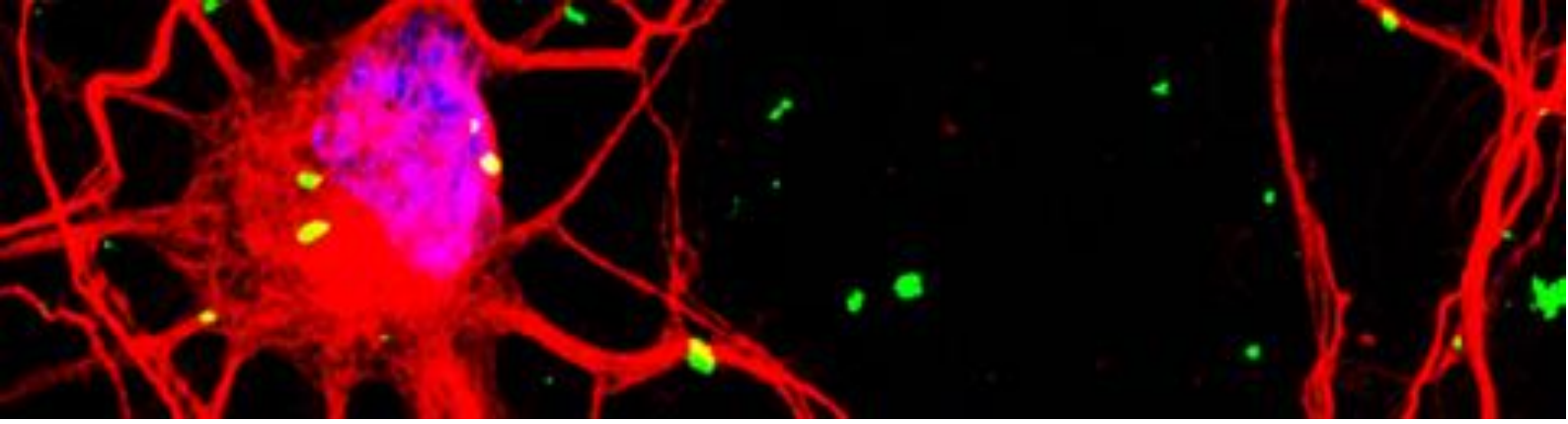


Corso di Dottorato in Neuroscienze
Curriculum Neuroscienze e Neurotecnologie
Ciclo XXXI

Titolo:

Cortical and subcortical neuronal substrates of social behavior

Autore: Marco Nigro
Supervisor: Francesco Papaleo

ISTITUTO ITALIANO DI TECNOLOGIA

UNIVERSITÀ DEGLI STUDI DI GENOVA

eulogy to Cullen B. Owens:

The blind men and the elephant is an old Indian parable that illustrates the inaccessibility of the nature of the truth. It has been adopted across many religions and cultures and interpreted in various ways. The story goes that a group of blind men attempts to touch the elephant in order to find out exactly what it is. As the different men touch different parts of the elephant's body, they come to different conclusion as to the true nature of the object of inquiry. For example, one man touches the leg and describes the object as pillar, while another man touches the tail and says it is a rope. Again another feels the trunk and deduces it is a tree branch, the ears feel like a hand fan, the belly a wall, and the tusk a solid pipe. Conflict between the men and their interpretation ensue.

This illustration seems fitting to the present thesis as a way to not only exemplify the power of neuroscience, but also to admit its limitations. Whereas we know that our subjective experience is true, it may not be the totally of the truth. In science we attempt to discover how something works, but our limitations are elucidate through only what is possible to perceive. Although different versions of the parable resolve the men's conflict in different ways, I choose to see a scientist's ending to the story, where the men stop arguing, begin listening and collaborate their experiences to come to know the "whole" elephant.

Illustration by Andrea Fiore

For my parents, girlfriend and friends, who always support me

Table of Contents

Chapter 1 General introduction 8
Medial Prefrontal Cortex 11
Oxytocin system 12
Dopamine system 13
Dysbindin mice 14
Scope of the Thesis 15
Chapter 2 Oxytocin-Dependent Emotion Recognition in Mice 18
Introduction 19
Results 20
Mice are able to discriminate unfamiliar conspecifics based on negative-valence emotional states 20
Mice are able to discriminate unfamiliar conspecifics based on positive-valence emotional states 24
Impact of sensory cues in the mouse emotion recognition task 27
Emotion discrimination abilities are a stable trait, not relatable to sociability 29
Endogenous release of oxytocin is necessary for emotion discrimination 30
PVN OXT projections to the central amygdala are an essential neural substrate for emotion recognition 34
Emotion recognition depends on OXTR levels in the CeA 39
Discussion 44
Materials and Methods 47
Emotion recognition task 47
“Classic” social interaction test and 3-chamber social interaction test 49
One-on-one social exploration tests 50
Sensory modality assessment 50
Place conditioning 51
Viral vectors 52
Corticosterone assay 52
Stereotaxic Injections 53
Drugs 53
Histology 53
Brain Autoradiography 54

Ex vivo electrophysiology 54
Statistical analyses 55
Chapter 3 Somatostatin interneurons within the prefrontal cortex control emotion recognition in mice 56
Introduction 57
Results 58
Mice can discriminate conspecifics based on the emotional state 58
Discrimination of a negative affective state 63
Emotion discrimination is a stable trait distinct from sociability 64
Enhanced mPFC neuronal activity during exploration of an emotionally-altered conspecific 64
Enhancement of mPFC neuronal activity is specifically linked with emotion discrimination 67
Photo-inhibition of mPFC PV+ interneurons do not affect emotion discrimination 70
Photo-inhibition of mPFC SOM+ interneurons abolish emotion discrimination 72
Photo-stimulation of SOM+ interneurons in the mPFC guides social discrimination 76
Discussion 78
Material and Methods 81
Emotion Recognition Task (ERT) 81
One-on-one social exploration tests 83
Sociability and social novelty tests 83
Sensory modality assessment 84
Conditioned place preference 86
Corticosterone assay 86
Viral vectors 87
Stereotaxic surgery, viral injections and tetrodes implants 87
Mice underwent stereotaxic surgery for fiberoptic implantation and for recording tetrodes implants 87
Optogenetic manipulations 88
In vivo recordings 88
Analysis 88
Definition of epochs of interest 88
Single neurons analysis 89
Histology 90
Statistics 90
Chapter 4 Oxytocin effects on social behavior are genetically modulated by astrocytic dopamine D2 receptor 91

Introduction 92
Results 93
Intranasal oxytocin rescues dysbindin reduced sociability 93
Intranasal oxytocin rescues PFC excitability 95
Dysbindin mice show altered activation of astrocytes 98
Discussion 102
Materials and Methods 105
Habituation social interaction test 105
Intranasal OXT Administration 106
Stereotaxic surgery, viral injections and tetrodes implants 106
Histology 107
In vivo recordings 107
Data analysis 108
Chapter 5 GENERAL CONCLUSION 109
Strength and Weakness 111
Future directions 113
BIBLIOGRAPHY 114

Chapter 1

General introduction

General introduction

Social cognition

Nowadays is increasing the interest in social cognitive neuroscience, which is an emergent and interdisciplinary field, devoted to the research of the neurobiological processes underlying social interactions and the behavioural alterations associated to neuropsychiatric disorders. The term “social cognition” refers to the mental operations that underlie social interaction and includes the perception and integration of social cues through a complex process involving attention, memory, motivation and emotion. Two of the more investigated aspects of social cognition are emotion recognition abilities and theory of mind (ToM), two partially overlapping but distinct cognitive domains. Emotion recognition refers to an individual’s ability to identify and discriminate between the basic emotional states of others, an ability that in human is mostly based on recognition of facial or vocal expression of emotions. Theory of mind refers to intellectual abilities that enable us to perceive that others have beliefs, desires, plans, hopes, information, and intentions that may differ from our own[1,2]. Collectively these abilities guide interpersonal skills that are important for communication, social interaction and emotion perception.

From an evolutionary point of view, social behavior is one of the most important properties of animal life and it plays a critical role in biological adaptations. Likewise, social interaction is a matter of survival for humans and many other animals. Humans can be considered among all primates, the most social, and success in social interactions is one of the major forces driving human evolution[3]. This function is essentially based on exchange of signals. Speech is the most obvious signal that characterized social communication in humans; there are many other more basic signals, which humans share with other social animals. For example, by monitoring eye gaze or by watching body movements humans can infer others intentions[4]. Many animals make use of similar signals to communicate. .

A successful social interaction requires at least two overlapping steps: (1) perception and processing of social signals related to other individuals’ emotional state and intentions; (2) formulation of appropriate responses to these signals. These processes must be dynamic and flexible, since the social context is continuously changing and updated with new information.

The neural substrates of social cognitive processing are complex and largely unknown. The “social brain” network, involving a range of cortical and subcortical regions and connective

pathways[5], varies depending on task demands, but is broadly thought to include limbic regions (such as the amygdala), the prefrontal cortex and temporoparietal junction, as well as the anterior cingulate and insular cortex[6,7]. Different FMRI human studies indicate that areas as the somatosensory and temporal cortices are involved in perception and processing of faces. In particular, the somatosensory cortex (SSC) has been involved in the recognition of facial emotions[8], and the temporal cortex (TC) in the visual processing of faces[9]. The amygdala, conventionally linked to processing of fear-related stimuli or threat detection, is currently known to play an important role in social cognition[10]. Lastly, the prefrontal cortex (PFC) is known to modulate decision making and executive control, to enable the choice of the most appropriate behavioral response, by integrating sensory and emotional cues[11]. Remarkably, alterations in the functionality of these regions have been highlighted in patients with neurological or psychiatric conditions[12–14] associated to social cognitive dysfunction, indicating a functional role of these regions as possible biological substrates for further investigation.

Comparative studies have pointed out that the neurobiological and molecular mechanisms underlying social behavior are highly conserved across species. The recent development of a variety of techniques that can be applied in animal research (especially in rodents), improving our knowledge of the study of the neurobiological and molecular mechanisms that underlie social behavior.

For example, a recent study developing a new method to map whole-brain activation, at the cellular level, highlighted a number of brain structures as the olfactory bulb, the hypothalamus, the lateral septum (LS),the amygdala, the nucleus accumbens (NAcc) and the PFC[15] , activated in mice after a social behavior paradigm. This study highlights the complexity of the network in the regulation social behavior, supporting the notion that the brain structures implicated in social processing are conserved among the animal species.

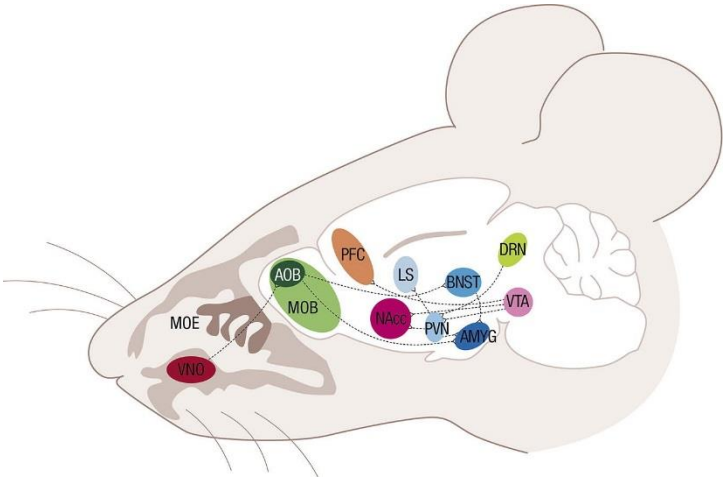


Fig1. Brain structures and neuronal circuits implicated in social cognition in rodents. In mice, olfactory signals from a social stimulus are perceived through the nasal organ area and transferred to cortical and subcortical brain areas to be processed and generate an output behavior. Figure adapted from Fernandez et al. 2018[16]

Medial Prefrontal Cortex

The medial PFC (mPFC) has long been considered the main area controlling high-order cognitive functions, such as planning, organizing, decision-making and problem solving. Specifically, the PFC plays a critical role in the ability to orchestrate behavior in accordance with internal states or intentions. Accordingly, humans with damages to frontal areas often show behavioral impairments that include inflexibility, perseveration, isolation and apathy[17] or antisocial behavior[18].

The mPFC is often referred to as a single brain region, but many subdivisions into distinct areas can be made, each defined by specific cytoarchitecture, cytochemistry, connectivity and functional properties. Defining and comparing the functional properties of these areas across species is complex: a large interspecies differences in the layering per area has in fact been described, rising the debate on whether or not rodents possess a region equivalent to the human PFC as they lack a granular zone in this area[19,20]. However, it has been noted that the formation of the general laminar pattern in the PFC shows a relation with phylogenesis; indeed, in “higher” mammalian species, such as primates and humans, PFC regions possess a granular layer IV, as well as an agranular layer. The lower is the species, the smaller is the proportion of granular PFC regions[19,20]. Thus, the concept of homologous structures with similar functions may apply. The mPFC in rodents is classified into three distinct neuroanatomical subregions based on connectivity and cytoarchitecture: the anterior cingulate (ACC), prelimbic (plPFC), and infralimbic (ilPFC) cortices[21]. These three regions have functional and connectional homology to the human Brodmann areas 24b, 32, and 25, respectively[22]. The mPFC is mainly composed by glutamatergic pyramidal neurons (~80–90%) and an array of local interneuron populations (~10–20%). The excitatory pyramidal neurons mediate output projections to other cortical areas such as sensory and association cortices, as well as to subcortical areas of the social circuitry including the striatum, the amygdala and the hypothalamus. The impact of the mPFC control can be then observed on neurotransmitter systems critical for social functioning such as dopamine (DA) and oxytocin[23]. This modulation is bidirectional, as the PFC also receives inputs from these neuromodulators through projections from subcortical structures.

Recently, by using a novel optogenetic tool to depolarize *in vivo* mPFC neurons for a long period of time, Yizhar and colleagues highlighted that altering the excitatory/inhibithory (E/I) balance, would lead to social dysfunction. By targeting both excitatory projection neurons and inhibitory parvalbumin interneurons in the mPFC of freely moving mice, they showed that elevation, but not reduction, of cellular E/I balance led to a profound impairment in social

information processing. This study indicated that a fine regulation of mPFC is required to maintain social abilities. In this thesis, further investigations will be presented in order to advance our understanding on the mechanisms that lead to profound cognitive and behavioral abnormalities.

Oxytocin system

The hypothalamus is social brain area, which has traditionally been involved in neuroendocrine functions[24]. In recent years, the paraventricular nucleus (PVN) of the hypothalamus has been found to be strongly implicated in social behavior. The PVN consists of two different neuronal populations, magnocellular and parvocellular neurons. Parvocellular neurons release mainly corticotrophin-releasing hormone, which plays a vital role in the stress response trough the hypothalamus–pituitary–adrenal axis. Magnocellular neurons synthetize large amounts of the neuropeptides oxytocin (OXT) and vasopressin[25]. These neuropeptides have long been known to play key roles in reproductive functions in mammals such as sexual behavior, parturition and maternal care[26]. More recently, animal studies have shown that they are also involved in other domains of social behavior in both males and females during non-lactating periods[27,28]. OXT is considered a “social” hormone because of the extensive literature documenting its ability to impact a variety of different social behaviors in many species, and because its production is dynamically regulated in response to specific social situations[29,29,30]. Oxytocinergic neurons in the PVN send their axons to several brain structures within the social network[31]. However, the brain mechanisms by which endogenous oxytocin produces its effects acting at specific brain sites have been not investigated.

An emerging and promising field of research is represented by the study of the behavioral effects produced by systemic or local administration of OXT. Behavioral studies in mice genetically modified that cannot produce OXT, or lack the OXT-receptor clearly indicate that the oxytocin system plays an essential role in maternal care, social cognition and affiliative behaviors[32]. These findings have led to a great deal of interest in OXT as a potential treatment for human social disorders, resulting in a large number of clinical trials to assess its therapeutic efficacy. Although the results of these trials have been encouraging in the context of autism and anxiety disorders[32], the effect of exogenous OXT are still controversial and highly variable[33–38]. Hence, the understanding of the brain circuits engaged by endogenous or exogenously supplied OXT and their role in specific behaviors remains incomplete. This topic represents one of the main aim in this thesis.

Dopamine system

Dopamine (DA) is a neurotransmitter that has been shown to play roles in various aspects of brain function, depending on the brain region involved[39]. Recent studies have demonstrated the involvement of the DA system in the regulation of cognitive and emotional functions[40] which may be partly associated with rewarding properties of social activity[41]. Dopaminergic neurons are located in midbrain cluster and recent evidences show a functional heterogeneity within these clusters, with subsets of DA neurons influencing specific behavioral outcomes[42,43]. DA signaling is processed thorough two distinct classes of receptors: D1-like (D1 and D5) are excitatory and have a low affinity for DA and D2-like (D2, D3 and D4) which are inhibitory and have a high affinity for DA, allowing them to respond to low DA concentrations present during phasic dopaminergic firing[44,45]. Dopamine D1 and D2 receptors often appear to yield the opposite effects in terms of behavioral outcomes, which are conserved across different species including rodents and non-human primates[40] and they operate via different intracellular signaling pathways[46]. Furthermore, these two classes of receptors are generally expressed on separate neurons, but despite these functional and anatomical differences in their expression, animal research has shown that both D1-R and D2-R are involved in the regulation of socially relevant behaviors, such as motivation[47], bonding[44] and aggression[48]. Indeed studies has been shown that dopamine can be important for the develop of E/I balance in the PFC, in this regards, a recent work of Petrelli et al. shows how a subset of cortical astrocyte are crucial for the regulation and homeostasis of DA during the postnatal development of the mPFC, allowing for optimal DA-mediated maturation of excitatory circuits.

Reduced prefrontal dopaminergic neurotransmission has been also shown to contribute to the negative and cognitive aspects observed in schizophrenia patients[49,50]. However, over the years new studies have reported a limited direct evidence of impairment within the dopaminergic system itself so the recent hypothesis has been reformulated in dopamine imbalance between different brain regions[51]. In particular impaired afferent circuits onto ventral tegmental area (VTA), where are localized dopaminergic neurons, lead to a dysregulated dopamine release in other brain region such as prefrontal cortex[51]. Nevertheless, this is still an open question whether different neuronal populations within the same structure contribute to the different behaviors, or whether the same neuronal population with different activity pattern dictates the behavioral response. Studies in animal models allowing the dissection of specific cell type contributions, which are crucial to address these questions.

Dysbindin mice

Neurodevelopmental disorders present a complex etiology involving susceptibility genes and environmental factors. Although multiple genes have been found to be associated with neurological or psychiatric disorders and the actual function or involvement of individual genes in the developmental aspects of brain structures formation are largely unknown. One of these candidate gene is dystrobrevin-binding protein 1 gene (DTNBP1) which encode for dysbindin-1 protein, which is widely expressed in human and mouse brain[52]. Dysbindin-1 has been implicated in the regulation of vesicle formation and synaptic release and it is a component of the biogenesis of lysosome-related organelles complex (BLOC-1) [52,53]. An animal model of dysbindin-1 functions is available in the sandy (sdy) mouse, which has a naturally occurring deletion mutation of exons 6 and 7 in the gene (DTNBP1) encoding for the mouse protein, resulting in loss of dysbindin-1. Recent studies in these mice show that dysbindin-1 is highly expressed during embryonic and early postnatal development compared to adulthood and BLOC-1 is involved in neurite outgrowth[54]. These findings highlight a potential role for dysbindin-1 in the normal development of brain structure and function. Indeed different studies, using cortical neuronal cultures taken from adult dysbindin-1 KO mice have been associated with alteration of various aspects of synaptic function [53,55], and the regulation of both dopamine and glutamate signaling in the brain[56]. In particular, in vitro experiments suggest that dysbindin-1 suppresses DA release[57] and cultured neurons have increased cell surface expression of the dopamine D2 receptor due to an increased membrane insertion rate[58]. In addition, layer II/III pyramidal neurons from the PFC of dysbindin-1 KO mice show increased activity at baseline, but decreased activity after D2 stimulation compared to wild-type mice and these effects may be due to D2-mediated alterations in the excitability of fast-spiking GABAergic interneurons[58,59]. In line with this results, dysbindin-1 KO mice display deficits in prepulse inhibition, social interaction, and diverse aspects of spatial memory [55,59]. In accordance with animals studies, patients with schizophrenia have lower expression levels of dysbindin-1 mRNA and protein in the prefrontal cortex and hippocampus[60–63] and they show cognitive disabilities[64,65]. Therefore, these data implicate reduced dysbindin-1 function in behavioral and neurobiological effects that are thought to play a role in the development of cognitive abnormalities. Overall, the dysbindin mouse is useful genetic model for the study of the cognitive domains.

Scope of the Thesis

Abnormalities in social functions represent a key feature of a number of neuropsychiatric disorders. In particular, alteration in social cues identification define autism spectrum disorders[66]. Similarly, patients with schizophrenia have marked impairments in processing non-verbal social affective information with poor performance in the emotion recognition test[1]. Relevantly, effective solutions for the treatment of deficits in non-verbal communication are still missing. This is also due to a poor understanding of the neurobiology (including brain circuits and related mechanisms) of social information processing. In this context, studies in mouse models, allowing the dissection of specific social domains and the investigation of the biological substrate can be very informative. Interestingly, despite the large array of behavioral tests available to assess mice social functions, such as social recognition, sexual maternal and aggressive behavior, there is no test available to assess facets of emotion discrimination in mice similar to the emotion recognition test used in humans. This was the main reason underlie the initial part of my my work (**chapther 2** and **3**) to set up a novel behavioral test aimed to investigate whether mice are able to discriminate unfamiliar conspecifics based on altered emotional states.

In particular in **chapter 2 "Oxytocin-dependent emotion recognition in mice"**, I will present a first validation of this novel paradigm, which allowed to reveal the implication of different endogenous OXT pathways in the expression of mice discriminatory behaviors towards unfamiliar conspecifics in positive or negative emotional states. In particular, manipulating OXT-ergic projections from PVN to different targets using the DREADD technology (designer receptors exclusively activated by designer drugs), we found that the release of OXT from the PVN to CeA is a necessary substrate for emotion recognition ability in mice. This finding was also supported by the evidence that rescuing CeA OXTR reduced levels in a mouse model of genetic liability, dysbindin-1, we were able to rescue their emotion recognition deficits. Our study provides and support the validity of this novel paradigm to explore social cognitive processes not previously investigated in mice, supporting more translational approaches between rodent and human social cognitive functions, for the investigation of circuits, genetics and neurochemical systems alteration involved in different psychiatric disorders.

My specific contribution to this work was the study of the sensory modalities used by mice to communicate emotions. In particular, to investigate which are the signals that mice perceive when evaluating emotional states in conspecifics. Investigation of visual, olfactory and auditory

cues (USVs) lead to the conclusion that olfactory and visual cues are both relevant in emotion communication.

In **chapter 3 "Somatostatin interneurons within the prefrontal cortex control emotion recognition in mice"** we studied the role of PFC, in particular the E/I balance of mPFC region during the emotion recognition test. We investigated the coding of emotionally-relevant information using a combination of approaches. We first assessed mPFC neuronal activity in freely moving animals with silicon probe implanted in the this brain region during the performance of the newly developed emotion recognition task for mice. We revealed that neurons in the mPFC are differentially activated during exploration of conspecifics depending on their affective state. To investigate the contribution of different PFC cell types in emotion recognition, we selectively manipulate the activity of somatostatin (SOM+) and parvalbumin (PV+) interneurons, by optogenetic technique. These results indicate that selective inhibition of mPFC somatostatin (SOM+) interneurons, but not parvalbumin (PV+), abolishes emotion discrimination. Conversely, activation of mPFC SOM+ interneurons is sufficient to induce emotion discrimination.

My contribution to this work was the recording and analysis of the mPFC electrophysiology data, which paved the way for the optogenetic investigation of the role of specific cell types in the recognition of emotions in conspecifics. In addition, I have contributed to different aspects of the setting up of the emotion recognition test.

The last work presented in **chapter 4 "Oxytocin effects on social behavior are genetically modulated by astrocytic dopamine D2 receptor"** was centered on the investigation of the mPFC modulation during freely social interaction. In this study, we used dysbindin-1 mutant mice in the social habituation paradigm, to investigate of intranasal OXT effect as potential treatment for social disabilities. In particular, we found in dysbindin-1 mice an alteration on the behavior and electrophysiology readouts in the mPFC, suggesting a valuable model to investigate the effect produced by OXT treatment. We revealed no alteration after the treatment in wild type and conversely a beneficial effect in dysbindin-1 mice, implying different mechanism might be involved in the two genetic background. Indeed, the gene expression profiles in the mPFC of wild type and dysbindin-1 mice show an up-regulation of different genes involved in astrogliosis. Overall, This study, linking genetic liability and astrocytes function, provides a new framework at the base of oxytocin behavioral effect.

My contribution to this work was to design, perform and analyze all different parts of behavioral, pharmacological*, in vivo* electrophysiological and astrocyte-related experiments.

Overall, the work presented in this thesis aims to uncover how different brain structures and systems contribute to the perception and coding of social stimuli. In **chapter 5,** I summarized how this research provide new contributions in the direction of better addressing the complexity of mice social behavior. I will also discuss potential applications of the presented findings in a clinical perspective and in relationship to specific genetic alterations and potential future direction to follow.

Chapter 2

Oxytocin-Dependent Emotion Recognition in Mice

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

(Submitted)

Abstract

Recognize and discriminate others' emotions is a fundamental social cognitive ability that influences development, survival and evolution of animals. The oxytocin (OXT) system has been linked to human emotion recognition, but mostly through the effects of exogenous OXT. Indeed, the implication of endogenous OXT pathways in emotion recognition remains elusive. By developing a new assay to measure the ability of mice to discriminate unfamiliar conspecifics based on negative- or positive-valence emotional states, here we revealed that endogenous OXT projections from the Hypothalamic Paraventricular Nucleus (PVN) to the Central Amygdala (CeA) are necessary for emotion recognition processing. OXT release to the Nucleus Accumbens, Prefrontal Cortex, and hippocampal CA2 was instead dispensable. Notably, mice emotion recognition was distinct from sociability or emotional contagion processes. Furthermore, AAV-mediated potentiation of CeA OXT signaling in a mouse model of cognitive liability was sufficient to rescue emotion recognition deficits. These findings demonstrate a central role of CeA OXT signaling in emotion recognition, and support the validity of this novel paradigm to explore social cognitive processes not previously investigated in mice.

Introduction

Social interactions depend on the ability to distinguish various expressions of emotions in others. This biologically innate process defined as “social cognition” has profound implications in everyone’s life[67,68]. Consistently, disturbances in social cognition are early and distinctive features of many neuropsychiatric, neurodevelopmental and neurodegenerative disorders[68]. Abnormalities in social cues identification define autism spectrum disorders[66]. Similarly, patients with schizophrenia have marked impairments in processing non-verbal social affective information while showing normal affect sharing and emotional experience[1]. Notably, social cognitive impairments in these individuals are not yet effectively cured, despite they have a more deleterious impact on daily functioning than non-social cognitive deficits[69].

The Oxytocin (OXT) system has been indicated as a major player in social information processing and social cognition, with implications in humans’ emotion recognition, empathy and trust, and as a promising target for the treatment of psychiatric disorders characterized by social and emotion processing dysfunctions[70–79]. Remarkably, our knowledge on the involvement of OXT system in emotion processing derives from genetic-association studies[69,80,81], or from the use of exogenous OXT, whose effects are still controversial and highly variable[33–38]. In particular, questions are still open on if/how much exogenous OXT arrive into the brain[82], if there are preferential effects on specific social functions[70], if there are differences in effects between healthy and patient groups[83,84], the differential effects of acute versus chronic exposure[85,86], and the modulation of its effects based on genetic background[70,80,81,81]. In this context, a better knowledge about endogenous OXT signaling pathways involved in emotion recognition processes, and their specific effects in relation to genetic backgrounds would be crucial.

To investigate the implication of endogenous OXT pathways in complex social cognitive processes, here we developed a paradigm to study emotion discrimination abilities in mice. Emotion recognition tasks are, indeed, the most-extensively used paradigms to assess human social cognition[1,68]. Moreover, several training programs targeting facial emotion perception have been implemented for individuals with schizophrenia and autism[87,88]. Previous studies indicate that rodents are sensitive to the affective states of others. This is demonstrated by the existence of so called social contagion [44,89,90], an automatic response evident when a rodent witness a familiar conspecific under pain or physical challenge[91], and by some sort of helping behavior[92] or consolatory behavior versus a previously stressed familiar conspecific[93]. Our new paradigm expand these previous tools, reliably addressing mice’ ability to discriminate

negative and positive emotions evoked in unfamiliar conspecifics, in a distinct way to emotional contagion or basic sociability.

By chemogenetic inhibition of OXT release from specific PVN projections to several brain target areas, we demonstrated the importance of OXT signaling to the CeA in recognizing emotions in others. This finding was further supported by the observation that rescuing decreased CeA OXT receptor (OXTR) levels in mutant mice with emotion recognition deficits was sufficient to restore this ability. Altogether, our results reveal the essential role of specific endogenous OXT pathways in the ability to discriminate emotions in others, and that genetic background and variation in OXTR within the CeA moderates these effects. This novel behavioral paradigm can then be used to explore social cognitive processes not previously investigated in mice.

Results

Mice are able to discriminate unfamiliar conspecifics based on negative-valence emotional states

To test whether mice could discriminate unfamiliar conspecifics in different emotional states, we placed an "observer" mouse in a cage containing two age- and sex-matched unfamiliar conspecifics ("demonstrators") in wire cups allowing visual, tactile, auditory and olfactory communication (Fig. 1a). The behavior of the observer mouse, concomitantly exposed to a neutral demonstrator and to a mouse in an altered emotional state, was then analyzed. Similar to emotion recognition tests performed in humans[68], this setting is focused on the social approach initiated by the observer mouse, avoiding potential confounders resulting from aggressive or sexual interactions.

In the "fear" condition, one of the two demonstrators was fear-conditioned to a tone cue at least one day before the test (Fig. 1b and online Methods). Thus, upon successive tone presentation, a negative-valence emotional state will be induced[94]. In order to measure the impact of the elicited emotion on observers' response before, during and after its induction, the tone was delivered during the second 2-minute trial of the task (Fig. 1b and online Methods). Consistently, we observed in the fear demonstrator a freezing response only during the 2-minutes tone presentation, associated with a reduction in rearing (Fig. S1a). No other behavioral parameters differed between the two demonstrators during the 6-minute test session (Fig. S1a). Thus, this experimental design allowed to tightly link observers' behavior with the emotional state evoked in one of the demonstrators.

We found that both male and female observers displayed increased social exploration, indicated by sniffing, towards the fear demonstrator compared to the neutral one (Fig. 1c-d and Fig. S2a). In particular, this effect was specific for the last 2-minute trial, and did not correlate with the freezing behavior shown by the fear demonstrator during tone presentation (Fig. 1e). A significant inverse correlation was instead evident during the 2-minute tone presentation (Fig. 1f) suggesting that demonstrator's freezing inhibited observer approach in trial 2, but did not influence the discriminatory behavior in trial 3.

In light of previous evidence[89,95–97], 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). Observers showed no freezing behavior, escape attempts or other stress-related behaviors, nor changes in rearing and grooming (Fig. 1g-h), and a normal decrease in locomotor activity (Fig. 1i) throughout the whole test. Moreover, the corticosterone levels of observer mice exposed to the fear paradigm or to two neutral demonstrators did not differ (Fig.1j). Notably, observers showed no discriminatory behaviors if exposed to two neutral demonstrators (Fig. S3). These findings suggest an ability of observer mice to discriminate unfamiliar conspecifics based on a negative emotional state, which was not associated to emotional contagion.

Figure 1

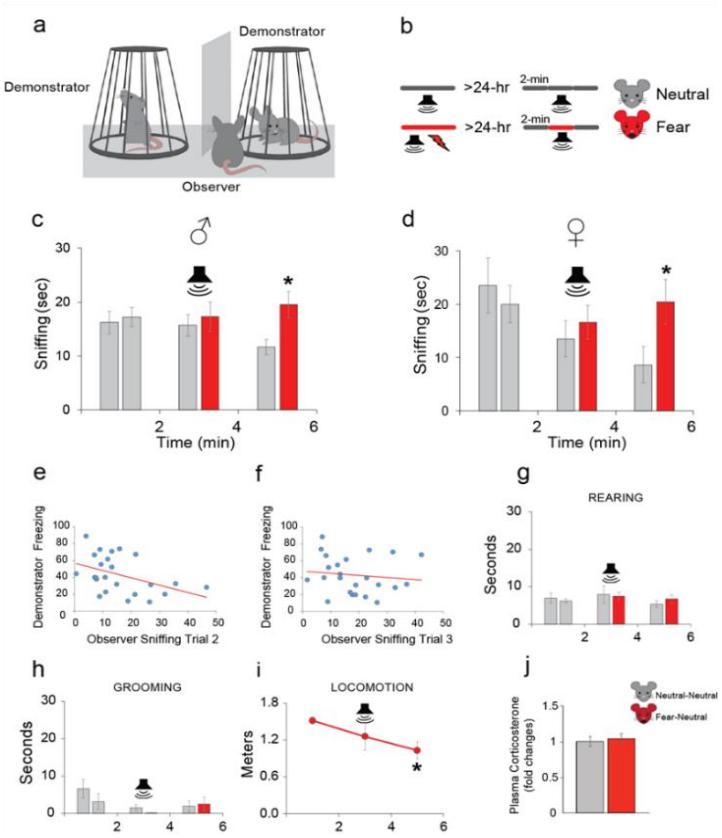


Figure 1. Mouse emotion recognition task for negative emotions. (**a**) Schematic drawing of the task setting. (**b**) Timeline of pre-test and test procedures to trigger in one of the demonstrator a "fear" emotional state by delivering the conditioned tone between the 2-4 minutes period of the testing phase. (**c-d**) Time (in seconds) spent sniffing demonstrators in neutral (grey bars) or tone-induced fearful (red bars) state displayed by (**c**) male and (**d**) female observer mice during the 6 minutes of the emotion recognition test, divided into three consecutive 2-minute time beans (last 2-min 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 for the 0-2 and 2-4 minutes test periods). *p<0.05 versus the exploration of the neutral demonstrator. N=8/15 observers per group. (**e-f**) Correlation analyses between the time spent freezing by the fear-conditioned demonstrator (in y axis) and time spent by the observer mouse sniffing the fear-conditioned demonstrator (in x axis) (e) in the time 2-4 minutes or (**f**) in the time 4-6 minutes of the emotion recognition test (Correlation for time 2-4 minutes r=-0.4310 and time 4-6 minutes r=-0.11). *p<0.005. N=24 observers. (**g-h**) Time (in seconds) spent in (**g**) rearing and (**h**) grooming close to demonstrators in neutral (grey bars) or tone-induced fearful states (red bars) displayed by the same observer mice during the 6 minutes of the emotion recognition test, divided into three consecutive 2-minute time beans (RM ANOVAs showed no significant differences). (**i**) Locomotor activity displayed by the same observer mice during the 6 minutes of the emotion recognition test, divided into three consecutive 2-minute time beans (RM ANOVA $F_{2,16}$=4.08, p=0.03.. *p<0.05 versus minutes 0-2. N=9 observers. (**j**) Blood corticosterone levels displayed by observer mice immediately after being tested in the emotion recognition task with two neutral demonstrators (grey bar) and 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.

Figure S1

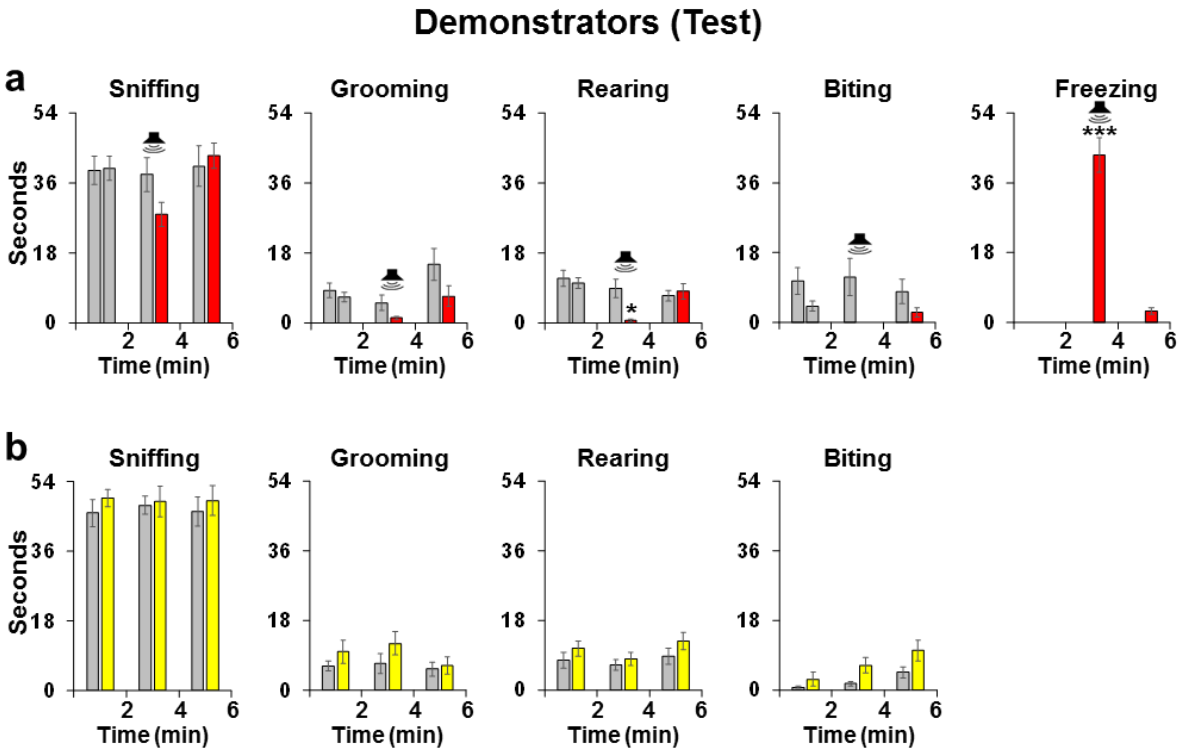


Figure S1. Demonstrators behavior during test. (**a**) Observable behaviors displayed by the neutral and fear demonstrator mice during the 6 minutes of the fear emotion recognition paradigm, divided into three consecutive 2-minute time beans. No demonstrator defecated or urinated during the whole test 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**) Observable behaviors displayed by the neutral and relief demonstrator mice during the 6 minutes of the relief emotion recognition paradigm, divided by three consecutive 2-minute time bean. No demonstrator defecated or urinated during the whole test session. No significant emotion-by-time statistical interaction was evident for sniffing (F2,36=0.09, p=0.92), grooming (F2,36=0.34, p=0.71), rearing (F2,36=0.31, p=0.73), and biting (F2,36=0.84, p=0.44. N=10 demonstrators per group.

Figure S2

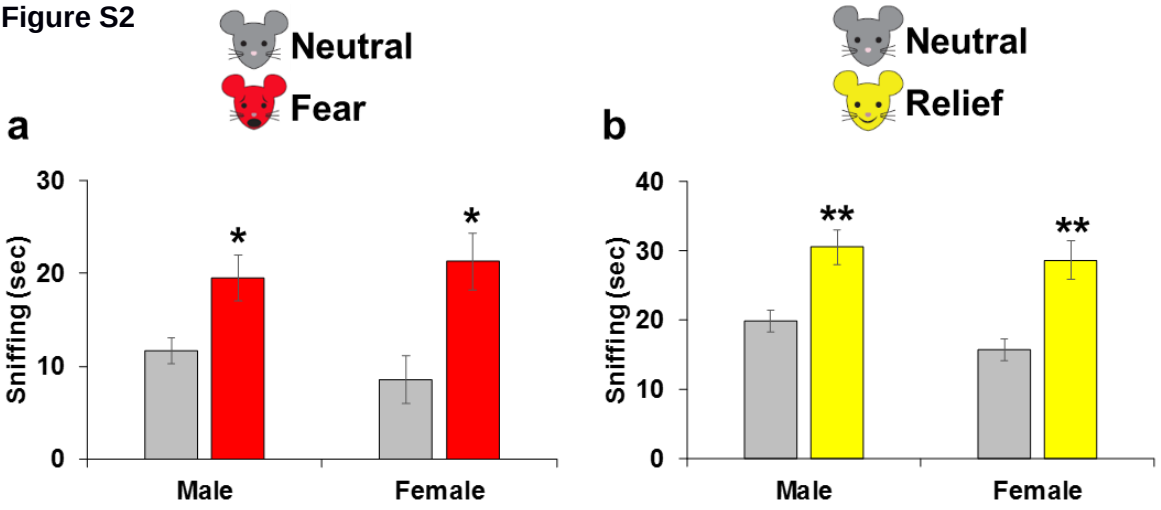


Figure S2. Equal emotion recognition abilities in male and female mice. Time spent sniffing showed by male and female observers towards demonstrators with (**a**) tone-induced fearful states (last 2-min RM ANOVA Male: $F_{1,15}$=6,51, p=0.022; Female: $F_{1,11}$=10,98, p=0.006), or (**b**) water-induced relief states (first 2-min RM ANOVA Male: $F_{1,14}$=15,07, p=0.001; Female: $F_{1,14}$=14,60, p=0.001), represented by red or yellow bars, respectively. *p<0.05 and **p<0.005 versus the neutral demonstrator. N=12-16 observers per group.

