the social response to different emotional states in conspecifics remains elusive. Here, using a combination of anatomical, genetic, and chemogenetic approaches, we investigated the contribution of endogenous OXT signaling in the ability of mice to discriminate unfamiliar conspecifics based on their emotional states. We found that OXTergic projections from the paraventricular nucleus of the hypothalamus (PVN) to the central amygdala (CeA) are crucial for the discrimination of both positively and negatively valenced emotional states. In contrast, blocking PVN OXT release into the nucleus accumbens, prefrontal cortex, and hippocampal CA2 did not alter this emotion discrimination. Furthermore, silencing each of these PVN OXT pathways did not influence basic social interaction. These findings were further supported by the demonstration that virally mediated enhancement of OXT signaling within the CeA was sufficient to rescue emotion discrimination deficits in a genetic mouse model of cognitive liability. Our results indicate that CeA OXT signaling plays a key role in emotion discrimination both in physiological and pathological conditions.



Graphical Abstract

Introduction

Social interactions are influenced by the ability to decipher expressions of emotions in others [1–3]. Disturbances in this capacity, defined as ‘‘social cognition’’ [2, 3], represent a distinctive feature of many psychiatric, neurodevelopmental, and neurodegenerative neurodegenerative disorders [3]. For instance, abnormalities in social cue identification define autism spectrum disorders [4, 5]. Similarly, patients with schizophrenia show marked impairments in the processing of non-verbal social affective information while showing normal affect sharing and emotional experience [6]. Despite the deleterious impact on the everyday life of these subjects [7], social cognitive impairments still lack an effective treatment.

The oxytocin (OXT) system is considered a major player in social information processing and social cognition [8–10]. Rodent studies implicate the OXT system in a number of social domains, such as processing of sensory stimuli, social recognition [11–16], social memory [17, 18], consolation [19], social reward [20, 21], response to fear [22–24], and sexual and parental behaviors [25, 26]. Intranasal administration of OXT in humans, despite controversial and variable effects [27–32], has been reported to modulate recognition of emotions, empathy, and trust [8, 9, 33–39]. Genetic-association studies also support an implication of the OXT system in emotion processing [10, 40, 41]. However, the role of the endogenous OXT system and its potential modulation by genetic background in the perception and processing of other’s emotions is still underexplored.

Here, in line with increasing evidence that higher-order social emotional processes can be studied in rodents [42], we implemented a behavioral setting that could approximate some features of human emotion recognition tasks [3, 43] or similar tasks used in non-human primates, dogs, sheep, and horses [44–47], to dissect the implication of the endogenous OXT system in emotion discrimination. In particular, using a chemogenetic approach, we explored the role of selected OXT projections from the paraventricular nucleus of the hypothalamus (PVN) in the ability of mice to discriminate unfamiliar conspecifics based on negatively or positively valenced emotional states. Furthermore, to start investigating the potential modulation by genetic background, we tested how a genetic variant (in dysbindin-1) with clinical relevance for cognitive and psychiatric liability [48–51] might modulate emotion discrimination through selective alterations in the endogenous OXT system.

Together, our results revealed an essential role of the PVNCeA OXT pathway in the discrimination of the emotions of others and that genetic background and variation in OXTR expression within the CeA can moderate these effects.

Results

Mice Discriminate Unfamiliar Conspecifics based on Negatively Valenced Emotional States

To test whether mice could discriminate unfamiliar conspecifics expressing different emotional states, we placed an ‘‘observer’’ mouse in a cage containing two age- and sex-matched unfamiliar conspecifics (‘‘demonstrators’’). The demonstrators were placed in wire cups to allow visual, tactile, auditory, and olfactory communication while avoiding aggressive or sexual interactions (Figure 1A). The task was thus centered on behaviors initiated by the observer when simultaneously exposed to a neutral demonstrator and to a mouse in an altered emotional state.

In the ‘‘fear’’ manipulation, one of the two demonstrators was fear conditioned to a tone cue at least 1 day before the test (Figure 1B). Upon presentation during the test, the tone would then evoke a negatively valenced emotional state in the conditioned mouse [52]. In particular, the tone was delivered during the second 2-min epoch of the test (Figure 1B) in order to assess observers’ responses before, during, and after the induction of the altered emotional state in the demonstrators. Consistently, we observed a freezing response in the fear demonstrator only during the 2-min tone presentation, associated with a reduction in rearing (Figure S1A). No other behavioral parameters differed between the two demonstrators during the 6-min test session (Figure S1A). Both male and female observers increased their sniffing, but not any other observable behavior, toward the fear-conditioned demonstrator compared to the neutral one (Figures 1C, 1D, 1G–1I, and S2A–S2C). This effect became evident after the 2-min tone presentation (Figures 1C, 1D, S2A, and S2C). Although no discriminatory behavior was observed during the tone presentation, we found an inverse correlation between the time the fear demonstrator spent freezing and the time of observer sniffing, suggesting that freezing per se might influence observer discrimination (Figure 1E). However, we found no correlation between the demonstrator freezing and the observer sniffing after the tone presentation, suggesting that demonstrator’s freezing did not affect the discriminatory behavior we observed (Figure 1F).

In light of previous evidence [23, 53, 54], we searched for signs of fear transfer from the emotionally altered demonstrator to the observer by quantifying freezing behavior, escape attempts, changes in locomotor activity, and other stress-related behaviors (i.e., rearing and grooming). During the 6-min test, we detected no sign of emotion contagion (Figures 1G–1I). Moreover, corticosterone levels of observer mice exposed to the fear paradigm or to two neutral demonstrators did not differ (Figure 1J). These findings suggest that mice can detect and socially respond to unfamiliar conspecifics in a negatively valenced emotional state.

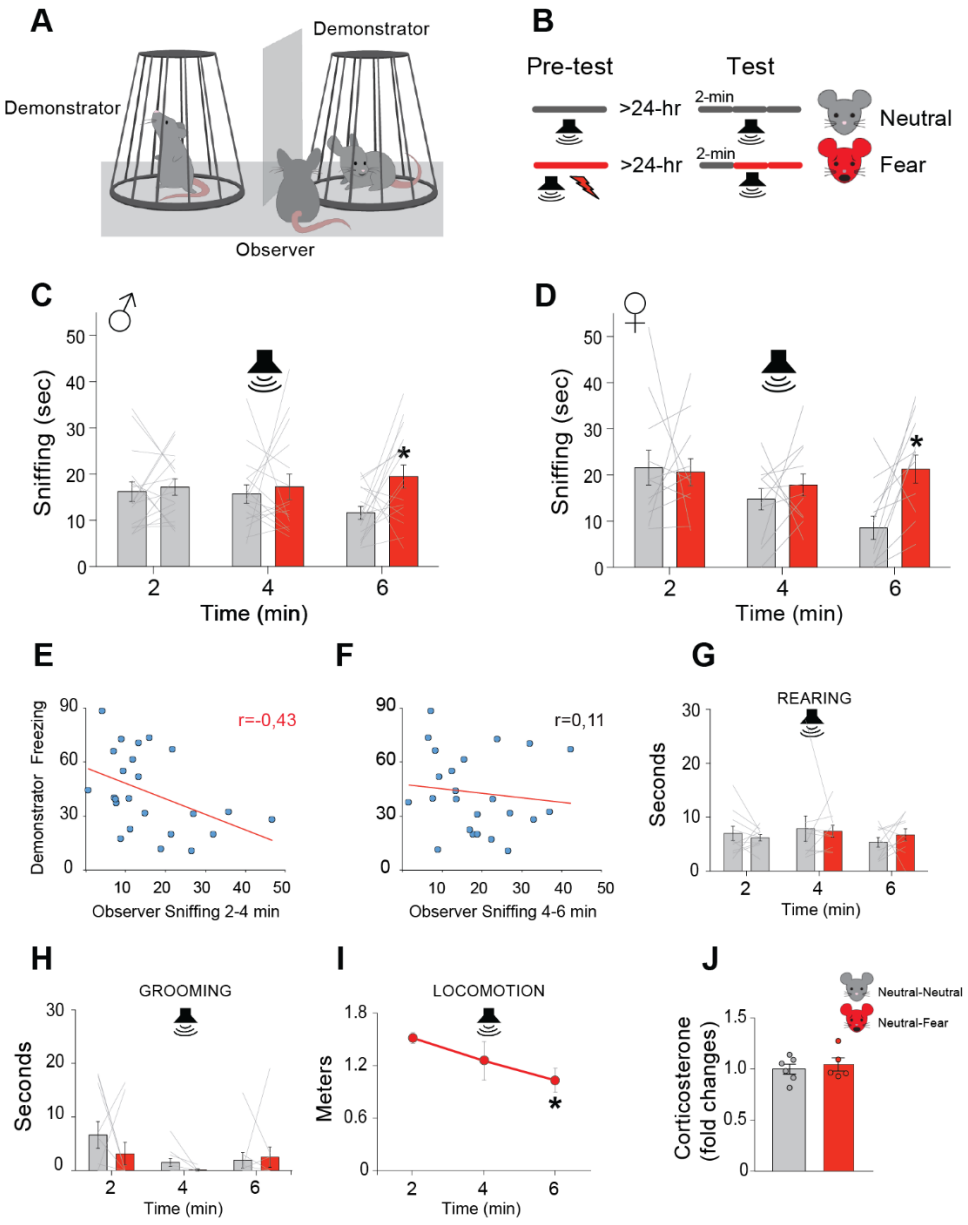


Figure 1. Mouse Emotion Discrimination for Fear. (**A**) Schematic drawing of the test setting. (**B**) Timeline of pre-test and test procedures to evoke in one of the demonstrators a ‘‘fear’’ state by delivering the conditioned tone in the 2- to 4-min epoch. (**C-D**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or tone-induced fear (red bars) state displayed by (**C**) male and (**D**) female observer mice during the 6-min test, divided into three consecutive 2-min epochs (last 2-min repeated measurements [RM] ANOVA for males $F_{1,15} = 6.51$, $p = 0.022$, and females $F_{1,11} = 10.98$, $p = 0.006$; no significant differences in the 0- to 2-min and 2- to 4-min epochs). *$p < 0.05$ versus the exploration of the neutral demonstrator. n = 8/15 observers per group. (**E-F**) Correlation analyses between the time the fear-conditioned demonstrator spent freezing (in y axis) and the time the observer spent sniffing the fear conditioned demonstrator (in x axis; **E**) in the 2- to 4-min epoch or (**F**) in the 4- to 6-min epoch of the test ($r = -0.4310$ for 2–4 min; $r = -0.11$ for 4–6 min). n = 24 observers. (G and H) Time (in seconds) spent (**G**) rearing and (**H**) grooming in proximity of the demonstrators in neutral (gray bars) or fear states (red bars) displayed by the same observer mice during the test (RM ANOVAs showed no significant differences). (**I**) Locomotor activity displayed by the same observer mice during the test (RM ANOVA $F_{2,16} = 4.08$; $p = 0.03$). *$p < 0.05$ versus 0–2 min. n = 9 observers. (**J**) Blood corticosterone levels displayed by observer mice immediately after being exposed to two neutral demonstrators (gray bar) or one neutral and one fear demonstrator (red bar). Data are expressed as fold changes compared to observers exposed to two neutral demonstrators (t test: df: 9; $p = 0.58$). n = 5/6 observers per group.
Error bars represent standard error of the mean. See also Figures S1 and S2.

Mice Discriminate Unfamiliar Conspecifics based on Positively Valenced Emotional States

We next investigated whether observer mice could discriminate unfamiliar conspecifics by detecting a positively valenced state. In particular, we exposed observer mice to a neutral demonstrator and to a demonstrator that received 1-h ad libitum access to water after 23 h of water deprivation (Figures 2A and 2B). Water was selected as a rewarding stimulus to avoid odorrelated cues that could differentiate the two demonstrators. We assumed that the relief from the distressing water deprivation would result in a positively valenced emotional state (‘‘relief’’). Consistently, we found that the 1-h ad libitum access to water resulted in a conditional place preference in mice that experienced the 23 h water deprivation, but not in mice in ad libitum water condition (Figures 2E and 2F). Moreover, 1-h ad libitum access to water after the 23-h deprivation reduced corticosterone levels in relief mice (Figure 2G). Furthermore, the relief manipulation induced no detectable behavioral alteration during the test compared to neutral demonstrators (Figure S1B).

Observers of both sexes showed increased social exploration toward the relief demonstrator compared to the neutral, selectively in the first 2 min of the task (Figures 2C, 2D, S2B, and S2D). No changes in rearing and grooming patterns toward the demonstrators and throughout the task were evident (Figures 2H and 2I). Moreover, observers showed the typical decrease in locomotor activity (Figure 2J) and did not show freezing behavior, escape attempts, or other stress-related behaviors during the entire test session. Furthermore, no alteration in corticosterone levels was detected between observers exposed to relief-neutral or neutral-neutral demonstrators (Figure S2I). These findings indicate that mice can detect and socially respond to unfamiliar conspecifics in a positively valenced emotional state.

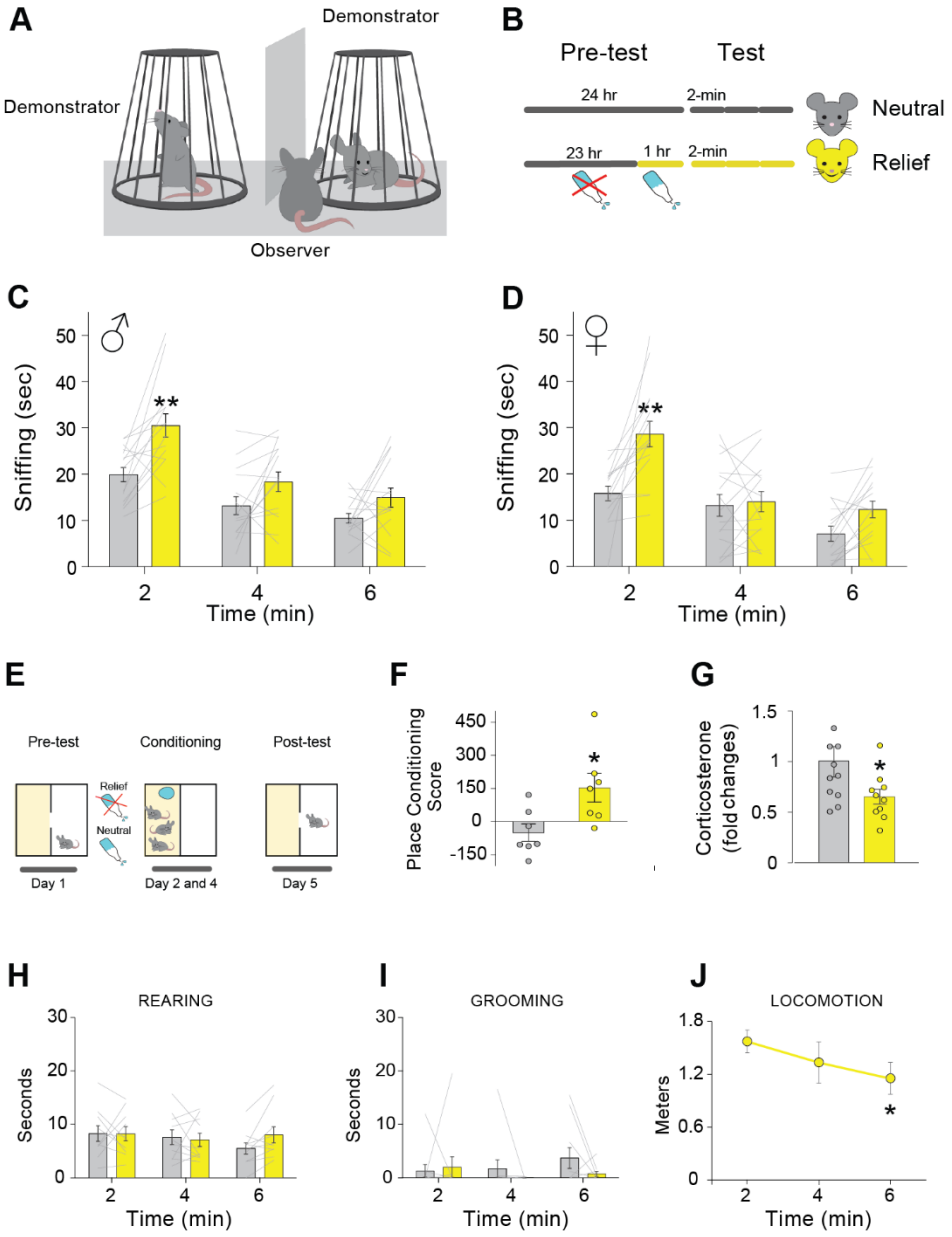


Figure 2. Mouse Emotion Discrimination for Relief. (**A**) Schematic drawing of the test setting. (**B**) Timeline of pre-test and test procedures to evoke in one of the two demonstrators a ‘‘relief’’ state during the testing phase. (**C-D**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or water-induced relief (yellow bars) states displayed by (**C**) male and (**D**) female observer mice during the 6 min of the test, divided into three consecutive 2-min epochs (first 2-min RM ANOVA for males $F_{1,14} = 15.07$, $p = 0.001$ and females $F_{1,14} = 14.60$, $p = 0.001$; no significant differences for the 2- to 4-min and 4- to 6-min epochs). **$p < 0.005$ versus the exploration of the neutral demonstrator. n= 15 observers per group. (**E**) Place conditioning procedure used to assess whether the relief manipulation was associated with a negative-, neutral-, or positive-valence affective state. (**F**) Place conditioning scores (in seconds) displayed by mice conditioned during a neutral (gray bar) or relief (yellow bar) state. For each mouse, a place conditioning score was calculated as the post- minus the pre-conditioning time spent in the conditioning-paired compartment of the apparatus. A positive score indicates place preference, a negative score a place aversion, and 0 no place conditioning (t test: df = 12; $p = 0.02$). *$p < 0.05$ versus the neutral control group. n = 7 per group. (**G**) Blood

corticosterone levels displayed by demonstrator mice immediately after a period of 24-h water deprivation (gray bar) or after a period of 1 h ad libitum access to water following 23-h water deprivation (yellow bar; t test: df: 19; p = 0.05). *p = 0.05 versus water deprived mice. n = 11 per group. (**H-I**) Time (in seconds) spent in (**H**) rearing and (**I**) grooming in proximity of demonstrators in neutral (gray bars) or relief (yellow bars) state displayed by observer mice during the test (RM ANOVAs showed no significant differences). (J) Locomotor activity displayed by the same observer mice during the test (RM ANOVA $F_{2,18} = 4.35$; p = 0.04). *p < 0.05 versus minute 0–2. n = 10 observers. Error bars represent standard error of the mean. See also Figures S1 and S2.

Sensory Modalities Implicated in Emotion Discrimination

Different sensory modalities might mediate social responses to emotional stimuli in different animal species [1, 55]. We explored the implication of visual, auditory, and olfactory cues in mice

emotion discrimination in our setting.

Preventing visual cues can reduce emotion contagion responses induced by the observation of a partner in distress [53, 54]. However, when we performed the test in complete darkness (Figure 3A), mice were still able to discriminate between a neutral and an emotionally altered conspecific (Figures 3B and 3C), similarly to that observed in standard lighting conditions (Figures 1 and 2). This suggests that mice can discriminate emotions even in the absence of visual cues. Notably, removal of visual cues anticipated observer discrimination of the fear demonstrator to the tone epoch (Figure 3B). In light of the negative correlation described (Figure 1E), this finding further suggests that observing a mouse freezing might negatively influence observers' social approach. Auditory cues in the form of ultrasonic vocalizations (USVs) might be used by rodents for social communication [1, 55–58]. However, when we recorded USVs in the fear and relief conditions (Figure 3D), we detected very few vocalizations in a negligible number of mice (3/12) and with no differences in frequency, duration, or amplitude among the different conditions (Figure 3E). This was further confirmed when we individually recorded neutral, fear, and relief demonstrators (Figure S1C). Our data support previous evidence [58] that, in contrast to rat use of USVs [57, 59–61], adult mice might not use USVs as the main modality by which they communicate emotional states. Despite this, we cannot exclude the possibility that other unidentified auditory cues might be involved (e.g., [62]).

Finally, we tested the impact of olfactory cues presenting the observers with cotton balls after being swiped throughout the body, head, and anogenital areas of a neutral, fear, or relief demonstrator immediately after the manipulations (Figure 3F). Observer mice showed an avoidance for the odor of a fear mouse (Figure 3G), in agreement with previous evidence [63–65],

and spent more time sniffing the relief odor compared to the neutral one (Figure 3H). These results confirm that olfactory cues convey information related to mouse emotional states but also indicate that observer responses are qualitatively different when demonstrators are physically present. Overall, this set of data indicates distinct implications of visual and olfactory social cues in the response to emotional stimuli.

Distinction between Emotion Discrimination and Sociability

The ability to discriminate between different states and the absolute quantity of social interaction (here referred to as ‘‘sociability’’) might be considered distinct constructs of social processes that are not necessarily interdependent [66]. Thus, we tested whether equivalent changes in sniffing behavior could be observable also in a context of no choice alternatives, presenting to observer mice a neutral, a fear, or a relief conspecific in a one-on-one free interaction setting. Social exploration levels did not differ between conditions (Figure 3I), showing the expected decrease of interaction over time [67]. Supporting this dichotomous effect, we revealed no correlation between emotion discrimination and the amount of social interaction in our setting (Figure S2H). Overall, this indicates that our setting (Figures 1 and 2) reveals aspects of social behavior not related to sociability.

Previous evidence measuring affective responses of an observer rat exposed to a demonstrator immediately after shock showed a similar decreased social exploration in one-on-one and one-on-two settings [16]. The discrepancy with our results, in which the one-on-one and one-on-two settings gave different results, might be due to the scalability feature of emotions [68], but not to mouse-rat differences. Indeed, using the same shock manipulations used in rats [16] in our setting (Figure 3J) recapitulated a similar general aversion during the whole test session (Figure 3K). This suggests that using higher-intensity emotional states might prove difficult to differentiate emotion discrimination from sociability measures.

Taken together, these data suggest that the paradigm used in this work (Figures 1 and 2) might reveal specific behavioral responses to mildly graded expression of emotions, possibly underestimated by previously used social interaction tests.

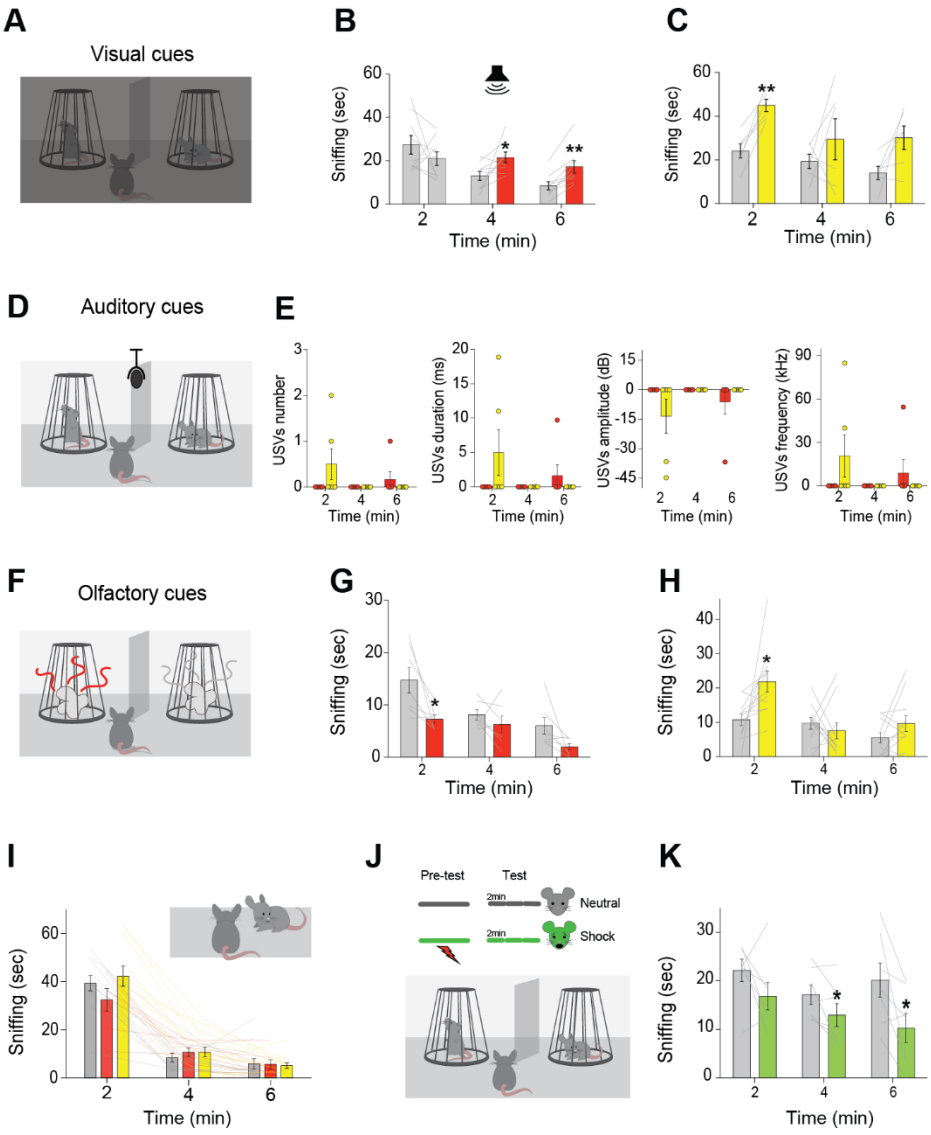


Figure 3. Analyses of Sensory Modalities and Distinction from Sociability. (**A**) Schematic drawing of the test setting performed in complete darkness. (**B-C**) Time (in seconds) spent by observer mice sniffing demonstrators during the 6 min of the (**B**) fear and (**C**) relief manipulations of the emotion discrimination test. Time spent sniffing neutral demonstrators is depicted in gray. Time spent sniffing (**B**) fear or (**C**) relief demonstrators are depicted in red and yellow, respectively (RM ANOVAs for fear, 2–4 min: $F_{1,8} = 5.63$, $p = 0.04$; 4–6 min: $F_{1,8} = 28.08$, $p = 0.0007$; for relief, 0–2 min: $F_{1,5} = 33.32$, $p = 0.002$). *$p < 0.05$ and **$p < 0.005$ versus the exploration of the neutral demonstrator. n = 6/9 observers per group. (**D**) Schematic drawing of the test setting to record USVs. (**E**) USV mean of number of calls, duration in milliseconds, amplitude in decibel, and frequency in KHz emitted by mice during the fear and relief settings (two-way ANOVAs showed no significant differences). n = 6 observers per group. (**F**) Schematic drawing of the test setting performed only with demonstrators' odors for fear and relief conditions. (**G-H**) Time (in seconds) spent by observer mice sniffing the odors from neutral (gray), fear (red), or relief (yellow) demonstrators during the 6 min of the (**G**) fear or (**H**) relief test (RM ANOVA for the fear manipulation, 0–2 min: $F_{1,6} = 9.15$, $p = 0.02$. No significant differences for the 2- to 4-min and 4- to 6-min test periods. RM ANOVA for the relief manipulation, 0–2 min: $F_{1,11} = 6.89$; $p = 0.02$. Similarly, no significant differences for the 2- to 4-min and 4- to 6-min test periods). *$p < 0.05$ versus the exploration of the neutral odor. n = 7/11 observers per group. (**I**) Schematic drawing of the one-on-one test setting and time (in seconds) spent by observer mice sniffing a single demonstrator in a neutral (gray), fear (red), or relief (yellow) state during a 6-min free interaction test. The tone to which only the fear demonstrator was fear conditioned

was delivered between 2 and 4 min of the test (ANOVAs revealed only a time effect with normal decreased exploration throughout the 6 min, $F_{2,56} = 132.01$; $p < 0.0001$). n = 12 observers. (**J**) Schematic drawing of the task setting and timeline of pre-test and test procedures to trigger in one of the demonstrator a ''shock'' emotional state. (**K**) Time (in seconds) spent by the observer mice sniffing demonstrators in neutral (gray bars) or shocked emotional state (green bars) during the test (RM ANOVAs, 0–2 min: $F_{1,6} = 2.40$, $p = 0.17$; 2–4 min: $F_{1,6} = 5.43$, $p = 0.05$; 4–6 min: $F_{1,6} = 8.11$, $p = 0.02$). *$p < 0.05$ versus the exploration of the neutral demonstrator. n = 7 observers.
Error bars represent standard error of the mean. See also Figure S1.

Endogenous Release of Oxytocin Is Necessary for Emotion Discrimination

The OXT system plays a pivotal role in social perception and cognition [8–10]. In particular, in humans, OXT has been associated with social cognitive functions, such as emotion recognition, empathy, and trust [8–10, 33, 35, 38, 39]. To assess whether the OXT system might be implicated in mouse emotion discrimination, we prevented OXT release from the PVN by bilateral injections of a recombinant adeno-associated virus (rAAV) expressing the inhibitory hM4D(Gi) DREADD (designer receptors exclusively activated by designer drugs) receptor under the control of an OXT promoter (Figures 4A, S3A, and S4A). In rodents, neurons in the PVN are the main source of OXT projections to the brain [22, 69]. We found that, in contrast to vehicle treatment (Veh), inhibition of PVN-OXT-projecting neurons (upon clozapine N-oxide [CNO] injection) abolished the ability of mice to discriminate either fear or relief states in conspecifics (Figures 4B, 4C, S3B, and S4F). This was equally evident in male (Figures 4B, 4C, and S4F) and female mice (Figure S3B) and was not associated with unspecific CNO effects, which suggests that CNO per se did not affect mouse responses during the test (Figures S3C and S3D). Notably, we were able to test the same mice in the different conditions as emotion discrimination ability was preserved when an observer was re-exposed to the same or to different paradigms (Figures S2E–S2G). Inhibiting PVN OXT projections produced selective effects on emotion discrimination. In fact, CNO administration had no effect on social exploration when the same mice were tested in a one-on-one free-interaction setting with an unfamiliar conspecific (Figure S4K). Overall, these data indicate a direct involvement of OXT release from PVN in the ability to discriminate different emotional states in conspecifics.

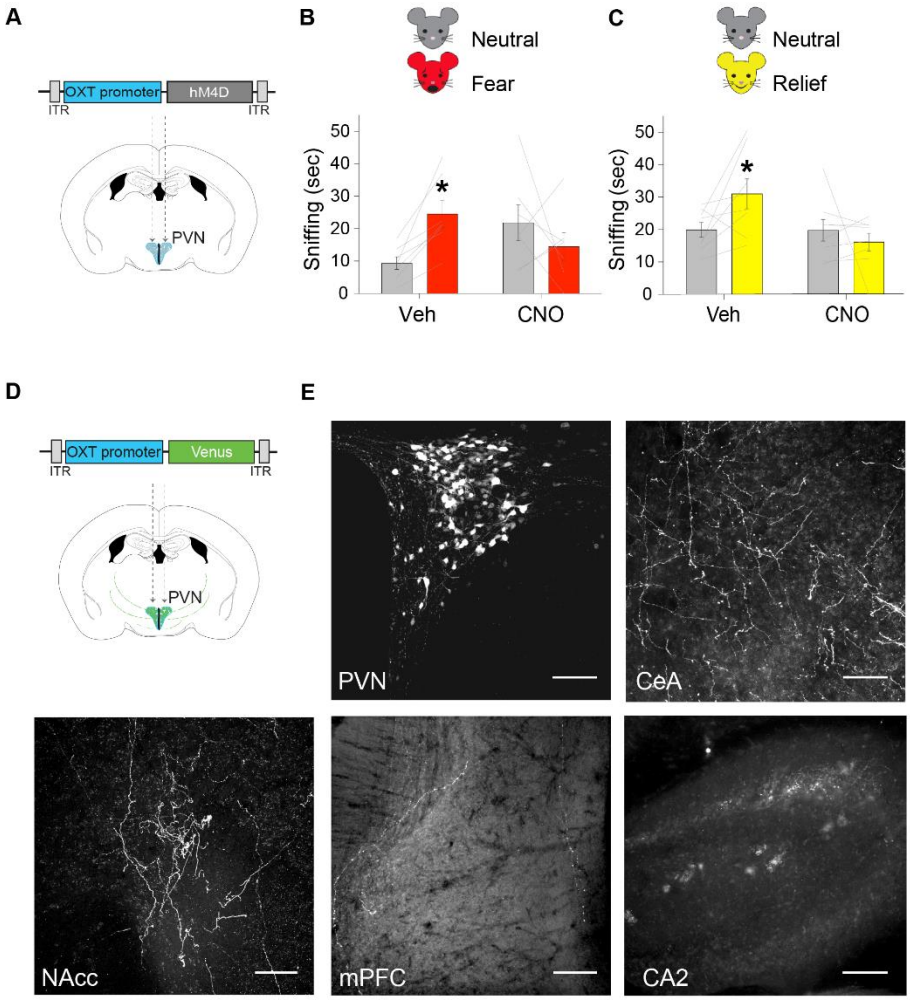


Figure 4. PVN OXT Release Is Necessary for Emotion Discrimination: Mapping of PVN OXT Projections in Mice. (**A**) Scheme of the viral vector used to infect the PVN OXT neurons. Injection placements in Figure S5. (**B-C**) Time (in seconds) spent sniffing the two demonstrators during the last 2 min (for the fear) or the first 2 min (for the relief) epochs of the emotion discrimination test, displayed by the same observer mice treated with vehicle or CNO (intraperitoneal [i.p.], 3 mg/kg in a volume of 10 mL/kg, 30 min before the test) and shown separately for each demonstrator's state. The sniffing time of the different conditions throughout the whole 6-min period of the test are reported in Figure S4. Time spent sniffing (**B**) fear (last 2-min RM ANOVA veh: $F_{1,6} = 19.07$, $p = 0.004$; CNO: $F_{1,6} = 0.85$, $p = 0.39$) or (**C**) relief demonstrators (first 2-min RM ANOVA veh: $F_{1,7} = 7.24$, $p = 0.031$; CNO: $F_{1,7} = 0.50$, $p = 0.05$) is represented by red or yellow bars, respectively. *$p < 0.05$ versus the neutral demonstrator within the same observer treatment. n = 6/8 observers per group. (**D**) Scheme of the viral vector used to infect the PVN OXT neurons. (E) Anatomy of OXT projections from the PVN to the central amygdala (CeA), nucleus accumbens (NAcc), medial prefrontal cortex (mPFC), and hippocampusCA2 (CA2). Scale bars: 100 mm.
Error bars represent standard error of the mean. See also Figures S3–S5.

PVN OXT Projections to the Central Amygdala Are an Essential Neural Substrate for Emotion Discrimination

To investigate the selective OXTergic circuits involved in emotion discrimination, we first visualized PVN OXT projections, injecting mice with a rAAV-expressing Venus under the control of an OXT promoter that allowed the fluorescent labeling of OXT-PVN neurons (Figures 4D and S5A). We focused on brain areas that have been identified as potential neuroanatomical substrates of emotion discrimination in humans [43, 70, 71] and that presented OXTergic innervation. We identified OXT-positive fibers in the central amygdala (CeA), nucleus accumbens (NAcc), hippocampal CA2, and medial prefrontal cortex (mPFC) (Figure 4E). Fewer fibers were evident in the insula, basolateral amygdala (BLA), and medial amygdala (MeA) (Figure S5B).

Next, to investigate the functional role of selective PVN OXT projections, we injected a retrogradely transported canineadenovirus- 2-expressing Cre recombinase (CAV2-Cre) into the target areas (CeA, NAcc, CA2 or mPFC) and also injected the PVN with a rAAV carrying a double-floxed inverted open reading frame (ORF) (DIO) of hM4D(Gi)DREADD receptor and mCherry under the control of the OXT promoter [69]. With this combination, we achieved DREADD(Gi)-mCherry expression exclusively in PVN OXT neurons projecting to the area injected with CAV2- Cre. We verified the regional and cell type specificity of virally mediated labeling of OXT neurons (Figures 5E and S5H–S5L). Due to the heterogeneity in fiber distribution in the different target areas, we controlled for the efficacy of DREADD-mediated inhibition in PVN back-labeled neurons from the different projection sites by performing ex vivo patch clamp electrophysiology recordings on PVN slices. We found a significant reduction in the number of evoked spikes after CNO application in back-labeled PVN neurons, which was equivalent for areas with intense OXTergic innervations (i.e., CeA and CA2) or with more sparse innervations (i.e., PFC; Figures 5F, 5G, and S5C–S5G).

In vivo, selective inhibition of OXT neurons projecting from the PVN to the CeA (Figures 5A, S4B, and S4G) was sufficient to recapitulate the deficits in emotion discrimination found by silencing all PVN projections (Figures 4B, 4C, and S3B). The same mice showed normal emotion discrimination abilities when treated with vehicle (Figures 5A and S4G). In contrast, selective inhibition of OXT neurons projecting from the PVN to the NAcc (Figures 5B, S4C, and S4H), the mPFC (Figures 5C, S4D, and S4I), and the CA2 (Figures 5D, S4E, and S4J) did not interfere with the ability to distinguish emotional states in conspecifics, indicating that OXT release from PVN

to these brain regions is dispensable for emotion discrimination. Finally, none of the OXT pathways manipulations altered the ability to interact with an unfamiliar conspecific in a one-on-one free-interaction setting (Figures S4K–S4O), further supporting the distinction between emotion discrimination and sociability measures. Overall, these findings demonstrate a pre-eminent contribution of the CeA in emotion discrimination abilities in mice and indicate that PVN OXTergic projections to the CeA are an essential neural substrate of this social cognitive function.

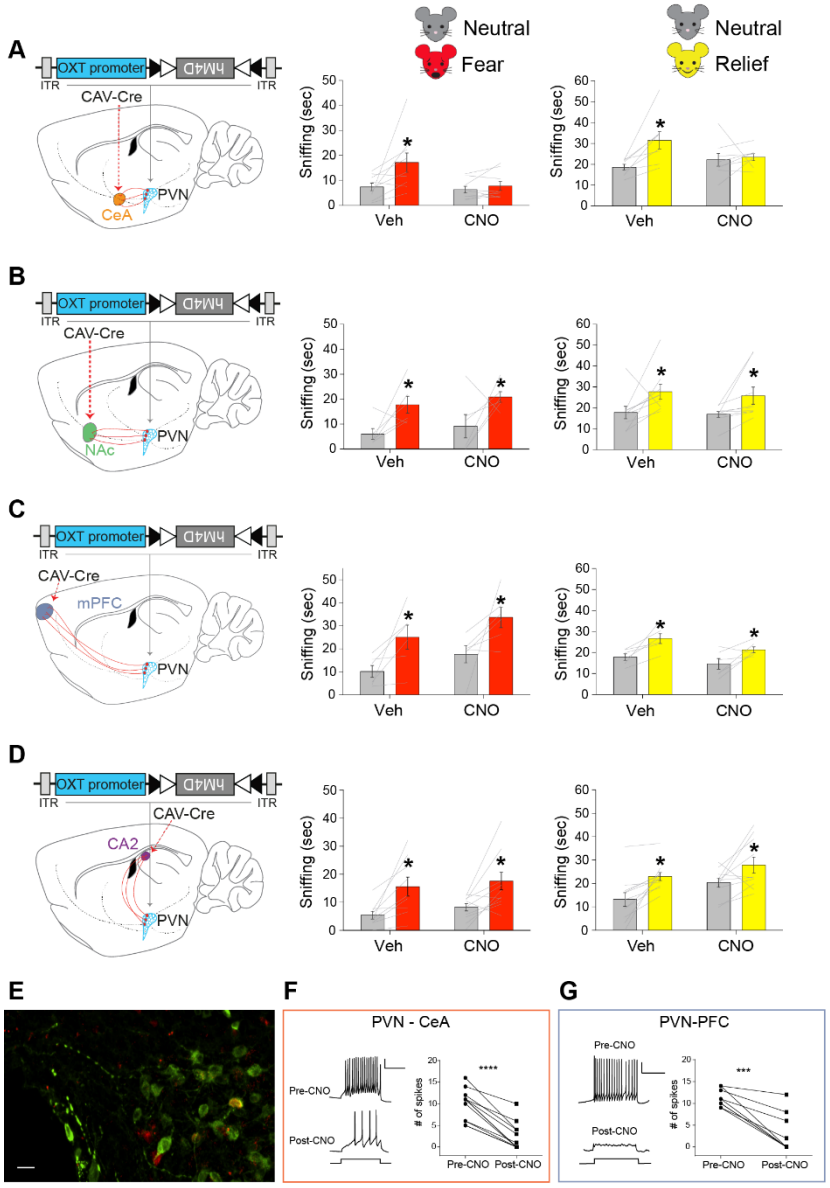


Figure 5. PVN-Central Amygdala OXT Projections Are Necessary for Emotion Discrimination in Mice. (**A–D**) Schemes showing the injection of viruses in the PVN and respective projection areas (CeA, NAcc, mPFC, or CA2). Placements are shown in Figure S5. Time (in seconds) spent sniffing the two demonstrators during the last 2-min (for

the fear) or the first 2-min (for the relief) epochs of the emotion discrimination test is shown, displayed by the same observermice treated with vehicle or CNO (i.p., 3 mg/kg in a volume of 10 mL/kg, 30 min before the test), and shown separately for each demonstrator's state. Time spent sniffing neutral demonstrators is depicted in gray. Time spent sniffing fear or relief demonstrators is represented by red or yellow bars, respectively. RM ANOVAs: (**A**) PVN-CeA (fear veh: $F_{1,8} = 5.76$, $p = 0.043$; CNO: $F_{1,8} = 1.57$, $p = 0.25$; relief veh: $F_{1,7} = 12.66$, $p = 0.009$; CNO: $F_{1,7} = 0.13$, $p = 0.73$); (**B**) PVN-NAcc (fear veh: $F_{1,5} = 6.02$, $p = 0.05$; CNO: $F_{1,5} = 7.40$, $p = 0.042$; relief veh: $F_{1,8} = 7.56$, $p = 0.025$; CNO: $F_{1,8} = 6.09$, $p = 0.039$); (**C**) PVNmPFC (fear veh: $F_{1,5} = 7.25$, $p = 0.043$; CNO: $F_{1,5} = 6.80$, $p = 0.048$; relief veh: $F_{1,5} = 11.86$, $p = 0.02$; CNO: $F_{1,5} = 7.10$, $p = 0.044$); (**D**) PVN-CA2 (fear veh: $F_{1,10} = 11.39$, $p = 0.009$; CNO: $F_{1,10} = 11.15$, $p = 0.007$; relief veh: $F_{1,10} = 16.91$, $p = 0.002$; CNO: $F_{1,10} = 6.60$, $p = 0.04$). n = 6/11 observers per group. *$p < 0.05$ versus the exploration of the neutral demonstrator. The sniffing time throughout the whole 6-min period of the different conditions is reported in the Figure S4. (**E**) Immunohistochemical staining for mCherry (red) and OXT (green) of back-labeled PVN neurons. Scale bar: 20 mm. (**F-G**) Electrophysiological validation of hM4D(Gi) action in PVN back-labeled neurons from (**F**) CeA and (**G**) mPFC. Example traces of membrane potential changes (left) and quantification (right) of single-cell data points of the number of spikes evoked by a depolarizing current step (duration: 1 s; amplitude: 20 pA; I_{Hold}: 0 pA) in PVN neurons, pre- and post-bath application of CNO in artificial cerebrospinal fluid (ACSF). Two-tailed paired t test: (**F**) PVN-CeA, n = 12 from 3 mice ($t = 9.033$; $df = 11$; $p < 0.0001$); (**G**) PVN-mPFC, n = 9 from 3 mice ($t = 6.596$; $df = 8$; $p = 0.0002$). Scale bars are 40 mV and 500 ms. ****$p < 0.0001$.
Error bars represent standard error of the mean. See also Figures S3–S5.

Emotion Discrimination Depends on OXTR Levels in the CeA

Altered amygdala reactivity in emotion discrimination has been consistently reported in autism and schizophrenia in association with genetic liability [72, 73]. In heterozygous knockout mice for dysbindin-1 (Dys+/_), a clinically relevant mouse model of cognitive and psychiatric liability [48–50], we identified reduced expression levels of OXT receptors (OXTR) in the CeA, but not in the BLA or MeA compared to wild-type littermates (Dys+/+; Figures 6A and 6B). We then assessed Dys+/_ mouse emotion discrimination abilities and observed deficits in both the fear and relief conditions (Figures 6C and 6D). In particular, we found that the impact of Dys mutation was selective for emotion discrimination, as Dys+/_ sociability and social novelty in the classic 3-chamber test were similar to Dys+/+ controls (Figures S6A and S6B). These data unravel a clinically relevant genetic variation, which concurrently leads to deficits in emotion discrimination and to changes in the CeA OXT system.

To test whether reduced OXTR levels in the CeA were responsible for Dys+/_ mouse emotion discrimination deficits, we increased the expression of OXTR selectively in the CeA of Dys+/- mice by bilateral injection of the AAV-EF1a-OXTRIRES- EYFP, expressing OXTR and EYFP (enhanced yellow fluorescent protein) under the control of the EF1a promoter [24] (Figures 6E and S6D). Increased OXTR levels within the CeA were confirmed by receptor autoradiography quantification (Figures S6C–S6F). Increasing OXTR levels in CeA of Dys+/- mice was sufficient to rescue their emotion discrimination deficits (Figures 6F and 6G). Altogether, these findings

strengthen the conclusion that appropriate OXTergic signaling within the CeA is critical to discriminate conspecifics based on their emotional state.

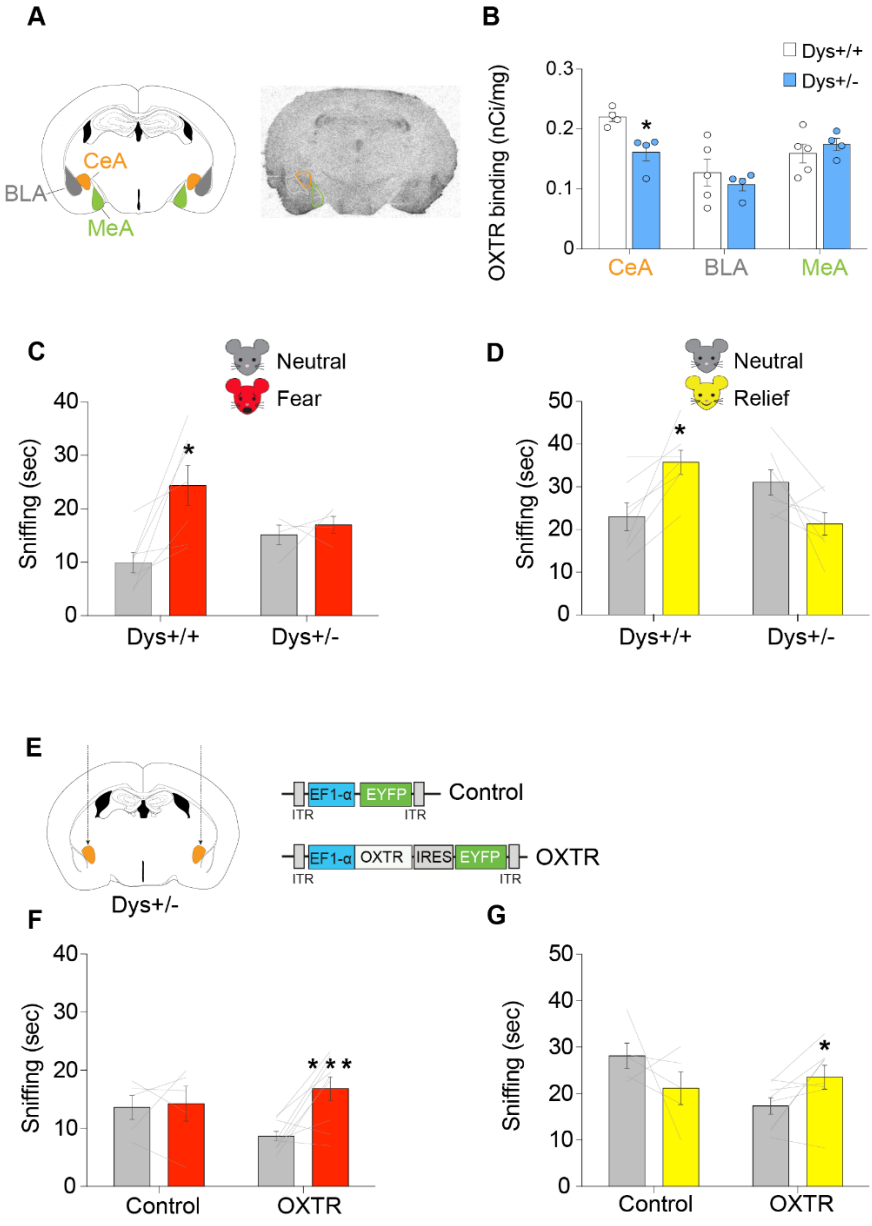


Figure 6. Emotion Discrimination Abilities are Genetically Modulated and depend on OXTR levels in the CeA. (**A**) Representative drawing and autoradiograph showing the ligand binding of 20 pmol/L I125-labeled OVTA, a potent and selective ligand for OXTR. Autoradiograms were obtained from coronal sections of brains of dysbindin-1 wild-type (Dys+/+) or heterozygous (Dys+/-) mice. CeA, central amygdala; BLA, basolateral amygdala; MeA, medial amygdala. (**B**) OXTR binding sites are expressed as nCi/mg of tissue equivalent. One-way ANOVA for CeA, $F_{1,6}$ = 12.5035, p = 0.01; BLA, $F_{1,7}$ = 0.53027, p = 0.5; and MeA, $F_{1,7}$ = 0.5609, p = 0.48. *p < 0.05 versus Dys+/+ mice. n = 4/5 for each group. (**C-D**) Time (in seconds) dysbindin-1 wild-type (Dys+/+) or heterozygous (Dys+/-) observer mice spent sniffing wild-type demonstrators in the two versions of the emotion discrimination test. Time spent sniffing neutral demonstrators is represented by gray bars. Time spent sniffing (**C**) fear demonstrators (only the last 2 min are displayed) or (**D**) water-induced relief state (only the first 2 min are displayed) are represented by red or yellow bars,

respectively. RM ANOVAs: (**C**) fear last 2 min, Dys+/+: $F_{1,6} = 12.41$, $p = 0.012$; Dys+/-: $F_{1,3} = 0.33$, $p = 0.61$; (**D**) relief first 2 min, Dys+/+: $F_{1,6} = 12.24$, $p = 0.012$; Dys+/-: $F_{1,6} = 5.11$, $p = 0.06$. n = 4/9 observers per group. *$p < 0.05$ versus the exploration of the neutral demonstrator. (**E**) Scheme of the viral vector used to infect the CeA of Dys+/- mice. Injection sites are shown in Figure S6. (**F-G**) Time (in seconds) spent sniffing each of the two wild-type demonstrators during the test displayed by dysbindin-1 heterozygous (Dys+/-) observer mice bilaterally injected with AAV-EF1a-EYFP (control) or AAV-EF1a-OXTREYFP (OXTR) and shown separately for each demonstrator's emotion. Time spent sniffing neutral demonstrators is represented by gray bars. Time spent sniffing demonstrators with (**F**) tone-induced fearful state (only the last 2 min are displayed) or (**G**) water-induced relief state (only the first 2 min are displayed) is represented by red or yellow bars, respectively. RM ANOVAs: (**F**) fear last 2 min, Dys+/- control virus: $F_{1,4} = 0.050$, $p = 0.83$; Dys+/- OXTR virus: $F_{1,8} = 13.63$, $p = 0.006$; (**G**) relief first 2 min, Dys+/- control virus: $F_{1,4} = 1.44$, $p = 0.30$; Dys+/- OXTR virus: $F_{1,7} = 6.11$, $p = 0.043$. n = 5/9 observers per group. ***$p < 0.0005$ and *$p < 0.05$ versus the exploration of the neutral demonstrator.
Error bars represent standard error of the mean. See also Figure S6.

Discussion

Combining behavioral, anatomical and genetic tools, this study reveals that the CeA, and in particular the OXTergic projections from the PVN to the CeA, is an essential neuronal substrate for the ability to socially respond to emotional states evoked in unfamiliar conspecifics. These mechanisms were revealed by a novel behavioral paradigm, which extends the opportunity to investigate previously unexplored aspects of social cognitive processes in mice.

Ethological Implications of Mice Emotion Discrimination

Using a two-choice setting, we reliably measured the ability of mice to distinguish conspecifics depending on their emotional state. This evidence supports and extends previous indications that rodents can perceive and react to altered emotional states in conspecifics [16, 19, 42, 53–55, 74–79]. Our data also extend previous rodent emotion-based tests, which mostly detect behavioral responses between familiar conspecifics, which may vary depending on the sex of the tested subjects [19, 23, 54, 80]. Indeed, here, we found that emotion discrimination ability is exerted toward unfamiliar conspecifics and it is similarly evident in male and female mice. Intriguingly, although observer mice showed similar social responses toward fear or relief demonstrators, the ethological meaning underlying these behaviors might differ. The effects of the fear manipulation are in line with studies indicating that rodents are sensitive to the distress of others and can assimilate pain or fear responses expressed by conspecifics undergoing or recently exposed to pain or footshock challenge [16, 19, 23, 53–55, 75–80]. However, rodents have been also reported to actively escape from aversive stimuli [23, 81], USVs calls induced by heavy distress [16, 82], and odors emitted by shocked, stressed, defeated, or sick conspecifics [63–65].

The approach behavior observed toward fear demonstrators might then be potentially related to affective responses, such as helping or consolatory behaviors previously described in rats [74, 79] or prairie voles [19]. Indeed, in our setting, the lack of pain or the minimal levels of distress induced by the recall of a fear memory might not ‘‘alarm’’ the observer but promote a ‘‘pro-social’’ approach. In agreement, exposing the observer to a recently shocked demonstrator induced avoidance. Consistent with the fast and transient discriminatory behavior we observed (up to 2 min), we did not detect any behavioral or physiological transfer of responses from the demonstrators (e.g., freezing, escape behaviors, or altered corticosterone levels) to observer mice. This suggests that emotional contagion was not the motivation for the observers' social approach. However, based on previous reports [53, 55, 80], we cannot exclude that some sort of emotion contagion might take place at a later time point. The attraction exerted by the relief condition is instead consistently supported by several complementary findings: the approach toward the relief mouse, its odor, the induced place preference conditioning, and the reduced corticosterone levels, all point to an attractive signal of this positively valenced state. This is in agreement with a number of different rewarding stimuli in mice, such as pleasant odors, sexual signals, and the intrinsic reward valence of social interaction [57, 83–87]. However, the approach behavior we observed was clearly not correlated to the rewarding properties of social interaction per se. This is supported by the distinction between emotion discrimination and one-on-one social interaction (Figures 1 and 2 versus 3I), the lack of correlation between emotion discrimination and absolute quantity of social interaction (Figure S2H), the selective effects of OXT pathways manipulation in emotion discrimination, but not in social interaction (Figures 4 and 5 versus S4), and the effects of dysbindin-1 genetic variants in emotion discrimination, but not in the 3-chamber task (Figures 6C and 6D versus S6A and S6B).

Altogether, our findings suggest that the two-choice emotion discrimination setting measures aspects of rodent social cognition that have been scarcely explored and that can complement currently available tools. Moreover, the ensemble of the presented results indicates the potential of this experimental setting to investigate a range of behavioral responses induced by emotional stimuli of different intensity and with potentially different ethological significance.

PVN-CeA OXT Projections as Key Modulator of Emotion Discrimination

We described the role of the endogenous OXT system, and specifically of OXT projections to CeA, NAcc, mPFC and hippocampal CA2, in the social response to emotional states evoked in unfamiliar conspecifics. Functional mapping of the selected PVN projections identified the CeA as a necessary site for OXT to allow discrimination of both negative and positive emotional states. These findings are in line with strong evidence across species, which implicate the amygdala as a critical hub in the processing of both positive and negative states [88–92]. In particular, given its access to primary sensory information, the CeA could orchestrate appropriate behavioral responses to salient stimuli with different valences [88, 93]. Consistently, a role for the CeA has been shown in modulating responses not just to threats or other aversive states [94–96] but also to reward-predictive cues and safety signals [97–101].

In humans, alterations in amygdala responses to positive and negative emotions have been reported in neuropsychiatric conditions, such as autism and schizophrenia [72, 73], in association with OXTR genetic variants [41, 102] or following intranasal OXT [33, 36, 38, 103]. However, the role of amygdala modulation, specifically in regard to the functionality of the endogenous OXT system, had not been elucidated. OXT has been shown to modulate emotion processing in rodents through an action in the anterior cingulate cortex, insula and lateral septum [16, 23, 24]. Our manipulations of endogenous PVN OXTergic projections add to this previous evidence, delineating OXT projections to the CeA as a neurobiological substrate for the ability to discriminate expression of emotions in others.

An intriguing question raised by our data is how the processing of both positive and negative states is achieved by OXT modulation in the CeA. The implication of CeA OXT signaling in processing threat responses [22, 94, 104] and social fear [105] would suggest a specific role in the detection of fear-mediated responses. The effect we found in the discrimination of both negative and positive states, however, supports the evidence of a generalized role of OXT in modulating CeA function in the response to socially communicated salient information, similar to evidence in humans [70, 90, 92, 106, 107]. Specific cell types and states of CeA neuronal subpopulations could be substrates of the perception and response to stimuli with different emotional valences. Distinct neuronal populations in sub-regions of the amygdala have indeed been shown to control specific behavioral responses to fear [104, 108], although evidence of OXT modulation of positively valenced cues are still poor.

Mechanistically, OXT has been documented to increase neuronal firing rates, mainly through inhibition of interneuron activity, which results in an increase in signal-noise ratio and subsequently enhanced information transfer [13, 22, 109–111]. The close proximity of highly enriched PVN-OXT fibers to GABAergic CeA neurons expressing OXTR [22] suggests that a similar modulation might occur during emotion discrimination in the CeA. In particular, an interaction between the OXT and the corticotropin-releasing factor (CRF) system could be one of the possible substrates, considering the high levels of CRF expression in CeA [112, 113] and the reported CRF-dependent PVN plasticity in transfer of fear [80]. Projection-specific manipulations dissecting CeA connectivity, as previously reported [114, 115], could help to elucidate how CeA neurons process stimuli of positive or negative valence to direct behavior.

Genetic Modulation of Emotion Discrimination through CeA OXTR Levels

The OXT system is regarded as a promising target for the treatment of social cognitive dysfunctions [31]. Altered levels of OXT and OXTR in specific regions of the brain have been reported in a number of animal models of psychiatric and neurodevelopmental disorders [116–119]. However, how these alterations might be involved in clinical manifestations and in particular in social cognitive endophenotypes remains unclear. The reduction in CeA OXTR levels in dysbindin-1 hypofunctioning mice has causally related OXT receptor-mediated mechanisms with the ability to discriminate negative and positive emotions in conspecifics. Genetic variants in dysbindin-1 are strongly associated with human intelligence [51]. Moreover, both mouse and human studies consistently indicate that reduced dysbindin-1 expression modulates higher-order cognitive functions [48–50]. Thus, our new findings extend the implication of dysbindin-1 genetics in the social cognitive domain and strengthen the evidence that OXT signaling within the CeA and, specifically, OXTR levels within this brain region constitute a crucial target to modulate emotion discrimination abilities.

From a therapeutic perspective, alterations in the endogenous OXT system can influence the response to exogenous OXT administration [40, 67, 120]. Notably, common functional genetic variants in dysbindin-1 have been recently shown to predict, in both humans and mice, cognitive responses to psychotropic drug treatments [50]. Thus, it is intriguing to speculate that dysbindin-1 genetic variants might also modulate social cognitive responses to exogenous OXT-related treatments. This aspect, together with the modulatory impact of endogenous and/or exogenous

OXT on emotion discrimination, represents an important subject for future studies, which may explore OXT pathways in the context of other genetic variants.

