



Neural Circuits Underlying Social Fear in Rodents: An Integrative Computational Model

Valerio Alfieri[†], Andrea Mattera^{*†} and Gianluca Baldassarre

Institute of Cognitive Sciences and Technologies, National Research Council, Rome, Italy

Social avoidance in rodents arises from a complex interplay between the prefrontal cortex and subcortical structures, such as the ventromedial hypothalamus and the dorsal periaqueductal gray matter. Experimental studies are revealing the contribution of these areas, but an integrative view and model of how they interact to produce adaptive behavior are still lacking. Here, we present a computational model of social avoidance, proposing a set of integrated hypotheses on the possible macro organization of the brain system underlying this phenomenon. The model is validated by accounting for several different empirical findings and produces predictions to be tested in future experiments.

Keywords: social fear, social anxiety disorder, ventromedial hypothalamus, medial prefrontal cortex, computational neuroscience

OPEN ACCESS

Edited by:

Conor J. Houghton,
University of Bristol, United Kingdom

Reviewed by:

Nancy Padilla Coreano,
Salk Institute for Biological Studies,
United States
Amiel Rosenkranz,
Rosalind Franklin University of
Medicine and Science, United States

*Correspondence:

Andrea Mattera
andrea.mattera@istc.cnr.it

[†]These authors have contributed
equally to this work and share first
authorship

Received: 21 December 2021

Accepted: 11 February 2022

Published: 08 March 2022

Citation:

Alfieri V, Mattera A and Baldassarre G
(2022) Neural Circuits Underlying
Social Fear in Rodents: An Integrative
Computational Model.
Front. Syst. Neurosci. 16:841085.
doi: 10.3389/fnsys.2022.841085

1. INTRODUCTION

The social anxiety disorder (SAD) is the third most frequent psychiatric disorder, also representing a risk factor for depression and addiction (Cohen et al., 2007; Stein and Stein, 2008; Leichsenring and Leweke, 2017; Dos Santos et al., 2019). In humans, SAD has been linked to alterations in the activity of prefrontal cortex, hyper-activation of the amygdala, and dysfunctions in many other subcortical regions, such as the bed nucleus of the stria terminalis (BNST), the striatum and the periaqueductal gray matter (PAG) (Berkowitz et al., 2007; Goldin et al., 2009; Labuschagne et al., 2012; Arnold Anteraper et al., 2014; Duval et al., 2015; Marazziti et al., 2015; Stein, 2015; Clauss et al., 2019).

Even though cognitive behavioral therapy is an effective solution for the treatment of SAD, it has limitations concerning the individual costs and the timing of the therapy (Pilling et al., 2013; Mayo-Wilson et al., 2014; Scaini et al., 2016; Dos Santos et al., 2019). Evidence suggests that pharmacological medications could complement a psychological approach, especially in cases when the severity of the impairment could cause other psychological and health risks, such as depression and suicide attempts (Hambrick et al., 2003; Vitiello, 2009; Kelly et al., 2014; Rao and Andrade, 2017). Unfortunately, current pharmacological treatments are based on serotonin and norepinephrine reuptake inhibitors, presenting multiple collateral effects and requiring at least some weeks to obtain a therapeutic response (Dos Santos et al., 2019). Animal research is very helpful for understanding the pathogenesis of SAD and for the development of effective drugs for its treatment (Wang et al., 2020). Indeed, animals characterised by rich inter-specific social interactions usually exhibit a high level of behavioral flexibility. Social behavior often entails repeated encounters between antagonist individuals. These interactions cause the defeated animals to experience different levels of physical and psychosocial stress and promote the adjustment of behavior to suitably cope with future encounters (Chen and Hong, 2018; Diaz and Lin, 2020). The shift toward a social avoidance strategy is presumably an adaptive mechanism aimed to diminish

future harm and facilitate alternative routes to the attainment of essential resources (Christoffel et al., 2015).