Figure S3

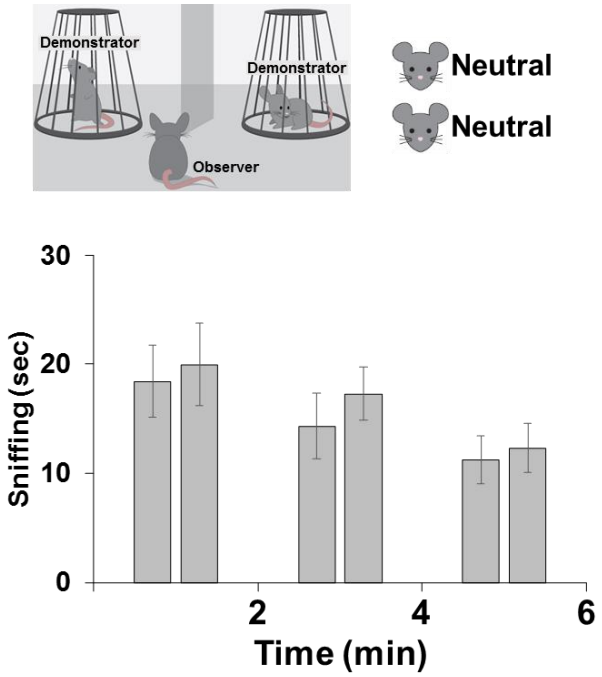


Figure S3. No discrimination towards two neutral demonstrators. Schematic drawing of the task setting, and time (in seconds) spent sniffing two demonstrators in neutral (grey bars) states displayed by observer mice during the 6 minutes of the emotion recognition test, divided into three consecutive 2-minute time beans. RM ANOVAs reveled no significant differences for the 0-2, 2-4, and 4-6 minutes test period.

Mice are able to discriminate unfamiliar conspecifics based on positive-valence emotional states

Emotion recognition paradigms to assess human social cognition include the presentation of positive-valence emotions. Thus, we investigated whether discriminatory behaviors in the observers could be detected towards positive-valence states. In particular, we exposed observer mice to a neutral demonstrator and to a demonstrator that received a 1-hour ad libitum access to water after 23-hours of water deprivation (Fig. 2a-b). Water was selected as a rewarding factor to avoid odor-related cues that could differentiate the two demonstrators. We assumed that the relief from the distressful water deprivation would result in a positive-valence emotional state (“relief”). Indeed, we found that the 1-hour ad libitum access to water resulted in a conditional place preference (Fig. 2e) in mice that experienced the 23-hours water deprivation, but not in mice in ad libitum water condition (Fig. 2f). Consistently, 1-hour ad libitum access to water after the 23-hour deprivation, reduced corticosterone levels (Fig.2g). During the emotion discrimination test, this manipulation did not induce any detectable behavioral difference between relief and neutral demonstrators (Fig. S1b).

Observers of both sexes showed increased social exploration towards the relief demonstrator compared to the neutral, selectively in the first two minutes of the task (Fig. 2c-d and Fig. S2b). No changes in rearing and grooming patterns towards the demonstrators and throughout the tasks were evident (Fig. 2h-i). Moreover, observers showed the typical decrease in locomotor activity (Fig. 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 (Fig.S4a). These data indicate that mice are able to discriminate unfamiliar conspecifics based on positive valence emotions.

Figure 2

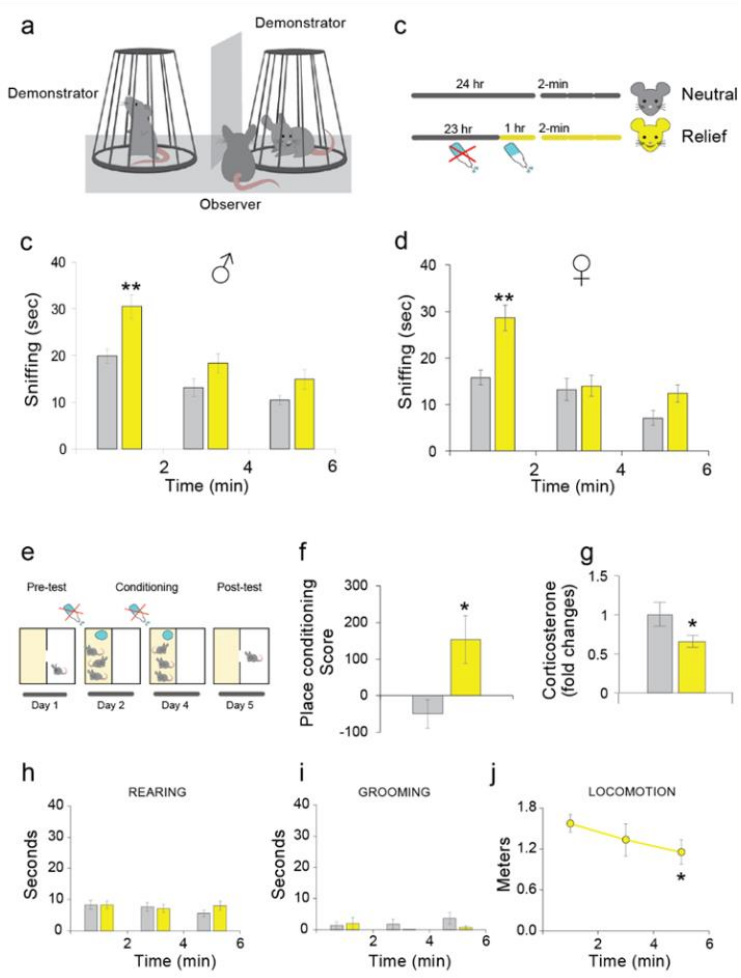


Figure 2. Mouse emotion recognition task for positive emotions. (**a**) Schematic drawing of the task setting. (**b**) Timeline of pre-test and test procedures to trigger in one of the demonstrator a "relief" emotional states during the testing phase. (**c-d**) Time (in seconds) spent sniffing demonstrators in neutral (grey bars) or water-induced relief (yellow bars) state displayed by (**c**) male and (**d**) female observer mice during the 6 minutes of the emotion recognition test, divided by three consecutive 2-minute time beans (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 2-4 and 4-6 minutes test periods).

p<0.005 versus the exploration of the neutral demonstrator. N=15 observers per group. (e**) Timeline of the Place Conditioning procedures used to assess if 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 (grey bar) or relief (yellow bar) emotional state. For each mouse, a place conditioning score was calculated as the post- minus the preconditioning time spent in the conditioning-paired compartment of the apparatus. A positive score indicates place preference, a negative score a place aversion, 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-hour water deprivation (grey bar) or after a period of 1-hour *ad libitum* access to water following 23-hour water deprivation (yellow bar). (T-test: df: 19; p=0.05). *p=0.05 versus water deprived mice. N=11 mice per group. (**h-i**) Time (in seconds) spent in (**h**) rearing and (**i**) grooming close to demonstrators in neutral (grey bars) or relief (yellow bars) state displayed by the same observer mice during the 6 minutes of the emotion recognition test, divided by three consecutive 2-minute time beans (RM ANOVAs showed no significant differences). (**j**) Locomotor activity displayed by the same observer mice during the 6 minutes of the emotion recognition test, divided by three consecutive 2-minute time beans. (RM ANOVA $F_{2,18}$=4.35, p=0.04). *p<0.05 versus minutes 0-2. N=10 observers.

Figure S4

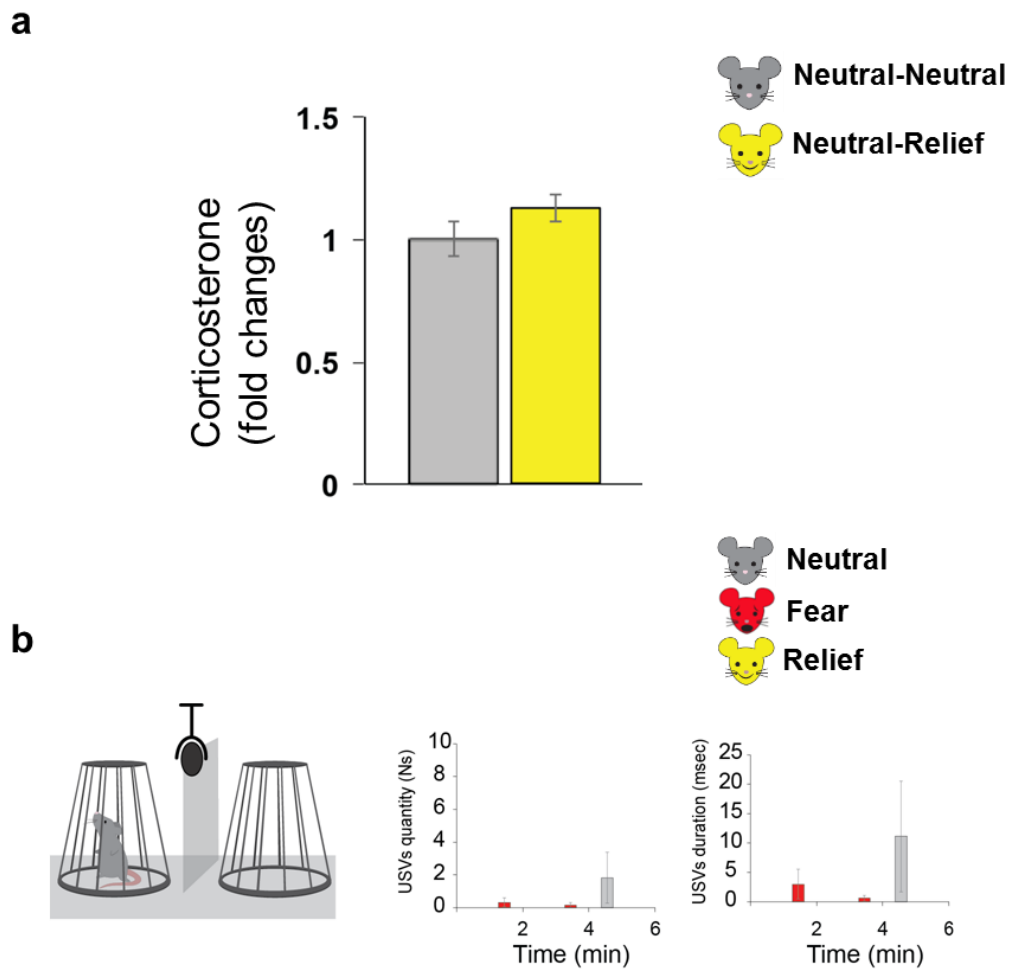


Figure S4. Corticosterone levels in observer mice (**a**) Blood corticosterone levels displayed by observer mice immediately after being tested in the emotional recognition task with two neutral demonstrators (grey bar), and one neutral and one relief demonstrators (yellow bar). Data are expressed as fold changes compared to observers exposed to two neutral demonstrators (T-test: df: 9; p=0.18). N=5/6 observers per group.
Ultrasonic vocalizations in single demonstrators. (**b**) Schematic drawing of the test setting to record USVs, mean number of USV calls per minute, and mean duration of USVs in milliseconds emitted by a single demonstrator mouse in neutral (grey bar), fear (red bar), or relief (yellow bar) emotional state. Two-Way ANOVAs showed no significant differences. N=6 demonstrators per group.

Impact of sensory cues in the mouse emotion recognition task

Human emotion discrimination paradigms mostly rely on visual detection of facial and body expressions. Rodents' visual system is developed enough to acquire information as evident from observational transfer of fear/pain paradigms[95,96]. However, knowledge about how emotions are expressed and perceived in mice is still poor. Thus, we investigated the impact of different sensory modalities (visual, auditory and olfactory) in the ability of mice to discriminate emotions in unfamiliar conspecifics.

To evaluate the use of visual cues we performed the test in complete darkness (Fig. 3a). We found that mice were still able to recognize altered emotional states in their conspecifics (Fig. 3b-c), similarly to what we observed in standard lighting conditions (Fig. 1-2). Removal of visual cues however, anticipated fear emotion discrimination to the tone trial (Fig. 3b) and tended to extended relief discrimination for the whole session (Fig. 3c). These data suggest that visual cues are dispensable to recognize emotions, but they can have a potential inhibitory effect on social approach and exploration.

To investigate the role of auditory information, we recorded ultrasonic vocalizations (USVs) during both fear and relief conditions (Fig 3d). We found very few vocalizations and no differences between neutral, fear or relief conditions (Fig. 3e-f and S4b). In agreement with previous evidence41, our data indicate that USVs in mice do not communicate altered emotional states, and that auditory information are not necessary for emotion discrimination.

Finally, we tested the impact of olfactory cues on observers performance, using a modified version of our experimental setting in which the observer was presented with cotton balls soaked with odors obtained respectively from neutral, fear or relief demonstrators (Fig. 3g). In contrast to emotion discrimination results (Fig. 1) and consistent with previous evidence [98–100], we found that observers avoided the odor from a fear demonstrator (Fig 3h). Instead, observers spent more time sniffing the relief odor compared to the neutral (Fig.3i), as similarly found in the relief emotion recognition (Fig. 2). These results suggest that olfactory cues convey information related to emotional states, but the discriminatory behavior that they trigger in observer mice are qualitatively different in the presence of the demonstrators. Overall, this set of data indicates distinct implications of both visual and olfactory social cues in the expression of mouse emotion discrimination.

Figure 3

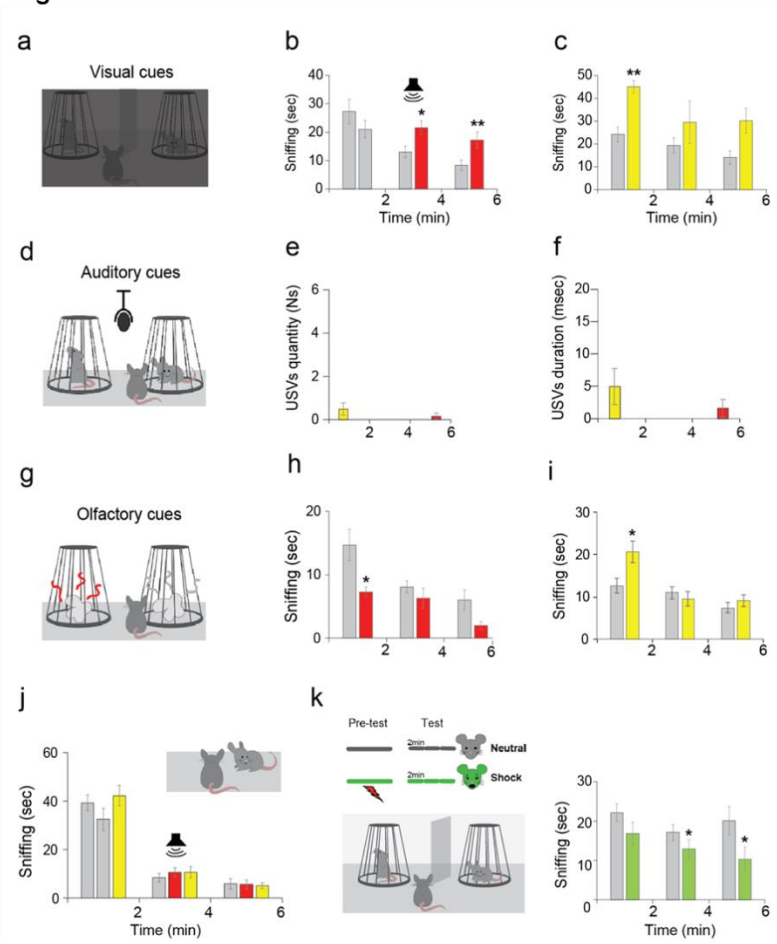


Figure 3. Sensory modalities implicated in the mouse emotion recognition and distinction to sociability. (**a**) Schematic drawing of the test setting performed in complete darkness. (**b-c**) Time (in seconds) spent by observer mice sniffing demonstrators during the 6 minutes of the (**b**) negative and (**c**) positive valence versions of the emotion recognition test. Time spent sniffing neutral demonstrators are depicted in grey. Time spent sniffing (**b**) fear or (**c**) relief demonstrator are depicted in red and yellow, respectively (RM ANOVAs for the fear manipulation, 2-4 minutes: $F_{1,8}$=5.63, p=0.04; 4-6 minutes: $F_{1,8}$=28.08, p=0.0007; for the relief manipulation, 0-2 minutes: $F_{1,5}$=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**) Mean number of USV calls per minute and (**f**) mean duration of USVs in milliseconds emitted by mice during the fear and relief emotion recognition tasks (Two-Way ANOVAs showed no significant differences). N=6 observers per group. (**g**) Schematic drawing of the test setting performed only with demonstrators' odors for fear and relief conditions. (**h-i**) Time (in seconds) spent by observer mice sniffing the odors from neutral (grey), fear (red), or relief (yellow) demonstrators during the 6 minutes of the (**h**) negative and (**i**) positive valence versions of the emotion recognition test (RM ANOVA for the fear manipulation, 0-2 minutes: $F_{1,6}$=9.15, p=0.02. No significant differences for the 2-4 and 4-6 minutes test periods. RM ANOVA for the relief manipulation, 0-2 minutes: $F_{1,20}$=4.25, p=0.052. Similarly, no significant differences for the 2-4 and 4-6 minutes test periods). *p<0.05 versus the exploration of the neutral odor. N=7/21 observers per group. (**j**) Schematic drawing of the one-on-one test setting and time (in seconds) spent by observer mice sniffing a single demonstrator in a neutral (grey), fear (red) or relief (yellow) state during a 6-minute free interaction test. The tone for which only the fear demonstrator was fear-conditioned was delivered between 2-4 minutes of the test (ANOVAs revealed only a time effect with normal decreased exploration throughout the 6 minutes, $F_{2,56}$=132.01, p<0.0001). N=12 observers. (**k**) 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. (**l**) Time (in seconds) spent by the observer mice sniffing demonstrators in neutral (grey bars) or shocked emotional state (green bars) during the 6 minutes of the emotion recognition test, divided into three consecutive 2-minute time beans (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.

Emotion discrimination abilities are a stable trait, not relatable to sociability

Human emotion recognition tasks measure the ability to label or discriminate emotions in others, which is distinct from sociability (defined as time spent interacting with others). We evaluated the sociability of an observer mouse towards a neutral, a fear, or a relief conspecific in a one-on-one setting and found that the levels of social exploration did not differ among conditions and showed the classic decrease over time (Fig. 3j). This indicates that mice ability to discriminate emotions is specifically highlighted when simultaneously exposing observer mice to two demonstrators in two different emotional states. Previous evidence measuring affective responses of an observer rat exposed to a demonstrator immediately after shock, showed equivalent social interaction data in one-on-one and one-on-two settings[40]. The discrepancy with our result might be due to the scalability feature of emotions[45] and not to species-specific differences. Indeed, exposing in our setting the observer to a demonstrator immediately after shock and a neutral one (Fig. 3k and40), generated a general aversion during the whole test session (Fig. 3k). Taken together, these data suggest that the emotion recognition setting (Fig. 1-2) is able to highlight specific behavioral responses to mildly graded expression of emotions, possibly underestimated by social interaction tests.

Human emotion recognition paradigms present strong test-retest reliability[46], which is a critical feature for longitudinal, drug response, psychobiological and clinical trial studies. In agreement, the ability to distinguish unfamiliar conspecifics based on emotional states remained unchanged when the same observer mouse was re-exposed to the same paradigm (Fig. S5c-e) or even if the same mouse was tested in the two different paradigms (Fig. S5b-d). This indicates that emotion discrimination in mice is a stable trait, and that this paradigm is well-suited to be used for Latin square design in mechanistic manipulations.

Figure S5

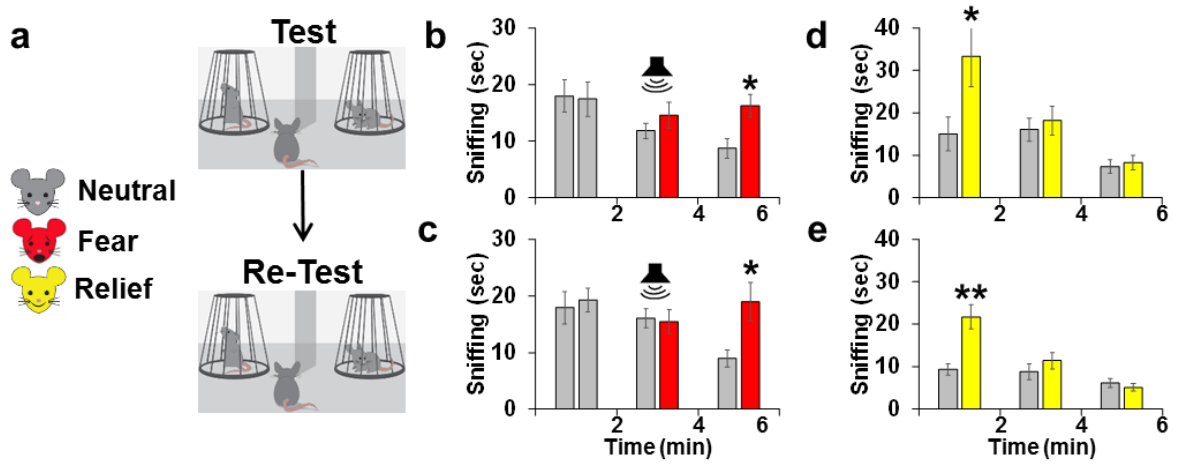


Figure S5. Reliability of the mouse emotion recognition task. (a) Schematic drawing of the test-retest reliability validation. (b, c, d, e) Time (in seconds) spent by observer mice sniffing the two demonstrators in the fear (b, c) and relief (d, e) paradigms during their (b, d) first and (c, e) second exposure to the 6-minute emotion recognition test. Time spent sniffing neutral demonstrators are depicted in grey, fear in red and relief in yellow. 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.

Endogenous release of oxytocin is necessary for emotion discrimination

We then asked whether this newly identified social cognitive ability in mice might be mediated by conserved neurobiological mechanisms consistent with those implicated in humans. The oxytocin (OXT) system plays a pivotal role in social perception and cognition[70,71,76]. In particular, in humans, OXT has been associated with social cognitive functions such as emotion recognition, empathy and trust[35,70,71,73,74,76].

To assess whether the central release of OXT is necessary to distinguish emotional states of others, we prevented the release of endogenous OXT from paraventricular nucleus of the hypothalamus (PVN) neurons by bilateral injections of a recombinant adeno-associated virus (rAAV) expressing the inhibitory hM4D(Gi) DREADD receptor under the control of the OXT promoter (Fig. 4a). In rodents, neurons in the PVN are the main source of OXT projections to brain regions[31,101]. We found that, in contrast to treatment with vehicle (Veh), inhibition of PVN OXT-projecting neurons (upon CNO injection) abolished the ability of mice to discriminate either fear or relief emotions in conspecifics (Fig. 4b-c and Fig. S6). This was equally evident in male (Fig. 4b-c) and female mice (Fig. S6). CNO treatment per se did not affect the ability of mice to discriminate between different emotional states in conspecifics (Fig. S7). Furthermore, inhibiting PVN OXT-projections produced selective effects in emotion

recognition, as no differences were induced by CNO on social exploration when the same mice were tested in a one-on-one free-interaction setting with an unfamiliar conspecific (Fig. S8a). Overall, these data indicate a direct involvement of PVN OXT release in the ability to discriminate different emotional states in conspecifics.

Figure 4

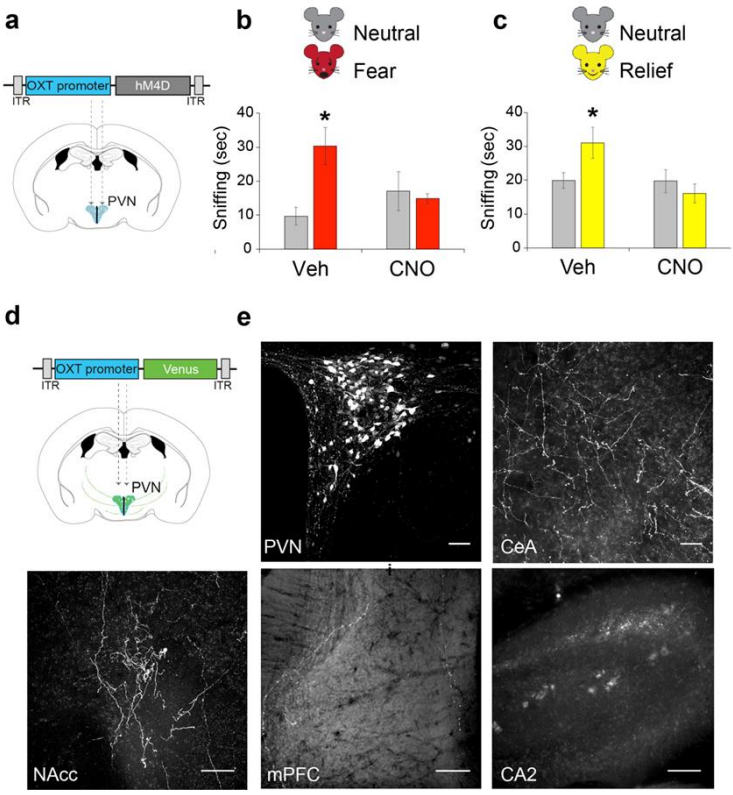


Fig4. PVN OXT release in emotion recognition and the anatomy of PVN OXT projections in mice. (A) Scheme of the viral vector used to infect the PVN OXT neurons. (B-C) Time (in seconds) spent sniffing each wire cage contatining two demonstrator mice during the first 2 minutes of the emotion recognition test displayed by the same observer mice treated with vehicle or CNO (i.p., 30 minutes before the test), and shown separately for each demonstrators emotion. Time spent sniffing neutral demonstrators are depicted in grey and represented by grey bars. Time spent sniffing demonstrators with (B) water-induced relief states (first 2-min RM ANOVA veh: $F_{(1,7)}$=7,24, p=0.031; CNO: $F_{(1,7)}$=0,50, p=0.50) , or (C) tone-induced fearful states (last 2-min RM ANOVA veh: $F_{(1,3)}$=21,76, p=0.018; CNO: $F_{(1,3)}$=0,10, p=0.77) are represented by yellow or red bars, respectively. *P<0.05 versus the neutral demonstrator within the same observer treatment. N=10 observers per group. (D) Scheme of the viral vector used to infect the PVN OXT neurons. (E) Anatomical OXT projections from PVN to central amygdala (CeA), hippocampus CA2 (CA2), and nucleus accumbens (NAcc). Scale bar: 100micronM.

Figure S6

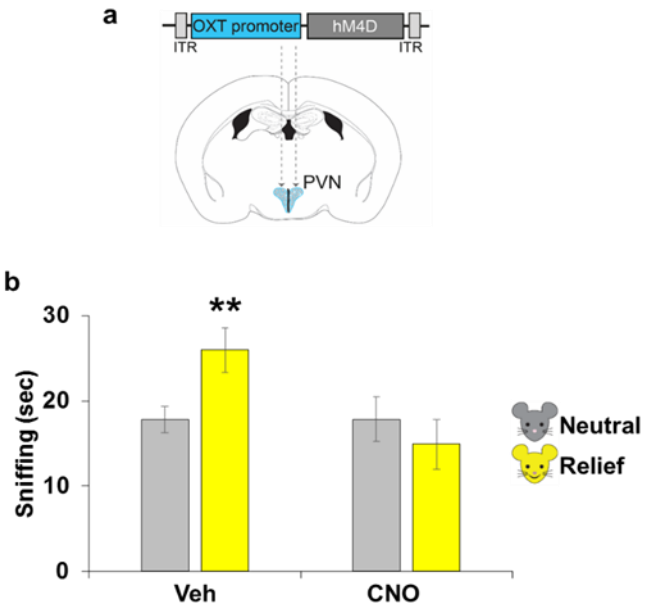


Figure S6. Inhibition of PVN OXT projecting neurons abolished emotion discriminate also in female mice. (a) Scheme of the viral vector used to infect the PVN OXT neurons. (b) Time (in seconds) spent sniffing each wire cage containing two demonstrator mice during the first 2 minutes of the emotion recognition test displayed by the same observer mice treated with vehicle or CNO (3 mg/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 are depicted in grey and 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) are represented by yellow bars. **P<0.005 versus the neutral demonstrator within the same observer treatment. N=10 observers per group.

Figure S7

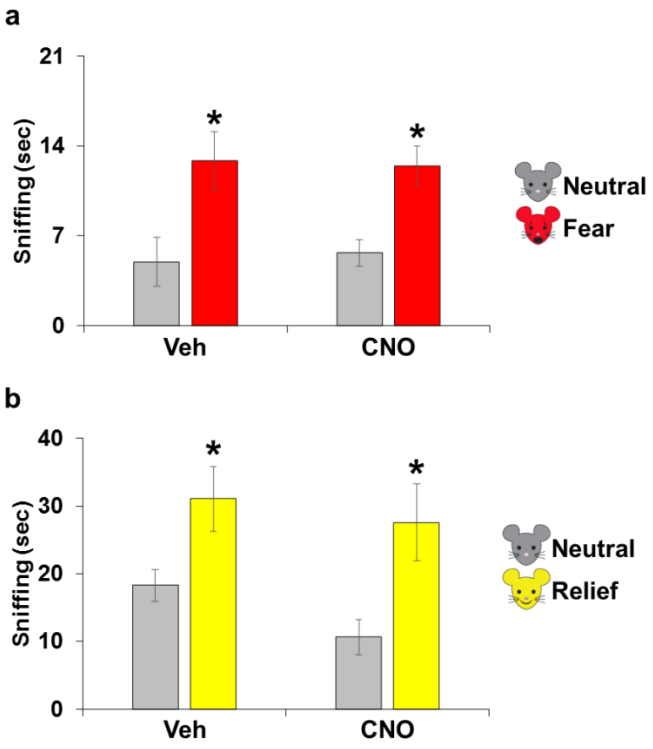


Figure S7. No effects of CNO treatment in emotion discrimination abilities. Time (in seconds) spent sniffing each wire cage containing two demonstrator mice during (a) the last 2 minutes and (b) the first 2 minutes of the emotion recognition test displayed by the same observer mice treated with vehicle or CNO (i.p., 3 mg/kg in a volume of 10 ml/kg, 30 minutes before the test), and shown separately for each demonstrator's emotion. Time spent sniffing neutral demonstrators are depicted in grey and represented by grey bars. Time spent sniffing demonstrators with (a) tone-induced fearful states (last 2-min RM ANOVA Veh: $F_{1,5}$=8,55, p=0.032; CNO: $F_{(1,5)}$=8,60, p=0.042), or (b) water-induced relief states (first 2-min RM ANOVA Veh: $F_{1,5}$=9,55, p=0.027; CNO: $F_{1,5}$=10,56, p=0.022) are represented by red or yellow bars, respectively. *P<0.05 versus the neutral demonstrator within the same observer treatment. N=6 observers per group.

Figure S8

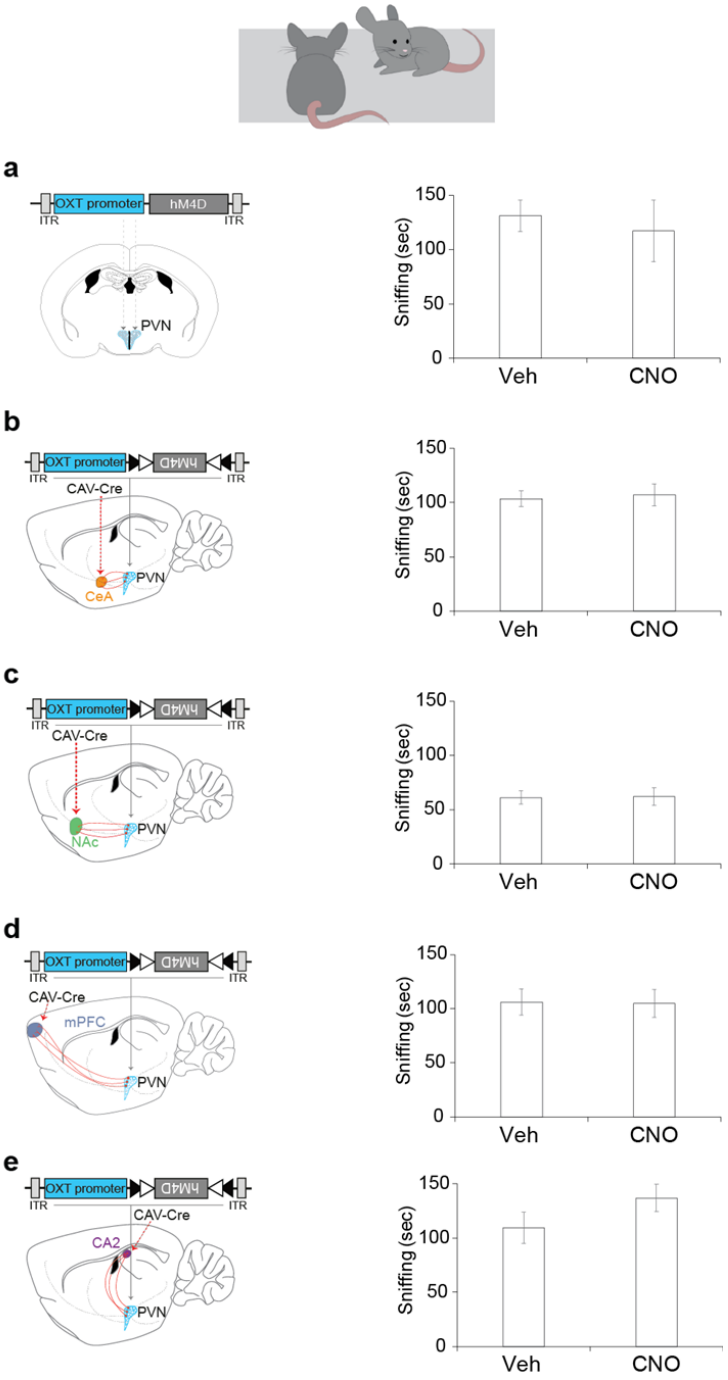


Figure S8. Inhibition of OXT projections did not alter free social interaction with an unfamiliar conspecific. (a-d) Schemes showing the injection of viruses in the PVN and respective projecting area. (e) Time (in seconds) spent sniffing an unfamiliar conspecific matched for sex and age during 4 minutes of free social interaction displayed by the same mice treated once with vehicle and another time with CNO (i.p., 30 minutes before the test). PVN (1-way ANOVA $F_{1,4}$=0,19, p=0.68); PVN-CeA (1-way ANOVA $F_{1,18}$=0,081, p=0.78); PVN-NAc (1-way ANOVA $F_{1,16}$=0,006, p=0.94); PVN-mPFC (1-way ANOVA $F_{1,10}$=0,005, p=0.94). N=4-10 per group.

PVN OXT projections to the central amygdala are an essential neural substrate for emotion recognition

To investigate the selective OXT-ergic circuits involved in emotion recognition, we first visualized PVN OXT projections in mice, injecting a rAAV expressing Venus under the control of an OXT promoter, that allowed the fluorescent labeling of OXT-PVN neurons (Fig. 4d). Our assessment focused on brain areas indicated as potential neuroanatomical substrates of emotion discrimination in humans[102,103] 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; Fig. 4e). Fewer fibers were evident in the insula, basolateral (BLA) and medial (MeA) amygdala (Fig. S9). To investigate the functional role of selective PVN OXT projections, we injected target areas with the retrogradely transported canine adenovirus 2 expressing Cre recombinase (CAV2-Cre). To interfere with the release of OXT from back-labeled PVN cells, we combined CAV2-Cre bilateral injections with the injection in the PVN of 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 promoter48. With this combination, we achieved DREADD(Gi)-mCherry expression exclusively in PVN OXT neurons projecting to the area injected with CAV2-Cre. Selective inhibition of OXT neurons projecting from the PVN to the CeA (Fig. 5a and Fig. S10b) was sufficient to recapitulate the deficits in emotion discrimination found by silencing all PVN projections (Fig. 4b-c). The same mice showed unimpaired emotion discrimination abilities when treated with vehicle (Fig. 5a). In contrast, selective inhibition of OXT neurons projecting from the PVN to the NAcc (Fig. 5b and Fig. S10c), the mPFC (Fig. 5c and Fig. S10d) and the CA2 (Fig. 5d and Fig. S10e) 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 recognition. Finally, none of the OXT pathways manipulations altered the ability to interact with an unfamiliar conspecific in a one-on-one free-interaction setting (Fig. S8b-e), indicating that the differences found in emotion discrimination are not due to alterations in other aspects of basic social interaction. We verified the regional and cell type specificity of virally-mediated labeling of OXT neurons (Fig. 5e and S11). Moreover, to control for the efficacy of DREADD-mediated inhibition in PVN back-labeled neurons from the different projection sites, we performed 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-labelled PVN neurons, which was equivalent for areas with intense OXTergic innervations (i.e CeA and CA2) or with more sporadic innervations (i.e. PFC; Fig. 5f-g and Fig. S12). Overall, these findings demonstrate a preponderant

contribution of the CeA in emotion recognition abilities in mice and indicate PVN OXT-ergic projections to the CeA as an essential neural substrate of such social cognitive function.

Figure S9

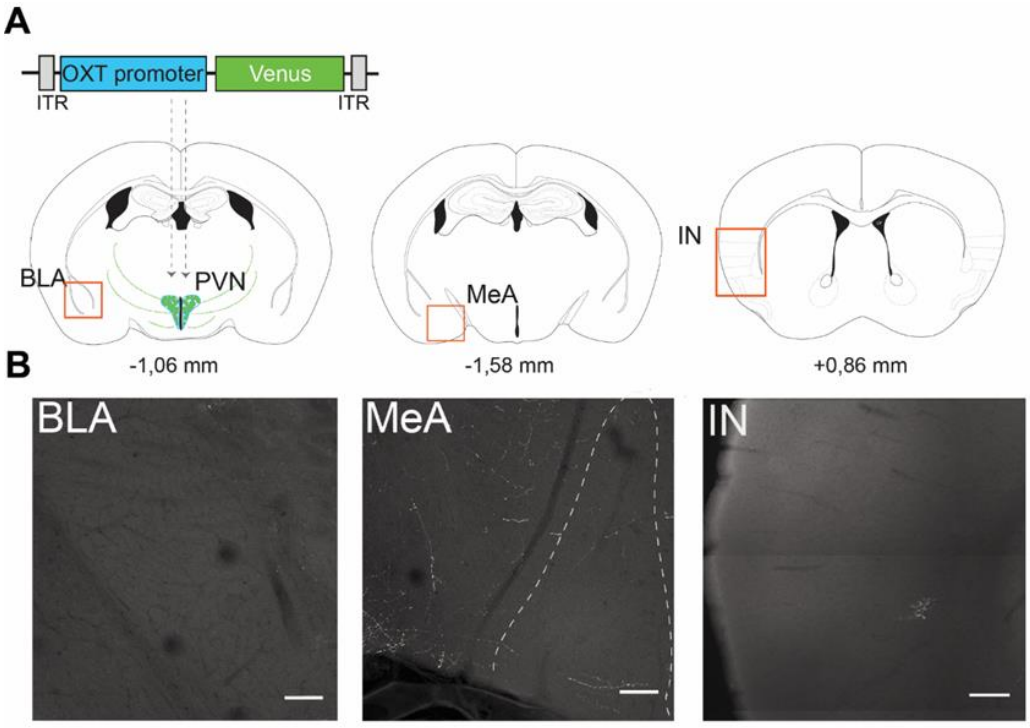


Figure S9. Anatomy of PVN OXT projections in mice. (A) Scheme of the viral vector used to infect the PVN OXT neurons. (B-C) Anatomical OXT projections from PVN to basolateral amygdala (BLA), medial amygdala (MeA), and insular cortex (IN). Scale bar: 100 µm.