Mouse Emotion Discrimination versus Human Emotion Recognition

The behavioral paradigm developed in this study was inspired by human emotion recognition tasks [3]. Emotion recognition tasks rely on the ability to discriminate basic expression of emotions in others and are the most extensively used paradigms to assess human social cognition [3, 43], with direct relevance for a number of pathologies, including autism and schizophrenia [5, 6]. Human emotion recognition paradigms include the presentation of positively and negatively valenced emotions [3]. Consistently, we adopted a two-choice discriminative setting, focused on behaviors initiated by the observer mouse and adopting manipulations that could induce both negative and positive emotional states in the demonstrators. Throughout the paper, we referred to ‘‘emotions’’ as subjective internal states of mice evoked by behavioral manipulations, aware that our definition is based on the assessment of behavioral output more than on the intangible and unmeasurable nature of emotions. Taking these important limitations in mind, our data provide some indication that, similar to human emotion recognition tasks [70, 90, 92, 121], mice can discriminate emotions in others in a way that is distinct from sociability. This ability shows strong test-retest reliability, and it is equally evident in male and female mice toward unfamiliar conspecifics. This ability appears to rely on a primary amygdala recruitment across all forms of emotion perception. Furthermore, we have demonstrated that emotion discrimination abilities in mice are dependent on the OXT system, in further agreement with human evidence where OXT has been associated with social cognitive functions, such as emotion recognition [8–10, 33–36, 38, 39]. Altogether, this indicates that our paradigm might approximate some features of human emotion recognition tasks [3, 43].

In conclusion, our data provide new insights into the role of endogenous OXT signaling in the ability to recognize emotions in unfamiliar conspecifics. Additionally, the evidence here reported also demonstrates an opportunity to reliably measure scarcely explored aspects of rodent social cognition. This could support more translational approaches between rodent and human social cognitive studies, with relevance to circuits, genetics, and neurochemical systems involved in different psychiatric disorders.

Materials and Methods

Mice

Males and females C57BL/6J mice, dysbindin-1 [49] heterozygous (Dys+/_) and their wild-type littermates (Dys+/+), and oxytocin receptor knockout (Oxtr_/_) [123], all 3-6 months-old, were used. Animals were housed two to four per cage in a climate-controlled (22 ± 2 C) and specific pathogen-free animal facility, with ad libitum access to food and water throughout, a standard environmental enrichment (material for nest and cardboard house), and with a 12-hour light/dark cycle (7pm/7am schedule). Experiments were run during the light phase (between 10am-5pm). All mice were handled on alternate days during the week preceding the first behavioral testing. Experimenters were blind to the mouse treatments and genotype during testing. Female mice were visually checked for estrus cycle immediately after the test and no correlation was found between estrus status and performance in the test. All procedures were approved by the Italian Ministry of Health (permits n. 230/2009-B and 107/2015-PR) and local Animal Use Committee and were conducted in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the European Community Council Directives.

Behavioral procedures

Emotion discrimination test

Habituation of the mice to the testing setting occurred on three consecutive days before the first experiment; each habituation session lasted 10 minutes. Test observer mice were habituated inside a Tecniplast cage (35.5x23.5x19 cm) with a separator and two cylindrical wire cups (10.5cm in height, bottom diameter 10.2cm, bars spaced 1 cm apart; Galaxy Cup, Spectrum Diversified Designs, Inc., Streetsboro, OH), around which they could freely move, as occurred during the test. A cup was placed on the top of the wire cups to prevent the observer mice from climbing and remaining on the top of them. The separator (11x14cm) between the two wire cups was wide enough to cover the reciprocal view of the demonstrators while leaving the observer mice free to move between the two sides of the cage. The wire cups, separators and experimental cages were replaced after each subject with clean copies to avoid scent carryover. Similarly, the rest of the apparatus was wiped down with water and dried with paper towels for each new subject. After each testing day, the wire cups, separators, and cubicles were wiped down with 70% ethanol and

allowed to air-dry. Testing cages were autoclaved as standardly performed in our animal facility. Demonstrator mice – matched by age and sex to the observers – were habituated inside the same Tecniplast cage (35.5x23.5x19 cm), under the wire cups for three consecutive times, ten minutes each. During both habituation and behavioral testing, the cages were placed inside soundproof cubicles (TSE Multi Conditioning Systems) homogeneously and dimly lit (6 ± 1 lux) to minimize gradients in light, temperature, sound and other environmental conditions that could produce a side preference. Digital cameras (imaging Source DMK 22AUC03 monochrome, Ugo Basile) were placed facing the long side of the cage and on top of the cage to record the three consecutive two-minute epochs from different angles, using the Anymaze program (Stoelting, Ireland). Behavioral scoring was performed a posteriori from videos by trained experimenters, blind to the manipulations of both the observers and demonstrators. Three independent persons scored the same data with an inter-rater reliability r score of 0.954. A sniffing event was considered when the observer touched with the nose the demonstrators' wire cup or when the observer's nose directly touched the demonstrator. The emotion discrimination experiments reported in this work were independently replicated by nine different researchers in three different laboratories.

Observers

Before the test, mice were habituated to the experimental setting as reported above. On the third day of habituation, mice were also habituated to the tone cue (4 kHz, 80 dB sound pressure level, three times for 30 s each with an intertrial interval of 90 s) without any conditioning. One hour prior to behavioral testing, mice were placed in the testing cage, in an experimental setting (i.e., separator and two wire cups), in a room adjacent to the testing room. Five minutes before the experiment, the testing cages containing the observer mice were gently moved into the testing cubicles. The 6-minute experiment began after placing one emotionally 'neutral' and one ''emotionally altered'' demonstrator under the wire cups. The order of insertion of the neutral or emotionally-altered demonstrator was randomly assigned.

Neutral demonstrators

In the days before the test, all neutral mice were habituated to the experimental setting as reported above. For the relief condition, neutral demonstrators underwent no manipulation the day before the test. For the fear condition, the day before the test, neutral demonstrators were habituated to the tone cue inside the cups as for the experimental setting and as done for the observer mice. On the testing day, neutral demonstrators were brought inside their home cages in the experimental

room one hour before the experiment began. Demonstrators were test-naive and used only once. In some cases, we re-used the same demonstrator for maximum two/ three times, with always at least one week between each consecutive test. No differences were observed in the performance of the observer mice depending on the demonstrators' previous experience.

Relief demonstrators

In the days before the test, mice were habituated to the experimental setting as reported above. Relief demonstrators were then water deprived 23 hours before the experiment. One hour before the test, ad libitum access to water was reestablished, and mice were brought inside the experimental room in their home cages. Food was available ad libitum all the time and some extra pellets were put inside the home cage during the 1-hour water reinsertion.

Fear demonstrators

In the days before the test, mice were habituated to the experimental setting as reported above. Fear demonstrators were fear conditioned (from one day to one week before the test) using the parameters and context previously described [124], and using the same tone delivered to the observers and neutral demonstrators during their habituation process. In particular, the conditioned stimulus was a tone (4 kHz, 80 dB sound pressure level, 30 s) and the unconditioned stimulus were three scrambled shocks (0.7 mA, 2 s duration, 90 s intershock interval) delivered through the grid floor that terminated simultaneously with the tone (2 s). The day of the test these mice were habituated, inside their home cages, in a room adjacent to the testing room for one hour prior to the test; they were consequently brought inside the experimental room one by one, before placing them under their designated wire cup. Fear mice were conditioned only once, in a separate room and using distinct apparatus (Ugo Basile SRL, Italy) from the one where the emotion recognition task would be performed. Fear demonstrators were generally used only once. In the case of a second exposure to the test, these demonstrators were just re-exposed to the same conditioned tone, at least one week apart from the previous exposure and maximum 1 month from the initial conditioning.

Shock demonstrators

This manipulation was performed for direct comparison with a rat protocol and was performed as previously described [16]. In particular, these demonstrator mice were exposed to two footshocks (1 mA, 5 s duration, 60 s intershock interval) immediately before the 6-minute test session. All other procedures were identical to the other demonstrators as described above.

''Classic'' social interaction test and 3-chamber social interaction test

Social interaction in freely interacting mice was performed as previously reported [67, 125]. Briefly, mice were individually placed in the testing cage 1 hour prior to testing. No previous single housing manipulation was adopted to avoid any home-cage territorial aggressive behaviors. Testing began when a stimulus mouse, matched for sex and age, was introduced into the testing cage for a 4-min period interaction.

Sociability and preference for social novelty in the 3-chamber task was tested as previously described [125]. In brief, the test consists of four phases of 10 minutes each. In phase 1 the subject mouse is placed into the center chamber with both doors closed. Then, in phase 2, both doors are open and the subject can freely explore all three empty compartments for another 10 minutes. Next, phase 3 consists of the ''Sociability'' test in which an empty wire cage is placed in one of the chambers and another wire cage with a stimulus mouse inside is placed in the other chamber. Finally, phase 4 consists of the ''Preference for Social Novelty'' test in which the empty wire cage is replaced with a novel mouse.

One-on-one social exploration tests

This test was similarly performed as previously described [16]. One hour prior to behavioral testing, each experimental subject was placed into a Tecniplast cage (35.5x23.5x19 cm) with shaved wood bedding and a wire lid, in a room adjacent to the testing room. Five minutes before the experiment, the testing cages containing the observer mice were gently moved into the testing sound proof cubicles. To begin the test, a demonstrator mouse was introduced to the cage for 6 minutes (as for the emotion discrimination task), and exploratory behaviors initiated by the test subject were timed by two independent experimenters blind to the manipulations. Demonstrator mice were used only once. Each observer was given tests on consecutive days: once with an unfamiliar naive conspecific, once with an unfamiliar fear conspecific (fear conditioning exactly as above), and once with an unfamiliar relief conspecific (manipulated exactly as above). Test order was counterbalanced.

Sensory modality assessment

In the ''complete darkness'' experiments, mice were tested as above, but eliminating all sources of light within the testing cage as well as in the entire testing room. Videos were recorded for

successive scoring either with an infrared thermal camera (FLIR A315, FLIR Systems) or with Imaging Source DMK 22AUC03 monochrome camera (Ugo Basile). The two camera settings gave the same experimental results.

For acoustic stimuli experiments, ultrasonic vocalisations (USVs) were recorded during the test phases performed as above in two different experimental settings: 1) exactly as reported above with one observer mouse and two demonstrators under the wire cups, and 2) with only one demonstrator present in the apparatus (and under the wire cup) for each emotional condition. This was done to make sure that the USVs recorded could be attributed to a single emotional state and/or to a communication between demonstrators and observer. Ultrasonic vocalizations recording and analysis were performed as previously described [67]. The ultrasonic microphone (Avisoft UltraSoundGate condenser microphone capsule CM16, Avisoft Bioacoustics, Berlin, Germany), sensitive to frequencies between 10 and 180 kHz, was mounted 20 cm above the cage to record for subsequent scoring of USV parameters. Vocalisations were recorded using AVISOFT RECORDER software version 3.2. Settings included sampling rate at 250 kHz; format 16 bit. For acoustical analysis, recordings were transferred to Avisoft SASLab Pro (Version 4.40) and a fast Fourier transformation (FFT) was conducted. Spectrograms were generated with an FFT-length of 1024 points and a time window overlap of 75% (100% Frame, Hamming window). The spectrogram was produced at a frequency resolution of 488 Hz and a time resolution of 1 ms. A lower cut-off frequency of 15 kHz was used to reduce background noise outside the relevant frequency band to 0 dB. Call detection was provided by an automatic threshold-based algorithm and a hold-time mechanism (hold time: 0.01 s). An experienced user checked the accuracy of call detection and obtained a 100% concordance between automated and observational detection. Parameters analyzed included number of calls, duration of calls and quantitative analyses of sound frequencies measured in terms of frequency and amplitude at the maximum of the spectrum.

For odor stimuli experiments, observers were tested as described above, but replacing the ‘‘demonstrator’’ with cotton balls, which had been swiped throughout the body, head and anogenital areas of demonstrators. Odors were separately collected from neutral, relief (after the 1 hour ad libitum access to water) and fear (immediately after the delivery of the conditioned tone cue) demonstrators. Each odor was always taken fresh from one single mouse (which was not reused) and used only once.

Place conditioning

Mice were tested in a well-established place conditioning paradigm, able to assess either positive or negative affective states in mice [124, 126]. The place conditioning paradigm was performed in a rectangular Plexiglas box (length, 42 cm; width, 21 cm; height, 21 cm) divided by a central partition into two chambers of equal size (21 3 21 3 21 cm). One compartment had black walls and a smooth Plexiglas floor, whereas the other one had vertical black and white striped (2 cm) walls and a slightly rough floor. During the test sessions, an aperture (4 3 4 cm) in the central partition allowed the mice to enter both sides of the apparatus, whereas during the conditioning sessions the individual compartments were closed off from each other. To measure time spent in each compartment a video tracking system (Anymaze) was used. The place conditioning experiment lasted 5 days and consisted of three phases: pre-conditioning test, conditioning phase and post-conditioning test. On day 1, each mouse was allowed to freely explore the entire apparatus for 20 min, and time spent in each of the two compartments was measured (pre-conditioning test). Conditioning sessions took place on days 2 and 4. Mice were divided in two groups: neutral and relief. Mice of the same home- cage were assigned to the same group. Mice were then divided in the two experimental groups with similar preconditioning time values in the two sides of place conditioning apparatus. As for the same manipulation in the emotion discrimination test, the relief group was assigned to receive a 23-hour water deprivation period before the two conditioning sessions on the day 2 and 4, when they were confined with their cage mates in one of the two compartments for 1 hour with free access to water and food (conditioning). Food in the home cage was available all the time. Other than the two 23-hr deprivation periods, water was available all the time. The neutral group was exposed to the same procedure but without any water deprivation. The post-conditioning test was performed on day 5 in the same conditions as the preconditioning test. For each mouse, a conditioning score was calculated as the post-conditioning time minus the pre-conditioning time (in seconds) spent in the conditioned compartment of the apparatus.

Corticosterone assay

Corticosterone concentration was analyzed from mice plasma. Immediately after the behavioral test, each mouse was sacrificed by decapitation. The blood was quickly collected in EDTA(0,5M)-coated tubes and centrifuged at 3000 rpm for 10 min; the supernatant obtained was stored at _20_C until the assay. The corticosterone concentration was detected by a commercially available Detect

X corticosterone enzyme-linked immunoassay (ELISA) kit (Arbor Assays, MI, USA; Cat N K014-H1) following the manufacturer's protocol. The level of corticosterone was expressed as fold changes compared to the control group average.

Viral vectors

The OXTp-Venus, OXTp-hM4D(Gi), OXTp-DIO-hM4D(Gi)-mCherry and EF1a-OXTR-IRES:EYFP AAV serotype 1/2 were cloned and produced as previously reported [22, 24, 69]. rAAV genomic titers were determined with QuickTiter AAV Quantitation Kit (Cell Biolabs, Inc., San Diego, California, USA) and RT-PCR using the ABI 7700 cycler (Applied Biosystems, California, USA). rAAVs titers were _1010 genomic copies per per ml. For the EF1a-OXTR-IRES:EYFP, AAV expression and spreading were assessed by injecting 0,5 ml in the CeA of OXTR deficient mice [123]. CAV2 equipped with Cre recombinase (titer: 2.5x1011 pp) was purchased from the Institute of Molecular Genetics in Montpellier CNRS, France [24, 122].

Stereotaxic Injections

Mice were anesthetized with 2% isoflurane in O2 by inhalation and mounted onto a stereotaxic frame (Kopf) linked to a digital reader. Mice were maintained on 1.5 - 2% isoflurane during the surgery. Brain coordinates of injections were chosen in accordance to the mouse brain atlas [127]: PVN (AP: _0.9 mm; L: ± 0.2 mm; DV: _4.5), CeA (AP: _1 mm; L: ± 2.2 mm; DV: _4.5 mm), NAcc (AP: +1.7 mm; L: ± 0.5 mm; DV: _4 mm), mPFC (AP: +1.9 mm; L: ± 0.25 mm; DV: _2.5 mm). Mice that had been injected with AAVs and/or CAV2 were allowed 1 month to recover and for the viral transgenes to adequately express before undergoing behavioral experiments. The injected volume viruses (rAAV and CAV2) were 75-100 nL volume, depending on the brain region. CAV2 was prediluted at the 1x109 ppl/ml concentration. Only mice with appropriate placements were included in the reported data (Figures S5H–S5L and S6F).

Drugs

At least 4 weeks after cerebral injections, we inhibited PVN OXT release by i.p. administration of Clozapine N-Oxide (CNO, #4936 Tocris Bioscience) dissolved in physiological saline (0.9% NaCl) at a dose of 3mg/kg in a volume of 10 ml/kg, 30 minutes before the emotion discrimination task. For control experiments, the same mice were injected with the same volume of saline.

Histology

At the end of the behavioral procedures mice were deeply anesthetized (urethane 20%) and transcardially perfused with 4% paraformaldehyde in PBS, pH 7.4. Brains were dissected, post fixed overnight and cryoprotected in 30% sucrose in PBS. 40-mm-thick coronal sections were cut using a Leica CM1900 microtome. For immunohistochemical studies free-floating sections of selected areas were washed in PBS three times for 10 minutes, permeabilized in PBS plus 0.4% Triton X-100 for 30 min, blocked by incubation in PBS plus4%normal goat serum (NGS), 0.2% Triton X-100 for 1 h (all at room temperature) and subsequently incubated with a GFP polyclonal antibody (1:1000, rabbit, Invitrogen, CatNo. A-11122), a dsRED polyclonal antibody (1:1000, rabbit, Clontech, CatNo. 632496), or an OXT monoclonal antibody (1:1000, mouse, PS38, kindly provided by Dr. Harold Gainer). Primary antisera were diluted in PBS plus 2%NGS overnight at 4°C for GFP antibody and overnight at room temperature for dsRED and OXT antibodies. Incubated slices were washed three times in PBS plus 1% NGS for 10 minutes at room temperature, incubated for 2 h at room temperature with a 1:1000 dilution of an Alexa Flour 488 goat anti-rabbit IgG (H+L) (1:1000, Thermo Fisher Scientific, CatNo.A11034) and Alexa Fluor 633 goat anti-rabbit IgG (H+L) (1:1000, Thermo Fisher Scientific, CatNo.A21071) in 1% NGS in PBS and subsequently washed three times in PBS for 10 min at room temperature. The sections were mounted on slides and coverslipped.

Imaging

All images were acquired on a Nikon 1 confocal laser scanning microscope. Digitalized images were analyzed using Fiji (NIMH, Bethesda MD, USA) and Adobe Photoshop CS5 (Adobe, Montain View, CA).

Brain Autoradiography

A separate cohort of naive mice was handled as described above and their brains were rapidly explanted, snap-frozen in isopentane at -25°C and moved at -80°C for storage. 14 mm-thick coronal sections were then cut with a cryostat and mounted on chromealum- gelatin-coated microscope slides. All slides were stored at -80°C until receptor autoradiography. The binding procedure and quantification of the resulting autoradiographic images were performed as previously described [67]. Briefly, sections were fixed with 0.2% paraformaldehyde and rinsed

twice with 0.1% bovine serum albumin in 50 mM Tris–HCl buffer (pH 7.4). OXT binding sites were detected by incubation (2 hours at room temperature in a humid chamber) with the radioiodinated OXTR antagonist ornithine vasotocin analog ([^{125}I]-OVTA, Perkin Elmer, MA, USA) at 0.02 nM in a medium containing 50mMTris–HCl, 0.025% bacitracin, 5mM MgCl2, 0.1% bovine serum albumin. Sections immediately adjacent to the ones used for [^{125}I]-OVTA binding were used to determine non-specific binding by addition of 2 mM OXT to the incubation solution. The unbound excess of ligand was washed out and slides were quickly dried under a stream of cool air and exposed to Biomax MR Films (Carestream) in an autoradiographic cassette for 72 hours. The final autoradiograms were digitalized by grayscale high-resolution scanning (600x600 dpi) and analysis was carried out using the ImageJ 1.47v software (NIH, USA). Target regions were identified by comparison with a reference mouse brain atlas [127]. Specific densitometric gray intensity was calculated by subtraction of the gray level of the respective non-specific binding section. Autoradiographic ^{125}I microscales (Amersham International, UK) also were exposed for 72 hours and a reference standard curve was generated. Levels of gray intensity were then converted to nCi/mg tissue equivalent by interpolation with the standard curve.

Ex vivo electrophysiology

Ex vivo patch clamp electrophysiology recordings were performed on PVN virus-injected slices. Mice were anesthetized with isoflurane and transcardially perfused with an ice-cold cutting solution containing: 200 mM sucrose, 4 mM MgCl2, 2.5 mM KCl, 1.25 mM NaH2PO4, 0.5 mM CaCl2, 25 mM NaHCO3 and 10 mM D-glucose (~300 mOsm, pH 7.4, oxygenated with 95% O2 and 5% CO2). Brains were removed and immersed in the cutting solution. Coronal slices (270 mm thick, VT1000S Leica Microsystem vibratome) were incubated for 2 min in a mannitol solution (225 mM mannitol, 2.5 mM KCl, 1.25 mM NaH2PO4, 8 mM MgSO4, 0.8 mM CaCl2, 25 mM NaHCO3 and 10 mM d-glucose (~300 mOsm, pH 7.4, oxygenated with 95% O2 and 5% CO2)) and then allowed to recover for 1 hour at 35_C in a solution containing: 117 mM NaCl, 2.5 mM KCl, 1.25 mM NaH2PO4, 3 mM MgCl2, 0.5 mM CaCl2, 25 mM NaHCO3 and 10 mM glucose (~310 mOsm, pH 7.4, oxygenated with 95% O2 and 5% CO2). Recordings were performed in magnocellular and parvocellular neurons of the PVN at room temperature in artificial cerebrospinal fluid (ACSF) with the following composition: 117 mM NaCl, 2.5 mM KCl, 1.25mMNaH2PO4, 1mMMgCl2, 2mMCaCl2, 25mMNaHCO3 and 10mMglucose (~310 mOsm,

pH 7.4, oxygenated with 95% O2 and 5% CO2). Patch pipettes were made from thick-wall borosilicate glass capillaries (B150-86-7.5, Sutter Instrument). Pipettes (5-7 mU) were filled with an intracellular solution containing: 130 mM K-gluconate, 10 mMHEPES, 7mM KCl, 0.6 mM EGTA, 4 mM Mg2ATP, 0.3 Mm Na3GTP, 10 mM Phosphocreatine. The pH was adjusted to 7.3 with HCl. Whole-cell recordings were performed on PVN neurons identified on a fluorescent-based approach. Once stable recording conditions were obtained (series resistances in the range of 10–25 mU), PVN neurons were identified electrophysiologically as magnocellular (presence of transient outward rectification) or parvocellular (lack of transient outward rectification) according to an established current-clamp protocol in literature [128]. Validation of iDREADDs was performed evoking spike firing in PVN neurons by injection of a small depolarizing current pulse (20 pA for 1 s) under current-clamp mode. Activation of iDREADDs was obtained using 10 mM Clozapine N-Oxide (CNO, #4936 Tocris Bioscience) applied in the bath for 15 min. Data, filtered at 0.1 Hz and 5 kHz and sampled at 10 kHz, were acquired with a patch-clamp amplifier (Multiclamp 700B, Molecular Devices) connected to a digital converter (Digidata 1440A, Molecular Devices) and analyzed using pClamp 10.2 software (Molecular Devices). All chemicals were purchased from Sigma, otherwise specified.

QUANTIFICATION AND STATISTICAL ANALYSIS

No statistical methods were used to predetermine sample sizes, although sample sizes were consistent with those of previous studies [19, 23, 80]. No explicit randomization method was used to allocate animals to different experimental groups, but samples always derived from mice of the same litter. Mice were tested and data were processed by investigators blinded to the treatment and genotype of the animals. All datasets were tested to fit a normal distribution using the Kolmogorov-Smirnov and Pearson's chi-square tests. All statistical parameters including the statistical tests used, exact value of n, what n represents, precision measures (mean ± SEM) and statistical significance are reported in the Figure Legends. Results are expressed as mean ± standard error of the mean (SEM) throughout. Each observer's behavior toward the two different demonstrator mice was analyzed using a within-groups Repeated-measures ANOVA (RM-ANOVA). The behavior of the two demonstrators and USVs recordings were analyzed by Two-Way ANOVAs with emotional state as between-subjects factors, and the within-session 2-min consecutive intervals as a repeated-measure within-subject factor. The behaviors of the observer

mice in the one-on-one setting were analyzed by Two- Way ANOVAs with the emotional state of the demonstrator as between-subjects factors, and the within-session 2-min consecutive intervals as a repeated-measure within-subject factor. Two or One-Way ANOVAs were used for autoradiography and social interactions when different genotypes and treatments were involved. Newman–Keul's post hoc test with multiple comparisons corrections was used for making comparisons within groups when the overall ANOVA showed statistically significant differences. Two-tailed paired t test or Two tailed Wilcoxon matched pairs signed rank test were used for electrophysiological experiment. The accepted value for significance was $p < 0.05$. Statistical analyses were performed using Statistica 13.2 (StatSoft).

Supplementary Figures

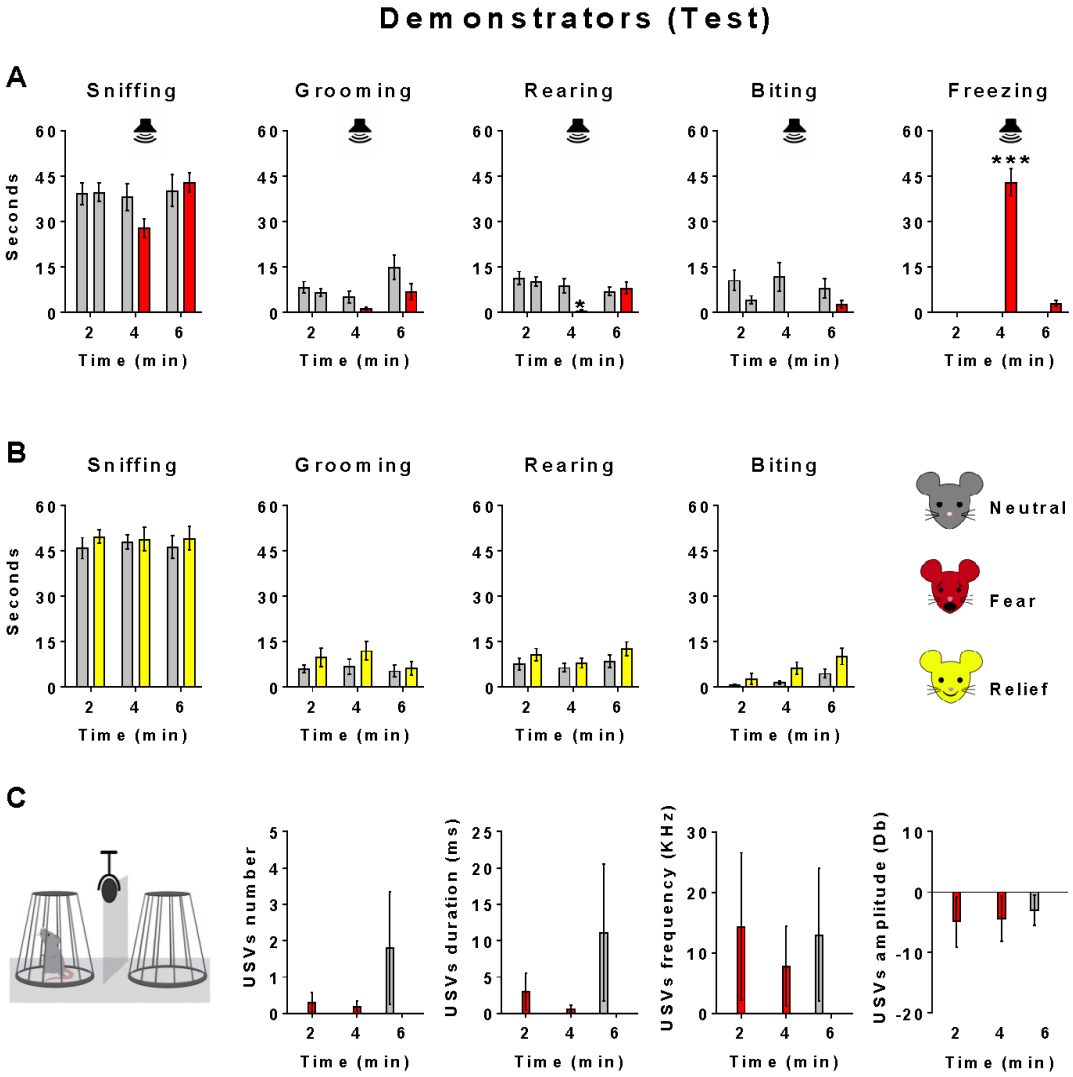


Figure S1. Demonstrators behavior during emotion discrimination test, Related to Figure 1, 2 and 3. (**A**) Behaviors displayed by neutral and fear demonstrator mice during the fear emotion discrimination paradigm. No demonstrator defecated or urinated during the whole test session. Emotion-by-time statistical interaction for sniffing ($F_{2,36}$=2.72, p=0.08), grooming ($F_{2,36}$=1.07, p=0.35), rearing ($F_{2,36}$=5.09, p=0.01), biting ($F_{2,36}$=1.28, p=0.29), and freezing ($F_{2,36}$=48.82, p<0.0001). *p<0.05, and ***p<0.0001 versus all other points. n=10 demonstrators per group. (**B**) Behaviors displayed by neutral and relief demonstrator mice during the relief emotion discrimination paradigm. No demonstrator defecated or urinated during the whole test session. No significant emotion-by-time statistical interaction was evident for sniffing ($F_{2,36}$=0.09, p=0.92), grooming ($F_{2,36}$=0.34, p=0.71), rearing ($F_{2,36}$=0.31, p=0.73), and biting ($F_{2,36}$=0.84, p=0.44). n=10 demonstrators per group. (**C**) Schematic drawing of the test setting to record USVs. Mean number and duration, peak frequency and amplitude at the maximum of the spectrum of USVs emitted by single demonstrator mice in neutral (grey bar), fear (red bar), or relief (yellow bar) emotional state are reported. Two-Way ANOVAs showed no significant differences. n=6 demonstrators per group.
Error bars represent standard error of the mean.

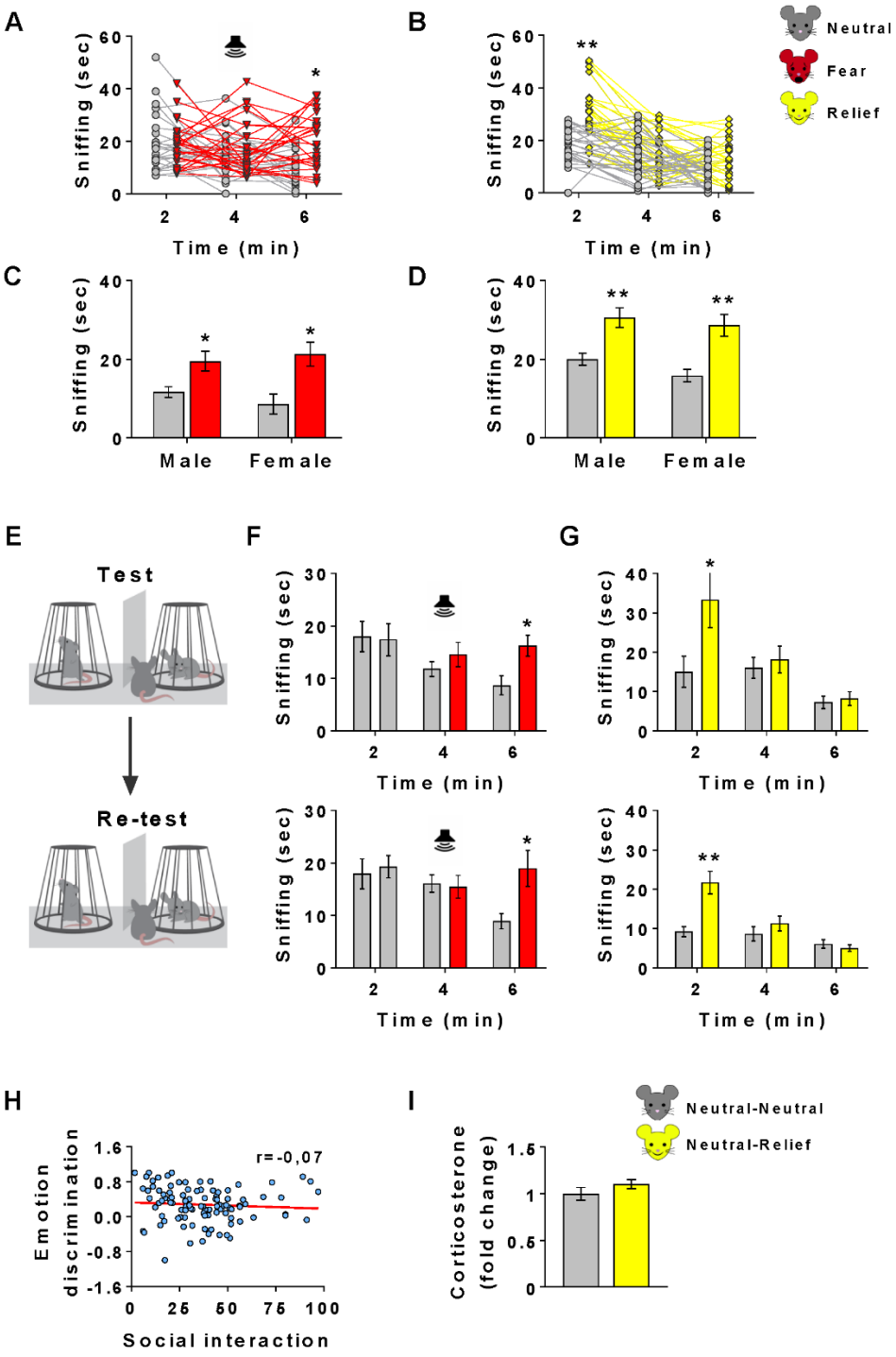


Figure S2. Characterization of the mouse emotion discrimination test, Related to Figure 1 and 2. (**A-B**) Time (in seconds) spent by each male and female observer mice sniffing the two demonstrators in the (**A**) fear and (**B**) relief paradigms. Emotion-by-time interaction for fear ($F_{2,108}$=5.00, p=0.008), or relief ($F_{2,116}$=6.71, p=0.002) paradigm. n=28/30 observers per group. *p<0.05 and **p<0.005 versus all other points. (**C-D**) Direct comparison between male and female emotion discrimination abilities in the (**C**) fear (last 2-min RM ANOVA Male: $F_{1,15}$=6.51, p=0.022; Female: $F_{1,11}$=10.98, p=0.006), or (**D**) relief (first 2-min RM ANOVA Male: $F_{1,14}$=15.07, p=0.001; Female: $F_{1,14}$=14.60, p=0.001) paradigm. *p<0.05 and **p<0.005 versus the neutral demonstrator. n=12/16 observers per group. (**E**) Schematic drawing of the test-retest reliability validation. (**F-G**) Time (in seconds) spent by observer mice

sniffing the two demonstrators in the (**F**) fear or (**G**) relief paradigms during their first and second exposure (one week later) to the emotion discrimination test. RM ANOVA for the "fear" manipulation, last 2-minute session, Test: $F_{1,13}$=6.10, p=0.028; Re-Test: $F_{1,13}$=8.59, p=0.012. RM ANOVA for the "relief" manipulation, first 2-minute session, Test: $F_{1,10}$=5.15, p=0.046; Re-Test: $F_{1,10}$=22.88, p=0.0008. *p<0.05, and **p<0.005 versus the exploration of the neutral demonstrator. n=11/14 observers per group. Time spent sniffing neutral, fear and relief demonstrators is depicted in grey, red and yellow, respectively. (**H**) Correlation analyses between the emotion discrimination index (in y axis) and social interaction (in x axis). No significant correlation was found (r=-0.07). n=101 observers. (**I**) Blood corticosterone levels displayed by observer mice immediately after being tested in the emotion discrimination test with two neutral demonstrators (grey bar), or one neutral and one relief demonstrators (yellow bar). Data are expressed as fold changes compared to observers exposed to two neutral demonstrators (T-test: df: 9; p=0.18). n=5/6 observers per group. Error bars represent standard error of the mean.

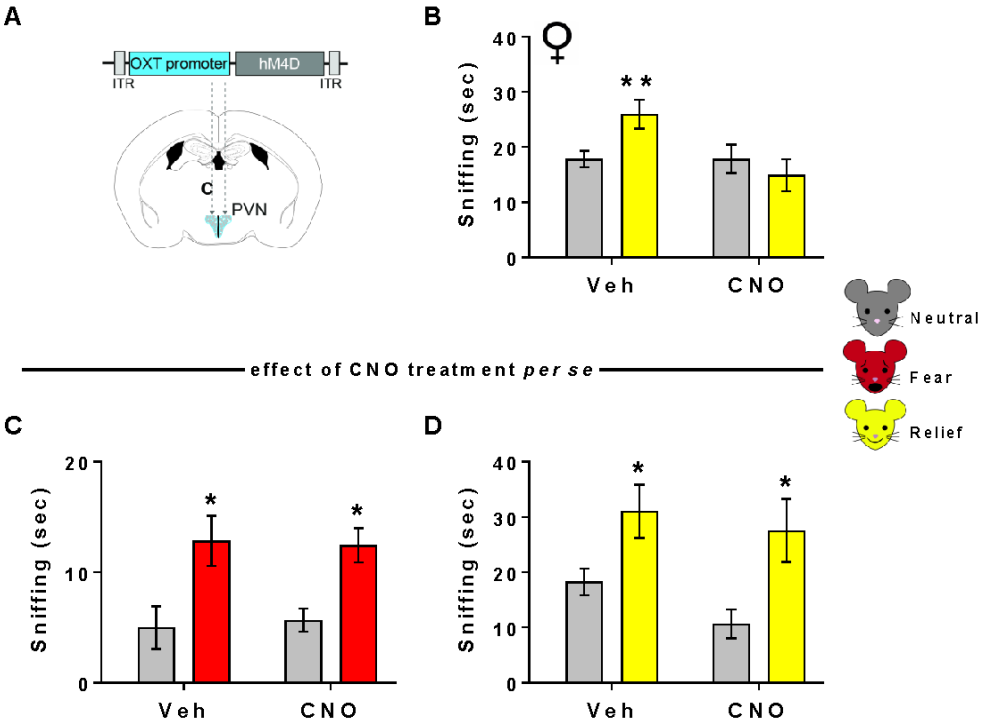


Figure S3. Control experiments for DREADDs manipulations, Related to Figure 4 and 5. (**A**) Scheme of the viral vector used to infect PVN OXT neurons. (**B**) Time (in seconds) spent sniffing the two demonstrators during the first 2 minutes of the emotion discrimination test, displayed by the same virus-injected observer female mice treated with vehicle or CNO (3mg/kg in a volume of 10 ml/kg, i.p., 30 minutes before the test), and shown separately for each demonstrator's emotion. Time spent sniffing neutral demonstrators is represented by grey bars. Time spent sniffing demonstrators with water-induced relief states (first 2-min RM ANOVA Veh: $F_{1,9}$=15.61, p=0.0033; CNO: $F_{1,9}$=1.04, p=0.33) is represented by yellow bars. **p<0.005 versus the neutral demonstrator within the same observer treatment. n=10 observers per group. (**C**-**D**) Assessment of CNO treatment *per se* in the emotion discrimination test. Time (in seconds) spent sniffing two demonstrator mice during (**C**) the last 2 minutes of the fear paradigm (RM ANOVA Veh: $F_{1,5}$=8.55, p=0.032; CNO: $F_{1,5}$=8.60, p=0.042) and (**D**) the first 2 minutes of the relief paradigm (RM ANOVA Veh: $F_{1,5}$=9.55, p=0.027; CNO: $F_{1,5}$=10.56, p=0.022) displayed by the same observer mice treated with vehicle or CNO (i.p., 3mg/kg in a volume of 10 ml/kg, 30 minutes before the test). Time spent sniffing neutral demonstrators is represented by grey bars. Time spent sniffing demonstrators with tone-induced fearful states or water-induced relief states is represented by red or yellow bars, respectively. *p<0.05 versus the neutral demonstrator within the same observer treatment. n=6 observers per group. Error bars represent standard error of the mean.

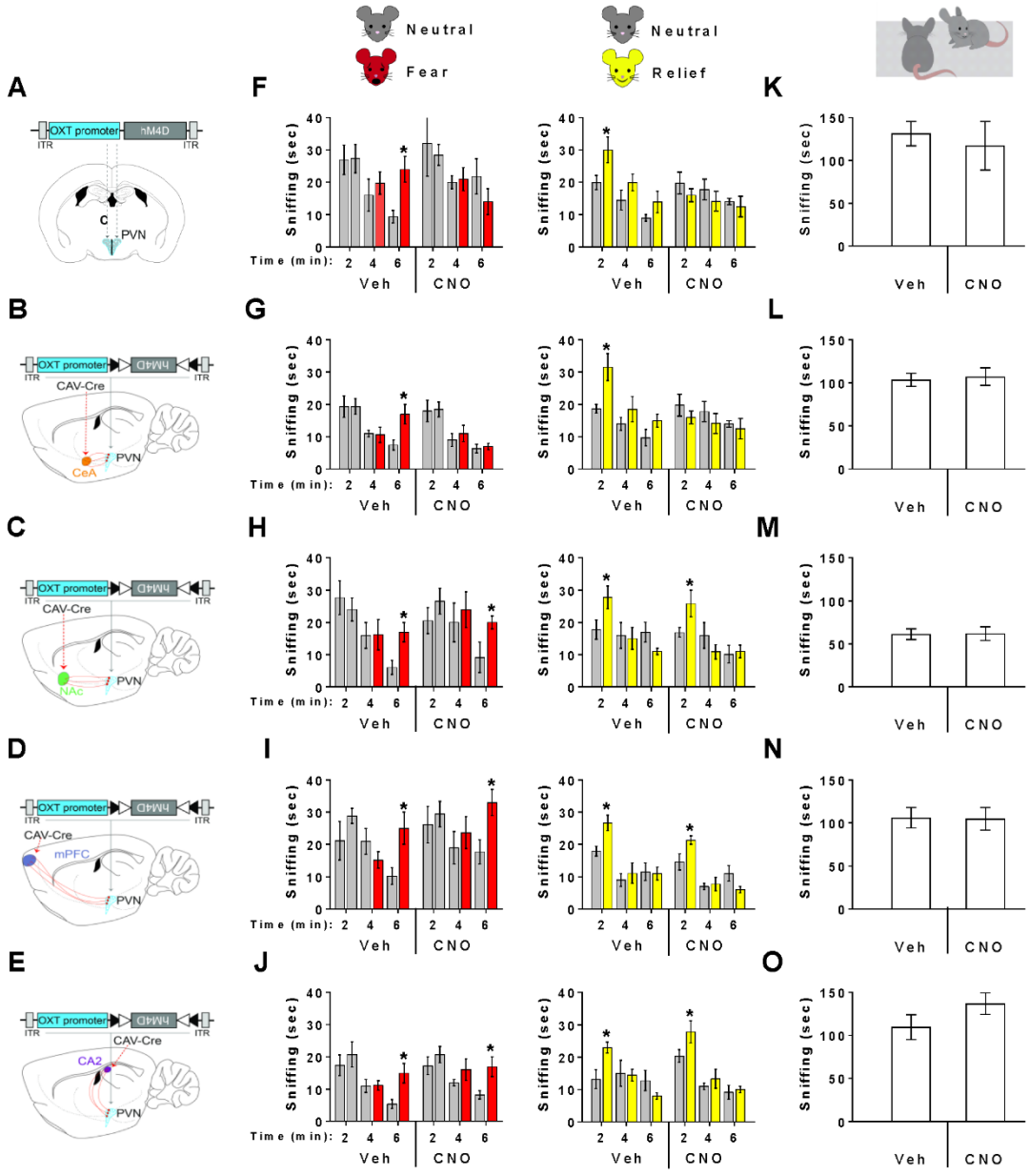


Figure S4. Effect of selective inhibition of PVN OXT projections in the emotion discrimination test and one-on-one social interaction, Related to Figure 4 and 5. (**A-E**) Scheme of the viral vector used to infect the PVN OXT neurons and respective projecting area (CeA, NAcc, mPFC or CA2). (**F-J**) Time (in seconds) spent sniffing the two demonstrators during the emotion discrimination test displayed by the same observer mice treated with vehicle or CNO, and shown separately for each demonstrator's emotion. Time spent sniffing neutral demonstrators is represented by grey bars. Time spent sniffing demonstrators with tone-induced fearful states, or water-induced relief states is represented by red or yellow bars, respectively. *$p<0.05$ versus the exploration of the neutral demonstrator. n=7/11 observers per group. (**K-O**) Time (in seconds) spent sniffing an unfamiliar age- and sex-matched conspecific during 4 minutes of free social interaction displayed by the same mice treated with vehicle or CNO. PVN (one-way ANOVA $F_{1,4}=0.19$, $p=0.68$); PVN-CeA (one-way ANOVA $F_{1,18}=0.081$, $p=0.78$); PVN-NAc (one-way ANOVA $F_{1,16}=0.006$, $p=0.94$); PVN-mPFC (one-way ANOVA $F_{1,10}=0.005$, $p=0.94$). n=4/10 per group. Error bars represent standard error of the mean.

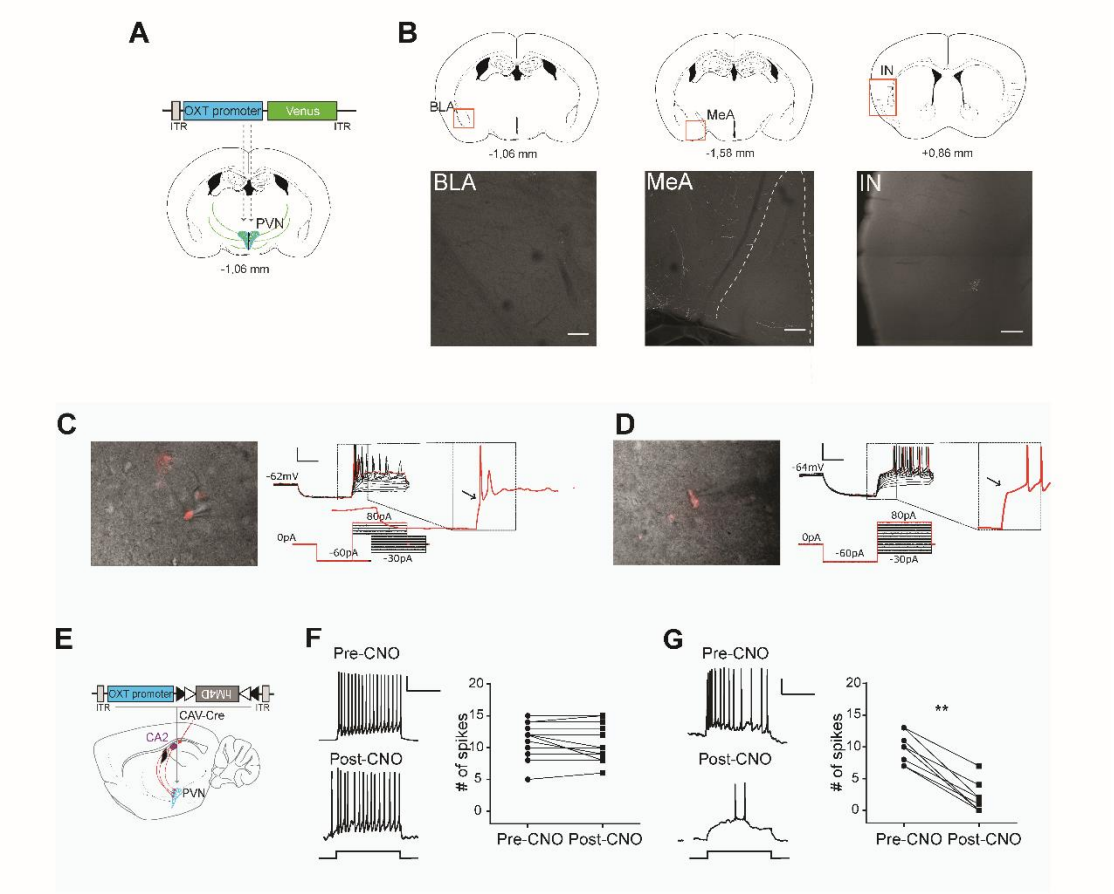

A
OXT promoter
Venus
ITR
ITR
PVN
-1,06 mm
B
BLA
-1,06 mm
MeA
-1,58 mm
IN
+0,86 mm
BLA
MeA
IN
C
-62mV
80pA
0pA
-60pA
-30pA
D
-64mV
80pA
0pA
-60pA
-30pA
E
OXT promoter
hM4Di
ITR
ITR
CAV-Cre
CA2
PVN
F
Pre-CNO
Post-CNO
# of spikes
Pre-CNO Post-CNO
G
Pre-CNO
Post-CNO
**
# of spikes
Pre-CNO Post-CNO
H
OTp-hM4D-mcherry
PVN
-0,83
-0,94
-1,06
-1,22
I
OTp-DIO-hM4D-mcherry
CAV-Cre
CeA
PVN
J
OTp-DIO-hM4D-mcherry
CAV-Cre
NAcc
PVN
K
OTp-DIO-hM4D-mcherry
CAV-Cre
mPFC
PVN
L
OTp-DIO-hM4D-mcherry
CAV-Cre
CA2
PVN

Figure S5. Anatomy of PVN OXT projections, *ex vivo* electrophysiological validation and placement of back-labeled PVN neurons, Related to Figure 4 and 5. (**A**) Scheme of the viral vector used to infect PVN OXT neurons. (**B**) OXT projections from the PVN to the basolateral amygdala (BLA), medial amygdala (MeA), and insular cortex (IN). Scale bar: 100 μm. (**C-D**) Representative IR-DIC image of back labeled (**C**) parvocellular and (**D**) magnocellular neurons during patch-clamp recordings. Right: in current clamp mode in response to depolarizing current pulses (I_{Hold}: 0 pA): (**C**) the parvocellular neuron lacks transient outward rectification (black arrow in the zoomed trace), whilst (**D**) the magnocellular neuron displays an inward rectification and a strong transient outward rectification (black arrow in zoomed trace). (**E**) Scheme of the viral vector used to infect the PVN OXT neurons and respective projecting area (CA2). (**F-G**) Example traces of membrane potential changes (left) and quantification (right) of the number of spikes evoked by a 1 sec of 20 pA current step (I_{Hold}: 0 pA) in (**F**) non infected neurons and in (**G**) PVN-CA2 neurons pre- and post- bath application of CNO. Scale bars: 40 mV and 500 ms. Two-tailed Wilcoxon matched-pairs signed rank test: (**F**) n=12 from 8 mice (p=0.25); (**G**) n=8 from 3 mice (p=0.0078). **p<0.01. (**H**) Placements of PVN injected with hM4D(Gi) DREADD receptor under the control of the OXT promoter. (**I-L**) Placements of selective PVN projecting neurons to the (**I**) CeA, (**J**) NAcc, (**K**) mPFC, and (**L**) CA2 obtained by bilateral injection in these areas of CAV2-Cre virus combined with bilateral injection in the PVN of rAAV carrying a double-floxed inverted open reading frame (ORF) (DIO) of the inhibitory hM4D(Gi) DREADD receptor and mCherry under the control of the OXT promoter.

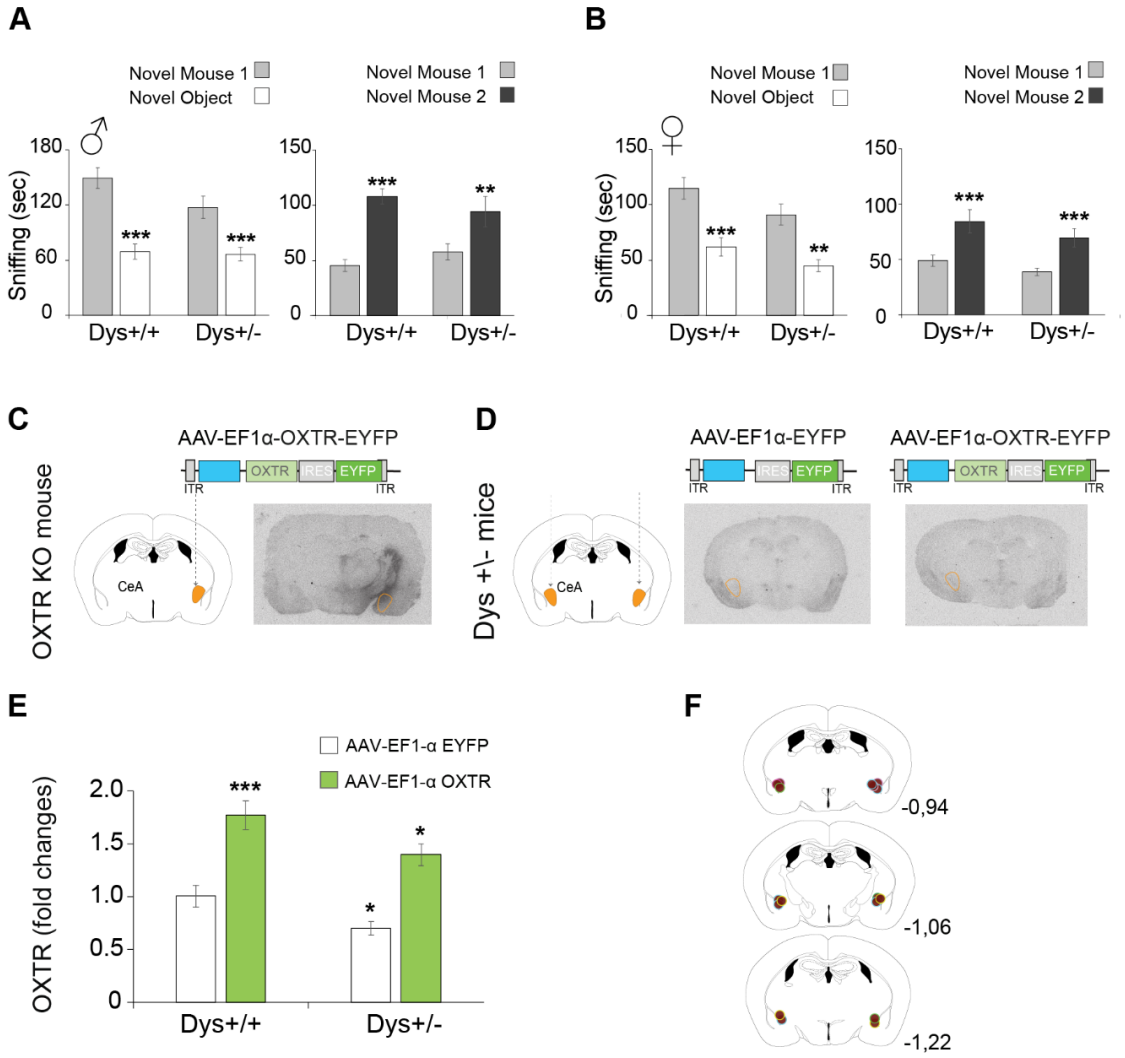


Figure S6. Dysbindin+/- show no deficits in sociability and social novelty in the 3-chamber task. Validation of AAV-mediated rescue of CeA OXTR levels, Related to Figure 6. (**A**) Time (in seconds) spent sniffing one unfamiliar mouse (novel mouse 1) and one novel object during the sociability phase of the 3-chamber task displayed by Dys+/+ and Dys+/- male and female littermates. RM ANOVA, Males Dys+/+: $F_{1,12}$=32.07, p=0.0001; Dys+/-: $F_{1,12}$=26.16, p=0.0003; Females Dys+/+: $F_{1,12}$=28.60, p=0.0002; Dys+/-: $F_{1,13}$=36.67, p=0.001. **p<0.005 and ***p<0.0005 versus the novel object within the same genotype. n=8/14 per group. (**B**) Time (in seconds) spent sniffing one unfamiliar mouse (novel mouse 2) and the now familiar mouse (novel mouse 1) during the social novelty phase of the 3-chamber task displayed by Dys+/+ and Dys+/- male and female littermates. RM ANOVA, Males Dys+/+: $F_{1,12}$=42.41, p<0.0001; Dys+/-: $F_{1,12}$=13.44, p=0.0032; Females Dys+/+: $F_{1,12}$=32.05, p=0.0001; Dys+/-: $F_{1,13}$=29.59, p=0.0001. **p<0.005 and ***p<0.0005 versus the novel mouse within the same genotype. n=8/14 per group. (**C**) Scheme of the viral vector used to unilaterally infect the CeA of OXTR-/- mice (left) and autoradiograms of OXT binding sites (right) that show efficiency and spreading of virus expression and spreading. (**D**) Scheme of the viral vectors used to bilaterally infect the CeA of Dys+/+ and Dys+/- littermates (left) and autoradiograms of OXT binding sites (right). Expression of OXTR was assessed by using the ligand binding of 20 pmol/l I^{125}-labeled OVTA. (**E**) Quantification of the autoradiographic I^{125} receptors in Dys+/+ and Dys+/- is expressed as fold changes from the Dys+/+ control group (Two-Way ANOVA genotype-by-treatment interaction $F_{2,13}$=12.86; p=0.0008). *p<0.05 and ***p<0.0005 versus Dys+/+ control virus group. n=3/6 per group. (**F**) Placements of Dys +/- mice injected in the CeA with AVV expressing OXTR and EYFP under the control of the EF1a promoter.
Error bars represent standard error of the mean.