The translational value of murine research has long been undermined by the lack of paradigms capable of reproducing the core symptoms of SAD without confounding effects. Indeed, most protocols also induce unspecific symptoms, e.g., anhedonia, general anxiety, and impaired locomotory activity (Huhman, 2006; Wang et al., 2020). In the last decade, however, it has been discovered that two protocols overcome these limitations. In the classical social defeat protocol an aggressive conspecific is placed in the same cage of the experimental mouse to allow a non-lethal conflict. If this procedure is repeated no more than 3 days and for a few minutes only at a time (sub-chronic defeat protocol), it causes a specific social impairment without manifestation of anxiety, stress or altered exploratory behavior (Franklin et al., 2017). An alternative protocol to avoid unspecific symptoms, also known as social fear conditioning, consists of pairing the social investigation of a conspecific with a foot shock, akin to what is done in auditory fear conditioning (Toth et al., 2012, 2013). After these protocols the defeated animal displays avoidance of conspecifics, measured as more time spent freezing and in defensive postures, less time spent investigating the conspecific and a higher number of sudden retreats from the investigation (Toth et al., 2012, 2013; Franklin et al., 2017; Xu et al., 2019; Krzywkowski et al., 2020).

A growing amount of research has begun to dissect the neural substrates of social avoidance in rodents. This led to the identification of distinct brain regions involved in the processing of defensive responses to aggressive conspecifics and their dissociation from the parallel and mostly non-overlapping neural circuits encoding predatory fear or auditory fear conditioning (Gross and Canteras, 2012; Silva et al., 2013, 2016a). The ventrolateral part of the ventromedial hypothalamus (VMHvl) is considered to be a central hub for the coordination of behavioral responses directed toward conspecifics, such as aggression, mating, approach, and defense (Lin et al., 2011; Gross and Canteras, 2012; Sakurai et al., 2016). The VMHvl receives projections from the medial amygdala (MeA), that relays conspecific-related information from pheromonal and olfactory cues captured by the olfactory systems (Gross and Canteras, 2012; Silva et al., 2016a), and information on the spatial context from the ventral hippocampus (vHIP, Chang and Gean, 2019).

Calcium imaging studies revealed that sub-chronic social defeat induces phasic responses in a VMHvl neural population associated with an internal state of social threat (Krzywkowski et al., 2020). Subsequent exposure to the context where the defeat took place reactivates the same population that was recruited during the traumatic experience, suggesting that VMHvl encodes a fear engram (Krzywkowski et al., 2020). Moreover, optogenetic activation of the neurons that were active during the defeat is sufficient to enact fear manifestations (Sakurai et al., 2016). The downstream target of the VMHvl is the dorsal part of the periaqueductal gray (dPAG; Silva et al. 2013). This is a midbrain nucleus that orchestrates the defensive behavior by coordinating the structures responsible for performing the motor action.

Besides VMHvl and dPAG, another region involved in the regulation of social interactions is the medial prefrontal cortex

(mPFC; Franklin et al. 2017; Xu et al. 2019; Wang et al. 2020). Modern studies aimed at explaining how cortical and subcortical systems interact to support social fear processes in rodents are drawing a puzzling picture. In particular, Franklin et al. (2017) found that sub-chronic social defeat induces a depotentiation of the connections from the mediodorsal thalamus (MDT) to the mPFC. In turn, the mPFC becomes less effective in the recruitment of the dPAG, resulting in social withdrawal. In agreement with this, the effect of sub-chronic social defeat can be mimicked through the inhibition of the output from the mPFC to the dPAG. This is in line with the experiments showing that subordinate mice in social ranking tests display reduced excitability in layer 5 neurons of the mPFC (Wang et al., 2011), whereas in dominant mice the connection between the MDT and the mPFC is potentiated (Zhou et al., 2017). Seemingly at odds with such evidence, Xu et al. (2019) found an increased firing of the pyramidal neurons in the mPFC during social avoidance. Consequently, lowering their activity through the manipulation of the upstream GABAergic neurons recovers social interactions (Xu et al., 2019). Overall, the fact that both an increase and a decrease of the activity of the mPFC excitatory neurons mediate social avoidance suggests that different sub-populations of the mPFC pyramidal neurons could exert an opposite influence on social behavior by acting as fear-ON and fear-OFF populations. Two lines of evidence further support this hypothesis. First, a study of aggressive behavior highlights a complex dual role of the mPFC, that is shown to be capable of inhibiting but also promoting aggression, possibly through the engagement of different sub-regions (Biro et al., 2017). Second, another study shows that during social exploration a neural population of the mPFC increases its activity while at the same time a second population turns off (Liang et al., 2018).