Figure 5

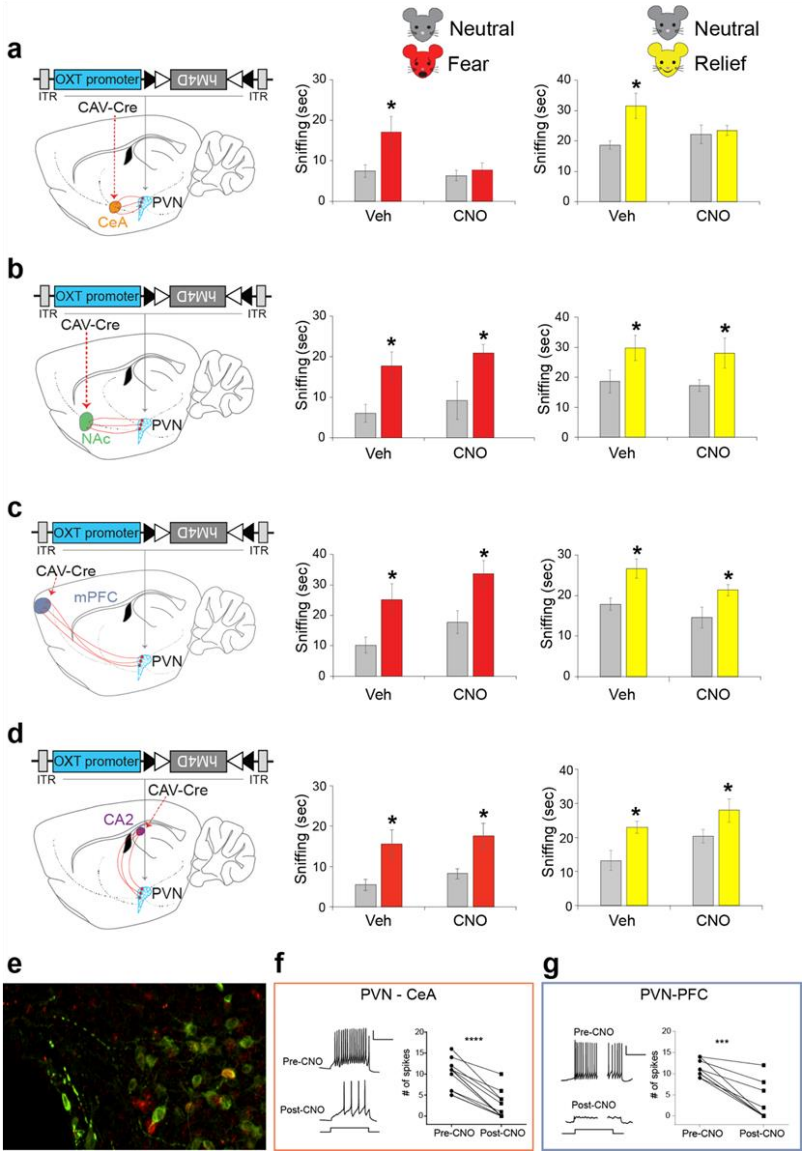


Fig5. PVN-Central Amygdala OXT projections are sufficient to mediate emotion recognition abilities in mice. (A, D, G) Schemes showing the injection of viruses in the PVN and respective projecting area. (B-C, E-F, H-I) Time (in seconds) spent sniffing each wire cage contatining two demonstrator mice during the first 2 minutes of the emotion recognition test displayed by the same observer mice treated with vehicle or CNO (i.p., 30 minutes before the test), and shown separately for each demonstrators emotion. Time spent sniffing neutral demonstrators are depicted in grey and represented by grey bars. Time spent sniffing demonstrators with (B, E, H) water-induced relief states, or (C, F, I) tone-induced fearful states are represented by yellow or red bars, respectively. PVN-CeA relief (first 2-min, RM ANOVA veh: F(1,7)=12,66, p=0.009; CNO: F(1,7)=0,13, p=0.73); PVN-CeA fear (last 2-min RM ANOVA veh: F(1,8)=5,76, p=0.043; CNO: F(1,8)=1,57, p=0.25); PVN-NAc relief (first 2-min, RM ANOVA veh: F(1,8)=7,56, p=0.025; CNO: F(1,8)=6,09, p=0.039); PVN-NAc fear (last 2-min RM ANOVA veh: F(1,5)=6,02, p=0.05; CNO: F(1,5)=7,40, p=0.042); PVN-CA2 relief (first 2-min, RM ANOVA veh: F(1,9)=0,018, p=0.90; CNO: F(1,9)=5,05, p=0.05); PVN-CA2 fear (last 2-min RM ANOVA veh: F(1,8)=11,39, p=0.010; CNO: F(1,7)=6,27, p=0.040). N=6-10 observers per group. *P<0.05 versus the neutral demonstrator within the same observer treatment.

Figure S10

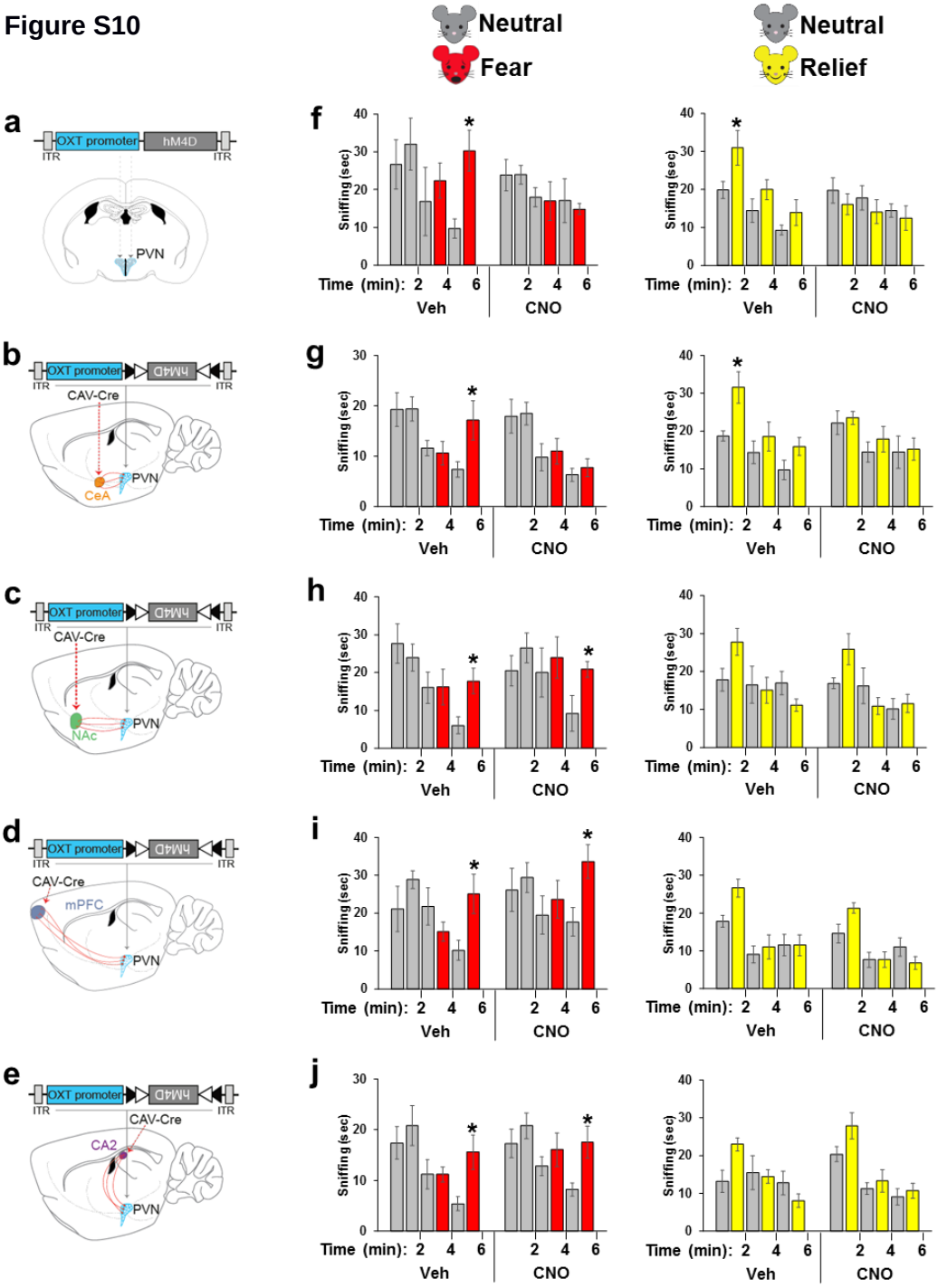


Figure S10. PVN-Central Amygdala OXT projections are necessary for emotion recognition in mice. (**a-e**) Scheme of the viral vector used to infect the PVN OXT neuronsand respective projecting area (CeA, NAcc, mPFC or CA2). (**b-c**) Time (in seconds) spent sniffing each of the two demonstrators, contained in the wire cage, during the emotion recognition test displayed by the same observer mice treated with vehicle or CNO (i.p., 30 minutes before the test), and shown separately for each demonstrator's emotion. Time spent sniffing neutral demonstrators are depicted in grey and represented by grey bars. Time spent sniffing demonstrators with tone-induced fearful states, or water-induced relief states are represented by red or yellow bars, respectively. N=5-11 observers per group. *p<0.05 versus the exploration of the neutral demonstrator.

Figure S11

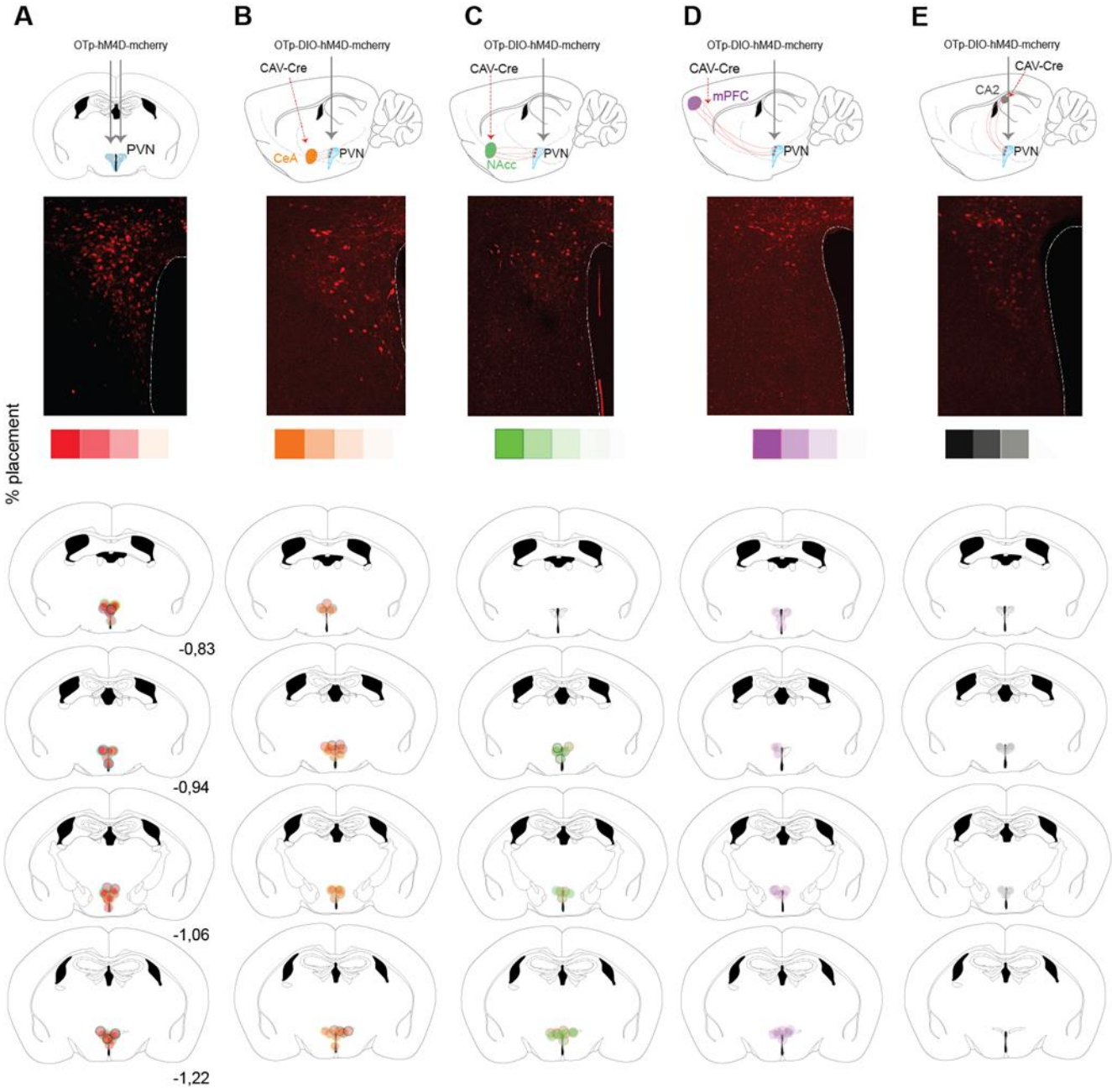


Figure S11. mcherry immunohistochemistry of retrogradelly labelled PVN neurons. Placements of selective inhibition of OXT neurons in (A) PVN by bilateral injection of hM4D DREADD receptor under the control of the OXT promoter, and in neurons projecting from the PVN to the (B) CeA, (C) NAcc, (D) mPFC, and (E) CA2 obtained by combined injection of CAV2-Cre bilateral injections with the injection in the PVN of rAAV carrying a double-floxed inverted open reading frame (ORF) (DIO) of the inhibitory hM4D DREADD receptor and mcherry under the control of the OXT promoter.

Figure S12

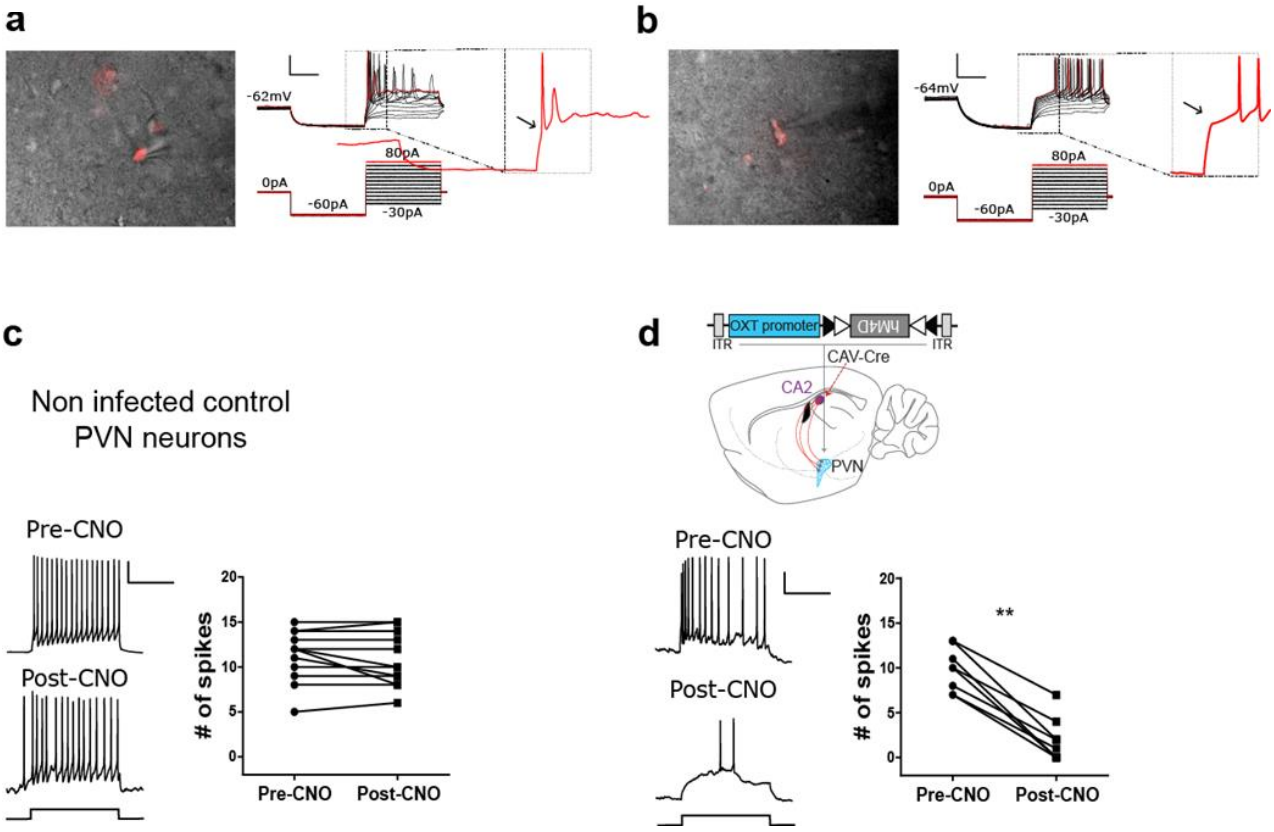


Figure S12. Electrophysiological validation of hM4D(Gi) action in PVN back labeled and control neurons. **(a)** Representative IR-DIC images of a back labeled parvocellular neuron in the PVN during patch-clamp recordings. Right: identification of the cell type by electrophysiological measurements under current clamp mode: the cell displays lack of transient outward rectification (black arrow in the zoomed trace) in response to depolarizing current pulses delivered at a hyperpolarized membrane potential. **(b)** Representative IR-DIC image of a magnocellular back labeled neuron in the PVN during patch-clamp recordings. Right: identification of the cell type by electrophysiological measurements under current clamp mode: the cell displays an inward rectification and a strong transient outward rectification (black arrow in zoomed trace) in response to depolarizing current pulses delivered at a hyperpolarized membrane potential. **(c)** 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 sec; amplitude: 20 pA) in control non infected neurons pre- and post- bath application of CNO in ACSF. Scale bars are 40 mV and 500 ms. Analyzed with two-tailed Wilcoxon matched-pairs signed rank test n=12 from 8 mice (p= 0.25). **(d)** 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 sec; amplitude: 20 pA) in PVN-CA2 back labeled neurons. Pre- and post- bath application of CNO in ACSF. Scale bars are 40 mV and 500 ms. **p< 0.01, two-tailed Wilcoxon matched-pairs signed rank test: n=8 from 3 mice (p= 0.0078).

Emotion recognition 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[104,105]. In heterozygous knockout mice for dysbindin-1 (Dys+/-), a clinically-relevant mouse model of cognitive and psychiatric liability[16,106,107] we identified reduced expression levels of OXT receptors (OXTR) compared to wild-type littermates (Dys+/+) in the CeA, but not in the basolateral (BLA) or medial (MeA)

amygdala (Fig. 6a-b). We then assessed Dys+/- mice emotion recognition abilities, and observed deficits in both the fear and relief conditions (Fig. 6c-d). In particular, we found that the impact of Dys mutation was selective for emotion recognition, as Dys +/- sociability and social memory in the classic 3-chamber test were similar to Dys +/+ controls (Fig. S13). These data unravel a clinically-relevant genetic variation which concurrently leads to deficits in emotion recognition abilities and in the CeA OXT system.

Next, to test if reduced OXTR levels in the CeA were responsible of Dys+/- mice emotion recognition deficits, we increased the expression of OXTR selectively in the CeA of Dys+/- mice by bilateral injection of a AAV-EF1a-OXTR-IRES-EYFP, expressing OXTR and EYFP (enhanced yellow fluorescent protein) under the control of the constitutively expressed EF1a promoter (Menon et al 2018; Fig. 6e, Fig. S14, Fig. S15). Increased OXTR levels within the CeA were confirmed by receptor autoradiography quantification (Fig S14c-d). Increasing OXTR levels in CeA of Dys +/- mice was sufficient to rescue their emotion recognition deficits (Fig. 6f-g). Altogether, these findings strengthen the conclusion that an appropriate OXTergic signaling within the CeA is a necessary condition for emotion discrimination abilities.

Figure 6

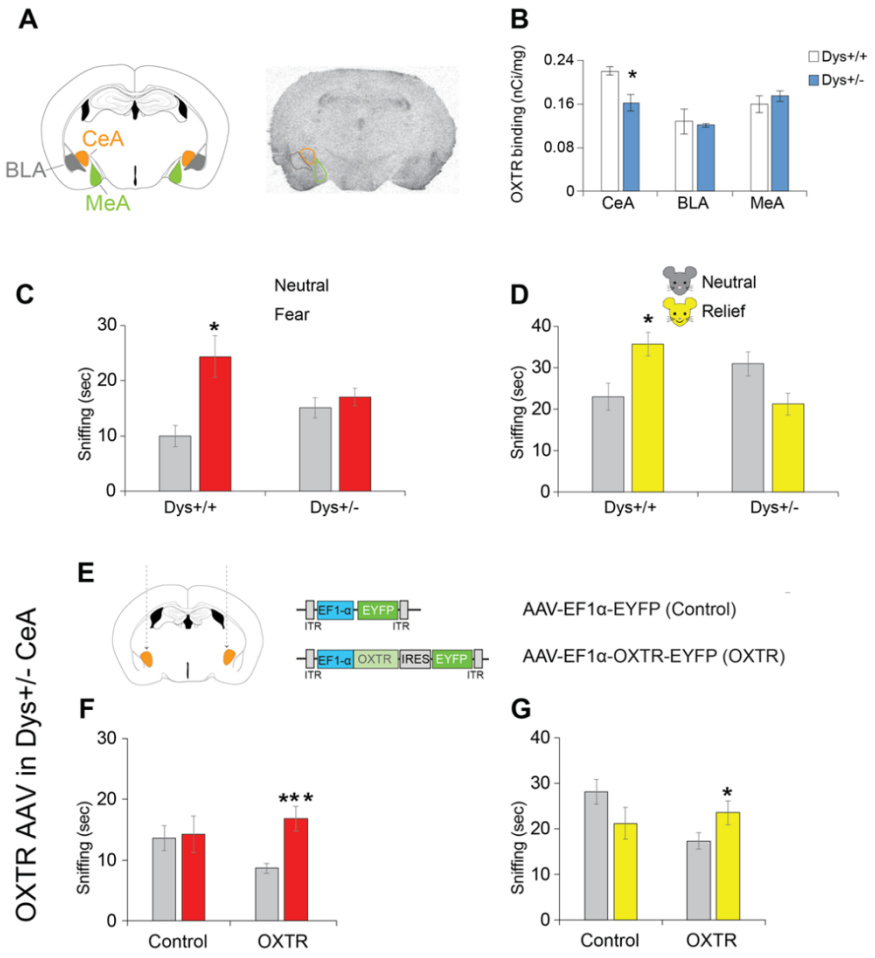


Figure 4. Intranasal OXT alters OXTR in the CeA and emotion recognition abilities depending on dysbindin-1 genetic background. (A) Representative drawing and autoradiograph showing the ligand binding of 20 pmol/l I125-labeled OVTA, a potent and selective ligand for OXTR. Autoradiograms were obtained from coronal sections of 3/4-month-old brains of mice chronically treated with intranasal OXT 0.3 IU/5 ml (OXT) or vehicle (Veh). CeA: central amygdala; BLA: basolateral amygdala. (B) Quantification of the autoradiographic I125receptors was obtained using NIH ImageJ- Software. Data is expressed as optical density. Two-Way ANOVA, genotype-by-treatment interaction: $F_{(2,21)}$=9,27, p=0.001. Ns=4-5 for each group; *p<0.05 vs veh-treated dys+/+ mice; #p<0.05 vs same genotype treated with vehicle. (C) OXTR mRNA expression in the CeA measured by real-time PCR in dysbindin-1 +/+, +/- and -/- mice chronically treated with intranasal OXT 0.3 IU/5 ml (OXT) or vehicle (Veh). Two-Way ANOVA, genotype-by-treatment interaction: $F_{(2,34)}$=6,15, p=0.005. Ns=4-8 for each group; *p<0.05 vs veh-treated dys+/+ mice; #p<0.05 vs same genotype treated with vehicle. (D-G) Time (in seconds) spent sniffing each wire cage containing two wild-type demonstrator mice during the first 2 minutes of the emotion recognition test displayed by dysbindin-1 wild-type (Dys+/+), heterozygous (Dys+/-) and homozygous (Dys-/-) knockout observer mice, and shown separately for each demonstrators emotion. Time spent sniffing neutral demonstrators are depicted in grey and represented by grey bars. Time spent sniffing demonstrators with (D, F) water-induced relief states, or (E, G) tone-induced fearful states are represented by yellow or red bars, respectively. (D-E) All observer mice were naïve or received two intranasal administrations of saline (0.9% NaCl) for 7-9 consecutive days in a volume of 5 microliter on both nostrils. RM ANOVAs, (D) Relief first 2-min, dys+/+: $F_{(1,6)}$=12,24, p=0.012; dys+/-: $F_{(1,6)}$=5,11, p=0.06; dys-/-: $F_{(1,6)}$=0,37, p=0.56. (E) Fear last 2-min, dys+/+: $F_{(1,6)}$=12,41, p=0.012; dys+/-: $F_{(1,3)}$=0,33, p=0.61; dys-/-: $F_{(1,8)}$=0,006, p=0.94. N=4-9 observers per group. *P<0.05 versus the neutral demonstrator within the same genotype. (F-G) All observer mice received two intranasal administrations of OXT (0.3 IU/5 microliter) for 7-9 consecutive days in a volume of 5 microliter on both nostrils. RM ANOVAs, (F) Relief first 2-min, dys+/+: $F_{(1,6)}$=1,83, p=0.23; dys+/-: $F_{(1,7)}$=9,76, p<0.02; dys-/-: $F_{(1,8)}$=5,82, p<0.05. (E) Fear last 2-min, dys+/+: $F_{(1,11)}$=0,001, p=0.98; dys+/-: $F_{(1,10)}$=1,73, p=0.22; dys-/-: $F_{(1,8)}$=8,89, p<0.02. N=7-12 observers per group. *P<0.05 versus the neutral demonstrator within the same genotype.

Figure S13

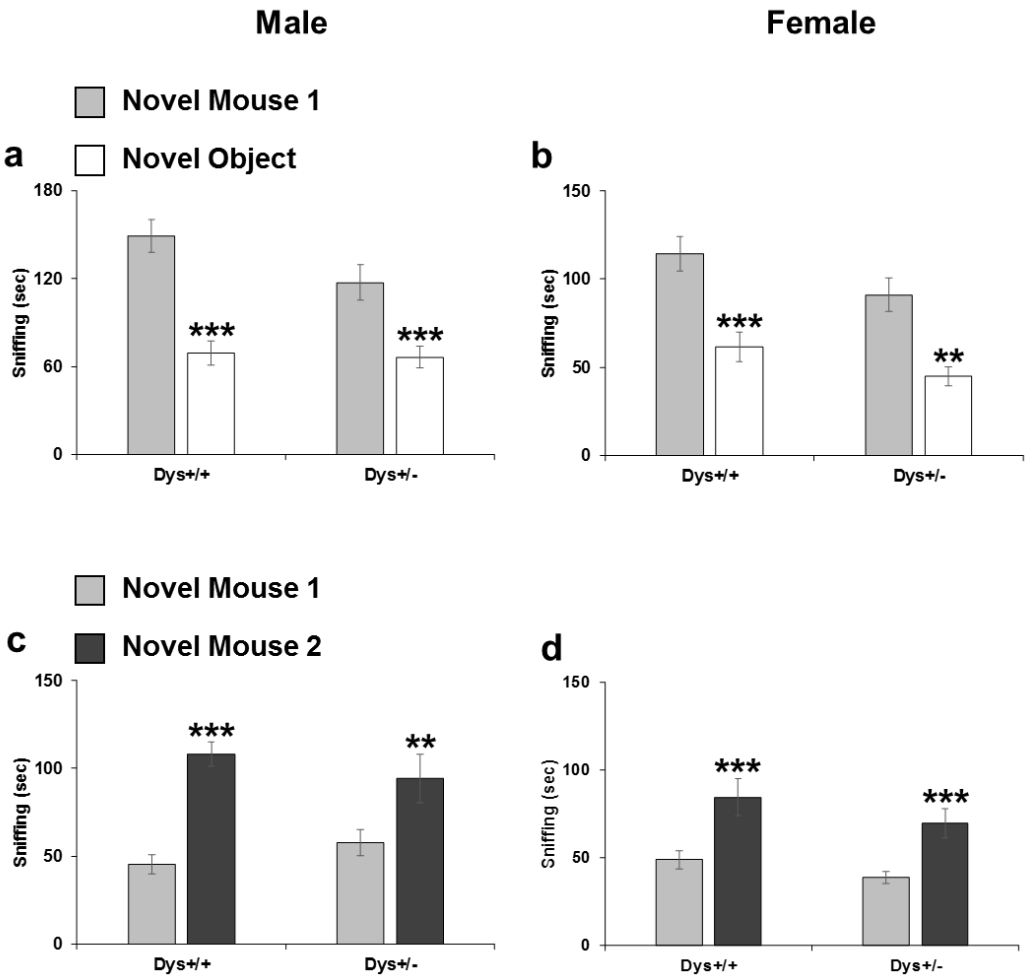


Figure S13. Dysbindin-1 heterozygous knockout mice show normal sociability and preference for social novelty in the 3-chamber task. (a-b) Time (in seconds) spent sniffing each wire cage containing one unfamiliar mouse (novel mouse 1) and one novel object during the sociability phase of the 3-chamber task displayed by Dys+/+ and Dys+/- (a) male and (b) 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. (c-d) Time (in seconds) spent sniffing each wire cage containing 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+/- (c) male and (d) 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<0.05, **P<0.005 and ***P<0.0005 versus the novel mouse within the same genotype. N=8-14 per group.

Figure S14

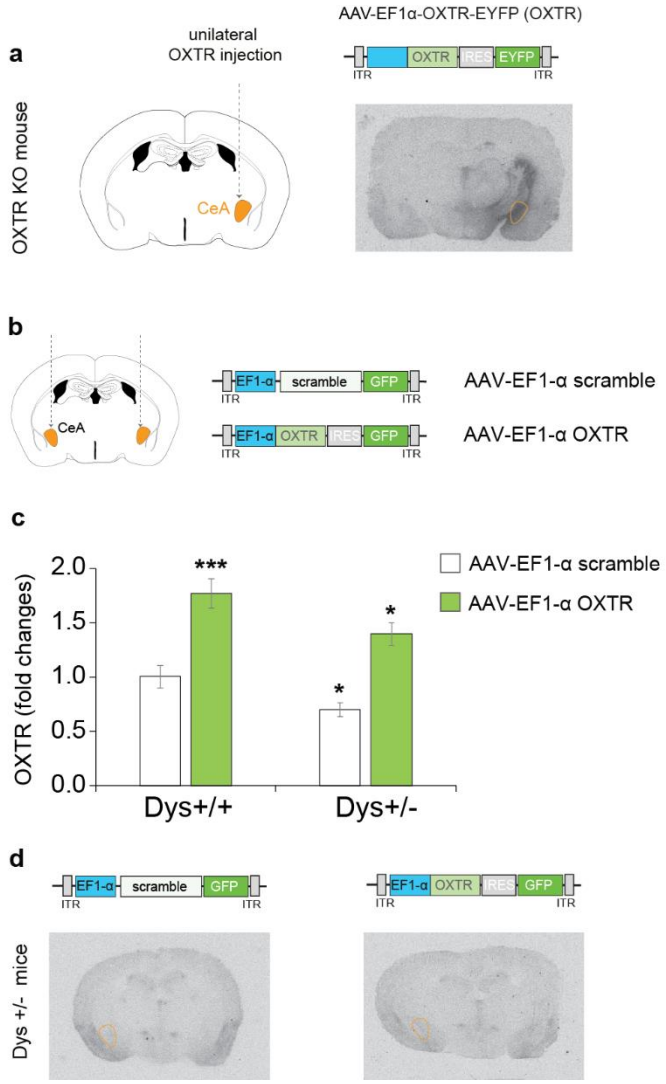


Figure S14. Validation of AAV virus expressing OXTR. The AAV-EF1α-OXTR-EYFP (OXTR) or AAV-EF1α-EYFP (Control) viruses were injected (**a**) unilaterally in the central amygdala (CeA) of OXTR knockout mice, and (**b**) bilaterally in the CeA of Dys+/+ and Dys+/- littermates determining an increased expression of OXTR assessed by autoradiography using the ligand binding of 20 pmol/l I125-labeled OVTA, a potent and selective ligand for OXTR. Autoradiograms were obtained from coronal sections. (**a** and **d**) autoradiography of OXT binding in OXTR-/- and Dys+/- mice, respectively, to demonstrate the viral OXTR expression under a background of no OXTR expression and in the dysbindin mice in which the behavioral deficits were rescued. (**c**) Quantification of the autoradiographic I125receptors in Dys+/+ and Dys+/- was obtained using NIH ImageJ- Software. OXTR binding sites are expressed as fold changes from the Dys+/+ control group. The *post-hoc* analyses on the Two-Way ANOVA genotype-by-treatment interaction revealed that both OXTR virus groups increased the OXTR protein levels in the CeA, while we confirmed reduced levels in the CeA of Dys+/- control virus mice. N=3-6 for each group; ***p<0.0005 and *p<0.05 versus Dys+/+ control virus group.

Figure S15

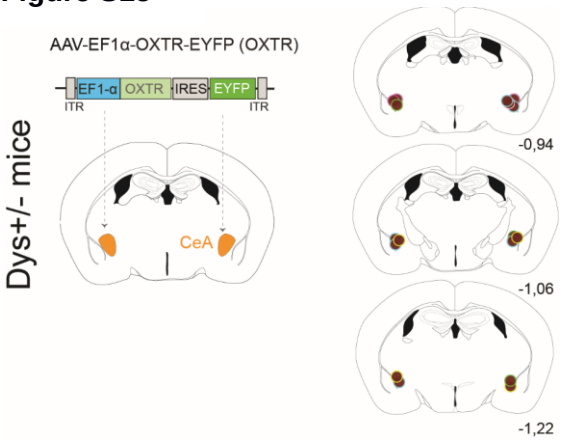


Figure S15. Placements of dys+/- CeA injected with AAV virus expressing OXTR. Schematic drawing of injection sites of the AVV expressing OXTR and EYFP under the control of the EF1a promoter in the CeA of Dys +/- mice.

Discussion

Combining behavioral, chemogenetic and genetic approaches, in this study we demonstrated that, similarly to humans[102,108,109], CeA is implicated in mice emotion recognition. Moreover, we demonstrated that OXTergic projection from the PVN to the CeA is an essential neuronal substrate of emotion discrimination. These mechanisms were revealed by a new paradigm measuring in mice the ability to discriminate unfamiliar conspecifics based on their emotional state, thus extending the opportunity to investigate previously unexplored social cognitive processes in mice, distinct from basal sociability and other social functions such as emotional contagion.

The developed paradigm revealed the ability of mice to discriminate unfamiliar conspecifics based on altered emotional states. Notably, observers' emotion discrimination was evident after (and not during) the induction of the altered emotional state in the demonstrators. This is in line with theories suggesting that emotional states are differentiated from simple reflex responses in their persistence after the disappearance of the triggering stimuli[110]. Previous protocols, able to induce emotion contagion/fear transfer, involved observers witnessing conspecifics in pain or exposed to foot shock[89,90,93,95,96]. In contrast, the manipulations we performed in demonstrator mice were designed to alter their emotional states causing minimal physical distress. Consistently, we did not detect in observer mice responses associable to emotional contagion, as freezing, state-matching, escape behaviors or altered corticosterone levels. This suggests the involvement of a cognitive process distinct from the automatic response to a conspecifics negative-emotional state[91]. Relevantly, most of previous emotion-based tests in rodents require familiarity between the observer and the demonstrator and show some sex-dependent effects[89,93,95]. In contrast, emotion recognition abilities in our paradigm were evident towards unfamiliar conspecifics and equally observable in male and female mice. Similarly, human emotion recognition tasks are mostly based on presentation of unfamiliar conspecifics and show similar performance between males and females. Finally, emotion discrimination ability in our paradigm was not predicted by sociability toward unfamiliar individuals in a classic 3-chamber test or towards emotionally-altered mice in a one-on-one setting. Altogether, this indicates that our paradigm parallel what is commonly measured in humans with emotion recognition tasks[68,111]and measure a scarcely explored aspect of rodents' social cognition, which complement currently available tools.

Observer mice showed similar preference towards demonstrators in negative- or positive-valence emotional states. The study of the sensory modalities involved in the processing of negative and positive emotions, however, provided some insight on possibly different ethological meaning of the two paradigms we used. In particular, while the relief condition

promoted the approach of the observer in all the modalities explored, the fear condition produced divergent responses. That is, social approach immediately after fear was evoked in a mouse by a tone, or visual cues were obscured, and aversion in presence of a fear odor, or of a recently shocked demonstrator. Rodents have been reported to actively escape from intense aversive stimuli[95], from aversive USVs calls induced by heavy distress[97,112], and from odors emitted by a shocked, heavily stressed, defeated, or sick conspecific[98–100]. Despite it is difficult drawing a conclusion on which signals guide mice specific behavioral responses, our data suggest that, while different in their emotional valence, fear and relief are not perceived by the observers as alarm signals, but as socially relevant states, promoting higher social interest than the neutral condition. The ensemble of these results indicate the potential of this setting to investigate a range of behavioral responses induced by emotional stimuli of different intensity, possibly recruiting different sensory modalities.

The development of this emotion recognition task for mice allowed us to uncover the primary role played by the endogenous OXT system, and specifically PVN-to-CeA OXT projections, in emotion recognition. The functional mapping of the selected PVN projections identified the CeA as a necessary and sufficient site for OXT to allow recognition of negative and positive emotions in conspecifics. Notably, in humans, positive and negative emotions are equally discriminated and equivalent brain regions represents positivity and negativity[113]. In particular, our findings are in line with several evidence in humans indicating that the amygdala is primarily recruited across all forms of emotion perception, and especially in decoding emotional expressions in others[102,108,109]. Moreover, alterations in amygdala responses to happy and fearful emotions have been reported in neuropsychiatric conditions such as autism and schizophrenia[104,105], in association with OXTR genetic variants[114], or following intranasal OXT[72,74,79,115]. However, the biological meaning of this altered amygdala activation and its implication in the context of the functionality of the endogenous OXT system were still uncertain. Thus, our manipulations of endogenous PVN OXT-ergic projections and preliminary evidence that PVN OXT neurons are unlikely to simultaneously project to different brain regions (not shown), delineate a first neurobiological substrate for the ability to discriminate expression of emotions in others.

Dysbindin-1 hypofunctioning mice showed emotion recognition deficits, for both negative and positive emotions, which depended on their decreased levels of OXTR in the CeA. Genetic variations in dysbindin-1 are strongly associated with human intelligence[116]. Moreover, both mice and human studies revealed that dysbindin-1 functional genetic variants that reduce its expression consistently modulate higher-order cognitive functions[16,106,107]. Thus, our current findings extend previous evidence, implicating dysbindin-1 genetics also in social cognitive

functions. In particular, these new genetic findings strength the evidence that the OXT signaling within the CeA and, specifically, OXTR levels within this brain region constitute a final target to modulate emotion recognition abilities. Notably, common functional genetic variants in dysbindin-1 can predict, in both humans and mice, cognitive responses to psychotropic drug treatments[106]. Thus, based on previous evidence suggesting that OXT treatments might have different outcomes depending on the status of the endogenous OXT system[80,81,117]and the current finding that dysbindin-1 alters the OXT CeA signaling, it is intriguing to speculate that dysbindin-1 genetic variants might also modulate social cognitive responses to exogenous OXT-related treatments.