References

1. Ferretti, V., and Papaleo, F. (2019). Understanding others: emotion recognition in humans and other animals. Genes Brain Behav. 18,e12544.
2. Dunbar, R.I. (2009). The social brain hypothesis and its implications for social evolution. Ann. Hum. Biol. 36, 562–572.
3. Henry, J.D., von Hippel, W., Molenberghs, P., Lee, T., and Sachdev, P.S. (2016). Clinical assessment of social cognitive function in neurological disorders. Nat. Rev. Neurol. 12, 28–39.
4. Kennedy, D.P., and Adolphs, R. (2012). Perception of emotions from facial expressions in high-functioning adults with autism. Neuropsychologia 50, 3313–3319.
5. Fletcher-Watson, S., McConnell, F., Manola, E., and McConachie, H. (2014). Interventions based on the theory of mind cognitive model for autism spectrum disorder (ASD). Cochrane Database Syst. Rev. (3), CD008785.
6. Kurtz, M.M., and Richardson, C.L. (2012). Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. Schizophr. Bull. 38, 1092–1104.
7. Fett, A.K., Viechtbauer, W., Dominguez, M.D., Penn, D.L., van Os, J., and Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci. Biobehav. Rev. 35, 573–588.
8. De Dreu, C.K., and Kret, M.E. (2016). Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense. Biol. Psychiatry 79, 165–173.
9. Guastella, A.J., Mitchell, P.B., and Mathews, F. (2008). Oxytocin enhances the encoding of positive social memories in humans. Biol. Psychiatry 64, 256–258.
10. Skuse, D.H., Lori, A., Cubells, J.F., Lee, I., Conneely, K.N., Puura, K., Lehtim€aki, T., Binder, E.B., and Young, L.J. (2014). Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills. Proc. Natl. Acad. Sci. USA 111, 1987–1992.
11. Choe, H.K., Reed, M.D., Benavidez, N., Montgomery, D., Soares, N., Yim, Y.S., and Choi, G.B. (2015). Oxytocin mediates entrainment of sensory stimuli to social cues of opposing valence. Neuron 87, 152–163.
12. Ferguson, J.N., Aldag, J.M., Insel, T.R., and Young, L.J. (2001). Oxytocin in the medial amygdala is essential for social recognition in the mouse. J. Neurosci. 21, 8278–8285.
13. Oettl, L.L., Ravi, N., Schneider, M., Scheller, M.F., Schneider, P., Mitre, M., da Silva Gouveia, M., Froemke, R.C., Chao, M.V., Young, W.S., et al. (2016). Oxytocin enhances social recognition by modulating cortical control of early olfactory processing. Neuron 90, 609–621.
14. Raam, T., McAvoy, K.M., Besnard, A., Veenema, A.H., and Sahay, A. (2017). Hippocampal oxytocin receptors are necessary for discrimination of social stimuli. Nat. Commun. 8, 2001.
15. Takayanagi, Y., Yoshida, M., Takashima, A., Takanami, K., Yoshida, S., Nishimori, K., Nishijima, I., Sakamoto, H., Yamagata, T., and Onaka, T. (2017). Activation of supraoptic oxytocin neurons by secretin facilitates social recognition. Biol. Psychiatry 81, 243–251.
16. Rogers-Carter, M.M., Varela, J.A., Gribbons, K.B., Pierce, A.F., McGoey, M.T., Ritchey, M., and Christianson, J.P. (2018). Insular cortex mediates approach and avoidance responses to social affective stimuli. Nat. Neurosci. 21, 404–414.
17. Ferguson, J.N., Young, L.J., Hearn, E.F., Matzuk, M.M., Insel, T.R., and Winslow, J.T. (2000). Social amnesia in mice lacking the oxytocin gene. Nat. Genet. 25, 284–288.
18. Lukas, M., Toth, I., Veenema, A.H., and Neumann, I.D. (2013). Oxytocin mediates rodent social memory within the lateral septum and the medial amygdala depending on the relevance of the social stimulus: male juvenile versus female adult conspecifics. Psychoneuroendocrinology 38, 916–926.
19. Burkett, J.P., Andari, E., Johnson, Z.V., Curry, D.C., de Waal, F.B., and Young, L.J. (2016). Oxytocin-dependent consolation behavior in rodents. Science 351, 375–378.
20. Do¨ len, G., Darvishzadeh, A., Huang, K.W., and Malenka, R.C. (2013). Social reward requires coordinated activity of nucleus accumbens oxytocin and serotonin. Nature 501, 179–184.

21. Hung, L.W., Neuner, S., Polepalli, J.S., Beier, K.T., Wright, M., Walsh, J.J., Lewis, E.M., Luo, L., Deisseroth, K., Do¨ len, G., and Malenka, R.C. (2017). Gating of social reward by oxytocin in the ventral tegmental area. Science 357, 1406–1411.
22. Knobloch, H.S., Charlet, A., Hoffmann, L.C., Eliava, M., Khrulev, S., Cetin, A.H., Osten, P., Schwarz, M.K., Seeburg, P.H., Stoop, R., and Grinevich, V. (2012). Evoked axonal oxytocin release in the central amygdala attenuates fear response. Neuron 73, 553–566.
23. Pisansky, M.T., Hanson, L.R., Gottesman, I.I., and Gewirtz, J.C. (2017). Oxytocin enhances observational fear in mice. Nat. Commun. 8, 2102.
24. Menon, R., Grund, T., Zoicas, I., Althammer, F., Fiedler, D., Biermeier, V., Bosch, O.J., Hiraoka, Y., Nishimori, K., Eliava, M., et al. (2018). Oxytocin signaling in the lateral septum prevents social fear during lactation. Curr. Biol. 28, 1066–1078.e6.
25. Nakajima, M., Go¨ rlich, A., and Heintz, N. (2014). Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. Cell 159, 295–305.
26. Scott, N., Prigge, M., Yizhar, O., and Kimchi, T. (2015). A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion. Nature 525, 519–522.
27. Guastella, A.J., Gray, K.M., Rinehart, N.J., Alvares, G.A., Tonge, B.J., Hickie, I.B., Keating, C.M., Cacciotti-Saija, C., and Einfeld, S.L. (2015). The effects of a course of intranasal oxytocin on social behaviors in youth diagnosed with autism spectrum disorders: a randomized controlled trial. J. Child Psychol. Psychiatry 56, 444–452.
28. Aoki, Y., Yahata, N., Watanabe, T., Takano, Y., Kawakubo, Y., Kuwabara, H., Iwashiro, N., Natsubori, T., Inoue, H., Suga, M., et al. (2014). Oxytocin improves behavioural and neural deficits in inferring others' social emotions in autism. Brain 137, 3073–3086.
29. Bakermans-Kranenburg, M.J., and van I Jzendoorn, M.H. (2013). Sniffing around oxytocin: review and meta-analyses of trials in healthy and clinical groups with implications for pharmacotherapy. Transl. Psychiatry 3, e258.
30. Cacciotti-Saija, C., Langdon, R., Ward, P.B., Hickie, I.B., Scott, E.M., Naismith, S.L., Moore, L., Alvares, G.A., Redoblado Hodge, M.A., and Guastella, A.J. (2015). A double-blind randomized controlled trial of oxytocin nasal spray and social cognition training for young people with early psychosis. Schizophr. Bull. 41, 483–493.
31. Young, L.J., and Barrett, C.E. (2015). Neuroscience. Can oxytocin treat autism? Science 347, 825–826.
32. Walum, H., Waldman, I.D., and Young, L.J. (2016). Statistical and methodological considerations for the interpretation of intranasal oxytocin studies. Biol. Psychiatry 79, 251–257.
33. Domes, G., Heinrichs, M., Gl€ascher, J., Bu¨ chel, C., Braus, D.F., and Herpertz, S.C. (2007). Oxytocin attenuates amygdala responses to emotional faces regardless of valence. Biol. Psychiatry 62, 1187–1190.
34. Domes, G., Sibold, M., Schulze, L., Lischke, A., Herpertz, S.C., and Heinrichs, M. (2013). Intranasal oxytocin increases covert attention to positive social cues. Psychol. Med. 43, 1747–1753.
35. Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. Nature 435, 673–676.
36. Leppanen, J., Ng, K.W., Tchanturia, K., and Treasure, J. (2017). Metaanalysis of the effects of intranasal oxytocin on interpretation and expression of emotions. Neurosci. Biobehav. Rev. 78, 125–144.
37. Marsh, A.A., Yu, H.H., Pine, D.S., and Blair, R.J. (2010). Oxytocin improves specific recognition of positive facial expressions. Psychopharmacology (Berl.) 209, 225–232.
38. Shahrestani, S., Kemp, A.H., and Guastella, A.J. (2013). The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. Neuropsychopharmacology 38, 1929–1936.
39. Van I Jzendoorn, M.H., and Bakermans-Kranenburg, M.J. (2012). A sniff of trust: meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. Psychoneuroendocrinology 37, 438–443.

40. Chen, F.S., Kumsta, R., Dvorak, F., Domes, G., Yim, O.S., Ebstein, R.P., and Heinrichs, M. (2015). Genetic modulation of oxytocin sensitivity: a pharmacogenetic approach. Transl. Psychiatry 5, e664.
41. Tost, H., Kolachana, B., Hakimi, S., Lemaitre, H., Verchinski, B.A., Mattay, V.S., Weinberger, D.R., and Meyer-Lindenberg, A. (2010). A common allele in the oxytocin receptor gene (OXTR) impacts prosocial temperament and human hypothalamic-limbic structure and function. Proc. Natl. Acad. Sci. USA 107, 13936–13941.
42. Meyza, K.Z., Bartal, I.B., Monfils, M.H., Panksepp, J.B., and Knapska, E. (2017). The roots of empathy: Through the lens of rodent models. Neurosci. Biobehav. Rev. 76 (Pt B), 216–234.
43. Green, M.F., Horan, W.P., and Lee, J. (2015). Social cognition in schizophrenia. Nat. Rev. Neurosci. 16, 620–631.
44. Barber, A.L., Randi, D., Mu¨ ller, C.A., and Huber, L. (2016). The processing of human emotional faces by pet and lab dogs: evidence for lateralization and experience effects. PLoS ONE 11, e0152393.
45. Parr, L.A., Waller, B.M., and Heintz, M. (2008). Facial expression categorization by chimpanzees using standardized stimuli. Emotion 8, 216–231.
46. Tate, A.J., Fischer, H., Leigh, A.E., and Kendrick, K.M. (2006). Behavioural and neurophysiological evidence for face identity and face emotion processing in animals. Philos. Trans. R. Soc. Lond. B Biol. Sci. 361, 2155–2172.
47. Wathan, J., Proops, L., Grounds, K., and McComb, K. (2016). Horses discriminate between facial expressions of conspecifics. Sci. Rep. 6, 38322.
48. Papaleo, F., Burdick, M.C., Callicott, J.H., and Weinberger, D.R. (2014). Epistatic interaction between COMT and DTNBP1 modulates prefrontal function in mice and in humans. Mol. Psychiatry 19, 311–316.
49. Papaleo, F., Yang, F., Garcia, S., Chen, J., Lu, B., Crawley, J.N., and Weinberger, D.R. (2012). Dysbindin-1 modulates prefrontal cortical activity and schizophrenia-like behaviors via dopamine/D2 pathways. Mol. Psychiatry 17, 85–98.
50. Scheggia, D., Mastrogiacomo, R., Mereu, M., Sannino, S., Straub, R.E., Armando, M., Manago` , F., Guadagna, S., Piras, F., Zhang, F., et al. (2018). Variations in Dysbindin-1 are associated with cognitive response to antipsychotic drug treatment. Nat. Commun. 9, 2265.
51. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., Coleman, J.R.I., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat. Genet. 50, 912–919.
52. LeDoux, J. (1996). Emotional networks and motor control: a fearful view. Prog. Brain Res. 107, 437–446.
53. Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H.E., Lin, S.Y., Rabah, D., Kinet, J.P., and Shin, H.S. (2010). Observational fear learning involves affective pain system and Cav1.2 Ca^{2+} channels in ACC. Nat. Neurosci. 13, 482–488.
54. Langford, D.J., Crager, S.E., Shehzad, Z., Smith, S.B., Sotocinal, S.G., Levenstadt, J.S., Chanda, M.L., Levitin, D.J., and Mogil, J.S. (2006). Social modulation of pain as evidence for empathy in mice. Science 312, 1967–1970.
55. Monfils, M.H., and Agee, L.A. (2019). Insights from social transmission of information in rodents. Genes Brain Behav. 18, e12534.
56. Scattoni, M.L., Crawley, J., and Ricceri, L. (2009). Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. Neurosci. Biobehav. Rev. 33, 508–515.
57. Wo¨ hr, M., and Schwarting, R.K. (2013). Affective communication in rodents: ultrasonic vocalizations as a tool for research on emotion and motivation. Cell Tissue Res. 354, 81–97.
58. Portfors, C.V. (2007). Types and functions of ultrasonic vocalizations in laboratory rats and mice. J. Am. Assoc. Lab. Anim. Sci. 46, 28–34.

59. Blanchard, R.J., Blanchard, D.C., Agullana, R., and Weiss, S.M. (1991). Twenty-two kHz alarm cries to presentation of a predator, by laboratory rats living in visible burrow systems. Physiol. Behav. 50, 967–972.
60. Brudzynski, S.M. (2013). Ethotransmission: communication of emotional states through ultrasonic vocalization in rats. Curr. Opin. Neurobiol. 23, 310–317.
61. Kim, E.J., Kim, E.S., Covey, E., and Kim, J.J. (2010). Social transmission of fear in rats: the role of 22-kHz ultrasonic distress vocalization. PLoS ONE 5, e15077.
62. Pereira, A.G., Cruz, A., Lima, S.Q., and Moita, M.A. (2012). Silence resulting from the cessation of movement signals danger. Curr. Biol. 22, R627– R628.
63. Boillat, M., Challet, L., Rossier, D., Kan, C., Carleton, A., and Rodriguez, I. (2015). The vomeronasal system mediates sick conspecific avoidance. Curr. Biol. 25, 251–255.
64. Sawyer, T.F. (1980). Androgen effects on responsiveness to aggression and stress-related odors of male mice. Physiol. Behav. 25, 183–187.
65. Zalaquett, C., and Thiessen, D. (1991). The effects of odors from stressed mice on conspecific behavior. Physiol. Behav. 50, 221–227.
66. NIH. Research domain criteria (RDoC). https://www.nimh.nih.gov/ research-priorities/rdoc/index.shtml.
67. Huang, H., Michetti, C., Busnelli, M., Manago` , F., Sannino, S., Scheggia, D., Giancardo, L., Sona, D., Murino, V., Chini, B., et al. (2014). Chronic and acute intranasal oxytocin produce divergent social effects in mice. Neuropsychopharmacology 39, 1102–1114.
68. Anderson, D.J., and Adolphs, R. (2014). A framework for studying emotions across species. Cell 157, 187–200.
69. Eliava, M., Melchior, M., Knobloch-Bollmann, H.S., Wahis, J., da Silva Gouveia, M., Tang, Y., Ciobanu, A.C., Triana Del Rio, R., Roth, L.C., Althammer, F., et al. (2016). A new population of parvocellular oxytocin neurons controlling magnocellular neuron activity and inflammatory pain processing. Neuron 89, 1291–1304.
70. Mormann, F., Niediek, J., Tudusciuc, O., Quesada, C.M., Coenen, V.A., Elger, C.E., and Adolphs, R. (2015). Neurons in the human amygdala encode face identity, but not gaze direction. Nat. Neurosci. 18, 1568– 1570.
71. Stanley, D.A., and Adolphs, R. (2013). Toward a neural basis for social behavior. Neuron 80, 816–826.
72. Tottenham, N., Hertzig, M.E., Gillespie-Lynch, K., Gilhooly, T., Millner, A.J., and Casey, B.J. (2014). Elevated amygdala response to faces and gaze aversion in autism spectrum disorder. Soc. Cogn. Affect. Neurosci. 9, 106–117.
73. Goghari, V.M., Sanford, N., Spilka, M.J., and Woodward, T.S. (2017). Task-related functional connectivity analysis of emotion discrimination in a family study of schizophrenia. Schizophr. Bull. 43, 1348–1362.
74. Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and prosocial behavior in rats. Science 334, 1427–1430. 75. Keum, S., Park, J., Kim, A., Park, J., Kim, K.K., Jeong, J., and Shin, H.S. (2016). Variability in empathic fear response among 11 inbred strains of mice. Genes Brain Behav. 15, 231–242.
75. Martin, L.J., Hathaway, G., Isbester, K., Mirali, S., Acland, E.L., Niederstrasser, N., Slepian, P.M., Trost, Z., Bartz, J.A., Sapolsky, R.M., et al. (2015). Reducing social stress elicits emotional contagion of pain in mouse and human strangers. Curr. Biol. 25, 326–332.
76. Sivaselvachandran, S., Acland, E.L., Abdallah, S., and Martin, L.J. (2018). Behavioral and mechanistic insight into rodent empathy. Neurosci. Biobehav. Rev. 91, 130–137.
77. Keum, S., Kim, A., Shin, J.J., Kim, J.H., Park, J., and Shin, H.S. (2018). A missense variant at the Nrxn3 locus enhances empathy fear in the mouse. Neuron 98, 588–601.e5.
78. Knapska, E., Mikosz, M., Werka, T., and Maren, S. (2009). Social modulation of learning in rats. Learn. Mem. 17, 35–42.

79. Sterley, T.L., Baimoukhametova, D., Fu¨ zesi, T., Zurek, A.A., Daviu, N., Rasiah, N.P., Rosenegger, D., and Bains, J.S. (2018). Social transmission and buffering of synaptic changes after stress. Nat. Neurosci. 21, 393–403.
80. Blanchard, R.J., and Blanchard, D.C. (1989). Attack and defense in rodents as ethoexperimental models for the study of emotion. Prog. Neuropsychopharmacol. Biol. Psychiatry 13 (Suppl), S3–S14.
81. Sadananda, M., Wo¨ hr, M., and Schwarting, R.K. (2008). Playback of 22-kHz and 50-kHz ultrasonic vocalizations induces differential c-fos expression in rat brain. Neurosci. Lett. 435, 17–23.
82. Mandairon, N., Poncelet, J., Bensafi, M., and Didier, A. (2009). Humans and mice express similar olfactory preferences. PLoS ONE 4, e4209.
83. Moncho-Bogani, J., Lanuza, E., Lorente, M.J., and Martinez-Garcia, F. (2004). Attraction to male pheromones and sexual behaviour show different regulatory mechanisms in female mice. Physiol. Behav. 81, 427–434.
84. McHenry, J.A., Otis, J.M., Rossi, M.A., Robinson, J.E., Kosyk, O., Miller, N.W., McElligott, Z.A., Budygin, E.A., Rubinow, D.R., and Stuber, G.D. (2017). Hormonal gain control of a medial preoptic area social reward circuit. Nat. Neurosci. 20, 449–458.
85. Panksepp, J.B., Jochman, K.A., Kim, J.U., Koy, J.J., Wilson, E.D., Chen, Q., Wilson, C.R., and Lahvis, G.P. (2007). Affiliative behavior, ultrasonic communication and social reward are influenced by genetic variation in adolescent mice. PLoS ONE 2, e351.
86. Venniro, M., Zhang, M., Caprioli, D., Hoots, J.K., Golden, S.A., Heins, C., Morales, M., Epstein, D.H., and Shaham, Y. (2018). Volitional social interaction prevents drug addiction in rat models. Nat. Neurosci. 21, 1520–1529.
87. Balleine, B.W., and Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. Trends Neurosci. 29, 272–279.
88. Corbit, L.H., and Balleine, B.W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. J. Neurosci. 25, 962–970.
89. Dricu, M., and Fru¨ hholz, S. (2016). Perceiving emotional expressions in others: Activation likelihood estimation meta-analyses of explicit evaluation, passive perception and incidental perception of emotions. Neurosci. Biobehav. Rev. 71, 810–828.
90. Morrison, S.E., and Salzman, C.D. (2010). Re-valuing the amygdala. Curr. Opin. Neurobiol. 20, 221–230.
91. Phelps, E.A., and LeDoux, J.E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48, 175–187.
92. Janak, P.H., and Tye, K.M. (2015). From circuits to behaviour in the amygdala. Nature 517, 284–292.
93. Rickenbacher, E., Perry, R.E., Sullivan, R.M., and Moita, M.A. (2017). Freezing suppression by oxytocin in central amygdala allows alternate defensive behaviours and mother-pup interactions. eLife 6, e24080.
94. Terburg, D., Scheggia, D., Triana Del Rio, R., Klumpers, F., Ciobanu, A.C., Morgan, B., Montoya, E.R., Bos, P.A., Giobellina, G., van den Burg, E.H., et al. (2018). The basolateral amygdala is essential for rapid escape: a human and rodent study. Cell 175, 723–735.e16.
95. Twining, R.C., Vantrease, J.E., Love, S., Padival, M., and Rosenkranz, J.A. (2017). An intra-amygdala circuit specifically regulates social fear learning. Nat. Neurosci. 20, 459–469.
96. Gallagher, M., Graham, P.W., and Holland, P.C. (1990). The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. J. Neurosci. 10, 1906–1911.
97. Genud-Gabai, R., Klavir, O., and Paz, R. (2013). Safety signals in the primate amygdala. J. Neurosci. 33, 17986–17994.
98. Murray, E.A. (2007). The amygdala, reward and emotion. Trends Cogn. Sci. 11, 489–497.

99. Rogan, M.T., Leon, K.S., Perez, D.L., and Kandel, E.R. (2005). Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. Neuron 46, 309–320.
100. Tye, K.M., Cone, J.J., Schairer, W.W., and Janak, P.H. (2010). Amygdala neural encoding of the absence of reward during extinction. J. Neurosci. 30, 116–125.
101. Haram, M., Bettella, F., Brandt, C.L., Quintana, D.S., Nerhus, M., Bjella, T., Djurovic, S., Westlye, L.T., Andreassen, O.A., Melle, I., and Tesli, M. (2016). Contribution of oxytocin receptor polymorphisms to amygdala activation in schizophrenia spectrum disorders. BJPsych Open 2, 353–358.
102. Gamer, M., Zurowski, B., and Bu¨ chel, C. (2010). Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. Proc. Natl. Acad. Sci. USA 107, 9400–9405.
103. Viviani, D., Charlet, A., van den Burg, E., Robinet, C., Hurni, N., Abatis, M., Magara, F., and Stoop, R. (2011). Oxytocin selectively gates fear responses through distinct outputs from the central amygdala. Science 333, 104–107.
104. Zoicas, I., Slattery, D.A., and Neumann, I.D. (2014). Brain oxytocin in social fear conditioning and its extinction: involvement of the lateral septum. Neuropsychopharmacology 39, 3027–3035.
105. Lindquist, K.A., Satpute, A.B., Wager, T.D., Weber, J., and Barrett, L.F. (2016). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. Cereb. Cortex 26, 1910–1922.
106. Liu, Y., Li, S., Lin, W., Li, W., Yan, X., Wang, X., Pan, X., Rutledge, R.B., and Ma, Y. (2019). Oxytocin modulates social value representations in the amygdala. Nat. Neurosci. 22, 633–641.
107. Huber, D., Veinante, P., and Stoop, R. (2005). Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. Science 308, 245–248.
108. Marlin, B.J., Mitre, M., D'amour, J.A., Chao, M.V., and Froemke, R.C. (2015). Oxytocin enables maternal behaviour by balancing cortical inhibition. Nature 520, 499–504.
109. Owen, S.F., Tuncdemir, S.N., Bader, P.L., Tirko, N.N., Fishell, G., and Tsien, R.W. (2013). Oxytocin enhances hippocampal spike transmission by modulating fast-spiking interneurons. Nature 500, 458–462.
110. Grinevich, V., and Stoop, R. (2018). Interplay between oxytocin and sensory systems in the orchestration of socio-emotional behaviors. Neuron 99, 887–904.
111. Ingallinesi, M., Rouibi, K., Le Moine, C., Papaleo, F., and Contarino, A. (2012). CRF2 receptor-deficiency eliminates opiate withdrawal distress without impairing stress coping. Mol. Psychiatry 17, 1283–1294.
112. Regev, L., Tsoory, M., Gil, S., and Chen, A. (2012). Site-specific genetic manipulation of amygdala corticotropin-releasing factor reveals its imperative role in mediating behavioral response to challenge. Biol. Psychiatry 71, 317–326.
113. Beyeler, A., Chang, C.J., Silvestre, M., L_eveˆ que, C., Namburi, P., Wildes, C.P., and Tye, K.M. (2018). Organization of valence-encoding and projection- defined neurons in the basolateral amygdala. Cell Rep. 22, 905–918.
114. Namburi, P., Beyeler, A., Yorozu, S., Calhoon, G.G., Halbert, S.A., Wichmann, R., Holden, S.S., Mertens, K.L., Anahtar, M., Felix-Ortiz, A.C., et al. (2015). A circuit mechanism for differentiating positive and negative associations. Nature 520, 675–678.
115. Schaller, F., Watrin, F., Sturny, R., Massacrier, A., Szepetowski, P., and Muscatelli, F. (2010). A single postnatal injection of oxytocin rescues the lethal feeding behaviour in mouse newborns deficient for the imprinted Magel2 gene. Hum. Mol. Genet. 19, 4895–4905.
116. Gigliucci, V., Leonzino, M., Busnelli, M., Luchetti, A., Palladino, V.S., D'Amato, F.R., and Chini, B. (2014). Region specific up-regulation of oxytocin receptors in the opioid oprm1 (-/-) mouse model of autism. Front Pediatr. 2, 91.
117. Pen˜ agarikano, O., La´ zaro, M.T., Lu, X.H., Gordon, A., Dong, H., Lam, H.A., Peles, E., Maidment, N.T., Murphy, N.P., Yang, X.W., et al. (2015). Exogenous and evoked oxytocin restores social behavior in the Cntnap2 mouse model of autism. Sci. Transl. Med. 7, 271ra8.

118. Francis, S.M., Sagar, A., Levin-Decanini, T., Liu, W., Carter, C.S., and Jacob, S. (2014). Oxytocin and vasopressin systems in genetic syndromes and neurodevelopmental disorders. Brain Res. 1580, 199–218.
119. Feng, C., Lori, A., Waldman, I.D., Binder, E.B., Haroon, E., and Rilling, J.K. (2015). A common oxytocin receptor gene (OXTR) polymorphism modulates intranasal oxytocin effects on the neural response to social cooperation in humans. Genes Brain Behav. 14, 516–525.
120. Cecilione, J.L., Rappaport, L.M., Verhulst, B., Carney, D.M., Blair, R.J.R., Brotman, M.A., Leibenluft, E., Pine, D.S., Roberson-Nay, R., and Hettema, J.M. (2017). Test-retest reliability of the facial expression labeling task. Psychol. Assess. 29, 1537–1542.
121. Bru, T., Salinas, S., and Kremer, E.J. (2010). An update on canine adenovirus type 2 and its vectors. Viruses 2, 2134–2153.
122. Takayanagi, Y., Yoshida, M., Bielsky, I.F., Ross, H.E., Kawamata, M., Onaka, T., Yanagisawa, T., Kimura, T., Matzuk, M.M., Young, L.J., and Nishimori, K. (2005). Pervasive social deficits, but normal parturition, in oxytocin receptor-deficient mice. Proc. Natl. Acad. Sci. USA 102, 16096–16101.
123. Scheggia, D., Zamberletti, E., Realini, N., Mereu, M., Contarini, G., Ferretti, V., Manago` , F., Margiani, G., Brunoro, R., Rubino, T., et al. (2018). Remote memories are enhanced by COMT activity through dysregulation of the endocannabinoid system in the prefrontal cortex. Mol. Psychiatry 23, 1040–1050.
124. Manago` , F., Mereu, M., Mastwal, S., Mastrogiacomo, R., Scheggia, D., Emanuele, M., De Luca, M.A., Weinberger, D.R., Wang, K.H., and Papaleo, F. (2016). Genetic disruption of Arc/Arg3.1 in mice causes alterations in dopamine and neurobehavioral phenotypes related to schizophrenia. Cell Rep. 16, 2116–2128.
125. Contarino,A.,Kitchener, P., Vall_ee,M., Papaleo, F., and Piazza,P.V. (2017). CRF1 receptor-deficiency increases cocaine reward. Neuropharmacology 117, 41–48.
126. Franklin, K.B.J., and Paxinos, G. (2007). The Mouse Brain in Stereotaxic Coordinates, Third Edition (Academic Press).
127. Luther, J.A., Halmos, K.C., and Tasker, J.G. (2000). A slow transient potassium current expressed in a subset of neurosecretory neurons of the hypothalamic paraventricular nucleus. J. Neurophysiol. 84, 1814– 1825.

CHAPTER 3

Oxytocin Discrepancies in Social Dynamics

Federica Maltese[1] and Francesco Papaleo[1]

[1]*Genetics of Cognition Laboratory, Neuroscience Area, Istituto Italiano di Tecnologia, Genova, Italy*

Published in *Neuron*. 2020 Aug 19; 107(4):591-593.

Comment on

Wireless Optogenetic Stimulation of Oxytocin Neurons in a Semi-natural Setup Dynamically Elevates Both Pro-social and Agonistic Behaviors.

Anpilov S, Shemesh Y, Eren N, Harony-Nicolas H, Benjamin A, Dine J, Oliveira VEM, Forkosh O, Karamihalev S, Hüttl RE, Feldman N, Berger R, Dagan A, Chen G, Neumann ID, Wagner S, Yizhar O, Chen A.

Neuron. 2020 Aug 19; 107(4):644-655

Abstract

Social group dynamics are highly complex. In this issue of Neuron, Anpilov et al. use a novel wireless optogenetic device to demonstrate that the repeated stimulation of oxytocin neurons modulates pro-social and agonistic behaviors in a time- and context-dependent manner.

Main Text

Social behavior refers to several processes associated with interactions with others; however, behaviors can differ greatly depending on social contexts and the goal of the interaction. This complexity may be reflected by the underlying brain mechanisms. Diverse neuronal processes may drive similar social responses, or the same neuronal pathway may differentially influence disparate social behaviors.

The development of cutting-edge techniques has enabled the study and manipulation of brain circuits in rodents with increasing spatial and temporal resolution. However, these technical achievements are not always paired with the proper assessment of distinct social behaviors. Mechanistic studies in rodents often rely on basic measurements within a single context to drive conclusions regarding social behaviors. The comprehensive investigation of cellular and circuital mechanisms must occur in parallel with the awareness of social process complexity; otherwise, the findings may remain incomplete or misleading.

Several paradigms can efficiently assess different social behaviors among rodents, including reciprocal exploration, communal nesting, sexual-parenting behaviors, territorial marking, aggressiveness, social memory, social hierarchies, social reward, emotion discrimination, consolation, pair-bonding, and harm avoidance. Similar to humans, rodents show complex group dynamics, forming multifaceted social structures that are shaped by the surrounding environment. Intriguingly, extended observations of groups of mice living in enriched environments have revealed that group dynamics may change over time, and individual differences may be identifiable that approximate human personality traits [1, 2]. These considerations highlight the importance of combining disparate and complimentary social assessments, using different timescales, to make appropriate inferences regarding social behaviors.

In this issue of Neuron, Anpilov et al. (2020) [3] developed a novel wireless optogenetic tool that was used to manipulate oxytocin (OXT) neurons in the paraventricular nucleus (PVN) of

the hypothalamus of male mice. Two different social contexts were used: (1) a semi-naturalistic arena, in which grouped mice live together for several days, and (2) a resident-intruder task. The social arena setting can be used to study the social dynamics of groups and individual behaviors across several days [1]. The resident-intruder test, in contrast, allows the assessment of environmentally induced agonistic behaviors during a discrete period. The wireless optogenetic device was magnetically activated approximately every 2 h, for 2 consecutive nights, in both social contexts. In the social arena, the magnetic activation was paired with food retrieval, whereas in the resident-intruder task, magnet activation was performed for two nights preceding a short interaction with a stranger male after 6 days of housing with a female mouse. The authors used a Stabilized Step-Function Opsin (SSFO), which allowed them to obtain a prolonged physiological like activation of targeted neurons through the increased sensitization to afferent inputs following a single light pulse, rather than inducing firing directly. The activation of PVN OXT neurons induced membrane depolarization for up to 30 min and increased OXT plasma levels. Moreover, this manipulation did not lead to OXT receptor (OXTR) downregulation, which normally occurs after exogenous OXT applications [4]. Thus, the remotely controlled system, together with SSFO, allowed the induction of a physiologic and prolonged OXT release in a complex social environment without interfering with social interactions.

Endogenous PVN OXT signaling has been reported to usually have ‘‘pro-social’’ effects, being necessary for social fear extinction [5] and for the discrimination of altered affective states in others [6]. Exogenous OXT treatments have been reported to produce both pro-social and anti-social responses and no effects at all [4, 7]. Anpilov et al. support this OXT complexity by showing that the same prolonged PVN OXT stimulation reduced agonistic behaviors in the resident-intruder test and increased them in the semi-naturalistic arena. These discrepant effects may depend on the diversity of social contexts and the salience of social cues. In the social arena, mice spend many days with their cagemates in an enriched environment [1]. In contrast, in the resident-intruder setting, following a prolonged pairing with a female, the resident mouse must interact in his homecage with a stranger male from another strain for a short period. Living in isolation or in groups, interacting with cagemates or with strangers, and having a mating experience are all factors that can influence the OXT system. Furthermore, in the social arena, all mice received the optogenetic stimulation, whereas only the resident mouse was stimulated in the resident-intruder test. Thus, the differential effects of opto-stimulated OXT neurons may depend on the distinct

basal status of the OXT system in each individual within the social interactions, which could be relevant when considering the development of personalized treatments targeting the OXT system. The opposing effects of OXT stimulation reported by Anpilov et al. appear to agree with the OXT salience hypothesis, which states that OXT effects can be different depending on the surrounding environment and the goal of the social interaction. However, their findings (reduced aggression versus a stranger and increased agonistic behaviors versus cagemates) appear to disagree with evidence on OXT differential impact on in-group versus out-group sociability. Some human studies have shown that OXT may increase trust and empathy toward in-group but not out-group members, and OXT may also increase aggression toward out-group members [8]. These latter studies utilized exogenous OXT, which, by downregulating the OXTR, could potentially produce opposite outcomes. The approach applied by Anpilov et al. demonstrated the importance of assessing and comparing different social settings to reach meaningful conclusions regarding the OXT modulation of social processes. Within the same social environment, Anpilov et al. found that prolonged OXT activation could differently modulate social dynamics in a time-dependent manner. The first day of PVN OXT manipulation within the enriched social arena resulted in increased affiliative behaviors. However, starting from the second day of activation, increased agonistic behaviors became evident. The reason for this time-sensitive differential effect of OXT stimulation remains unclear. Nevertheless, the wireless device activation was linked to the feeders, which were accessible for only a limited time to mimic competition for food. Under aversive conditions, OXT may increase responses toward threat signals more than those associated with positive social cues [9]. This result also agrees with the allostatic OXT theory, in which the OXT system may be able to predict environmental changes to adapt its responses depending on survival and adaptation needs [10]. However, prolonged OXT circuit activation had no effects on identity domains (IDs), categorized by Forkosh et al., (2019) [1]. This result indicated that protracted OXT stimulation cannot change the overall behavioral status of a subject, and ID features are likely not sensitive enough to detect discrete changes in social dynamics. Similarly, the authors found that grooming behaviors were not correlated with modifications in-group dynamics; therefore, the role of OXT-mediated grooming in the context of social stress remains unclear.

This study emphasized the importance of combining advanced biological and computational tools, using different behavioral contexts and temporal investigations, to better define social brain functioning. The optogenetic strategy developed by Anpilov et al., combined

with their enriched behavioral set-up, might present new opportunities for the investigation of social dynamics in a more ethologically relevant manner.

This work introduces further outstanding questions that remain to be addressed. Depending on the social behavior investigated, OXT effects may be sex-dependent. Thus, investigating the impacts of prolonged OXT stimulation on female group dynamics or larger groups of mice, in which both males and females are present, would be highly relevant. Another question is how group dynamics would be influenced if the wireless manipulation was only applied to one or a few mice within each group. This is connected with the impacts of personal traits on OXT action, which still need to be understood. Indeed, OXT effects can vary across individuals based on their own personal traits [9]. Moreover, the consideration of individual features will be crucial to achieve biologically supported interpretations of acquired data and to better guide the development of personalized clinical interventions. Finally, the mechanisms underlying the salience hypothesis of OXT remain to be uncovered. OXT has been reported to interact with many neurotransmitters, including acetylcholine, glutamate, GABA, dopamine, and serotonin. For example, in cooperation with the dopamine system, OXT might increase the salience of social cues by modulating attentional orienting responses [9]. Furthermore, OXT has an almost identical structure to vasopressin (AVP), potentially activating the same receptor populations. In particular, the binding of OXT with AVPRs or OXTR may vary depending on the behavior and context, and some magnocellular neurons in the PVN might co-release both AVP and OXT. Because opposite effects have been reported for AVP and OXT during different behaviors, the mechanisms underlying OXT-AVP crosstalk will be important to probe and may be implicated in the ‘‘apparent’’ dual role for the OXT modulation of social behaviors. Uncovering the interplay between OXT and other neurotransmitters and neuromodulators during disparate social contexts represents an essential step toward examining the role of OXT in the social brain.

References

1. Forkosh, O., Karamihalev, S., Roeh, S., Alon, U., Anpilov, S., Touma, C., Nussbaumer, M., Flachskamm, C., Kaplick, P.M., Shemesh, Y., and Chen, A. (2019). Identity domains capture individual differences from across the behavioral repertoire. Nat. Neurosci. 22, 2023–2028.
2. Freund, J., Brandmaier, A.M., Lewejohann, L., Kirste, I., Kritzler, M., Kr€uger, A., Sachser, N., Lindenberger, U., and Kempermann, G. (2013). Emergence of individuality in genetically identical mice. Science 340, 756–759.
3. Anpilov, S., Shemesh, Y., Eren, N., Harony- Nicolas, H., Benjamin, A., Dine, J., Oliveira, V.E.M., Forkosh, O., Karamihalev, S., Hüttl, R.E., et al. (2020). Wireless Optogenetic Stimulation of Oxytocin Neurons in a Semi-natural Setup Dynamically Elevates Both Pro-social and Agonistic Behaviors. Neuron 107, 644–655.
4. Huang, H., Michetti, C., Busnelli, M., Manago` , F., Sannino, S., Scheggia, D., Giancardo, L., Sona, D., Murino, V., Chini, B., et al. (2014). Chronic and acute intranasal oxytocin produce divergent social effects in mice. Neuropsychopharmacology 39, 1102–1114.
5. Menon, R., Grund, T., Zoicas, I., Althammer, F., Fiedler, D., Biermeier, V., Bosch, O.J., Hiraoka, Y., Nishimori, K., Eliava, M., et al. (2018). Oxytocin Signaling in the Lateral Septum Prevents Social Fear during Lactation. Curr Biol 28, 1066–1078.
6. Ferretti, V., Maltese, F., Contarini, G., Nigro, M., Bonavia, A., Huang, H., Gigliucci, V., Morelli, G., Scheggia, D., Manago, F., et al. (2019). Oxytocin Signaling in the Central Amygdala Modulates Emotion Discrimination in Mice. Curr Biol 29, 1938–1953.
7. Quintana, D.S. (2018). Revisiting non-significant effects of intranasal oxytocin using equivalence testing. Psychoneuroendocrinology 87, 127–130.
8. De Dreu, C.K., Greer, L.L., Handgraaf, M.J., Shalvi, S., Van Kleef, G.A., Baas, M., Ten Velden, F.S., Van Dijk, E., and Feith, S.W. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. Science 328, 1408–1411.
9. Shamay-Tsoory, S.G., and Abu-Akel, A. (2016). The Social Salience Hypothesis of Oxytocin. Biol. Psychiatry 79, 194–202.
10. Quintana, D.S., and Guastella, A.J. (2020). An Allostatic Theory of Oxytocin. Trends Cogn. Sci. 24, 515–528.

CHAPTER 4

Somatostatin interneurons in the prefrontal cortex control affective state discrimination in mice

Diego Scheggia[1,5], Francesca Managò[1,5], **Federica Maltese**[1], Stefania Bruni[1], Marco Nigro[1], Daniel Dautan[1], Patrick Latuske[1,2], Gabriella Contarini[1], Marta Gomez-Gonzalo[3], Linda Maria Requie[3], Valentina Ferretti[1], Giulia Castellani[1], Daniele Mauro[1], Alessandra Bonavia[1], Giorgio Carmignoto[3], Ofer Yizhar[4], Francesco Papaleo[1,6]

[1]*Genetics of Cognition laboratory, Neuroscience area, Istituto Italiano di Tecnologia, Genova, Italy.*
[2]*Central Nervous System Diseases Research, Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany.*
[3]*Neuroscience Institute, Italian National Research Council (CNR), Padua, Italy.*
[4]*Department of Neurobiology, Weizmann Institute of Science, Rehovot, Israel.*
[5]These authors contributed equally to this work.

Published in *Nature Neuroscience* (2020) 23, pages 47–60.

Abstract

The prefrontal cortex (PFC) is implicated in processing of the affective state of others through non-verbal communication. This social cognitive function is thought to rely on an intact cortical neuronal excitatory and inhibitory balance. Here combining in vivo electrophysiology with a behavioral task for affective state discrimination in mice, we show a differential activation of medial PFC (mPFC) neurons during social exploration that depends on the affective state of the conspecific. Optogenetic manipulations revealed a double dissociation between the role of interneurons in social cognition. Specifically, inhibition of mPFC somatostatin (SOM+), but not of parvalbumin (PV+) interneurons, abolishes affective state discrimination. Accordingly, synchronized activation of mPFC SOM+ interneurons selectively induces social discrimination. As visualized by in vivo singlecell microendoscopic Ca^{2+} imaging, an increased synchronous activity of mPFC SOM+ interneurons, guiding inhibition of pyramidal neurons, is associated with affective state discrimination. Our findings provide new insights into the neurobiological mechanisms of affective state discrimination.

Introduction

Understanding the emotions of others by perception of facial and body expressions is an ability that crucially affects everyday life (1). Impairments in recognition of emotions are common in many neurodegenerative, neuropsychiatric, and neurodevelopmental disorders (2,3). For example, emotion recognition deficits are core features of autism spectrum disorders (ASDs) (4) and are strongly evident in schizophrenia (5). These social cognitive impairments might have a more deleterious impact on daily functioning than non-social cognitive deficits (6). Moreover, the management of these social cognitive deficits remains inadequate, highlighting the need for a deeper understanding of the mechanisms underlying the ability to recognize affective state in others.

The 'social brain', identified by human neuroimaging studies, refers to a network that controls social cognitive processes in which limbic and frontal regions may play a key role (7,8). In particular, the top-down control of social cognitive functions may be orchestrated by the prefrontal cortex (PFC) over the limbic system (9,10). Indeed, damage to the medial PFC (mPFC) is associated with impaired recognition of emotions (11,12). Thus, the PFC is an attractive brain region for the study of neurobiological mechanisms underlying such behavior (12). However, our understanding of the PFC neural circuits underpinning the recognition of emotion remains incomplete, mainly owing to the resolution level of manipulations allowed in humans and the lack of translational models.

Accumulating evidence indicates that the balance of neuronal excitation and inhibition governs cortical functions (13,14). Perturbations in this balance are commonly invoked as a possible final shared pathway in the etiology of ASD and schizophrenia15. For example, in humans, reduced density of interneurons (16,17) and altered GABAergic signaling (18,19) are common findings in the brains of patients with ASD. In line with these findings in humans, disruption of the excitatory and inhibitory balance in the mPFC of mice led to social exploration deficits and sociability impairments (20). Moreover, other rodent studies implicated the PFC in different social functions, such as social interaction (20,21), vicarious freezing (22,23), social hierarchy (24,25), and affiliative behavior (26). However, the involvement of PFC circuits and related excitatory and inhibitory balance in the ability to detect and process the expression of affective state in others remains uncertain.

Here, we proposed that inhibitory neuron subpopulations within the mPFC could differentially contribute to the processing of affective state discrimination. To explore mPFC circuits involved in affective state discrimination in a cell type-specific manner, we used a rodent task approximating features of the human 'emotion recognition task'(4). In particular, the task is designed to study the ability of mice to discriminate conspecifics based on their affective state. Using in vivo electrophysiology and microendoscope imaging, we demonstrated that the mPFC differentially responds to conspecifics in an altered affective state. By optogenetic perturbation experiments and microendoscope imaging we then dissected the involvement of different mPFC neuronal subpopulations in affective state discrimination. Together, our data support a model in which, in the mPFC, synchronized activation of somatostatin (SOM+), but not of parvalbumin (PV+) interneurons or pyramidal neurons, is a primary substrate for the expression of affective state discrimination.

Results

Mice can discriminate conspecifics based on altered affective states

In the affective state discrimination task (ADT), we tested whether a mouse ('observer') could distinguish between two unfamiliar conspecifics ('demonstrators') based on their affective state. The observer was placed in front of two demonstrators, matched for sex and age, inside inverted wire cups, divided by a black wall (Fig. 1a). To induce changes in the affective state, one demonstrator ('relief ') was exposed to a procedure consisting in 23 h of water deprivation, followed by 60 min of water restoration before the test. The other demonstrator was a naive mouse with ad libitum water access ('neutral', Fig. 1a). The relief procedure led to conditioned place preference and reduced corticosterone levels27, suggesting the induction of a positive affective state. After habituation to the testing arena with empty cups, we presented the relieved and neutral unfamiliar demonstrators to a naive observer; as a result, observers displayed increased sniffing towards the relieved conspecific (Fig. 1b and Extended Data Fig. 1a) than towards the neutral demonstrator and spent more time in the related zone (Fig. 1c). This behavior was evident during the first 2 min of observation (Fig. 1b,c and Extended Data Fig. 1a). When demonstrators were familiar cagemates, the observers showed more persistent discrimination that lasted up to 4 min (Extended Data Fig. 2a–f). Observers made a similar number of visits to each demonstrator

(Extended Data Fig. 1b), but during the first 2 min, made longer visits to the relieved mouse (Extended Data Fig. 1c). Furthermore, the latency to make the first visit towards the relieved demonstrator was consistently shorter than towards the neutral demonstrator (Fig. 1d). We did not detect other differences in the behaviors of the observers, such as grooming and rearing (Extended Data Fig. 1f). No differences in affective state discrimination were observed between sex-matched females and males (Fig. 1k and Extended Data Fig. 1e). Neutral and relieved demonstrators displayed no observable behavioral differences (Supplementary Fig. 1a).

Next, we tested whether this discrimination could be extended to a different, negative, affective state. We used the same ADT setting, but the observers were presented with a demonstrator that underwent a mild stress protocol, consisting of 15 min of acute restraint before the beginning of the ADT, and a neutral demonstrator (Fig. 1e). We observed increased exploration towards the stressed demonstrator (Fig. 1f and Extended Data Fig. 1g) and more time spent in the related zone (Fig. 1g). Additionally, mice first entered the zone related to the stressed demonstrator (Fig. 1h), and made longer visits to this zone during the first 2 min of the test (Extended Data Fig. 1i). The total number of visits did not differ between the stressed and neutral demonstrators (Extended Data Fig. 1h). Also in this condition, when the observer and the demonstrators were cagemates, the discrimination was longer (Extended Data Fig. 2g–l). Moreover, observer mice had a closer approach towards the stressed demonstrators as the average distance of their head from the demonstrators during exploration was shorter for the stressed conspecific than for the neutral one (Extended Data Fig. 1j). The observer did not show differences in other behaviors, such as grooming and rearing (Extended Data Fig. 1l). No differences between male and female mice were evident in the discrimination of sex-matched stressed demonstrators (Fig. 1n and Extended Data Fig. 1k). During the test, stressed demonstrators displayed more grooming than neutral mice, but similar rearing, sniffing, biting, and freezing behaviors (Supplementary Fig. 1b). However, grooming was not correlated with ability of the observers to discriminate (Extended Data Fig. 1m). Further, no signs of transfer of behavioral responses between observer and demonstrator, and no escape behaviors or altered corticosterone levels in the observer mice, were present throughout the test session in both relief and stress conditions (Extended Data Fig. 1f,l,n). Overall, these results showed that mice can similarly discriminate others based on both positive and negative affective states.

Affective state discrimination is a stable trait distinct from sociability

To evaluate the reliability of affective state discrimination, we replicated the ADT several times in naive mice with different experimenters and in later optogenetics and electrophysiological experiments, replicating our initial findings in a large group of animals (n = 96 'relief ', n = 93 'stress'). Data from ADTs conducted in naïve animals and in mice implanted with electrodes and under 'light off' conditions were pooled and are shown as percent exploration towards the manipulated mouse (relief and stress, Fig. 1i,l). The affective state discrimination was a reliably observable behavior, with only 12% of tested mice not discriminating between a mouse in an altered or a neutral affective state (relief, 10 of 96, Fig. 1j; stress, 13 of 93; Fig. 1m). The scores of exploration towards the relieved and stressed demonstrators fit a normal distribution (D'Agostino and Pearson normality test, stressed: n = 93, K_2 = 1.54, P = 0.46; relieved: n = 96, K_2 = 1.83, P = 0.39). Moreover, affective state discrimination abilities were stable; when re-exposed to the same (Extended Data Fig. 1o) or to a different paradigm (Extended Data Fig. 1p), observers showed similar behavior. These data showed that affective state discrimination is a stable trait in mice.

Notably, if observer mice were tested in a one-on-one free social interaction setting in a novel environment with unfamiliar neutral, relief, or stress demonstrators (Fig. 1o), they spent a similar amount of time in social interaction with mice in altered or neutral affective states. Further, we also observed a classic habituation curve that was not influenced by the affective state of the demonstrator (Fig. 1o). This finding suggests that the discrimination revealed by the ADT is not due to a generalized increase in social exploration (an index of sociability), but is rather a specific measure of discrimination of affective states.

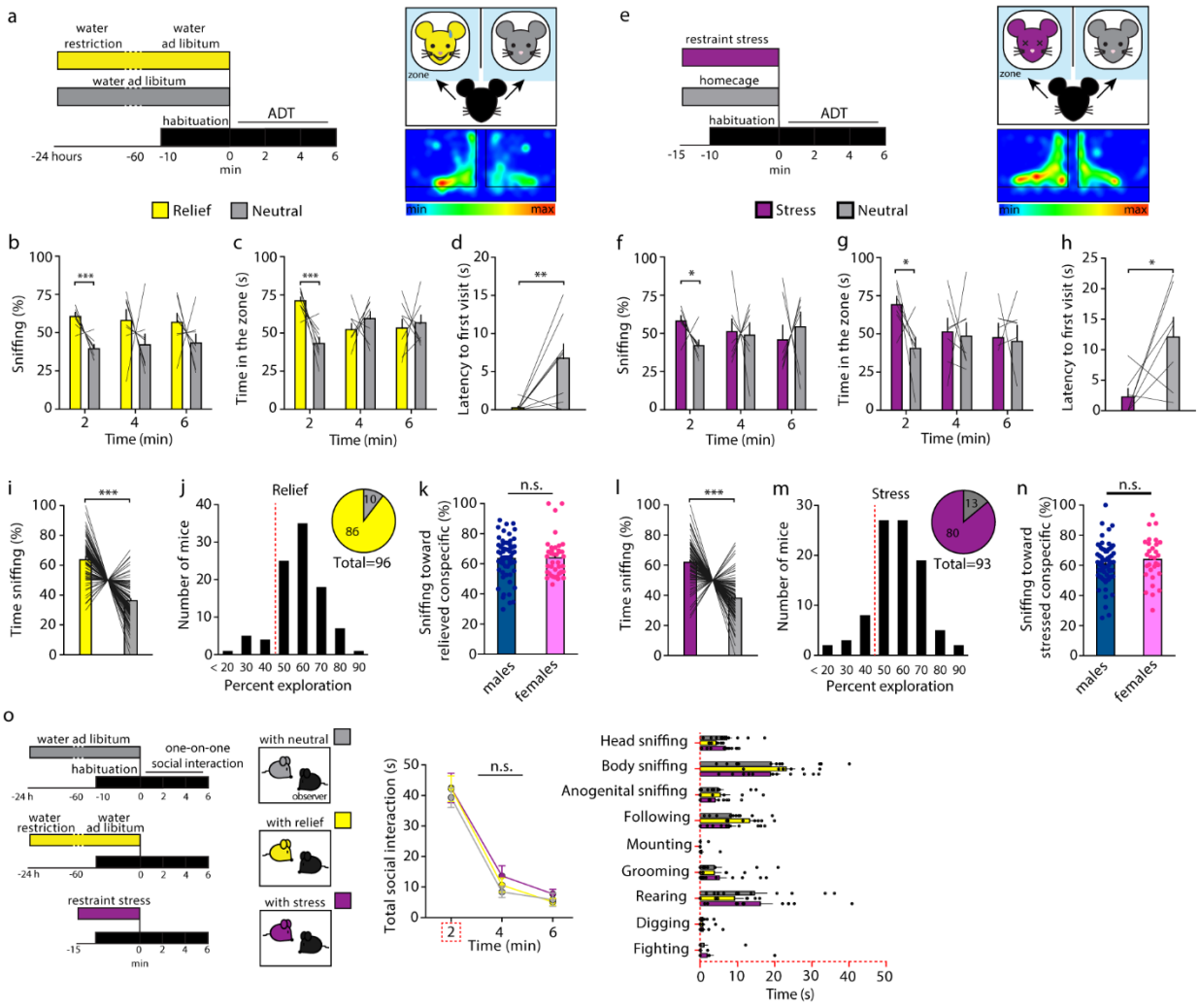


Fig. 1 | Mice can discriminate conspecifics based on their affective state. a–d, Data derived from n = 8 male mice. **a**, Left, experimental design of the ADT. After 23 h of water deprivation, one demonstrator was given access to water for 1 h before the test (relief, yellow), while the other demonstrator had ad libitum water access (neutral, gray). Right, schematic illustration of the testing arena with stressed and neutral demonstrators (counterbalanced left and right across experiments) and graphical representation of the amount of time observers spent in different parts of the apparatus (with blue as the shortest and red as the longest time). **b**,**c**, Increased exploration behavior (sniffing) towards the relieved (two-tailed multiple t-test, Bonferroni correction, 2 min: t = 6.22, d.f. = 14, P < 0.0001) (b) and increased time spent in the zone with the relieved demonstrator compared to that spent with the neutral demonstrator during the first 2 min of testing (two-tailed multiple t-test, Bonferroni correction, 2 min: t = 5.68, d.f. = 14, P = 0.0001) (**c**). **d**, Shorter latency to visit the relieved mice than to visit the neutral demonstrator (two-tailed paired t-test: t = 3.16, d.f. = 7, P = 0.015). **e–h**, Data derived from n = 7 mice. e, In the stress protocol, one demonstrator (stress, purple) was subjected to the restraint stress test for 15 min immediately before the beginning of the ADT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. Increased sniffing of the stressed demonstrator (two-tailed multiple t-test, Bonferroni correction, 2 min: t = 3.22, d.f. = 12, P = 0.021) (**f**) and time spent in the zone with the stressed compared with the neutral demonstrator (two-tailed multiple t-test, Bonferroni correction, 2 min: t = 3.23, d.f. = 12, P = 0.021) (**g**). **h**, The latency to the first visit was significantly lower for the stressed demonstrator than for the neutral demonstrator (two-tailed paired t-test: t = 2.43, d.f. = 6, P = 0.050). **i**, The ADT was replicated several times, and the percentage of explorations of demonstrators was pooled (n = 98 mice for relief manipulation and n = 93 for stress). Observers explored the relieved more than the neutral demonstrators during the first 2 min of the ADT (two-tailed paired t-test: t = 10.12, d.f. = 97, P < 0.0001). **j**,**k**, Exploration of the relieved demonstrator was higher than chance in a large number of mice (86 of 96, one-sample t-test against chance, defined as 50%: t = 10.12, d.f. = 97, P < 0.0001) (**j**), and did not change depending on gender (59 male, 39 female; two-tailed unpaired t-test: t = 0.10, d.f. =

96, P = 0.91) (**k**). **l**, More exploration of the stressed than of the neutral demonstrator in several replications of the ADT (two-tailed paired t-test: t = 8.22, d.f. = 92, P < 0.0001). **m**,**n**, Exploration of the stressed demonstrator was higher than chance in a large number of mice (80 of 93, one-sample t-test against chance, defined as 50%: t = 8.22, d.f. = 92, P < 0.0001) (**m**) and did not change depending on gender (63 male, 30 female; two-tailed unpaired t-test: t = 0.92, d.f. = 91, P = 0.357) (**n**). **o**, Mice did not show any difference in sociability when presented with a neutral, relieved, or stressed mouse in a standard free social interaction test (two-way ANOVA, time × affective state (neutral, relief, stress), $F_{(4,56)} = 0.21$, P = 0.930, n = 12 mice exposed to the three conditions). Bar and line graphs show mean ± s.e.m. *P < 0.05. **P < 0.005. ***P < 0.0005. n.s., not significant.

Enhanced mPFC neuronal activity during exploration of a conspecific in an altered affective state

To investigate the possible recruitment of the mouse mPFC in affective state discrimination, we implanted observers with a linear multi-electrode array in the mPFC and carried out chronic electrophysiological recordings during the ADT (Fig. 2 and Extended Data Fig. 3a,d). In the relief versus neutral condition, we recorded 57 well-isolated units, with the majority from the prelimbic cortex. Based on spike waveform features, such as spike width and firing frequency (28), we classified these units into narrow-spiking putative inhibitory interneurons and wide-spiking putative pyramidal neurons (Extended Data Fig. 3b). Among the recorded units, 39 out of 57 cells discharged during the direct exploration of the demonstrators (response to social exploration, Fig. 2a). The remaining cells (n = 18) showed no variations in their firing rate either before or during the social exploration and did not display stronger activation for one of the two affective states (‘no response’, Fig. 2a). The majority of responsive units discharged before and after the beginning of the exploration (85%), a smaller group only after the interaction started (13%), and a few units (2%) were activated only before the mouse started to explore one of the two demonstrators (Fig. 2a). Moreover, 79% of responsive neurons displayed sustained activity throughout the duration of social exploration (Fig. 2a). About 79% of the narrow-spiking cells had stronger discharge when the mouse explored the relieved conspecific than when the mouse explored the neutral conspecific throughout the test, whereas only 33% of the wide-spiking cells responded preferentially to the relieved demonstrator (Fig. 2b). In particular, the population of narrow-spiking neurons displayed a higher firing rate when the observer explored the relieved demonstrator than when the observer explored the neutral demonstrator until the end of the social exploration, when the firing drastically decreased (Fig. 2c). The increase in narrow-spiking neuronal activity associated with the exploration of the manipulated demonstrator began before the start of the interaction and was sustained during the whole exploration (Fig. 2c). By contrast, the population of wide-spiking cells did not show a preference for the relieved mouse in the first

2 min of the ADT (Fig. 2d). Moreover, a direct comparison indicated a higher preference of discharge for the relieved over the neutral demonstrator in the narrow-spiking neurons than in the wide-spiking neurons (Extended Data Fig. 3c).

A similar activity pattern was displayed by mPFC neurons in observers recorded during the stress ADT (Fig. 2e). A large number of these units were classified as narrow-spiking (n = 52), and the remaining (n = 31) were classified as wide-spiking neurons (Extended Data Fig. 3e). Out of the 83 recorded units, 66 responded to social exploration (Fig. 2e). The majority (77%) discharged just before and during the exploration, some of them discharged only after the beginning of the interaction (20%), and a few units (3%) increased their activity only before the mouse started to explore one of the two demonstrators (Fig. 2e). A large number of responsive units showed a sustained activity (83%) rather than a transient response (Fig. 2e). Moreover, as observed for the relief ADT, a higher number of narrow-spiking cells (84%) were strongly engaged when the mouse explored the stressed rather than the neutral conspecific, whereas only 29% of the wide-spiking cells responded more to the manipulated demonstrator (Fig. 2f). At a population level, the firing rate of narrow-spiking neurons was increased (Fig. 2g), whereas that of wide-spiking cells was decreased (Fig. 2h), during the ADT compared to the habituation without demonstrators. Additionally, for narrow-spiking neurons the increased firing rate when exploring a stressed demonstrator compared to the neutral was much more evident than in wide-spiking cells (Fig. 2g,h and Extended Data Fig. 3f).

Consistent with human functional magnetic resonance imaging (fMRI) (29) and brain lesion studies (11), these results show that the mouse mPFC is engaged during affective state discrimination with the majority of cells responding to the expression of altered states. Moreover, we found that, in the mPFC, narrow-spiking putative inhibitory interneurons show a higher engagement and preference for conspecifics in either a positive or negative affective state than wide-spiking neurons.

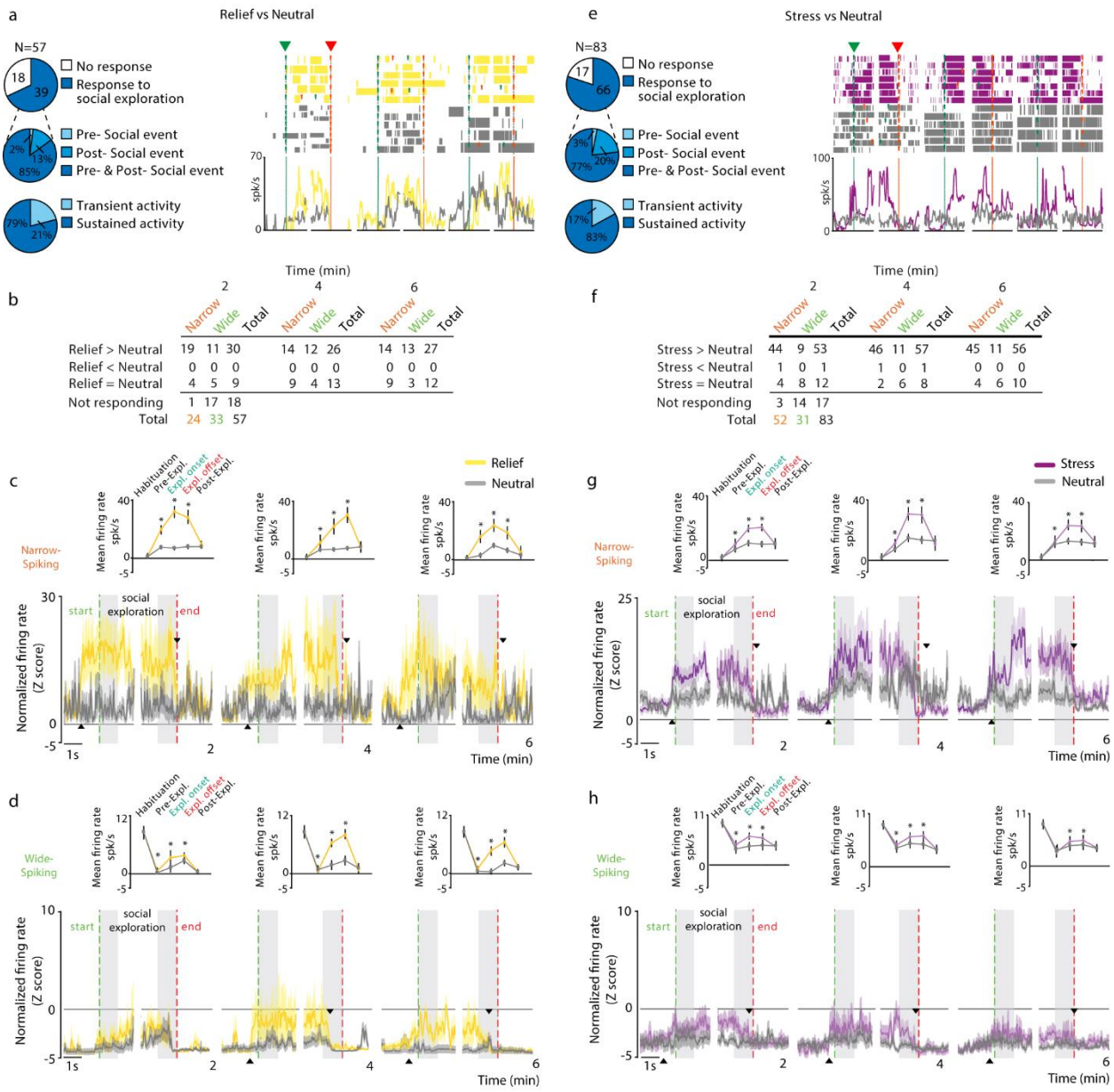


Fig. 2 | Enhanced neuronal activity during exploration of conspecifics in an altered affective state. a,**e**, Left: characteristics of responsive and nonresponsive cells to social exploration. Right: examples of a responsive cell recorded during the exploration of (**a**) a relieved (yellow) versus a neutral (gray) demonstrator and (**e**) a stressed (purple) versus a neutral demonstrator. Rasters (top) and polylines (bottom) aligned on the beginning (green dotted line) and end (red dotted line) of each exploration of the relieved, stressed or neutral demonstrators in the same session. **b**, **f**, Number of responsive neurons (**b**, relief, $n = 57$ neurons from $n = 6$ male mice; **f**, stress, $n = 83$ neurons from $n = 9$ male mice) showing a preference for altered (relief or stress), neutral or both affective states, and non-responsive neurons, divided in the two subpopulations of narrow-spiking and wide-spiking cells. **c**, Top: population responses calculated as an average of the activity of narrow-spiking neurons during habituation, pre-exploration, exploration onset, exploration offset and post-exploration periods towards a relieved or neutral demonstrator (mean ± s.e.m.). *$P < 0.05$ versus exploration of the neutral mouse. Bottom: population firing rate of narrow-spiking neurons normalized to the habituation during the pre-social exploration, start social exploration (green dashed line), social exploration, end social exploration (red dashed line) and post-social exploration epochs. Increased firing rate during exploration of the relieved demonstrator (start social exploration and social exploration) compared to the pre-social exploration epoch in all 6 min tests ($P < 0.0001$) and compared to the firing rate when exploring the neutral demonstrator ($P <$

0.0001, 3 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(2,46)} = 107.79$, $P < 0.0001$; 4 min, $F_{(2,46)} = 23.25$, $P < 0.0001$; 6 min, $F_{(2,46)} = 17.5$, $P < 0.0001$). Drop of firing rate after the end of social exploration of a relieved demonstrator (post-social exploration epoch, 2 min, $P = 0.000441$; 4 min, $P = 0.003244$; 6 min, $P = 0.000176$), which was still higher than the firing rate during exploration of a neutral demonstrator in all 6 min tests (2 min, $P = 0.000002$; 4 min, $P = 0.000197$; 6 min, $P = 0.000003$; 2 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(1,23)} = 16.95$, $P = 0.000420$; 4 min, $F_{(1,23)} = 19.94$, $P = 0.000176$; 6 min, $F_{(1,23)} = 7.51$, $P = 0.011666$). **d**, Top: same analyses as in **c**, but for wide-spiking neurons ($n = 33$). *$P < 0.05$ versus exploration of the neutral mouse. Bottom: population firing rate of wide-spiking neurons, shown as in **c**. Increased firing rate during start social exploration and social exploration of the relieved demonstrator compared to the pre-social exploration epoch (2 min, $P = 0.005246$; 4 min, $P = 0.004034$; 6 min, $P = 0{,}000019$) and compared to the firing rate when exploring the neutral demonstrator in all 6 min tests (2 min, $P = 0.005246$; 4 min, $P = 0.019708$; 6 min, $P = 0{,}000417$; 3 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(2,64)} = 7.45$, $P = 0.001234$; 4 min, $F_{(2,64)} = 11.27$, $P = 0.000064$; 6 min, $F_{(2,64)} = 17.46$, $P = 0.000001$). Drop of activity after the end of the exploration epoch (post-social exploration epoch, 2 min, $P = 0.020214$; 4 min, $P = 0.002099$; 6 min, $P = 0{,}2087696$), which was still higher than the firing activity when exploring the neutral demonstrator in all 6 min tests (2 min, $P = 0.028718$; 4 min, $P = 0.040627$; 6 min, $P = 0.0408769$; 2 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(1,32)} = 6.54$, $P = 0.015444$; 4 min, $F_{(1,32)} = 5.03$, $P = 0.031810$; 6 min, $F_{(1,32)} = 7.98$, $P = 0.008054$). Colored shaded regions around each line represent 1 s.e.m. Gray shaded areas were used for statistical analysis of the population response. The black arrows indicate the time of onset (upwards arrow) and end (downwards arrow) of significant separation of the activity between the two compared conditions. **g**, Top: population responses calculated as an average of the activity of narrow-spiking neurons ($n = 52$) during different phases of the ADT (mean ± s.e.m.). *$P < 0.05$ versus exploration of the neutral mouse. Bottom: population firing rate of narrow-spiking neurons normalized to the habituation. Increased firing rate at the beginning of exploration of the stressed demonstrator compared to the pre-social interaction epoch (2 min, $P = 0.00$; 4 min, $P = 0.000092$; 6 min, $P < 0{,}0001$) and compared to the firing rate when exploring the neutral demonstrator in all 6 min test (2 min, $P = 0.020214$; 4 min, $P = 0.000132$; 6 min, $P = 0{,}010969$; 3 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(2,\ 102)} = 38.46$, $P < 0.0001$; 4 min, $F_{(2,\ 102)} = 32.85$, $P < 0.0001$; 6 min, $F_{(2,\ 102)} = 70.58$, $P < 0.0001$). Drop of activity on end social exploration (2 min, $P < 0.0001$; 4 min, $P = 0.000027$; 6 min, $P = 0{,}000819$), which was higher than firing activity when exploring the neutral demonstrator in all 6 min tests (2 min, $P < 0.0001$; 4 min, $P < 0.0001$; 6 min, $P < 0{,}0001$; 2 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(1,51)} = 82.6$, $P < 0.0001$; 4 min, $F_{(1,51)} = 19.49$, $P = 0.000053$; 6 min, $F_{(1,51)} = 29.45$, $P < 0.0001$). **h**, Top: same analyses as in **g**, but for wide-spiking cells ($n = 31$). Bottom: increased normalized population activity of wide-spiking cells before and during exploration of the stressed demonstrator compared to during the pre-social interaction epoch (2 min, $P = 0.330727$; 4 min, $P = 0.000521$; 6 min, $P = 0.012764$) and compared to during exploration of the neutral demonstrator (2 min, $P = 0.0475371$; 4 min, $P = 0.008066$; 6 min, $P = 0{,}03386$; 3 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(2,60)} = 11.49$, $P = 0.00006$; 4 min, $F_{(2,60)} = 23.62$, $P = 0.00006$; 6 min, $F_{(2,60)} = 24.40$, $P < 0.0001$). The increased firing rate during the social exploration epoch with a stressed conspecific dropped after the end of the exploration in all 6 min tests (2 min, $P = 0.000699$; 4 min, $P = 0.012473$; 6 min, $P = 0.002589$) and was higher than the firing activity during the exploration of the neutral demonstrator (2 min, $P = 0.000106$; 4 min, $P = 0.010444$; 6 min, $P = 0.000004$; 2 × 2 repeated-measures ANOVA, exploration epoch and affective state, 2 min, $F_{(1,30)} = 19.55$, $P = 0.000119$; 4 min, $F_{(1,30)} = 33.94$, $P < 0.0001$; 6 min, $F_{(1,30)} = 7.83.01$, $P = 0.008855$).

Enhancement of mPFC neuronal activity is specifically linked to affective state discrimination

To verify whether the observed increase in neuronal activity was specifically related to the exploration of conspecifics in different affective states, we used two neutral demonstrators in the ADT (Extended Data Fig. 3g). Observer mice similarly explored both demonstrators, showing no observable bias between the two (Extended Data Fig. 3g). Similar to what we observed with demonstrators with altered affective states, a large putative

interneurons (*n* = 55, Extended Data Fig. 3h), whereas fewer were classified into wide-spiking neurons (*n* = 27). Moreover, population responses show that narrow-spiking neurons increased, whereas wide-spiking neurons decreased, their activity in the presence of the demonstrators compared to habituation without demonstrators (Extended Data Fig. 3i,j). However, no differences in neuronal activity were evident during exploration of the two neutral demonstrators (Extended Data Fig. 3i,j). This finding supports the link between the increased mPFC neuronal activity and the exploration of a mouse with altered affective states.

Next, we investigated which sensory modality might trigger affective state discrimination and its related mPFC neuronal activation. No significant ultrasonic vocalization (USV) calls were recorded in the ADT (Supplementary Fig. 2a), and no differences in the emission of USVs were evident when demonstrators were tested separately (Supplementary Fig. 2a). This result indicates marginal involvement of auditory cues in mouse affective state discrimination, consistent with previous literature showing that adult mice do not engage USV with conspecifics of the same sex (30). Observer mice tested in complete darkness showed the same increase in exploration towards both relieved (Supplementary Fig. 2b) and stressed demonstrators (Supplementary Fig. 2c). However, showing the videos of mice in neutral, relief, or stress states to the observers was sufficient to detect a discrimination in the relief ADT (Supplementary Fig. 2d,e). Therefore, visual cues can convey affective information, but they might be dispensable.

We then tested odor cues by collecting separately odors from neutral, relief, and stress demonstrators by gently brushing a cotton ball all over the body of the mice (especially including the nose, body, and anogenital parts). Then, odors were placed under the inverted wired cup instead of the demonstrators. In contrast to the results obtained in the ADT (Fig. 1), observers showed marked avoidance of the odor of a stressed demonstrator (Supplementary Fig. 2g), but preferred to explore the odor from the relieved demonstrators (Supplementary Fig. 2f). The neuronal activity patterns in the mPFC during exploration of odors of manipulated and neutral demonstrators was similar to those in the presence of the demonstrator mice (Supplementary Fig. 3a–c, f–h). Again, compared to wide-spiking neurons, the narrow-spiking neurons had a higher discharge and higher degree of preference for the odor from demonstrators in altered affective states than demonstrators in a neutral one (Supplementary Fig. 3d,i). However, the degree of firing preference for odors associated with an altered affective state was smaller than when the actual demonstrators were present (Supplementary Fig. 3e, j).

Overall, these results indicate that the increased firing rate in mPFC cells (especially narrow-spiking cells) are associated with the exploration of altered affective states.

Photoinhibition of mPFC PV+ interneurons does not affect affective state discrimination

Our electrophysiological findings suggested a major engagement of putative interneurons in the mPFC. The most abundant subpopulation of interneurons in the mPFC is represented by parvalbumin-positive (PV+) cells (13). To investigate whether silencing mPFC PV+ cells might affect the discrimination of affective states, we bilaterally injected an adeno-associated virus (AAV) carrying a Cre-dependent halorhodopsin (AAV-EF1a-DIO-eNpHR3.0- eYFP) into the mPFC of PV-Cre mice and implanted optic fibers terminating dorsally to this area (Fig. 3a and Extended Data Fig. 4a). We silenced mPFC PV+ cell activity with continuous green light using a laser stimulation protocol similar to that used in previous in vivo studies (31), during the first 2 min of the ADT (Fig. 3b,e). This targeted the time window in which we observed increased exploration of the demonstrators with altered affective states. Mice were tested on consecutive weeks, with the protocol (relief and stress) and treatment (light off and light on) counterbalanced. Photoinhibition of PV+ cells reduced the general investigation of demonstrators by observers (Fig. 3c,f), which is an index of sociability. However, photoinhibition of PV+ cells did not modify the affective state discrimination of the observers in either the relief (Fig. 3d) or stress (Fig. 3g) paradigm. Mice also first visited the zones related to the relief (Extended Data Fig. 4c) or stress (Extended Data 4f) demonstrators, and the inhibition of PV+ cells did not modify this behavior (Extended Data Fig. 4c,f) or the total number of visits to each demonstrator (Extended Data Fig. 4d,g). Finally, photoinhibition of PV+ cells did not induce any gross motor deficits (Extended Data Fig. 4h,i). These results suggest that PV+ cells in the mPFC might be involved in sociability, but not in affective state discrimination.

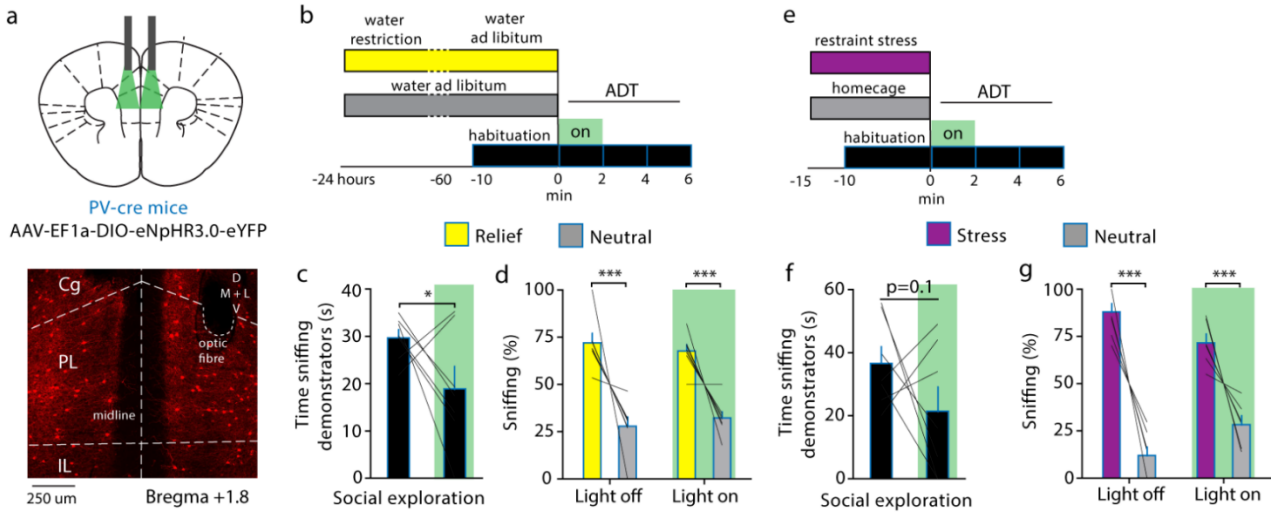


Fig. 3 | Photoinhibition of PV+ interneurons does not affect affective state discrimination. a, Top: PV-Cre male mice were bilaterally injected in the mPFC with AAV-EF1a-DIO-eNpHR3.0-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. Bottom: representative image of a coronal mPFC section (Cg, cingulate; PL, prelimbic area; IL, infralimbic area; D, dorsal; V, ventral; M, medial; L, lateral). Findings were replicated in two independent experiments with similar results. Data were derived from $n = 7$ mice in each condition (relief and stress). **b**, Mice were tested in the ADT with one relieved and one neutral demonstrator. Photoinhibition was performed for 2 min from the beginning of the test using continuous green light ($\lambda = 532$ nm). **c**, Reduced social investigation during optical inhibition of PV+ cells (two-tailed unpaired *t*-test: $t = 2.12$, d.f. = 12, $P = 0.029$). **d**, Increased sniffing (expressed as a percentage) of the relieved demonstrator compared to the neutral demonstrator during the first 2 min of testing without light delivery (two-way repeated-measures ANOVA, affective state (relief, neutral): $F_{(1,12)} = 203.8$, $P < 0.0001$). **e**, Mice were tested in the ADT with one stressed and one neutral demonstrator, and photoinhibition was performed during the first 2 min. **f**, Tendency for reduced social investigation during optical inhibition of PV+ cells (light off, 36.6 ± 5.30; light on, 21.44 ± 7.85, two-tailed unpaired *t*-test: $t = 1.6$, d.f. = 12, $P = 0.1$). **g**, Increased exploration of the stressed demonstrator compared to the neutral demonstrator in both light-off and light-on conditions (two-way repeated-measures ANOVA, affective state × light (off, on): $F_{(1,12)} = 7.83$, $P = 0.016$). Bar and line graphs show mean ± s.e.m. $^{*}P < 0.05$. $^{***}P < 0.0005$.

Photoinhibition of mPFC SOM+ interneurons abolishes affective state discrimination

SOM+ cells constitute another major subtype of local GABAergic interneurons in the cerebral cortex (13). To investigate their possible involvement, we bilaterally injected AAV-EF1a-DIO-eNpHR3.0-eYFP into the mPFC of SOM-Cre mice and implanted chronic optic fibers terminating dorsally to this area (Fig. 4a and Extended Data Fig. 5a). Optical inhibition of SOM+ cells, using the same light delivery protocol (Fig. 4b,f) as that adopted for PV+ cells, abolished affective state discrimination (Fig. 4d,h), producing no effects on social exploration (Fig. 4c,g). This effect was temporary and reversible. Indeed, after cessation of photoinhibition, we again observed an increased exploration of the relieved (Fig. 4e) and stressed (Fig. 4i) demonstrators. When tested in the light off condition, these same mice showed affective state discrimination as observed in naive mice (Fig. 4d,h). Moreover, testing these animals with the same light delivery

protocol, but with a neutral demonstrator and an object or with a familiar and a novel demonstrator, as commonly used in the classic three-chambers test (32), did not influence sociability or social novelty discriminations (Fig. 4j,k). Furthermore, optogenetic inhibition of SOM+ cells did not modify discrimination of social odors (Extended Data Fig. 5b,c), or discrimination between two different rewarding palatable food stimuli (Fig. 4l). Ex vivo recording experiments confirmed continuous green light effectively and similarly photoinhibited PV+ and SOM+ interneurons (Extended Data Fig. 6a–c). Moreover, in vivo electrophysiological recordings combined with optogenetics showed effective inhibition of neurons for the entire light stimulation period (Extended Data Fig. 6d). Overall, these results indicate that SOM+ interneurons are specifically required for the ability to discriminate conspecifics based on their affective state.

Next, we aimed to control the photoinhibition of SOM+ cells using a closed-loop system such that the presence of the observer in the zone related to the relieved (Fig. 4m and Extended Data Fig. 5f) or stressed (Fig. 4q and Extended Data Fig. 5j) demonstrators triggered the optical inhibition of these cells. Similar to naive mice (Fig. 1d,h), before stimulation, observers first visited the relieved (Fig. 4n) and stressed (Fig. 4r) demonstrators. In line with the above experiments, time-locked photoinhibition of SOM+ interneurons when exploring the demonstrators in altered affective state abolished the discrimination (relief, Fig. 4o,p and Extended Data Fig. 5g; stress, Fig. 4s,t and Extended Data Fig. 5k), whereas the number of visits to each demonstrator was not affected (Extended Data Fig. 5h,l). To rule out the possibility that SOM+ inhibition was aversive per se, we tested these same mice with two neutral demonstrators, pairing light delivery with the exploration of only one demonstrator. This manipulation did not induce any bias in the exploration (Extended Data Fig. 5n). Furthermore, SOM+ photoinhibition did not induce motor deficits (Extended Data Fig. 5i,m) and light delivery in mice without viral expression did not induce any place avoidance or motor abnormalities (Supplementary Fig. 4).

Overall, these experiments indicate that SOM+ interneurons in the mPFC are necessary for affective state discrimination in mice.

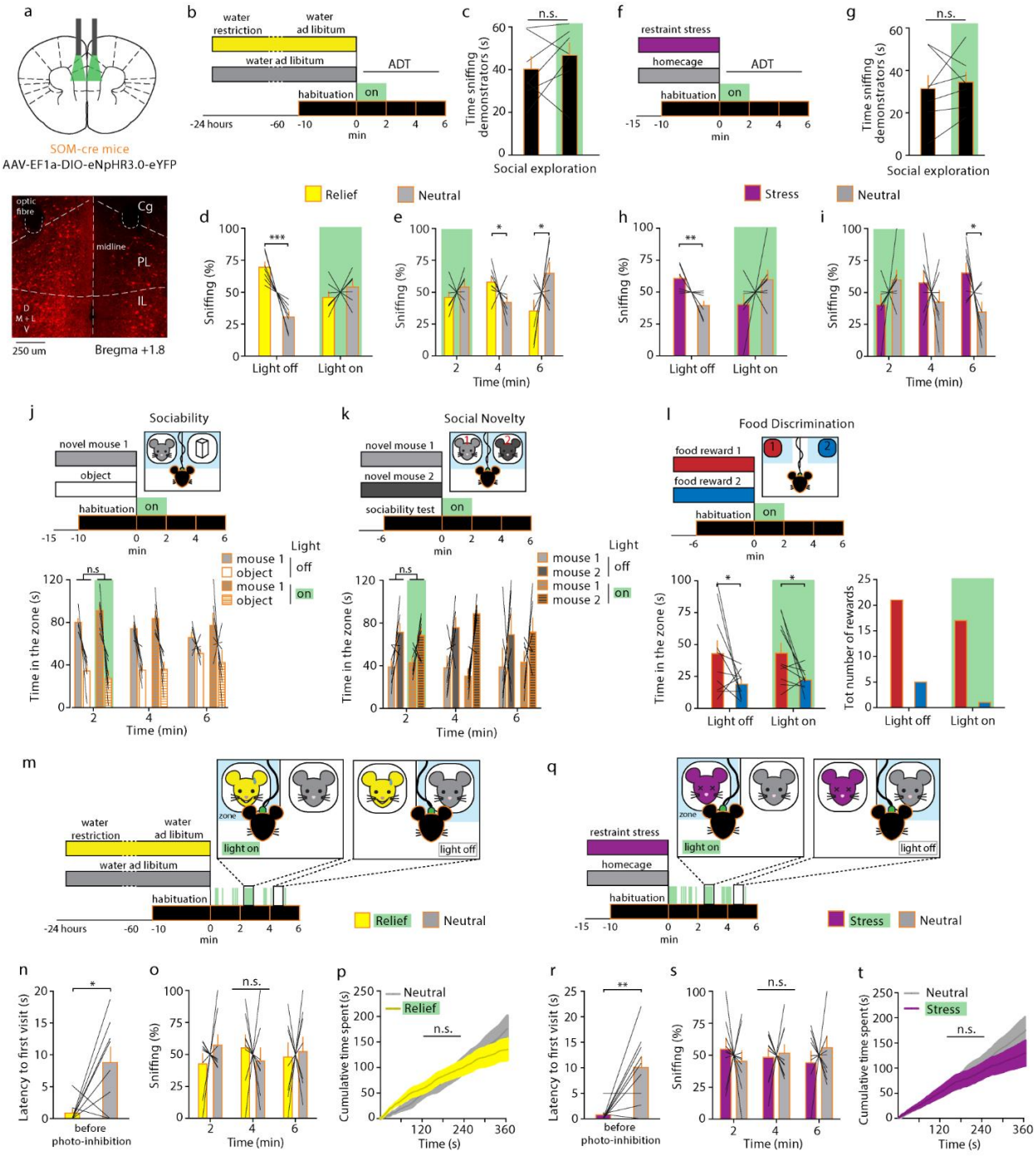


Fig. 4 | Photoinhibition of mPFC SOM+ interneurons abolishes affective state discrimination. a, Top: SOM-Cre male mice were injected in the mPFC with AAV-EF1a-DIO-eNpHR3.0-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. Bottom: representative image of coronal mPFC section. Findings were replicated in four independent experiments with similar results. **b**, Mice were tested in the ADT with one relieved and one neutral demonstrator. Photoinhibition was performed for 2 min from the beginning of the test using continuous green light (λ = 532 nm). Data derived from $n = 7$ mice. **c**, Total exploration of the demonstrators was not affected by photoinhibition of SOM+ cells (two-tailed unpaired *t*-test: $t = 0.78$, d.f. = 12, $P = 0.449$). **d**, Increased exploration of the relieved demonstrator compared to the neutral demonstrator during the first 2 min of testing, which was abolished in the light-on condition (two-way repeated-measures ANOVA, affective state × light (off, on): $F_{(1,12)} = 18.4$, $P = 0.001$). **e**, Immediately after photoinhibition of SOM+ neurons, mice explored the relieved demonstrator more than

the neutral demonstrator (two-tailed multiple *t*-test, Bonferroni correction, 4 min: $t = 2.34$, d.f. = 12, $P = 0.037$; 6 min: $t = 2.38$, d.f. = 12, $P = 0.034$). **f**, Mice were tested in the ADT with one stressed and one neutral demonstrator and photoinhibition was performed for the first 2 min. Data derived from $n = 7$ mice. **g**, Total exploration towards the demonstrators was not affected by photoinhibition of SOM+ cells (two-tailed unpaired *t*-test: $t = 0.38$, d.f. = 12, $P = 0.707$). **h**, Increased sniffing of the stressed demonstrator compared to the neutral demonstrator during the first 2 min of testing (light off) was abolished by photoinhibition of SOM+ cells (light on, two-way repeated-measures ANOVA, affective state × light (off, on): $F_{(1,12)} = 8.01$, $P = 0.015$). **i**, Following cessation of photoinhibition, as in the light-off condition, mice explored the stressed demonstrator significantly more than the neutral demonstrator in the last 2 min of the ADT (two-tailed multiple *t*-test, Bonferroni correction, 2 min: $t = 2.78$, d.f. = 12, $P = 0.048$). **j**,**k**, Data derived from $n = 6$ mice. **j**, SOM+ cell photoinhibition did not modify sociability (preference for a novel mouse compared to a novel object, three-way ANOVA, mouse versus object, $F_{(1,60)} = 98.97$, $P < 0.0001$; light off versus light on, $F_{(1,60)} = 0.47$, $P = 0.493$) (**j**) or social novelty (preference for a novel mouse compared to a familiar one, three-way ANOVA, mouse 1 versus mouse 2, $F_{(1,60)} = 29.82$, $P < 0.0001$; light off versus light on, $F_{(1,60)} = 0.11$, $P = 0.735$) (**k**). **l**, Photoinhibition of SOM+ cells did not modify discrimination between two palatable food rewards (food reward 1, purified 5-TUL rodent diet pellets (red); food reward 2, sucrose 5-TUT pellets (blue); left, time in the zone, two-way repeated-measures ANOVA, food reward 1 versus food reward 2: $F_{(1,18)} = 9.45$, $P = 0.006$; light off versus light on, $F_{(1,18)} = 0.10$, $P = 0.748$; right, total number of eaten food rewards, 5-TUL preference in eight of ten mice, Fisher's exact test, light off, $P = 0.02$, and in nine of ten mice, light on, $P = 0.0011$; $n = 10$ mice). **m**, Exploration of the relieved demonstrator (observer in the zone, light blue) was paired to SOM+ cell photoinhibition throughout the test. Data derived from $n = 8$ mice. **n**, Mice firstly explored the relieved demonstrators (two-tailed paired *t*-test: $t = 2.69$, d.f. = 7, $P = 0.031$). **o**,**p**, Sniffing of relieved and neutral mice (two-tailed multiple *t*-test, Bonferroni correction, 2 min: $t = 1.27$, d.f. = 14, $P = 0.673$) (**o**) and time spent in the zones did not differ when photoinhibition was paired to exploration of the relieved demonstrator (two-way repeated-measures ANOVA, affective state (neutral, relief): $F_{(1,7)} = 0.00$, $P = 0.978$) (**p**). **q**, Exploration of the stressed demonstrator was paired to SOM+ cell photoinhibition throughout the test. Data derived from $n = 9$ mice. **r**, Mice firstly explored the stressed demonstrators (two-tailed paired *t*-test, $t = 3.61$, d.f. = 8, $P = 0.006$). **s**,**t**, No difference in sniffing (two-tailed multiple *t*-test, Bonferroni correction, 2 min, $t = 0.78$, d.f. = 16, $P = 0.755$) and time spent (two-way repeated-measures ANOVA, affective state (neutral, stress): $F_{(1,8)} = 0.28$, $P = 0.610$) with a stressed or neutral demonstrator during inhibition of SOM+ cells. Bar graphs show mean ± s.e.m. *$P < 0.05$. **$P < 0.005$. ***$P < 0.0005$. n.s., not significant.

Photostimulation of SOM+ interneurons in the mPFC guides social discrimination

Next, we asked whether stimulation of mPFC SOM+ interneurons could be sufficient to induce discrimination between conspecifics in a neutral state. To test this hypothesis, we bilaterally injected a Cre-dependent channelrhodopsin-2 vector (AAV-EF1a-DIO-ChR2-eYFP) into the mPFC of SOM-Cre mice and implanted chronic optic fibers terminating dorsal to the injection site (Fig. 5a and Supplementary Fig. 5a). We tested these observer mice in the ADT while presenting two neutral demonstrators. Exploration of one of the two neutral demonstrators was paired photostimulation of mPFC SOM+ cells in the observer (Fig. 5b and Supplementary Fig. 5b). This protocol induced increased exploration towards the demonstrator paired with light, measured as increased sniffing (Fig. 5c) and time spent in the related zone (Fig. 5d and Supplementary Fig. 5b), producing no effects on the number of visits in each zone (Fig. 5e). Photostimulation of SOM+ cells also reduced the distance between the observer head and the explored demonstrator, indicating that SOM+ activation induced a closer approach (Fig. 5f). Photostimulation of SOM+ neurons did not induce motor abnormalities (Supplementary Fig. 5c,d), and it was not rewarding

per se, as shown by real-time place preference (RTPP) and optical self-stimulation protocols (Fig. 5j–m and Supplementary Fig. 5e–g). When tested with two neutral demonstrators without photostimulation, these same observers did not show any discrimination (Fig. 5g,h). Moreover, to check whether photoinduced discrimination was specifically related to social cues, we repeated the same experiment using two identical objects instead of the demonstrator mice. The activation of mPFC SOM+ cells did not produce any difference in exploration of the objects (Fig. 5i). These findings indicate that the stimulation of SOM+ interneurons in the mPFC is sufficient to induce social discrimination. Altogether, our results demonstrate that, within the mPFC, SOM+ interneurons are a key modulator of affective state discrimination.

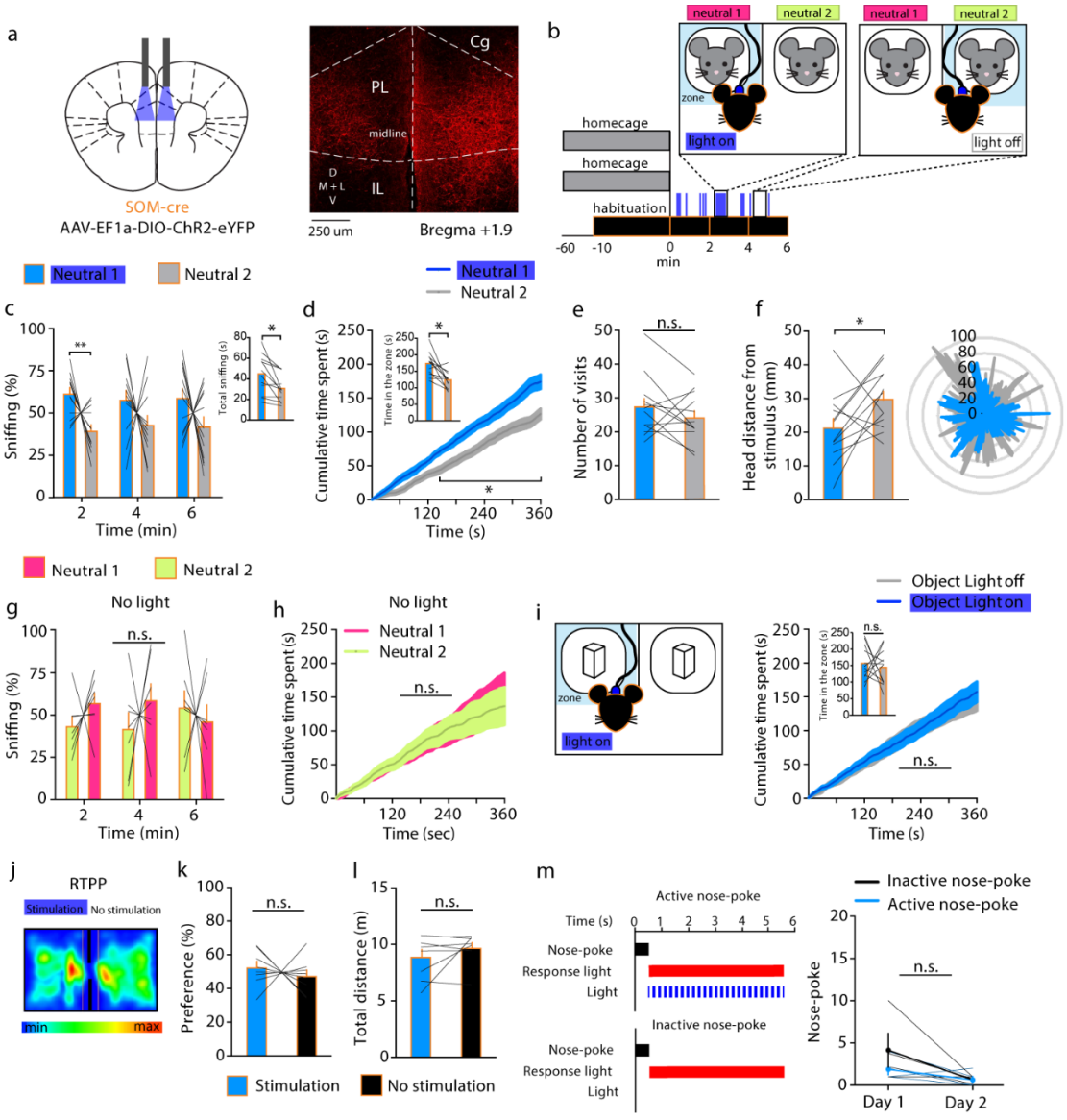


Fig. 5 | Photostimulation of SOM+ interneurons in the mPFC guides social discrimination. a, Left: SOM-Cre male mice were bilaterally injected in the mPFC with AAV-EF1a-DIO-ChR2-eYFP and implanted bilaterally with chronic optic fibers terminating dorsal to the injection area. Right: representative image of coronal mPFC section. Findings were replicated in three independent experiments with similar results. Data derived from $n = 12$ mice. **b**, Mice were tested in the ADT with two naive, non-manipulated neutral demonstrators. Photostimulation with blue light was paired to exploration of one of the two demonstrators (counterbalanced, left or right, across observers). **c**,**d**, Increased sniffing (two-tailed multiple *t*-test, Bonferroni correction, 2 min, $t = 3.95$, d.f. = 22, $P = 0.002$; inset, total time sniffing, two-tailed paired *t*-test, $t = 3.35$, d.f. = 11, $P = 0.006$) (**c**) and time spent in the zone (cumulative time spent, two-way ANOVA, time × light (off, on): $F_{(359, 3,949)} = 4.97$, $P < 0.0001$; inset, total time in the zone, two-tailed paired *t*-test, $t = 3.86$, d.f. = 11, $P = 0.002$) of the demonstrator paired to the photostimulation compared to the unpaired demonstrator (**d**). **e**, No difference in the number of visits in the zone paired with the light compared to the unpaired zone (two-tailed paired *t*-test, $t = 0.10$, d.f. = 11, $P = 0.291$). **f**, Left: head distance of the observers was shorter to demonstrators paired with light (two-tailed paired *t*-test, $t = 2.09$, d.f. = 11, $P = 0.05$). Right: schematic of second-by-second (over 360 s) head distance from demonstrators. **g**,**h**, Data derived from $n = 7$ mice. Mice tested with two neutral demonstrators without light stimulation showed no difference in sniffing (two-tailed multiple *t*-test, Bonferroni correction, 2 min: $t = 1.51$, d.f. = 12, $P = 0.463$) (**g**) and time spent in each zone (two-way ANOVA, affective state (neutral 1, neutral 2): $F_{(1,10)} = 2.52$, $P = 0.996$) (**h**). **i**, Photostimulation coupled to exploration of an object did not induce object discrimination (two-way ANOVA, time × light (off, on): $F_{(359, 7,898)} = 0.26$, $P > 0.999$, $n = 12$ mice). **j**, Graphical representation of the amount of time mice spent in the stimulation or no-stimulation compartment during RTPP (with blue as the shortest and red as the longest time). **k**,**l**, Data derived from $n = 8$ mice. Mice showed no preference between the two compartments (two-tailed paired *t*-test, $t = 0.67$, d.f. = 7, $P = 0.522$) (**k**) and no difference in distance traveled during RTPP (two-tailed paired *t*-test, $t = 0.63$, d.f. = 7, $P = 0.543$) (**l**). **m**, Left: schematic representation of optical self-stimulation of SOM+ cells. Right: activation of SOM+ cells is not sufficient to induce self-stimulation (two-way ANOVA, active versus inactive nose poke: $F_{(1,3)} = 2.27$, $P = 0.228$, $n = 4$ mice). Bar and line graphs show mean ± s.e.m. $^*P < 0.05$. $^{**}P < 0.005$. n.s., not significant.

Photonhibition of mPFC pyramidal neurons does not abolish affective state discrimination

Interneurons dynamically modulate the activity of pyramidal neurons, which are the major cell type in the mPFC. Our electrophysiological recordings indicated that wide-spiking putative pyramidal cells might also show increased activity when exploring a demonstrator in an altered affective state. However, this increased activation was smaller than that in narrow-spiking cells and did not overcome wide-spiking activity in the habituation phase. Therefore, we tested a possible involvement of pyramidal cell activation in ADTs by injecting C57/BL6 mice bilaterally with AAV-CaMKIIa-eNpHR3.0-eYFP and implanting the optical fibers dorsally to the injection site (Fig. 6a). We then tested these mice in the relief and stress ADTs, inhibiting pyramidal neurons during the first 2 min period (Fig. 6b,e). Photoinhibition of mPFC pyramidal neurons did not modify affective state discrimination of the relieved demonstrator (Fig. 6d), but it partially reduced discrimination between the stressed and the neutral demonstrator (Fig. 6g). In both conditions, inhibition of mPFC pyramidal neurons did not influence the total amount of time spent exploring the two demonstrators (Fig. 6c,f). These results indicate that silencing mPFC pyramidal cells is not sufficient to completely abolish discrimination of affective states and also fails to influence general sociability indexes. Moreover, these data also suggest that the affective state-related firing

pattern that we observed in wide-spiking cells (Fig. 2d,h) could not depend uniquely on pyramidal neurons. Notably, wide-spiking neurons are a heterogeneous neuronal population that might also include SOM+ interneurons (14,33).

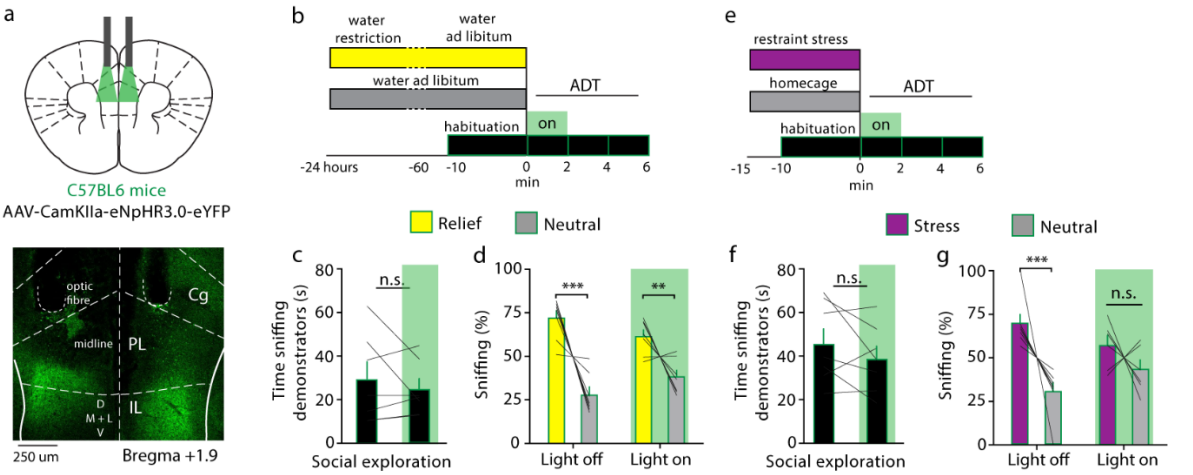


Fig. 6 | Photoinhibition of pyramidal cells in the mPFC does not abolish affective state discrimination. a, Top: C57BL6 male mice were bilaterally injected in the mPFC with AAV-CamKIIa-eNpHR3.0-eYFP and implanted with optic fibers terminating dorsal to the injection area. Bottom: representative image of coronal mPFC section. Findings were replicated in three independent experiments with similar results. Data derived from $n = 7$ mice. **b**, Mice were tested in the ADT with one relieved and one neutral demonstrator. Photoinhibition was performed for 2 min from the beginning of the test using continuous green light ($\lambda = 532$ nm). **c**, Total exploration towards the demonstrators was not affected by photoinhibition of pyramidal cells (two-tailed unpaired *t*-test, $t = 0.52$, d.f. = 12, $P = 0.609$). **d**, Increased sniffing (expressed as a percentage) of the relieved demonstrator compared to the neutral demonstrator during the first 2 min (two-way repeated-measures ANOVA, affective state (neutral, relief): $F_{(1,12)} = 32.37$, $P = 0.0001$). No effect of light delivery was found (two-way repeated-measures ANOVA, light (off, on): $F_{(1,12)} = 1.71$, $P = 0.215$). **e**, Mice were tested in the ADT with one stressed and one neutral demonstrator and photoinhibition was performed for the first 2 min. Data derived from $n = 7$ mice. **f**, No effects of photoinhibition of pyramidal cells on total social exploration were seen (two-tailed unpaired *t*-test, $t = 1.34$, d.f. = 6, $P = 0.227$). **g**, Increased exploration of the stressed demonstrator compared to the neutral demonstrator in the light-off condition, but not in the light-on condition (two-way repeated-measures ANOVA, affective state (neutral, stress) × light (off, on): $F_{(1,12)} = 10.39$, $P = 0.007$; Bonferroni multiple comparison, light off $P < 0.0001$, light on $P = 0.447$). Bar and line graphs show mean ± s.e.m. $^{**}P < 0.005$. $^{***}P < 0.0005$. n.s., not significant.

Affective state discrimination activates SOM+ and inhibits pyramidal neurons

To clarify how mPFC SOM+ neurons physiologically respond during affective state discrimination and how their activity is integrated in the mPFC network, we next investigated the dynamics of individual SOM+ and pyramidal neurons during the ADT. We performed in vivo microendoscope imaging of neurons expressing a genetically encoded calcium indicator. To visualize changes in intracellular calcium concentration indicative of neural activity, we selectively expressed GCaMP6m (AAV-Syn.Flex. GCaMP6m) in SOM+ and GCaMP6f (AAV-

CamKIIa-GCaMP6f) in pyramidal neurons (Fig. 7a). We found an overall increase in the rate of Ca^{2+} events in the ADT compared to the habituation for SOM+ neurons in the stress procedure (Fig. 7b,c). By contrast, a decreased number of Ca^{2+} events was evident for pyramidal neurons in the neutral–neutral condition (Fig. 7d,e). More importantly, the direct quantification of the activity of individual neurons within each ADT paradigm (that is, only neutrals, relief versus neutral, and stress versus neutral) revealed that SOM+ cells were more active when the observer was in contact with either the relieved or stressed than when the observer was in contact with the neutral demonstrator (Fig. 7f). Furthermore, both the relief and stress ADTs triggered a similar increase in activated SOM+ cells (Fig. 7g). By contrast, pyramidal neurons were less active when the observer was in contact with the stressed demonstrator than when in contact with the neutral demonstrator, whereas no differences were evident in the relief ADT (Fig. 7h). Moreover, we showed an increased percentage of inhibited pyramidal cells during contact with a mouse in an altered affective state (relief versus neutral and stress versus neutral) than during contact with a mouse in a neutral state (neutral versus neutral, Fig. 7i).

Affective state discrimination triggers the synchronous activity of SOM+ interneurons

Neuronal coordination is thought to play a key role in information propagation and processing, and SOM+ interneurons are crucial for rhythm generation (34,35). We therefore assessed the possible impact of the synchronization of SOM+ interneurons or pyramidal neurons on affective state discrimination. A comparison of the synchronous activity of neurons (SI) between the habituation and ADT revealed no difference between SOM+ interneurons and pyramidal neurons in the condition with only neutral demonstrators (Fig. 7j). By contrast, SOM+ neurons showed a higher synchronization than pyramidal neurons in both the relief and stress ADT (Fig. 7j). Furthermore, compared to the habituation phase, the synchronized SOM+ pairs increased (Fig. 7k) only in presence of demonstrators in altered affective state. By contrast, the synchronized pyramidal pairs were lower in the ADT than in the habituation in all conditions (Fig. 7k). Finally, the pairwise synchronization of SOM+ neurons was higher in the relief and stress ADTs than in the only-neutral condition (Fig.7l), whereas pyramidal neurons showed the opposite result (Fig. 7l). These data show that the coding of an altered affective state in a conspecific is associated with an increase in both the activation and synchronization of SOM+ neurons, and with an increased inhibition and reduced synchronization of pyramidal neurons (Fig. 7m). The calcium imaging data

support the optogenetic evidence that simultaneously inhibiting SOM+ neurons could abolish affective state discrimination by preventing the synchronized activation of SOM+ neurons, whereas inhibiting pyramidal neurons does not abolish affective state discrimination because it simulates the physiological inhibition of pyramidal cells by altered affective states. We then directly tested the hypothesis that a synchronous inhibition of SOM+ neurons might lead to a higher disinhibition of pyramidal neurons. In vivo calcium imaging of mPFC pyramidal neurons (AAV-CaMKIIa-GCamP6f) revealed that the majority of pyramidal neurons increased their Ca^{2+} events during simultaneous photoinhibition of SOM+ neurons (AAV-DIO-eNpHR3.0, Fig. 7n,o). In addition, optogenetic-assisted tetrode recordings in SOM-Cre mice injected with eNpHR3.0 confirmed that slow-firing putative pyramidal cells can be disinhibited for the entire period of SOM+ photoinhibition (Fig. 7p). Conversely, photostimulation of SOM+ neurons in the mPFC consistently inhibited pyramidal neurons (Supplementary Fig. 6a–e). These data indicate that an overall synchronous inhibition of SOM+ neurons prevalently disinhibits mPFC pyramidal neurons.

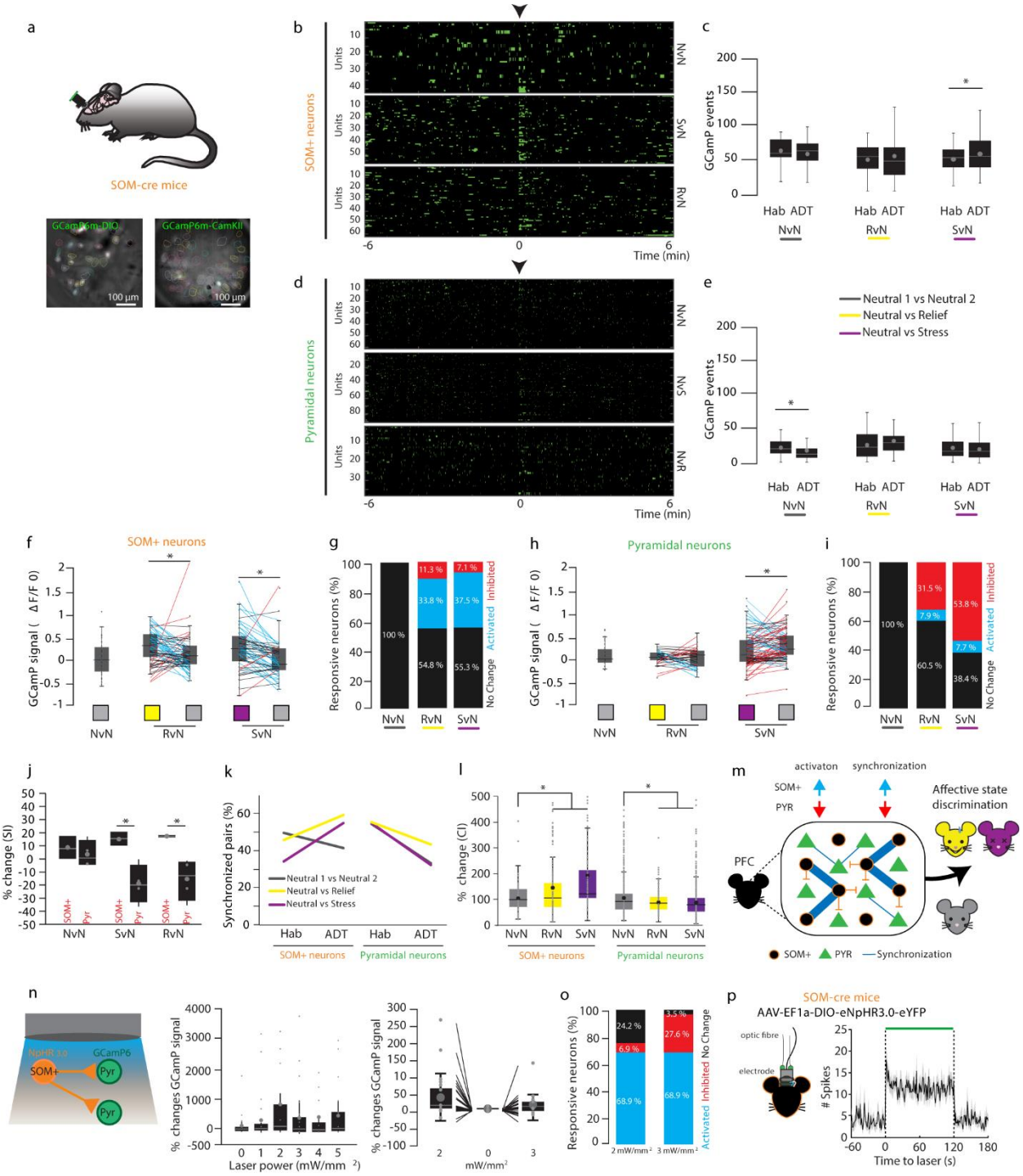


Fig. 7 | Concurrent activation of SOM+ interneurons in the mPFC inhibits pyramidal neurons and drives affective state discrimination. a, SOM-Cre male mice were injected in the mPFC with AAV.Flex.GCaMP6m (n = 4 mice) to target SOM+ interneurons or with AAV-CaMKIIa-GCaMP6f to record calcium dynamics in pyramidal neurons (n = 3 mice). Mice were implanted with a ProView lens to enable large-scale neuronal Ca^{2+} imaging using a miniature fluorescence microscope. **b**,**d**, Cluster traces of Ca^{2+} transient events from SOM+ interneurons (**b**) and pyramidal neurons (**d**) during habituation (time −6 to 0 min) and ADT testing (time 0 to +6 min; the black arrow represents the beginning of the test) in three different conditions (with two neutral demonstrators, a relieved and a neutral demonstrator, and a stressed and a neutral demonstrator). **c**, No variation was seen in the total number of GCaMP events in SOM+ interneurons between habituation and test in the condition with two neutral demonstrators

(two-tailed paired *t*-test, $t = 0.746$, d.f. = 41, $P = 0.45$) or in the condition with a relieved and a neutral demonstrator (two-tailed paired *t*-test, $t = -0.146$, d.f. = 61, $P = 0.884$), and an increase was noted during the condition with a stressed and a neutral demonstrator (paired *t*-test, $t = -1.98$, d.f. = 56, $P = 0.0262$). **e**, We found a decreased total number of GCaMP events in pyramidal neurons between habituation and test in the neutral versus neutral (two-tailed paired *t*-test, $t = 3.44$, d.f. = 60, $P = 0.0005$), but not in the relief versus neutral (two-tailed paired *t*-test, $t = -1.22$, d.f. = 37, $P = 0.229$) or stressed versus neutral (two-tailed paired *t*-test, $t = 0.919$, d.f. = 90, $P = 0.360$) condition. **f**, Increased Ca^{2+} signal from SOM+ interneurons during single contacts with each demonstrator, normalized by 200 ms, when exploring a conspecific in an altered affective state compared to a neutral state in the relief versus neutral (two-tailed paired *t*-test, $t = 2.21$, d.f. = 61, $P = 0.015$) and stress versus neutral (two-tailed paired *t*-test, $t = 4.242$, d.f. = 55, $P = 0.00001$) condition. **g**, Activated or inhibited neurons are those above or below 2 s.d., calculated from the average Ca^{2+} signal during contact in neutral versus neutral condition within each animal. During contact with the relieved demonstrator, 33.8% of the total neurons were activated, 11.3% were inhibited and 54.8% presented no changes. During contact with the stressed demonstrator, 37.5% of the total neurons were activated, 7.1% were inhibited and 55.3% were not responding (total number of recorded neurons: neutral versus neutral = 42, relief versus neutral = 62, stress versus neutral = 56). **h**, The Ca^{2+} signal from pyramidal neurons during contact with each demonstrator, normalized by 200 ms, shows no changes between the neutral and relieved state (two-tailed paired *t* test, $t = 0.2662$, d.f. = 37, $P = 0.791$), and a decrease between neutral and stressed states (two-tailed paired *t*-test, $t = -2.68$, d.f. = 90, $P = 0.004$). **i**, Neuronal activity analyzed as in **g**: 31.5% of the total recorded neurons were inhibited during contact with the relieved demonstrator, 7.9% were activated and 60.5% showed no significant changes. During contact with a stressed demonstrator, 53.8% of the recorded neurons were inhibited, 7.7% were activated and 38.4% did not respond (total number of recorded neurons: neutral versus neutral = 61, relief versus neutral = 38, stress versus neutral = 91). **j**, Percentage of change in the synchronization index (SI) (between the test and the habituation period) of the entire recorded population during exposure to two neutral demonstrators, one neutral and one relieved demonstrator, and one neutral and one stressed demonstrator. The population of SOM+ neurons increased their synchronization in the presence of mice with altered affective states, whereas the population of pyramidal neurons decreased their synchronization in similar conditions (one-way ANOVA, neutral versus neutral, $F_{(1,5)} = 0.52$, $P = 0.5117$; relief versus neutral, $F_{(1,5)} = 11.77$, $P = 0.0265$; stress versus neutral, $F_{(1,5)} = 8.77$, $P = 0.0415$). **k**, Increased number of synchronized SOM+ pairs (in each animal) only in the presence of affectively altered mice (neutral versus neutral, 49.11% versus 41.17%; relief versus neutral, 45.74% versus 59.10%; stress versus neutral, 34.10% versus 54.91%), and reduced number of synchronized pyramidal pairs in the ADT compared to the habituation and in the ADT in all conditions (habituation versus test; neutral versus neutral (NvN), 54.23% versus 33.01%; relief versus neutral (NvR), 55.41% versus 43.34%; stress versus neutral (NvS), 54.93% versus 32.31%). **l**, Percentage of change in the correlation index (CI) between the test and the habituation revealed a significant increase in synchronized SOM+ pairs (one-way ANOVA, $F_{(2, 535)} = 12.55$, $P = 0.000001$, post hoc Bonferroni test, NvN versus NvS, $P = 0.0001$; NvN versus NvR, $P = 0.001$; NvR versus NvS, $P = 0.079$) and a significant decrease in synchronized pyramidal pairs (one-way ANOVA, $F_{(2, 2,635)} = 29.1$, $P = 0.00001$, post hoc Bonferroni test, NvN versus NvS, $P = 0.0001$; NvN versus NvR, $P = 0.0001$ NvR versus NvS, $P = 1.0$) in the presence of an affectively altered mouse compared to the neutral versus neutral condition. Box-and-whisker plots display the median (center line), mean (circle), third quartile (box) and lowest to highest observation (whiskers). **m**, Graphical model of the suggested mechanism implicating mPFC SOM+ interneurons in affective state discrimination. **n**, SOM-Cre mice injected in the mPFC with AAV-CaMKIIa-GCaMP6f and AAV-DIO-eNpHR3.0 to record the calcium dynamics of pyramidal neurons following inhibition of SOM+ neurons ($n = 3$ mice). Diagram illustrating the state of mPFC pyramidal neurons during photoinhibition of SOM+ interneurons in the mPFC. Recordings of Ca^{2+} events during baseline and following photoinhibition of SOM+ interneurons revealed an increase in the event rate proportional to the light power used (0, 1, 2, 3, 4, 5 mW mm−2). Data are shown as the average percentage change between the photostimulation period and the baseline. Pairwise analysis of the variation from the whole neuronal population ($n = 26$ neurons) revealed a significant increase in event rate following 2 or 3 mW mm−2 compared to the absence of stimulation at 0 mW mm−2 (one-way ANOVA, $F_{(2,86)} = 6.24$, $P = 0.003$, Bonferroni post hoc test, 0–2 mW mm−2, $P = 0.003$; 0–3 mW mm−2, $P = 0.791$; 2–3 mW mm−2, $P = 0.065$). Box-and-whisker plots display the median (center line), mean (circle), individual observations (small circles), third quartile (box) and lowest to highest observation (whiskers). Pyr, pyramidal. **o**, During photoinhibition of SOM+ interneurons with either 2 or 3 mW mm−2, the same pyramidal neurons (68.9%) showed an increase in their event rate, whereas a variable small minority of pyramidal neurons were inhibited or unaffected. **p**, Optogenetic-assisted tetrode recordings in the mPFC in AAV-EF1a-DIO-eNpHR3.0-eYFP-injected SOM-Cre mice; representative example (one neuron with eight stimulation repetitions) of a peristimulus time histogram showing the firing rate (mean ± s.e.m.) of slow-firing putative pyramidal cells, before, during and after the delivery of the green light pulse (120 s).