An intriguing question raised by our data is how is the specificity of OXT action to emotion discrimination achieved in the CeA. The implication of CeA OXT signaling in conditioned threat responses[31,118] would suggest a specific ability in the detection of fear-mediated responses. The effects we found in the discrimination of both negative and positive states, however, indicates a more generalized role of the CeA in the detection of socially communicated salient information, as similarly evident in humans[102,108,109,113]. OXT has been documented to increase neuronal firing rates, mainly through modulation of interneuron activity, resulting in lower background activity and enhanced information transfer[119–121]. The close proximity of PVN-OXT fibers to GABAergic CeA neurons expressing OXTR[31,118] suggests that a similar modulation might occur during emotion recognition in the CeA. Distinct neuronal populations in sub-regions of the CeA have been shown to control specific behavioral responses to fear[31,118]. Understanding the specific cell-type identity and state of CeA neuronal subpopulations subtending the perception and response to emotional cues of different valence is an important line of research for future studies.

In conclusion, the data here presented provide significant new insights into the role of endogenous OXT signaling in the ability to recognize emotions in unfamiliar conspecifics. This was achieved by the development of a new method to address previously scarcely explored aspects of rodents' 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

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. Males and females C57BL/6J mice, and dysbindin-11 heterozygous (Dys+/-) and their wild-type littermates (Dys+/+), all of 3-6 month-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 (within 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. 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 discriminations reported in this work were performed and independently replicated by nine different researchers and independently replicated in three different laboratories.

Emotion recognition task

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) to 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, sex and genotype 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) were placed facing the long side of the cage and on top of the cage to record from different angles the three consecutive two-minute trials, using the Anymaze program (Stoelting, Ireland).

Observers. Before the test, mice were habituated to the experimental setting as reported above. The third day of habituation, mice were also habituated to the tone cue (4 kHz, 80 dB sound pressure level, three times for 30 seconds each with an intertrial interval of 90 seconds) without any conditioning. One hour prior to behavioral testing, mice were placed in the testing cage, in 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 in the testing cubicles. After having placed one emotionally 'neutral' and one "emotionally altered" demonstrator under the wire cups, the 6-minute experiment began. 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-naïve 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. 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 ad libitum all the time and some extra pellets were put inside the home cage during the 1-hour water reinsertion.

‘Fear’ demonstrators. The days before the test, mice were habituated to the experimental setting as reported above. ‘Fear’ demonstrators were fear conditioned using the parameters and context previously described2, 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 sec) 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 sec). 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 and in a separate room and apparatus (Ugo Basile SRL, Italy) respect to where the emotion recognition task would be performed. Fear demonstrator were 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 described3. 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 tested as previously reported4. Briefly, mice were individually placed in the testing cage 1 h prior to testing. No previous single housing manipulation was adopted to avoid any instauration of home-cage territory and 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 described5.

One-on-one social exploration tests

This test was similarly performed as previously described3. 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 in the testing sound proof cubicles. To begin the test a demonstrator mouse was introduced to the cage for 6 minutes (as for the emotion recognition task), and exploratory behaviors initiated by the test subject were timed by two independent experimenters blind to the manipulations. Demonstrators 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 cameras setting 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. 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 analysed for each test day included number of calls and duration of calls. Quantitative analyses of sound frequencies measured in terms of frequency and amplitude at the maximum of the spectrum were not performed because of the paucity of emitted USVs in all conditions performed.

For odor stimuli experiments, observers were tested as described above, but presenting as "demonstrator" only cotton balls impregnated with the odor 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 by gently brushing the cotton ball all over the body of the mice (especially including the nose, body and anogenital parts). Each odor was always freshly taken 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[6, 7]. 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 (21×21×21 cm) as previously described[6]. 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 (4×4 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: preconditioning 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 (preconditioning 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 recognition 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 time. Other than the two 23-hr deprivation periods, water was available all time. The neutral group was exposed to the same procedure but without any water deprivation. Post conditioning test was performed on day 5 in the same condition of the preconditioning test. For each mouse, a conditioning score was calculated as the post conditioning time minus the preconditioning time (in seconds) spent in the conditioned compartment of the apparatus.

Viral vectors

Generation of viral vectors. The OXTp-Venus, OXTp-hM4D(Gi), and OXTp-DIO-hM4D(Gi)-mCherry AAV serotype 1/2 were cloned and produced as reported previously8, 9. 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 µL. CAV2 equipped with Cre recombinase was (titer: 2.5x1011 pp) purchased from the Institute of Molecular Genetics in Montpellier CNRS, France10. The EF1a-OXTR-IRES:EYFP AAV 1/2 was cloned from the mOXTR-PCDNA3.1 plasmid kindly provided by Dr. M. Busnelli and Dr. B. Chini. mOXTR fragment (ORF) was released by BamHI and EcoRI restriction and inserted in an AAV 2 backbone carrying elongation factor 1. Next, the IRES:EYFP sequence was introduced via EcoRI and EcoRV ligation sites. E.coli TOP 10 cells were transfected with the helper plasmids and plasmids of both viruses and further steps of the viral production and purification were conducted in analogy to previous work8, 9. 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 µL.

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 2500 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.

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 (Paxinos and Watson, 1998): 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 pre-diluted at the 1x109 ppl/ml concentration. Only mice with appropriate placements were included in the reported data (Fig. S10, S11).

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 3 mg/kg in a volume of 10 ml/kg , 30 minutes before the emotional recognition 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-µm-thick coronal sections were cut using a Leica VT1000S 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 plus 4% 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, Invitrogen, CatNo. A-11122), or a dsRED polyclonal antibody (1:1000 Clontech, CatNo. 632496). Primary antisera were diluted in PBS plus 2% NGS overnight at 4°C for GFP antibody and 48 h at 4°C for dsRED antibody. 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 a Alexa Flour 488 goat anti-rabbit IgG (H+L) (1:1000, Molecular Probes®, CatNo.A11034) and Alexa Fluor 633 goat anti-rabbit IgG (H+L) (1:1000, Molecular Probes®, CatNo. A21071) in PBS plus 1% NGS, and subsequently washed there 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 naïve 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. 14µm-thick coronal sections were then cut with a cryostat and mounted on chrome-alum-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 described4.

Ex vivo electrophysiology

Virus-injected 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 25 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 µm 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 25 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 25 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, 1.25 mM NaH2PO4, 1 mM MgCl2, 2 mM CaCl2, 25 mM NaHCO3 and 25 mM glucose (~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 mΩ) were filled with an intracellular solution containing: 130 mM K-gluconate, 10 mM HEPES, 7 mM 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 mΩ), 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 literature11. Validation of iDREADDs was performed evoking spike firing in PVN neurons by injection of a small depolarizing current pulse (20 pA for 1 second) under current-clamp mode. Activation of iDREADDs was obtained using 10 μM 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) and analyzed using pClamp 10.2 software (Molecular Devices). All chemicals were purchased from Sigma, otherwise specified.

Statistical analyses

Results are expressed as mean±standard error of the mean (s.e.m.) throughout. Each observer's behavior towards the two different demonstrator mice was analyzed using a within-groups Repeated Measures ANOVA (RM-ANOVA). The behaviors of the two demonstrators and recorded USVs 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. The accepted value for significance was $p<0.05$. Statistical analyses were performed using Statistica 13.2 (StatSoft).

Chapter 3

Somatostatin interneurons within the prefrontal cortex control emotion recognition in mice

Diego Scheggi, Francesca Managò, Federica Maltese, Stefania Bruni, ***Marco Nigro****, Gabriella Contarini, Valentina Ferretti, Alessandra Bonavia, Ofer Yizhar, Francesco Papaleo*

(Submitted)

Abstract

The prefrontal cortex (PFC) has been implicated in the processing of the emotional state of others through non-verbal communication. These social cognitive functions are altered in psychiatric disorders such as autism and schizophrenia, and are hypothesized to rely on an intact cortical neuronal excitation/inhibition balance. Here, we show by combining in vivo electrophysiology with a behavioral task for emotion recognition in mice, that neurons in the medial prefrontal cortex (mPFC) are differentially activated during exploration of conspecifics depending on their affective state. Using optogenetics manipulations, we reveal that selective inhibition of mPFC somatostatin (SOM+) interneurons, but not parvalbumin (PV+), abolishes emotion discrimination. Conversely, activation of mPFC SOM+ interneurons is sufficient to induce emotion discrimination. Our findings provide new insight into the neurobiological mechanisms of emotion recognition. The prefrontal cortex (PFC) has been implicated in the processing of the emotional state of others through non-verbal communication. These social cognitive functions are altered in psychiatric disorders such as autism and schizophrenia, and are hypothesized to rely on an intact cortical neuronal excitation/inhibition balance. Here, we show by combining *in vivo* electrophysiology with a behavioral task for emotion recognition in mice, that neurons in the medial prefrontal cortex (mPFC) are differentially activated during exploration of conspecifics depending on their affective state. Using optogenetics manipulations, we reveal that selective inhibition of mPFC somatostatin (SOM+) interneurons, but not parvalbumin (PV+), abolishes emotion discrimination. Conversely, activation of mPFC SOM+ interneurons is sufficient to induce emotion discrimination. Our findings provide new insight into the neurobiological mechanisms of emotion recognition.

Introduction

Understanding others' emotions by perception of facial and body expressions is an ability crucially affecting our everyday life[68]. Impairments in these recognition measures are common in many neurodegenerative, neuropsychiatric and neurodevelopmental disorders. For example, emotion recognition deficits are core features of Autism Spectrum Disorders[66], and are strongly evident in schizophrenia[122] and in patients with prefrontal lesions[123]. Notably, these social cognitive impairments might have more deleterious impact on daily functioning than non-social cognitive deficits[69]. However, effective cures are still missing.

The "social brain", identified by human imaging studies, refers to a network subserving social cognitive processes, in which limbic and frontal regions are suggested to play a key role[124,125]. In particular, the top-down control of social cognitive functions is thought to be orchestrated by the prefrontal cortex (PFC) over the limbic system[126,127]. Indeed, damage of the medial PFC is associated with significantly impaired emotion recognition[128,129]. Thus, the PFC become an attractive brain regional target[129]. However, our understanding of PFC neural circuits and mechanisms underpinning emotion recognition remains incomplete, mainly due to the resolution level of manipulations allowed in humans and the lack of translational models.

Balance of neuronal excitation and inhibition governs cortical functions[130]. Perturbations in this balance are commonly invoked as a possible final shared pathway in the etiology of autism and schizophrenia[131]. For example, in humans, reduced interneurons[132,133] and alteration of GABAergic signaling[134,135] is a common feature of the autistic brain. In line with these human findings, disruption of the excitatory/inhibitory balance in the mPFC of mice led to social exploration deficits and sociability impairments[136]. Moreover, other rodents studies implicated the PFC in different social functioning such as social interaction[136–138], vicarious freezing[139,140], social hierarchy[141,142], and affiliative behavior[143]. However, the implication of PFC circuits and related excitatory/inhibitory balance in the ability to detect and process expression of emotions in others is still uncertain.

Here, we hypothesized that neuronal sub-populations within the mPFC could contribute differently in the processing of emotion discrimination. In humans, this social cognitive process is assessed by "emotion recognition tasks" that measure the ability to discriminate basic expression of emotions in others[144]. To explore mPFC circuits involved in emotion discrimination in a cell-specific manner, we devised a rodent equivalent of the human "emotion recognition task" (ERT). The ERT is designed to study the ability of mice to discriminate conspecifics based on their emotional state. This was objectively quantified as discriminatory behaviors between an emotionally-altered and a naïve conspecific. By in vivo electrophysiology, we first demonstrated the engagement of the mPFC in such discriminations.

We then conducted an optogenetics dissection of the involvement of different mPFC neuronal sub-populations. In particular, we found that mPFC interneurons expressing somatostatin (SOM+), in contrast to interneurons expressing parvalbumin (PV+), are both necessary and sufficient for the expression of emotion recognition.

Results

Mice can discriminate conspecifics based on the emotional state

In the emotion recognition task (ERT) we tested whether a mouse ("observer"), could distinguish between two unfamiliar conspecifics ("demonstrators"), based on their affective state. The observer was put in front of demonstrators, matched for sex and age, placed inside inverted wire cups, which were divided by a black wall (Fig. 1a). To induce changes in affective state, one of the two demonstrators ("relief") was exposed to a procedure consisting of 60 minutes of water restoration before the test, following 23 hours of water deprivation, and a naïve mouse with ad libitum water access ("neutral", Fig. 1a). We assumed that this procedure was associated with a positive-valence emotional state, because it resulted in a conditioned place preference (Supplementary Fig. 1a), and reduced corticosterone levels (Supplementary Fig. 1b). After habituation to the testing arena with empty cups, we presented to a naïve observer a relieved and a neutral unfamiliar demonstrator, which resulted in higher sniffing towards the relieved conspecific (Fig. 1b and Supplementary Fig. 2a) as well as more time spent in the related zone, compared to the neutral demonstrator (Fig. 1c). This behavior was evident during the first two minutes of observation (Fig. 1b,c and Supplementary Fig. 2a). Moreover, when demonstrators were familiar cage-mates, the observers showed a more persistent discrimination that lasted four minutes (Supplementary Fig. 3). Demonstrators showed no other observable behavioral differences (Supplementary Fig. 1c). Observers made similar number of visits to each demonstrator (Supplementary Fig. 2c). However, during the first 2 minutes, observers made on average longer visits to the relieved mouse (Supplementary Fig. 2d) and the latency to make the first visit was consistently lower towards the relieved compared to the neutral demonstrator (Fig 1d). We did not detect differences in other behaviors such as grooming and rearing in the observers (Supplementary Fig. 2g). Additionally, we tested female mice that revealed no differences for emotion discrimination compared to males (Fig. 1k and Supplementary Fig. 2f). Overall, this behavioral analysis suggest that mice are capable to discriminate between conspecifics in different affective states.

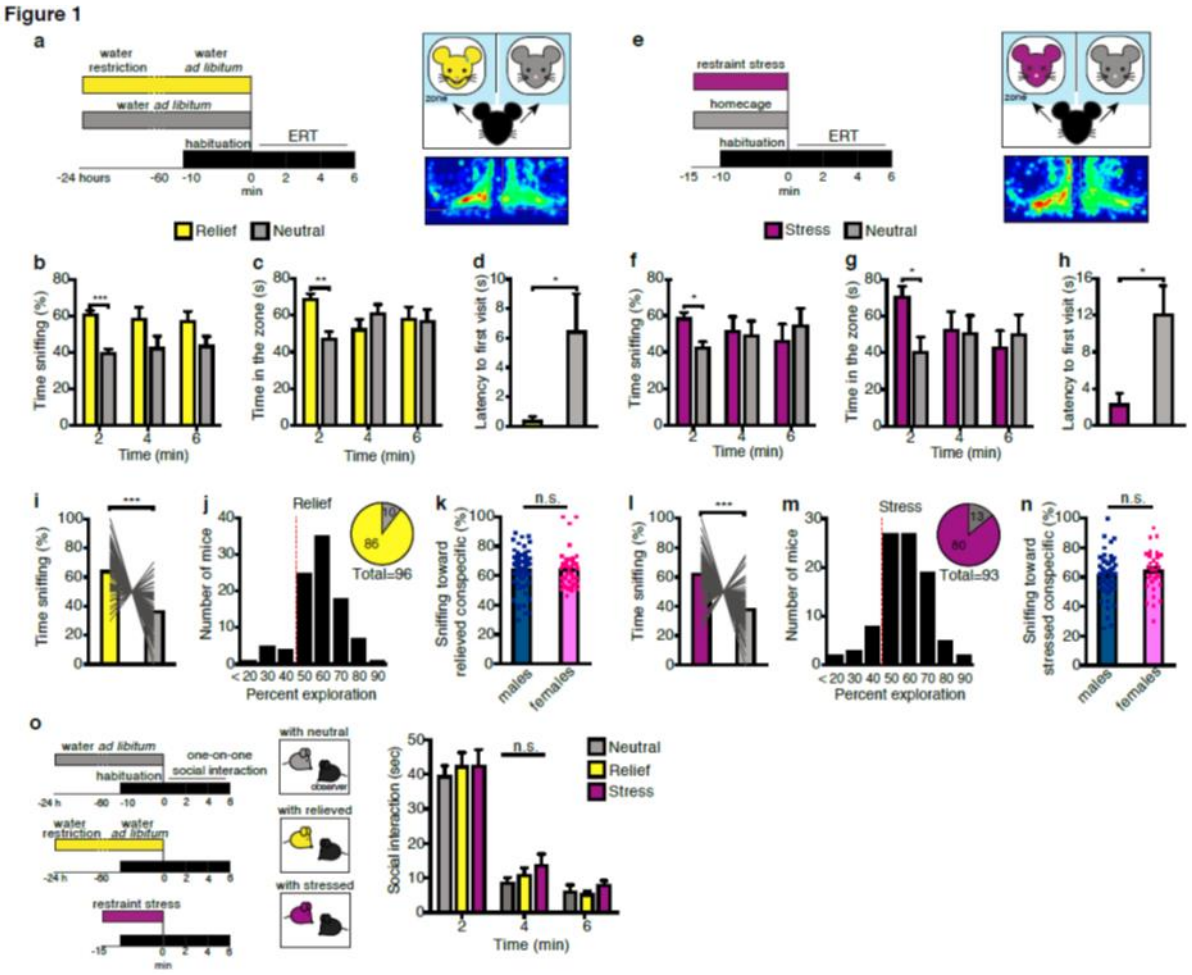
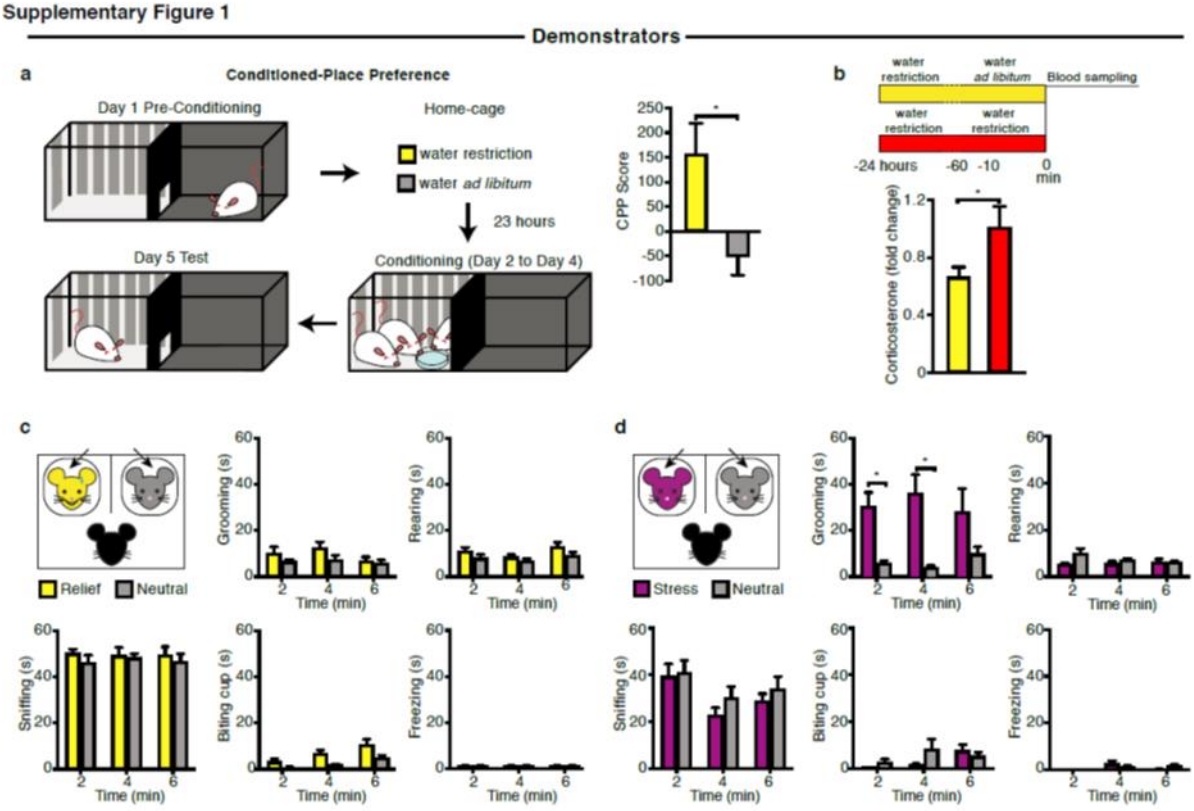


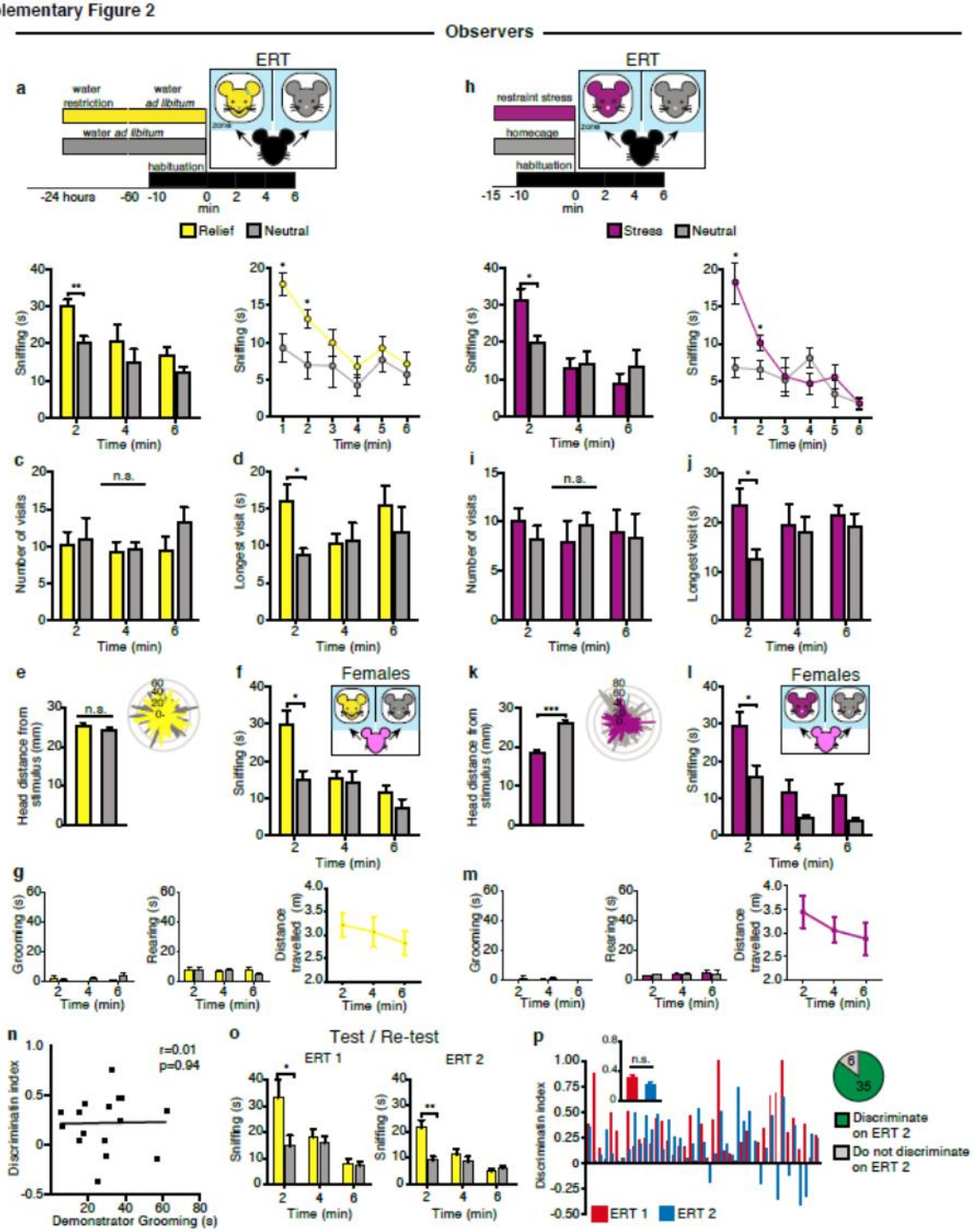
Figure 1. Mice can discriminate conspecifics based on their emotional state. (a) *Left***,** experimental design of the ERT. One demonstrator was given water access for 1 hour before the test, after 23 hours of water deprivation ("relief", yellow), while the other demonstrator had *ad libitum* water access ("neutral", grey). *Right*, schematic illustration of testing arena with stressed and neutral demonstrators (counterbalanced left and right across experiments) and graphical representation of amount of time observers spent in different parts of the apparatus (with blue as the shortest and red the longest time). **(b)** Increased exploration behavior (sniffing) to the relieved (Multiple t-test, Bonferroni correction, 2 min: t=6.22, df=14, p<0.0005; n=8 mice) and **(c)** increased time spent with the relieved demonstrator compared to the neutral, during the first 2 minutes of testing (Multiple t-test, Bonferroni correction, 2 min: t=4.28, df=10, p<0.005; n=6 mice). **(d)** Observers first visited the relieved mice, as latency to the first visit was significantly lower compared to the neutral (Unpaired t-test: t=2.31, df=10, p<0.05; n=6 mice). **(e)** In the stress protocol one demonstrator ("stress", purple) was subjected to restraint stress test for 15 minutes immediately before the beginning of ERT. The other demonstrator ("neutral", grey) waited undisturbed in the home-cage. **(f)** Increased sniffing to the stressed demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=3.22, df=12, p<0.05; n=7 mice) and **(g)** time spent with the stressed compared to the neutral (Multiple t-test, Bonferroni correction, 2 min: t=2.89, df=10, p<0.05; n=6 mice). **(h)** Latency to the first visit the stressed demonstrator was significantly lower compared to the neutral (Unpaired t-test: t=2.89, df=12, p<0.05; n=7 mice). **(i)** ERT was replicated several times and percentage of exploration towards demonstrators was pooled together (n=96 mice for relief manipulation and n=93 for the stress one). Observers explored more the relieved compared to neutral demonstrator during the first two minutes of ERT (Unpaired t-test: t=14.32, df=194, p<0.0005). **(j)** Exploration to the relieved demonstrator was higher than chance in a large number of mice, (86/96, one-sample t-test against chance, defined as 50%: t=10.12, df=97, p<0.0005) and **(k)** did not change depending on gender (male/females: 59/39; Unpaired t-test: t=0.10, df=96, p=0.91). **(l)** Increased exploration of the stressed compared to neutral demonstrator in several replication of the ERT (Unpaired t-test: t=11.63, df=184, p<0.0005). **(m)** Exploration of the stressed demonstrator was higher than chance in a large number of mice (80/93, one-sample t-test against chance, defined as 50%: t=8.22, df=92, p<0.0005) and **(n)** did not change depending on gender (male/females: 63/30; Unpaired t-test: t=0.92, df=91, p=0.35). **(o)** Mice did not show any difference in sociability when presented with a neutral or relieved or a stressed mouse in a standard free social interaction test (two-way ANOVA, time X group, $F_{(4,56)}$=0.21, p=0.93).

Supplementary Figure 1



Supplementary Figure 1. Observable behaviors of demonstrator mice. Related to Figure 1 **(a)** *Left*, Experimental design for the Conditioned-Place Preference (CPP) test, used to assess if the "relief" manipulation was associated with a negative- or positive-valence affective state. For each mouse a CPP score was calculated (post-conditioning – pre-conditioning time spent in the conditioning-paired compartment of the apparatus), with positive scores indicating place preference. Water-restricted mice ("relief", yellow) showed place preference of compartment conditioned with water compared to mice with *ad libitum* water access ("neutral", gray; unpaired t-test: t=2.67, df=14, p<0.05; n=7 mice). **(b)** Increased plasma corticosterone levels in water-restricted animals (red) compared to mice that received 1-hour of water restoration, following 23 hours of water restriction (yellow; unpaired t-test: t=2.00, df=19, p=0.05; n=10-11 mice). (**c**) Observable behaviors displayed by the neutral and relieved demonstrator mice during the 6 minutes of the ERT, divided by three consecutive 2-minute time beans. No significant emotion-by-time statistical interaction was evident for sniffing, grooming, rearing, biting, and freezing (n=10 demonstrators per group). **(d)** Observable behaviors displayed by the neutral and stressed demonstrator mice during the ERT. Stressed demonstrators showed increased grooming behavior (Multiple t-test, Bonferroni correction, 2 min: t=3.75, df=14, p<0.05; 4 min: t=3.70, df=14, p<0.05; n=8 mice). No significant emotion-by-time statistical interaction was evident for sniffing, rearing, biting, and freezing (n=8 demonstrators per group). *p<0.05, **p<0.005.

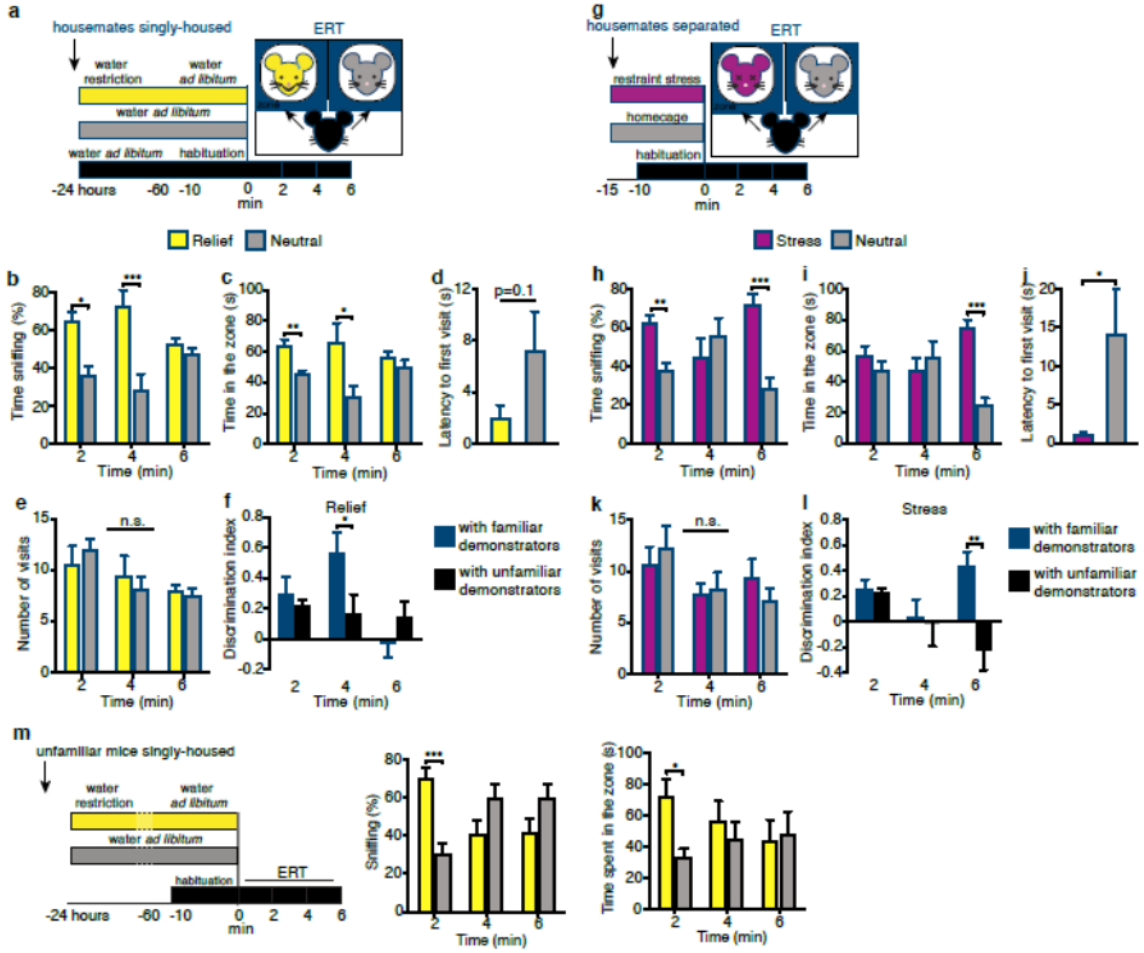
Supplementary Figure 2



Supplementary Figure 2. Mice can make discrimination of conspecifics based on the emotional state. Related to Figure 1 **(a)** *Top***,** Experimental design of the ERT. 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). Mice showed increased exploration behavior, measured as direct sniffing (in seconds), to the relieved compared to neutral demonstrator (left, showed in 120-seconds beams, 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<0.05, 120s: t=2.77, df=14, p<0.05; n=8). **(c)** Average number of visits to each zone did not differ. **(d)** Observers made longer visits in the zone related to the relief demonstrators (Multiple t-test, Bonferroni correction, 2 min: t=2.79, df=10, p<0.05; n=6). **(e)** Average distance of observers' head to the relieved and the neutral demonstrators did not differ during ERT. **(f)** Female mice showed increased sniffing to the relieved compared to neutral, sex-matched,

demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=3.32, df=18, p<0.05; n=10). **(g)** Grooming and rearing behaviors, and locomotor activity displayed by the observers during the ERT with the neutral and the relieved demonstrators. **(h)** *Top***,** In the stress protocol one demonstrator ("stress", purple) was subjected to restraint stress test for 15 minutes culminating in the beginning of ERT. 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, Multiple t-test, Bonferroni correction, 2 min: t=3.13, df=10, p<0.05; right, showed in 60-seconds beams, 60s: t=3.67, df=10, p<0.05, 120s: t=2.12, df=10, p<0.05; n=6). **(i)** Average number of visits to each zone did not differ. **(j)** Observers made longer visits in the zone related to the stressed demonstrators (Multiple t-test, Bonferroni correction, 2 min: t=2.56, df=10, p<0.05; n=6). **(k)** Average distance of observer mice during exploration to the stressed demonstrator was shorter compared to neutral (Unpaired t-test, Bonferroni correction, t=6.11, df=718, p<0.0005; n=6). **(l)** Female mice showed increased sniffing to the stressed compared to neutral, sex-matched, demonstrators (Multiple t-test, Bonferroni correction, 2 min: t=2.69, df=18, p<0.05; n=11). **(m)** Grooming and rearing behaviors, and locomotor activity displayed by the observers during the ERT with the neutral and the stressed demonstrators. **(n)** No correlation between discrimination index and grooming behavior of the stressed demonstrators. **(o)** First and second testing in the same ERT ("relief") showed similar behavioral pattern with increased sniffing towards the relieved demonstrator compared to the neutral (ERT 1: Multiple t-test, Bonferroni correction, 2 min: t=2.25, df=20, p<0.05; ERT 2: Multiple t-test, Bonferroni correction, 2 min: t=3.99, df=20, p<0.05; n=11). **(p)** For each observer tested in the ERT with both protocol (relief and stress) a discrimination index was calculated to compared performance on ERT 1 (red) and ERT 2 (blue; discrimination index = exploration of "relief"/"stress" - exploration of "neutral" / total time of exploration). Positive index means discrimination between "emotionally-altered" and "neutral". Of 41 mice tested in ERT 1 and ERT 2 only 6 did not show a positive discrimination index on second testing. Average discrimination index did not differ between ERT 1 and ERT 2.

Supplementary Figure 3



Supplementary Figure 3. Emotion discrimination in mice is enhanced between familiar of conspecifics. Related to Figure 1. **(a)** *Top***,** Experimental design of the ERT with cage-mates 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”, grey). **(b)** Increased sniffing to the relieved compared to neutral demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=3.60, df=12, p<0.05, 4 min: t=3.35, df=12; n=7). **(c)** Increased time spent in the zone related to the relieved demonstrator compared to the neutral (Multiple t-test, Bonferroni correction, 2 min: t=3.21, df=12, p<0.05, 4 min: t=2.37, df=12; n=7). **(d)** Shorter latency to make the first visit to the relieved demonstrator compared to the neutral (1.95±1.0 relief, 7.14±3.0, Unpaired t-test, t=1.58, df=12, p=0.13). **(e)** Average number of visits to each zone did not differ. **(f)** When tested with cage-mates, discrimination of relieved versus neutral demonstrators was longer as discrimination index was increased compared to mice tested with unfamiliar demonstrators (Multiple t-test, Bonferroni correction, 4 min: t=2.07, df=13, p=0.05, n=7). **(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 ERT. The other demonstrator (“neutral”, grey) waited undisturbed in his home-cage**. (h)** Increased sniffing to the stressed compared to neutral demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=4.27, df=12, p<0.005; 6 min: t=5.16, df=12, p<0.0005; n=7). **(i)** Increased time spent in the zone related to the stressed demonstrator compared to the neutral (Multiple t-test, Bonferroni correction, 6 min: t=6.13, df=12, p<0.0005, n=7). **(j)** Shorter latency to make the first visit to the stressed demonstrator compared to the neutral (Unpaired t-test, t=2.15, df=12, p<0.05). **(k)** Average number of visits to each zone did not differ. **(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 (Multiple t-test, Bonferroni correction, 6 min: t=3.45, df=11, p<0.005, n=6-7). **(m)** To rule out that social isolation 23 hours before testing (to allow water restriction of one cage-mate – “relief”) could have affected experiments with familiar mice, we tested singly-housed observers with unfamiliar demonstrators. Mice showed increased sniffing towards the relieved demonstrators (Multiple t-test, Bonferroni correction, 6 min: t=5.48, df=12, p<0.0005, n=7) and increased time spent in the related zone (Multiple t-test, Bonferroni correction, 6 min: t=2.86, df=12, p<0.05, n=7), only during the first 2 minutes of ERT, and not further, as showed in **b.**

Discrimination of a negative affective state

We next tested if emotion discrimination could be extended to a different, negative-affective state. To do this, we tested the ability of the observers to discriminate between one demonstrator that underwent a mild stress protocol, consisting of 15 minutes of acute restraint before the beginning of the ERT, and a neutral demonstrator (Fig 1e). We observed increased exploration towards the stressed demonstrator (Fig. 1f and Supplementary Fig. 2h) and higher time spent in the related zone (Fig. 1g). Also in this case, mice first entered the zone related to the “emotionally-altered” demonstrator (Fig. 1h) and made longer visits to the stressed demonstrator during the first 2 minutes of the test (Supplementary Fig. 2j), while the total number of visits did not differ between the stressed and neutral demonstrator (Supplementary Fig. 2i). Also in stress condition, when the demonstrator was a cage-mate of the observer, the discrimination was longer (Supplementary Fig. 3). Moreover, observer mice had a closer approach towards the stressed demonstrators as the distance of their head from the demonstrators during exploration was shorter for the stressed conspecific compared to the neutral (Supplementary Fig. 2k). Observer did not show differences in other behaviors such as grooming and rearing (Supplementary Fig. 2m). No differences between male and female mice were evident in the discrimination of the stressed demonstrators (Fig. 1n and Supplementary Fig. 2l). During the test, stressed demonstrators showed higher grooming compared to neutral mice, but similar rearing, sniffing, biting and freezing behaviors (Supplementary Fig. 1d). However, grooming was not correlated to observers’ ability to discriminate (Supplementary

Fig. 2n). Overall, these results showed that mice can similarly discriminate others based on both positive- and negative-valence states.

Emotion discrimination is a stable trait distinct from sociability

To evaluate the reliability of the emotion discrimination, we replicated the ERT several times in naïve mice, with different experimenters, in two different laboratories, 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 ERT conducted in naïve animals and in mice implanted with electrodes and under “light-OFF” condition were pooled and showed as percent exploration towards the emotionally-altered mouse (relief and stress, Fig. 1i,l). The emotion discrimination was a reliably observable behavior, with only 12% of tested mice not discriminating between an emotionally-altered and a neutral demonstrator (relief: 10/96, Fig. 1j; stress: 13/93, Fig. 1m). The scores of exploration towards the relieved and the stressed demonstrators were found to fit a normal distribution (D’Agostino and Pearson normality test, stressed: n=93, K2=1.54, p=0.46; relieved: n=96, K2=1.83, p=0.39). Moreover, emotion discrimination abilities were stable, as when re-exposed to the same (Supplementary Fig. 2o) or to a different paradigm observers showed similar behavior (Supplementary Fig. 2p). Thus, these data show that emotion recognition is a stable trait in mice.

Notably, if observer mice were tested in a one-on-one free social interaction setting with a neutral, a relief or a stress mouse (Fig. 1o), they spent a similar amount of time in social interaction with the emotionally-altered and neutral demonstrators and showed a classic habituation curve that was not influenced by the affective state of the demonstrator (Fig. 1o). This suggests that the discrimination revealed by the ERT is not due to a generalized increase of social exploration (an index of sociability), but rather is a more specific measure of emotion discrimination.