Discussion

In this study, we revealed a specific association between affective state discrimination and mPFC neuronal activity. In particular, using bidirectional optogenetic manipulations, we demonstrated that SOM+, but not PV+ interneurons, in the mPFC are crucial for affective state discrimination. Moreover, by means of single-cell calcium imaging, we demonstrated that an increase in SOM+ cells synchrony and a decrease in pyramidal cell synchrony are correlated with affective state discrimination.

In agreement with evidence from human studies (7–12), our findings in mice refine the involvement of the mPFC inhibitory circuits in the processing and reaction to altered affective states in others. In particular, the activity of mPFC neurons increased during the exploration of conspecifics with an altered affective state. The increased mPFC firing when approaching a mouse is consistent with previous evidence showing increased social-dependent firing that did not occur with inanimate objects or a novel environment (36). However, in contrast to the habituation pattern of the standard social approach (36), the increased firing towards a mouse in an altered affective state remained sustained throughout the test. Furthermore, behaviors measured in the ADT and related mPFC neuronal firing were qualitatively or quantitatively different in the presence of the demonstrators or only of their odor cues. Compared to other complex social paradigms in rodents (22,37,38), the behaviors we observed differed in the nature and intensity of the social interaction, possibly linked to differences in the age of the animals (juvenile versus adult), species (rats versus mice), familiarity (Extended Data Fig. 2), experimental setting (novel environment versus home cage), and the intensity of the affective state (27). Notably, the intensity of our manipulations (stress and relief) was milder than that used in other studies (37,38), and did not induce any increased exploration in a classical one-on-one social interaction test. Overall, our findings highlight the distinct social functions assessed by the ADT and suggest that distinct patterns of mPFC microcircuit recruitment might support specific social cognitive functions.

We revealed a double dissociation between the roles of mPFC interneurons in social functions. Manipulation of SOM+ cell activity altered affective state discrimination, but did not alter social preference in general. By contrast, inhibition of mPFC PV+ neurons altered general sociability without affecting discrimination of affective states. Indeed, increasing mPFC pyramidal excitation, directly or through the inhibition of PV+ interneurons, can reduce sociability and social interaction in a classic three-chamber test (20). SOM+ interneurons can influence the activity of

pyramidal cells both by inhibiting PV-expressing interneurons (39), which provide perisomatic inhibition to pyramidal neurons (40), and by providing direct inhibitory inputs to distal dendritic branches (40,41). Thus, the effects of SOM+ neurons on affective state discrimination could rely either on their inhibition of PV+ cell activity or on their direct effects on the synaptic integration of pyramidal neurons. However, a direct inhibition of PV+ cells had no effect on affective state discrimination. Moreover, activation but not inhibition of PV+ cells could rescue social deficits in an animal model of autism (42). Thus, we could assume that SOM+ interneurons contribute to the correct integration of information flow from other brain structures, as recently suggested for working memory functions through interactions with the ventral hippocampus (31). This interpretation is in line with our evidence that, at the microcircuit level, synchronized activity of mPFC SOM+ neurons is correlated with affective state discrimination. This could be associated with the final integration of inputs arising from several other brain regions. Among them, a possible substrate could be excitatory synaptic inputs from the basolateral region of the amygdala onto the mPFC (43). The amygdala has been most consistently associated with the control of both negative and positive affective states (44,45). Intriguingly, previous evidence showed that, within the PFC, SOM+ neurons, but not PV+ neurons, are highly enriched in oxytocin receptors (46), and that the oxytocin system is strongly implicated in social functions (47,48) and emotion recognition (27,49,50). As such, it is tempting to speculate that mPFC SOM+ interneurons constitute a major site of integration for socially valenced information from subcortical structures. This role of mPFC SOM+ neurons appeared to be specific for the recognition and reaction to the affective states of mediating the discrimination of affective states versus the general extrinsic avoidance or seeking of aversive and rewarding stimuli as well as other social processes.

Our in vivo microcircuit investigation indicates that pyramidal neurons were mostly inhibited during affective state discrimination. Moreover, we observed a reduction in the synchronized activation of pyramidal neurons associated with affective state discrimination. Consistently, photostimulation of mPFC SOM+ interneurons, which induced social discrimination, inhibited mPFC pyramidal neurons. In line with these data, it is conceivable that the direct inhibition of mPFC pyramidal neurons was not sufficient to abolish affective state discrimination since it emulated the natural effect of activated SOM+ interneurons during affective state discrimination. Indeed, we were able to abolish discrimination of affective states only by photoinhibition of mPFC SOM+ cells, and we demonstrated that this manipulation mostly

activated mPFC pyramidal neurons. Taken together, these findings demonstrate a previously unreported role for mPFC SOM+ interneurons in specific aspects of social cognition.

In summary, we used a mouse model inspired by a well-established human task for emotion recognition, and uncovered a selective and pivotal role for mPFC SOM+ interneurons in the ability to discriminate the expression of affective states in others. This finding provides a deeper mechanistic understanding of social cognition with relevance to the development of social cognitive dysfunctions in psychiatric disorders such as ASD and schizophrenia. Indeed, our findings hold both fundamental and clinical relevance, refining the idea of excitatory–inhibitory imbalance as a core behavioral dysfunction in these disorders and providing a convergent and selective target for manipulation of emotion recognition abilities.

Material and Methods

Mice

All procedures were approved by the Italian Ministry of Health (permits 230/2009-B, 107/2015-PR, and 749/2017-PR) and local Animal Use Committee and were conducted in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the European Community Council Directives. Routine veterinary care and animal maintenance was provided by dedicated and trained personnel. Males and female C57BL/6 J animals aged 3–6 months old were used. Founders of the *PvalbCre*, B6.129P2- *Pvalbtm1(cre)Arbr/J*, id #017320, RRID:IMSR_JAX:017320 (called PV-Cre line) and *SomCre*, *Somtm2.1(cre)Zjh/J* mice, id #013044, RRID:IMSR_JAX:013044 (called SOM-Cre line) were purchased from The Jackson Laboratory and then bred and expanded in our animal facility for successive testing. Mouse genotypes were identified by polymerase chain reaction (PCR) analysis of ear DNA. Distinct cohorts of naïve mice were used for each experiment. Animals were housed two to four per cage in a climate-controlled facility (22 ± 2 °C), with ad libitum access to food and water throughout, and with a 12 h light–dark cycle (19:00–07:00 schedule). Experiments were run during the light phase (within 10:00–17:00). All mice were handled on alternate days during the week preceding the first behavioral testing. Female mice were visually checked for estrus cycle immediately after the test and no correlation was found between estrus status and performance in the test.

Behavioral paradigms

Affective Discrimination Test. Testing mice (observers) were habituated (3 nonconsecutive days, for 10 min) inside a standard mouse cage (Tecniplast, 35.5 × 23.5 × 19 cm) equipped with a dark separator between two cylindrical wire cups that hosted the demonstrators (10.5 cm in height, bottom diameter 10.2 cm, bars spaced 1 cm apart; Galaxy Cup, Spectrum Diversified Designs). The separator (11 × 14 cm) between the two wire cups was wide enough to cover the reciprocal view of the demonstrators while leaving the observer mice free to move between the two sides of the cage. A plastic cylinder was placed on the top of the wire cups to prevent the observer mice from climbing and remaining on the top of them. The cups, separators, and experimental cages were replaced after each subject with clean copies to avoid scent carryover. Similarly, the rest of the apparatus was wiped with water and dried with paper towels for each new subject. After each testing day, the wire cups, separators, and cubicles were wiped down with 70% ethanol and allowed to air dry. Testing cages were autoclaved as performed as standard in our animal facility. Habituation to the testing setting occurred on 3 consecutive days before the first experiment; each habituation session lasted 10 min. During both habituation and behavioral testing, the cages were placed inside a dimly lit (6 ± 1 lx) soundproof cubicle (Med Associates). Demonstrator mice, matched by age, sex, and strain to the observers, were habituated, without the observer, inside the same cage under the wire cups. A recent report showed that encoding of social cues can depend on spatial location (51). However, we always counterbalanced the presentation of the neutral versus affectively altered demonstrators in the two sides of the testing arena and no differences were observed in the affective state discrimination performance depending on the spatial location of the demonstrators (two-way repeatedmeasures ANOVA, location of demonstrators: $F(1,49) = 0.25$, $P = 0.61$). Similarly, the optogenetics manipulations were as effective if the affectively altered mouse was on the left or right side (two-way repeated-measures ANOVA, location of demonstrators: $F(1,43) = 1.14$, $P = 0.29$). Therefore, we cannot exclude that some sort of spatial selectivity of neural activity would be present during our task, but this was not crucial for the behaviors that we observed in the ADT.

Observers. On the day of testing, 10 min before the experiment, observer mice were gently moved into the dimly lit testing cubicles. For the optogenetics and in vivo electrophysiology experiments, observer mice were connected to an optic fiber or headstage during this period. Then, one neutral

demonstrator and one affectively altered demonstrator (relief or stress) mouse were placed under the wire cups, and the 6 min experiment started.

Neutral demonstrators. All neutral mice were habituated to the experimental setting as described above. For both relief and stress conditions, neutral demonstrators did not receive any manipulation and were left undisturbed, with ad libitum water access, in their home cage. On the day of testing, neutral demonstrators were brought, inside their home cages, into the experimental room 1 h before the experiment began. All demonstrators were group-housed, separately from the cages of stressed and relieved demonstrators. Demonstrators were test naive and used a maximum of two or three times, always with at least 1 week between each consecutive test. No differences were observed in the performance of the observer mice depending on the previous experience of the demonstrator.

Relief demonstrators. Mice were habituated to the experimental setting as reported above. Relieved demonstrators were water-deprived 23 h before the experiment. Ad libitum access to water was given 1 h before the test, and mice were brought inside the experimental room in their home cages. Food was ad libitum all the time and some extra pellets were put inside the home cage during the 1 h water restoration.

Stress demonstrators. Mice were subjected to a mild stress consisting of the restraint tube test, a standard procedure to induce physiological stress in rodents (52), for 15 min before the beginning of the ADT. These mice were then immediately moved to the testing arena. Digital cameras (Imaging Source, DMK 22AUC03 monochrome) were placed facing the long side of the cage and on top of the cage to record the test from different angles using a behavioral tracking system (Anymaze 6.0, Stoelting). These videos were used offline by experimenters blind to the manipulations of both the observers and demonstrators for a posteriori scoring of behaviors: sniffing, grooming, rearing, freezing, time spent in the zones, visits in the zones, latency to make the first visit, average length of visits, and locomotion parameters (distance traveled and average speed). Manual behavioral scoring was performed with videos at half speed by two or three independent trained experimenters (inter-rater reliability r score >0.90) blind to the manipulations and genotypes.

One-on-one social exploration tests. This test was similarly performed as previously described (47). One hour prior to behavioral testing, each experimental subject was placed into a plastic cage

(Tecniplast, 35.5 × 23.5 × 19 cm) with shaved wood bedding and a wire lid, in a room adjacent to the testing room. The testing cages containing the observer mice were gently moved into the testing soundproof cubicles 5 min before the experiment. To begin the test a demonstrator mouse was introduced to the cage for 6 min (as for the ADT), and exploratory behaviors initiated by the test subject were scored by two independent experimenters blind to the manipulations. Demonstrator mice were used only once. Each observer was given tests on consecutive days: once with an unfamiliar naive conspecific (neutral), once with an unfamiliar relieved conspecific (relief, as described above) and once with an unfamiliar stressed conspecific (stress, as described above). Test order was counterbalanced.

Sociability and social novelty tests. We adapted a widely used standard test for assaying sociability in mice to our setting (32,53). The session started with the observer in the same testing cage as that used for the ADT for a 6 min habituation period. After habituation, the observer was presented for another 6 min with a white or black plastic object (novel object) contained in one of the two wire cups, placed in one side of the chamber. Simultaneously, an adult conspecific mouse (novel mouse 1) that had no previous contact with the observer, was placed in the wire cup in the other side of the chamber. To measure sociability (the tendency of the subject mouse to spend time with a conspecific, compared with time spent with an object), a discrimination index was calculated (time spent with novel mouse 1 – time spent with novel object / total time spent with novel mouse 1 and novel object). Following the sociability test, the object was replaced with a novel mouse (novel mouse 2) and the observer was tested for another 6 min to assess the preference for social novelty. This is defined as more time spent in the chamber with novel mouse 2 than time in the chamber with novel mouse 1. Most mice prefer to spend more time near the unfamiliar novel mouse (novel mouse 2). To assess social novelty, we calculated a discrimination index for each mouse (time spent with novel mouse 2 − time spent with novel mouse 1 / total time spent with novel mouse 1 and novel mouse 2).

RTPP. Mice were placed into a custom-made two-chambered box (35.5 × 23.5 × 19 cm) and allowed to explore the two compartments for 20 min. Using Anymaze (version 6.0, Stoelting) connected to a blue light laser, light stimulation at 30 Hz (5 ms pulse width) was delivered during the duration of their time spent in the light stimulation side of the chamber. Mice received no light

stimulation on the no stimulation side. The experimental animals were counterbalanced for the stimulation side. Preference in each experiment was determined as the percentage of time spent in the light stimulation side out of the total explored time during the test.

Optical self-stimulation. Successful nose pokes on the active portal were rewarded with 5 s of light stimulation (5 ms, 30 Hz) and a cue light turned on for 5 s. An inactive nose poke resulted in a cue light on for 5 s, but did not trigger light stimulation. This procedure was repeated for 2 d.

Sensory modality assessment. See Chapter 2, page 46-47.

Food discrimination test. To test discrimination of palatable food rewards the same setting and procedure of the ADT were used. Two small transparent plastic bowls were placed in the two opposite compartments of the arena. On the day of testing, two different food rewards were placed in the bowls: a sucrose reward (50 pellets, 14 mg; Test Diet, 5-TUT), and a purified reward (50 pellets, 14 mg; Test Diet, 5-TUL). A Digital camera (Imaging Source, DMK 22AUC03 monochrome) placed on top of the cage recorded the test using a behavioral tracking system (Anymaze 6.0, Stoelting). Videos were used offline to calculate the time spent in the zones with with food rewards (6 × 6 cm area around the bowls). To assess the preference of the mice the number of pellets eaten during the test was also counted.

Corticosterone assay. See Chapter 2, page 48-49.

Viral injections, optic fibers, and tetrode implants

Viral vectors. AAV5-EF1a- DIO-eNpHR3.0-eYFP.WPRE.hGH (Addgene, 20949, qTiter 1.95 × 1013 GC per ml), AAV5-EF1a-DIO-hChR2(H134R)-eYFP.WPRE.hGH (Addgene, 20298P, ddTiter 2.76 × 1013 GC per ml) and AAV1-CamKIIa-eNpHR3.0-eYFP.WPRE. hGH (addgene, 26971P, qtiter 5.12 × 1012 GC per ml) were purchased from the University of Pennsylvania Viral Vector Core.

Surgical procedures. C57BL/6J, SOM-Cre, and PV-Cre mice were naive and 2–3 months old at the time of surgery. All mice were anesthetized with a mix of isoflurane (2%) and oxygen (1.5%) by inhalation and mounted into a stereotaxic frame (Kopf Instruments) linked to a digital

micromanipulator. Brain coordinates of viral injection in the mPFC were chosen in accordance with the mouse brain atlas (54): anterior–posterior (AP), +1.9 mm; medial–lateral (ML): ± 0.30 mm; and dorsal–ventral (DV): −2.4 mm. The volume of AAV injection was 0.4 µl per hemisphere. We infused virus through a glass micropipette connected to a 10 µl Hamilton syringe. After infusion, the pipette was kept in place for 5 min and then slowly withdrawn over 5 min. After virus injection mice were given 3 weeks to recover and for the viral transgenes to adequately express before undergoing optic fiber implantation and behavioral experiments. Mice underwent stereotaxic surgery for fiberoptic implantation 3 weeks after viral injection. The skull was exposed and the two previous holes were used to target the mPFC. Dual fiberoptic cannulae (200 µm, 0.37 numerical aperture, fiber distance 0.7 mm; Doric Lenses) were lowered 2 mm from the surface of the skull at roughly 400 µm dorsal to the virus injection site, and implants were secured to the skull with MetaBond and dental cement. For in vivo electrophysiological recording, mice were implanted with silicon probes carrying a 16-channel linear multi-electrode array in the right mPFC (Neuronexus, A4x4-3mm-100-125-177-Z16) in accordance with the mouse brain atlas (54): AP, +1.9 mm; ML, ± 0.30 mm; and DV, −2.5 mm. Before the permanent attachment to the skull, the tetrodes were protected with Kwik-Kast silicone elastomer (World Precision Instruments) and secured using dental acrylic. After electrodes and fiberoptic implantation, mice were allowed to recover for 7–10 d depending on the general health.

In vivo optogenetic behavior

During behavioral testing, fiberoptic cannulae were connected to patch cords (Doric Lenses), which were in turn connected to blue (447 nm) or green light (532 nm) lasers (CNI Laser) using a 1 × 2 intensity division fiberoptic rotary joint (Doric Lenses) located above the cubicle containing the testing arena. Laser power was adjusted such that the light exiting the fiberoptic cable was approximately 4.5 mW. For photoinhibition experiments we used continuous green light and laser stimulations were controlled by a microcontroller board (Arduino Mega 2560, Arduino). For photostimulation experiments we used either 5 s continuous light pulses or 30 Hz, 5 ms pulses of blue light. Ex vivo experiments demonstrated that the 5 s continuous pulse was efficient in exciting SOM+ cells for at least 3 s (Supplementary Fig. 6d, e). Behavioral data with the two light protocols did not differ, and were therefore pooled. To control optical inhibition or stimulation with a closed-loop system dependent on mice behavior during ADTs, a behavioral tracking system (Anymaze

6.0, Stoelting) detected online the location of the observer mouse. When the observer entered an area (14 × 12 cm) that included the demonstrator to be paired with optogenetical manipulations (the relieved and the stress demonstrators; one of the two naïve neutral demonstrators) the software triggered the laser.

In vivo electrophysiology. *Surgery and acquisition*. SOM-Cre mice (n = 5) underwent a second surgery for the implantation of a Microdrive for tetrode recording of extracellular activity 2–3 weeks following bilateral injection of AAVDIO- eNpHR3.0. Briefly, a Microdrive (Axona) composed of four twisted tetrodes (tungsten 12 μm, M408870, CFW) with a gold-cyanide adjusted impedance of 0.4 MΩ and two optic fibers (Doric Lenses) targeting the right and left mPFC (same coordinates as above). The implant was fixed to the skull using two anchor screws and adhesive dental cement to reduce movement artifacts. The extracellular activity was acquired at 32 kHz, band pass-filtered at 300–9000 Hz, and digitalized (16 bits per sample, Digital Lynx X, Neuralynx). Two ground screws over the cerebellum served as a reference signal. Following recovery, animals were handled and tetrodes lowered daily until they reached the adequate coordinates (25–50 μm d−1). When the tetrodes were located in the dorsal PFC, the signal was acquired using a 16-channel analog headstage (HS-18-MM, Neuralynx) and a digital Lynx X system (Neuralynx). On each acquisition day, the optic fiber located on the implants was connected to a green laser (532 nm, CNI Laser) by a 1 × 2 fiberoptic rotary joint (Doric Lenses). The laser power was set to 4.5 mW, measured at the tip of each fiber. Each recording consisted of a 10 min baseline following 120 s intervals of photoinhibition (120 s on, 120 s off) for a duration of 30 min. The laser was controlled via Transistor–transistor logic (TTL) pulses, which were recorded using the digital inputs of the recording system.

Analyses. Action potentials were extracted by calculating the root mean square of the first 60 s from the band pass-filtered signal (500–5000 Hz). Whenever the signal exceeded the baseline by two times the mean noise level an action potential was detected. The waveforms recorded on each wire of the tetrode were projected to a lower dimensional space using principal component analyses and automatically clustered using the spike-sorting algorithm (55) (KlustaKwik). Clusters were manually refined with the graphical user interface Klusters (56). Only clusters with a clear refractory period in their spike-time autocorrelation were kept for further analysis. To detect laser-

responsive units, a peristimulus time histogram for each laser pulse was calculated (60 s baseline, 120 s photoinhibition, 60 s post-stimulus, bin size 1 s). Histograms of individual pulses were averaged to obtain the mean laser response of a unit (spikes per s). A unit was considered responsive when the firing rate deviated by more than 2 s.d. after the laser onset from the mean baseline period (60 s before laser onset).

Electrophysiology

Recordings were carried out mainly from the prelimbic cortex by means of a 16-channel Neuralynx Digital X system (NeuraLynx). A Saturn-X Commutator was connected through a TETH-XNT-MM extension cable to the data acquisition system and the headstage to the flex tether, allowing mice to exhibit more natural behavior and unrestricted movements. Unit signals were filtered between 300 and 9000 Hz, digitized at 32 kHz, and stored on a personal computer using a Cheetah data acquisition system (Neuralynx). The anatomical location of the recording region was determined based on the location of a marking lesion. A digital camera (The Imaging Source, DMK 22AUC03 monochrome) was mounted on the side of the testing arena, to record mice behaviors using a behavioral tracking system (Anymaze 6.0, Stoelting). All quantitative analyses of neuronal data were performed offline using NeuroExplorer (version 4.135, Plexon). The raw signals related to all temporal intervals (baseline, and 2 min, 4 min, and 6 min) of a recording session were merged together. The wide-band activity was high pass-filtered offline (300 Hz). This was followed by waveform detection: a negative threshold corresponding to threefold standard deviation from the mean peak height was applied to all simultaneously recorded channels. In this way, the absolute value of the threshold on the raw signal amplitude given in mV was different across the channels, depending on the specific signal-to-noise ratio. After removal of artifacts, all the detected waveforms were considered as multi-units (MUA). Single units were then isolated using principal component and template matching techniques provided by Offline Sorter (version 3.3.5, Plexon). For each isolated unit, we verified that the projection of its spikes in the three-dimensional space formed by the first two principal components and the acquisition time remained stable for the entire duration of the session, and that the percentage of interspike intervals of less than 1 ms was less than 0.5%. With the type of probes used in this study, the recording stability was extremely good: in case of relevant changes in the neuronal activity during acquisition, the entire penetration was discarded and the data were not included in the dataset. To distinguish

between excitatory neurons and inhibitory interneurons, which account for most of the inhibited neurons, we classified recorded units into narrow-spiking putative inhibitory interneurons and wide-spiking putative pyramidal neurons based on spike waveform features, such as spike width and firing frequency (28). In particular, we used three parameters: the length of the period during which the voltage potential of the waveform was negative (indicating membrane depolarization), the length of the period during which the voltage potential of the waveform was positive (indicating membrane hyperpolarization), and the average firing frequency during the entire recording session (57,58).

Behavior. Video images were analyzed offline and manually scored with half speed. A valid exploration trial was defined as a greater than 1 s event of sniffing towards the demonstrator when inside the zone or contact with one of the two demonstrators with an interval from the previous visit of more than 2 s.

Definition of epochs of interest. All the neurons included in the present work were recorded for a variable number of trials depending on the number of times the observer decided to explore the relieved or the stressed rather than the neutral conspecific. Based on the timestamps related to the main behavioral events (start and end of the social exploration), we defined five different epochs of interest for statistical analysis of neuronal responses: the pre-social exploration or preodor exploration epoch, from 2 s to 1 s before the onset of the social exploration or odor exploration; the start social exploration or start odor exploration epoch, corresponding to 1 s before the onset of the social exploration or odor exploration; the social exploration or odor exploration epoch, including the 1 s after the onset of the social exploration or odor exploration; the end social exploration or end odor exploration epoch, ranging from 1 s before the social or odor exploration to the end of the social exploration or odor exploration; and the post-social exploration or post-odor exploration epoch, ranging from 1 s to 2 s after the end of the social exploration or odor exploration.

Single-neuron analysis. After identification of single units that remained stable over the entire duration of the experiment, neurons were defined as task-related if they significantly varied their discharge during at least one of the epochs of interest (see above), investigated by means of the

following repeated-measures analyses of variance (ANOVAs (with significance classed as $P < 0.05$)):

1) Neural response to the beginning of a social exploration. The activity of neurons during three sessions was analyzed separately by identical 3 × 2 ANOVAs, with factor 'epoch' (three levels: baseline, pre-social exploration, and post-social exploration) and 'affective state' (two levels: relief or stress, and neutral) followed by Bonferroni post-hoc tests ($P < 0.05$) in case of significant interaction effects as our goal was to identify not only possible activity changes induced by the social exploration, but also possible differences when the observer explores the relieved or stressed demonstrator rather than the neutral mouse. Neurons that matched the criteria defined above were classified as responsive cells (activated only before or only after the social event, or during both epochs).

2) Neural response to the end of a social exploration. Possible modulation of single-neuron activity corresponding to the end of a social exploration during the three sessions was analyzed separately by identical 2 × 2 ANOVAs, with factor 'epoch' (two levels: pre-end social interaction, and post-end social interaction) and 'affective state' (two levels: relief or stress, and neutral) followed by Bonferroni post-hoc tests ($P < 0.05$). Subsequently, all of these neurons were involved in further statistical analysis. To quantify the preference shown by single neurons for an altered state, the preference index (PI) was calculated as follows:

$$PI = \frac{(Rp - Rn)}{(Rp + Rn)}$$

where *Rp* and *Rn* are the mean firing rates of the neuron in its preferred and nonpreferred conditions, respectively. The selection of the trials to be included in the calculation of *Rp* and *Rn* was based on the results of single-neuron statistical analysis, in terms of main or interaction effects. PI values ranged from 0 (discharge identical between the compared conditions) to 1 (complete selectivity for the preferred condition).

Population analyses. Population analyses were performed on all recorded neurons, classified on the basis of the results of the above described analyses, and considering single-neuron responses calculated as averaged activity (reported as z-scores) in 20 ms bins across trials of the same condition. *Z*-scores were calculated by generating a peristimulus time histogram of the binned neuronal data and summing all action potentials of all the neurons in each bin. The average of an

initial habituation count preceding all three recording sessions was subtracted from this total of spikes. Finally, to identify the start or end of population selectivity for specific variables (that is, relieved or stressed, or neutral demonstrators), paired sample *t*-tests were used to establish the first or last of a series of at least five consecutive 80 ms bins (slid forward in steps of 20 ms) in which the activity significantly differed (uncorrected $P < 0.05$) between the two compared conditions.

In vivo large-scale calcium imaging

Stereotaxic surgeries. For large-scale calcium imaging, SOM-Cre mice (2–3 months old, $n = 4$ mice) were injected, as described above, with GCaMP6m (59) (AAV1.Syn.Flex.GCaMP6m.WPRE.SV40, 100838-AAV1, titer 2.1 × 1013 vg ml−1, diluted 1:5 with saline; ADDGENE) or with GCaMP6f (AAV1.CaMKIIa.GCaMP6f.WPRE.bGHpA, titer 1.68 × 1013 vg ml−1; 1000-002613, Inscopix) and halorhodopsin (AAV5-EF1a-DIO-eNpHR3.0-mCherry-WPRE, AV4832e, titer 3.3 × 1013 vg ml−1, UNC Vector Core, 2–3 months old, $n = 3$ mice). For simultaneous inhibition of SOM+ and Ca^{2+} imaging of pyramidal neurons, the two viruses were mixed in a 1:3 ratio. For all experiments, the coordinates for virus injection in the prefrontal cortex were: AP, +1.9 mm and ML, +0.3 mm; at three different DV coordinates, −2.1, −2.5, and −2.9 mm, from the dura. Mice underwent a second surgery for the implantation of a 0.5 mm outer diameter gradient-index (GRIN) lens (ProView lenses, Inscopix) at the same coordinates as the virus injections, 10–14 d following virus injections. Before lowering the GRIN lens, a pre-track was performed using a blunted needle (25 gauge) to reduce tissue damage and minimize brain pressure. GRIN lens insertion was monitored and positioned at the focal point presenting the highest Ca^{2+} signal and was maintained in position using a thin layer of Kwik-Sil (World Precision Instruments), a blunted skull screw, and adhesive dental cement (Sun Medical Super-Bond C&B Kit, Dental Leader). During recovery, the exposed lens was protected with biocompatible silicone. The distance between the exposed GRIN lens and the miniaturized microscope was defined 10–14 days following implantation using a baseplate (Inscopix) and affixed at the optimal focal plane using adhesive dental cement.

Calcium signal acquisition. The nVista miniature microscope (nVista 3.0, Inscopix) with an integrated blue LED (475 nm, average power 1 mW mm−2) was used to image the GCaMP6 calcium indicator before (habituation) and during the ADT. The miniature microscope was

positioned on the head of the animal 2–3 days before behavioral testing, and the focal plane was adjusted using inbuilt tools in the Acquisition Software (version 1.2.0, Inscopix). Optimum laser power, imaging gain, and focal distance were selected for each animal and conserved across all experiments. The habituation period was conducted in the same apparatus as the ADT with no animals for 6 min, where Ca^{2+} signal was acquired at a frame rate of 20 frames per s. Then, two demonstrator mice were placed under their respective wire cups as described above. Ca^{2+} signals were collected for the entire 6 min period of the ADT. For all demonstrators (two neutral demonstrators, one relieved and one neutral demonstrator, and one stressed and one neutral demonstrator), the signal was processed using data analysis software (Inscopix Data Processing Software version 1.2, Inscopix). All recordings of the ADT signals were analyzed at the same time as their respective habituation period (time series) and spatially down-sampled (×4), and movements were corrected over all recording acquisition windows to reduced movement artifacts. Dropped frames were corrected by a smooth correction of the edges. For the simultaneous optogenetic stimulation and Ca^{2+} imaging recording in freely moving mice, we used the miniature microscope nVoke (nVoke 1.0 system, Inscopix) with an integrated blue LED (EX-LED, 435–460 nm, average power 0.7 mW mm−2) and red LED (OG-LED, 590–650 nm, light power 0–5 mW mm−2). Before recording, the optimal focal plane was manually adjusted for each mouse, and the LED excitation power adjusted and conserved across all experiments. Mice were then placed in their home cage and received six randomized stimulations of 2 min at different photoinhibition power (0, 1, 2, 3, 4 and 5 mW mm−2). Every photoinhibition period was preceded by a baseline period of 2 min, during which only Ca^{2+} signal was acquired. The Ca^{2+} signal was then transferred for analyses using data analysis software (IDPS, Inscopix). All acquired videos were first rearranged on progressive stimulation power (0 to 5 mW mm−2), analyzed all together (time series) and spatially downsampled (×4), and movements were corrected over all recording acquisition windows to reduced movement artifacts. Dropped frames were corrected by a smooth correction (average) of the edges.

Calcium signal analyses. For all experiments, the raw-calcium signal and artifact variation of calcium were then compensated using $\Delta F/F0$, where the fluorescence (F) was compared to the background (F0) in each frame. The first selection of neurons was done using manual regions of interest (ROIs), for which a selection tool was used to delimitate the borders of putative GCaMP-

expressing neurons. The Ca^{2+} signal for all ROIs was converted into an xls file before processing. To remove the photo-bleaching effect, the signal was first corrected using moving average methods over the entire recording. Then the signal was down-sampled to 5 frames per s to fit the decay time of the GCaMP signal (175–250 ms) and normalized using the z-score correction over the habituation or ADT period. At that point, a second selection of ROIs was done, for which only ROIs that presented at least one significant increase in Ca^{2+} ($\Delta F/F0$), further defined as a calcium event (Ca^{2+} signal ≥ mean + 2 s.d. for more than 200 ms), during habituation and test were used. The behavioral contacts were manually scored and the timestamps of the onset or end of the contact with neutral, relieved, or stressed animal were used to define the peri-event time histogram if the contact duration lasted for more than 1 s. Calcium signals ($\Delta F/F0$) and calcium events (2 s.d.) were measured during the entire habituation period and the ADT period. The calcium signal for each ROI was aligned to behavioral events and an individual z-score was calculated for the immediate period before the contact (1 s) for each 200 ms bin of the contact period. To define a variation in the calcium signal as related to the affective state of a mouse, we used the mean and the standard deviation of the neutral-only demonstrators condition as the reference. During the relieved versus neutral and stressed versus neutral conditions, a significant variation (± 2 s.d. of the neutrals only condition) lasting more than 50% of the contact period was considered an increase (+2 s.d.) or a reduction (−2 s.d.) of the calcium activity during contact for each individual neuron. An ROI with an average variation in the calcium signal of below ± 2 s.d. was considered to present no changes during contact.

SI and CI of the Ca^{2+} events. For habituation and ADT recording, population synchronization (SI) was obtained by defining the proportions of neurons presenting calcium events within the same bin (200 ms). All bins presenting no events in all simultaneously recorded neurons were discarded from the average. For ADT only, a CI (ref. 60) was calculated between each possible pair of neurons as follows:

$$CI = \frac{Nab \times \Delta t}{Na \times Nb \times 2T}$$

Where Δt is the synchronization window (200 ms), *Na* is the total number of spikes of the neuron *a*, *Nb* is the total number of spikes of the neuron *b*, *T* is the recording duration (6 min), and $Nab \times \Delta t$ is the total number of spike pairs for which a spike from neuron a falls within Δt of a spike from

neuron *b*. We ran a permutation analysis to test the Poisson distribution for the CI of all pairs of neurons recorded to define the CI associated with random distribution (Clr). Then we used for each population (SOM+ and pyramidal) their respective CIr to define whether a pair of neurons was synchronized during the relieved versus neutral or stressed versus neutral ADTs. We then calculated the percentage of synchronized pairs over the total possible pairs in the habituation and in the ADT. Furthermore, we calculated the percentage change of the CI between the habituation and the ADT for all pairs that have been considered synchronized only in the ADT.

Histology

At the end of the behavioral procedures, we checked viral expression and the position of the optic fibers. Mice were deeply anesthetized (urethane 20%) and transcardially perfused with 4% paraformaldehyde in PBS at pH 7.4. Brains were dissected, fixed overnight, and cryoprotected in 30% sucrose in PBS. Coronal sections (40 μm) were cut using a HM450 microtome (Thermo Fisher Scientific). For immunohistochemical studies free-floating sections of selected areas were washed in PBS three times for 10 min, permeabilized in PBS plus 0.4% Triton X-100 for 30 min, blocked by incubation in PBS plus 4% normal goat serum (NGS) and 0.2% Triton X-100 for 30 min (all at room temperature, 20–23 °C), and subsequently incubated with a GFP polyclonal antibody (1:1000, Invitrogen, A-11122). Primary antiserum was diluted in PBS plus 2% NGS and 0.1% Triton X-100 overnight at 4 °C. Incubated slices were washed three times in PBS plus 1% NGS for 10 min at room temperature, incubated for 2 h at room temperature with a 1:1000 dilution of Alexa Flour 546 goat anti-rabbit IgG (1:1000, Molecular Probes, A11035) in PBS plus 2% NGS and 0.1% Triton X-100, and subsequently washed three times in PBS for 10 min at room temperature. The sections were mounted on slides and covered with cover slips. All images were acquired on a Nikon 1 confocal laser scanning microscope (Nikon). Digitalized images were analyzed using Fiji (version 2.0.0, NIMH) and Adobe Photoshop CS5 (version 12.0.4, Adobe).

Brain slice preparations and ex vivo electrophysiological recordings

Transverse slices (240 μm) containing the mPFC were prepared from 3–4 month old PV-Cre and SOM-Cre mice bilaterally injected with AAV-EF1a-DIO-eNpHR3.0-eYFP or AAV-EF1a-DIO-ChR2-eYFP into the mPFC as performed for the in vivo studies. After mice were anesthetized with isoflurane and decapitated, the brain was quickly removed and placed in ice-cold standard

artificial cerebrospinal fluid (sACSF) containing 125 mM NaCl, 2 mM KCl, 2 mM CaCl2, 1 mM MgCl2, 25 mM glucose, 25 mM NaHCO3, and 1.25 NaH2PO4, at pH 7.4, with 95% O2 and 5% CO2). Slices were cut with a vibratome (Leica, Vibratome VT1000S) in an ice cold solution containing 130 mM k-gluconate, 15 mM KCl, 0.2 mM EGTA, 20 mM HEPES, 25 mM glucose, and 2 mM kynurenic acid). Slices were then transferred for 1 min in a room-temperature solution containing 225 mM d-mannitol, 2.5 mM KCl, 1.25 mM NaH2PO4, 26 mM NaHCO3, 25 mM glucose, 0.8 mM CaCl2, and 8 mM MgCl2, with 95% O2 and 5% CO2. Finally, slices were transferred in ACSF at 30 °C for 15 min and then maintained at room temperature for the entire experiment. Brain slices were continuously perfused in a submerged chamber with a recording solution containing 120 mM NaCl, 2 mM KCl, 1 mM NaH2PO4, 26 mM NaHCO3, 1 mM MgCl2, 2 mM CaCl2, 10 mM glucose, at pH 7.4, with 95% O2 and 5% CO2). Cells were visualized with an Olympus FV1000 microscope (Olympus Optical). Electrophysiological recordings from pyramidal cells and YFP-expressing interneurons were made. Patch electrodes had resistances of 3–4 MΩ and were filled with an internal solution containing 135 mM k-gluconate, 70 mM KCl, 10 mM Hepes, 1 mM MgCl2, and 2 mM Na2ATP (pH 7.4 adjusted with KOH, 280–290 mOsm l−1). Recordings were obtained using a multiclamp- 700B amplifier (Molecular Devices). Signals were filtered at 2 kHz and acquired at 10 kHz sampling rate with a DigiData 1440A interface board and pClamp 10 software. YFP-eNpHR3.0-expressing and YFP-ChR2-expressing interneurons were visualized with a single-photon laser scanning confocal microscope (FV1000, Olympus) equipped with a laser tuned at 488 nm (to image YFP fluorescence). Specific neuron classes were confirmed by the firing pattern on intracellular current pulse injections. To characterize the inhibitory effect of the 532 nm light on YFP-eNpHR3.0-expressing interneurons, we placed a monopolar fiberoptic connected to a 532 nm laser in proximity to an mPFC slice containing YFPeNpHR3.0- expressing interneurons. Owing to the positioning of the fiberoptic tip at the border of the microscope objective lens, the distance and the angle of the light with respect to the recording area could not be optimal. We performed whole cell current clamp recordings from YFP-eNpHR3.0-expressing interneurons while injecting current pulses (4 ms duration, 2 Hz) at the intensity that was sufficient to evoke an action potential with a success rate of 100%. Then, we turned on the green light for two minutes while applying the current pulses. These procedures were repetitively applied at different laser powers and the mean laser power inhibiting the action potential firing in both SOM+ and PV+ interneurons was determined. Photostimulation of ChR2-

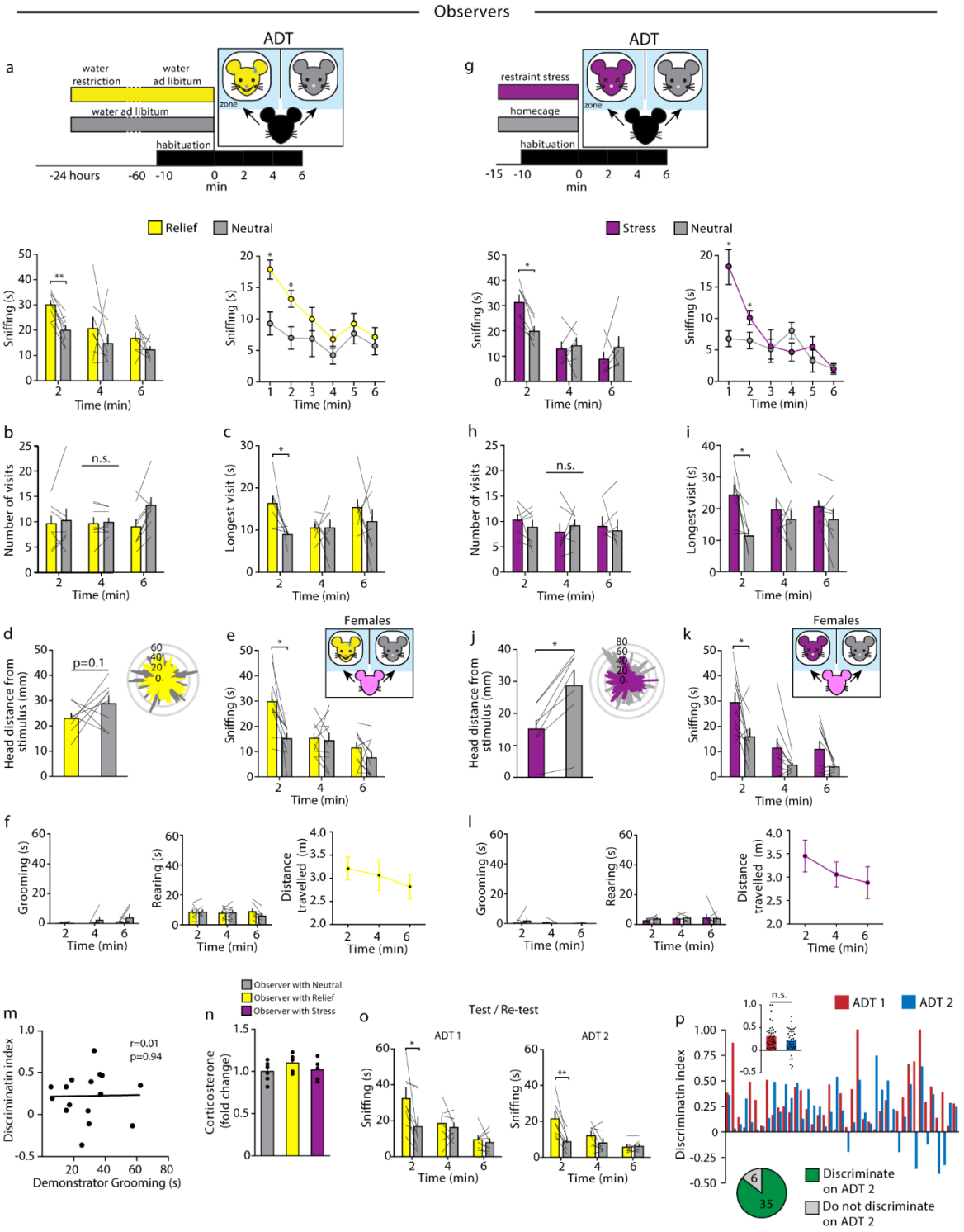
expressing SOM+ interneurons in brain slices consisted of 5 s light pulses (473 nm) delivered by a blue module laser diode (Cobolt), which was collimated and coupled under the objective with an optical fiber with a power of 2.25 mW. Patch-clamp recordings of YFP-ChR2-expressing SOM+ interneurons were performed in the cell-attached configuration and the frequency of action potentials was calculated in time bins of 1 s before, during, and after the 5 s light pulse. To study in brain slices the inhibitory effect on pyramidal cells of 5 s photostimulation of YFP-ChR2-expressing SOM+ interneurons, we used two experimental configurations. First, to study the putative GABA-mediated currents we performed whole-cell recordings from pyramidal neurons voltage clamped at −40 or −50 mV (that is, a potential that was distant from the chloride reversal potential), while turning on the blue module laser diode with an intensity of 2.25 mW. Then, we changed to a current clamp configuration and we injected onto the pyramidal cells (at −60 mV) 4 ms current pulses at suboptimal intensity that induced either a subthreshold depolarization or an action potential. The probability of successful action potentials was calculated by dividing the number of action potentials observed by the number of current pulses injected. This probability was calculated at basal conditions, during the 5 s photostimulation and after switching off the blue light.

Statistics

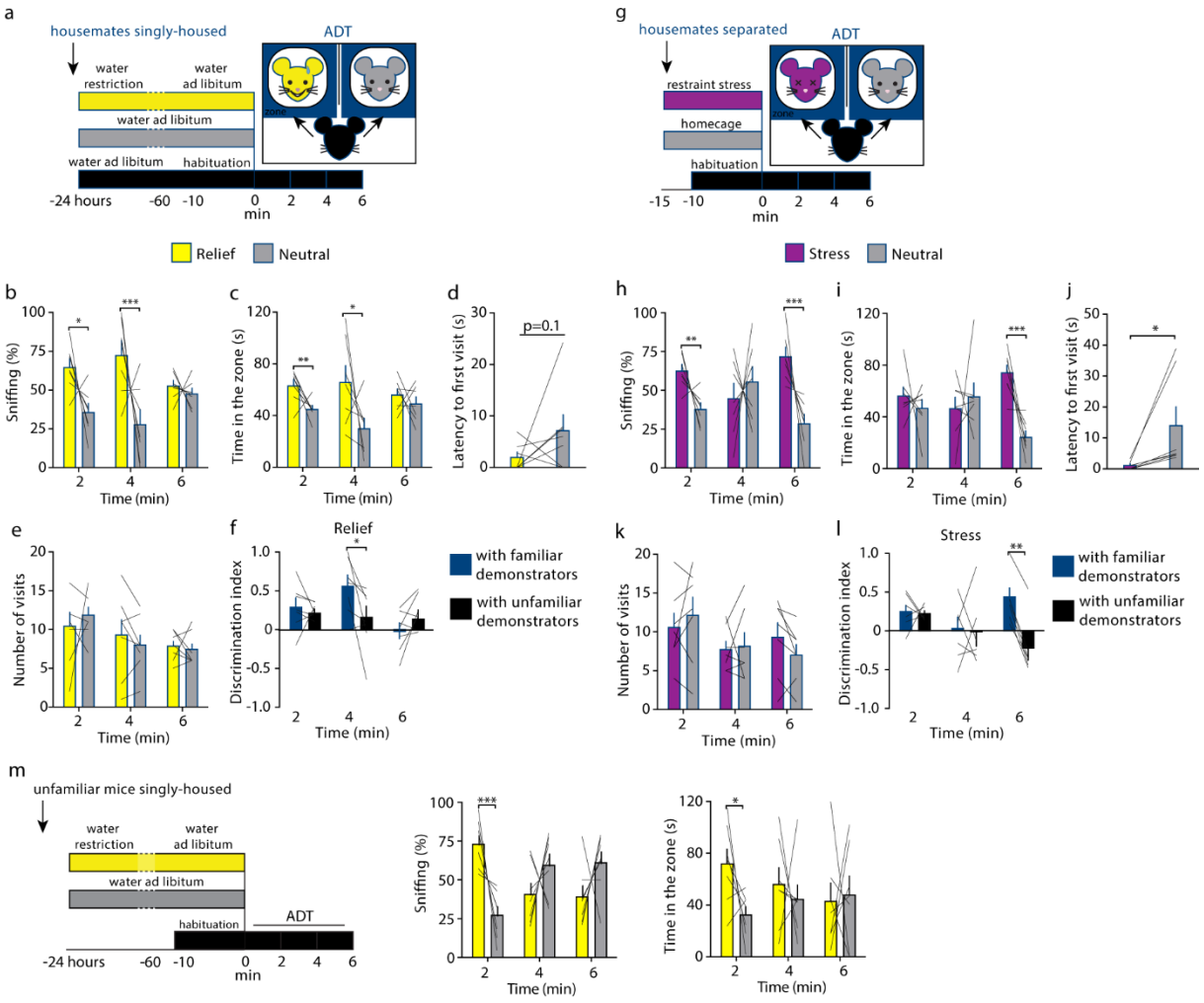
Results are expressed as mean ± s.e.m. throughout the manuscript. For the analyses of behaviors in the different tasks, the multiple *t*-test and two-way repeated-measures ANOVA were used, as appropriate, followed by Bonferroni correction. Sniffing behavior was calculated as a percentage to allow direct comparison between mice strains and different experimental conditions. The behaviors of the two demonstrators and recorded USVs were analyzed using a multiple *t*-test, followed by Bonferroni correction. The behaviors of the observer mice in the one-on-one setting were analyzed using a multiple *t*-test, followed by Bonferroni correction. For in vivo electrophysiological recording analysis we used: a 3 × 2 ANOVAs with factor epoch (three levels: baseline, pre-social exploration, and post-social exploration) and affective state (two levels: relief or stress, and neutral), followed by Bonferroni post-hoc tests; 2 × 2 ANOVAs, with factor epoch (two levels: pre-end social exploration and post-end social exploration) and affective state (two levels: relief or stress, and neutral), followed by Bonferroni posthoc tests. For the preference index an independent-samples *t*-test was used. The accepted value for significance was $P < 0.05$.

Statistical analyses were performed using GraphPad Prism 7. Numbers of mice are reported in the figure legends. Data distribution was tested using the D'Agostino and Pearson normality test. The experiments reported in this work were repeated independently two to four times, using mice from at least four different generations. No statistical methods were used to predetermine sample size for single experiments. The animal numbers were based on estimation from previous studies, including our own published studies (27,53). Mice were excluded post-hoc when optic fiber placement or viral expression patterns were not appropriate (outside the target region or week viral expression) or when implants were lost during tests lasting for multiple days. For all behavioral tests, electrophysiological recordings and calcium imaging littermates were randomly assigned to the different groups. Specific randomization in the organization of the experimental conditions is described in the results and figure legends. Experimenters were not blinded during data acquisition, but all analyses were performed with blinding of the experimental conditions.

Extended Data

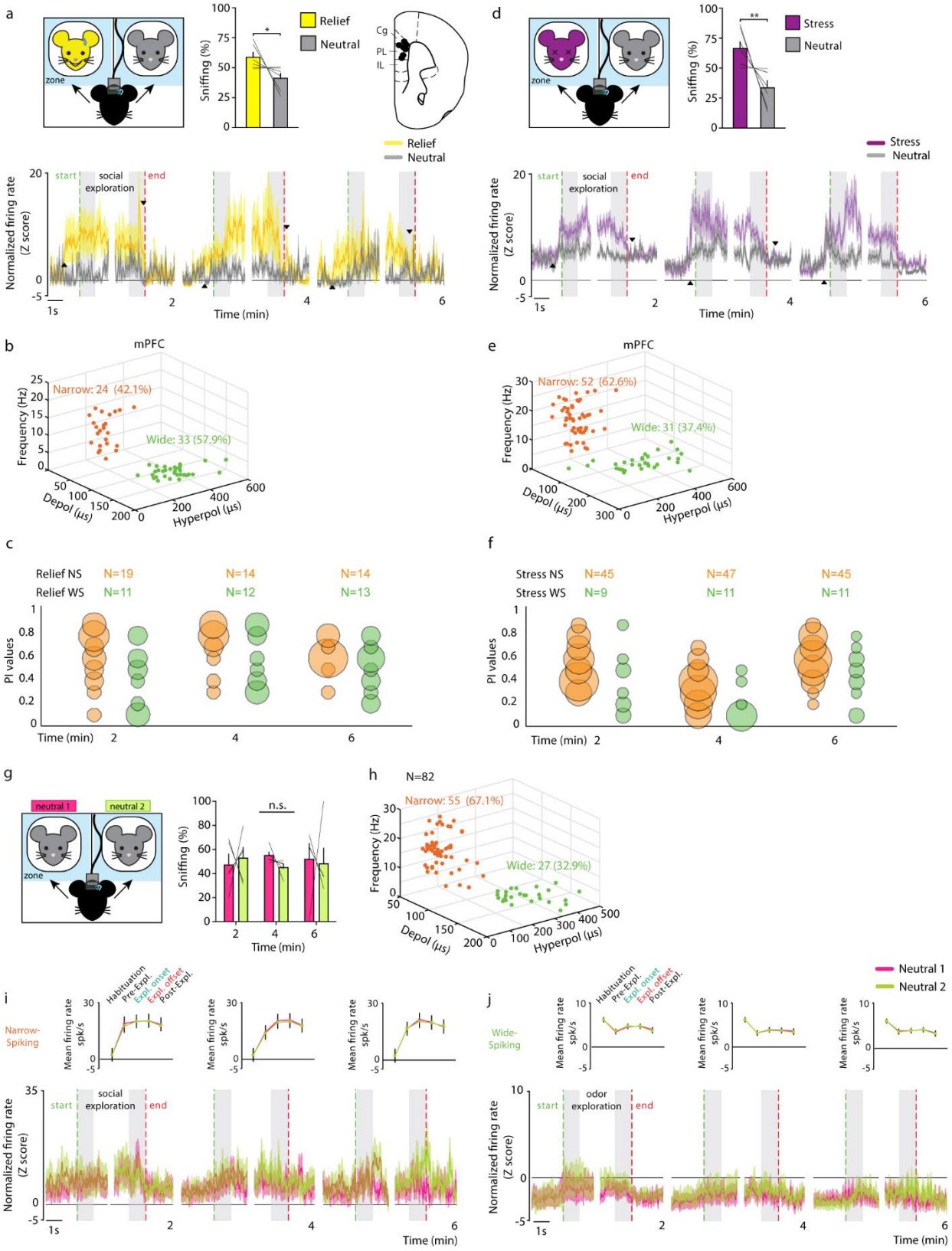


Extended Data Fig. 1 | Mice can discriminate conspecifics based on the affective state. Related to Fig. 1. a-f, Data derived from *n*=8 mice. **a**, *Top*, Experimental design of the ADT. One demonstrator was given water access for 1 hour before the test, after 23 hours of water deprivation ('relief', yellow), while the other demonstrator had *ad libitum* water access ('neutral', grey). Increased exploration (sniffing, in seconds) to the relieved compared to neutral demonstrator (left, showed in 120-seconds beams, two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.85, df=14, p=0.005; right, showed in 60-seconds beams, 60s: t=3.6, df=14, p=0.017, 120s: t=2.77, df=14, p=0.014). **b**, No difference in average number of visits to each zone (two-tailed multiple t-test, Bonferroni correction, t=0.22, df=14, p>0.999) **c**, Longer visits of observers in the zone related to the relief demonstrators (two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.89, df=14, p=0.004). **d**, Average distance of observers' head to the relieved and the neutral demonstrators did not differ during ADT (two-tailed paired t-test, t=1.22, df=7, p=0.261). **e**, In female mice increased sniffing to the relieved compared to neutral, sex-matched, demonstrator (two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.32, df=18, p=0.011; *n*=10 mice). **f**, No difference in grooming (two-tailed multiple t-test, Bonferroni correction, t=0.45, df=14, p=0.657) and rearing behaviors (t=0.00, df=14, p=0.998), and locomotor activity (one-way ANOVA, $F_{(2,21)}$=1.59, p=0.226), displayed by the observers during the ADT with the neutral and the relieved demonstrators. g-l, Data derived from *n*=7 mice. **g**, *Top*, In the stress protocol one demonstrator ('stress', purple) was subjected to restraint stress test for 15 minutes culminating in the beginning of ADT. The other demonstrator ('neutral', grey) waited undisturbed in his home-cage. *Bottom*, mice showed increased sniffing to the stressed compared to neutral demonstrator (left, showed in 120-seconds beams, two-tailed multiple t-test, Bonferroni correction, 2 min: t=2.11, df=12, p=0.05; right, showed in 60-seconds beams, 60s: t=3.67, df=10, p=0.02, 120s: t=2.12, df=10, p=0.05; *n*=6 mice). **h**, Average number of visits to each zone did not differ (two-tailed multiple t-test, Bonferroni correction, t=0.81, df=12, p>0.999). **i**, Observers made longer visits in the zone related to the stressed demonstrators (two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.46, df=12, p=0.017). **j**, Shorter average head distance of observer mice to the stressed demonstrator compared to neutral during the ADT (two-tailed paired t-test, t=5.31, df=7, p=0.001). **k**, Female mice showed increased sniffing to the stressed compared to neutral, sex-matched, demonstrators (two-tailed multiple t-test, Bonferroni correction, 2 min: t=2.69, df=18, p=0.044; *n*=11 mice). **l**, No difference in grooming (two-tailed multiple t-test, Bonferroni correction, t=0.84, df=12, p=0.708) and rearing behaviors (t=1.20, df=12, p=0.578), and locomotor activity (one-way ANOVA, $F_{(2,18)}$=0.72, p=0.498) displayed by the observers during the ADT with the neutral and the stressed demonstrators. **m**, No correlation between discrimination index, calculated for the first two minutes of ADT, and grooming behavior of the stressed demonstrators (*n*=16, Pearson's *r*=0.017, linear regression showed no significant deviation from zero, two-tailed p=0.949). **n**, No difference of average plasma corticosterone levels after ADT with two neutral demonstrators (gray), one relieved and one neutral (yellow), one stressed and one neutral demonstrator (purple) (one-way ANOVA, $F_{(2,13)}$=1.08, p=0.367). **o**, First and second testing in the same ADT ('relief') showed similar behavioral pattern with increased sniffing towards the relieved demonstrator compared to the neutral (ADT 1: two-tailed multiple t-test, Bonferroni correction, 2 min: t=2.25, df=20, p=0.035; ADT 2: t=3.99, df=20, p=0.002; *n*=11). **p**, For each observer tested in the ADT with both protocol (relief and stress) a discrimination index was calculated to compared performance on ADT 1 (red) and ADT 2 (blue; discrimination index = exploration of "relief"/"stress" - exploration of "neutral" / total time of exploration). Positive index means discrimination between "affectively-altered" and "neutral". Of 41 mice tested in ADT 1 and ADT 2 only 6 did not show a positive discrimination index on second testing. Average discrimination index did not differ between ADT 1 and ADT 2 (two-tailed unpaired t-test, t=1.73, df=80, p=0.089; *n*=41 mice). Bar and line graphs show mean ± s.e.m. * p<0.05, ** p<0.005. n.s. not significant.