Enhanced mPFC neuronal activity during exploration of an emotionally-altered conspecific

To investigate the possible recruitment of mice mPFC in emotion discrimination, we implanted tetrodes in this region on the observer mice and we carried out chronic electrophysiological recordings during the ERT (Fig. 2a). In the “relief” versus “neutral” condition, we recorded 57 well-isolated units, with the majority recorded from the prelimbic cortex (PL). We classified these units into narrow-spiking (NS, Fig. 2b) putative inhibitory interneurons and wide-spiking

(WS, Fig. 2b) putative pyramidal neurons, based on spike waveform features, such as spike width and firing frequency[145]. In particular, we used three parameters: depolarization phase phase at half amplitude, duration of the hyperpolarization phase at half amplitude, and the average firing frequency during the entire recording session[146,147]. We found that among the recorded units, 39 out of 57 displayed a different activation during the direct exploration of the demonstrators (response to social exploration, Fig. 2b). The remaining cells (n=18) did not show any variation in their firing rate either before or during the social exploration, and did not display any stronger activation for one of the two emotional state in all of the three intervals we recorded ("no response", Fig. 2b). 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 only few units (2%) activated before the mouse started to explore one of the two demonstrators (Fig. 2b). Moreover, the majority of these cells discharged stronger when the mouse explored the relieved rather than the neutral conspecific and this behavior was observed throughout the 6 minutes of the test (table Fig. 2b). Finally, we observed that 79% of responsive neurons displayed a sustained activity for all the length of the social exploration (Fig. 2b). In particular, these neurons displayed a higher firing rate until the end of the social exploration that drastically decreased when the interaction ended. These responsive cells showed increased neuronal activation when exploring the relief demonstrator, compared to the neutral, which disappeared at the end of a social event (Fig. 2c). The same pattern of discharge was observed at the population level (Fig. 2d).

A similar mPFC neuronal activation patter was evident in observer mice performing the stress ERT paradigm. Of 83-isolated units, 66 responded to social exploration (Fig. 2f). We classified 52 of these units into NS putative interneurons and 31 WS putative pyramidal cells (Fig. 2f). The majority (77%) discharged just before and during the exploration, some of them only after the beginning of the interaction (20%), and only few units (3%) activated before the mouse started to explore one of the two demonstrators (Fig. 2f). An higher number of responsive units showed a sustained activity (85%) rather than a transient response. As for the relief condition, exploration of the stressed demonstrator led to higher firing of cells compared to the firing rates of the same units during exploration of the neutral demonstrator (Fig. 2g,h). Overall, these data indicate that exploration of a demonstrator in either a positively- or negatively-altered affective states leads to increased activation of mPFC neurons in mice. Consistent with human fMRI30 and brain lesions study[6], these results show that the mouse mPFC is engaged during emotion recognition, with the majority of cells responding to the expression of emotionally-altered states.

Figure 2

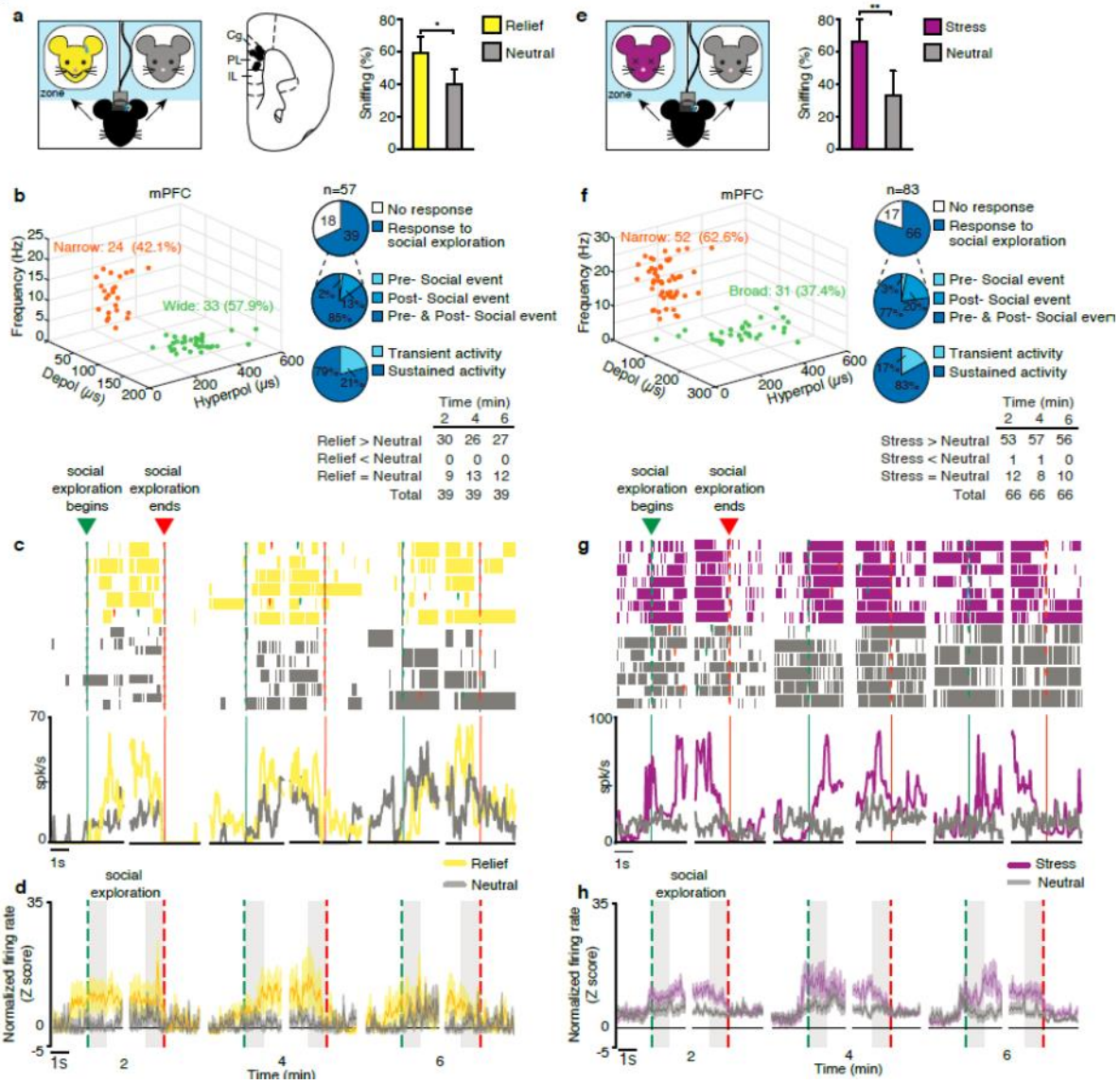


Figure 2. Enhanced neuronal activity during exploration of an emotionally altered conspecifics. (a) Left, mice were implanted with electrodes for chronic electrophysiological multi-unit and single-units recordings and tested in the ERT. Middle, electrodes placement in the mPFC (Cg, Cingulate; PL, Prelimbic area; IL, Infralimbic area). Right, chronic recording electrodes did not modify emotion discrimination. Mice showed increased exploration to the relieved demonstrator compared to the neutral during the first 2 minutes of testing (Unpaired t-test: t=2.33, df=10, p<0.05; n=6 mice). (b) Left, recorded cells were classified based on three properties: depolarization half-width, hyperpolarization half-width, and mean firing frequency. A hierarchical clustering method was used to separate cells into two populations: wide spike-width (putative pyramidal cell, green) or narrow spike-widths (putative interneurons; orange). (c) Top, Rasters and polylines aligned on the beginning (green dotted line) and the end (red dotted line) of each exploration of the relieved or neutral demonstrator in the same session, which were separated by a variable time interval of the duration of at least of 1 second. Rasters and histograms of single-neuron response when the observer explores different demonstrators are shown in different colors. Bottom, Examples of two responsive cells recorded in response to the exploration of relieved (yellow) and neutral (grey) demonstrators. (d) Population activity of all recorded neurons during the relief ERT (n=57) before, during and after social exploration of the two demonstrators throughout the 6-min experiment. Lines indicate the average discharge intensity of neurons when observers explored the relieved (yellow) and the neutral (gray) demonstrators aligned as the single neurons example in c. Colored shaded regions around each line represent 1 SEM. Gray shaded areas represent the windows used for statistical analysis of the populations response that highlight a stronger neuronal activation of both the populations during exploration of the relieved compared to neutral demonstrator that disappears when the exploration ends.(e) Mice showed increased exploration to the stressed demonstrator compared to the neutral

during the first 2 minutes of testing (n=7 mice, Unpaired t-test: t=4.12, df=12, p<0.005). (f) Left, recorded cells were classified as wide-spiking (putative pyramidal cell, green) or narrow-spiking (putative interneurons; orange). (g) Rasters and histograms of single-neuron response when the observer explored the stressed (purple) or the neutral (gray) demonstrators. Bottom, Examples of two responsive cells recorded in response to the exploration of stressed and neutral demonstrators. (d) Population activity of all recorded neurons in the stress ERT (n=83) before, during and after social exploration throughout the experiment. Purple and gray lines indicate the average discharge intensity of neurons when observers explored the stressed and the neutral demonstrators aligned as the single neurons example in g.

Enhancement of mPFC neuronal activity is specifically linked with emotion discrimination

To control for the specificity of the observed increase in neuronal activity during exploration of conspecifics in different affective states, we repeated the same experiments, but using two neutral demonstrators (Supplementary Fig. 4a). Observer mice similarly explored both demonstrators showing no discrimination of the two (Supplementary Fig. 4a). Moreover, no differences in neuronal activity were evident during exploration of the two neutral demonstrators (Supplementary Fig. 4c). This suggest that the observed increase of mPFC neuronal activation was specific to exploration of affective states of demonstrator mice.

We next investigated what sensory modality might trigger emotion discrimination and its related mPFC neuronal activation. No significant ultrasonic vocalization (USV) calls were recorded in the ERT (Supplementary Fig. 4d), and no differences in the emission of USVs were evident when “neutral”, “relief” and “stress” demonstrators were tested separately (Supplementary Fig. 4d). This indicate a marginal involvement of auditory cues (i.e. USVs) in mouse emotion recognition, consistent with previous literature showing that adult mice do not engage USV with conspecifics of same sex[148]. Similarly, we observed that visual cues were not essential, as observer mice tested in complete darkness showed the same increase of exploration towards both relieved (Supplementary Fig. 4e) and stressed demonstrators (Supplementary Fig. 4f). Thus, we checked if presentation of just odor cues could be sufficient to trigger the same activation pattern we revealed with the demonstrator mice (Fig. 2). Here, we recorded neuronal activity in mPFC when the observers were presented to the odors of an emotionally-altered and a neutral demonstrator (Supplementary Fig. 4g,j). To do this, odors were separately collected 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 demonstrator. Notably, in contrast to the results obtained in the ERT (Fig. 1), observers showed a marked avoidance towards the odor of a “stress” demonstrator (Supplementary Fig. 4j), while observers preferred to explore the odor from a “relief” demonstrator (Supplementary Fig. 4g). In these conditions, no differences in neuronal activation related to the exploration of the two

odors were evident (Supplementary Fig. 4i,l). Overall, these results revealed that odors alone did not elicit a similar behavioral and electrophysiological pattern as compared to demonstrator mice, confirming that the increase of mPFC neuronal activity (Fig. 2) was specifically elicited by the different affective states of demonstrators.

Supplementary Figure 4. Neutral demonstrators and odors did not recapitulate the activity pattern elicited by emotion discrimination. Related to Figure 2. **(a)** Mice have been implanted with recording electrodes in the mPFC and tested in the ERT with two naïve "neutral" demonstrators. Observers equally explored the two demonstrator and did not show observable discrimination. **(b)** We recorded 83 units and 55 were classified into narrow-spiking putative inhibitory interneurons and 27 into wide-spiking putative pyramidal neurons, based on spike waveform features, such as spike width and firing frequency[149]. **(c)** Population activity of recorded neurons (n=83) before, during and after social exploration throughout the experiment representing the average discharge intensity of neurons when observers explored neutral 1 (pink) and the neutral 2 (light green) demonstrators aligned on the beginning (green line) and the end (red line) of each exploration of the demonstrators in the same session, which were separated by a variable time interval of the duration at least of 1 second. Colored shaded regions around each line represent 1 SEM. Gray shaded areas represent the windows used for statistical analysis of the populations response that highlight a stronger neuronal activation of both the populations during exploration of the relieved compared to neutral demonstrator that disappears when the exploration ends. **(d)** *Top*, we measured ultra-sonic vocalization (USV) during the ERT with a relieved and one neutral demonstrator (yellow), and with a stressed and one neutral (purple). The number and the duration of calls was negligible in both conditions. *Bottom,* we measured USV of a relieved a stressed and a neutral demonstrator, separately and without observer. Also in this case the number and the duration of calls was negligible. **(e and f)** We tested observers in the ERT in the darkness. A visible light has been replaced with an infrared light to allow camera recording. All the other settings remained the same as described in Figure 1. Observers showed increased sniffing towards both the relieved **(e)** (Multiple t-test, Bonferroni correction, 2 min: $t=7.15$, $df=10$, $p<0.0005$; 6 min: $t=3.59$, $df=10$, $p<0.05$; n=7 mice) and **(f)** the stressed demonstrator (Multiple t-test, Bonferroni correction, 2 min: $t=6.13$, $df=10$, $p<0.0005$; 4 min: $t=3.08$, $df=10$, $p<0.05$). **(g)** Mice showed a preference for exploration of the odor of the relieved demonstrator (Multiple t-test, Bonferroni correction, 2 min: $t=3.57$, $df=42$, $p<0.005$). **(h)** We recorded 28 units, 18 were classified into narrow-spiking and 19 into wide-spiking. **(c)** Population activity of recorded neurons showed no difference between exploration of the odor of the relieved compared to the neutral demonstrator. **(j)** Mice showed a marked avoidance for the odor of the stressed demonstrator (Multiple t-test, Bonferroni correction, 2 min: $t=12.35$, $df=14$, $p<0.0005$; 4 min: $t=20.89$, $df=14$, $p<0.0005$). **(k)** We recorded 40 units, 21 were classified into narrow-spiking and 19 into wide-spiking. **(l)** Population activity of recorded neurons showed no difference between exploration of the odor of the relieved compared to the neutral demonstrator.

Supplementary Figure 4

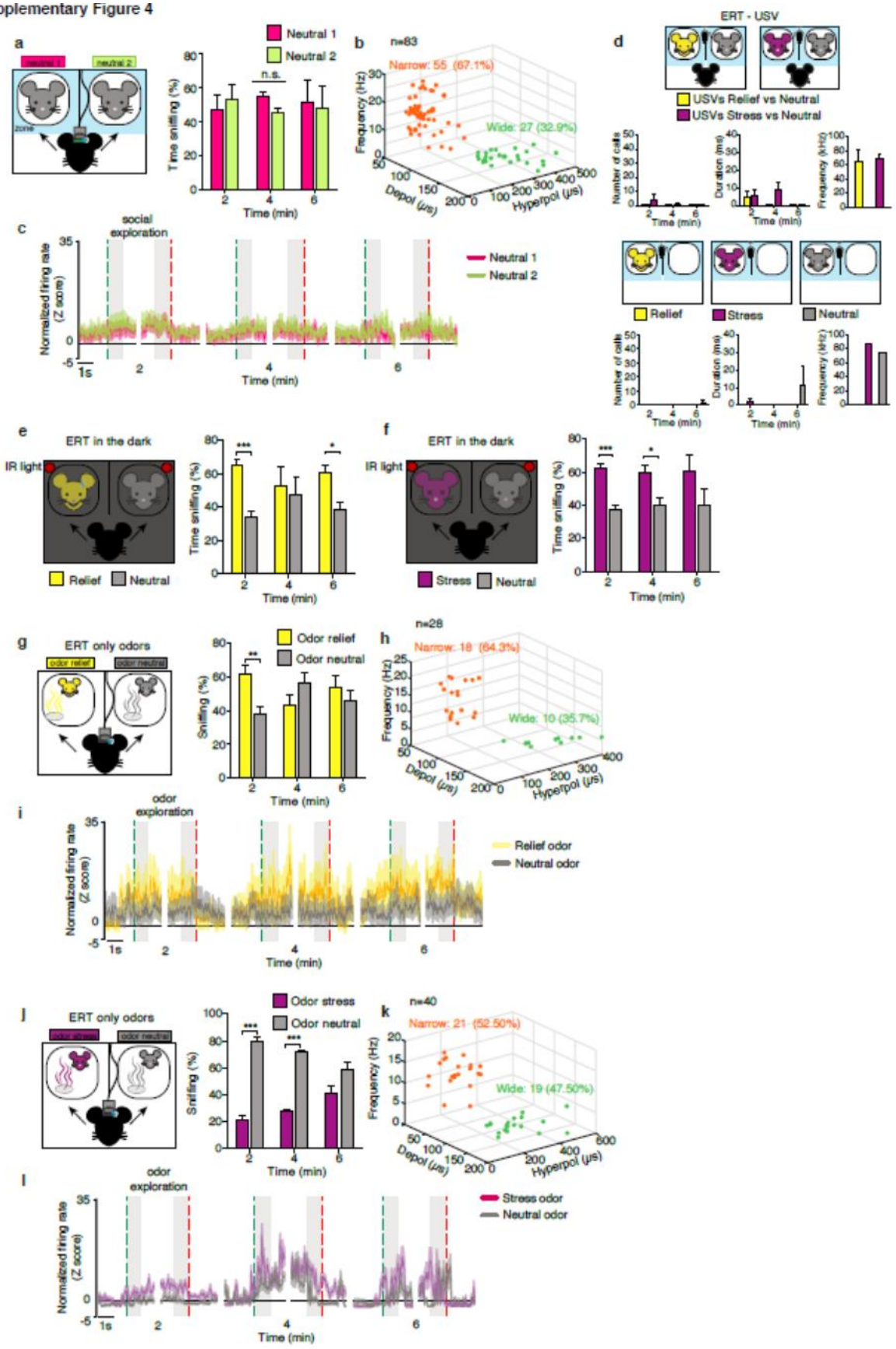


Photo-inhibition of mPFC PV+ interneurons do not affect emotion discrimination

Our electrophysiological findings during emotion discrimination revealed a major engagement of putative interneurons in the mPFC. The most abundant subpopulation of interneurons in the mPFC is represented by the parvalbumin-positive (PV+) cells[130]. Thus, we first investigated whether PV+ cells were necessary for emotion discrimination during the ERT. To specifically inhibit PV+ cells, we bilaterally injected a double-floxed inverted open-reading-frame (DIO) adeno-associated viruses (AAV) encoding eNpHR3.0 coupled to an eYFP tag into the mPFC of PV::cre transgenic mice, and implanted chronic optic fibers terminating dorsal to this area (Fig. 3a). We did not distinguish between different subtypes of PV+ interneurons, such as basket or chandelier cells. We targeted PV+ interneurons across all layers from L2 to L5 (Fig. 3a). We optically silenced mPFC PV+ cells activity with continuous green light during the first 2 minutes of the ERT (Fig. 3b,f), to target the time window in which we observed an increased exploration to the emotionally-altered demonstrator (Fig. 1). Mice were tested on consecutive weeks, with protocol (relief and stress) and treatment (light off and light on) counterbalanced. Photo-inhibition of PV+ cells reduced observers general investigation of demonstrators (Fig. 3c,g), an index of sociability. However, PV+ photo-inhibition did not modify observers' emotion discrimination in both the relief (Fig. 3d,e) and the stress (Fig. 3h,i) paradigms. PV-cre transgenic mice also first visited the zones related to the relief (Supplementary Fig. 5a) and stress demonstrators (Supplementary Fig. 5c), and the inhibition of PV+ cells did not modify this behavior (Supplementary Fig. 5a,c), as well as the total number of visits to each demonstrator (Supplementary Fig. 5b,d). These results suggest that PV+ cells in the mPFC might be involved in sociability, but not in emotion discrimination.

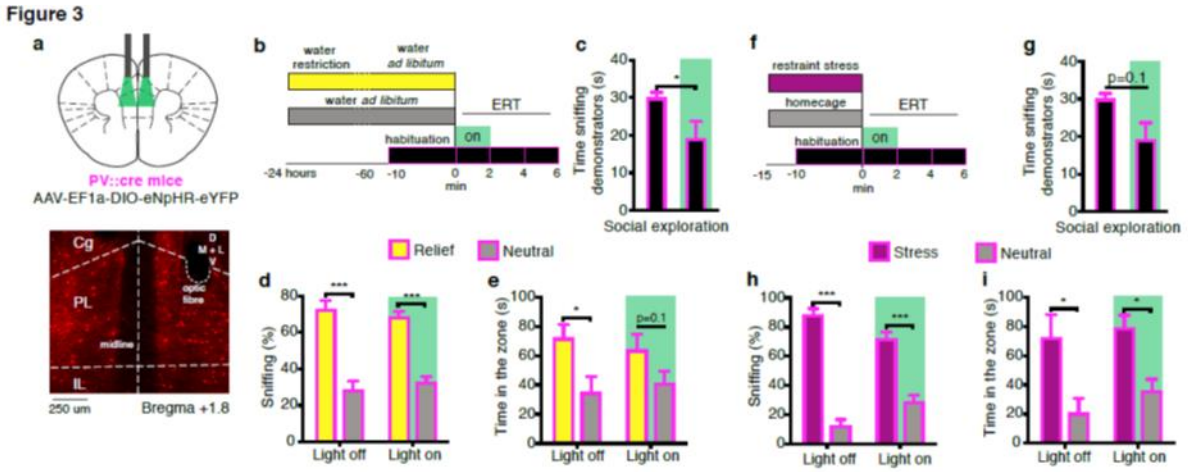
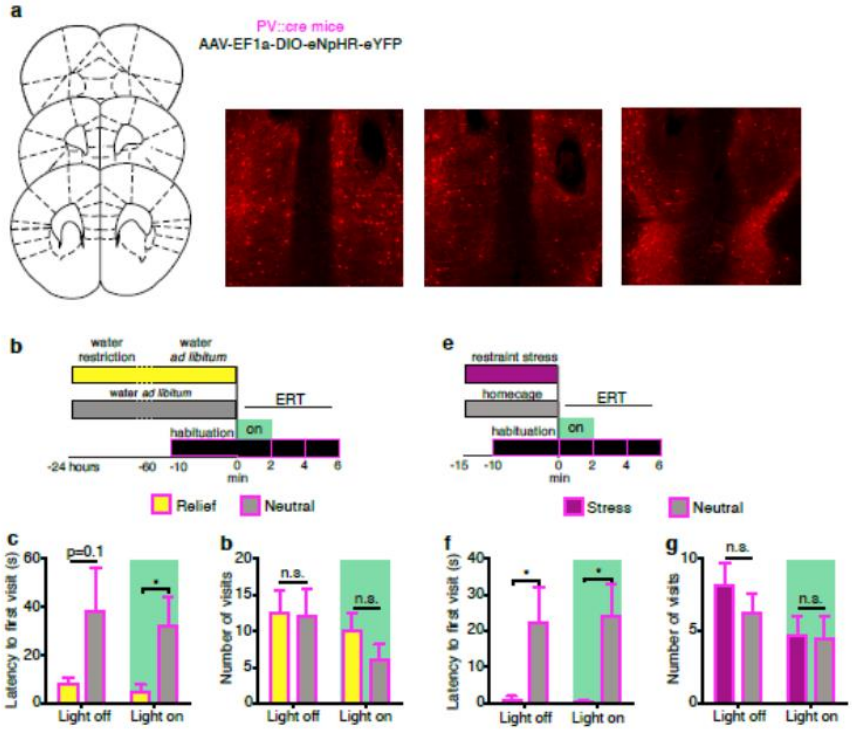


Figure 3. Photo-inhibition of PV+ interneurons did not affect emotion discrimination. (a) *Top***,** PV::cre mice were bilaterally injected in the mPFC with AAV-EF1a-DIO-eNpHR-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. *Bottom*, representative image of coronal mPFC section. **(b)** Mice were tested in the ERT with one relieved and one neutral demonstrator. Photo-inhibition was performed for 2 minutes, from the beginning of the test, using continuous green light (λ=532 nm). **(c)** Reduced social investigation during optical inhibition of PV+ cells (Unpaired t-test: t=2.12, df=12, p<0.05; n=6 mice). **(d)** Increased sniffing (expressed in %) to the relieved demonstrator compared to the neutral during the first 2 minutes of testing without light delivery ("Light off"; Multiple t-test, Bonferroni correction, 2 min: t=5.93, df=12, p<0.0005; n=7 mice) and during light stimulation ("Light on"; t=6.96, df=12, p<0.0005). **(e)** Increased time spent in the zone related to the relieved demonstrator compared to the neutral in Light off condition (Multiple t-test, Bonferroni correction, 2 min: t=2.45, df=12, p<0.05). With Light on mice spent more time, although not significantly (Multiple t-test, Bonferroni correction, 2 min: t=1.56, df=12, p=0.1), into the zone of the relieved mouse (63.12±10.55 relief, 40.7±8.03 neutral). **(f)** Mice were tested in the ERT with one stressed and one neutral demonstrators and photo-inhibition was performed for the first 2 minutes. **(g)** Tendency for Reduced social investigation during optical inhibition of PV+ cells (Light off: 36.6±5.30, Light on:21.44±7.85, Unpaired t-test: t=1.6, df=12, p=0.1; n=7 mice). **(h)** Increased exploration to the stressed demonstrator compared to the neutral in both "Light off" (Multiple t-test, Bonferroni correction, 2 min: t=11.46, df=12, p<0.0005, n=7) and "Light on" conditions (Multiple t-test, Bonferroni correction, 2 min: t=6.25, df=12, p<0.0005, n=7). **(i)** Higher time spent into the zone related to the relieved demonstrator compared to the neutral in both Light off (Multiple t-test, Bonferroni correction, 2 min: t=2.65, df=12, p<0.05, n=7) and Light on conditions (Multiple t-test, Bonferroni correction, 2 min: t=3.33, df=12, p<0.05, n=7).

Supplementary Figure 5



Supplementary Figure 5. Photo-inhibition of PV+ interneurons do not affect emotion discrimination. Related to Figure 3. (a) Left, Red areas represent the minimum (darker color) and the maximum (lighter color) expression of AAV-EF1a-DIO-eNpHR-eYFP in PV::cre mice. Right, representative images of viral expression in the mPFC (in rostro-caudal order) after injection with AAV-EF1a-DIO-eNpHR-eYFP. (b) PV::cre mice were tested in the ERT 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) PV-cre mice made the first visit to the relieved demonstrator both in light off and light on conditions (Multiple t-test, Bonferroni correction, off: $t=1.68$, $df=12$, $p=0.1$, on: $t=2.11$, $df=12$, $p<0.05$; $n=7$). (d) Optical inhibition of PV+ did not modify the number of visits to each demonstrator. (e) PV::cre mice were tested in the ERT with one stressed and one neutral demonstrator. Photo-inhibition was performed for 2 minutes, from the beginning of the test, using continuous green light. (f) PV-cre mice made the first visit to the stressed demonstrator both in light off and light on conditions (Multiple t-test, Bonferroni correction, off: $t=2.21$, $df=12$, $p<0.05$, on: $t=2.64$, $df=12$, $p<0.05$; $n=7$). (g) Optical inhibition of PV+ did not modify the number of visits to each demonstrator.

Photo-inhibition of mPFC SOM+ interneurons abolish emotion discrimination

Somatostatin-expressing cells constitute another major subtype of local GABAergic interneurons in the cerebral cortex[130]. To investigate the possible involvement of these interneurons, we bilaterally injected an AAV-EF1a-DIO-eNpHR-eYFP into the mPFC of SOM::cre transgenic mice and implanted a chronic optic fibers terminating dorsal to this area (Fig. 4a and Supplementary Fig. 6a). We targeted SOM+ interneurons across all layers from L2 to L5 (Fig. 4a and Supplementary Fig. 6a). Optical inhibition of SOM+ cells, with continuous green light during the first 2 minutes of the ERT, abolished emotion discrimination (Fig. 4d,h), without affecting social exploration (Fig. 4c,g). This effect was temporary and reversible, as

after cessation of photo-inhibition, we observed again an increased exploration of the relieved (Fig. 4e) and the stressed demonstrator (Fig. 4i). When tested on “light off” condition, these same mice showed the expected discrimination between the neutral and the stressed or relieved demonstrators (Fig. 4d,h). Moreover, testing these animals with the same illumination protocol, but presented with a neutral demonstrator and an object, or with a familiar and a novel demonstrators, as commonly used in the classic three-chambers test[150], did not influence sociability and social novelty discriminations (Fig. 4j,k). Further, photo-inhibition of SOM+ did not modify odor discrimination (Supplementary Fig. 6b,c). Overall, these results indicate that SOM+ are selectively implicated in the ability to discriminate conspecifics based on their affective state.

We next set out to control the photo-inhibition of SOM+ using a closed-loop system, such that the presence of the observer in the “zone" related to the relieved or the stressed demonstrators triggered the optical inhibition of mPFC SOM+ cells (Fig. 4l,p and Supplementary Fig. 6d,i). Similarly to what we observed in naïve mice (Fig. 1d), observers first visited the relieved and the stressed demonstrators (Fig. 4m,q). In line with above experiments, photo-inhibition of SOM+ paired to the exploration of emotionally-altered demonstrators abolished emotion discrimination during the first 2 minutes of testing (Fig. 4n,r and Supplementary Fig. 6e,j). However, time-locked photo-inhibition further modified the exploration over time, as observers started to explore and spend more time with the neutral compared to the emotionally-altered demonstrator (Fig. 4o,s and Supplementary Fig. 6e,j), while the number of visits to each demonstrators were not affected (Supplementary Fig. 6f,k). 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 discrimination (Supplementary Fig. 6n). Furthermore, photo-inhibition of SOM+ did not induce any gross motor deficits (Supplementary Fig. 6g-m) and light delivery in non-infected mice did not induce any place avoidance or motor abnormalities (Supplementary Fig. 7). Overall, these experiments indicate that SOM+ interneurons in the mPFC are necessary for emotion discrimination in mice.

Figure 4

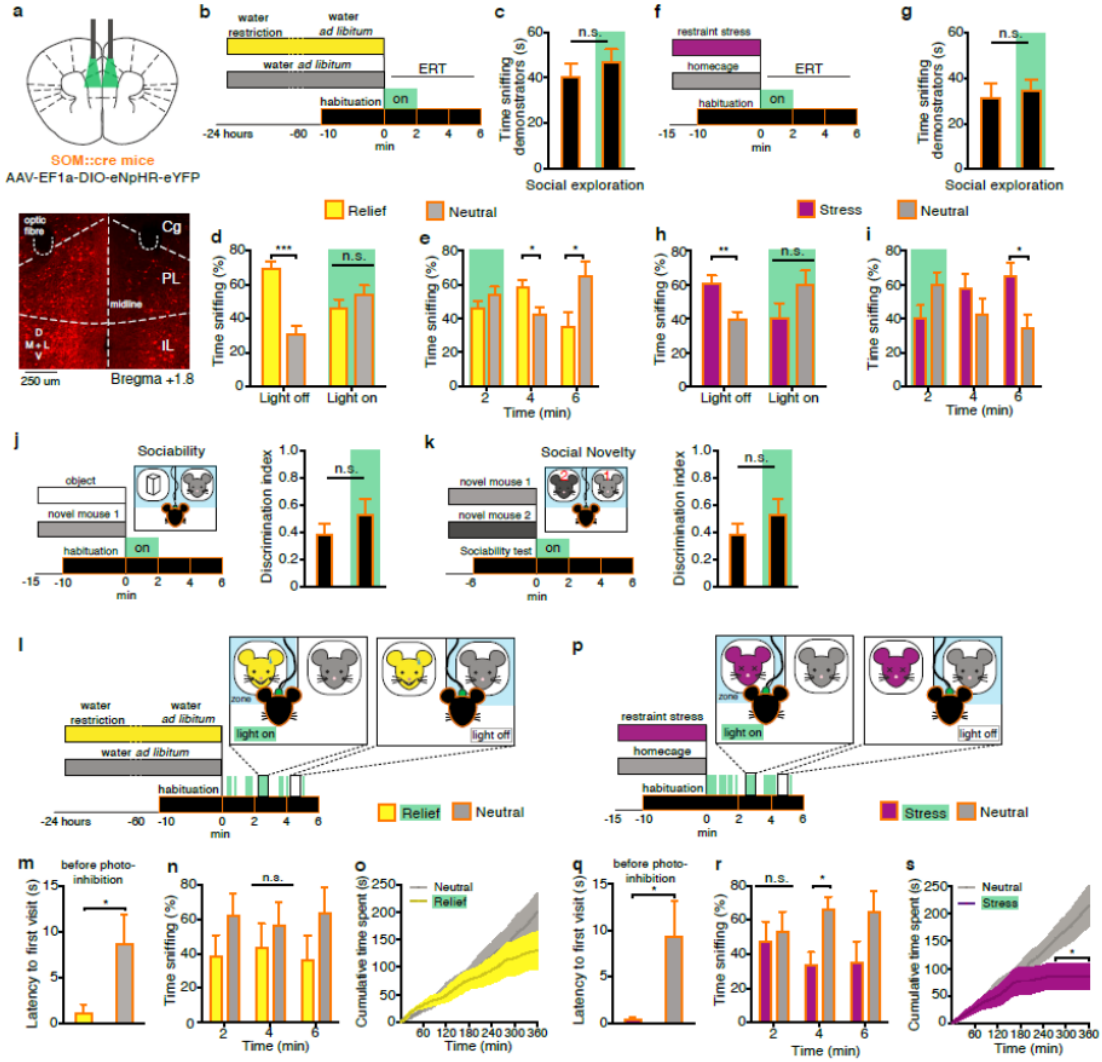
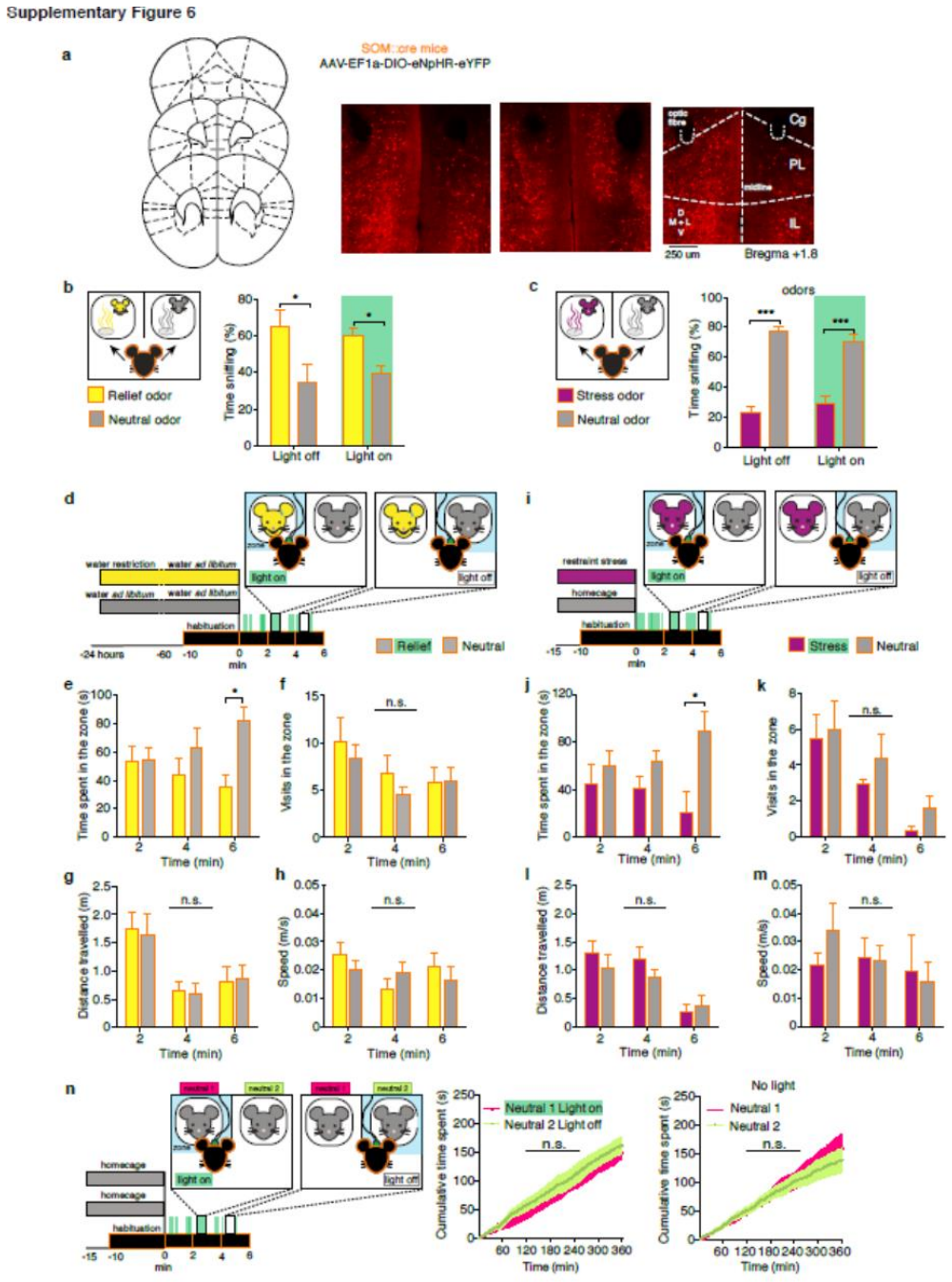


Figure 4. Photo-inhibition of mPFC SOM+ interneurons abolished emotion discrimination. (a) Top, SOM::cre mice were injected in the mPFC with AAV-EF1a-DIO-eNpHR-eYFP and implanted bilaterally with optic fibers terminating dorsal to the injection area. Bottom, representative image of coronal mPFC section. (b) Mice were tested in the ERT 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) Total exploration towards the demonstrators was not affected by SOM+ photo-inhibition. (d) Increased exploration to the relieved demonstrator compared to the neutral during the first 2 minutes of testing (Multiple t-test, Bonferroni correction, 2 min: t=6.63, df=12, p<0.0005; n=7 mice), which was abolished in the light on condition (t=1.26, df=12, p=0.22). (e) Immediately after photo-inhibition of SOM+ mice explored more the relieved demonstrator (Multiple t-test, Bonferroni correction, 2 min: t=2.34, df=12, p<0.05; n=7 mice). (f) Mice were tested in the ERT with one stress and one neutral demonstrators and photo-inhibition was performed for the first 2 minutes. (g) Total exploration towards the demonstrators was not affected by SOM+ photo-inhibition. (h) Increased sniffing to the stressed demonstrator compared to the neutral during the first 2 minutes of testing (Multiple t-test, Bonferroni correction, 2 min: t=3.97, df=12, p<0.005; n=7 mice), but this effect was abolished by photo-inhibition of SOM+ (t=1.76, df=12, p=0.10). (i) Following cessation of photo-inhibition, as in the light off condition mice significantly explored more the stress demonstrators, in the last 2 minutes of testing (Multiple t-test, Bonferroni correction, 2 min: t=2.78, df=12, p<0.05; n=7 mice). (j and k) SOM+ photo-inhibition did not modify either sociability (preference to spend more time with a novel mouse than with novel object) or social novelty (preference to spend more time with a novel than with a familiar mouse). (l) Exploration of the relieved demonstrator was paired to SOM+ photo-inhibition throughout the test. (m) Mice firstly explored the relief demonstrators (Unpaired t-test: t=2.16, df=8, p<0.05; n=5 mice). (n and o) Photo-inhibition of SOM+ paired to exploration of the relieved demonstrator abolished emotion discrimination, as sniffing of relieved and neutral mice (n) and time spent in the zones (o) were not significantly different. (p) Exploration of stressed demonstrator was paired to SOM+ photo-

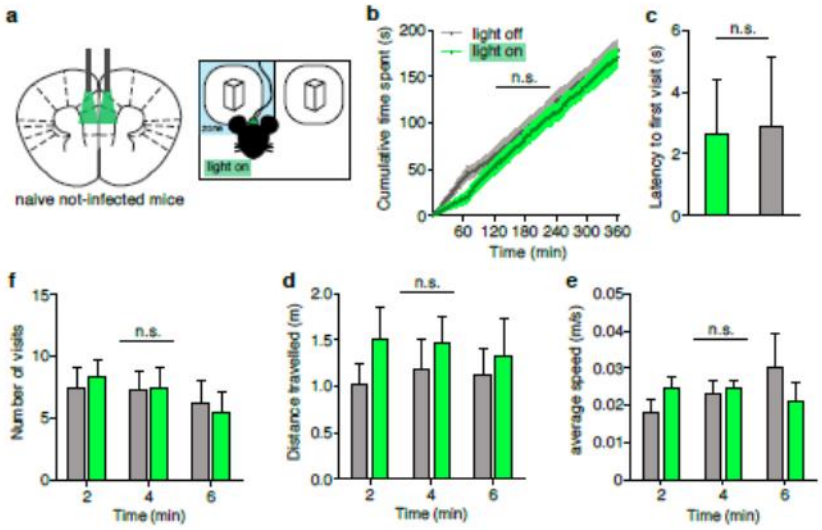
inhibition throughout the test. (q) Mice firstly explored the stressed demonstrators (Unpaired t-test: t=2.32, df=8, p<0.05; n=5 mice). (r and s) Photo-inhibition of SOM+ paired to exploration of the stressed demonstrator abolished emotion discrimination (Multiple t-test, Bonferroni correction, 2 min: t=0.34, df=10, p=0.74; n=6 mice), as sniffing of the demonstrators (r) and time spent in the zones (s) were not significantly different. However, photo-inhibition further modified exploration of observer over time, which was increased toward the neutral (Multiple t-test, Bonferroni correction, sniffing: t=3.10, df=10, p<0.05, n=6 mice; cumulative time spent: two-way ANOVA RM, time x group (neutral, stress): F(359,1436)=5.76, p<0.0005, n=5).



Supplementary Figure 6. Photoinhibition of SOM+ interneurons abolish emotion discrimination. Related to Figure 4. **(a)** *Top,* representative images of viral expression in the mPFC (in rostro-caudal order) after injection with AAV-EF1a-DIO-eNpHR-eYFP. *Bottom*, reconstruction of viral expression and location of optical fibers. Red areas represent the expression (higher expression = darker color) of AAV-EF1a-DIO-eNpHR-eYFP in SOM::cre mice. **(b)**

Increased exploration toward the odor of the relieved demonstrators compared to the neutral (Multiple t-test, Bonferroni correction, off: t=3.00, df=29, p<0.05), and no effects of SOM+ photo-inhibition. **(c)** Avoidance of the odor of the stressed demonstrators (Multiple t-test, Bonferroni correction, off: t=9.89, df=20, p<0.0005), which SOM+ photo-inhibition did not change. **(d and e)** SOM+ photoinhibition with continuous green light for two minutes did not induce any gross motor change in both relief and stress ERT. **(f)** Exploration of the relieved demonstrator was paired to SOM+ photo-inhibition throughout the test (n=5 mice). **(g)** No preference to spend more time with the relieved demonstrator during photo-inhibition of SOM+, on the first two minutes of ERT, and increased time spent with the neutral demonstrator in the last two minutes (Multiple t-test, Bonferroni correction, 6 min: t=3.56, df=8, p<0.05; n=5 mice). **(h)** No change of number of visits to each demonstrator during photoinhibition of SOM+. **(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+. **(l)** No difference of number of visits to stressed and neutral demonstator during inhibition of SOM+. **(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 ERT (counterbalanced, left or right, across observers, continuous green light). and did not induce social discrimination or avoidance. No discrimination of the two neutral demonstrators without light stimulation ("No light").

Supplementary Figure 7



Supplementary Figure 7. Green light do not induce place avoidance. Related to Figure 4. (a) Naïve C57BL/6J mice were implanted bilaterally with fiberoptic implants and tested in the ERT setting with two objects, for 6 minutes. Exploration of one object was paired to continuous green light delivery. (b) Illumination of mPFC with green light did not induce place avoidance as mice spent similar time with both objects. (c) Illumination of mPFC with green light did not modify latency to make the first visit (d) number of visits to each zone and (d and e) did not induce gross motor deficits.

Photo-stimulation of SOM+ interneurons in the mPFC guides social discrimination

We next asked whether, conversely, stimulation of these SOM+ interneurons could be sufficient to induce a discrimination between conspecifics not expressing any altered emotion. To test this, we injected a cre-dependent channel-rhodopsin-2 vector (AAV-EF1a-DIO-ChR2-eYFP) into the mPFC of SOM::cre transgenic mice and implanted chronic optic fibers terminating dorsal to the injection site (Fig. 5a). We tested these observer mice in the ERT while presenting two neutral naïve demonstrators. Exploration of one of the two neutral

demonstrators was paired to photo-stimulation of mPFC SOM+ cells in the observer. This protocol induced an increase of exploration towards the demonstrator paired with light delivery, measured as increased sniffing (Fig. 5c) and time spent in the related zone (Fig. 5d), without affecting the number of visits in each zone (Fig. 5e). Photo-stimulation of SOM+ also reduced the distance between the observer head and the explored demonstrator, indicating that SOM+ activation induced a closer approach (Fig. 5f). As expected, when tested with two neutral demonstrators without photo-stimulation, these same observers did not show any discrimination (Fig. 5g,h). Moreover, to check whether the formation of a discrimination was specifically related to social cues, we repeated the same experiment, but now using two identical objects instead of the demonstrator mice. The activation of mPFC SOM+ cells did not produce any object discrimination (Fig. 5i). These findings indicate that the stimulation of SOM+ interneurons in the mPFC is sufficient to induce a social discrimination. Altogether, our results demonstrate that within the mPFC the SOM+ sub-population of interneurons are a key modulator of emotion discrimination.

Figure 5

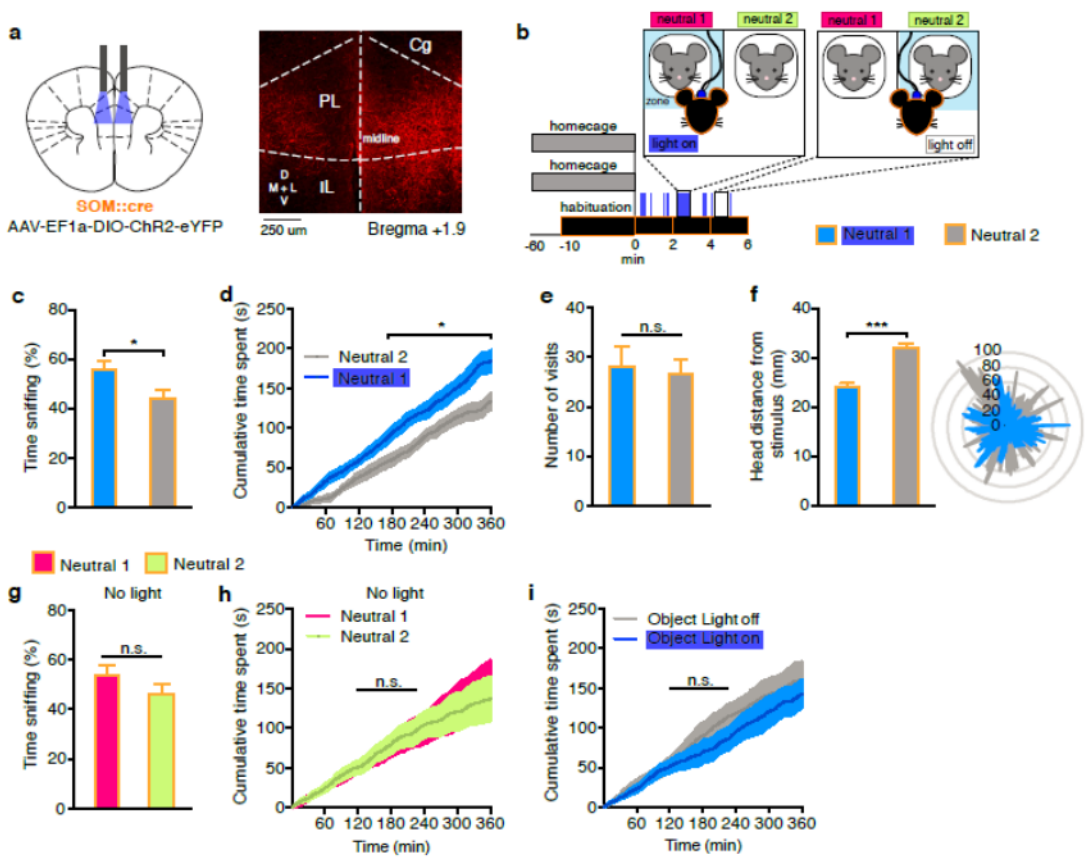


Figure 5. SOM+ interneurons in the mPFC are sufficient to induce social discrimination. (a) Left, SOM-cre transgenic mice were injected in the mPFC with AAV-EF1a-DIO-ChR2-eYFP, and implanted bilaterally with chronic optic fibers terminating dorsal to the injection area. Right, representative image of coronal mPFC section. (b) Mice were tested in the ERT with two, naïve non-manipulated, neutral demonstrators. Photo-stimulation was paired to

exploration of one of the two demonstrators (counterbalanced, left of right, across observers) using 5s-pulses of continuous blue light. (c) Increased sniffing (unpaired t-test: t=3.03, df=14, p<0.05; n=7 mice) and (d) time spent in the zone (two-way ANOVA, time x treatment: F(359,2513)=301.5, p<0.0005) of the demonstrator paired to the photo-stimulation compared to the unpaired demonstrator. (e) No difference in the number of visits in the zone paired with the light compared to the unpaired. (f) Left, head distance of the observers was shorter to demonstrators paired with light (two-way ANOVA, time x distance: F(359,3600)=1.43, p<0.0005). Right, schematic of second-by-second (over 360s) head distance from demonstrators (gray, neutral-unpaired; blue, neutral-light paired). (g) Mice tested with two neutral demonstrators without any light stimulation showed no difference of sniffing and (h) time spent in each zone. (i) Photo-stimulation coupled to exploration of an object did not induce object discrimination.

Discussion

In this study, we revealed a specific association between emotion discrimination and mPFC neuronal activity using a behavioral paradigm that allows a quantitative assessment of the ability of mice to discriminate either positive- or negative-emotional valence in conspecifics. Next, using bidirectional optogenetic manipulations, we demonstrated that SOM+ interneurons but not PV+ interneurons in the mPFC are crucial for emotion recognition. Increased exploratory behavior towards a mouse, manipulated to express an altered affective state, was not observable in a classical one-on-one social interaction test. Moreover, the optogenetic manipulations of the mPFC revealed double dissociation of the roles of interneurons in social functions. Manipulation of SOM+ neurons altered the discrimination of an emotionally altered mouse but did not alter social preference more generally, while optogenetic inhibition of mPFC PV+ neurons altered general sociability without affecting discrimination of emotions. Overall, these findings demonstrate that our paradigm selectively measures discrimination of emotions in a distinct way from general sociability and social novelty.

Rodents actively avoid intense aversive stimuli[151,152], including aversive USV calls induced by heavy distress[153,154] and odors emitted by a shocked, heavily stressed, defeated, or sick conspecific[155–157]. However, in our setting, USVs were not involved, and the discriminations were qualitatively different in the presence of the demonstrators or only of their odor cues. Neuronal recording during odor discrimination did not recapitulate the same activation pattern evoked by the presence of the relieved and stressed demonstrators. Moreover, familiarity enhanced the exploration towards the emotionally altered demonstrators, which is similarly found in humans during ERTs[158]. These results suggest that avoidance of aversive stimuli or seeking extrinsic rewards were not the sole reason for the discrimination of emotions that we observed. Thus, the discriminatory behavior we measured in this study does not reflect a simple reaction to “alarm” or “attractive” sensory cues but rather stems from specific recognition of expression of emotional states in others.

Discrimination between neutral and emotionally altered demonstrators was fast, short-lasting and did not correlate with observable behaviors of the demonstrators. Indeed, the

demonstrators were not under any physical distress during the discrimination but rather carried over an altered emotional state from manipulations performed before the test. No signs of transfer of behavioral responses between observer and demonstrator, no escape behaviors or altered corticosterone levels in the observer mice were present throughout the test session. Moreover, emotion discrimination was equally evident in the absence of visual cues. All these features differentiate our paradigm from previous settings designed to study emotional contagion and vicarious learning[151,159–162]. In particular, social observational learning such as the transfer of fear or pain requires the direct visual observation of a demonstrator during physical challenge for a longer period[160,163]. Furthermore, in contrast to previous paradigms specifically designed to address helping behaviors, such as liberating a trapped conspecific[164], or consolatory behavior, such as allogrooming a distressed conspecific[159], our task allows the observer to engage in sniffing exploration only. Altogether, these findings indicate that our task can measure social cognitive abilities distinct from those addressed by previous paradigms, tackling features closer to those measured by human ERTs[165,166] and complementing current available tools to address the mechanisms underlying higher-order social processes.

Our findings show that the activity of mPFC neurons increases during the exploration of conspecifics manipulated to induce an altered emotional state. The increased mPFC firing when approaching a mouse is consistent with previous evidence showing increased social-dependent firing that did not occur towards an inanimate object or an empty chamber[167]. However, in contrast with the habituation pattern evident in a normal social approach[167], the increased firing towards an emotionally altered mouse remained sustained for the entire test session. Moreover, in contrast to our emotion discrimination data, increasing mPFC excitation but not inhibition can reduce sociability and social interaction in a classical 3-chamber test[168]. These findings further support the distinct social function assessed by the ERT and suggest that the timing and nature of mPFC firing might differentially code for separate social cognitive functions. SOM+ interneurons in the mPFC play an essential role in the process of discriminating a mouse manipulated to induce an altered emotional state. Indeed (1) temporally specific inhibition of these cells selectively abolished discrimination of emotions but not sociability or social odor discrimination, (2) activation of SOM+ neurons was sufficient to induce social discrimination, and (3) these effects were not evident towards inanimate objects. SOM+ interneurons can disinhibit pyramidal cells by inhibiting PV-expressing interneurons[169]. Thus, SOM+-dependent effects on emotion recognition could ultimately rely upon their inhibition of PV+ cell activity. However, SOM+ are low threshold spiking interneurons and have electrophysiological properties[170] that could allow them to be activated more readily than PV+ in the mPFC by excitatory inputs. In support of the latter scenario, we did not find any effect

on emotion recognition following the inhibition of PV+ cells. In agreement with this finding, activation but not inhibition of PV+ cells could rescue social deficits in an animal model of autism[171]. Together, these findings show a previously unreported role for SOM+ cells in specific aspects of social cognition.

We revealed a previously unexpected differential role of SOM+ and PV+ interneurons with respect to social investigation and discrimination of emotions. These interneurons differ in their physiological properties as well as in their connectivity with principal neurons[172]. Indeed, PV+ neurons mainly provide perisomatic inhibition to pyramidal neurons[173], while SOM+ cells preferentially target distal dendritic branches[173,174]. Thus, we could assume that PV+ provide strong network inhibition, while SOM+ could contribute to the correct integration of information flow from other brain structures as also recently suggested for working memory functions through interactions with the ventral hippocampus[175]. Another potential source of socially related information is the basolateral region of the amygdala (BLA), which is reciprocally connected with the mPFC. Excitatory synaptic inputs from the BLA onto principal and PV+ neurons in the mPFC display marked short-term depression of principal neurons while providing strongly facilitating inputs to SOM+ cells in the mPFC[170]. The amygdala is one of the brain regions most consistently associated in the processing of both negative and positive emotional discrimination[176,177]. Furthermore, interestingly, within the PFC, SOM+ neurons are highly enriched in oxytocin receptors[178], and the oxytocin system has been strongly implicated in social functions[179,180] and particularly in emotion recognition[181–183]. Thus, information on the emotional valence from subcortical structures could be integrated by SOM+ neurons within the mPFC local network for implementation in cognitive processes required for discrimination of emotions.

Using a reversed translation of a human task to a mouse model, we uncovered a selective and pivotal role for mPFC SOM+ interneurons in the ability to discriminate the expression of emotions in others. This mechanism and the behavioral assessment employed create new opportunities for the study of potential treatment for the still incurable social cognitive dysfunctions evident in a number of psychiatric disorders such as autism and schizophrenia. Indeed, our findings hold both fundamental and clinical relevance, supporting the implication of a cortical excitatory and inhibitory imbalance in core behavioral dysfunctions in these disorders and providing a convergent and selective target to manipulate emotion recognition abilities.

Material and Methods

Mice. All procedures were approved by the Italian Ministry of Health (permits n. 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 animals' maintenance was provided by dedicated and trained personnel. Three to 6-month-old males and females C57BL/6J animals 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 transgenic mice, id #013044, RRID:IMSR_JAX:013044 (called SOM-cre line) were purchased from the Jackson Laboratory (Bar Harbor, USA) and then breed and expanded in our animal facility for successive testing. Mouse genotypes were identified by PCR analysis of tail 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-hour light/dark cycle (7pm/7am schedule). Experiments were run during the light phase (within 10am-5pm). All mice were handled on alternate days during the week preceding the first behavioral testing. Behavioral scoring was performed a posteriori from videos by two-three independent trained experimenters (inter-rater reliability r score > 0.90), blind to the manipulations and genotypes. 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. The results reported in this work were performed and independently replicated in more than XY different batches of mice coming from at least 4 different generations.

Emotion Recognition Task (ERT)

Testing mice (“observers”) were habituated inside a standard mouse cage (Tecniplast, 35.5x23.5x19 cm) equipped with a dark separator in the middle of the 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 during the test. 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. 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 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 standardly performed in our animal facility. Habituation to the testing setting occurred on three consecutive days before the first experiment; each habituation session lasted 10 minutes. Demonstrator mice – matched by age, sex and strain to the observers – were habituated, without observer, inside the same cage under the wire cups. During both habituation and behavioral testing, the cages were placed inside a dimly lit (6±1 lux) soundproof cubicle (Med Associates).

Observers. Before the test, mice were habituated to the experimental setting as reported above. Ten minutes before the experiment, observer mice were gently moved in the dimly lit testing cubicles. For the optogenetics and in vivo electrophysiology experiments, observer mice were connected to optic fiber or headstage cable for 10 minutes before testing. Then, one emotionally "neutral" and one "relief" or "stress" demonstrator mice were place under the wire cups, and the 6-minutes 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, in the experimental room one hour before the experiment began. All demonstrators were group-housed, separately from cages of stressed and relieved demonstrators. Demonstrators were test-naïve and used 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. Mice were habituated to the experimental setting as reported above. Relieved demonstrators were water-deprived 23 hours before the experiment. One hour before the test ad libitum access to water was given, and mice were brought inside 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-hour water restoration.

"Stress" demonstrators. Mice were subjected to a mild stress consisting in Restraint tube test, a standard procedure to induce physiological stress in rodents59, for 15 minutes before the beginning of the ERT. Then were 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, Ireland). These videos were used 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 travelled, average speed).

One-on-one social exploration tests

This test was similarly performed as previously described36. One hour prior to behavioral testing, each experimental subject was placed into a plastic cage (Tecniplast, 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 in the testing soundproof cubicles. To begin the test a demonstrator mouse was introduced to the cage for 6 minutes (as for the ERT), and exploratory behaviors initiated by the test subject were timed by two independent experimenters blind to the manipulations. Demonstrators 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 to our setting a widely employed standard test for assaying sociability in mice32. The session started with the observer in the same testing cage used for the ERT for a 6-minute habituation period. After habituation, the observer was presented, for other 6 minutes, 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), which has had no previous contact with the observer, was placed in the wire cup in the other side chamber. To measure sociability, the tendency of the subject mouse to spend time with a conspecific, as 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 mouse1 and novel object). Following sociability test, the object was replaced with a novel mouse (novel mouse 2) and the observer was tested for other 6 minutes to assess the preference for social novelty. This is defined as more time 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 completely unfamiliar 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 mouse1 and novel mouse 2).

Sensory modality assessment

For testing ERT in the darkness, mice were tested as above, but eliminating all sources of light within the testing cage as well as in the 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 cameras setting gave the same experimental results.

For auditory stimuli testing, ultrasonic vocalisations (USVs) were recorded during the test phases performed as above in two different experimental settings: 1) standard setting 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 condition ("relief", or "stress" or "neutral)". This was done to check whether the USVs recorded could be attributed to a single emotional state and/or to a communication between demonstrators and observer. 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 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 analysed for each test day included number of calls and duration of calls. Quantitative analyses of sound frequencies measured in terms of frequency and amplitude at the maximum of the spectrum were not performed because of the paucity of emitted USVs in all conditions performed.

For olfactory stimuli testing, observers were tested as described above, but presenting as "demonstrator" only cotton balls impregnated with the odor of demonstrators. Odors were separately and freshly collected from "neutral", "relief" (after the 1-hour ad libitum access to water) and "stress" (immediately after the restraint tube stress) demonstrators by gently brushing the cotton ball all over the body of the mice (especially including the nose, body and

anogenital parts). Each odor was always taken from one single mouse (which was not reused) and used only once.

For testing ERT in the darkness, mice were tested as above, but eliminating all sources of light within the testing cage as well as in the 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 cameras setting gave the same experimental results.

For auditory stimuli testing, ultrasonic vocalisations (USVs) were recorded during the test phases performed as above in two different experimental settings: 1) standard setting 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 condition ("relief", or "stress" or "neutral)". This was done to check whether the USVs recorded could be attributed to a single emotional state and/or to a communication between demonstrators and observer. 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 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 analysed for each test day included number of calls and duration of calls. Quantitative analyses of sound frequencies measured in terms of frequency and amplitude at the maximum of the spectrum were not performed because of the paucity of emitted USVs in all conditions performed.

For olfactory stimuli testing, observers were tested as described above, but presenting as "demonstrator" only cotton balls impregnated with the odor of demonstrators. Odors were separately and freshly collected from "neutral", "relief" (after the 1-hour ad libitum access to water) and "stress" (immediately after the restraint tube stress) demonstrators by gently brushing the cotton ball all over the body of the mice (especially including the nose, body and

anogenital parts). Each odor was always taken from one single mouse (which was not reused) and used only once.

Conditioned place preference

Mice were tested in a well-established conditioned place preference (CPP) paradigm able to assess either positive or negative affective states in mice60,61. The test 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 (21×21×21 cm) as previously described60. 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, a door (4×4 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 6.0; Stoelting, Ireland) was used. The place conditioning 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 pre-conditioning time values in the two sides of place conditioning apparatus. As for the same manipulation in the ERT, 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 time. Other than the two 23-hr deprivation periods, water was available all time. The “neutral” group was exposed to the same procedure but without any water deprivation. Post-conditioning test was performed on day 5 in the same condition of the pre-conditioning test. For each mouse, a conditioning score was calculated as the post conditioning time minus the preconditioning 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 2500 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

AAV5-EF1a-DIO-eNpHR-eYFP.WPRE.hGH (Addgene 20949, qTiter 1.95e13 GC/ml), AAV5-EF1a-DIO-hChR2(H134R)-eYFP.WPRE.hGH (Addgene 20298P, ddTiter 2.76e13 GC/ml) and AAV1-CamKIIa-eNpHR3.0-eYFP.WPRE.hGH (addgene 26971P, qtiter 5.12e12 GC/ml) were purchased from the University of Pennsylvania Viral Vector Core.

Stereotaxic surgery, viral injections and tetrodes implants

C57BL/6J, SOM-cre and PV-cre transgenic mice were naïve and 2 to 3-months old at the time of surgery. All mice were anesthetized with 2% isoflurane in O2 by inhalation and mounted into a stereotaxic frame (Kopf) linked to a digital reader. Mice were maintained on 1.5 - 2% isoflurane during the surgery. Brain coordinates of viral injection in the mPFC was chosen in accordance to the mouse brain atlas (Paxinos and Watson, 1998): AP: +1.9 mm; ML: □ 0.30 mm; DV: -2.5 mm. Volume of AAV injection was 0.4 □L per hemisphere. We infused virus through a glass micropipette connected to a 10-µL Hamilton syringe. After infusion, injector was kept in place for 5 min and then slowly withdrawn over 5 min. After virus injection mice were allowed 3 weeks to recover and for the viral transgenes to adequately express before undergoing optic fibre implantation and behavioral experiments.

Mice underwent stereotaxic surgery for fiberoptic implantation and for recording tetrodes implants

The skull was exposed and two holes were drilled to target the mPFC in accordance to the mouse brain atlas (Paxinos and Watson, 1998): AP: +1.9 mm; L: □ 0.30 mm; DV: -2.5 mm. For fiber optic implantation, bilateral fiberoptic cannula (200 µm, 0.37 NA; Doric Lenses) were lowered 2 mm from skull, to be 500 µm dorsal to the virus injection site, and secured to the skull with MetaBond and dental cement. For in vivo electrophysiological recording, mice were implanted with silicon probes carrying four tetrodes in the right mPFC (Neuronexus A4x4-3mm-100-125-177-Z16). Prior to 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 7 to 10 days depending on the general health.

Optogenetic manipulations

During behavioral testing, fiberoptic cannulae were connected to patch cords (Doric Lenses), which were in turn connected to blue or green light lasers (CNI laser) using a 1x2 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 fiber optic cable was 4.5 mW. For photo-inhibition experiments we used continuous green light (532 nm, CNI laser). For photo-stimulation experiments we used 5-s pulses of blue light (447 nm, CNI laser). To control optical inhibition or stimulation with a closed-loop system dependent on mice behavior during ERT, a behavioral tracking system (Anymaze 6.0, Stoelting, Ireland) detected online the location of the observer mouse in the testing arena and triggered the laser.

In vivo recordings

Neuronexus silicon probes carrying four tetrodes were implanted in the right mPFC (:+1,8 mm anterior, + 0,2 mm lateral from bregma, and -2,5 mm ventral from the brain surface) under general anesthesia. After 1 week of recovery from surgery, recordings were carried out mainly from the prelimbic cortex by means of a 16 channels Neuralynx Digital X system (NeuroLynx). 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 (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, Ireland). All quantitative analyses of neuronal data were performed offline using dedicated software (Plexon). Both putative pyramidal cells and putative interneurons were included in the analysis.

Analysis

Behavior. Video images were analyzed offline with movie maker. A valid exploration trial was defined as a >1 sec exploration of one of the two demonstrators with an interval from the previous visit >2sec.

Definition of epochs of interest

All the neurons included in the present work were recorded for a variable numbers of trials depending on the number of times the observer decided to explore the relief rather than the neutral conspecific. Based on the timestamps related to the main behavioral events (start and end of the social exploration), we defined four different epochs of interest for statistical analysis

of neuronal responses: (1) pre social interaction epoch, corresponding to 1 sec before the onset of the social interaction; (2) post social interaction epoch, including the 1 sec after the onset of the social interaction; (3) end social interaction epoch, ranging from 1 s before to the end of the social interaction; (4) post social interaction epoch, ranging from 1 s after the end of the social interaction.

Single neurons 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 ANOVAs (with significance criterion of $p < 0.05$):

1) Neural response to the beginning of a social interaction. Neurons' activity during three sessions were analyzed separately by means of identical 2x2 Two-way ANOVAs, with factor Epoch (2 levels: pre social interaction, post-social interaction) and Emotional state (2 levels: relief/stress, neutral) followed by Bonferroni post hoc tests ($p < 0.05$) in case of significant interaction effects as our goal was not only that of identifying possible activity changes induced by the social interaction, but also possible differences when the observer explores the relief/stress demonstrator rather than the neutral mouse. Neurons matched the above mentioned criteria were classified as responsive cells (activated only pre, or only post the social event, or during both the epochs).

2) Neural response to the end of a social interaction. Possible modulation of single neuron activity in correspondence of the end of a social interaction during the three sessions were analyzed separately by means of identical 2x2 Two-way ANOVAs, with factor Epoch (2 levels: pre- end social interaction, post- end social interaction) and Emotional state (2 levels: relief/stress, neutral) followed by Bonferroni post hoc tests ($p < 0.05$). Subsequently, all of these neurons were involved in further statistical analysis.

Population analyses

Population analyses were performed on all of recorded neurons, classified on the basis of the results of the above described analyses, and taking into account single-neuron responses calculated as averaged activity (spk/s) in 20 ms bins across trials of the same condition. The same epochs used for single unit data were used for population analyses as well.

Histology

At the end of the behavioral procedures we checked viral expression and position of the optic fibers. 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-µm-thick coronal sections were cut using a Leica VT1000S 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 plus 4% 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, Invitrogen, CatNo. A-11122). Primary antisera were diluted in PBS plus 2% NGS overnight at 4°C for GFP antibody. 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 Alexa Flour 488 goat anti-rabbit IgG (H+L) (1:1000, Molecular Probes®, CatNo.A11034) in PBS plus 1% NGS, and subsequently washed three times in PBS for 10 min at room temperature. The sections were mounted on slides and coverslipped. 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).

Statistics

Results are expressed as mean±standard error of the mean (s.e.m.) throughout the manuscript. Behaviors of observers towards the two different demonstrator mice was analyzed using a Multiple t-test, followed by Bonferroni correction. Sniffing behavior was calculated in percentage to allow direct comparison between mice strains and different manipulations. 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 Two-way ANOVAs with factor Epoch (2 levels: pre social interaction, post-social interaction) and Emotional state (2 levels: “relief”/”stress”, “neutral”) followed by Bonferroni post hoc tests. The accepted value for significance was $p<0.05$. Statistical analyses were performed using GraphPad Prism 7.

Chapter 4

Oxytocin effects on social behavior are genetically modulated by astrocytic dopamine D2 receptor

Nigro M*., Ferretti V. , Bruni S., Huang H., Losi G., Lia A., Monari S. , Carmignoto G., Papaleo F.*

(in preparation)

Abstract

Reduced sociability represents one of the earlier symptoms associated with several neuropsychiatric disorders. Yet, the brain mechanisms implicated in such reductions are poorly understood and effective treatment are still missing. Oxytocin (OXT) is an endogenous neuropeptide, which has recently received an extraordinary attention as potential treatment for social deficits. However, the physiological mechanisms underlying exogenous oxytocin effects remain unclear. In this study, we found that mice with reduction in dysbindin-1 gene show decreased sociability toward an unfamiliar conspecific associated with altered mPFC firing rates recorded during social interaction. Remarkably, intranasal oxytocin administration restored both behavioral and physiological alterations in dysbindin-1 mice while produced no effects in wild type mice. Relevantly, we identified in the mPFC of dysbindin-1 mice a higher astrocytes activity. We tested the hypothesis that oxytocin beneficial effect in dysbindin-1 mice could be mediated by a specific action on astrocytes. We showed that oxytocin treatment had no effect in wild types, but led to a specific recovery in dysbindin-1 mice astrocytes dysregulation mediated by an action at D2-receptors. Together, our data indicate that oxytocin treatment is able to restore social deficits and altered PFC function in a model of genetic liability. Moreover, we report for the first time a possible role of astrocytes in PFC behavioral control, and as a critical site for oxytocin action.

Introduction

The ability to interact with others has profound impact on individual's life. Relevantly, in a number of neuropsychiatric conditions social deficits are among the earlier deficit to appear and the most difficult to treat.

Oxytocin (OXT) is an highly conserved neuropeptide that has been implicated both in humans[29] and other animals[27,28] in a wide range of social behaviors, including sexual behavior, parturition and maternal care[26]. This neuropeptide, considered a robust facilitator of social abilities[184–186], produced in clinical trials inconsistent effects[33–38], strongly indicating the need to understand the brain mechanisms underlying OXT's effects on social functioning.

The medial prefrontal cortex (mPFC) is a brain structure implicated in the processing of social information, and in the selection of adaptive responses, coordinating cortical and subcortical inputs. Alterations in mPFC functioning and social deficits are among the principal hallmarks of neuropsychiatric disorders as autism and schizophrenia, characterized by a strong genetic component[187,188]. In this study we took advantage of a mutant mouse, presenting reduced expression of the dystrobrevin-binding protein 1 gene (*DTNBP1*), which encode for dysbindin-1 protein. Genetic variations in DTNBP1 have been previously reported to impact human and mice cognitive abilities[64,65,189]. In this study using a combination of electrophysiological and behavioural approaches, we identified in dysbindin-1 mice (Dys+/-) social deficits and altered mPFC excitability during social exploration of an unfamiliar conspecific. We found that both effects were rescued by the intranasal administration of oxytocin, before test.

Astrocytes have been shown to play a key role in the modulation and control of neuronal activity[190,191]. Relevantly, in dysbindin-1 mice we observed an increased astrocytes activity in the mPFC. We hypothesized that oxytocin effects in dysbindin-1 mice could be mediated by an action on astrocytes. Accordingly, we found that oxytocin application in slices was able to reduce the increased astrocyte Ca++ activity observed in dysbindin-1 mice, while producing no effect in wild type astrocytes. Based on the evidence of a role of D2 receptors in regulating astrocytes function[16,192], and of increased D2 receptors in the mPFC of dysbindin-1 mice, we tested the hypothesis that D2 receptors in the astrocytes could mediate OXT effects: AAV-mediated selective down-regulation of astrocytes D2 receptor in mPFC of dysbindin-1 mice abolished the effect of exogenous OXT on mice sociability, indicating a crucial role of astrocytes in the regulation of behavioral oxytocin effects.

Results

Intranasal oxytocin rescues dysbindin-1 reduced sociability

In mice, the exposure to a new conspecific spontaneously leads to social exploration. To assess dysbindin-1 mice social abilities, we exposed a test mouse, habituated for 1 hour to a new cage, to a matched -age and -sex conspecific, and we allowed them to freely interact for for 60 seconds *per* trials (Fig1a). We analyzed social behaviour of the test mouse, which included nose-to-nose, following, ano-genital sniffing, body sniffing. As shown, wild type mice gradually decreased social exploration over trials, indicating increasing familiarity with the new individual (p=0, one-way analysis of variance (ANOVA), Dunnett's pairwise comparison to control; wild type baseline versus dysbindin-1 baseline, S1-S4 $P < 0.05$; over the first 10 interaction; Fig. 1b). The analysis of trials and genotype effects on social interaction by two-way analysis of variance (ANOVA) revealed trials ($F(3, 90) = 29{,}41$ $P < 0{,}0001$), genotype ($F(2, 30) = 14{,}78$, $P < 0{,}0001$), and genotype x trials significant interaction ($F(6, 90) = 3{,}649$, $p=0{,}0027$) indicating reduced levels of dysbindin-1 significantly reduced mice sociability. In addition total social interaction time differences, we also noted that social duration was significantly shorter in dysbindin-1 mice compared to wild type mice in the trial 1 (Turkey's comparison to control; wild type versus dysbindin-1, $P < 0.05$). To investigate the recruitment of mPFC in social interaction, we implanted wild type and dysbindin-1 mice with tetrodes in the mPFC. Before performing the electrophysiological recordings we assessed mice social behavior during the first two trials of the social interaction test. We choose to perform the experiments in these trials as they appeared the most indicative of mice sociability in previous experiments with no implants (Fig 1b). Confirming previous behavioral data, dysbindin-1 mice show significantly reduced social investigation towards the unfamiliar conspecific, compared to wild type mice (one-way ANOVA, Turkey's comparison to control; wild type versus dysbindin-1, $P < 0.05$; Fig 1b').

We then tested the effect of a single dose (60 IU) of intranasal oxytocin, administered 5 minutes before the beginning of the test. Oxytocin treated wild type mice showed a normal pattern of social interaction over trials (Fig 1c), which was not affected by OXT administration. Differently, dysbindin-1 mice presented reduced social exploration when administered with vehicle and significantly increase social interaction when treated with OXT as shown in fig. 1c' (one-way ANOVA, Turkey's comparison test to control; dysbindin-1 vehicle versus dysbindin-1 OXT acute, $P < 0.05$). To verify whether oxytocin induced effects were specific for social abilities,

we measured other behavioral parameters previously described as altered in dysbindin-1 mice such as hyperactivity, and we found no changes due to oxytocin treatment (data no show).

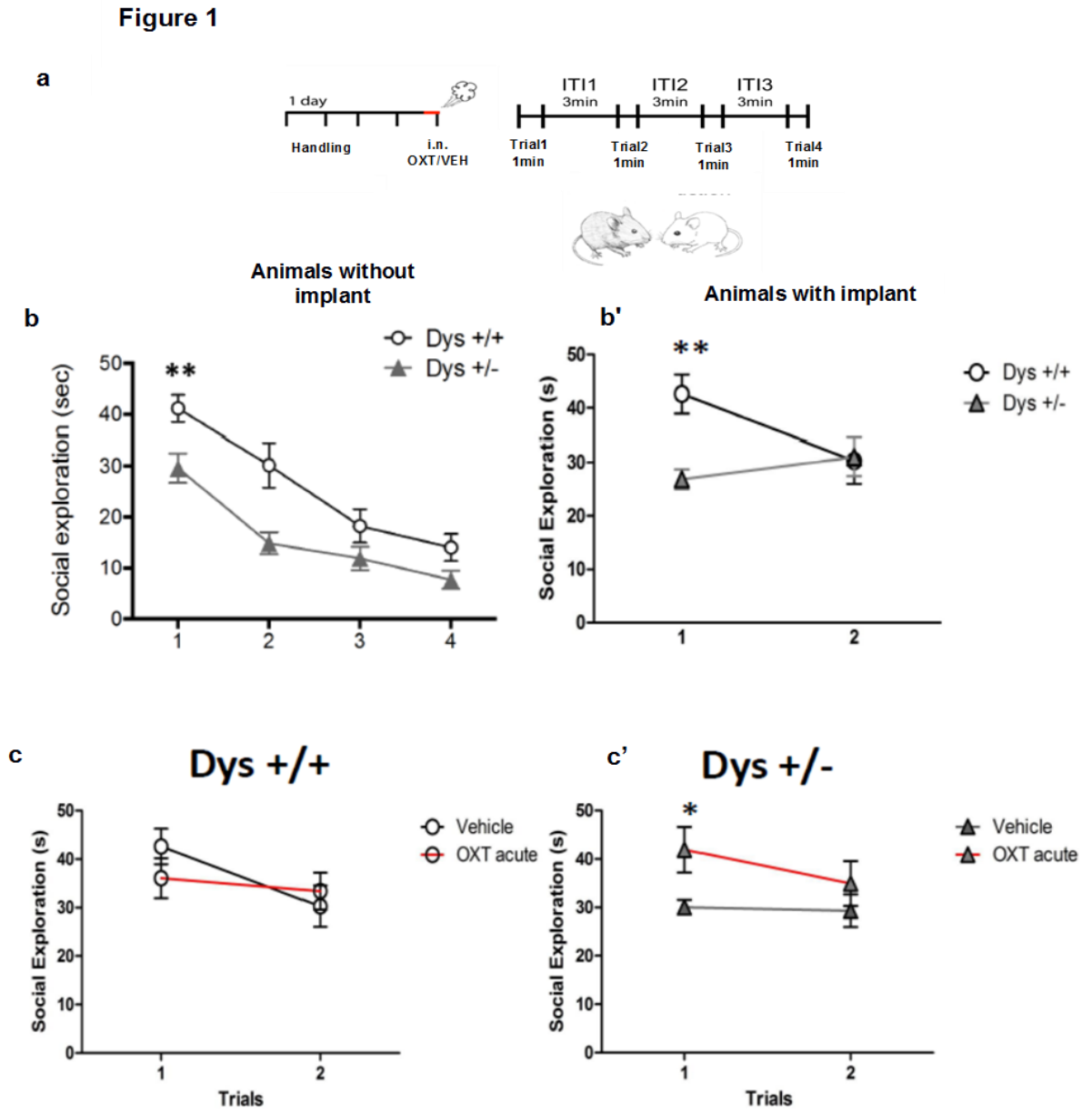


Fig. 1 Acute intranasal OXT treatment increase social interaction between unfamiliar males. (a) Experimental design for the social habituation test. Duration of occurrence of various social behaviors in male mice exposed to an unfamiliar conspecific (including nose-to-body, nose-to-nose, nose-to-back sniffing, going on top of another mouse and following) in male mice towards unfamiliar male. (b) social investigation of the unfamiliar male in mice not implanted and implanted with tetrodes (b'). The first two trials between the two groups; one-way analysis of variance (ANOVA), Dunnett's pairwise comparison to control; wild type (Dys+/+) baseline versus dysbindin-1 (Dys+/-) baseline, $P < 0.05$; (c) single acute dose of saline (VEH in black) or 60 IU/5 µL (in red), 5 min before the test in implanted mice. The wild type did not show any effect after the treatment in both trials; instead, dysbindin-1 mice (c') show higher exploration respect to saline group in trial 1 one-way ANOVA, Turkey's comparison test to control; dysbindin-1 vehicle versus dysbindin-1 OXT acute, $P < 0.05$.