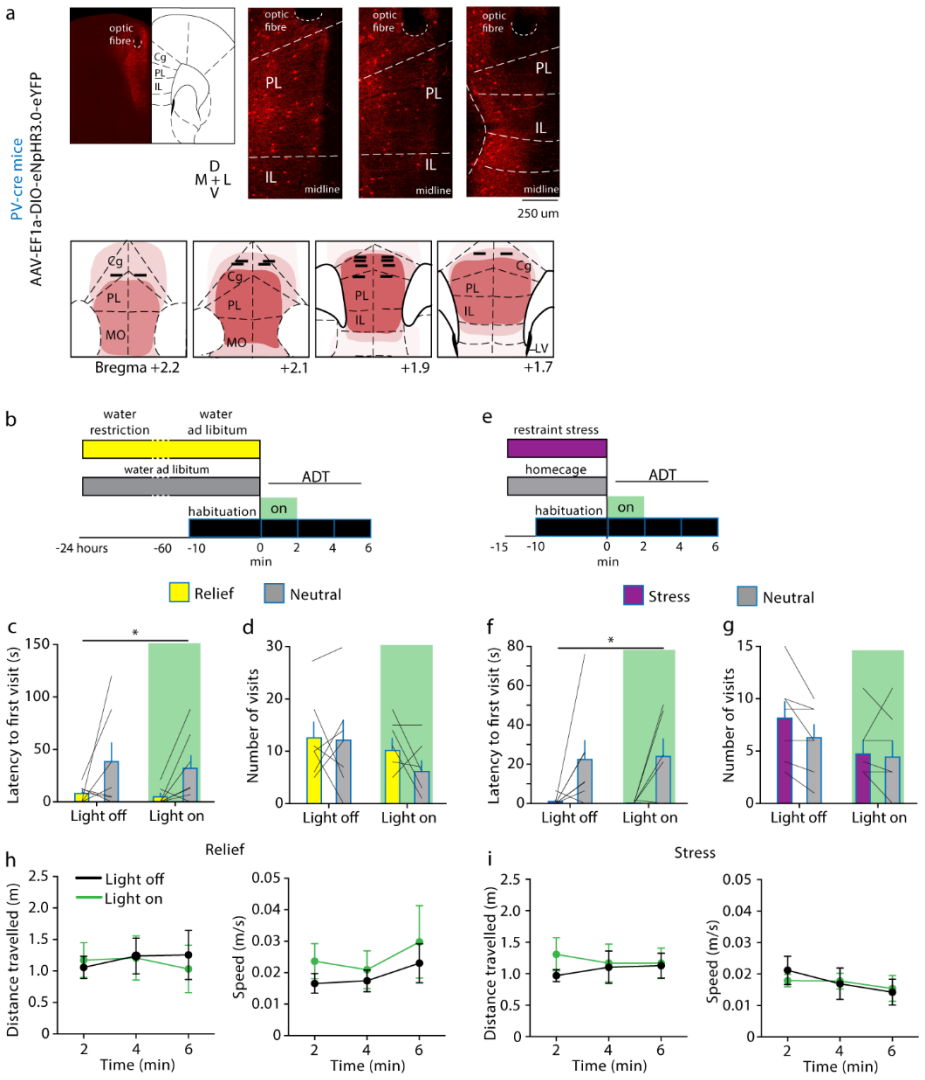


Extended Data Fig. 2 | Affective state discrimination is enhanced between familiar conspecifics. Related to Fig. 1. a-f, Data derived from *n*=7 mice. **a**, *Top*, Experimental design of the ADT with cagemates demonstrators. Observer and demonstrators were singly-housed 23 hours before testing. One demonstrator was given water access for 1 hour before the test, after 23 hours of water deprivation ("relief", yellow), while the other demonstrator had *ad libitum* water access ("neutral", gray). **b**, Increased sniffing to the relieved compared to neutral demonstrator (two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.60, df=12, p=0.010, 4 min: t=3.35, df=12, p=0.017). **c**, Increased time spent in the zone related to the relieved demonstrator compared to the neutral (two-tailed multiple t-test, Bonferroni correction, 2 min: t=3.21, df=12, p=0.022, 4 min: t=2.37, df=12, p=0.035). **d**, Latency to make the first visit to the relieved demonstrator compared to the neutral (1.95±1.0 relief, 7.14±3.0, two-tailed paired t-test, t=1.37, df=6, p=0.13). **e**, Average number of visits to each zone did not differ (two-tailed multiple t-test, Bonferroni correction: t=0.63, df=12, p>0,999). **f**, When tested with cagemates, discrimination of relieved versus neutral demonstrators was longer as discrimination index was increased compared to mice tested with unfamiliar demonstrators (two-tailed multiple t-test, Bonferroni correction, 4 min: t=2.07, df=13, p=0.05). g-l, Data derived from *n*=7 mice. **g**, In the stress protocol using cage-mates, one demonstrator ("stress", purple) was subjected to restraint stress test for 15 minutes culminating in the beginning of ADT. The other demonstrator ("neutral", grey) waited undisturbed in his home-cage.. **h**, Increased sniffing to the stressed compared to neutral demonstrator (two-tailed multiple t-test, Bonferroni correction, 2 min: t=4.27, df=12, p=0.003; 6 min: t=5.16, df=12, p=0.0007). **i**, Increased time spent in the zone related to the stressed demonstrator compared to the neutral (two-tailed multiple t-test, Bonferroni correction, 6 min: t=6.13, df=12, p=0.0001). **j**, Shorter latency to make the first visit to the stressed demonstrator compared to the neutral (paired t-test, t=2.31, df=6, p=0.05). **k**, Average number of visits to each zone did not differ (two-tailed multiple t-test, Bonferroni correction: t=0.53, df=12, p>0,999). **l**, When tested with cage-mates, discrimination of the stressed versus the neutral demonstrators was longer as discrimination index was increased compared to mice tested with unfamiliar

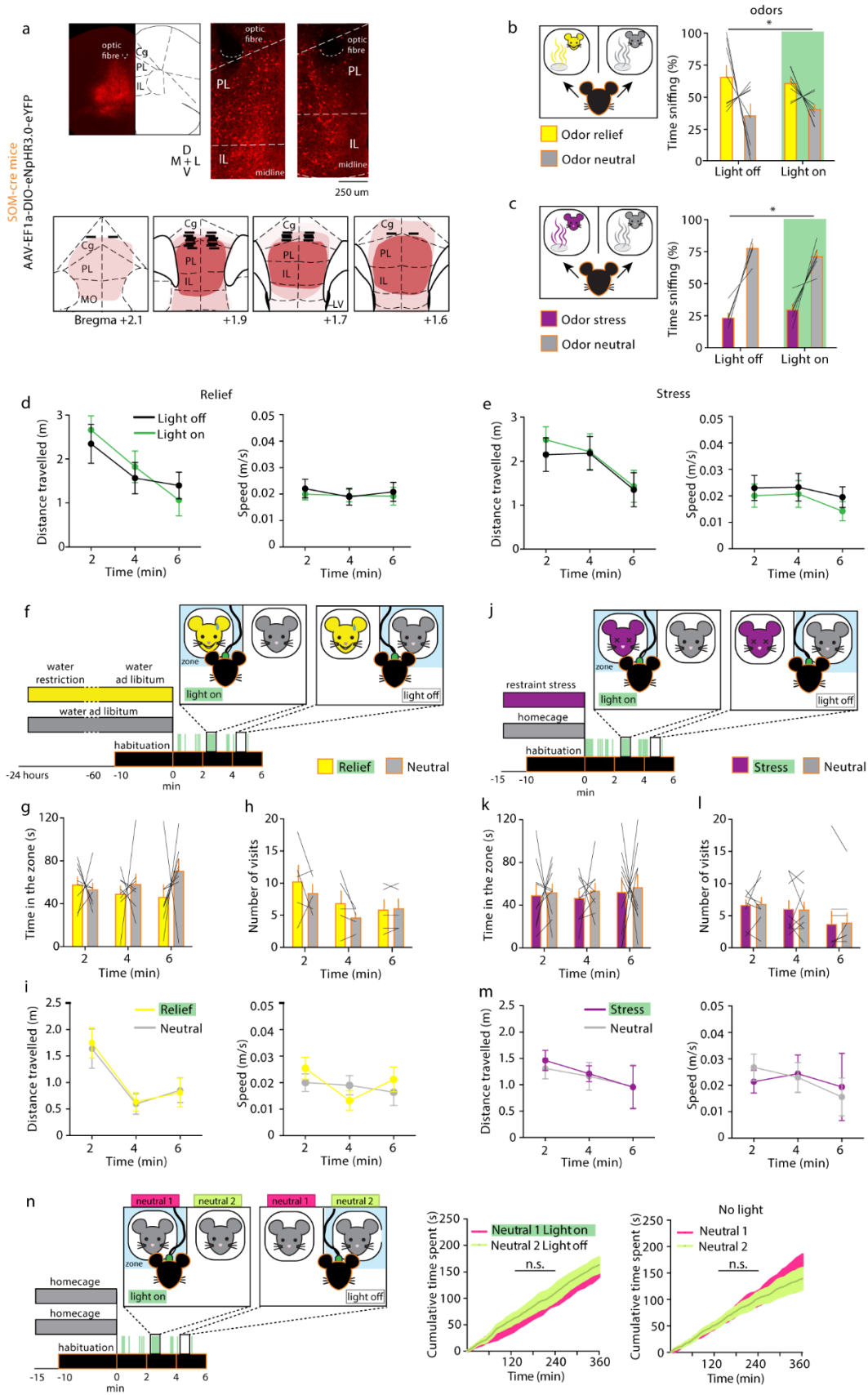
demonstrators (two-tailed multiple t-test, Bonferroni correction, 6 min: t=3.45, df=11, p=0.01). **m**, To rule out that social isolation 23 hours before testing (to allow water restriction of one cage-mate – "relief") could have affected ADT with familiar mice, we tested singly-housed observers with unfamiliar demonstrators. Mice showed increased sniffing towards the relieved demonstrators (two-tailed multiple t-test, Bonferroni correction, 6 min: t=5.48, df=12, p=0.0004, *n*=7) and increased time spent in the related zone (two-tailed multiple t-test, Bonferroni correction, 6 min: t=2.86, df=12, p=0.041), only during the first 2 minutes of ADT, and not further, as showed in (**b**). Bar graphs show mean ± s.e.m. * p<0.05, ** p<0.005. *** p<0.0005.


a
zone
Sniffing (%)
Relief
Neutral
*
Cg
PL
IL
Relief
Neutral
Normalized firing rate (Z score)
start
social exploration
end
1s
Time (min)
d
zone
Sniffing (%)
Stress
Neutral
**
Stress
Neutral
Normalized firing rate (Z score)
start
social exploration
end
1s
Time (min)
b
mPFC
Narrow: 24 (42.1%)
Wide: 33 (57.9%)
Frequency (Hz)
Depol (µs)
Hyperpol (µs)
e
mPFC
Narrow: 52 (62.6%)
Wide: 31 (37.4%)
Frequency (Hz)
Depol (µs)
Hyperpol (µs)
c
Relief NS N=19 N=14 N=14
Relief WS N=11 N=12 N=13
PI values
Time (min)
f
Stress NS N=45 N=47 N=45
Stress WS N=9 N=11 N=11
PI values
Time (min)
g
neutral 1
neutral 2
zone
Sniffing (%)
n.s.
Time (min)
h
N=82
Narrow: 55 (67.1%)
Wide: 27 (32.9%)
Frequency (Hz)
Depol (µs)
Hyperpol (µs)
i
Narrow-Spiking
Habituation
Pre-Expl.
Expl. onset
Expl. offset
Post-Expl.
Mean firing rate spk/s
Normalized firing rate (Z score)
start
social exploration
end
1s
Time (min)
j
Wide-Spiking
Neutral 1
Neutral 2
Habituation
Pre-Expl.
Expl. onset
Expl. offset
Post-Expl.
Mean firing rate spk/s
Normalized firing rate (Z score)
start
odor exploration
end
1s
Time (min)

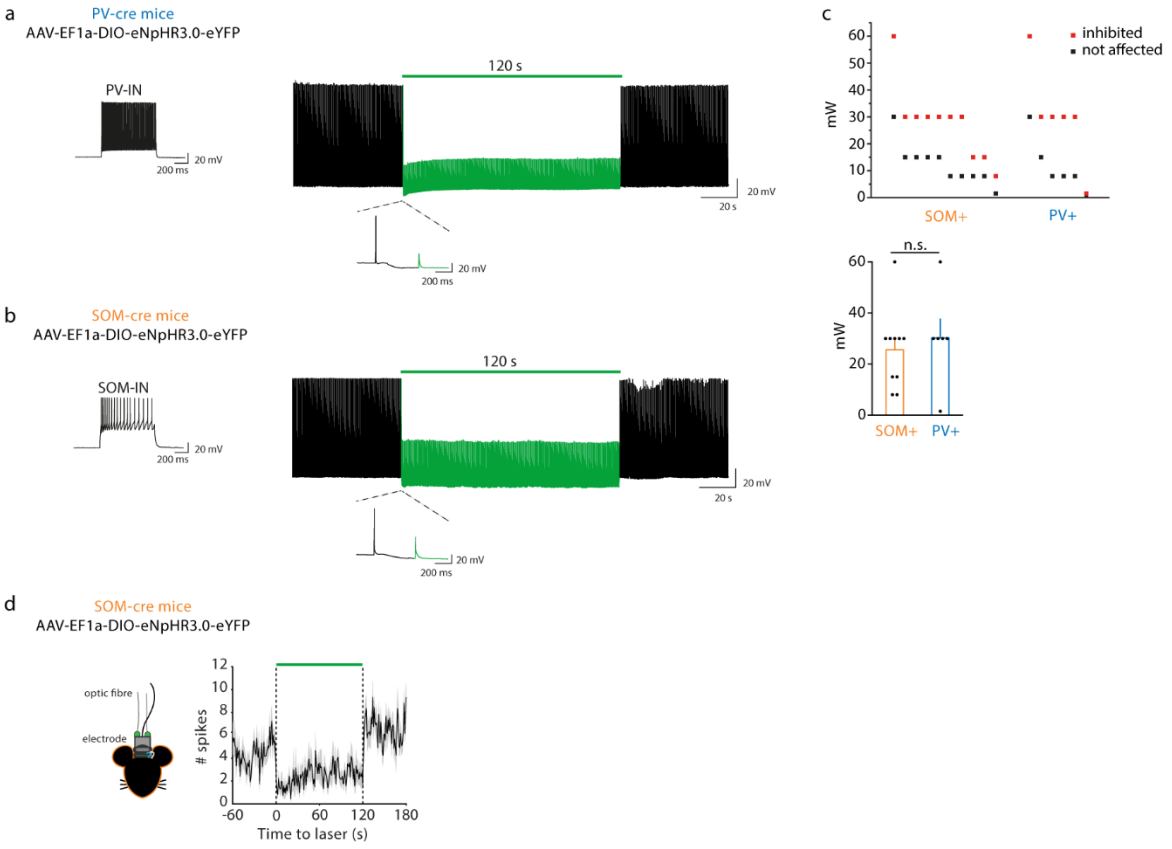
Extended Data Fig. 3 | Enhanced neuronal activity during exploration of conspecifics in an altered, but not neutral, affective state. Related to Fig. 2. **a**, *Up left*, mice implanted with chronic recording electrodes showed increased exploration of the relieved demonstrator compared to the neutral one, during the first 2 minutes of testing (two-tailed unpaired t-test: t=2.33, df=10, p=0.0418; *n*=6 mice). *Up right*, electrode placement in the mPFC (Cg, cingulate; PL, prelimbic area; IL, infralimbic area). *Bottom*, increased average population (N=57) activity during the epoch before the exploration and the epoch in which the observer explored the relieved demonstrator (yellow), compared to the pre social exploration epoch (min2: p<0.0001; min4: p<0.0001; min6: p=0.000696) and compared to exploration of the neutral (gray, min2: p=0.020031, min4: p=0.016927; min6: p=0.47572; 3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,112)}$=52.75, p<0.0001; min4: $F_{(2,112)}$ =54.75, p<0.0001; min6: $F_{(2,112)}$=24.79, p<0.0001). The black arrows indicate the time of onset (upward arrow) and end (downward arrow) of significant separation of the activity between the two compared conditions. Increased average population activity during the end social exploration epoch of the relieved demonstrator compared to that during the following epoch (post social exploration epoch, min2: p=0.045877; min4: p=0.000218; min6: p=0.049170) and compared to the neutral (min2: p=0.000275; min4: p=0.001698; min6: p=0.001408 2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1,56)}$=7.07, p=0.010161; min4: $F_{(1,56)}$=7.17, p=0.009716; min6: $F_{(1,56)}$=4.7, p= 0.034324). **b**, Hierarchical clustering method was used to separate recorded cells, during the ADT with one relieved and one neutral demonstrator, into the following populations: wide spiking (putative pyramidal cell, green, N=33) or narrow spiking (putative interneurons; orange, N=24). **c**, Frequency distribution of preference indexes (PIs) for the relief or neutral affective state in the NS (orange) and WS (green) neuronal population during the exploration of conspecifics during all the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=8). Independentsamples t-tests, min2: t=3.39015, p<0.002095; min4: t = 2.1174, p=0.04478; min6: 2.2808, p=0.031349. **d**, *Up*, mice showed more exploration of the stressed demonstrator than of the neutral demonstrator during the first 2 minutes of testing (*n*=7 mice, two-tailed unpaired t-test: t=4.12, df=12, p=0.0014). *Bottom*, increased average population (n=83) activity during the start social exploration epoch and the social exploration epoch towards the stressed demonstrator (purple), compared to the pre social exploration epoch (min2: p<0.0001; min4: p=0.000003; min6: p<0.0001) and compared to exploration of the neutral (gray, min2: p=0.006666; min4: p=0.001223; min6: p=0.000027; 3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,164)}$=45.13, p<0.0001; min4: $F_{(2,164)}$=60.42, p<0.0001; min6: $F_{(2,164)}$=74.44, p<0.0001). Increased average population activity during end social exploration epoch of the stressed demonstrator compared to the following epoch (post social exploration epoch, min2: p<0.0001; min4: p=0.000002; min6: p=0.000001) and to the exploration of the neutral demonstrator (min2: p<0.0001; min4: p<0.0001; min6: p<0.0001; 2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1,82)}$=116.08, p<0.0001; min4: $F_{(1,82)}$=48.95, p<0.0001; min6: $F_{(1,82)}$=21.86, p<0.0001). **e**, Classification of recorded cells, during the ADT with one stressed and one neutral demonstrator, in NS (N=52) and WS as described in b. **f**, Frequency distribution of preference indexes (PIs) for the stress or neutral affective state in the NS (orange) and WS (green) neuronal population during the exploration of conspecifics during all the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=13). Independent-samples two-tailed t-tests, min2: t = 2.62995, p=0.011206; min4: t = 4.55371, p=0.000029; min6: t = 3.43784, p=0.001137. **g**, Mice have been implanted with recording electrodes in the mPFC and tested in the ADT with two naïve "neutral" demonstrators. Mice equally explored the two demonstrator and did not show observable discrimination. **h**, Classification of the recorded cells (N=82) in NS (N=55) and WS cells (N=27) as described in b during the ADT with two neutral demonstrators. **i**, *Top*, population responses calculated as an average of the activity of NS neurons (N=55) during "Habituation", "Pre-exploration", "Exploration onset", "Exploration offset", and "Post-exploration" periods towards neutral demonstrator 1 (pink) and neutral demonstrator 2 (green) (mean ± s.e.m.). *p<0.05 versus exploration of the neutral mouse. *Bottom*, normalized population activity of NS neurons (n=55) during exploration of both neutral demonstrators was stronger compared to the entire pre-exploration period (min2: p=0.0417281; min4: p=0.016125; min6: p=0.047255) without any difference between the two mice (min2: p=0.135825; min4: p=0.789647; min6: p=0.666470; 3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,108)}$=0.61, p=0.55; min4: $F_{(2,108)}$=0.67, p=0.51; min6: $F_{(2,108)}$=2.06, p=0.07). Other legends as in (c). **j**, Same analyses as in (i) but for WS (N=27). *Top*, population responses calculated as an average of the activity of WS neurons during "Habituation", "Pre-exploration", "Exploration onset", "Exploration offset", and "Post-exploration periods" towards two neutral demonstrators (mean ± s.e.m.). *p<0.05 versus exploration of the neutral mouse. *Bottom*, Normalized population activity of WS neurons (n=27) during exploration of both neutral demonstrators was stronger compared to the pre-exploration period (min2: p=0.028877; min4: p=0.0499772; min6: p=0.00326), without any difference between the two mice (min2: p=0.356287; min4: p=0.598312; min6: p=0.583072; 3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,52)}$=0.9, p=0.4; min4: $F_{(2,52)}$=0.8, p=0.45; min6: $F_{(2,52)}$=2.8, p=0.07).



Extended Data Fig. 4 | Photoinhibition of PV+ interneurons does not affect affective state discrimination. Related to Fig. 3. **a**, *Top*, representative images of viral expression in the mPFC after injection with AAV-EF1a-DIO-eNpHR3.0-eYFP. *Bottom*, reconstruction of viral expression and location of optical fibers. Red areas represent the expression (higher expression = darker color) of AAV-EF1a-DIO-eNpHR3.0-eYFP in PV-cre mice. Findings were replicated in two independent experiments with similar results. Data derived from *n*=7 mice. **b**, PV-cre mice were tested in the ADT with one relieved and one neutral demonstrator. Photo-inhibition was performed for 2 minutes, from the beginning of the test, using continuous green light. **c**, No effect of PV+ photoinhibition on latency to made the first visit to the relieved demonstrator (two-way RM ANOVA, affective state (relief, neutral): $F_{(1,12)}$=7.18, p=0.020). **d**, Optical inhibition of PV+ did not modify the number of visits to each demonstrator (two-way RM ANOVA, affective state (relief, neutral) x light (off,on): $F_{(1,24)}$=0.38, p=0.541). **e**, PV-cre mice were tested in the ADT with one stressed and one neutral demonstrator. Photoinhibition was performed for 2 minutes, from the beginning of the test, using continuous green light. Data derived from *n*=7 mice. **f**, No effect of PV+ photoinhibition on latency to made the first visit to the stressed demonstrator (two-way RM ANOVA, affective state (stress, neutral): $F_{(1,12)}$=12.75, p=0.003). **g**, Optical inhibition of PV+ did not modify the number of visits to each demonstrator (two-way RM ANOVA, affective state (stress, neutral) x light (off,on): $F_{(1,24)}$=0.30, p=0.587). **h** and **i**, Optical inhibition of PV+ did not induce gross motor deficits during ADT in both relief (two-tailed multiple t-test, Bonferroni correction, distance travelled: t=0.35, df=12, p=0.730; average speed: t=1.10, df=12, p=0.290) and stress conditions (distance travelled: t=1.19, df=12, p=0.254; average speed: t=0.47, df=12, p=0.640). Bar and line graphs show mean ± s.e.m. * p<0.05.

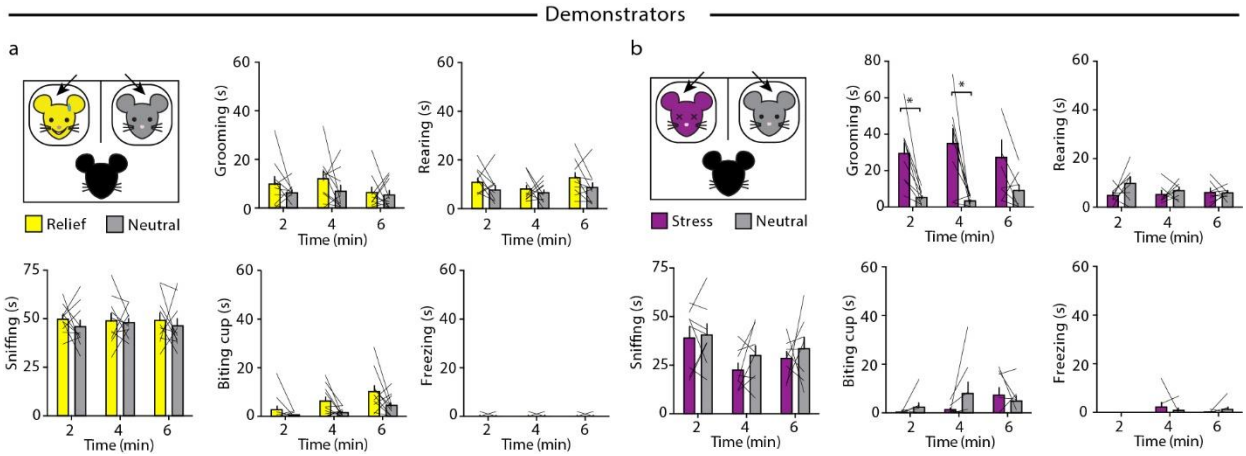

a
SOM-cre mice
AAV-EF1a-DIO-eNpHR3.0-eYFP
optic fibre
Cg
PL
IL
MO
LV
midline
D
M+L
V
250 um
Bregma +2.1
+1.9
+1.7
+1.6
b
odors
Odor relief
Odor neutral
Time sniffing (%)
Light off
Light on
c
Odor stress
Odor neutral
d
Relief
Light off
Light on
Distance travelled (m)
Speed (m/s)
Time (min)
e
Stress
f
water restriction
water ad libitum
habituation
-24 hours
min
zone
light on
light off
Relief
Neutral
j
restraint stress
homecage
Stress
g
Time in the zone (s)
h
Number of visits
k
l
i
m
n
neutral 1
neutral 2
Neutral 1 Light on
Neutral 2 Light off
n.s.
Cumulative time spent (s)
No light
Neutral 1
Neutral 2

Extended Data Fig. 5 | Photoinhibition of SOM+ interneurons abolishes affective state discrimination. Related to Fig. 4. **a**, *Top*, representative images of viral expression in the mPFC (in rostro-caudal order) after injection with AAV-EF1a-DIO-eNpHR3.0-eYFP. *Bottom*, reconstruction of viral expression and location of optical fibers. Red areas represent the expression (higher expression = darker color) of AAV-EF1a-DIO-eNpHR3.0-eYFP in SOM-cre mice. Findings were replicated in four independent experiments with similar results. **b**, Increased exploration toward the odor of the relieved demonstrators compared to the neutral in light off and light on conditions (two-way RM ANOVA, affective state (relief, neutral): $F_{(1,14)}$=12.65, p<0.005, *n*=8 mice), and no effects of SOM+ photo-inhibition ($F_{(1,14)}$=0, p=0.99). **c**, Avoidance of the odor of the stressed demonstrators in in light off and light on conditions (two-way RM ANOVA, affective state (stress, neutral): $F_{(1,10)}$=76.21, p<0.0001, n=6 mice), which SOM+ photo-inhibition did not change ($F_{(1,10)}$=0.33, p=0.576). d and e, SOM+ photoinhibition with continuous green light for two minutes did not induce any gross motor change in both relief (twotailed multiple t-test, Bonferroni correction, distance travelled: t=0.57, df=12, p=0.576; average speed: t=0.50, df=12, p=0.625) and stress conditions (distance travelled: t=0.69, df=12, p=0.498; average speed: t=0.45, df=12, p=0.658). Data derived from *n*=7 mice in each condition (relief, stress). **f**, Exploration of the relieved demonstrator was paired to SOM+ photo-inhibition throughout the test (*n*=8 mice). **g**, No preference to spend more time with the relieved demonstrator during photo-inhibition of SOM+, on the first two minutes of ADT (two-tailed multiple t-test, Bonferroni correction: t=0.47, df=14, p>0.999). **h**, No change of number of visits to each demonstrator during photoinhibition of SOM+ (two-tailed multiple t-test, Bonferroni correction: t=0.88, df=14, p>0.999). **i**, SOM+ photoinhibition paired to exploration of the relieved demonstrators did not induce any gross motor changes. **j**, Exploration of the relieved demonstrator was paired to SOM+ photo-inhibition throughout the test (*n*=9 mice). **k**, No difference in time spent with the two demonstrators during photoinhibition of SOM+ (two-tailed multiple t-test, Bonferroni correction: t=0.19, df=16, p>0.999). **l**, No difference of number of visits to stressed and neutral demonstrator during inhibition of SOM+ (two-tailed multiple t-test, Bonferroni correction: t=0.11, df=16, p>0.999). **m**, SOM+ photoinhibition paired to exploration of the stressed demonstrators did not induce any gross motor changes. **n**, Exploration of one naïve "neutral" demonstrator ("neutral 1") was paired to SOM+ photoinhibition throughout the ADT (counterbalanced, left or right, across observers, continuous green light). Data derived from *n*=8 mice. SOM+ photoinhibition did not induce social discrimination or avoidance conditions (two-way RM ANOVA, time x light (on,off): $F_{(359,2513)}$=0.21, p>0.999). No discrimination of the two neutral demonstrators without light stimulation ("No light", two-way RM ANOVA, time x light (on,off): $F_{(359,2513)}$=0.19, p>0.999). Bar and line graphs show mean ± s.e.m. * p<0.05.

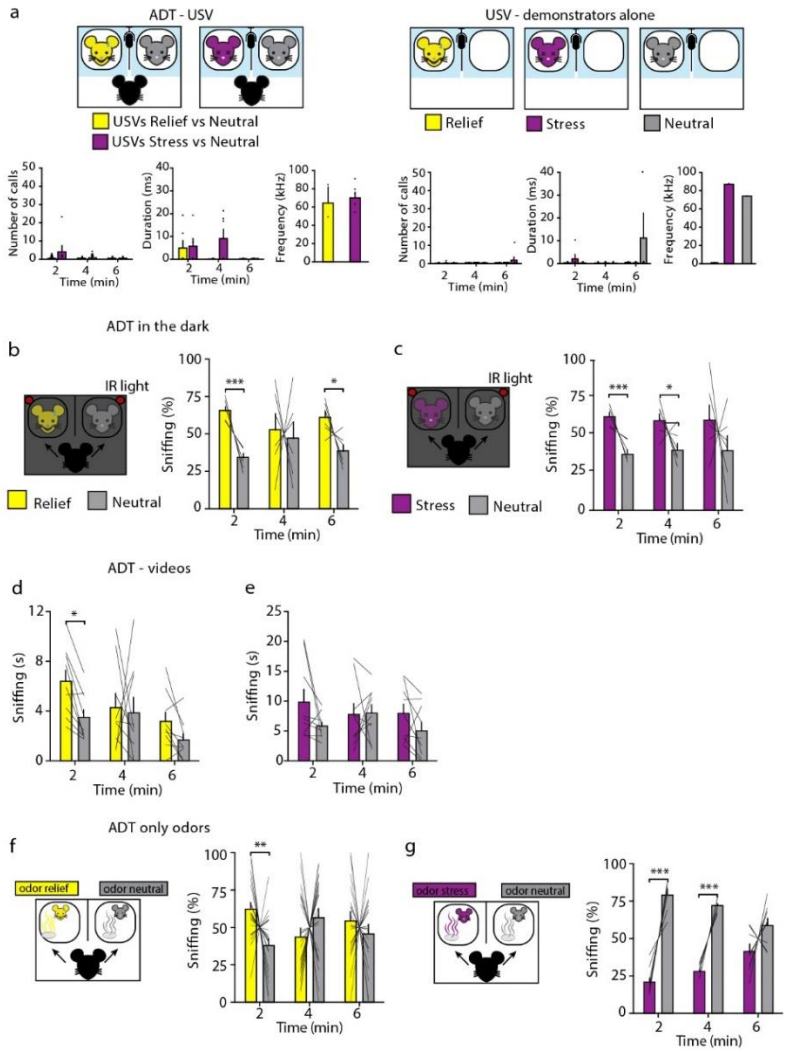


Extended Data Fig. 6 | Optogenetic light pulses inhibit AP firing in YFP-eNpHR3.0-expressing SOM+ and PV+ interneurons. a, Left, AP firing induced in an YFP-eNpHR3.0-expressing PV+ interneuron from a brain slice by a steady-state depolarizing current pulse (800 ms duration). Right, representative experiment showing the APs successfully evoked by 4 ms current pulses (injected at 2 Hz) in the absence of 532 nm light (black trace) and the depolarizations evoked by 4 ms current pulses during the 120 s 532 nm light pulse (green line) which fail to reach AP threshold (red trace). *Inset*, enlarged time-scale showing an AP (black trace) in the absence and a subthreshold depolarization (red trace) in the presence of green light. **b**, Same as in (**a**), but from an YFP-eNpHR3.0-expressing SOM+ interneuron. **c**, Summary of the laser power (mW) that inhibits (as in **a**, **b**) or fails to inhibit the AP firing for each interneuron (n=10 SOM+ cells from n=2 mice, n=6 PV+ cells from n=2 mice). The bar chart shows the mean (+/- s.e.m.) laser power inhibiting AP firing from the recorded cells (SOM+ cells, 27.8 ± 4.45; PV+ cells, 30.25 ± 7.55; Mann-Whitney Rank Sum Test, p=0.708). Error bars show s.e.m. **d**, Optogenetic-assisted tetrode recordings in the mPFC in AAV-EF1a-DIO-eNpHR3.0-eYFP injected SOM-cre mice; representative example (one neuron with 11 stimulations) of a peristimulus time histogram showing the firing rate (mean +/- s.e.m.) of an inhibited putative SOM interneuron, before during and after the delivery of the green light pulse (for 120 seconds).

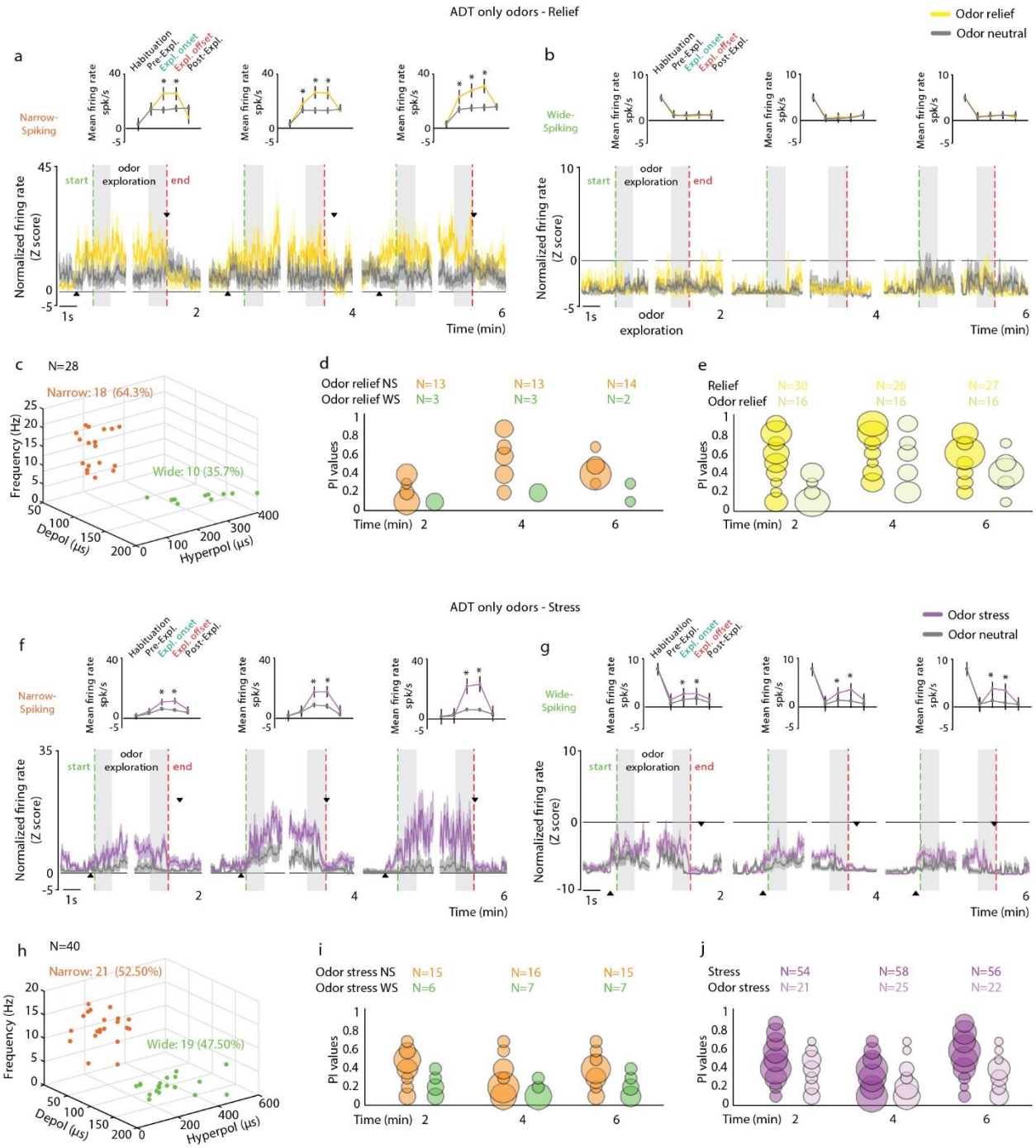
Supplementary Figures



Supplementary Figure 1. Observable behaviors of demonstrator mice. Related to Figure 1. a, Observable behaviors displayed by the neutral and relieved demonstrator mice during the 6 minutes of the ADT, divided by three consecutive 2-minute time beans. Data derived from n=10 demonstrator mice per group. No significant affective state-by-time statistical interaction was evident for sniffing (two-way RM ANOVA, $F_{(2,36)}$=0.08, p=0.916), grooming ($F_{(2,36)}$=0.34, p=0.712), rearing ($F_{(2,36)}$=0.31, p=0.732), biting ($F_{(2,36)}$=0.84, p=0.439), and freezing ($F_{(2,36)}$=0.00, p>0.999). **b**, Observable behaviors displayed by the neutral and stressed demonstrator mice during the ADT. Data derived from n=8 demonstrator mice per group. Stressed demonstrators showed increased grooming behavior compared to neutral ones (two-way RM ANOVA, affective state (stress, neutral) $F_{(1,18)}$=5.83, p=0.026). No significant affective state-by-time statistical interaction was evident for sniffing (two-way RM ANOVA, $F_{(2,28)}$=0.24, p=0.781), rearing ($F_{(2,28)}$=1.88, p=0.171), biting ($F_{(2,28)}$=2.30, p=0.118), and freezing ($F_{(2,28)}$=0.87, p=0.426). Bar graphs show mean ± s.e.m. *p<0.05.

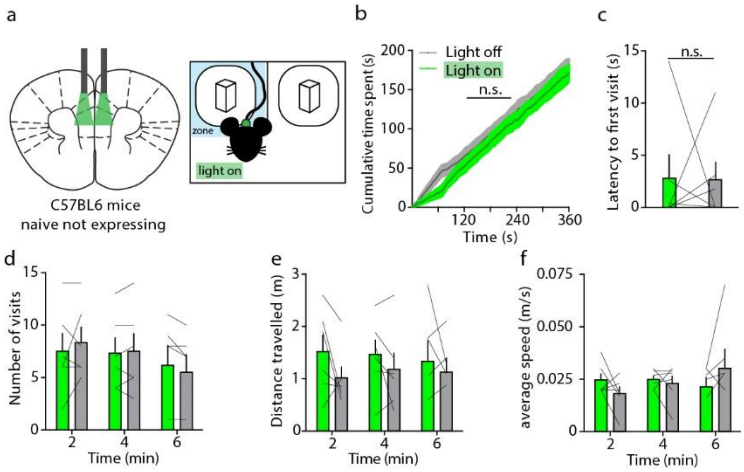


Supplementary Figure 2. Sensory modalities involved in the ADT. Related to Figure 1. a, Left, we measured ultra-sonic vocalization (USV) during the ADT with a relieved and a neutral (yellow), and with a stressed and a neutral (purple) demonstrator. Data derived from n=6 per group. The number (two-way RM ANOVA, affective state x time, $F_{(2,20)}$=0.84, p=0.442) and the duration of calls ($F_{(2,20)}$=1.77, p=0.195) was negligible and did not differ between conditions. Frequency (in Hz) of USV calls collected from two mice in relief and four mice in stress ADT. Right, we measured USV of a relieved, a stressed and a neutral demonstrator, separately and without observer. The number (two-way RM ANOVA, affective state x time, $F_{(4,30)}$=1.04, p=0.400) and the duration of calls ($F_{(4,30)}$=1.08, p=0.380) were negligible and did not differ between conditions. USV frequency collected from one mouse in each condition. **b and c**, We tested observers in the ADT in the darkness using either an infrared light (IR) to allow camera recording or using a thermocamera. All the other settings remained the same as described in Figure 1. Data derived from n=7 per group. Observers showed increased sniffing towards both the relieved (b) (Multiple t-test, Bonferroni correction, 2 min: t=7.15, df=10, p=0.00009; 6 min: t=3.59, df=10, p=0.014) and (c) the stressed demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=6.13, df=10, p=0.0003; 4 min: t=3.08, df=10, p=0.034). **d and e**, Video of demonstrators were recorded after relief or stress manipulation and neutral mice, these videos were presented to the observer during the ADT for 6 minutes. Data derived from n=9 per group d, Observers made more sniffing nearby the video of the relief demonstrator compared to the neutral (Multiple t-test, Bonferroni correction, 2 min: t=2.57, df=16, p=0.020), (e) while we found no difference presenting videos of the stress and neutral demonstrators (Multiple t-test, Bonferroni correction, 2 min: t=1.76, df=16, p=0.291). **f**, Mice showed a preference for exploration of the odor of the relieved demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=3.57, df=42, p=0.001; n=22 mice). **g**, Mice showed a marked avoidance for the odor of the stressed demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=12.35, df=14, p<0.0001; 4 min: t=20.89, df=14, p<0.0001; n=8 mice). Bar graphs show mean ± s.e.m. * p<0.05, ** p<0.005. *** p<0.0005.

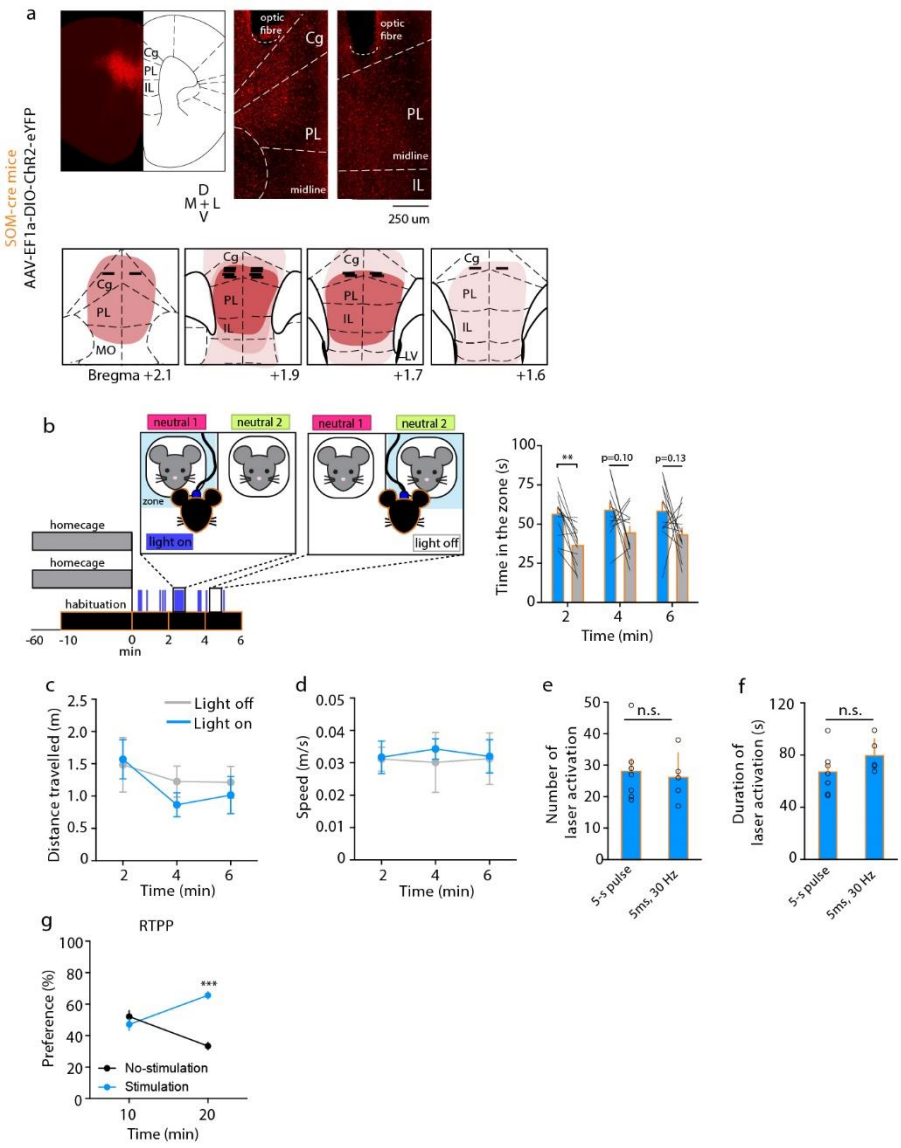


Supplementary Figure 3. Exposure to demonstrator odors only, did not recapitulate the activity pattern elicited by ADT. Related to Figure 2. a, Top, population responses calculated as an average of the activity of NS neurons (N=18) during "Habituation", "Pre-exploration", "Exploration onset", "Exploration offset", and "Post-exploration" periods towards the odor of a relieved or neutral demonstrator (mean ± s.e.m.). *p<0.05 versus exploration of the neutral odor. Bottom, population firing rate of NS neurons normalized on the habituation during the pre-odor exploration, "start" odor exploration (green dashed line), "odor exploration", "end" odor exploration (red dashed line), and post-odor exploration epochs (see methods). The firing rate increased when starting to sniff the relieved odor,

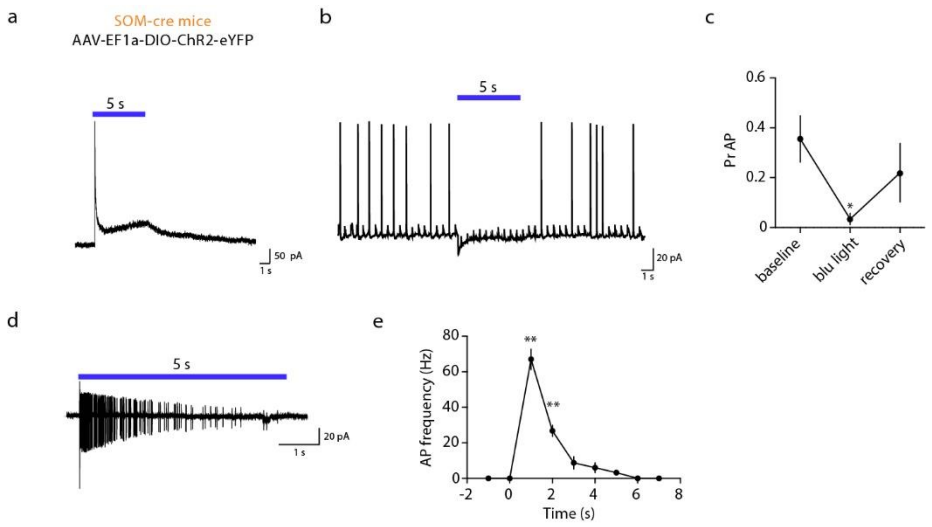
compared to the pre odor exploration in all 6-minute test (min2: p=0.000015; min4: p=0.000578; min6: p=0.000136) and compared to firing activity when sniffing the neutral one in all 6-minute test (min2: p=0.000085; min4: p=0.000639; min6: p=0.0043075; 3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,34)}$=7.83, p=0.001588; min4: $F_{(2,34)}$=11.18, p=0.000186; min6: $F_{(2,34)}$=18.47, p=0.000004). The increased firing rate during the exploration of the relieved odor dropped after the end of the sniffing (2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1,17)}$=46.96, p=0.000003; min4: $F_{(1,17)}$=7.18, p=0.0158; min6: $F_{(1,17)}$=11.01, p=0.0040). Colored shaded regions around each line represent 1 s.e.m. Gray shaded areas are used for statistical analysis of the population responses. The black arrows indicate the time of onset (upward arrow) and end (downward arrow) of significant separation of the activity between the two compared conditions. **b**, Same analyses as in (a) but for WS cells (N=10). Top, Population of WS neurons (N=10) did not show any preference for one of the two odors. Bottom, population firing rate of WS neurons normalized on the habituation during the pre-odor exploration, start odor exploration, odor exploration, end odor exploration, and post-odor exploration epochs. This population of neurons did not show any increase of their activity but only in the last session of the test during the exploration of both relieved and neutral odor compared to the pre-odor exploration epoch (min2: p=0.5365; min4: p=0.1977; min6: p=0.0385, 3x2 RM ANOVA; exploration epoch and affective state, min2: $F_{(2,18)}$=0.64, p=0.5368; min4: $F_{(2,18)}$=1.64, p=0.2210; min6: $F_{(2,18)}$= 2.26, p=0.1363). The increased firing rate during the exploration of both relieved and neutral odor dropped after the end of the sniffing in in the last 2 min of the test (min2: p=0.3207; min4: p=0.2484; min6: p=0.0362; 2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1, 9)}$=1.10, p=0.3207; min4: $F_{(1, 9)}$=1.37, p=0.27; min6: $F_{(1,39)}$=0.16, p=0.70). **c**, We recorded 28 units during the test with relief and neutral odors, 18 were classified into NS and 10 into WS. d, Frequency distribution of preference indexes (PIs) for the relief or neutral affective state in the NS (orange) and WS (green) neuronal populations during odors exploration throughout the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=9). Independent-samples two-tailed t-tests, relieved and neutral NS vs relieved and neutral WS, min2: t=2.20925, p=0.0432; min4: t=2.62086, p=0.0201; min6: t=3.24046, p=0.0059. **e**, Frequency distribution of PIs for the affective or neutral affective state during the exploration of a relieved demonstrator ì (yellow) and only its odor (light yellow) during all the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=12). Independent-samples two-tailed t-tests, relieved and neutral demonstrators vs relieved and neutral odor, min2: t=3.85401, p=0.0003; min4: t = 2.21293, p=0.0326; min6: t=3.71309, p=0.0006. **f**, Top, population responses calculated as an average of the activity of NS neurons (N=21) during “Habituation”, “Pre-exploration”, “Exploration onset”, “Exploration offset”, and “Post-exploration” periods towards odor of a stressed or neutral demonstrator (mean ± s.e.m.). *p<0.05 versus exploration of the neutral odor. Bottom, population firing rate of NS neurons normalized on the habituation. The firing rate increased when starting to sniff both stressed and neutral odors, compared to the pre-exploration epoch in all 6-minute test (min2: p=0.0047; min4: p=0.000007; min6: p<0.0001) and the intensity of discharge was stronger when sniffing the stressed odor compared to the neutral one in all 6-minute test (3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,40)}$=12.06, p=0.000082; min4: $F_{(2,40)}$=14.34, p<0.00001; min6: $F_{(2,40)}$=13.43, p=0.000035). The increased firing rate during the exploration of both the odors dropped after the end of the sniffing (min2: p=0.000007; min4: p=0.0112; min6: p=0.0385), and also in this case the activity of the neural population was higher in the case of sniffing the stressed odor compared to the neutral (2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1,20)}$=5.05, p=0.0361; min4: $F_{(1,20)}$=10.40, p=0.0042; min6: $F_{(1,20)}$=19.01, p<0.0001). **g**, Same analyses as in (f) but for WS cells (N=19). The firing rate increased when starting to sniff both of the odors, compared to the pre-exploration epoch in all 6-minute test (min2: p=0.0125; min4: p=0.0043; min6: p=0.0143), although the intensity of the discharge was higher for the stressed odor (3x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(2,36)}$=8.28, p=0.001099; min4: $F_{(2,36)}$=26.31, p<0.0001; min6: $F_{(2,36)}$=13.07, p=0.000054). The increased firing rate during the exploration of the stress and neutral odor dropped after the end of the sniffing (min2: p=0.0031; min4: p=0.000009; min6: p=0.000004), especially when the mice sniffed the stressed odor (2x2 RM ANOVA, exploration epoch and affective state, min2: $F_{(1, 18)}$=9.37, p=0.0067; min4: $F_{(1,18)}$=33.26, p=0.000023; min6: $F_{(1,18)}$=20.98, p=0.0002). **h**, We recorded 40 units, 21 were classified into NS and 19 into WS. **i**, Frequency distribution of preference indexes (PIs) for the stress or neutral affective state in the NS and WS neuronal populations during the exploration of odors during all the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=7). Independent-samples two-tailed t-tests, stressed and neutral NS vs stressed and neutral WS, min2: t=3.14889, p=0.0052; min4: t=2.77003, p=0.0108; min6: t=2.64569, p=0.0155). **j**, Frequency distribution of preference indexes (PIs) for the affective or neutral affective state during the exploration of a stress demonstrator (purple) and a stress odor (light purple) during all the test. The size of each circle is proportional to the number of single neurons (from N=1 to N=14). Independent-samples two-tailed t-tests, stress and neutral demonstrators vs stress and neutral odor, min2: t= 2.89848, p=0.0049; min4: t=3.66078, p=0.000223; min6: t=5.64016, p=0.00001).



Supplementary Figure 4. Green light delivery to the mPFC does not induce place avoidance. Related to Figure 4. a, Naïve C57BL6 mice were implanted bilaterally with fiberoptic implants and tested in the ADT setting with two objects, for 6 minutes. Exploration of one object was paired to continuous green light delivery. Data derived from n=6 mice. **b**, Illumination of mPFC with green light did not induce place avoidance as mice spent similar time with both objects (two-way RM ANOVA, time x light (on,off): $F(359,3590)=0.61$, $p>0.999$). **c**, Illumination of mPFC with green light did not modify latency to make the first visit (two-tailed paired t-test: $t=0.05$, $df=5$, $p=0.958$), **d**, number of visits to each zone (two-tailed multiple t-test, Bonferroni correction: $t=0.37$, $df=10$, $p>0.999$) and (e and f) did not induce gross motor deficits (two-tailed multiple t-test, Bonferroni correction, distance travelled: $t=1.24$, $df=10$, $p=0.721$; average speed: $t=1.39$, $df=10$, $p=0.576$). Bar and line graphs show mean ± s.e.m.



Supplementary Figure 5. Photostimulation of SOM+ interneurons guides social discrimination. Related to Figure 5. a, Top, representative images of viral expression in the mPFC after injection with AAV-EF1a-DIO-ChR2-eYFP. Bottom, reconstruction of viral expression and location of optical fibers. Red areas represent the expression (higher expression = darker color) of AAV-EF1a-DIO-ChR2-eYFP in SOM-cre mice. Findings were replicated in three independent experiments with similar results. Data derived from n=12 mice. **b**, Mice were tested in the ADT with two, naïve non-manipulated, neutral demonstrators. Photostimulation was paired to exploration of one of the two demonstrators (counterbalanced, left or right, across observers). Mice spent more time in the zone related to the mouse paired with SOM+ activation (Multiple t-test, Bonferroni correction, 2 min: t=3.95, df=22, p=0.002). **c and d**, optical stimulation of SOM+ did not induce gross motor deficits (two-tailed multiple t-test, Bonferroni correction, distance travelled: t=0.25, df=22, p=0.802; average speed: t=0.46, df=22, p=0.649). **e and f**, Mice were tested using 5s-pulses of continuous blue light or 5ms, 30Hz stimulation. e,The number of light deliveries did not significantly change between the two protocols (two-tailed unpaired t-test, t=0.35, df=10, p=0.730) and (f) the duration of illumination was similar (two-tailed unpaired t-test, t=1.36, df=10, p=0.201). Because we found no differences between the two protocols data were pooled together. **g**, SOM+ activation did not induce place preference during the first 10 minutes of testing. However, long lasting stimulation of SOM+ resulted in a marked place avoidance of the compartment paired with stimulation (Multiple t-test, Bonferroni correction, 2 min: t=9.70, df=14, p<0.0001; n=8 mice). Bar and line graphs show mean ± s.e.m. ** p<0.005, ***p<0.005, n.s. not significant.



Supplementary Figure 6. Ex vivo recordings of mPFC pyramidal and SOM+ neurons following 5-sec blue light stimulation of YFP-ChR2-expressing SOM+ interneurons. a, Averaged electrophysiological trace of the hyperpolarizing current evoked in pyramidal neurons (Vc -40 / -50 mV, n= 4 independent experiments) from brain slices by 5 s 473 nm light pulse stimulation of YFP-ChR2-expressing SOM+ interneurons. **b**, Representative electrophysiological trace from a pyramidal neuron (Vc – 60 mV) showing the reduction of AP firing success rate of depolarizing current pulses at 2 Hz during the 5 s blue light pulse. **c**, Mean probability of the AP firing before, during and after the 5 s blue light pulse in pyramidal neurons (n= 4 independent experiments; before, 0.36 ± 0.09; during, 0.03 ± 0.02; after, 0.22 ± 0.12). *p< 0.05 (during, p=0.024; after, p=0.196, two-sided paired t-test). Error bars show s.e.m. **d**, Representative cell-attached recording from an YFP-ChR2-expressing SOM+ interneuron showing adaptive firing during a 5 s 473 nm light pulse. **e**, Histogram showing the mean AP firing frequency before, during and after the blue light pulse in 1s time bins (n= 4 independent experiments; 1st s, 67 ± 5.8; 2nd s, 26.75 ± 3.2; 3rd s, 8.75 ± 3.54; 4th s, 6 ± 2.74; 5th s, 3.25 ± 1.49). **p<0.01, 1st s, p=0.0014 ; 2nd s, p=0.0036; 3rd s, p=0.09; 4th s, p=0.116; 5th s, p=0.118; two-sided paired t-test). Error bars show s.e.m.

References

1. Henry, J. D., von Hippel, W., Molenberghs, P., Lee, T. & Sachdev, P. S. Clinical assessment of social cognitive function in neurological disorders. *Nat. Rev. Neurol.* **12**, 28–39 (2016).
2. Heide, J. et al. Differential response to risperidone in schizophrenia patients by KCNH2 genotype and drug metabolizer status. *Am. J. Psychiatry* **173**, 53–59 (2016).
3. Barak, B. & Feng, G. Neurobiology of social behavior abnormalities in autism and Williams syndrome. *Nat. Neurosci.* **19**, 647–655 (2016).
4. Kennedy, D. P. & Adolphs, R. Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia* **50**, 3313–3319 (2012).
5. Green, M. F., Horan, W. P. & Lee, J. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* **16**, 620–631 (2015).
6. Fett, A.-K. J. et al. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neurosci. Biobehav. Rev.* **35**, 573–588 (2011).
7. Adolphs, R. Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* **1**, 21–62 (2002).
8. Adolphs, R. et al. A mechanism for impaired fear recognition after amygdala damage. *Nature* **433**, 68 (2005).
9. Frith, C. D. & Frith, U. Mechanisms of social cognition. *Annu. Rev. Psychol.* **63**, 287–313 (2011).
10. Adolphs, R. The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* **60**, 693–716 (2009).
11. Monte, O. D. et al. A voxel-based lesion study on facial emotion recognition after penetrating brain injury. *Soc. Cogn. Affect. Neurosci.* **8**, 632–639 (2013).
12. Hiser, J. & Koenigs, M. The multifaceted role of the ventromedial prefrontal cortex in emotion, decision making, social cognition, and psychopathology. *Biol. Psychiatry* **15**, 638–647 (2018).
13. Isaacson, J. S. & Scanziani, M. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).
14. Kepecs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014).
15. Rubenstein, J. L. R. & Merzenich, M. M. Model of autism: increased ratio of excitation / inhibition in key neural systems. *Brain* **2**, 255–267 (2003).
16. Hussman, J. P. Suppressed GABAergic inhibition as a common factor in suspected etiologies of autism. *J. Autism Dev. Disord.* **31**, 247–248 (2001).
17. Hashemi, E., Ariza, J., Rogers, H., Noctor, S. C. & Martínez-Cerdeño, V. The number of parvalbumin-expressing interneurons is decreased in the medial prefrontal cortex in autism. *Cereb. Cortex* **27**, 1931–1943 (2017).
18. Fatemi, S. H. et al. mRNA and protein levels for GABAAα4, α5, β1 and GABABR1 receptors are altered in brains from subjects with autism. *J. Autism Dev. Disord.* **40**, 743–750 (2010).
19. Harada, M. et al. Non-invasive evaluation of the GABAergic/glutamatergic system in autistic patients observed by MEGA-editing proton MR spectroscopy using a clinical 3 Tesla instrument. *J. Autism Dev. Disord.* **41**, 447–454 (2011).
20. Yizhar, O. et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).
21. Franklin, T. B. et al. Prefrontal cortical control of a brainstem social behavior circuit. *Nat. Neurosci.* **20**, 260–270 (2017).
22. Keum, S. et al. A missense variant at the *Nrxn3* locus enhances empathy fear in the mouse. *Neuron* **98**, 588–601.e5 (2018). Nature
23. Carrillo, M. et al. Emotional mirror neurons in the rat's anterior cingulate cortex. *Curr. Biol.* **29**, 1301–1312.e6 (2019).
24. Wang, F. Bidirectional control of social hierarchy by synaptic efficacy in medial prefrontal cortex. *Science* **334**, 693–697 (2011).