Intranasal oxytocin rescues PFC excitability

We then investigated the effects of intranasal oxytocin on PFC activity during social exploration in wild type and dysbindin-1 mice. Behavioural videos time-locked to PFC signals were collected. mPFC neurons were selected on the base of the quality and stability of signals, and their activity was analyzed in time windows of two second: during the baseline (the period preceding the introduction of the stimulus mouse), before and after social approach (which was defined as the first physical contact of the test mouse sniffing the stimulus mouse) in trial one (T1), during the inter-trial period (when the stimulus mouse was removed from the test cage) and before and after social approach in the second trial (T2; Fig 2a).

In the mPFC of the wild type vehicle group, we recorded 171 well-isolated neurons, mainly in prelimbic cortex, and 74% of them modulate their firing in proximity of a social event. The majority of responsive units (46%) observed increase firing rate preceding social approach (main effect of social trial, mean difference 15,61 ± 0,9644 spk/s Student's unpaired t test $p < 0.001$ compared to baseline), followed by a significant drop in firing rates at the onset of social exploration (main effect of social interaction, mean difference 11,31 ± 1,066 spk/s, Student's unpaired t test $p <0.001$) as shown in the population analysis and representative neuron in Fig 2b. When the stimulus mouse was removed, firing rates returned to baseline levels indicating that mPFC activity was primarily driven by active exploration. From the dysbindin-1 vehicle group we selected in total less than 50% of social-related neurons (N=47), showing higher baseline activity respect to wild type mice (wild type vs dysbindin-1 baseline, mean difference 6,229 ± 0,7151 spk/s, Student's unpaired t test $p <0.001$), suggesting higher mPFC state in basal condition. Similarly to controls, neurons displayed increased frequencies when the stimulus mouse was introduced in the cage (main effect of social trial, mean difference 2,461 ± 0,6315 spk/s Student's unpaired t test $p < 0.001$ compared to baseline; Fig 2b'). However, we found a different pattern of activity in proximity of the social approach, as we did not observe a decrease, but a sustained firing rate after the onset of social behavior (main effect of social interaction, mean difference -1,025 ± 0,5861 spk/s, Student's unpaired t test p=0,0818) in both trials, as shown in the population analysis in Figure 2b'. Consistently with previous evidence on dysbindin-1 PFC excitability[193], we found higher basal firing rate also during the inter-trial compare to controls (Student's unpaired t test $p <0.001$).

Then, we studied the effect of intranasal oxytocin administration on wild-type PFC activity, and we found that the same proportion of neurons were socially-modulated compared to vehicle treated mice (79%). Despite a decrease in frequencies, we found that no effect was produced by OXT in the pattern of activity of wild-type neurons in proximity of social exploration (main

effect of OXT treatment, mean difference -3,751 ± 0,2481 spk/s, Student's unpaired t test p <0.001). Differently, in dysbindin-1 mice, oxytocin administration increased the number of socially-responsive neurons (84%); further, in Trial 1, we found OXT induced a rescue of the altered pattern of activity observed in vehicle group, producing the same pattern observed in wild type mice (main effect of OXT treatment, mean difference -5,449 ± 0,6581 spk/s, Student's unpaired t test p <0.001; Fig 3a').

These results indicate that mPFC activity is strongly associated to mice sociability. Moreover, they suggest that the reduction in firing rates observed at the onset of social exploration could be required for the codifying/processing of social information. Dysbindin-1 genetic reduction might have an impact on this process by altering PFC excitability, which we found to be restored by intranasal oxytocin administration.

Figure 2

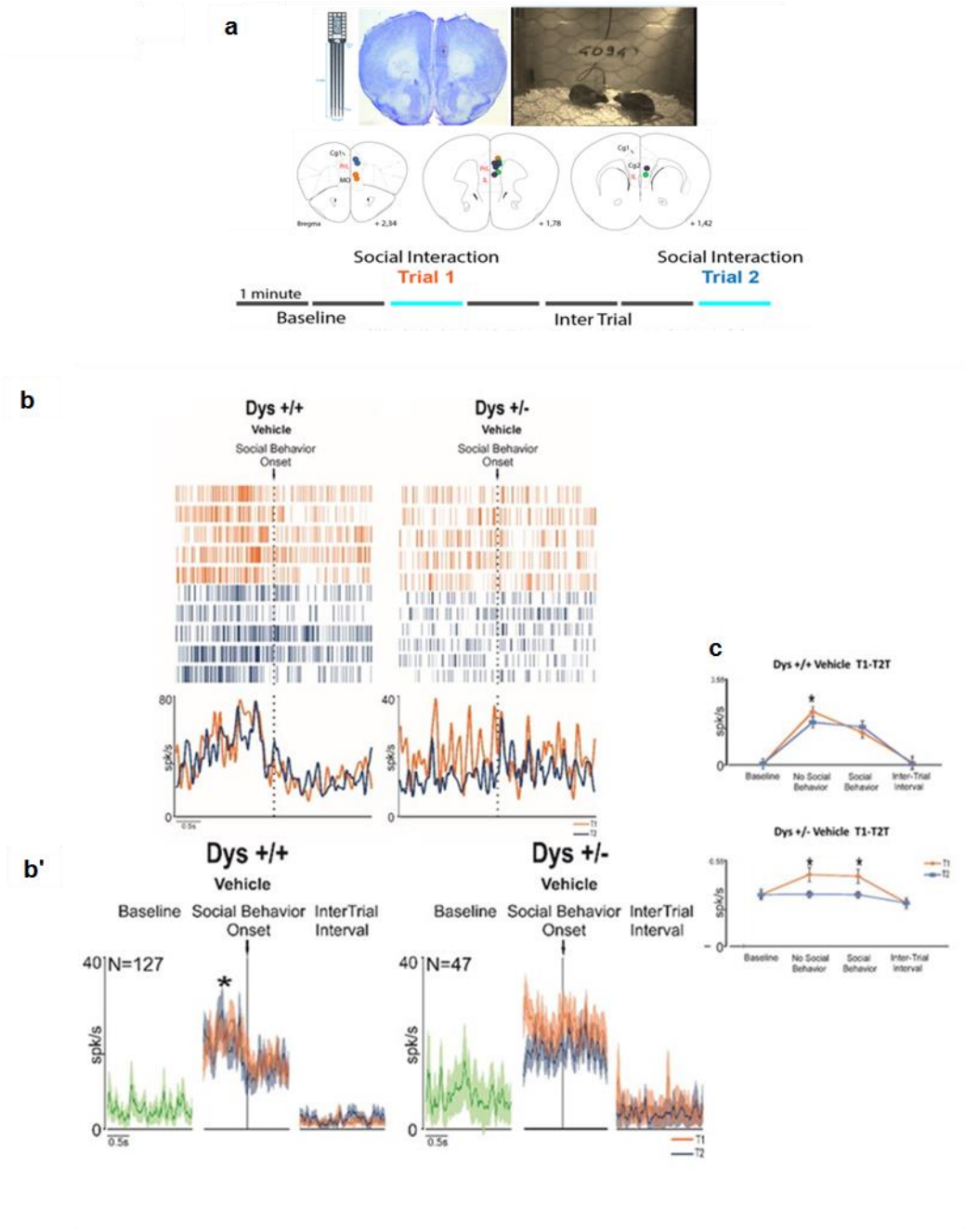


Figure 2. Sustained neuronal activity in dysbindin-1mice during social exploration. (a) top, mice were implanted with electrodes for chronic electrophysiological multi-unit and single-units recordings and tested in the social habituation test. Middle, electrodes placement in the mPFC (Cg, Cingulate; PL, Prelimbic area; IL, Infralimbic area). Bottom, experimental design used for the social habituation test associated to electrophysiology recordings. (b) Dysbindin-1 mice showed increased activity compared to wild type during the baseline (purple) period and show altered modulation during social interaction (Unpaired t-test: t=2.33, df=10, p<0.001; n=6 mice). Rasters and polylines aligned on the beginning (green arrows) of each exploration of unfamiliar conspecific in single trial, which were separated by a variable time interval of the duration of at least of 1 second. Rasters and histograms of single-neuron response when the tested mouse explores stimulus in two trials are shown in different colors as well as IEI (orange and blue respectively). Top, Examples of two responsive cells recorded in response to the exploration of stimulus. Bottom, Population activity of all recorded neurons during the social habituation (n=89) before, during and after test of the stimulus throughout the 1-min experiment. Lines indicate the average discharge intensity of neurons when tested mouse explored the stimulus, trial 1 in orange and trial 2 in blue, aligned as the single neurons example

in (b). Colored shaded regions around each line represent 1 SEM. (c) Dysbindin-1 mice showed sustained activity during no social or social event compared to the WT during the 2 trials of testing (n=7 mice, Unpaired t-test: t=4.12, df=12, p<0.001).

Figure 3

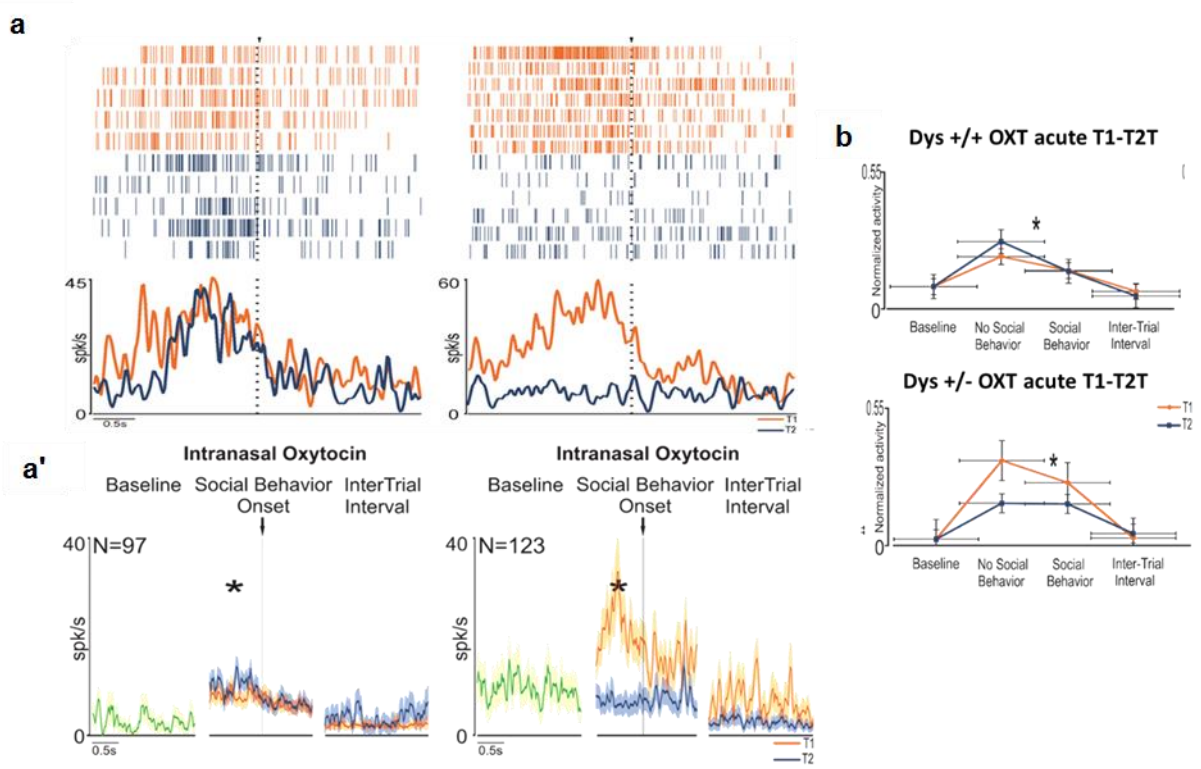


Figure 3. Intranasal OXT treatment induce the same activity pattern elicited in wild type mice. (a) We recorded 97 and 123 units in wild type and dysbinin-1 mice respectively. Bottom, population activity of recorded neurons (n=97) before, during and after trials throughout the experiment representing the average discharge intensity of neurons when tested mice explored unfamiliar conspecific in trial 1 (orange) and trial 2 (blue) aligned on the beginning (black line) of each exploration of the demonstrators in the same session, which were separated by a variable time interval of the duration at least of 1 second. Colored shaded regions around each line represent 1 SEM (t-test, Bonferroni correction, t=7.15, df=10, p<0.001; 6 min: t=3.59, df=10, p<0.001; n=7 mice) in trial 1 and trial 2. Dysbindin-1 showed a similar pattern of activity after a single dose of OXT 60IU/5µL in trial 1 (n=4 mice, Unpaired t-test: t=4.12, df=12, p<0.001).

Dysbindin-1 mice show altered activation of astrocytes

The specific pro-social effect of oxytocin in dysbindin-1 mice, lead us to hypothesize that dysbindin-1 genetics could determine a physiological substrate for a preferential effect of oxytocin on mice sociability. We, then, investigated the different gene expression profiles in the mPFC of wild type and dysbindin-1 mice. Remarkably, the meta-analysis revealed an up regulation of different genes involved in astrogliosis in dysbindin-1 mice compare to control group (Fig. 4).

To confirm these results, and verify possible altered astrocytes activity in dysbindin-1 mice, we performed immunostaining of an astroglial marker, the glial fibrillary acid protein (GFAP) in

mPFC slices and found a significant increase of GFAP intensity in the dysbindin-1 mice compared to wild type (p= 0,0489, Fig. 5). We characterized at functional level this increased reactivity, by measuring astrocytes dynamics in basal condition and after oxytocin application. We imaged Ca2+ influx in PFC slices obtained from wild type and dysbindin-1 mice expressing the encoded Ca2+ indicator GCaMP6 selectively in astrocytes (GCaMP6::GFAP). As expected, in dysbindin-1 mPFC, we observed an increase number of Ca2+ peaks compared to wild type (Fig 6). Notably, a single application of OXT was able to reduce dysbindin-1 Ca2+ elevation in astrocytes, specifically in the microdomain. These results demonstrated that dysbindin-1 reduction is associated with alteration in astrocytes reactivity, which are restored by oxytocin application.

Interesting, the suppression of neuroinflammation has been described to be regulated through D2 receptor on astrocyte. Moreover, variations in D2R signaling have been shown to produce alterations in astrocytes reactivity[16]. Hence, we tested the hypothesis that D2-Rs in the astrocytes could mediate oxytocin effects in dysbindin-1 mice. To selectively target D2-receptors in astrocytes, we injected dysbindin-1 D2 receptors floxed mice with an AAV expressing Cre under the GFAP promoter (GFAP - Cre) in the mPFC of wt and dysbindin-1 D2-floxed mice. Control mice received a bilateral mPFC injection of an AAV expressing GFP under the GFAP promoter. We found that, the rescuing effect produced by intranasal oxytocin in dysbindin-1 mice was abolished when D2 receptors in the astrocytes were decreased, as shown in Figure 7 ($F(1,10) = 16.17$ $P < 0.001$ two-way analysis of variance (ANOVA), Newman-Keuls post hoc, compared to the control group).

Overall, these results indicated that the effect of oxytocin in dysbindin-1 mice sociability could be mediated by a D2 receptors astrocytes action.

Figure 4

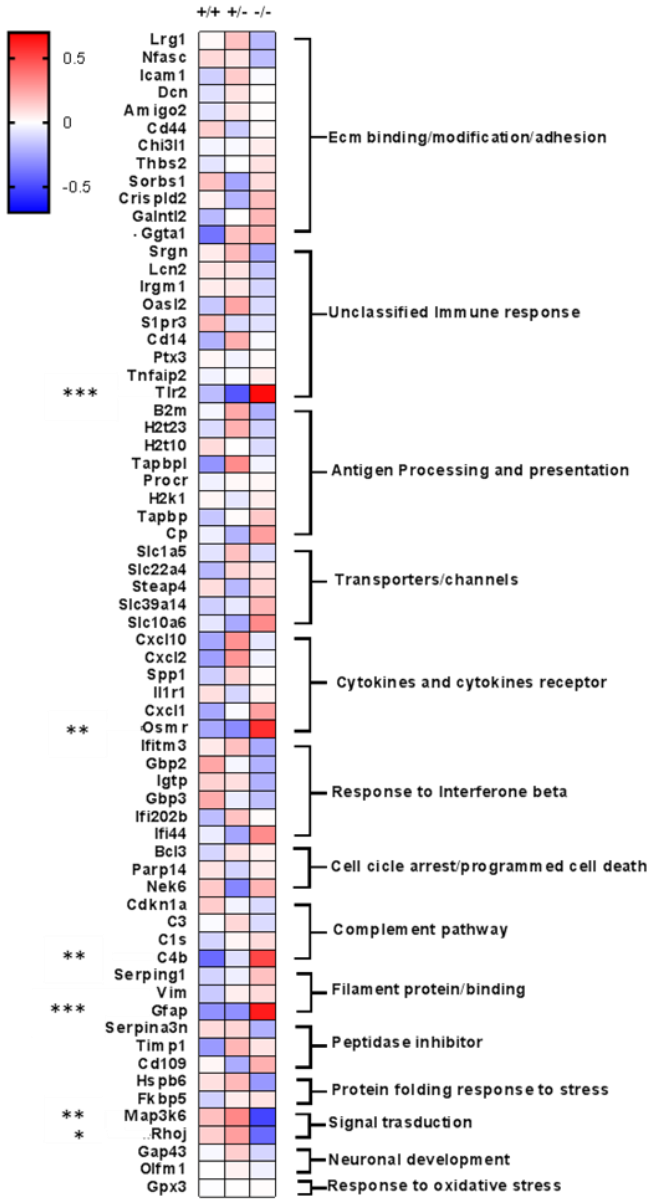


Fig. 4. Heat map of inflammatory markers selected by a microarray screening from the mPFC of wt and dysbindin-1 mice. The heat map is based on hierarchical clustering ($P > 0.001$) of genes involved in inflammation states. All gene expression levels were transformed to scores ranging from −0.5 to 0.5 and were colored blue, white, or red to represent low, moderate, or high expression levels, respectively. The relative expression levels were scaled based on their mean and do not represent expression levels in comparison with controls. Dysbindin-1 heterozygous and knock-out show an higher expression for genes involved inflammatory process compare to the wild-type.

Figure 5

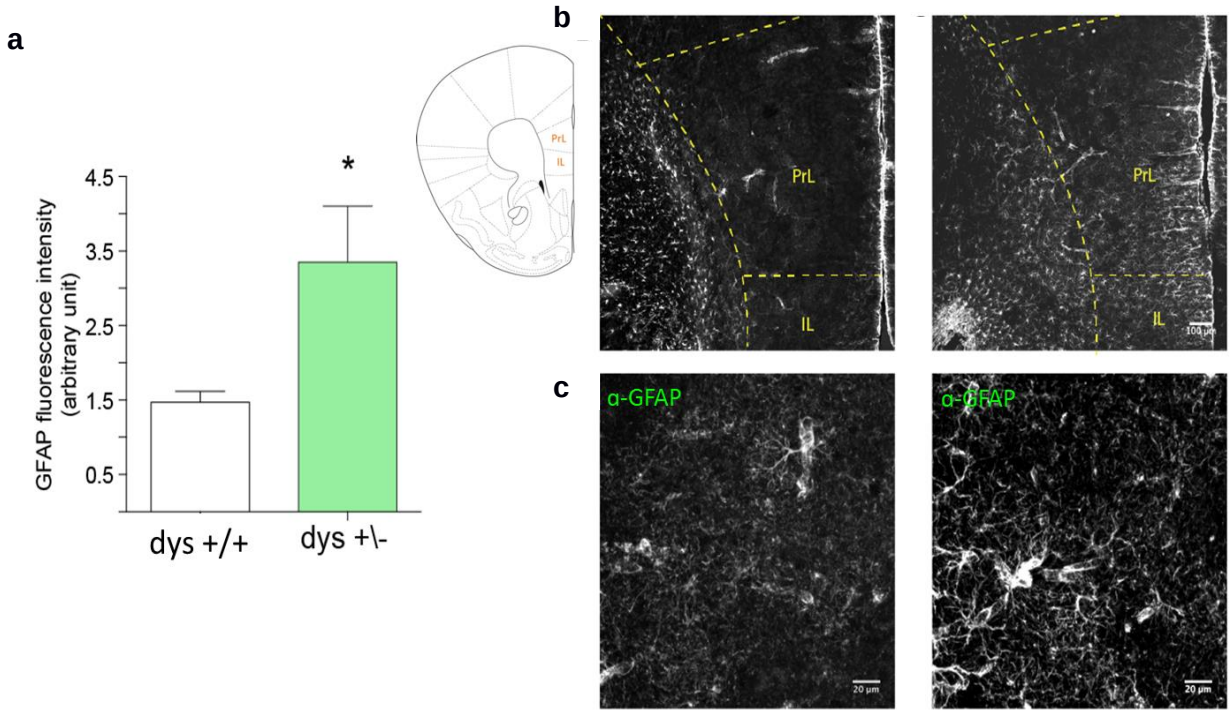


Fig. 5. Dysbindin-1 mice show more GFAP positive astrocytes in mPFC. Dys +/+ (4) and Dys +/- (4) mice was imaged at confocal microscope (Nikon A1) after immunohistochemestry for GFAP (NB300-141 Rabbit Polyclonal GFAP Antibody). (a) Quantification of GFAP signal in wild type (white) and dysbinin-1 (light green). (b) Representative images of mPFC in wild type on the right and dysbindin-1 on the left. Scale bar 100µm. (c) Higher magnification of the focus plan on the top. Scale bar 20µm.

Figure 6

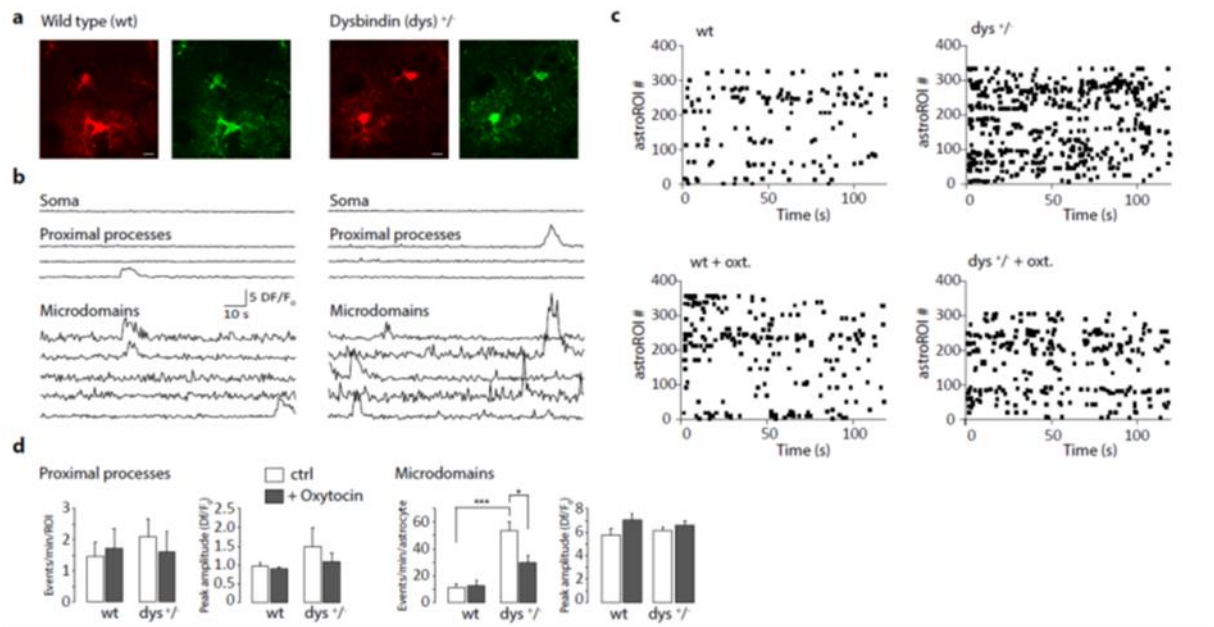


Fig. 6. Higher calcium signal in astrocyte of Dysbindin-1 mice. a. Two-photon images of Tomato (red) and GCamp6 (green) expressing astrocytes in PFC brain slices from wild type and dysbindin-1mice. Scale bars, 10 µm. b. Representative Ca2+ signal traces from the main astrocytic compartments. c. Raster plots of Ca2+ transients from all GCaMP6 astrocyte microdomains in wild type and dysbindin-1 mice, before and after slice perfusion with oxytocin (oxt.; 1µM). d. Mean number of events per minute per ROI or astrocyte and mean Ca2+ transients peak amplitude for wild type and dysbindin-1 astrocytes (12 astrocytes for both wild type and dysbindin-1 mice, 4 animals each, *p<0.05;***p<0.001; Student t test).

Figure 7

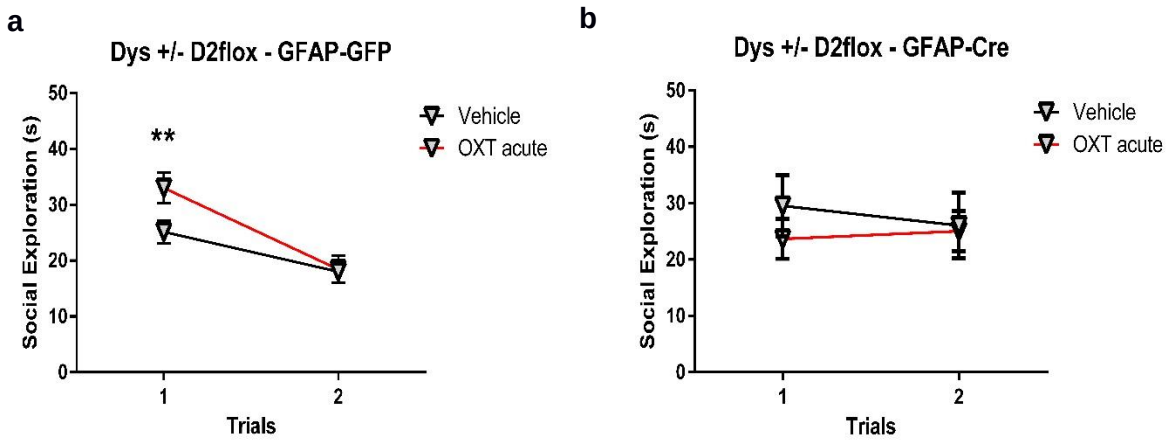


Fig. 7 D2-Rs in astrocytes mediate the positive effect of acute OXT treatment in dysbindin-1 mice. (a) Social habituation test in D2 floxed xdysbindin-1 mice injected with the GFAP-GFP AAV (a) or with the GFAP-Cre AAV(b). The analysis of first two trials in the two groups indicate that the effect of OXT on sociability depends on D2 receptors levels in the astrocytes (two-way analysis of variance (ANOVA), Newman-Keuls post hoc test Dysbindin-1 x D2flox GFAP-GFP vehicle (n=8 black line) vs OXT treatment (n=8 red line), $p < 0.001$; over the first 10 interaction bouts; (b) dysbindin-1 D2flox GFAP-Cre vehicle (n=6 black line) versus dysbindin-1 D2flox GFAP-Cre OXT (n=6 red line), $p = 0.77$).

Discussion

In this study we focused on the effects produced by reduction of dysbindin-1, the gene product encoded by DTNBP1, whose polymorphisms have been considered risk factors for schizophrenia onset[64,65], on mice sociability. In dysbindin-1 mice, in which we previously described alteration of D2 signaling in the mPFC[16,16], we identified reduced sociability and an altered pattern of mPFC excitability during social interaction.

The neuropeptide oxytocin (OXT) is considered a robust facilitator of social abilities[26,185,186]. This notion is supported by observations both in humans and mice of beneficial effects of intranasal oxytocin (OXT) on processing of social information[29,29,30]. However, the physiological mechanisms underlying oxytocin effects are still unclear.

Combining the social habituation test with in vivo electrophysiology, we revealed in dysbindin-1 mice a reduction of social exploration toward unfamiliar conspecific, as well as an alteration in mPFC excitability; Increased cortical activity is consistent with a number of findings in which alteration in mPFC functioning is associated with neurological or psychiatric disorders[12–14]. Earlier studies have found an increase of activity in mPFC during exploration of unfamiliar conspecific in the three-chamber test, in which mice are physically separated by wire cup from

the stimuli mice or objects[194]. This differs from our setting in which the animals are physically in contact. Thus, our results might suggest that the brain activity mediating direct contact during social interaction, or exploration while the subject is confined, might imply different social processes and different information coding modalities. In this study, we show that the mPFC in wild type is activated before the onset of social exploration and inhibited during the active exploration. Differently, dysbindin-1 mice mPFC show a higher, persistent activation, not modulated by social exploration, which correlates with a lower sociability (Fig.1b-2b).

Remarkably, we found that intranasal OXT treatment produced no effect in wild type mice, but was able to rescue social behavior and mPFC alterations in mutant mice (Fig.1c-3a) inducing a pattern of activation similar to the one observed in wild types controls. A number of studies have shown the beneficial effect of OXT in response of social situation[29,30]. Relevantly, Galbusera et al. recently showed that OXT administered via intranasal route, elicits a transient activation of cortical regions[195] supporting our finding of a specific recruitment of the mPFC after intranasal OXT administration. In addition, OXT has been reported to induce the top-down recruitment of GABAergic interneurons in the cortex [119] and to increase signal-to-noise ratios in the mPFC, improving the temporal precision and fidelity of information transfer, elevating inhibitory tone[120,121]. Our results suggest a similar action could explain the effect of OXT to recapitulate in dysbindin-1 mice mPFC a firing rate pattern similar to the one observed in wild type mice.

Our findings also indicated that the alterations found in the E/I balance of mPFC of dysbindin-1, might involve alterations in astrocytes function[190,191]. Several studies reported that glia cells contribute and regulate PFC homeostasis[196–198]. Using gene profiling, astrocytes immunoreactivity and Ca^{2+} imaging, we showed that dysbindin-1 mice, together with abnormal neuronal activity and behavioural phenotype, present increased astrocytes activity (Fig.5-6). In particular, we found higher gene expression of proinflammatory cytokines and GFAP expression in mPFC of dysbindin-1 mice (Fig.4). This result was also confirmed by our immunohistochemistry study showing increased GFAP expression in dysbindin-1 compared to controls. Interestingly, application of OXT to mPFC slices, decreased dysbindin-1 astrocytes reactivity leaving unaltered wild type levels. (Fig.6). Notably, increased of neuroinflammation has been implicated in a number of psychiatric disorders (e.g., mood and anxiety disorders) raising the possibility that inflammatory mechanisms are critical for linking psychosocial factors and related disorders to physical health[199].

Recently, suppression of neuroinflammation state has been described to be regulated through D2 receptors located on astrocytes, via an intracellular pathway that negatively regulate

inflammatory path [16,192]. Relevantly for our study, recent findings also reported that activation of the oxytocin receptor can increase the D2 signaling in the D2R-OXTR heteroreceptor complex, via a facilitatory allosteric receptor-receptor interaction[200,201] suggesting D2 receptors in astrocytes could mediate oxytocin responses observed in astrocytes. This hypothesis seems to be supported by our results showing that silencing astrocytic D2 receptors in the mPFC of disbindin-1 mice, we abolished the beneficial effects produced by intranasal OXT administration (Fig.7). This result could indicate D2 receptors in mPFC astrocytes as a possible mechanism of action of OXT induced behavioral effects.

Materials and Methods

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. Males dysbindin-1 heterozygous (Dys+/-) and their wild-type littermates (Dys+/+), all of 3-6 month-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 (within 2-5pm). 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. Behavioral scoring was performed a posteriori from videos by trained experimenters, blind to the manipulations of both the observers and demonstrators. A sniffing event was considered when the test mouse touched with the nose directly the unfamiliar stimulus.

Habituation social interaction test

Naïve mice were tested as similarly reported previously[202,203] in 2150E Tecniplast cages (35.5 × 23.5 × 19 cm). Male mice were individually placed in the testing cage 1 h prior to the testing. No previous singlehousing manipulation was adopted to avoid any instauration of home-cage territory and aggressive behaviors. Testing began 5 min after the intranasal treatment when a stimulus male mouse was introduced into the testing cage for a 1-min interaction. At the end of the 1-min trial, we removed the stimulus animal and returned it to an individual holding cage. We repeated this sequence for four trials with 3-min inter-trial intervals introducing the same stimulus mouse in all three trials, instead for the in vivo electrophysiology we recorded and performed only two trials with the same inter-trial-interval. Videos of behaviors were recorded and subsequently scored offline. After each testing day, testing cages were autoclaved as standardly performed in our animal facility. 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) were placed facing the long side of the cage and on top of the cage to record from different angles the three consecutive two-minute trials, using the Anymaze program (Stoelting, Ireland).

Intranasal OXT Administration

OXT (Novartis Pharma AG, Switzerland) was dissolved in saline (0.9% NaCl) and administered intranasally in a volume of 5 µl to each mouse in doses of 0.3 IU/5 µl (OXT 0.3 IU). An amount of 1 IU of our solution contained 1.667 µg of synthetic OXT. Thus, 0.3 IU corresponded to 0.5001 µg (≈4.96e-10 mol) of OXT, for each administration of 5 µl (ie OXT 0.3 IU≈19 µg/kg or 11 IU/kg). This dose were chosen in order to be much lower than subcutaneous OXT doses (ie 250 µg/kg) used in mice that could have produced peripheral effects[204]. The doses we used are also similar to the higher range of intranasal OXT doses recently given to adolescent prairie voles[205]; even though in the previous study they used a much higher volume of injection (ie 25 µl) compared with our study (ie, 5 µl). Control mice received the same volume of saline (VEH). A 200-µl Eppendorf pipette with gel-loading tips (Costar) were used for administration. Drops of the 5 µl solution were gently placed equally on both nostrils of each mouse, which were taken in when they reflexively inhaled. Administration was rapid (less than 30 s) and handling was consistent across treatment groups. For the acute intranasal treatments, mice were administered with OXT only once, 5 min before the test. The delay of only 5 min was chosen based on evidence indicating that intranasal administration of OXT has very rapid pharmacokinetics with effects expected to appear already within a few minutes[206].

Stereotaxic surgery, viral injections and tetrodes implants

Wild type and dysbindin-1 mice were naïve and 2 to 3-months old at the time of surgery. All mice were anesthetized with 2% isoflurane in O2 by inhalation and mounted into a stereotaxic frame (Kopf) linked to a digital reader. Mice were maintained on 1.5 - 2% isoflurane during the surgery. Brain coordinates of viral injection in the mPFC was chosen in accordance to the mouse brain atlas (Paxinos and Watson, 1998): AP: +1.9 mm; ML: ± 0.2 mm; DV: -2.9 / -2.5 mm. Volume of AAV injection was 0.5 µL per hemisphere. We infused virus through a glass micropipette connected to a 10-µL Hamilton syringe. After infusion, injector was kept in place for 2 min and then slowly withdrawn to second injection spot. After virus injection mice were allowed 3 weeks to recover and for the viral transgenes to adequately express before undergoing behavioral experiments.

The skull was exposed and one holes was drilled to target the mPFC in accordance to the mouse brain atlas (Paxinos and Watson, 1998): AP: +1.9 mm; L: + 0.20 mm; DV: -2.6 mm. For in vivo electrophysiological recording, mice were implanted with silicon probes carrying four tetrodes in the right mPFC (Neuronexus A4x4-3mm-100-125-177-Z16). Prior to 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 7 to 10 days depending on the general health.

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-µm-thick coronal sections were cut using a Leica VT1000S 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 plus 4% normal goat serum (NGS), 0.2% Triton X-100 for 1 h (all at room temperature) and subsequently incubated with a GFAP polyclonal antibody (1:1000). Primary antisera were diluted in PBS plus 2% NGS overnight at 4°C for GFAP antibody. 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 a Alexa Flour 488 goat anti-rabbit IgG (H+L) (1:1000, Molecular Probes®, CatNo.A11034) in PBS plus 1% NGS, and subsequently washed there 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).

In vivo recordings

Neuronexus silicon probes carrying four tetrodes were implanted in the right mPFC (:+1,8 mm anterior, + 0,2 mm lateral from bregma, and -2,5 mm ventral from the brain surface) under general anesthesia. After 1 week of recovery from surgery, recordings were carried out mainly from the prelimbic cortex by means of a 16 channels Neuralynx Digital X system (NeuroLynx). 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 (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, Ireland). All quantitative analyses of neuronal data were performed offline using dedicated software (Plexon). Both putative pyramidal cells and putative interneurons were included in the analysis.

Data analysis

Neuronal firing rate. Single units were isolated using standard principal component and template matching techniques, provided by dedicate offline sorting software (Plexon). For each isolated unit, we verified that the projection of its spikes in the 3D 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 <1 ms was <0.5%. In case of relevant changes in the neuronal activity during acquisition, the entire animal was discarded and the data not included in the dataset.

Statistical Analysis. For the analysis of neural data, after identification of single units that remained stable over the duration of the experiment, we carried out Student's t-test applied to each trial, using a significance criterion of $p<0.05$ uncorrected. The null hypothesis was that a neuron does not vary its spontaneous activity during one of the two epochs of interest, PRE and POST of a social event, compared with a baseline epoch. In this study, the PRE and POST epochs correspond to a temporal window of 2 s previous and 2 s after the onset of a social event, respectively, while the baseline epoch corresponds to the first two second of the each recording session, in which the stimulus mouse had not been introduced yet to the testing cage. Therefore, all neurons showing a significant main effect in this analysis in at least one epoch of the first or the second trial were analyzed also at population level. Population analyses were performed taking into account single neuron responses expressed in terms of spk/s. For each neuron, the mean activity was calculated every 20 ms bins in all the recorded social events of all the experimental conditions to be compared by means and analyzed with different Student's t test with a significance criterion $p<0.05$ uncorrected.

Chapter 5

GENERAL CONCLUSION

General conclusion

The aim of this thesis was to clarify the implication of different brain pathways and systems in different aspects of mice social behavior. The work presented in the first two chapters of the thesis was dedicated to the setting a new behavioral tool able to highlight social cognition abilities in rodents. Emotion recognition tasks are the most extensively used paradigms to assess human social cognition, but lack of a parallel in rodents study. To fill this gap, we developed and validated a new behavioral test to assess these abilities in mice, highlighting, in particular, the ability to discriminate unfamiliar conspecific based on their emotional state. We showed that the ability to recognize emotions in conspecifics could be distinguished by previously reported emotion-associated processes, as vicarious freezing and emotion contagion. We investigated the sensory modalities mice use to communicate and perceive emotional states in conspecific, and we showed a prominent implication of visual and olfactory cues. More studies will be required to understand the specificity of these signals, and the way they lead to social responses in conspecifics. One of the main advantages of this new tool is the possibility to achieve causal information about the involvement of specific brain circuits in mediating emotion recognition process. In particular, the use of this setting allowed us to study the role of the endogenous oxytocin system in the processing of conspecifics emotions. In particular, we provide significant new insights on the role of the PVN-CeA oxytocin pathway. 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.

In humans, damages of the mPFC have been associated with impaired emotion recognition ability [128,129]. Combining *in vivo* electrophysiology with the emotion recognition test, we investigated the role of the mPFC. In particular, electrophysiology recordings revealed a mainly activation of interneurons during the exploration of mice in altered emotional state. By optogenetically manipulating two distinct classes of interneurons we revealed that SOM+ neurons play a critical role in the ability to recognize emotions, while optogenetic inhibition of mPFC PV+ neurons produced no effects. Overall, these results support the implication of a cortical excitatory and inhibitory imbalance as core behavioral dysfunctions in social cognitive deficits with possible relevance to disorders characterized by these social deficits such as autism and schizophrenia.

In the last period of my thesis, I focused more on the investigation of the behavioral and physiological effects produced by exogenous oxytocin administration. In particular, taking

advantage of the effects oxytocin produced in a mouse model of genetic liability, we investigated the physiological mechanisms of exogenous oxytocin action in the mPFC. We found that acute intranasal OXT administration was able to rescue social deficits in dysbindin-1 knockout mice and to ameliorate the neuronal and astrocytic pattern of activity within the mPFC. All together, these results further support the importance of mPFC E/I balance in social behavior. Additionally, these findings point to a potential unexpected role of astrocytes as possible key regulators of oxytocin effects.