25. Zhou, T. et al. History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science* **357**, 162–168 (2017).
26. Amadei, E. A. et al. Dynamic corticostriatal activity biases social bonding in monogamous female prairie voles. *Nature* **546**, 297–301 (2017).
27. Ferretti, V. et al. Oxytocin signaling in the central amygdala modulates emotion discrimination in mice. *Curr. Biol.* **29**, 1938–1953.e6 (2019).
28. Stark, E. et al. Inhibition-induced theta resonance in cortical circuits. *Neuron* **80**, 1263–1276 (2013).
29. Spunt, R. P., Ellsworth, E. & Adolphs, R. The neural basis of understanding the expression of the emotions in man and animals. *Soc. Cogn. Affect. Neurosci.* **12**, 95–105 (2017).
30. Portfors, C. V. Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J. Am. Assoc. Lab. Anim. Sci.* **46**, 28–34 (2007).
31. Abbas, A. I. et al. Somatostatin interneurons facilitate hippocampal-prefrontal synchrony and prefrontal spatial encoding. *Neuron* **100**, 926–939 (2018).
32. Crawley, J. N. Mouse behavioral assays relevant to the symptoms of autism. *Brain Pathol.* **17**, 448–459 (2007).
33. Liguz-Lecznar, M., Urban-Ciecko, J. & Kossut, M. Somatostatin and somatostatin-containing neurons in shaping neuronal activity and plasticity. *Front. Neural Circuits* **10**, 1–15 (2016).
34. Stark, E. et al. Pyramidal cell-interneuron interactions underlie hippocampal
35. ripple oscillations. *Neuron* **83**, 467–480 (2014).
36. Veit, J., Hakim, R., Jadi, M. P., Sejnowski, T. J. & Adesnik, H. Cortical gamma band synchronization through somatostatin interneurons. *Nat. Neurosci.* **20**, 951–959 (2017).
37. Lee, E. et al. Enhanced neuronal activity in the medial prefrontal cortex during social approach behavior. *J. Neurosci.* **36**, 6926–6936 (2016).
38. Sterley, T.-L. et al. Social transmission and buffering of synaptic changes after stress. *Nat. Neurosci.* **21**, 393–403 (2018).
39. Rogers-Carter, M. M. et al. Insular cortex mediates approach and avoidance responses to social affective stimuli. *Nat. Neurosci.* **21**, 404–414 (2018).
40. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).
41. Kubota, Y. Untangling GABAergic wiring in the cortical microcircuit. *Curr. Opin. Neurobiol.* **26**, 7 14 (2014).
42. Marlin, J. J. & Carter, A. G. GABA-A receptor inhibition of local calcium signaling in spines and dendrites. *J. Neurosci.* **34**, 15898–15911 (2014).
43. Selimbeyoglu, A. et al. Modulation of prefrontal cortex excitation / inhibition balance rescues social behavior in *CNTNAP2*-deficient mice. *Sci. Transl. Med.* **9**, eaah6733 (2017).
44. McGarry, L. M. & Carter, A. G. Inhibitory gating of basolateral amygdala inputs to the prefrontal cortex. *J. Neurosci.* **36**, 9391–9406 (2016).
45. Namburi, P. et al. A circuit mechanism for differentiating positive and negative associations. *Nature* **520**, 675–678 (2015).
46. Terburg, D. et al. The basolateral amygdala is essential for rapid escape: a human and rodent study. *Cell* **175**, 723–735.e16 (2018).
47. Nakajima, M., Görlich, A. & Heintz, N. Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. *Cell* **159**, 295–305 (2014).
48. Huang, H. et al. Chronic and acute intranasal oxytocin produce divergent social effects in mice. *Neuropsychopharmacology* **39**, 1102–1114 (2014).
49. Raam, T., McAvoy, K. M., Besnard, A., Veenema, A. H. & Sahay, A. Hippocampal oxytocin receptors are necessary for discrimination of social stimuli. *Nat. Commun.* **8**, 2001 (2017).

50. Oettl, L. L. et al. Oxytocin enhances social recognition by modulating cortical control of early olfactory processing. *Neuron* **90**, 609–621 (2016).
51. Domes, G. et al. Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol. Psychiatry* **62**, 1187–1190 (2007).
52. Murugan, M. et al. Combined social and spatial coding in a descending projection from the prefrontal cortex. *Cell* **171**, 1663–1677.e16 (2017).
53. Brockhurst, J., Cheleuitte-Nieves, C., Buckmaster, C. L., Schatzberg, A. F. & Lyons, D. M. Stress inoculation modeled in mice. *Transl. Psychiatry* **5**, e537 (2015).
54. Managò, F. et al. Genetic disruption of *Arc*/*Arg3.1* in mice causes alterations in dopamine and neurobehavioral phenotypes related to schizophrenia. *Cell Rep.* **16**, 2116–2128 (2016).
55. Paxinos, G. & Franklin, K. B. J. *The Mouse Brain in Stereotaxic Coordinates*, 2nd edn. (Elsevier/Academic Press, 2001).
56. Rossant, C. et al. Spike sorting for large, dense electrode arrays. *Nat. Neurosci.* **19**, 634–641 (2016).
57. Hazan, L., Zugaro, M. & Buzsáki, G. Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *J. Neurosci. Methods* **155**, 207–216 (2006).
58. Barthó, P. et al. Ongoing network state controls the length of sleep spindles via inhibitory activity. *Neuron* **82**, 1367–1379 (2014).
59. Likhtik, E., Pelletier, J. G., Popescu, A. T. & Paré, D. Identification of basolateral amygdala projection cells and interneurons using extracellular recordings. *J. Neurophysiol.* **96**, 3257–3265 (2006).
60. Chen, T.-W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
61. Cutts, C. S. & Eglen, S. J. Detecting pairwise correlations in spike trains: An objective comparison of methods and applicati

CHAPTER 5

Self-Experience Impact on Emotion Discrimination and Modulation by Corticotropin-Releasing Factor

Federica Maltese[1], Ana Marta Capaz[1], Francesco Papaleo[1]

[1]*Genetics of Cognition laboratory, Neuroscience area, Istituto Italiano di Tecnologia, Genova, Italy.*

Abstract

Self-experience of emotional states is an important modulator of empathy and recognition of emotions in others. By using a paradigm enabling to measure in mice the ability to discriminate conspecifics based on their altered affective state, we started to address whether self-experience could modulate this capability and the neuronal mechanisms underlying these processes. Observer mice simultaneously exposed to a neutral and a stressed demonstrator showed either a preference or an avoidance for stressed demonstrators, only if they had previously received the same stress experience. In contrast, naïve observers with no self-experience of stress showed only preference for stressed demonstrators over neutral ones. Notably, we found that self-experience modulation of emotion discrimination is estrus cycle-dependent in females and dominance-dependent in males. A major player modulating responses to stress is the corticotropin-releasing factor (CRF). CRF neurons are present in the medial prefrontal cortex (mPFC) which is an important hub of emotion discrimination. Using optogenetic manipulations, we targeted CRF neurons within the mPFC during emotion discrimination, with or without stressful self-experience. Our findings suggest that CRF neurons are involved in discrimination of negatively-, but not positively-, valenced affective states. Moreover, mPFC-CRF neurons might be implicated in dominance-dependent mechanisms underlying self-experience effects on emotion discrimination. Overall, our results will elucidate how sharing the same negatively-valenced experience has an impact on responses to others' affective states and the implication of the CRF system in these processes.

Introduction

Social cognition is an important aspect of social interaction that allows detection and interpretation of social signals from conspecifics [1, 2]. Among social cognitive processes, emotion recognition is the ability to perceive and process others' affective states [3], while empathy is the capacity to share feelings of others and to respond with care to the distress in others [4]. Many studies showed a positive correlation between affective empathy and emotion recognition accuracy [5-7]. However, there are also evidence of a negative correlation between the two processes [8, 9], or even no correlation at all [10]. These discrepancies might depend on the inter-individual subjective influence of previous emotional experience (hereafter "self-experience") when facing emotional states in others [11]. Indeed, observing another person in a particular emotional state can trigger memories of a similar feeling that took place in the past [4] (Zaki & Ochsner, 2015), and that can affect both empathic responses to others and recognition of altered emotional states in others. In particular, self-experience could provoke self-oriented feeling of distress, which may disrupt the emotion recognition process by decreasing attention to affective cues displayed by others [8, 9, 12, 13]. In alternative, self-experience could prompt other-oriented emphatic concern, which may increase accuracy in emotion recognition by motivating attention toward others [14]. Thus, self-experience might either increase or decrease the sensibility to others' emotions in an individual basis. However, the mechanisms of such individual variations are poorly understood.

Emotion recognition, empathic and prosocial behaviors are not uniquely evident in humans, but can be reveled in other mammals including non-human primates and rodents [15-18]. In particular, rodents are able to discriminate conspecifics based on their emotional state and to show empathic-like behaviors such as observational fear, emotional contagion of pain, social buffering, helping and consolation behaviors [19-26]. However, the relation between emotion recognition and empathy has never been investigated in rodents. In particular, whether rodents' ability to discriminate others' emotions can be influenced by self-experience and the neuronal mechanisms underlying these processes are still not known.

Copying with stressful components of social interactions is an important aspect in individual propensity to respond to others' internal states [27]. Such coping is compromised in a variety of psychiatric and neurodevelopmental disorders, such as social anxiety disorder, autism spectrum disorders and schizophrenia [28-30]. A master regulator of stress-coping responses is the

corticotropin-releasing factor (CRF) [31-33], a neuropeptide mostly produced and secreted from the paraventricular nucleus (PVN) of the hypothalamus. The CRF system is implicated in emotional and cognitive components of stress responses [34-36], and in different aspects of social behaviors such as social interaction [37], social memory [38], social defeat [39], pair bonding [40], defensive and aggressive responses [41], and sexual behavior [42]. Moreover, human studies have revealed that the CRF system modulates the effect of stress on empathyc behaviors [43, 44].

CRF neurons are also found in the medial prefrontal cortex (mPFC) [45, 46], which is strongly implicated in emotion recognition and other socio-cognitive abilities [20, 47-50]. Moreover, the mPFC structure and function can be affected by stress [51, 52], and it plays a critical role in the regulation of response to stress by promoting adaptation and survival [53-55]. Despite this, whether mPFC CRF neurons might modulate responses to others' altered affective states and the impact of self-experience have been unexplored.

In the present study, we investigated whether sharing the same stressful experience could affect mice ability to discriminate altered states in conspecifics. We found that self-experience of a stressful event led to altered emotion discrimination abilities and this effect was partially mediated by estrus cycle in females and social hierarchy in males. Moreover, using opotogenetic manipulation of mPFC-CRF neurons, we investigated whether these neurons might modulate mice ability to discriminate altered emotional states in conspecifics. Our findings suggest that CRF neurons are involved in discrimination of negatively-, but not positively-, valenced affective state. Moreover, we reported preliminary data suggesting that mPFC-CRF neurons might be implicated in dominance-dependent mechanisms underlying self-experience effects on emotion discrimination.

Overall, our results show that sharing the same negatively-valenced experience has an aversive effect towards others' affective state and suggest an implication of the CRF system in this process.

Results

Self-experience differentiates individual responses to negative affective state in others

We first tested whether sharing experience of an altered affective state could have an impact on mice ability to discriminate conspecific based on the same altered state. In the Emotion Discrimination Task (EDT) we previously developed [19, 20], an observer mouse is placed in front of two conspecifics (demonstrators) in different emotional states in order to discriminate between them. The two demonstrators, matched with the observer by the same sex and age, were located inside inverted wire cups, divided by a black wall (Fig. 1A and E). To induce an altered affective state with negative valence, one demonstrator ('stress') underwent a mild stress protocol, consisting of 15 min of acute restraint immediately before the EDT (Fig. 1A and E). The other demonstrator was a naïve mouse directly taken from its homecage ('neutral', Fig. 1A and E). As we previously reported [20], naïve observers displayed increased sniffing towards stressed conspecifics than towards neutral demonstrators (Fig. 1B and Supplementary Fig. 1A-B). No differences in emotion discrimination were observed between sex-matched females and males (Fig. 1C and Supplementary Fig. 1A-B). In the self-experience group, observers experienced the same stressful challenge as stressed demonstrators 24 hours before the EDT test (Fig. 1E). The day of the test, demonstrator mice were manipulated exactly as described above. We found that self-experience overall abolished emotion discrimination towards a stressed mouse (Fig. 1F and Supplementary Fig. 1E-F). This was equally evident in males and females mice (Fig. 1G and Supplementary Fig. 1 E-F).

We next explored individual differences in emotion discrimination. In naïve observers, the scores of exploration towards stressed demonstrators fit a normal distribution (D'Agostino and Pearson normality test, $K2 = 0.42$, $P = 0.81$), with the 89% of mice showing a preference towards stressed demonstrators (107 out of 120, Fig. 1D). A similar normal distribution was found in mice that experienced stress (D'Agostino and Pearson normality test, $K2 = 1.28$, $P = 0.53$), but the manipulation induced in 55% of them (51 out of 93) an increased sniffing towards stressed mice, while the other 45% (42 out of 93) displayed increased approach towards neutral demonstrators (Fig. 1H). Moreover, the discrimination index (DI) was differently distributed in the two populations (Fig. 1I). In particular, the majority of naïve observers (92 out of 120) had positive values ($DI > 0.10$) of DI (Fig. 1I), which indicates a preference towards stressed demonstrators. In

the self-experience group, instead, DI was distributed from negative to positive values, with almost half of the mice (47 out of 93) approaching more stressed demonstrators (DI > 0.10), while the other half (36 out of 93) interacted more with the neutral ones (DI < -0.10; Fig 1I). In both conditions, a small proportion of mice did not show any preference for neither demonstrators (-0.10 < DI < +0.10; 20 out of 120 in naïve observers, 10 out of 93 in self-experienced mice; Fig.1I).

Finally, as measure of stress level in observer mice, we assessed the amount of grooming during the entire test and found that mice with self-experience had higher levels compared to mice in the normal condition (Fig. 1J). However, this increased grooming in self-experienced mice was not correlated with their discrimination index (Fig.1K), indicating that this behavior is not directly linked to preference or avoidance of stressed demonstrators. We also measured grooming behavior in demonstrators and found that in both conditions stressed mice displayed more grooming than neutral ones (Supplementary Fig.1 C and G), but this was not correlated with observers' discrimination behavior (Supplementary Fig.1 D and H).

Overall, these results showed that experience of an affective state modulates in a subjective way mice ability to discriminate conspecifics in the same altered state compared to neutral demonstrators.

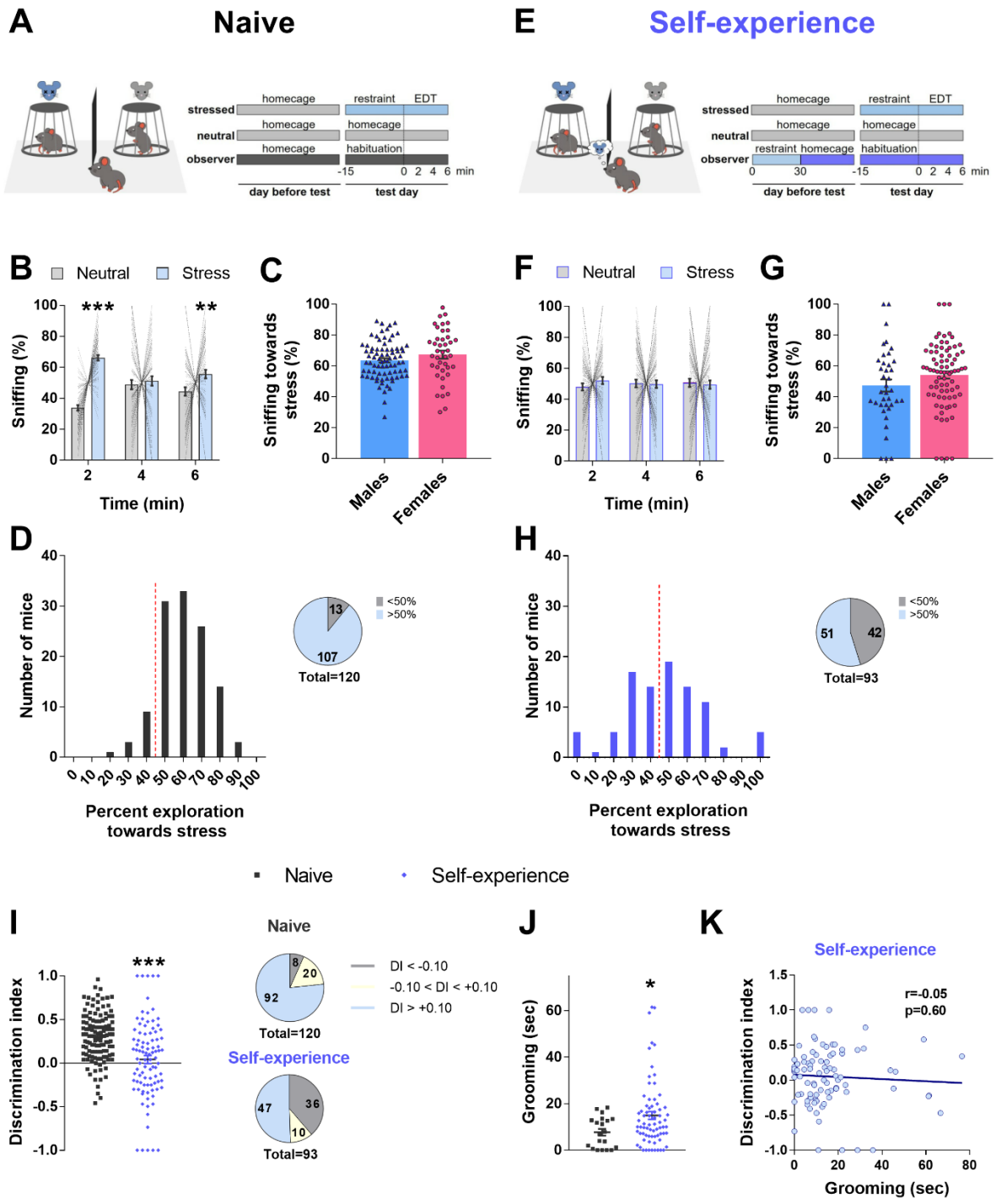


Figure 1. Self-experience affects emotion discrimination. (**A**) Experimental design of the EDT in naive condition. Observers did not undergo any manipulation. One demonstrator (stress, blue light) was subjected to the restraint stress for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. (**B**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by naïve observer mice (two-tailed multiple t-test, Holm-Sidak correction. 2 min: t = 8.851 , d.f. = 402, P < 0.0001; 4 min: t = 0.63, d.f = 402, P = 0.53; 6 min: t = 3.062, d.f.= 402, P= 0.004). N=120 observers. (**C**) No difference between sexes was found during the first 2 min of the test (79 males, 41 females; two-tailed unpaired t-test: t = 1.415, d.f. = 118, P = 0.16). (**D**) Observers

distribution based on exploration towards stressed demonstrator (107 of 120 higher than 50%, one-sample t-test against chance, defined as 50%: t = 11.77, d.f. = 238, P < 0.0001) (**E**) Experimental design of the EDT in self-experience condition. Observers were subjected to restraint stress for 30 min, 24 hrs before the EDT. One demonstrator (stress, blue light) was subjected to restraint stress for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. (**F**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by observer mice 24 hrs after self-experience (two-tailed multiple t-test, Holm-Sidak correction. 2 min: t = 1.161, d.f. = 552, P = 0.57; 4 min: t = 0.11, d.f. = 552, P = 0.93; 6 min: t = 0.32, d.f. = 552, P = 0.93). N=93 observers. (**G**) No difference between sexes was found during the first 2 min of the test (38 males, 79 females; two-tailed unpaired t-test: t = 1.567, d.f. = 115, P = 0.12). (**H**) Observers distribution based on exploration towards stressed demonstrator (51 of 93 higher than 50%, one-sample t-test against chance, defined as 50%: t = 0.886, d.f. = 184, P= 0.37). (**I-J**) Mice in naïve (dark gray) and self-experience (blue) condition significantly differ based on (**I**) discrimination index ([sniffing towards stressed – sniffing towards neutral] / [sniffing towards stressed + sniffing towards neutral]) during the first 2 min of the test (two-tailed unpaired t-test: t = 5.168, d.f. = 211, P < 0.0001), and (**J**) based on amount of time (in seconds) spent grooming across the 6 min test (two-tailed unpaired t-test: t = 2.192, d.f. = 90, P = 0.03). N=93 naïve observers and N=120 self-experienced mice. (**K**) No significant correlation was found between discrimination index (in y axis) and grooming (in x axis) within self-experience group (r = -0.05; P = 0.60) N=89 mice.
Bar and line graphs show mean ± s.e.m. *P < 0.05. **P < 0.005. ***P < 0.0005.
See also Supplementary Figure 1.

Estrus phase prevents self-experience modulation of emotion discrimination

Many evidence suggests the influence of menstrual cycle on emotion recognition and empathy in women [56-58]. Moreover, in female rodents it has been shown that emotional contagion can be dependent on estrus cycle [59]. Thus, we hypothesized that sexual hormones could influence self-experience mediated mechanisms of emotion discrimination. We tested whether female mice in different phases of the estrus cycle had diverse behavioral outcomes after self-experience. In mice, estrus cycle averages 4–5 days and it is generally divided into 4 phases: proestrus, estrus, metestrus and diestrus [60]. We monitored estrus cycle every day by using the vaginal cytological evaluation method [61], until we were sure that mice would be in the desired phase. As previously done in literature [59], mice were tested in the estrus or diestrus phases, which are the longest within the cycle and have opposite hormonal configuration [61]. In our previous studies, we checked estrus cycle in naïve observers and found that it does not affect emotion discrimination abilities [19, 20]. Here, self-experienced females in the estrus phase showed increased sniffing towards stressed demonstrators compared to neutral mice (Fig. 2A-B). Diestral females, instead, showed overall any preference, with half of them preferring stressed demonstrators and the other half avoiding stressed stimuli (Fig. 2C and E). Nevertheless, during the 2 minutes in the middle of the test, females in diestrus showed a significant avoidance of stressed demonstrators (Fig. 2C). Moreover, discrimination index was differently distributed in the two groups (Fig. 2E). We also checked emotion discrimination behavior dividing observers

depending on the cycle phase in which they were when stress experience occurred. We did not find any difference in the EDT performance of these mice (Supplementary Fig. 2A-C), showing that estrus cycle modulation of self-experience does not occur during experience of stress. Finally, we checked grooming behavior in both estral and diestral mice, but no difference was detected (Fig. 2D).

Overall, these data showed that estrus phase prevents self-experience modulation of emotion discrimination abilities, partially explaining this process.

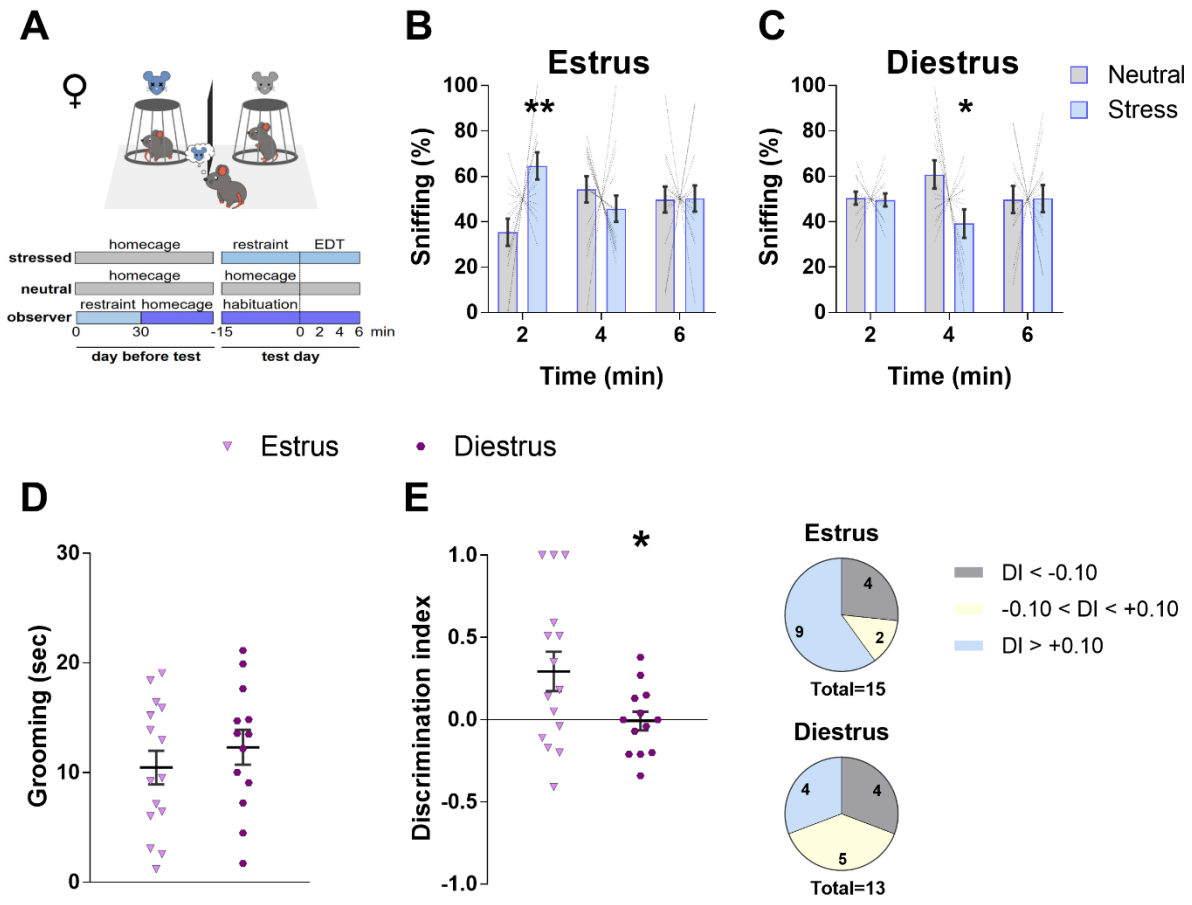


Figure 2. Estrus phase protects females from self-experience alterations of emotion discrimination. (**A**) Experimental design of the EDT in self-experience condition. Observers were subjected to restraint stress for 30 min, 24 hrs before the EDT. One demonstrator (stress, blue light) was subjected to restraint stress for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. (**B-C**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by observer mice in estrus (**B**) or diestrus (**C**) 24 hrs after self-experience (two-tailed multiple t-test, Holm-Sidak correction. Estrus. 2 min: t = 3.548, d.f. = 84, P = 0.002; 4 min: t = 1.025, d.f. = 84, P= 0.52; 6 min: t = 0.04, d.f. = 84, P= 0.96. Diestrus. 2 min: t = 0.10, d.f. = 72, P= 0.99; 4 min: t = 2.9, d.f. = 72, P= 0.01; 6 min: t = 0.04, d.f. = 72, P= 0.99). (**D**) Time (in seconds) spent grooming across the 6 min test displayed by observer mice in estrus (light purple) or diestrus (purple). No difference between the two

groups was found (two-tailed unpaired t-test: t = 0.83, d.f. = 26, P = 0.41). (**E**) Mice in estrus (light purple) and diestrus (purple) significantly differ during the first 2 min of the test based on discrimination index ([sniffing towards stressed – sniffing towards neutral] / [sniffing towards stressed + sniffing towards neutral]) (two-tailed unpaired t-test: t = 2.161, d.f. = 26, P = 0.04). N=15 estral observers and N=13 diestral observers.
Bar and line graphs show mean ± s.e.m. *P < 0.05. **P < 0.005.
See also Supplementary Figure 2.

Dominant mice are more susceptible to self-experience modulation of emotion discrimination

We next tested the same hypothesis about sexual hormones on male mice. Many evidence suggests that testosterone levels can affect emotion recognition abilities [62-64]. Moreover, it has been shown that social dominance, which reflects testosterone levels [65, 66], has a modulatory effect on social fear transmission in male rodents [67]. Thus, we tested whether dominance could influence emotion discrimination abilities after self-experience. We used the well-validated tube test to assess hierarchy in mice [68], in order to divide observers in dominants and subordinates. In naïve condition, both dominant and subordinate mice showed the same increased sniffing towards stressed demonstrators compared to neutral ones (Fig. 3A-C). In the self-experience condition, dominant mice avoided stressed demonstrators compared to neutral ones (Fig. 3D-E). Instead, subordinate males were mixed in mice that preferred stressed demonstrators and mice that avoided them (Fig. 3F and H). No difference was found when comparing the different ranks within the subordinates group (Supplementary Fig. 2D-E). Behavioral differences between dominants and subordinates are also shown by the discrimination index, which is differently distributed in the two groups (Fig. 3H). Finally, we checked grooming behavior in both dominant and subordinate mice, but no difference was detected (Fig. 3G).

Overall, these data showed that dominant mice are more vulnerable to self-experience modulation of emotion discrimination, and that dominance status partially predicts self-experience impact on emotion discrimination abilities.

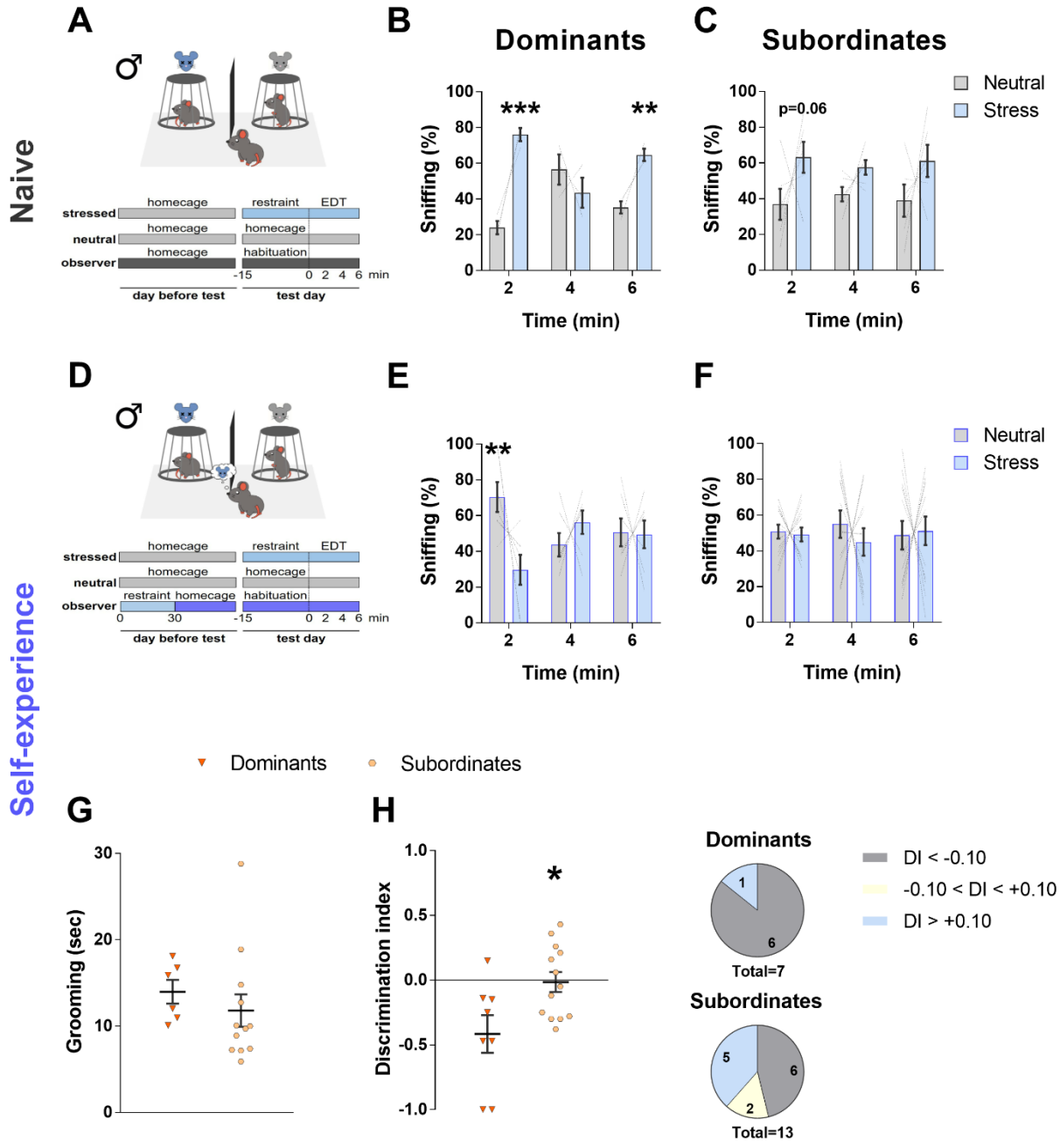


Figure 3. Dominant mice are more sensible to self-experience modulation of emotion discrimination. (**A**) Experimental design of the EDT in naïve condition. Observers did not undergo any manipulation. One demonstrator (stress, blue light) was subjected to the restraint stress for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. Hierarchy was assessed later by using the tube test. (**B-C**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by dominant (**B**) or subordinate (**C**) naïve mice (two-tailed multiple t-test, Holm-Sidak correction. Dominants. 2 min: t = 6.492, d.f. = 12, P < 0.0001; 4 min: t = 1.621, d.f. = 12, P= 0.13; 6 min: t = 3.666, d.f. = 12, P= 0.006. Subordinates. 2 min: t = 2.444, d.f. = 30, P= 0.06; 4 min: t = 1.407, d.f. = 30, P= 0.17; 6 min: t = 2.079, d.f. = 30, P= 0.09). N=3 dominant observers and N=6 subordinate observers. (**D**) Experimental design of the EDT in self-experience condition. Observers were subjected to restraint stress for 30 min, 24 hrs before the EDT. One demonstrator (stress, blue light) was subjected to the restraint

stress test for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. Hierarchy was assessed later by using the tube test. (**E-F**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by dominant (**E**) or subordinate (**F**) mice 24 hrs after self-experience (two-tailed multiple t-test, Holm-Sidak correction. Dominants. 2 min: t = 3.8, d.f. = 36, P= 0.0016; 4 min: t = 1.167, d.f. = 36, P= 0.43; 6 min: t = 0.117, d.f. = 36, P= 0.90. Subordinates, 2 min: t = 0.161, d.f. = 72, P= 0.96; 4 min: t = 1.04, d.f. = 72, P= 0.65; 6 min: t = 0.25, d.f. = 72, P= 0.96). (**G**) Time (in seconds) spent grooming across the 6 min test displayed by dominant (orange) or subordinates (light orange) mice. No difference between the two groups was found (two-tailed unpaired t-test: t = 0.761, d.f. = 16, P = 0.45). (**H**) Dominant (orange) and subordinate (light orange) mice significantly differ during the first 2 min of the test based on discrimination index ([sniffing towards stressed – sniffing towards neutral] / [sniffing towards stressed + sniffing towards neutral]) (two-tailed unpaired t-test: t = 2.672, d.f. = 19, P = 0.015). N=7 dominant observers and N=13 subordinate observers.
Bar and line graphs show mean ± s.e.m. *P < 0.05. **P < 0.005. ***P < 0.0005.
See also Supplementary Figure 2.

CRF neurons are important for discrimination of negatively-valenced affective states

The CRF system is an important modulator of emotional and cognitive responses to stress [34-36]. In particular, CRF positive cells are found within the mPFC, and they are mostly GABAergic interneurons densely distributed in layers 2/3 and sparsely in layer 5 [45, 46]. To investigate their possible involvement in emotion discrimination, we bilaterally injected an adeno-associated virus (AAV) carrying a Cre-dependent halorhodopsin (AAV-EF1a-DIO-eNpHR3.0-eYFP) into the mPFC of CRF-Cre mice, which were subsequently implanted with optic fibers terminating dorsally to this area (Fig. 4A and Supplementary Fig.4A). CRF+ neurons activity was silenced during the first 2 min of the EDT by delivering continuous green light (532nm), using the same laser stimulation protocol applied in our previous study [20] (Fig. 4B and Supplementary Fig.4A). In this way, we targeted the time window in which naïve observers increase exploration towards demonstrators in altered affective states (Fig 1B; 19, 20). Mice were tested on consecutive weeks, with treatments (light off and light on) counterbalanced. We surprisingly found that naïve observers were not able to discriminate demonstrators in both the light OFF and ON condition (Fig. 4C-D). CRF neurons could overexpress the cre-recombinase enzyme as shown in other mouse Cre lines [69-73], potentially leading to altered neuronal signaling [70, 71]. Thus, we tested naïve CRF-Cre observers (without any virus injection or lens implantation) and found that they failed in discriminating between a stressed and a neutral demonstrator (Supplementary Fig. 3). In contrast, CRF-Cre mice were able to discriminate conspecifics based on a positively-valenced affective state (relief), and even optogenetic silencing of mPFC CRF neurons modify this ability (Supplementary Fig. 4).

Overall, these results suggest that CRF+ cells might be involved in discrimination of negatively, but not positively-, valenced affective state.

mPFC CRF neurons modulate dominance-dependent mechanisms of self-experience

We next tested whether silencing mPFC-CRF neurons could affect self-experience modulation of emotion discrimination. Also in this case, mPFC CRF+ cells were inhibited by delivering continuous green light during the first 2 minutes of the test, which was carried out 24 hours after the self-experience of stress protocol (Fig. 4E). In both light OFF and light ON condition, we found overall no discrimination between stressed and neutral demonstrators (Fig. 4F-G). In line with the findings reported in Figure 1I, discrimination indexes in both light OFF and light ON condition were distributed from negative to positive values, with only the 30-40% of observes showing increased sniffing towards stressed demonstrators (Fig. 4H). However, by separating male observers in dominants and subordinates, based on the social hierarchy assessed by the tube test, we found that, while replicating in the light OFF condition the same results found in Figure 3, in the light ON condition there was a change in mice performance compared with the OFF condition (Fig. 4I-J). In particular, dominant mice did not avoid stressed demonstrators anymore, while subordinate observers showed an inversion in their behavioral outcome by starting to avoid stressed demonstrators (Fig. 4I-J).

These results suggest that CRF+ neurons within the mPFC play a role in social dominance dependent mechanisms that mediate self-experience impacts on mice emotion discrimination abilities.

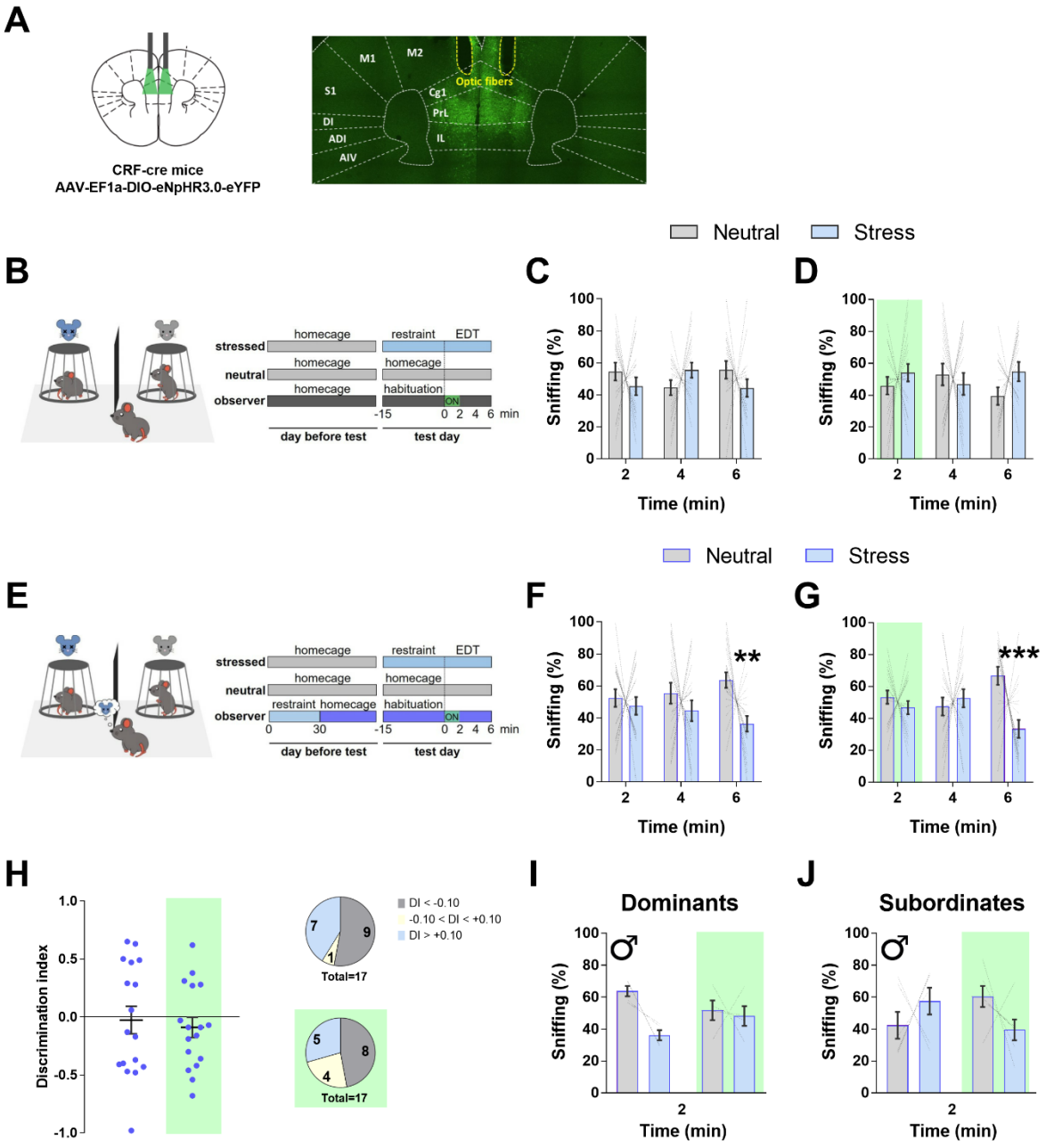


Figure 4. CRF modulates discrimination of negatively-valenced affective state and dominant-dependent mechanisms of self-experience. (**A**) left: CRF-Cre mice were injected in the mPFC with AAV EF1a-DIO-eNpHR3.0-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. Right: representative image of coronal mPFC section. Findings were replicated in three independent experiments with similar results. (**B**) Experimental design of the EDT in naïve condition. Observers did not undergo any manipulation. Mice were tested in the EDT with one stressed and one neutral demonstrator. Photoinhibition was performed for 2 min from the beginning of the test using continuous green light (λ = 532 nm). (**C-D**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by naïve observer mice tested in both the light OFF (**C**) and light ON (**D**) condition (two-tailed multiple t-test, Holm-Sidak correction. Light OFF. 2 min: t = 1.177, d.f. = 96, P= 0.37; 4 min: t = 1.415, d.f. = 96, P= 0.37; 6 min: t = 1.471, d.f. = 96, P= 0.37. Light ON. 2 min: t = 0.916, d.f. = 96, P= 0.59; 4 min: t = 0.66, d.f. = 96, P=

0.59; 6 min: t = 1.738, d.f. = 96, P= 0.23). N=17 observers. (**E**) Experimental design of the EDT in self-experience condition. Observers were subjected to restraint stress for 30 min, 24 hrs before the EDT. Mice were tested in the EDT with one stressed and one neutral demonstrator. Photoinhibition was performed for 2 min from the beginning of the test using continuous green light (λ = 532 nm). (**F-G**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed 24h after self-experience by observer mice in both the light OFF (**F**) and light ON (**G**) condition (two-tailed multiple t-test, Holm-Sidak correction. Light OFF. 2 min: t = 0.593, d.f. = 102, P= 0.55; 4 min: t = 1.32, d.f. = 102, P= 0.34; 6 min: t = 3.313, d.f. = 102, P= 0.004. Light ON. 2 min: t = 0.865, d.f. = 102, P= 0.62; 4 min: t = 0.689, d.f. = 102, P= 0.62; 6 min: t = 4.402, d.f. = 102, P < 0.0001). N=18 observers. (**H**) Discrimination index ([sniffing towards stressed – sniffing towards neutral] / [sniffing towards stressed + sniffing towards neutral]) of CRF-cre self-experienced observers in light OFF and light ON condition during the first 2 min of the test (two-tailed unpaired t-test: t = 0.413, d.f. = 16, P = 0.68). N=18 observers (**I-J**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the first 2 min of EDT displayed 24 hrs after self-experience by dominant (**I**) and subordinate (**J**) observers in both the light OFF and light ON condition (two-tailed multiple t-test, Holm-Sidak correction. Dominants. light OFF: t = 2.244, d.f. = 32, P= 0.12; light ON: t = 0.28, d.f. = 32, P= 0.77. Subordinates. light OFF: t = 1.518, d.f. = 32, P= 0.25; light ON: t = 2.083, d.f. = 32, P= 0.12). N=4 dominant observers and N=6 subordinate observers.
Bar and line graphs show mean ± s.e.m. **P < 0.005. ***P < 0.0005.
See also Supplementary Figure 3 and 4.

Discussion

In this study, we revealed that self-experience of an aversive experience leads to altered discrimination of the same negatively-valenced state in mice. In particular, we found that self-experience affected mice differently, partially depending on estrus cycle in females and hierarchy status in males. Moreover, we reported first evidence of CRF modulation of emotion discrimination and self-experience processes.

In agreement with evidence in human studies, our findings support the indication that self-experience can affect individual responses to altered affective states in others others [8]. In particular, in humans it has been shown that, while observing others in a particular state of mind, recalling of a similar experience or felling that occurred in the past can lead to two different reactions. First, self-experience can lead to empathic concern that drives the attention to others and their expressive cues [11]. In contrast, self-experience could lead to personal distress that shifts the attention away from others' and their facial expressions [8, 9, 12, 13]. Thus, self-experience in humans induce different reactions that are subjective to each individual. We found a similar pattern of responses in mice, where a previous experience of a stressful event triggered in some mice the propensity to approach a stressed conspecific over a neutral one, while other observers avoided the stressed mouse. We might speculate that mice that spend more time with the stressed conspecific might be more engaged in other-concern. Instead, mice that avoid stressed demonstrators might perceive more self-distress, in line with evidence showing that personal distress is related with

keeping greater physical distance and avoiding exposure to others' suffering [74, 75]. However, further investigations are needed in order to better understand these behavioral responses.

In literature, there are evidence of self-experience modulation of empathic-like behaviors in both mice [25, 76] and rats [77-79], as well as modulation of pro-social responses in rats (e.g. helping behavior) [80, 81]. In particular, in some models of emotional contagion it has been shown that vicarious freezing can occur only when observers had experienced before the same aversive event [76-79]. Nevertheless, other studies have shown that previous experience of the same episode is not strictly necessary to elicit emotional contagion [77, 82-88], but it could modulate it by increasing the individual response to the same experience [25]. Despite this evidence, our study is the first providing initial evidence of self-experience modulation of emotion discrimination in rodents.

In our model, observers and demonstrators were always unfamiliar to each other. In contrast, models of emotional contagion have assessed self-experience impact between cagemates [77-79]. In particular, in emotional contagion models it has been reported a modulatory role when observers and subjects are siblings or are mating partners with each other [24]. However, it has also been shown that self-experience modulation of empathic behaviors could not be sensitive to whether or not observers and demonstrators are cage-mates or strangers to each other [76]. We have previously shown that familiarity increase emotion discrimination abilities in mice [20]. However, additional experiments will be required to evaluate if familiarity might modulate emotion discrimination also after self-experience.

Another open question is whether a non-specific prior stressor could affect emotion discrimination abilities. For example, Sanders et al. (2013) [76] have shown that only the same stressful experience could trigger empathic responses. Thus, it would be important to assess the impact of stress itself on emotion discrimination abilities.

Many humans' studies have showed that facial emotion processing and empathic behaviors can be modulated by sexual hormones [56-58, 63, 89, 90]. In particular, menstrual cycle has been highly correlated to emotion recognition accuracy, with different behavioral outcomes depending on the specific phase of the cycle [56-58]. By testing the hypothesis that estrus cycle in female mice could modulate self-experience effects on emotion discrimination, we found that observers in estrus, the phase in which they are sexually receptive, are not affected by self-experience and their emotion discrimination performance is similar to no self-experience condition. This is in line

with a rat model of observational fear learning in which it has been shown that social transfer of fear, which usually promotes defensive behavior to potentially dangerous situations in the environment, did not occur in estral females [59]. Our hypothesis is that during sexually receptivity stage, females are less sensible to self-distress thus showing a normal behavior in discrimination of stressed state. However, further investigation is needed in order to better understand the ethological meaning behind estrus cycle modulation of self-experience.

Many human studies have reported a modulatory effect of testosterone levels on empathy and emotion recognition processes [63, 89, 90]. Testosterone levels can be reflected by social hierarchy ranks [65, 66], which are important contributors to individual differences that might shape how individuals respond to environmental challenges [91]. Indeed, several studies have demonstrated the influence of social status in cognitive performance [92, 93]. We found that self-experience in dominant mice completely reverted preference in exploring a stressed mouse, making them avoid stressed demonstrators. In agreement, dominant male mice show higher anxiety and increased susceptibility to chronic social defeat than subordinate mice [94]. Moreover, considering that subordinates mice live an every-day social defeat within the home-cage, we could hypothesize that these mice are more habituated to stressful situations than dominants. Indeed, it has been reported that subordinate mice have higher basal corticosterone levels [95]. An open question is whether, in a familiarity context, emotion discrimination performance after self-experience might change depending of the dominance relation between observer and demonstrators. For example, in a rat model of social fear learning it has been shown that subordinate subjects show increased responses when fear-associations are transmitted by a dominant cage-mate [67]. Thus, further investigations are needed in order to clarify how dominance status modulate self-experience impact on emotion discrimination.

The neuronal mechanisms underlying self-experience effect on emotion recognition are poorly understood. We focused on the mPFC, which is known to be highly involved in social cognition [96, 97]. In particular, in our previous work, we revealed a specific association between affective state discrimination and mPFC neuronal activity [20]. Here we targeted the CRF system, which is implicated in emotional and cognitive components of stress responses [34-36], and in different aspects of social behaviors [41]. Few studies have directly investigated the neural circuit mechanisms underlying mPFC CRF neurons. In particular, it has been shown their involvement in mediating opioid reward [98], cognitive dysfunction [99] and behavioral responses under stressful

situations [46]. To our knowledge, this is the first time CRF neurons have been studied in the context of empathic behaviors. We found that silencing of mPFC CRF neurons led to emotion discrimination impairments. However, the same deficiencies were already present in our control group of CRF-cre mice injected with halorhodopsin and implanted with optic fibers in the mPFC, but did not receive delivery of green light. This evidence highlighted possible basal deficits in emotion discrimination abilities of CRF-cre mouse line. In agreement, naïve CRF-cre mice failed in discriminating between a stressed and a neutral demonstrator. Moreover, CRF-Cre mice behaved normally in the positively-valenced version of our emotion discrimination task (relief; 19, 20). This evidence are in line with the specific involvement of CRF in stress copying behaviors. From our knowledge, there are no data in literature reporting behavioral alterations in CRF-cre mice [35, 46, 98, 100, 101], even though it is not surprising that a cre line could have this kind of impairments. For example, in a choline acetyltransferase (ChAT)-cre line it has been reported not only behavioral deficits compared to wild-types, but also altered expression of the ChAT gene [102]. This kind of alterations could be due to overexpression of the cre-recombinase enzyme [69-73], which could potentially lead to altered neuronal signaling [70, 71]. Thus, to better understand the basal deficits we found, a further characterization of CRF-cre line within the context of emphatic behaviors must be undertaken. Nevertheless, in the self-experience paradigm, we could replicate the same results found in wild-type mice, especially those related to social dominance. Indeed, in the light OFF control group we found that dominant mice were significantly avoiding stressed demonstrators, and preliminary results showed that this was reverted by optogenetic inhibition of mPFC CRF neurons. The implication of mPFC circuits in dominance-dependent mechanisms is supported by evidence reporting differential brain gene expression in the mPFC between dominants and subordinates [103]. Overall, the results here reported showed an involvement of CRF neurons in discrimination of negatively-, but not positively-, valenced affective states, and first evidence of a possible modulation of self-experience processes.

In summary, our data provide new insights into the role of self-experience in modulating the ability to recognize and respond to others' emotions. This represents an important opportunity to reliably measure scarcely explored aspects of social cognition, supporting more translational approaches between rodent and human social cognitive studies, with relevance to circuits and neurochemical systems involved in different psychiatric disorders such as autism spectrum disorders and schizophrenia. Moreover, our preliminary results offer evidence of CRF involvement

in emotion discrimination and self-experience processes, providing a possible target for manipulation of emotion recognition abilities.

Materials and methods

Mice

All procedures were approved by the Italian Ministry of Health (permits 230/2009-B, 107/2015-PR, and 749/2017-PR) and local Animal Use Committee and were conducted in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the European Community Council Directives. Routine veterinary care and animal maintenance was provided by dedicated and trained personnel. Males and female C57BL/6 J animals aged 3–6 months old were used. Founders of the CRF-cre line (B6(Cg)-Crhtm1(cre)Zjh/J, id #012704) were purchased from The Jackson Laboratory and then bred and expanded in our animal facility for successive testing. Mouse genotypes were identified by polymerase chain reaction (PCR) analysis of ear DNA. Distinct cohorts of naive mice were used for each experiment. Animals were housed two to four per cage in a climate-controlled facility (22 ± 2 °C), with ad libitum access to food and water throughout, and with a 12 h light–dark cycle (19:00–07:00 schedule). Experiments were run during the light phase (within 10:00–17:00). All mice were handled on alternate days during the week preceding the first behavioral testing.

Behavioral paradigms

Emotion Discrimination Test. Testing mice (observers) were habituated and tested inside a standard mouse cage (Tecniplast, 35.5 × 23.5 × 19 cm) equipped with a dark separator between two cylindrical wire cups that hosted the demonstrators (10.5 cm in height, bottom diameter 10.2 cm, bars spaced 1 cm apart; Galaxy Cup, Spectrum Diversified Designs). The separator (11 × 14 cm) between the two wire cups was wide enough to cover the reciprocal view of the demonstrators while leaving the observer mice free to move between the two sides of the cage. A plastic cylinder was placed on the top of the wire cups to prevent the observer mice from climbing and remaining on the top of them. The cups, separators, and experimental cages were replaced after each subject with clean copies to avoid scent carryover. Similarly, the rest of the apparatus was wiped with water and dried with paper towels for each new subject. After each testing day, the wire cups,

separators, and cubicles were wiped down with 50% ethanol and allowed to air dry. Testing cages were autoclaved as performed as standard in our animal facility. Habituation to the testing setting occurred on 3 consecutive days before the first experiment; each habituation session lasted 10 min. During both habituation and behavioral testing, the cages were placed inside a dimly lit (6 ± 1 lx) soundproof cubicle (Med Associates). Demonstrator mice, matched by age, sex, and strain to the observers, were habituated, without the observer, inside the same cage under the wire cups.

Observers.

Naïve condition: On the day of testing, 15 min before the experiment, observer mice were gently moved into the dimly lit testing cubicles. For the optogenetics experiments, observer mice were connected to an optic fiber during this period. Then, one neutral demonstrator and one emotionally altered demonstrator (relief or stress) mouse were placed under the wire cups, and the 6 min experiment started.

Self-experience condition: 24 h before the experiment, observer mice were subjected to a stress consisting of the restraint test, a standard procedure to induce physiological stress in rodents [104], for 30 min. All mice in the same cage were subjected to the same manipulation at the same time. On the day of testing, 15 min before the experiment, observer mice were gently moved into the dimly lit testing cubicles. For the optogenetics experiments, observer mice were connected to an optic fiber during this period. Then, one neutral demonstrator and one stressed mouse were placed under the wire cups, and the 6 min experiment started.

Neutral demonstrators. All neutral mice were habituated to the experimental setting as described above. For both relief and stress conditions, neutral demonstrators did not receive any manipulation and were left undisturbed, with ad libitum water access, in their home cage. On the day of testing, neutral demonstrators were brought, inside their home cages, into the experimental room 1 h before the experiment began. All demonstrators were group-housed, separately from the cages of stressed and relieved demonstrators. Demonstrators were test naive and used a maximum of two or three times, always with at least 1 week between each consecutive test. No differences were observed in the performance of the observer mice depending on the previous experience of the demonstrator.

Relief demonstrators. Mice were habituated to the experimental setting as reported above. Relieved demonstrators were water-deprived 23 h before the experiment. Ad libitum access to water was given 1 h before the test, and mice were brought inside the experimental room in their

home cages. Food was ad libitum all the time and some extra pellets were put inside the home cage during the 1 h water restoration.

Stress demonstrators. Mice were subjected to a mild stress consisting of the restraint test for 15 min before the beginning of the EDT. These mice were then immediately moved to the testing arena, under the wire cup.

Digital cameras (Imaging Source, DMK 22AUC03 monochrome) were placed facing the long side of the cage and on top of the cage to record the test from different angles using a behavioral tracking system (Anymaze 6.0, Stoelting). These videos were used offline by experimenters blind to the manipulations of both the observers and demonstrators for a posteriori scoring of sniffing and grooming.

Tube Test. After EDT, social hierarchy of male mice was assessed as previously described [68]. Briefly, this test consisted of two phases: training and testing. Training phase lasted two consecutive days during which mice were trained to walk through a Plexiglas tube ten times, five times from each side. In the testing phase, cagemate mice were tested in a pair-wise fashion in the tube, and mouse rank was assessed by the number of times each mouse wins or loose. The rank was considered stable when all mice maintained the same ranking for four consecutive days. After each trial, the mice were put back into the home cage and left for 2 min before starting the next trial, in order to reduce the potential immediate impact of recent winning or losing.

Estrus cycle assessment

Estrus cycle assessment was carried out daily in observer females using the cytological vaginal evaluation method as previously described [61]. Briefly, vaginal cells samples were taken between 10:00 and 12:00 a.m., and stained with 1% Toloudine blue. Stained smears were then analyzed under a light microscope. Observers were tested during the estrus phase, characterized by the presence of numerous un-nucleated cornified cells in the vaginal smear, or during the diestrus phase characterized by a smear enriched of leukocytes and with sparse nucleated cells.

Viral injections and optic fibers implants

Viral vectors. AAV5-EF1a- DIO-eNpHR3.0-eYFP.WPRE.hGH (Addgene, 20949, qTiter 1.95 × 1013 GC per ml) was purchased from the University of Pennsylvania Viral Vector Core.

Surgical procedures. CRF-Cre mice were naive and 2–3 months old at the time of surgery. All mice were anesthetized with a mix of isoflurane (2%) and oxygen (1.5%) by inhalation and mounted into a stereotaxic frame (Kopf Instruments) linked to a digital micromanipulator. Brain coordinates of viral injection in the mPFC were chosen in accordance with the mouse brain atlas: anterior–posterior (AP), +1.9 mm; medial–lateral (ML): ± 0.25 mm; and dorsal–ventral (DV): −2.4 mm. The volume of AAV injection was 250nl per hemisphere. We infused virus through a glass micropipette connected to a 10 μl Hamilton syringe. After infusion, the pipette was kept in place for 10 min and then slowly withdrawn. After virus injection mice were given 2 weeks to recover and for the viral transgenes to adequately express before undergoing optic fiber implantation and behavioral experiments. Mice underwent stereotaxic surgery for fiber optic implantation 2 weeks after viral injection. The skull was exposed and the two previous holes were used to target the mPFC. Dual fiberoptic cannulae (200 μm, 0.37 numerical aperture, fiber distance 0.7 mm; Doric Lenses) were lowered 2 mm from the surface of the skull at roughly 400 μm dorsal to the virus injection site, and implants were secured to the skull with MetaBond and dental cement. After optic fiber implantation, mice were allowed to recover for 7–10 d depending on the general health.

In vivo optogenetic behavior

During behavioral testing, fiber optic cannulae were connected to patch cords (Doric Lenses), which were in turn connected to green light (532 nm) lasers (CNI Laser) using a 1 × 2 intensity division fiberoptic rotary joint (Doric Lenses) located above the cubicle containing the testing arena. Laser power was adjusted such that the light exiting the fiberoptic cable was approximately 4.5 mW. For photoinhibition experiments we used continuous green light, delivered during the first two min of the task. Laser activation was controlled by a microcontroller board (Arduino Mega 2560, Arduino).

Histology

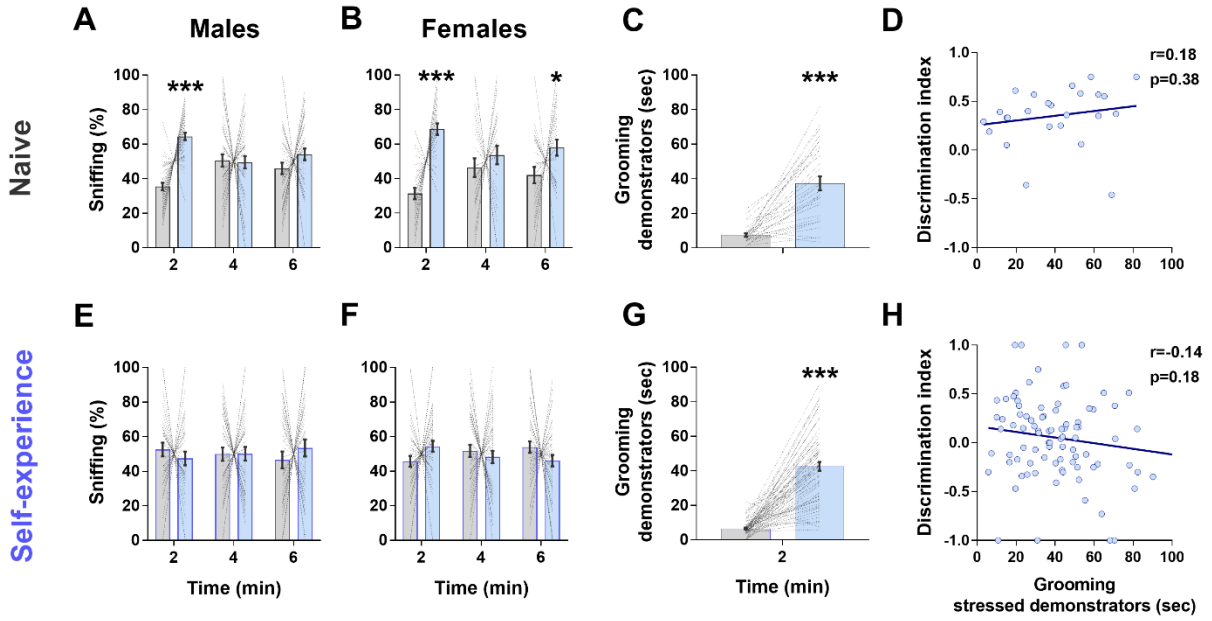
At the end of the behavioral procedures, we checked viral expression and the position of the optic fibers. Mice were deeply anesthetized (urethane 20%) and transcardially perfused with 4% paraformaldehyde in PBS at pH 7.4. Brains were dissected, fixed overnight, and cryoprotected in 30% sucrose in PBS. Coronal sections (40 μm) were cut using a HM450 microtome (Thermo Fisher Scientific). For immunohistochemical studies free-floating sections of selected areas were

washed in PBS three times for 10 min, permeabilized in PBS plus 0.4% Triton X-100 for 30 min, blocked by incubation in PBS plus 4% normal goat serum (NGS) and 0.2% Triton X-100 for 30 min (all at room temperature, 20–23 °C), and subsequently incubated with a GFP polyclonal antibody (1:1000, Invitrogen, A-11122). Primary antiserum was diluted in PBS plus 2% NGS and 0.1% Triton X-100 overnight at 4 °C. Incubated slices were washed three times in PBS plus 1% NGS for 10 min at room temperature, incubated for 2 h at room temperature with a 1:1000 dilution of Alexa Flour 488 goat anti-rabbit IgG (1:1000, Molecular Probes, A11034) in PBS plus 2% NGS and 0.1% Triton X-100, and subsequently washed three times in PBS for 10 min at room temperature. The sections were mounted on slides and covered with cover slips. All images were acquired on a Nikon 1 confocal laser scanning microscope (Nikon). Digitalized images were analyzed using Fiji (version 2.0.0, NIMH).

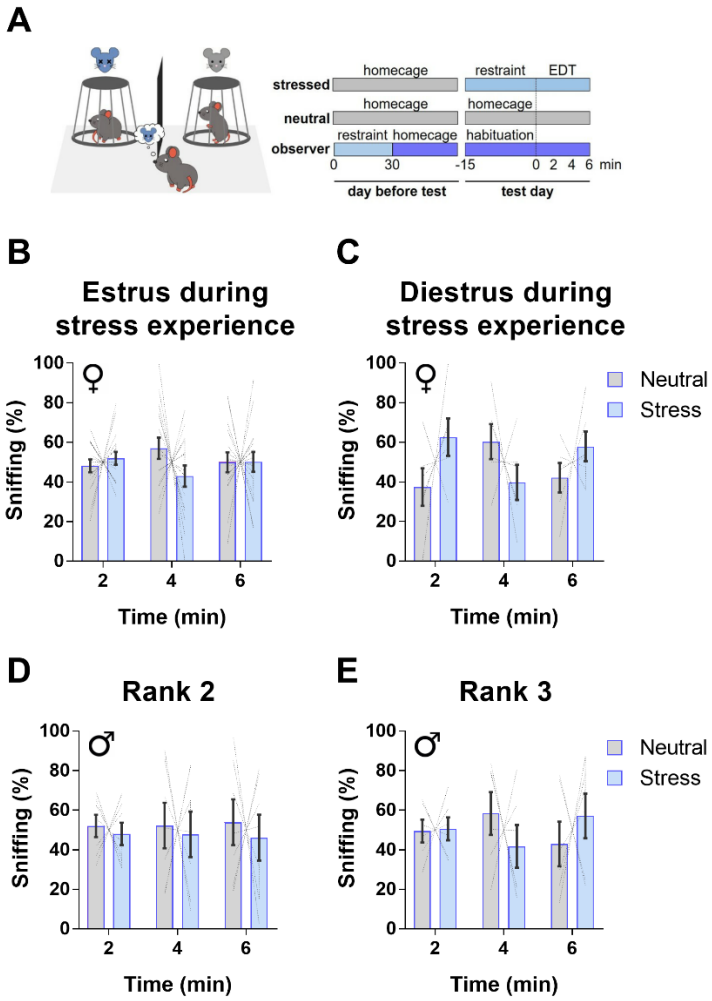
Statistics

Results are expressed as mean ± s.e.m. throughout the manuscript. For the analyses of emotion discrimination behavior, the multiple t-test followed by Holm-Sidak correction was used. Sniffing behavior was calculated as a percentage to allow direct comparison between different experimental conditions. Discrimination index and grooming behavior of observers in different conditions were analyzed using paired or unpaired t-test, depending on the samples. The accepted value for significance was $P < 0.05$. Statistical analyses were performed using GraphPad Prism 7. Numbers of mice are reported in the figure legends. Data distribution was tested using the D'Agostino and Pearson normality test. The experiments reported in this work were repeated independently two to four times, using mice from at least three different generations. No statistical methods were used to predetermine sample size for single experiments. The animal numbers were based on estimation from previous studies, including our own published studies [19, 20]. Mice were excluded post-hoc when optic fiber placement or viral expression patterns were not appropriate (outside the target region or week viral expression). For all behavioral tests, littermates were randomly assigned to the different groups. Specific randomization in the organization of the experimental conditions is described in the results and figure legends. Experimenters were not blinded during data acquisition, but all analyses were performed with blinding of the experimental conditions.

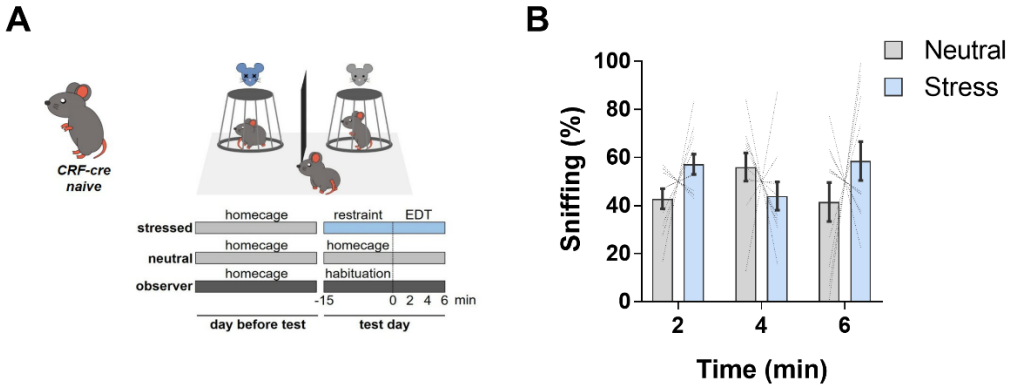
Supplementary Figures



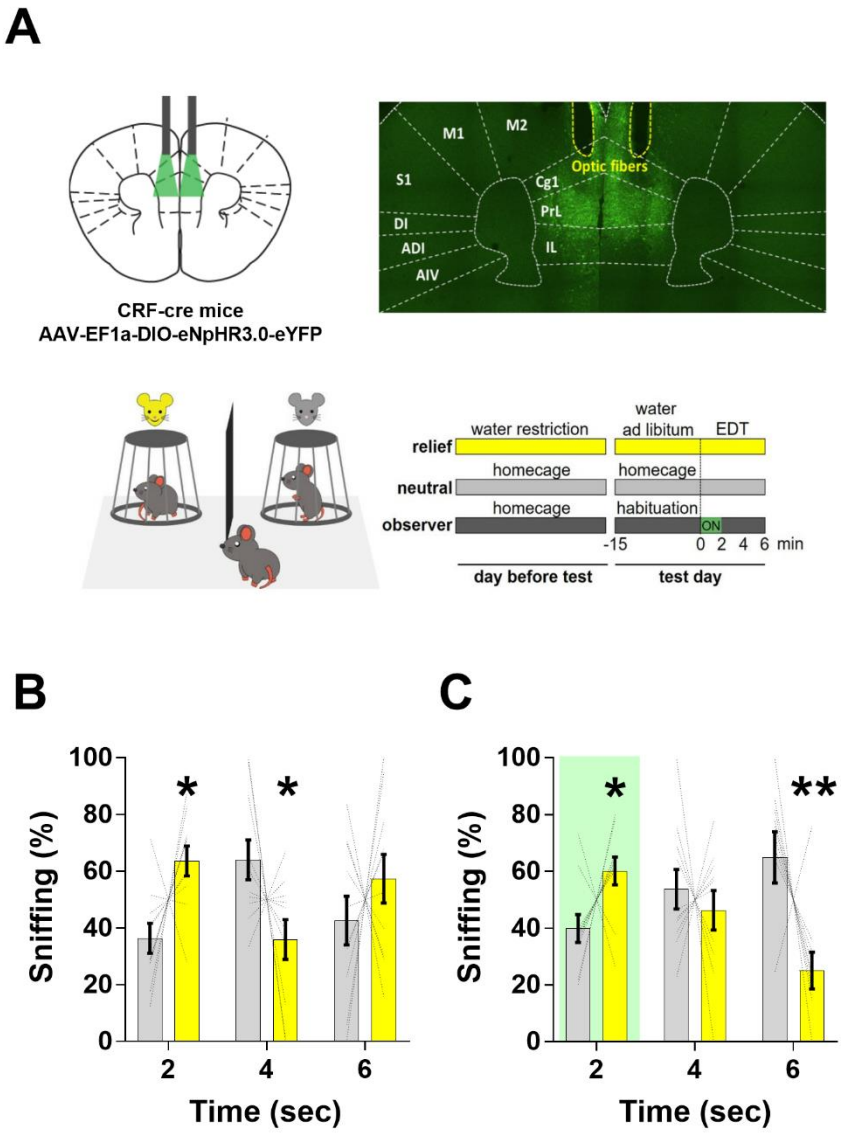
Supplementary Figure 1. Emotion discrimination abilities in naïve and self-experienced males and females, and demonstrators' grooming behavior. (**A-B**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by male (**A**) and female (**B**) naïve observers (two-tailed multiple t-test, Holm-Sidak correction. Male. 2 min: t = 6.666, d.f. = 240, P < 0.0001; 4 min: t = 0.23, d.f. = 240, P= 0.81; 6 min: t = 1.851, d.f. = 240, P= 0.13. Female. 2 min: t = 5.797, d.f. = 156, P < 0.0001; 4 min: t = 1.139, d.f. = 156, P= 0.25; 6 min: t = 2.472, d.f. = 156, P= 0.03). N=41 male observers and N=27 female observers. (**C**) Time (in seconds) spent grooming during the first 2 min of the test displayed by neutral (gray bars) or stressed (light blue bars) demonstrators in naïve condition (two-tailed unpaired t-test: t = 7.658, d.f. = 62, P < 0.0001). N=32 neutral demonstrators and N=32 stressed demonstrators. (**D**) No significant correlation was found (r = 0.18, p = 0.38) between observers' discrimination index (in y axis) and stressed demonstrators' grooming (in x axis) during the first 2 min of the test. N=25 observers and N=25 stressed demonstrators. (**E-F**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by male (**E**) and female (**F**) observer mice 24h after self-experience (two-tailed multiple t-test Holm-Sidak correction. Males. 2 min: t = 0.855, d.f. = 222, P= 0.63; 4 min: t = 0.027, d.f. = 222, P= 0.97; 6 min: t = 1.138, d.f. = 222, P= 0.58. Females. 2 min: t = 1.829, d.f. = 258, P= 0.19; 4 min: t = 0.746, d.f. = 258, P= 0.45; 6 min: t = 1.647, d.f. = 258, P= 0.19). N=38 male observers and N=44 female observers. (**G**) Time (in seconds) spent grooming during the first 2 min of the test displayed by neutral (gray bars) or stressed (light blue bars) demonstrators in self-experience condition (two-tailed unpaired t-test: t = 13.8, d.f. = 142, P < 0.0001). N=72 neutral demonstrators and N=72 stressed demonstrators. (**H**) No significant correlation was found (r=-0.14, p = 0.18) between observers' discrimination index (in y axis) and stressed demonstrators' grooming (in x axis) during the first 2 min of the test. N=89 observers and N=89 stressed demonstrators.
Bar and line graphs show mean ± s.e.m. *P < 0.05. ***P < 0.0005. Related to Figure 1.



Supplementary Figure 2. Estrus cycle and social hierarchy in self-experience modulation of emotion discrimination. (**A**) Experimental design of the EDT in self-experience condition. Observers were subjected to restraint stress for 30 min, 24 hrs before the EDT. One demonstrator (stress, blue light) was subjected to the restraint stress test for 15 min immediately before the beginning of the EDT. The other demonstrator (neutral, gray) waited undisturbed in the home cage. (**B-C**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test, which occurred 24 hrs after self-experience, displayed by observer mice that were in estrus (**B**) or diestrus (**C**) when self-experience of stress occurred (two-tailed multiple t-test, Holm-Sidak correction. Estrus during self-experience. 2 min: t = 0.566, d.f. = 90, P= 0.81; 4 min: t = 2.12, d.f. = 90, P= 0.10; 6 min: t = 0.013, d.f. = 90, P= 0.98. Diestrus during self-experience. 2 min: t = 2.064, d.f. = 30, P= 0.13; 4 min: t = 1.693, d.f. = 30, P= 0.19; 6 min: t = 1.284, d.f. = 30, P= 0.20). N=16 estral observers and N=6 diestral observers. (**D-E**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) state during the 6-min test displayed by rank 2 (**D**) or rank 3 (**E**) self-experienced subordinate mice (two-tailed multiple t-test, Holm-Sidak correction. Rank 2. 2 min: t = 0.274, d.f. = 36, P= 0.94; 4 min: t = 0.304, d.f. = 36, P= 0.94; 6 min: t = 0.55, d.f. = 36, P= 0.92. Rank 3. 2 min: t = 0.08, d.f. = 30, P= 0.93; 4 min: t = 1.214, d.f. = 30, P= 0.55; 6 min: t = 1.045, d.f. = 30, P= 0.55). N=7 rank 2 observers and N=6 rank 3 observers.
Bar and line graphs show mean ± s.e.m. Related to Figure 2 and 3.



Supplementary Figure 3. CRF-cre mice show emotion discrimination impairments. (**A**) Experimental design of CRF-Cre mice (without virus injection or lens implantation) tested in the EDT in naïve condition. Observers did not undergo any manipulation. Mice were tested in the EDT with one stressed and one neutral demonstrator. (**B**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or stressed (blue light bars) during the 6-min test, divided into three consecutive 2-min epochs, displayed by CRF-cre naïve observers (two-tailed multiple t-test, Holm-Sidak correction. 2 min: t = 1.53, d.f. = 54, P = 0.24; 4 min: t = 1.285, d.f. = 54, P = 0.24; 6 min: t = 7.82, d.f. = 54, P = 0.20). N=10 observers.
Bar and line graphs show mean ± s.e.m. Related to Figure 4.



Supplementary Figure 4. mPFC CRF neurons do not modulate discrimination of positively-valenced affective state. (**A**) up: CRF-Cre mice were injected in the mPFC with AAV-EF1a-DIO-eNpHR3.0-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. bottom: Experimental design of the EDT relief in naïve condition. Observers did not undergo any manipulation. After 23 h of water deprivation, one demonstrator was given access to water for 1 h before the test (relief, yellow), while the other demonstrator had ad libitum water access (neutral, gray). Photoinhibition was performed for 2 min from the beginning of the test using continuous green light (λ = 532 nm). (**B**-**C**) Time (in seconds) spent sniffing demonstrators in neutral (gray bars) or relieved (yellow bars) state during the 6-min test, divided into three consecutive 2-min epochs, displayed by observer mice tested in both the light OFF (**B**) and light ON (**C**) condition (two-tailed multiple t-test, Holm-Sidak correction. Light OFF. 2 min: t = 2.572, d.f. = 54, P= 0.03; 4 min: t = 2.653, d.f. = 54, P= 0.03; 6 min: t = 1.398, d.f. = 54, P= 0.16. Light ON. 2 min: t = 2.029, d.f. = 54, P= 0.05; 4 min: t = 0.750, d.f. = 54, P= 0.45; 6 min: t = 4.005, d.f. = 54, P= 0.0005). N=10 observers. Bar and line graphs show mean ± s.e.m. *P < 0.05. **P < 0.005. Related to Figure 4.

References

1. Adolphs, R. (2001). The neurobiology of social cognition. Curr Opin Neurobiol 11, 231-239.
2. Millan, M.J., Agid, Y., Brune, M., Bullmore, E.T., Carter, C.S., Clayton, N.S., Connor, R., Davis, S., Deakin, B., DeRubeis, R.J., et al. (2012). Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat Rev Drug Discov 11, 141-168.
3. Adolphs, R. (2002). Neural systems for recognizing emotion. Curr Opin Neurobiol 12, 169-177.
4. Preston, S.D., and de Waal, F.B. (2002). Empathy: Its ultimate and proximate bases. Behav Brain Sci 25, 1-20; discussion 20-71.
5. Chikovani, G., Babuadze, L., Iashvili, N., Gvalia, T., and Surguladze, S. (2015). Empathy costs: Negative emotional bias in high empathisers. Psychiatry Res 229, 340-346.
6. Melchers, M.C., Li, M., Haas, B.W., Reuter, M., Bischoff, L., and Montag, C. (2016). Similar Personality Patterns Are Associated with Empathy in Four Different Countries. Front Psychol 7, 290.
7. Olderbak, S., and Wilhelm, O. (2017). Emotion perception and empathy: An individual differences test of relations. Emotion 17, 1092-1106.
8. Israelashvili, J., Sauter, D.A., and Fischer, A.H. (2020). Different faces of empathy: Feelings of similarity disrupt recognition of negative emotions. J Exp Soc Psychol 87, 103912.
9. Israelashvili, J., Sauter, D., and Fischer, A. (2020). Two facets of affective empathy: concern and distress have opposite relationships to emotion recognition. Cogn Emot 34, 1112-1122.
10. Gallant, C., and Good, D. (2020). Examining the "reading the mind in the eyes test" as an assessment of subtle differences in affective theory of mind after concussion. Clin Neuropsychol 34, 296-317.
11. Cowan, D.G., Vanman, E.J., and Nielsen, M. (2014). Motivated empathy: the mechanics of the empathic gaze. Cogn Emot 28, 1522-1530.
12. Perry, D., Hendler, T., and Shamay-Tsoory, S.G. (2011). Projecting memories: the role of the hippocampus in emotional mentalizing. Neuroimage 54, 1669-1676.
13. Todd, R.M., Cunningham, W.A., Anderson, A.K., and Thompson, E. (2012). Affect-biased attention as emotion regulation. Trends Cogn Sci 16, 365-372.
14. Zaki, J. (2014). Empathy: a motivated account. Psychol Bull 140, 1608-1647.
15. Ferretti, V., and Papaleo, F. (2019). Understanding others: Emotion recognition in humans and other animals. Genes Brain Behav 18, e12544.
16. Panksepp, J.B., and Lahvis, G.P. (2011). Rodent empathy and affective neuroscience. Neurosci Biobehav Rev 35, 1864-1875.
17. de Waal, F.B.M., and Preston, S.D. (2017). Mammalian empathy: behavioural manifestations and neural basis. Nat Rev Neurosci 18, 498-509.
18. Decety, J. (2011). The neuroevolution of empathy. Ann N Y Acad Sci 1231, 35-45.
19. Ferretti, V., Maltese, F., Contarini, G., Nigro, M., Bonavia, A., Huang, H., Gigliucci, V., Morelli, G., Scheggia, D., Manago, F., et al. (2019). Oxytocin Signaling in the Central Amygdala Modulates Emotion Discrimination in Mice. Curr Biol 29, 1938-1953 e1936.
20. Scheggia, D., Manago, F., Maltese, F., Bruni, S., Nigro, M., Dautan, D., Latuske, P., Contarini, G., Gomez-Gonzalo, M., Requie, L.M., et al. (2020). Somatostatin interneurons in the prefrontal cortex control affective state discrimination in mice. Nat Neurosci 23, 47-60.
21. Ben-Ami Bartal, I., Decety, J., and Mason, P. (2011). Empathy and pro-social behavior in rats. Science 334, 1427-1430.
22. Burkett, J.P., Andari, E., Johnson, Z.V., Curry, D.C., de Waal, F.B., and Young, L.J. (2016). Oxytocin-dependent consolation behavior in rodents. Science 351, 375-378.
23. Church, R.M. (1959). Emotional reactions of rats to the pain of others. J Comp Physiol Psychol 52, 132-134.

24. Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H.E., Lin, S.Y., Rabah, D., Kinet, J.P., and Shin, H.S. (2010). Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. Nat Neurosci 13, 482-488.
25. Langford, D.J., Crager, S.E., Shehzad, Z., Smith, S.B., Sotocinal, S.G., Levenstadt, J.S., Chanda, M.L., Levitin, D.J., and Mogil, J.S. (2006). Social modulation of pain as evidence for empathy in mice. Science 312, 1967-1970.
26. Smith, M.L., Asada, N., and Malenka, R.C. (2021). Anterior cingulate inputs to nucleus accumbens control the social transfer of pain and analgesia. Science 371, 153-159.
27. Iarocci, G., Yager, J., and Elfers, T. (2007). What gene-environment interactions can tell us about social competence in typical and atypical populations. Brain Cogn 65, 112-127.
28. Horan, W.P., and Blanchard, J.J. (2003). Emotional responses to psychosocial stress in schizophrenia: the role of individual differences in affective traits and coping. Schizophr Res 60, 271-283.
29. Heimberg, R.G., Hofmann, S.G., Liebowitz, M.R., Schneier, F.R., Smits, J.A., Stein, M.B., Hinton, D.E., and Craske, M.G. (2014). Social anxiety disorder in DSM-5. Depress Anxiety 31, 472-479.
30. Mehling, M.H., and Tasse, M.J. (2016). Severity of Autism Spectrum Disorders: Current Conceptualization, and Transition to DSM-5. J Autism Dev Disord 46, 2000-2016.
31. Aguilera, G., and Liu, Y. (2012). The molecular physiology of CRH neurons. Front Neuroendocrinol 33, 67-84.
32. Bao, A.M., Hestiantoro, A., Van Someren, E.J., Swaab, D.F., and Zhou, J.N. (2005). Colocalization of corticotropin-releasing hormone and oestrogen receptor-alpha in the paraventricular nucleus of the hypothalamus in mood disorders. Brain 128, 1301-1313.
33. Ingallinesi, M., Rouibi, K., Le Moine, C., Papaleo, F., and Contarino, A. (2012). CRF2 receptor-deficiency eliminates opiate withdrawal distress without impairing stress coping. Mol Psychiatry 17, 1283-1294.
34. Dedic, N., Chen, A., and Deussing, J.M. (2018). The CRF Family of Neuropeptides and their Receptors - Mediators of the Central Stress Response. Curr Mol Pharmacol 11, 4-31.
35. Sanford, C.A., Soden, M.E., Baird, M.A., Miller, S.M., Schulkin, J., Palmiter, R.D., Clark, M., and Zweifel, L.S. (2017). A Central Amygdala CRF Circuit Facilitates Learning about Weak Threats. Neuron 93, 164-178.
36. Sterley, T.L., Baimoukhametova, D., Fuzesi, T., Zurek, A.A., Daviu, N., Rasiah, N.P., Rosenegger, D., and Bains, J.S. (2018). Social transmission and buffering of synaptic changes after stress. Nat Neurosci 21, 393-403.
37. Gehlert, D.R., Shekhar, A., Morin, S.M., Hipskind, P.A., Zink, C., Gackenheimer, S.L., Shaw, J., Fitz, S.D., and Sajdyk, T.J. (2005). Stress and central Urocortin increase anxiety-like behavior in the social interaction test via the CRF1 receptor. Eur J Pharmacol 509, 145-153.
38. Heinrichs, S.C. (2003). Modulation of social learning in rats by brain corticotropin-releasing factor. Brain Res 994, 107-114.
39. Cooper, M.A., and Huhman, K.L. (2005). Corticotropin-releasing factor type II (CRF-sub-2) receptors in the bed nucleus of the stria terminalis modulate conditioned defeat in Syrian hamsters (Mesocricetus auratus). Behav Neurosci 119, 1042-1051.
40. Lim, M.M., Liu, Y., Ryabinin, A.E., Bai, Y., Wang, Z., and Young, L.J. (2007). CRF receptors in the nucleus accumbens modulate partner preference in prairie voles. Horm Behav 51, 508-515.
41. Hostetler, C.M., and Ryabinin, A.E. (2013). The CRF system and social behavior: a review. Front Neurosci 7, 92.
42. Kobayashi, T., Kiyokawa, Y., Arata, S., Takeuchi, Y., and Mori, Y. (2013). c-Fos expression during the modulation of sexual behavior by an alarm pheromone. Behav Brain Res 237, 230-237.
43. Grimm, S., Wirth, K., Fan, Y., Weigand, A., Gartner, M., Feeser, M., Dziobek, I., Bajbouj, M., and Aust, S. (2017). The interaction of corticotropin-releasing hormone receptor gene and early life stress on emotional empathy. Behav Brain Res 329, 180-185.

44. Hsu, D.T., Mickey, B.J., Langenecker, S.A., Heitzeg, M.M., Love, T.M., Wang, H., Kennedy, S.E., Pecina, M., Shafir, T., Hodgkinson, C.A., et al. (2012). Variation in the corticotropin-releasing hormone receptor 1 (CRHR1) gene influences fMRI signal responses during emotional stimulus processing. J Neurosci 32, 3253-3260.
45. Kubota, Y., Shigematsu, N., Karube, F., Sekigawa, A., Kato, S., Yamaguchi, N., Hirai, Y., Morishima, M., and Kawaguchi, Y. (2011). Selective coexpression of multiple chemical markers defines discrete populations of neocortical GABAergic neurons. Cereb Cortex 21, 1803-1817.
46. Chen, P., Lou, S., Huang, Z.H., Wang, Z., Shan, Q.H., Wang, Y., Yang, Y., Li, X., Gong, H., Jin, Y., et al. (2020). Prefrontal Cortex Corticotropin-Releasing Factor Neurons Control Behavioral Style Selection under Challenging Situations. Neuron 106, 301-315 e307.
47. Dalley, J.W., Cardinal, R.N., and Robbins, T.W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev 28, 771-784.
48. Frith, C.D., and Frith, U. (2012). Mechanisms of social cognition. Annu Rev Psychol 63, 287-313.
49. Hiser, J., and Koenigs, M. (2018). The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making, Social Cognition, and Psychopathology. Biol Psychiatry 83, 638-647.
50. Kawasaki, H., Adolphs, R., Oya, H., Kovach, C., Damasio, H., Kaufman, O., and Howard, M., 3rd (2005). Analysis of single-unit responses to emotional scenes in human ventromedial prefrontal cortex. J Cogn Neurosci 17, 1509-1518.
51. Yuen, E.Y., Liu, W., Karatsoreos, I.N., Feng, J., McEwen, B.S., and Yan, Z. (2009). Acute stress enhances glutamatergic transmission in prefrontal cortex and facilitates working memory. Proc Natl Acad Sci U S A 106, 14075-14079.
52. Liston, C., Miller, M.M., Goldwater, D.S., Radley, J.J., Rocher, A.B., Hof, P.R., Morrison, J.H., and McEwen, B.S. (2006). Stress-induced alterations in prefrontal cortical dendritic morphology predict selective impairments in perceptual attentional set-shifting. J Neurosci 26, 7870-7874.
53. McKlveen, J.M., Myers, B., and Herman, J.P. (2015). The medial prefrontal cortex: coordinator of autonomic, neuroendocrine and behavioural responses to stress. J Neuroendocrinol 27, 446-456.
54. Arnsten, A.F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. Nat Rev Neurosci 10, 410-422.
55. McEwen, B.S., and Morrison, J.H. (2013). The brain on stress: vulnerability and plasticity of the prefrontal cortex over the life course. Neuron 79, 16-29.
56. Derntl, B., Hack, R.L., Kryspin-Exner, I., and Habel, U. (2013). Association of menstrual cycle phase with the core components of empathy. Horm Behav 63, 97-104.
57. Guapo, V.G., Graeff, F.G., Zani, A.C., Labate, C.M., dos Reis, R.M., and Del-Ben, C.M. (2009). Effects of sex hormonal levels and phases of the menstrual cycle in the processing of emotional faces. Psychoneuroendocrinology 34, 1087-1094.
58. Pearson, R., and Lewis, M.B. (2005). Fear recognition across the menstrual cycle. Horm Behav 47, 267-271.
59. Mikosz, M., Nowak, A., Werka, T., and Knapska, E. (2015). Sex differences in social modulation of learning in rats. Sci Rep 5, 18114.
60. Cora, M.C., Kooistra, L., and Travlos, G. (2015). Vaginal Cytology of the Laboratory Rat and Mouse: Review and Criteria for the Staging of the Estrous Cycle Using Stained Vaginal Smears. Toxicol Pathol 43, 776-793.
61. McLean, A.C., Valenzuela, N., Fai, S., and Bennett, S.A. (2012). Performing vaginal lavage, crystal violet staining, and vaginal cytological evaluation for mouse estrous cycle staging identification. J Vis Exp, e4389.
62. Ackermann, S., Spalek, K., Rasch, B., Gschwind, L., Coynel, D., Fastenrath, M., Papassotiropoulos, A., and de Quervain, D.J. (2012). Testosterone levels in healthy men are related to amygdala reactivity and memory performance. Psychoneuroendocrinology 37, 1417-1424.
63. Bos, P.A., Hofman, D., Hermans, E.J., Montoya, E.R., Baron-Cohen, S., and van Honk, J. (2016). Testosterone reduces functional connectivity during the 'Reading the Mind in the Eyes' Test. Psychoneuroendocrinology 68, 194-201.

64. Vongas, J.G., and Al Hajj, R. (2017). The effects of competition and implicit power motive on men's testosterone, emotion recognition, and aggression. Horm Behav 92, 57-71.
65. Ely, D.L. (1981). Hypertension, social rank, and aortic arteriosclerosis in CBA/J mice. Physiol Behav 26, 655-661.
66. Machida, T., Yonezawa, Y., and Noumura, T. (1981). Age-associated changes in plasma testosterone levels in male mice and their relation to social dominance or subordinance. Horm Behav 15, 238-245.
67. Jones, C.E., and Monfils, M.H. (2016). Dominance status predicts social fear transmission in laboratory rats. Anim Cogn 19, 1051-1069.
68. Fan, Z., Zhu, H., Zhou, T., Wang, S., Wu, Y., and Hu, H. (2019). Using the tube test to measure social hierarchy in mice. Nat Protoc 14, 819-831.
69. Schmidt, E.E., Taylor, D.S., Prigge, J.R., Barnett, S., and Capecchi, M.R. (2000). Illegitimate Cre-dependent chromosome rearrangements in transgenic mouse spermatids. Proc Natl Acad Sci U S A 97, 13702-13707.
70. Schmidt-Supprian, M., Wunderlich, F.T., and Rajewsky, K. (2007). Excision of the Frt-flanked neo (R) cassette from the CD19cre knock-in transgene reduces Cre-mediated recombination. Transgenic Res 16, 657-660.
71. Harno, E., Cottrell, E.C., and White, A. (2013). Metabolic pitfalls of CNS Cre-based technology. Cell Metab 18, 21-28.
72. Kim, T.H., Richards, K., Heng, J., Petrou, S., and Reid, C.A. (2013). Two lines of transgenic mice expressing cre-recombinase exhibit increased seizure susceptibility. Epilepsy Res 104, 11-16.
73. Giusti, S.A., Vercelli, C.A., Vogl, A.M., Kolarz, A.W., Pino, N.S., Deussing, J.M., and Refojo, D. (2014). Behavioral phenotyping of Nestin-Cre mice: implications for genetic mouse models of psychiatric disorders. J Psychiatr Res 55, 87-95.
74. Davis, M.H., Mitchell, K.V., Hall, J.A., Lothert, J., Snapp, T., and Meyer, M. (1999). Empathy, expectations, and situational preferences: personality influences on the decision to participate in volunteer helping behaviors. J Pers 67, 469-503.
75. Grynberg, D., and Lopez-Perez, B. (2018). Facing others' misfortune: Personal distress mediates the association between maladaptive emotion regulation and social avoidance. PLoS One 13, e0194248.
76. Sanders, J., Mayford, M., and Jeste, D. (2013). Empathic fear responses in mice are triggered by recognition of a shared experience. PLoS One 8, e74609.
77. Masuda, A.A., S. (2009). Social transmission of avoidance behavior under situational change in learned and unlearned rats. PLoS One 4.
78. Kim, E.J., Kim, E.S., Covey, E., and Kim, J.J. (2010). Social transmission of fear in rats: the role of 22-kHz ultrasonic distress vocalization. PLoS One 5, e15077.
79. Atsak, P., Orre, M., Bakker, P., Cerliani, L., Roozendaal, B., Gazzola, V., Moita, M., and Keysers, C. (2011). Experience modulates vicarious freezing in rats: a model for empathy. PLoS One 6, e21855.
80. Sato, N., Tan, L., Tate, K., and Okada, M. (2015). Rats demonstrate helping behavior toward a soaked conspecific. Anim Cogn 18, 1039-1047.
81. Cox, S.S., and Reichel, C.M. (2020). Rats display empathic behavior independent of the opportunity for social interaction. Neuropsychopharmacology 45, 1097-1104.
82. Kim, A., Keum, S., and Shin, H.S. (2018). Observational fear behavior in rodents as a model for empathy. Genes Brain Behav, e12521.
83. Knapska, E., Mikosz, M., Werka, T., and Maren, S. (2010). Social modulation of learning in rats. Learn Mem 17, 35-42.
84. Pereira, A.G., Cruz, A., Lima, S.Q., and Moita, M.A. (2012). Silence resulting from the cessation of movement signals danger. Curr Biol 22, R627-628.
85. Jones, C.E., Riha, P.D., Gore, A.C., and Monfils, M.H. (2014). Social transmission of Pavlovian fear: fear-conditioning by-proxy in related female rats. Anim Cogn 17, 827-834.

86. Bruchey, A.K., Jones, C.E., and Monfils, M.H. (2010). Fear conditioning by-proxy: social transmission of fear during memory retrieval. Behav Brain Res 214, 80-84.
87. Guzman, Y.F., Tronson, N.C., Guedea, A., Huh, K.H., Gao, C., and Radulovic, J. (2009). Social modeling of conditioned fear in mice by non-fearful conspecifics. Behav Brain Res 201, 173-178.
88. Bredy, T.W., and Barad, M. (2009). Social modulation of associative fear learning by pheromone communication. Learn Mem 16, 12-18.
89. Zilioli, S., Caldbick, E., and Watson, N.V. (2014). Testosterone reactivity to facial display of emotions in men and women. Horm Behav 65, 461-468.
90. Osorio, F.L., de Paula Cassis, J.M., Machado de Sousa, J.P., Poli-Neto, O., and Martin-Santos, R. (2018). Sex Hormones and Processing of Facial Expressions of Emotion: A Systematic Literature Review. Front Psychol 9, 529.
91. Forkosh, O., Karamihalev, S., Roeh, S., Alon, U., Anpilov, S., Touma, C., Nussbaumer, M., Flachskamm, C., Kaplick, P.M., Shemesh, Y., et al. (2019). Identity domains capture individual differences from across the behavioral repertoire. Nat Neurosci 22, 2023-2028.
92. Colas-Zelin, D., Light, K.R., Kolata, S., Wass, C., Denman-Brice, A., Rios, C., Szalk, K., and Matzel, L.D. (2012). The imposition of, but not the propensity for, social subordination impairs exploratory behaviors and general cognitive abilities. Behav Brain Res 232, 294-305.
93. Matzel, L.D., Kolata, S., Light, K., and Sauce, B. (2017). The tendency for social submission predicts superior cognitive performance in previously isolated male mice. Behav Processes 134, 12-21.
94. Larrieu, T., Cherix, A., Duque, A., Rodrigues, J., Lei, H., Gruetter, R., and Sandi, C. (2017). Hierarchical Status Predicts Behavioral Vulnerability and Nucleus Accumbens Metabolic Profile Following Chronic Social Defeat Stress. Curr Biol 27, 2202-2210 e2204.
95. Bartolomucci, A. (2007). Social stress, immune functions and disease in rodents. Front Neuroendocrinol 28, 28-49.
96. Millan, M.J., and Bales, K.L. (2013). Towards improved animal models for evaluating social cognition and its disruption in schizophrenia: the CNTRICS initiative. Neurosci Biobehav Rev 37, 2166-2180.
97. Maat, A., van Montfort, S.J., de Nijs, J., Derks, E.M., Kahn, R.S., Linszen, D.H., van Os, J., Wiersma, D., Bruggeman, R., Cahn, W., et al. (2015). Emotion processing in schizophrenia is state and trait dependent. Schizophr Res 161, 392-398.
98. Kai, Y., Li, Y., Sun, T., Yin, W., Mao, Y., Li, J., Xie, W., Chen, S., Wang, L., Li, J., et al. (2018). A medial prefrontal cortex-nucleus acumens corticotropin-releasing factor circuitry for neuropathic pain-increased susceptibility to opioid reward. Transl Psychiatry 8, 100.
99. Hupalo, S., Martin, A.J., Green, R.K., Devilbiss, D.M., and Berridge, C.W. (2019). Prefrontal Corticotropin-Releasing Factor (CRF) Neurons Act Locally to Modulate Frontostriatal Cognition and Circuit Function. J Neurosci 39, 2080-2090.
100. Kim, J., Lee, S., Fang, Y.Y., Shin, A., Park, S., Hashikawa, K., Bhat, S., Kim, D., Sohn, J.W., Lin, D., et al. (2019). Rapid, biphasic CRF neuronal responses encode positive and negative valence. Nat Neurosci 22, 576-585.
101. Hartley, N.D., Gaulden, A.D., Baldi, R., Winters, N.D., Salimando, G.J., Rosas-Vidal, L.E., Jameson, A., Winder, D.G., and Patel, S. (2019). Dynamic remodeling of a basolateral-to-central amygdala glutamatergic circuit across fear states. Nat Neurosci 22, 2000-2012.
102. Chen, E., Lallai, V., Sherafat, Y., Grimes, N.P., Pushkin, A.N., Fowler, J.P., and Fowler, C.D. (2018). Altered Baseline and Nicotine-Mediated Behavioral and Cholinergic Profiles in ChAT-Cre Mouse Lines. J Neurosci 38, 2177-2188.
103. Palle, A., Zorzo, C., Luskey, V.E., McGreevy, K.R., Fernandez, S., and Trejo, J.L. (2019). Social dominance differentially alters gene expression in the medial prefrontal cortex without affecting adult hippocampal neurogenesis or stress and anxiety-like behavior. FASEB J 33, 6995-7008.
104. Brockhurst, J., Cheleuitte-Nieves, C., Buckmaster, C.L., Schatzberg, A.F., and Lyons, D.M. (2015). Stress inoculation modeled in mice. Transl Psychiatry 5, e537.

CHAPTER 6

General Conclusions and Future perspectives

The aim of this thesis was to investigate in mice emotion recognition, the ability to perceive and understand others' emotional state [1], in order to dissect the neuronal mechanisms underlying these process. We developed a novel behavioral setting, the Emotion discrimination Task (EDT), with which we could detect mice ability to discriminate conspecifics based on their altered affective state. In particular, we found that mice prefer to spend more time sniffing demonstrators in a positively- (relief) or negatively- (fear and stress) valenced state compared to neutral mice (Fig. 1A). Moreover, a comprehensive investigation on the sensory modalities guiding emotion discrimination revealed that visual and olfactory cues are implicated in this process. We demonstrated also that self-experience of an aversive state can influence in an individual way the responses to the same altered affective states in others. In particular, we found that following self-experience, a subpopulation of mice prefer sniffing more stressed demonstrators, while another subpopulation avoid them. We revealed that sexual hormones could modulate this process, showing that females in estrus are not affected by self-experience, while dominant mice are more sensible to it (Fig. 1B). All these evidence are in line with human modulation of emotion recognition and are possibly related to emphatic behaviors [2-5].

Overall, we were able to establish a mouse paradigm to assess emotion discrimination, approximating key features of equivalent tasks used in humans [6]. Our EDT is already well accepted in the scientific community and other laboratories already started using it [7]. Notably, emotion recognition processes are impaired in many psychiatric and neurodevelopmental disorders such as schizophrenia and autism spectrum disorders [8-12]. Thus, application of our task could be useful in understanding the neuropathology underlying social dysfunctions in these disorders.

Investigation of emotion recognition and empathy neural basis has found large interest in the last decades. In particular, validations of rodents' models for these processes could open the way to uncover with a cell-, circuit- and system-specificity the brain mechanisms implicated. This would allow for thorough preclinical screening of potential therapeutic targets and agents. However, there are still several open questions concerning the reliability of studying social cognition in rodents. Indeed, in rodents, these processes appear to be mediated by rapid automatic responses, with no evidence of conscious processing that is a crucial component of human social cognition. To overcome these limitations, continuous systemic comparisons between patterns of human and rodent brain activation is crucial. The evidence reported in this thesis, together with others in literature [13-17], are promising indications that confirm overlapping substrates (e.g.

amygdala and mPFC) between humans' and rodents' emotional processes. Indeed, the paradigm here developed, opened the way to mechanistic studies enabling to investigate the neuronal mechanisms underlying emotion recognition processes. In particular, given that recognition of socially and emotionally relevant information requires a network of neural assembles [18-21], different brain circuits were investigated.

First, using a combination of anatomical and chemogenetic approaches, we found that OXTergic projections from the paraventricular nucleus of the hypothalamus (PVN) to the central amygdala (CeA) are crucial for the discrimination of both positively and negatively valenced emotional states (Fig. 1C). This finding was further supported by using a genetic mouse model of cognitive liability (dysbindin-1 mouse model), which we found to have low OXT receptor (OXTR) levels in the CeA. We demonstrated that virally mediated enhancement of CeA OXTR was sufficient to rescue emotion discrimination deficits of these mice. Oxytocin plays a pivotal role in social functions, and in humans, it has been associated with social cognitive behaviors such as emotion recognition, empathy and trust [22-24]. Our findings provide first evidence of a specific involvement of the endogenous OXT system in emotion discrimination. In particular, we demonstrated that OXT signaling within the CeA plays a key role in emotion discrimination both in physiological and pathological conditions.

Second, using in vivo electrophysiology, we found that medial prefrontal cortex (mPFC) neurons activation during social exploration was differently associated with the affective state of the conspecific. Moreover, optogenetic manipulations revealed a specific involvement of somatostatin (SOM+) interneurons during discrimination of altered emotional states (Fig. 1E-F). In particular, by using in vivo single cell microendoscopic Ca2+ imaging, we found that an increased synchronous activity of mPFC SOM+ interneurons is associated with affective state discrimination. The mPFC represents an important core of emotion recognition as reported by a number of fMRI studies in humans [18, 20, 25-27]. Thus, we extended the notion that the mPFC is critical for emotion discrimination, and more specifically, we disentangled the selective mPFC microcircuit needed to support this process.

Finally, our preliminary findings on the CRF system suggest its involvement in discrimination of negatively-, but not positively-, valenced affective states. Moreover, we reported first evidence that mPFC-CRF neurons might be implicated in dominance-dependent mechanisms underlying self-experience modulation of emotion discrimination (Fig. 1D). This evidence are in

agreement with the notion that the CRF system is implicated in both emotional and cognitive components of responses to stress [28-30]. Moreover, our findings further elucidate mPFC mechanisms involved in emotion discrimination and its modulation by self-experience.

The data reported in this thesis highlighted subjective modulations of emotion recognition and empathic behaviors. Among all the possible individual aspects, genetic factors have been strongly implicated in modulation of emotion-processing in rodents [31-36], as well as in social cognition deficits associated with several psychiatric disorders [37-40]. Thus, a deeper investigation on genetic factors' contribution to social cognitive abilities will be required to facilitate at-risk subpopulations' identification and consequently to develop medical approaches based on the genetic background of each individual.

Overall, the findings reported in this thesis unravel new insights into the behavioral and neuronal mechanisms underlying emotion recognition, providing selective targets for manipulation of this ability with clinical relevance.

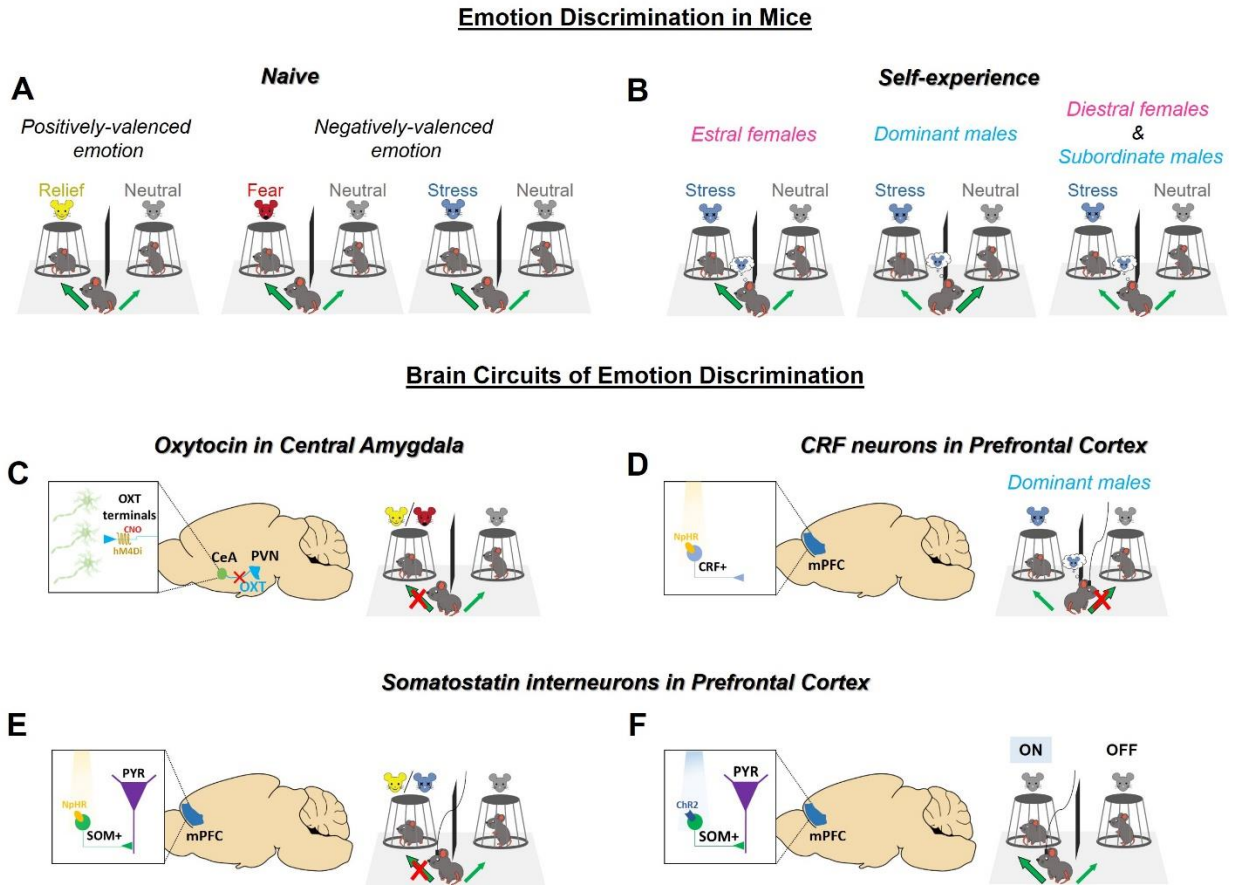


Figure 1. Brain circuits of Emotion Discrimination in mice. (**A**) Graphical representation of the Emotion Discrimination task designed for mice. Naïve observers exposed simultaneously to two demonstators, one in a neutral state and the other in either positively- (relief) or negatively- (fear and stress) valenced state, spend more time sniffing the emotionally altered mouse. (**B**) Graphical representation of the Emotion Discrimination task in observer mice that had experienced a stressfull event. *Left*, estral females spend more time sniffing the demonstrator that experienced the same stressfull event. *Middle*, dominant males spend more time sniffing the neutral demonstrator. *Right*, diestral females and subordinate males do not show any preference for either the neutral demonstrator or the mouse that have experienced the same stressfull event. (**C**) Schematic representation of chemogenetic manipulation of OXTergic projections from the paraventricular nucleus of the hypothalamus (PVN) to the central amygdala (CeA). Silencing these projections leads to impairments in mice ability to discriminate both positively (relief) and negatively (fear) valenced emotional states. (**D**) Schematic representation of optogenetic manipulation of mPFC CRF neurons. Silencing these neurons leads to abolishment of dominance-dependent mechanisms underlying self-experience modulation of emotion discrimination. (**E-F**) Schematic representation of optogenetic manipulation of mPFC SOM neurons. (**E**) Silencing these neurons leads to impairments in mice ability to discriminate both positively (relief) and negatively (stress) valenced emotional states. (**F**) Exposing observers to two neutral mice and activating mPFC SOM neurons everytime they enter the zone of one of the two demonstrators leads to develop a preference towards this neutral mouse.

References

1. Adolphs, R. (2002). Neural systems for recognizing emotion. Curr Opin Neurobiol 12, 169-177.
2. Israelashvili, J., Sauter, D., and Fischer, A. (2020). Two facets of affective empathy: concern and distress have opposite relationships to emotion recognition. Cogn Emot 34, 1112-1122.
3. Millan, M.J., and Bales, K.L. (2013). Towards improved animal models for evaluating social cognition and its disruption in schizophrenia: the CNTRICS initiative. Neurosci Biobehav Rev 37, 2166-2180.
4. Olderbak, S., and Wilhelm, O. (2017). Emotion perception and empathy: An individual differences test of relations. Emotion 17, 1092-1106.
5. Chikovani, G., Babuadze, L., Iashvili, N., Gvalia, T., and Surguladze, S. (2015). Empathy costs: Negative emotional bias in high empathisers. Psychiatry Res 229, 340-346.
6. Ekman, P., Levenson, R.W., and Friesen, W.V. (1983). Autonomic nervous system activity distinguishes among emotions. Science 221, 1208-1210.
7. Smith, M.L., Asada, N., and Malenka, R.C. (2021). Anterior cingulate inputs to nucleus accumbens control the social transfer of pain and analgesia. Science 371, 153-159.
8. Gur, R.E., McGrath, C., Chan, R.M., Schroeder, L., Turner, T., Turetsky, B.I., Kohler, C., Alsop, D., Maldjian, J., Ragland, J.D., et al. (2002). An fMRI study of facial emotion processing in patients with schizophrenia. Am J Psychiatry 159, 1992-1999.
9. Green, M.F., Horan, W.P., and Lee, J. (2015). Social cognition in schizophrenia. Nat Rev Neurosci 16, 620-631.
10. Black, M.H., Chen, N.T.M., Iyer, K.K., Lipp, O.V., Bolte, S., Falkmer, M., Tan, T., and Girdler, S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. Neurosci Biobehav Rev 80, 488-515.
11. Harms, M.B., Martin, A., and Wallace, G.L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. Neuropsychol Rev 20, 290-322.
12. Xavier, J., Vignaud, V., Ruggiero, R., Bodeau, N., Cohen, D., and Chaby, L. (2015). A Multidimensional Approach to the Study of Emotion Recognition in Autism Spectrum Disorders. Front Psychol 6, 1954.
13. Jeon, D., Kim, S., Chetana, M., Jo, D., Ruley, H.E., Lin, S.Y., Rabah, D., Kinet, J.P., and Shin, H.S. (2010). Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. Nat Neurosci 13, 482-488.
14. Jones, C.E., and Monfils, M.H. (2016). Dominance status predicts social fear transmission in laboratory rats. Anim Cogn 19, 1051-1069.
15. Twining, R.C., Vantrease, J.E., Love, S., Padival, M., and Rosenkranz, J.A. (2017). An intra-amygdala circuit specifically regulates social fear learning. Nat Neurosci 20, 459-469.
16. Rogers-Carter, M.M., Varela, J.A., Gribbons, K.B., Pierce, A.F., McGoey, M.T., Ritchey, M., and Christianson, J.P. (2018). Insular cortex mediates approach and avoidance responses to social affective stimuli. Nat Neurosci 21, 404-414.
17. Allsop, S.A., Wichmann, R., Mills, F., Burgos-Robles, A., Chang, C.J., Felix-Ortiz, A.C., Vienne, A., Beyeler, A., Izadmehr, E.M., Glober, G., et al. (2018). Corticoamygdala Transfer of Socially Derived Information Gates Observational Learning. Cell 173, 1329-1342 e1318.
18. Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annu Rev Psychol 60, 693-716.
19. de Waal, F.B.M., and Preston, S.D. (2017). Mammalian empathy: behavioural manifestations and neural basis. Nat Rev Neurosci 18, 498-509.
20. Frith, C.D., and Frith, U. (2012). Mechanisms of social cognition. Annu Rev Psychol 63, 287-313.
21. Vuilleumier, P., and Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. Neuropsychologia 45, 174-194.

22. De Dreu, C.K., and Kret, M.E. (2016). Oxytocin Conditions Intergroup Relations Through Upregulated In-Group Empathy, Cooperation, Conformity, and Defense. Biol Psychiatry 79, 165-173.
23. Guastella, A.J., Mitchell, P.B., and Mathews, F. (2008). Oxytocin enhances the encoding of positive social memories in humans. Biol Psychiatry 64, 256-258.
24. Domes, G., Sibold, M., Schulze, L., Lischke, A., Herpertz, S.C., and Heinrichs, M. (2013). Intranasal oxytocin increases covert attention to positive social cues. Psychol Med 43, 1747-1753.
25. Dalley, J.W., Cardinal, R.N., and Robbins, T.W. (2004). Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. Neurosci Biobehav Rev 28, 771-784.
26. Kawasaki, H., Adolphs, R., Oya, H., Kovach, C., Damasio, H., Kaufman, O., and Howard, M., 3rd (2005). Analysis of single-unit responses to emotional scenes in human ventromedial prefrontal cortex. J Cogn Neurosci 17, 1509-1518.
27. Hiser, J., and Koenigs, M. (2018). The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making, Social Cognition, and Psychopathology. Biol Psychiatry 83, 638-647.
28. Dedic, N., Chen, A., and Deussing, J.M. (2018). The CRF Family of Neuropeptides and their Receptors - Mediators of the Central Stress Response. Curr Mol Pharmacol 11, 4-31.
29. Sanford, C.A., Soden, M.E., Baird, M.A., Miller, S.M., Schulkin, J., Palmiter, R.D., Clark, M., and Zweifel, L.S. (2017). A Central Amygdala CRF Circuit Facilitates Learning about Weak Threats. Neuron 93, 164-178.
30. Sterley, T.L., Baimoukhametova, D., Fuzesi, T., Zurek, A.A., Daviu, N., Rasiah, N.P., Rosenegger, D., and Bains, J.S. (2018). Social transmission and buffering of synaptic changes after stress. Nat Neurosci 21, 393-403.
31. Laviola, G., Zoratto, F., Ingiosi, D., Carito, V., Huzard, D., Fiore, M., and Macri, S. (2017). Low empathy-like behaviour in male mice associates with impaired sociability, emotional memory, physiological stress reactivity and variations in neurobiological regulations. PLoS One 12, e0188907.
32. Zoratto, F., Sbriccoli, M., Martinelli, A., Glennon, J.C., Macri, S., and Laviola, G. (2018). Intranasal oxytocin administration promotes emotional contagion and reduces aggression in a mouse model of callousness. Neuropharmacology 143, 250-267.
33. Keum, S., Park, J., Kim, A., Park, J., Kim, K.K., Jeong, J., and Shin, H.S. (2016). Variability in empathic fear response among 11 inbred strains of mice. Genes Brain Behav 15, 231-242.
34. Keum, S., Kim, A., Shin, J.J., Kim, J.H., Park, J., and Shin, H.S. (2018). A Missense Variant at the Nrxn3 Locus Enhances Empathy Fear in the Mouse. Neuron 98, 588-601 e585.
35. Barak, B., and Feng, G. (2016). Neurobiology of social behavior abnormalities in autism and Williams syndrome. Nat Neurosci 19, 647-655.
36. Meyza, K., Nikolaev, T., Kondrakiewicz, K., Blanchard, D.C., Blanchard, R.J., and Knapska, E. (2015). Neuronal correlates of asocial behavior in a BTBR T (+) Itpr3(tf)/J mouse model of autism. Front Behav Neurosci 9, 199.
37. Savage, J.E., Rose, R.J., Pulkkinen, L., Silventoinen, K., Korhonen, T., Kaprio, J., Gillespie, N., and Dick, D.M. (2018). Early maturation and substance use across adolescence and young adulthood: A longitudinal study of Finnish twins. Dev Psychopathol 30, 79-92.
38. Davies, G., Lam, M., Harris, S.E., Trampush, J.W., Luciano, M., Hill, W.D., Hagenaars, S.P., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun 9, 2098.
39. Warrier, V., Toro, R., Chakrabarti, B., i, P.-B.a.g., Borglum, A.D., Grove, J., andMe Research, T., Hinds, D.A., Bourgeron, T., and Baron-Cohen, S. (2018). Genome-wide analyses of self-reported empathy: correlations with autism, schizophrenia, and anorexia nervosa. Transl Psychiatry 8, 35.
40. Germine, L., Robinson, E.B., Smoller, J.W., Calkins, M.E., Moore, T.M., Hakonarson, H., Daly, M.J., Lee, P.H., Holmes, A.J., Buckner, R.L., et al. (2016). Association between polygenic risk for schizophrenia, neurocognition and social cognition across development. Transl Psychiatry 6, e924.

AKNOWLEDGEMENTS

I would like to express my gratitude for my supervisor Dr. Francesco Papaleo, for sharing his knowledge and enthusiasm, for his constructive criticism and for encouraging my ideas and research. I would like to thank also my other “mentors” Valentina Ferretti, Diego Scheggia and Francesca Managò, for their teaching, support and friendship. Their guidance helped me growing as a scientist during my Ph.D. I would like to thank the Animal Facility staff for excellent technical assistance and the care they have for our colonies. Finally, I would like to thank my fellow labmates for sharing with me this inspiring and exciting Ph.D experience and for their sincere friendship: Gabriella Trigilio, Gabriella Contarini, Giulia Castellani, Francesca Scarsi and Giada Pacinelli.