Strength and Weakness

The findings presented in **chapter 2** demonstrate the validity of this novel paradigm to explore social cognitive processes in mice. Previous studies on social cognitive function in rodents, required familiarity between the observer and the demonstrator or involved observers witnessing conspecific in pain or exposed to a foot shock[95,96,99,101,102]. Differently, the manipulation that we performed was able to detect emotion recognition abilities toward unfamiliar conspecific without a direct emotion contagion, as we did not detect changes in corticosterone levels and escape or other stressed behavior, such as state-matching or freezeing. Moreover, this manipulation that we induced to alter the emotional state, causes minimal physically distress to the animal. However, further studies are essential in order to better understand how the visual and olfactory social cues are integrated in the brain circuits for promoting social interest toward relevant states. Another strength of methods used is the successful application of DREADD technology and ontogenetic manipulations to gain an understanding of the role of specific endogenous OXT pathways, as well as PFC microcircuits. Further experiments would be needed to explore if OXT PVN-CeA pathways might modulates cortical responses to emotion recognition or if the opposite direction is true. A possibility to start addressing this issue could be the combination of the DREDD technology in PVN projections with electrophysiology recording in mPFC.

The results presented in **chapter 3** were obtained combining in vivo electrophysiology with optogentic techniques demonstrate unreported role of mPFC interneurons in specific aspects of social cognition. Indeed, related to the rodent abilities for emotion recognition as previously discuss in **chapter 2**, using in vivo electrophysiology approach, we show that also at cortical level the emotional state elicit different mPFC activity, showing the social cue are differently integrated in the local mPFC network. However, we used silicon probes carrying four fixed tetrodes, which did not allow us to regulate deepness or electrodes position, complicating our understanding of unique frontal cytoarchitecture. Nevertheless, the optogenetic manipulations

of interneurons support the electrophysiology data and highlight the implication of a cortical inhibitory/excitatory balance in social function. Further analysis of electrophysiology and behavioral data in dysbindin-1 mice may confirm cortical alteration in a model of social deficits (data not included in this thesis). Moreover, since SOM+ neurons are highly enriched in oxytocin receptors[178], further investigation of the influence of OXT treatment at cortical level in wild type and dysibindin-1 mice may be important to better understand OXT effect in the modulation of emotional stimuli.

The findings presented in **chapter 4** combined multidisciplinary techniques and revealed how oxytocin treatment is able to restore social deficits and altered PFC function in dysbindin-1 mouse model, showing a possible role of astrocytic dopamine D2 receptors. Dysbindin-1 mutant mice are interesting model for studying the genetic component in the developing alteration of cortical and behavioral phenotype related to dopaminergic system, since this system is highly involved in psychiatric diseases. Initially in this study, we showed lower sociability and higher mPFC activity in dysbindin-1 mice compare to the wild type; however further analysis of the electrophysiology data is essential to identify the principal cells (i.e. pyramidal or interneurons) involved in the social interaction and to characterized if the OXT treatment change the activity of these cells, since OXT receptor is highly express in SOM+ cells. One strength of this work was to combine genetic data in mPFC with pharmacological response to highlight possible implication of astrocytes function in the ability to modulate social behavior. However, it is still missing a better characterization of changes in astrocyte/neuronal physiology after reducing astrocytic D2R expression. Another weak point is related to the electrophysiological result presented in **chapter 3**. We observed an increased of mPFC during exploration of demonstrator in the ERT, instead in freely moving animals the mPFC show an opposite modulation thus, arising questions which inputs drive the initial activation and subsequently inhibition during physical contact in the social interaction test and why mPFC is inhibited during the interaction and activated during exploration of conspecific. Further studies on mPFC circuits will be needed to answer these questions.

Future directions

Using mutant mice is a valuable source of information in determining both physiological and behavioral alterations. Our results in the emotion recognition test presented in **chapters 2-3** encourage further studies to investigate the connection between amygdala and cortical regions for emotions processing, and a region that could be mediating this information's flow could be the basolateral amygdala (BLA). In fact, an interconnected sheath of GABAergic neurons is found interposed between the BLA and CeA[207,208]. Importantly, the BLA is reciprocally connected with cortical regions, especially the prefrontal cortex. Based on this, manipulating BLA projections with DREDD or optogenetically technique in distinct phases of social cues presentation should be done, in order to find a possible converging point for the integration of social signals between cortical and subcortical areas.

Notably, the selectively optogenetic inhibition of SOM+ neurons in specific target layer of mPFC (i.e layer V important for the output generation) or specific area, such as prelimbic or infralimbic region could improve our knowledge on the role of prefrontal cortex circuits in social cognition functions.

The results presented in **chapter 4** indicate a potential role of astrocytes in modulation of brain activity after intranasal OXT administration. Future patch-clamp recording and Ca++ imaging experiments on mPFC slices should investigate the physiologic changes in the local network after the suppression of astrocytic D2 receptor. Nevertheless, it is still widely debated whether and how astrocytes release chemical transmitters in vivo[209], therefore the functional relevance of a direct neurotransmitter release and metabolism from astrocytes in the brain physiology will require further investigations. Furthermore only a few reports have suggested a link between mPFC activity and abnormal social behaviors in rodents [136,141], so future investigation are necessary to better identified the role of mPFC and its balance for the output behavior.

In this thesis we described that genetic variations reducing dysbindin-1 expression, which has been reported to impact human cognitive abilities[64,65] and responses to antipsychotic drugs[189], can also predict social responses to intranasal OXT . Future studies will be needed to better address the crosstalk between the dopaminergic and oxytonergic systems in the context of social behavior and their modulation by dysbindin genetic variants.

BIBLIOGRAPHY

1. Green, M. F., Horan, W. P. & Lee, J. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* **16**, 620–631 (2015).
2. Poletti, M., Enrici, I. & Adenzato, M. Cognitive and affective Theory of Mind in neurodegenerative diseases: neuropsychological, neuroanatomical and neurochemical levels. *Neurosci. Biobehav. Rev.* **36**, 2147–2164 (2012).
3. Frith, C. D. & Frith, U. Implicit and Explicit Processes in Social Cognition. *Neuron* **60**, 503–510 (2008).
4. Frith, C. D. & Frith, U. How we predict what other people are going to do. *Brain Res.* **1079**, 36–46 (2006).
5. Van Overwalle, F. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* **30**, 829–858 (2009).
6. Forbes, C. E. & Grafman, J. The role of the human prefrontal cortex in social cognition and moral judgment. *Annu. Rev. Neurosci.* **33**, 299–324 (2010).
7. Lavin, C. *et al.* The anterior cingulate cortex: an integrative hub for human socially-driven interactions. *Front. Neurosci.* **7**, 64 (2013).
8. Adolphs, R., Damasio, H., Tranel, D., Cooper, G. & Damasio, A. R. A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J. Neurosci. Off. J. Soc. Neurosci.* **20**, 2683–2690 (2000).
9. Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H. & Livingstone, M. S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
10. Adolphs, R. What does the amygdala contribute to social cognition? *Ann. N. Y. Acad. Sci.* **1191**, 42–61 (2010).

11. Nelson, E. E. & Guyer, A. E. The development of the ventral prefrontal cortex and social flexibility. *Dev. Cogn. Neurosci.* **1**, 233–245 (2011).
12. Batista, S. *et al.* Theory of Mind and Executive Functions are Dissociated in Multiple Sclerosis. *Arch. Clin. Neuropsychol. Off. J. Natl. Acad. Neuropsychol.* 1–11 (2017). doi:10.1093/arclin/acx101
13. Cusi, A. M., Macqueen, G. M. & McKinnon, M. C. Patients with bipolar disorder show impaired performance on complex tests of social cognition. *Psychiatry Res.* **200**, 258–264 (2012).
14. Patriquin, M. A., DeRamus, T., Libero, L. E., Laird, A. & Kana, R. K. Neuroanatomical and neurofunctional markers of social cognition in autism spectrum disorder. *Hum. Brain Mapp.* **37**, 3957–3978 (2016).
15. Kim, Y. *et al.* Mapping social behavior-induced brain activation at cellular resolution in the mouse. *Cell Rep.* **10**, 292–305 (2015).
16. Fernández, M., Mollinedo-Gajate, I. & Peñagarikano, O. Neural Circuits for Social Cognition: Implications for Autism. *Neuroscience* **370**, 148–162 (2018).
17. Barrash, J., Tranel, D. & Anderson, S. W. Acquired personality disturbances associated with bilateral damage to the ventromedial prefrontal region. *Dev. Neuropsychol.* **18**, 355–381 (2000).
18. Anderson, S. W., Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat. Neurosci.* **2**, 1032–1037 (1999).
19. Elston, G. N. Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb. Cortex N. Y. N 1991* **13**, 1124–1138 (2003).
20. Ongür, D. & Price, J. L. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex N. Y. N 1991* **10**, 206–219 (2000).
21. Vertes, R. P. Differential projections of the infralimbic and prelimbic cortex in the rat. *Synap. N. Y. N* **51**, 32–58 (2004).

22. Gabbott, P. L. A., Warner, T. A., Jays, P. R. L., Salway, P. & Busby, S. J. Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. *J. Comp. Neurol.* **492**, 145–177 (2005).
23. Schubert, D., Martens, G. J. M. & Kolk, S. M. Molecular underpinnings of prefrontal cortex development in rodents provide insights into the etiology of neurodevelopmental disorders. *Mol. Psychiatry* **20**, 795–809 (2015).
24. Clarke, I. J. Hypothalamus as an endocrine organ. *Compr. Physiol.* **5**, 217–253 (2015).
25. Swanson, L. W. & Sawchenko, P. E. Hypothalamic integration: organization of the paraventricular and supraoptic nuclei. *Annu. Rev. Neurosci.* **6**, 269–324 (1983).
26. Bosch, O. J. & Neumann, I. D. Both oxytocin and vasopressin are mediators of maternal care and aggression in rodents: from central release to sites of action. *Horm. Behav.* **61**, 293–303 (2012).
27. Lim, M. M. & Young, L. J. Neuropeptidergic regulation of affiliative behavior and social bonding in animals. *Horm. Behav.* **50**, 506–517 (2006).
28. Neumann, I. D. Brain oxytocin: a key regulator of emotional and social behaviours in both females and males. *J. Neuroendocrinol.* **20**, 858–865 (2008).
29. Guastella, A. J. & Hickie, I. B. Oxytocin Treatment, Circuitry, and Autism: A Critical Review of the Literature Placing Oxytocin Into the Autism Context. *Biol. Psychiatry* **79**, 234–242 (2016).
30. Insel, T. R. The challenge of translation in social neuroscience: a review of oxytocin, vasopressin, and affiliative behavior. *Neuron* **65**, 768–779 (2010).
31. Knobloch, H. S. *et al.* Evoked axonal oxytocin release in the central amygdala attenuates fear response. *Neuron* **73**, 553–566 (2012).
32. Li, K., Nakajima, M., Ibañez-Tallon, I. & Heintz, N. A Cortical Circuit for Sexually Dimorphic Oxytocin-Dependent Anxiety Behaviors. *Cell* **167**, 60-72.e11 (2016).

33. AJ, 18. Guastella. 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 (2015).
34. Y, 19. Aoki. Oxytocin improves behavioural and neural deficits in inferring others'. in *social emotions in autism. Brain : a journal of neurology 137* 3073–3086 (2014).
35. 20. Bakermans-Kranenburg MJ, van I. Sniffing around oxytocin: review and meta-analyses of trials in healthy and clinical groups with implications for pharmacotherapy. *Transl Psychiatry* **3**, (2013).
36. C, 21. Cacciotti-Saija. 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 (2015).
37. 22. Young LJ, B. C. Neuroscience. Can oxytocin treat autism? *Science* **347**, 825–826 (2015).
38. 23. Walum H, Y. L., Waldman ID. Statistical and Methodological Considerations for the Interpretation of Intranasal Oxytocin Studies. *Biol. Psychiatry* **79**, 251–257 (2016).
39. Phillips, A. G., Vacca, G. & Ahn, S. A top-down perspective on dopamine, motivation and memory. *Pharmacol. Biochem. Behav.* **90**, 236–249 (2008).
40. Yamaguchi, Y., Lee, Y.-A., Kato, A., Jas, E. & Goto, Y. The Roles of Dopamine D2 Receptor in the Social Hierarchy of Rodents and Primates. *Sci. Rep.* **7**, 43348 (2017).
41. Robinson, D. L., Zitzman, D. L. & Williams, S. K. Mesolimbic dopamine transients in motivated behaviors: focus on maternal behavior. *Front. Psychiatry* **2**, 23 (2011).
42. Lammel, S., Lim, B. K. & Malenka, R. C. Reward and aversion in a heterogeneous midbrain dopamine system. *Neuropharmacology* **76**, 351–359 (2014).
43. Poulin, J.-F. *et al.* Mapping projections of molecularly defined dopamine neuron subtypes using intersectional genetic approaches. *Nat. Neurosci.* **21**, 1260 (2018).

44. Burkett, J. P. & Young, L. J. The behavioral, anatomical and pharmacological parallels between social attachment, love and addiction. *Psychopharmacology (Berl.)* **224**, 1–26 (2012).
45. Girault, J.-A. & Greengard, P. The neurobiology of dopamine signaling. *Arch. Neurol.* **61**, 641–644 (2004).
46. Vallone, D., Picetti, R. & Borrelli, E. Structure and function of dopamine receptors. *Neurosci. Biobehav. Rev.* **24**, 125–132 (2000).
47. Ikemoto, S., Glazier, B. S., Murphy, J. M. & McBride, W. J. Role of dopamine D1 and D2 receptors in the nucleus accumbens in mediating reward. *J. Neurosci. Off. J. Soc. Neurosci.* **17**, 8580–8587 (1997).
48. Couppis, M. H., Kennedy, C. H. & Stanwood, G. D. Differences in aggressive behavior and in the mesocorticolimbic DA system between A/J and BALB/cJ mice. *Synap. N. Y. N* **62**, 715–724 (2008).
49. Davis, K. L., Kahn, R. S., Ko, G. & Davidson, M. Dopamine in schizophrenia: a review and reconceptualization. *Am. J. Psychiatry* **148**, 1474–1486 (1991).
50. Simpson, E. H., Kellendonk, C. & Kandel, E. A possible role for the striatum in the pathogenesis of the cognitive symptoms of schizophrenia. *Neuron* **65**, 585–596 (2010).
51. Grace, A. A. Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nat. Rev. Neurosci.* **17**, 524–532 (2016).
52. Talbot, K. *et al.* Dysbindin-1 and Its Protein Family. in *Handbook of Neurochemistry and Molecular Neurobiology: Schizophrenia* (eds. Lajtha, A., Javitt, D. & Kantrowitz, J.) 107–241 (Springer US, 2009). doi:10.1007/978-0-387-30410-6_5
53. Dickman, D. K. & Davis, G. W. The schizophrenia susceptibility gene dysbindin controls synaptic homeostasis. *Science* **326**, 1127–1130 (2009).
54. Ghiani, C. A. *et al.* The dysbindin-containing complex (BLOC-1) in brain: developmental regulation, interaction with SNARE proteins and role in neurite outgrowth. *Mol. Psychiatry* **15**, 115, 204–215 (2010).

55. Cox, M. M. *et al.* Neurobehavioral abnormalities in the dysbindin-1 mutant, sandy, on a C57BL/6J genetic background. *Genes Brain Behav.* **8**, 390–397 (2009).
56. Papaleo, F. & Weinberger, D. R. Dysbindin and Schizophrenia: it's dopamine and glutamate all over again. *Biol. Psychiatry* **69**, 2–4 (2011).
57. Kumamoto, N. *et al.* Hyperactivation of midbrain dopaminergic system in schizophrenia could be attributed to the down-regulation of dysbindin. *Biochem. Biophys. Res. Commun.* **345**, 904–909 (2006).
58. Ji, Y. *et al.* Role of dysbindin in dopamine receptor trafficking and cortical GABA function. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19593–19598 (2009).
59. Dysbindin-1 modulates prefrontal cortical activity and schizophrenia-like behaviors via dopamine/D2 pathways. *Mol. Psychiatry* **17**, 85–98 (2012).
60. Talbot, K. *et al.* Dysbindin-1 is reduced in intrinsic, glutamatergic terminals of the hippocampal formation in schizophrenia. *J. Clin. Invest.* **113**, 1353–1363 (2004).
61. Tang, J. *et al.* Dysbindin-1 in dorsolateral prefrontal cortex of schizophrenia cases is reduced in an isoform-specific manner unrelated to dysbindin-1 mRNA expression. *Hum. Mol. Genet.* **18**, 3851–3863 (2009).
62. Weickert, C. S. *et al.* Human dysbindin (DTNBP1) gene expression in normal brain and in schizophrenic prefrontal cortex and midbrain. *Arch. Gen. Psychiatry* **61**, 544–555 (2004).
63. Weickert, C. S., Rothmond, D. A., Hyde, T. M., Kleinman, J. E. & Straub, R. E. Reduced DTNBP1 (dysbindin-1) mRNA in the hippocampal formation of schizophrenia patients. *Schizophr. Res.* **98**, 105–110 (2008).
64. Burdick, K. E. *et al.* Genetic variation in DTNBP1 influences general cognitive ability. *Hum. Mol. Genet.* **15**, 1563–1568 (2006).

65. Fallgatter, A. J. *et al.* DTNBP1 (dysbindin) gene variants modulate prefrontal brain function in healthy individuals. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **31**, 2002–2010 (2006).
66. DP, 3. Kennedy. Adolphs R. Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia* **50**, 3313–3319 (2012).
67. Dunbar RI. The social brain hypothesis and its implications for social evolution. *Ann. Hum. Biol.* **36**, 562–572 (2009).
68. Sachdev PS. Clinical assessment of social cognitive function in neurological disorders. *Nat. Rev. Neurol.* **12**, 28–39 (2016).
69. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neurosci. Biobehav. Rev.* **35**, 573–588 (2011).
70. Skuse, D. H. *et al.* Common polymorphism in the oxytocin receptor gene (OXTR) is associated with human social recognition skills. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 1987–1992 (2014).
71. Guastella, A. J., Mitchell, P. B. & Mathews, F. Oxytocin enhances the encoding of positive social memories in humans. *Biol. Psychiatry* **64**, 256–258 (2008).
72. Domes, G. *et al.* Oxytocin attenuates amygdala responses to emotional faces regardless of valence. *Biol. Psychiatry* **62**, 1187–1190 (2007).
73. Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U. & Fehr, E. Oxytocin increases trust in humans. *Nature* **435**, 673–676 (2005).
74. Shahrestani, S., Kemp, A. H. & Guastella, A. J. The impact of a single administration of intranasal oxytocin on the recognition of basic emotions in humans: a meta-analysis. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **38**, 1929–1936 (2013).
75. Van IJzendoorn, M. H. & Bakermans-Kranenburg, M. J. 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 (2012).

76. De Dreu CK, Kret ME. Oxytocin Conditions Intergroup Relations Through Upregulated In-Group Empathy, Cooperation, Conformity, and Defense. *Biol. Psychiatry* **79**, 165–173 (2016).
77. Oxytocin improves specific recognition of positive facial expressions. *Psychopharmacology (Berl.)* **209**, 225–232 (2010).
78. Intranasal oxytocin increases covert attention to positive social cues. *Psychol. Med.* **43**, 1747–1753 (2013).
79. Meta-analysis of the effects of intranasal oxytocin on interpretation and expression of emotions. *Neurosci. Biobehav. Rev.* **78**, 125–144 (2017).
80. FS, 16. Chen. Genetic modulation of oxytocin sensitivity: a pharmacogenetic approach. *Transl Psychiatry* **5**, (2015).
81. H, 17. Tost. A common allele in the oxytocin receptor gene (OXTR) impacts prosocial temperament and human hypothalamic-limbic structure and function. in *Proceedings of the National Academy of Sciences of the United States of America 107* 13936–13941 (2010).
82. MR, 24. Lee. Oxytocin by intranasal and intravenous routes reaches the cerebrospinal fluid in rhesus macaques: determination using a novel oxytocin assay. *Mol. Psychiatry* **23**, 115–122 (2018).
83. 25. Bartz JA, B. N., Zaki J. Ochsner KN. Social effects of oxytocin in humans: context and person matter. in *Trends in cognitive sciences* **15**, 301–309 (2011).
84. Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nat. Rev. Neurosci.* **12**, 524–538 (2011).
85. Watanabe, T. *et al.* Diminished Medial Prefrontal Activity behind Autistic Social Judgments of Incongruent Information. *PLOS ONE* **7**, e39561 (2012).
86. H, 27. Huang. Chronic and acute intranasal oxytocin produce divergent social effects in mice. Neuropsychopharmacology : official publication of. *Am. Coll. Neuropsychopharmacol.* **39**, 1102–1114 (2014).

87. 30. Kurtz MM, R. C. Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. *Schizophr Bull* **38**, 1092–1104 (2012).
88. Interventions based on the Theory of Mind cognitive model for autism spectrum disorder (ASD). The Cochrane database of systematic. *reviews* **8785**, (2014).
89. DJ, 33. Langford. Social modulation of pain as evidence for empathy in mice. *Science* **312**, (1967).
90. TL, 35. Sterley. *Social transmission and buffering of synaptic changes after stress. Nature neuroscience*. (2018).
91. 36. Prochazkova E, K. M. Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neurosci. Biobehav. Rev.* **80**, 99–114 (2017).
92. 37. Ben-Ami Bartal I, M. P., Decety J. Empathy and pro-social behavior in rats. *Science* **334**, 1427–1430 (2011).
93. 32. Burkett JP, J. Z., Andari E. Curry DC, de Waal FB, Young LJ. Oxytocin-dependent consolation behavior in rodents. *Science* **351**, 375–378 (2016).
94. J, 38. LeDoux. Emotional networks and motor control: a fearful view. *Prog Brain Res* **107**, 437–446 (1996).
95. Gewirtz JC. Oxytocin enhances observational fear in mice. *Nat. Commun.* **8**, (2017).
96. D, 39. Jeon. Observational fear learning involves affective pain system. in *and Cav1.2 Ca2+ channels in ACC. Nature neuroscience* **13**, 482–488 (2010).
97. Rogers-Carter, M. M. Insular cortex mediates approach and avoidance responses to social affective stimuli. *Nat. Neurosci.* **21**, 404–414 (2018).
98. The vomeronasal system mediates sick conspecific avoidance. *Curr Biol* **25**, 251–255 (2015).
99. TF, 43. Sawyer. Androgen effects on responsiveness to aggression and stress-related odors of male mice. *Physiol Behav* **25**, 183–187 (1980).
100. 44. Zalaquett C, T. D. The effects of odors from stressed mice on conspecific behavior. *Physiol Behav* **50**, 221–227 (1991).

101. M, 48. Eliava. A New Population of Parvocellular Oxytocin Neurons Controlling Magnocellular Neuron Activity and Inflammatory Pain Processing. *Neuron* **89**, 1291–1304 (2016).

102. F, 49. Mormann. Neurons in the human amygdala encode face identity, but not gaze direction. *Nat. Neurosci.* **18**, 1568–1570 (2015).

103. 50. Stanley DA, A. R. Toward a neural basis for social behavior. *Neuron* **80**, 816–826 (2013).

104. Millner AJ, Casey BJ. Elevated amygdala response to faces and gaze aversion in autism spectrum disorder. *Soc. Cogn. Affect. Neurosci.* **9**, 106–117 (2014).

105. 52. Goghari VM, S. N. *Spilka MJ, Woodward TS. Task-Related Functional Connectivity Analysis of Emotion Discrimination in a Family Study of Schizophrenia. Schizophrenia bulletin*. (2017).

106. D, 55. Scheggia. Variations in Dysbindin-1 are associated with cognitive response to antipsychotic drug treatment. *Nat. Commun. 9* **2265**, (2018).

107. Epistatic interaction between COMT and DTNBP1 modulates prefrontal function in mice and in humans. *Mol. Psychiatry* **19**, 311–316 (2014).

108. 56. Dricu M, F. S. 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 (2016).

109. 57. Phelps EA, L. J. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**, 175–187 (2005).

110. DJ, 45. Anderson. Adolphs R. A framework for studying emotions across species. *Cell* **157**, 187–200 (2014).

111. 4. Green MF, L. J., Horan WP. Social cognition in schizophrenia. Nature reviews. *Neuroscience* **16**, 620–631 (2015).

112. 59. Sadananda M, W. M. Schwarting RK. Playback of 22-kHz and 50-kHz ultrasonic vocalizations induces differential c-fos expression in rat brain. *Neurosci Lett* **435**, 17–23 (2008).

113. Barrett LF. The Brain Basis of Positive and Negative Affect: Evidence from a Meta-Analysis of the Human Neuroimaging Literature. *Cereb Cortex* **26**, (1910).

114. M, 61. Haram. Contribution of oxytocin receptor polymorphisms to amygdala activation in schizophrenia spectrum disorders. *BJPsych Open* **2**, 353–358 (2016).

115. 62. Gamer M, B. C., Zurowski B. Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans. in *Proceedings of the National Academy of Sciences of the United States of America 107* 9400–9405 (2010).

116. JE, 63. Savage. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat Genet* **50**, 912–919 (2018).

117. Haroon E, Rilling JK. 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 (2015).

118. D, 64. Viviani. Oxytocin selectively gates fear responses through distinct outputs from the central amygdala. *Science* **333**, 104–107 (2011).

119. Oettl, L.-L. *et al.* Oxytocin Enhances Social Recognition by Modulating Cortical Control of Early Olfactory Processing. *Neuron* **90**, 609–621 (2016).

120. Owen, S. F. *et al.* Oxytocin enhances hippocampal spike transmission by modulating fast-spiking interneurons. *Nature* **500**, 458–462 (2013).

121. Marlin, B. J., Mitre, M., D'amour, J. A., Chao, M. V. & Froemke, R. C. Oxytocin enables maternal behaviour by balancing cortical inhibition. *Nature* **520**, 499–504 (2015).

122. Green, M. F. *et al.* Social Cognition in Schizophrenia: An NIMH Workshop on Definitions, Assessment, and Research Opportunities. *Schizophr. Bull.* **34**, 1211–1220 (2008).

123. A voxel-based lesion study on facial emotion recognition after penetrating brain injury. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3739908/. (Accessed: 5th December 2018)

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

125. 9. Adolphs, R. A mechanism for impaired fear recognition after amygdala damage. *Nature* **433**, 68 (2005).

126. 10. Frith, U., C. D. &. Frith. Mechanisms of Social Cognition. *Annu Rev Psychol* **63**, 287–313 (2011).

127. 11. Adolphs, R. The Social Brain: Neural Basis of Social Knowledge. *Annu Rev Psychol* **60**, 693–716 (2009).

128. 6. Monte, O. D. A voxel-based lesion study on facial emotion recognition after penetrating brain injury. *Soc Cogn Affect Neurosci* **8**, 632–639 (2013).

129. 12. Hiser, M., J. &. Koenigs. The Multifaceted Role of the Ventromedial Prefrontal Cortex in Emotion, Decision Making. *Soc. Cogn. Psychopathol. Biol. Psychiatry* (2018). doi:doi:10.1016/j.biopsych.2017.10.030

130. 13. Isaacson, M., J. S. &. Scanziani. How inhibition shapes cortical activity. *Neuron* **72**, 231–243 (2011).

131. 14. Rubenstein, M. M., J. L. R. &. Merzenich. Model of autism : increased ratio of excitation / inhibition in key neural systems. *Brain* **2**, 255–267 (2003).

132. 15. Hussman, J. P. Suppressed GABAergic inhibition as a common factor in suspected etiologies of autism. *J Autism Dev Disord* **31**, 247–8 (2001).

133. The Number of Parvalbumin-Expressing. in *Interneurons Is Decreased in the Medial Prefrontal Cortex in Autism. Cereb. Cortex* **27**, 1931–1943 (2017).

134. 17. Fatemi, S. H. 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).

135. 18. Harada, M. 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).

136. Yizhar, O. *et al.* Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).

137. 20. Selimbeyoglu, A. Modulation of prefrontal cortex excitation / inhibition balance rescues social behavior in CNTNAP2 -deficient mice. *Sci Transl Med 9* **6733**, (2017).

138. 21. Franklin, T. B. Prefrontal cortical control of a brainstem social behavior circuit. *Nat Neurosci* **20**, 260–270 (2017).

139. 22. Jeon, D. Observational fear learning involves affective pain system and Ca v 1.2 Ca 2+ channels in ACC. *Nat Neurosci* **13**, 482–488 (2010).

140. 23. Keum, S. A Missense Variant at the Nrxn3 Locus Enhances Empathy Fear in the Mouse. *Neuron* **98**, 588–601 (2018).

141. Wang, F., Kessels, H. W. & Hu, H. The mouse that roared: neural mechanisms of social hierarchy. *Trends Neurosci.* **37**, 674–682 (2014).

142. 25. Zhou, T. History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science* **80**, 162–168 (2017).

143. 26. Amadei, E. A. Dynamic corticostriatal activity biases social bonding in monogamous female prairie voles. *Nature* **546**, 297–301 (2017).

144. 5. Kennedy, R., D. P. &. Adolphs. Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia* **50**, 3313–9 (2012).

145. 27. Stark, E. Inhibition-Induced theta resonance in cortical circuits. *Neuron* **80**, 1263–1276 (2013).

146. 28. Barthó, P. Ongoing network state controls the length of sleep spindles via inhibitory activity. *Neuron* **82**, 1367–79 (2014).

147. Identification of basolateral amygdala projection cells and interneurons using extracellular recordings. *J Neurophysiol* **96**, 3257–65 (2006).

148. 31. Portfors, C. V. Types and functions of ultrasonic vocalizations in laboratory rats and mice. *J Am Assoc Lab Anim Sci* **46**, 28–34 (2007).

149. Stark, E. *et al.* Inhibition-Induced theta resonance in cortical circuits. *Neuron* **80**, 1263–1276 (2013).

150. 32. Crawley, J. N. Mouse behavioral assays relevant to the symptoms of autism. *Brain Pathol* **17**, 448–459 (2007).

151. Pisanksy, M.T., Hanson, L.R., Gottesman, I.I., & Gewirtz, J. C. Oxytocin enhances observational fear in mice. *Nat. Commun.* 1–11 (2018). doi:10.1038/s41467-017-02279-5

152. Blanchard, R. J. & Blanchard, D. C. Attack and defense in rodents as ethoexperimental models for the study of emotion. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **13 Suppl**, S3-14 (1989).

153. Sadananda, M., Wöhr, M. & Schwarting, R. K. W. Playback of 22-kHz and 50-kHz ultrasonic vocalizations induces differential c-fos expression in rat brain. *Neurosci. Lett.* **435**, 17–23 (2008).

154. Rogers-Carter, M. M. *et al.* Insular cortex mediates approach and avoidance responses to social affective stimuli. *Nat. Neurosci.* **21**, 404–414 (2018).

155. Zalaquett, C. & Thiessen, D. The effects of odors from stressed mice on conspecific behavior. *Physiol. Behav.* **50**, 221–7 (1991).

156. Boillat, M. *et al.* The vomeronasal system mediates sick conspecific avoidance. *Curr. Biol. CB* **25**, 251–255 (2015).

157. Sawyer, T. F. Androgen effects on responsiveness to aggression and stress-related odors of male mice. *Physiol. Behav.* **25**, 183–7 (1980).

158. Huynh, C. M., Vicente, G. I. & Peissig, J. J. The Effects of Familiarity on Genuine Emotion Recognition. *J. Vis.* **10**, 628–628 (2010).

159. Burkett, J. P. *et al.* Oxytocin-dependent consolation behavior in rodents. *Science* **351**, 375–8 (2016).

160. Langford, D. J. Social Modulation of Pain as Evidence for Empathy in Mice. *Science* **312**, 1967–1970 (2006).

161. Sterley, T. L. *et al.* Social transmission and buffering of synaptic changes after stress. *Nat. Neurosci.* **21**, 393–403 (2018).

162. Jeon, D. *et al.* Observational fear learning involves affective pain system and Cav1.2 Ca2+ channels in ACC. *Nat. Neurosci.* **13**, 482–488 (2010).

163. Monfils, M.-H. & Agee, L. Insights from social transmission of information in rodents. *Genes Brain Behav.* e12534 (2018). doi:10.1111/gbb.12534

164. Bartal, I. B.-A., Decety, J. & Mason, P. Empathy and Pro-Social Behavior in Rats. *Science* **334**, 1427–1430 (2011).

165. 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).

166. Green, M. F., Horan, W. P. & Lee, J. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* **16**, 620–631 (2015).

167. Lee, E. *et al.* Enhanced Neuronal Activity in the Medial Prefrontal Cortex during Social Approach Behavior. *J. Neurosci.* **36**, 6926–6936 (2016).

168. Yizhar, O. *et al.* Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).

169. 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).

170. McGarry, L. M. & Carter, A. G. Inhibitory Gating of Basolateral Amygdala Inputs to the Prefrontal Cortex. *J. Neurosci.* **36**, 9391–9406 (2016).

171. Selimbeyoglu, A. *et al.* Modulation of prefrontal cortex excitation / inhibition balance rescues social behavior in CNTNAP2 -deficient mice. *Sci. Transl. Med.* **9**, eaah6733 (2017).

172. Kepecs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014).

173. Kubota, Y. Untangling GABAergic wiring in the cortical microcircuit. *Curr. Opin. Neurobiol.* **26**, 7–14 (2014).

174. Marlin, J. J. & Carter, A. G. GABA-A receptor inhibition of local calcium signaling in spines and dendrites. *J. Neurosci. Off. J. Soc. Neurosci.* **34**, 15898–911 (2014).

175. Abbas, A. I. *et al.* Somatostatin Interneurons Facilitate Hippocampal-Prefrontal Synchrony and Prefrontal Spatial Encoding. *Neuron* (2018). doi:10.1016/j.neuron.2018.09.029

176. Jennings, J. H. *et al.* Distinct extended amygdala circuits for divergent motivational states. *Nature* **496**, 224 (2013).

177. Namburi, P. *et al.* A circuit mechanism for differentiating positive and negative associations. *Nature* **520**, 675–678 (2015).

178. Nakajima, M., Görlich, A. & Heintz, N. Oxytocin modulates female sociosexual behavior through a specific class of prefrontal cortical interneurons. *Cell* **159**, 295–305 (2014).

179. Huang, H. *et al.* Chronic and acute intranasal oxytocin produce divergent social effects in mice. *Neuropsychopharmacology* **39**, (2014).

180. Raam, T., McAvoy, K. M., Besnard, A., Veenema, A. & Sahay, A. Hippocampal oxytocin receptors are necessary for discrimination of social stimuli. *Nat. Commun.* **8**, 1–14 (2017).

181. Oettl, L. L. *et al.* Oxytocin Enhances Social Recognition by Modulating Cortical Control of Early Olfactory Processing. *Neuron* **90**, 609–621 (2016).

182. Domes, G. *et al.* Oxytocin Attenuates Amygdala Responses to Emotional Faces Regardless of Valence. *Biol. Psychiatry* **62**, 1187–1190 (2007).

183. Shin, N. Y. *et al.* Effects of oxytocin on neural response to facial expressions in patients with schizophrenia. *Neuropsychopharmacology* **40**, 1919–1927 (2015).

184. Gordon, I., Martin, C., Feldman, R. & Leckman, J. F. Oxytocin and social motivation. *Dev. Cogn. Neurosci.* **1**, 471–493 (2011).

185. Panksepp, J. & Panksepp, J. B. Toward a cross-species understanding of empathy. *Trends Neurosci.* **36**, (2013).

186. Zak, P. J., Stanton, A. A. & Ahmadi, S. Oxytocin Increases Generosity in Humans. *PLOS ONE* **2**, e1128 (2007).

187. Loos, M. *et al.* Neuregulin-3 in the mouse medial prefrontal cortex regulates impulsive action. *Biol. Psychiatry* **76**, 648–655 (2014).

188. Niwa, M. *et al.* Knockdown of DISC1 by in utero gene transfer disturbs postnatal dopaminergic maturation in the frontal cortex and leads to adult behavioral deficits. *Neuron* **65**, 480–489 (2010).

189. Scheggia, D. *et al.* Variations in Dysbindin-1 are associated with cognitive response to antipsychotic drug treatment. *Nat. Commun.* **9**, 2265 (2018).

190. Henneberger, C. & Rusakov, D. A. Synaptic plasticity and Ca2+ signalling in astrocytes. *Neuron Glia Biol.* **6**, 141–146 (2010).

191. Perea, G., Navarrete, M. & Araque, A. Tripartite synapses: astrocytes process and control synaptic information. *Trends Neurosci.* **32**, 421–431 (2009).

192. Zhang, Y. *et al.* Activation of Dopamine D2 Receptor Suppresses Neuroinflammation through αB-crystalline by Inhibition of NF-κB Nuclear Translocation in Experimental ICH Mice Model. *Stroke J. Cereb. Circ.* **46**, 2637–2646 (2015).

193. Yuan, Q. *et al.* Regulation of Brain-Derived Neurotrophic Factor Exocytosis and Gamma-Aminobutyric Acidergic Interneuron Synapse by the Schizophrenia Susceptibility Gene Dysbindin-1. *Biol. Psychiatry* **80**, 312–322 (2016).

194. Lee, E. *et al.* Enhanced Neuronal Activity in the Medial Prefrontal Cortex during Social Approach Behavior. *J. Neurosci. Off. J. Soc. Neurosci.* **36**, 6926–6936 (2016).

195. Galbusera, A. *et al.* Intranasal Oxytocin and Vasopressin Modulate Divergent Brainwide Functional Substrates. *Neuropsychopharmacology* **42**, 1420–1434 (2017).

196. Verkhratsky, A. & Nedergaard, M. The homeostatic astroglia emerges from evolutionary specialization of neural cells. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, (2016).

197. Dossi, E., Vasile, F. & Rouach, N. Human astrocytes in the diseased brain. *Brain Res. Bull.* **136**, 139–156 (2018).

198. Croft, W., Dobson, K. L. & Bellamy, T. C. Plasticity of Neuron-Glial Transmission: Equipping Glia for Long-Term Integration of Network Activity. *Neural Plast.* **2015**, (2015).

199. Miller, G., Chen, E. & Cole, S. W. Health psychology: developing biologically plausible models linking the social world and physical health. *Annu. Rev. Psychol.* **60**, 501–524 (2009).

200. de la Mora, M. P. *et al.* Signaling in dopamine D2 receptor-oxytocin receptor heterocomplexes and its relevance for the anxiolytic effects of dopamine and oxytocin interactions in the amygdala of the rat. *Biochim. Biophys. Acta BBA - Mol. Basis Dis.* **1862**, 2075–2085 (2016).

201. Romero-Fernandez, W., Borroto-Escuela, D. O., Agnati, L. F. & Fuxe, K. Evidence for the existence of dopamine d2-oxytocin receptor heteromers in the ventral and dorsal striatum with facilitatory receptor–receptor interactions. *Mol. Psychiatry* **18**, 849–850 (2013).

202. Scearce-Levie, K. *et al.* Abnormal social behaviors in mice lacking Fgf17. *Genes Brain Behav.* **7**, 344–354 (2008).

203. Ferguson, J. N. *et al.* Social amnesia in mice lacking the oxytocin gene. *Nat. Genet.* **25**, 284–288 (2000).

204. Sala, M. *et al.* Pharmacologic rescue of impaired cognitive flexibility, social deficits, increased aggression, and seizure susceptibility in oxytocin receptor null mice: a neurobehavioral model of autism. *Biol. Psychiatry* **69**, 875–882 (2011).

205. Bales, K. L. *et al.* Chronic intranasal oxytocin causes long-term impairments in partner preference formation in male prairie voles. *Biol. Psychiatry* **74**, 180–188 (2013).

206. Veening, J. G. & Olivier, B. Intranasal administration of oxytocin: behavioral and clinical effects, a review. *Neurosci. Biobehav. Rev.* **37**, 1445–1465 (2013).

207. Marowsky, A., Yanagawa, Y., Obata, K. & Vogt, K. E. A specialized subclass of interneurons mediates dopaminergic facilitation of amygdala function. *Neuron* **48**, 1025–1037 (2005).

208. Ehrlich, I. *et al.* Amygdala inhibitory circuits and the control of fear memory. *Neuron* **62**, 757–771 (2009).

209. Hamilton, N. B. & Attwell, D. Do astrocytes really exocytose neurotransmitters? *Nat. Rev. Neurosci.* **11**, 227–238 (2010).