Reconciling these multiple pieces of evidence requires the integration of all the available information in the same framework to clarify the possible roles and interactions that cortical and subcortical structures express in social avoidance. The aim of this work is hence to present a system-level computational model able to aggregate the different experimental findings in a coherent scheme.

To our knowledge, while classical fear conditioning has been widely investigated with the use of computational models (Burgos and Murillo-Rodríguez, 2007; Mannella et al., 2008; Krasne et al., 2011; Anastasio, 2013; John et al., 2013; Moustafa et al., 2013; Carrere and Alexandre, 2015; Li et al., 2016; Bennett et al., 2019; Mattered et al., 2020), the model proposed here is the first to account for social avoidance. The model is based on a number of hypotheses, grounded on literature, concerning the involved brain areas, neural populations, and connections between them underlying social avoidance. The goal of the model construction is to operationalize such hypotheses and integrate them in a coherent whole (also giving a possible explanation to the “puzzling picture” illustrated above), to validate this with the qualitative reproduction of several current experiments, and then to derive new predictions testable in future experiments (Shen and McNaughton, 1996).

The rest of this article is organized as follows. First, we describe how we developed the model based on reasoned

hypotheses grounded on the empirical literature. Second, we validate the model by reproducing some important experimental findings on sub-chronic social defeat and social fear conditioning. Third, we manipulate the model to produce new testable predictions. Last, we discuss our findings in light of the existing literature.

2. METHODS

2.1. Firing Rate Units

The model is formed by firing rate leaky units, each representing a population of neurons. The voltage V_{post} of a post-synaptic unit is regulated through the following differential equation:

$$\tau \dot{V}_{post} = -V_{post} + I + \sum_{pre} w_{post,pre} F(V_{pre}) \quad (1)$$

where τ is the time constant, I is the external input to the unit (representing the “defeat,” the “conspicuous,” and the “context;” see **Figure 1**), $w_{post,pre}$ is the connection weight between the presynaptic unit pre and unit $post$. F is the activation of the unit, computed with the hyperbolic tangent function $\tanh(x)$, and represents the firing rate of a population of neurons (Burgos and Murillo-Rodríguez, 2007; Moustafa et al., 2013; Carrere and Alexandre, 2015; Mannella et al., 2016; Bennett et al., 2019; Mattera et al., 2020):

$$F(V) = [\tanh(V)]^+ \quad (2)$$

where $[x]^+$ is the positive function ($[x]^+ = x$ if $x \geq 0$, and $[x]^+ = 0$ if $x < 0$). The \tanh function restricts the range of the units activation in the interval $[0,1]$. The equation was approximated with the Euler method with discrete time steps Δt (the values of the model parameters are listed in **Supplementary Table S1**).

2.2. Model Architecture

Figure 1 illustrates the brain areas and their connections forming the architecture of the model; these areas and connections are grounded in the biological literature on social fear. This architecture is not meant to be comprehensive of all the mouse brain areas involved in social behavior, but only of those necessary to explain the key target experiments of section 3.3. In the decision of which areas to include or not into the model we used the following “criteria of coherence” by selecting only studies that: (1) focused on mice; (2) investigated areas activated during an aversive social encounter (Silva et al., 2013; Krzywkowski et al., 2020) or after paradigms of social sub-chronic defeat or foot-shock social fear conditioning (Franklin et al. 2017; Xu et al. 2019; Krzywkowski et al. 2020; conditions that avoid non-specific symptoms and confounding factors) when the defeated mouse interacted with the conspecific. This allowed us to identify and simulate the murine brain areas recruited in naive mice at the time of sub-chronic social defeat, and those activated in defeated mice when they subsequently avoid the conspecifics. The information coming from those articles points out the VMHvl, dPAG and mPFC as areas strongly involved in the specific symptoms, thus we focused on them

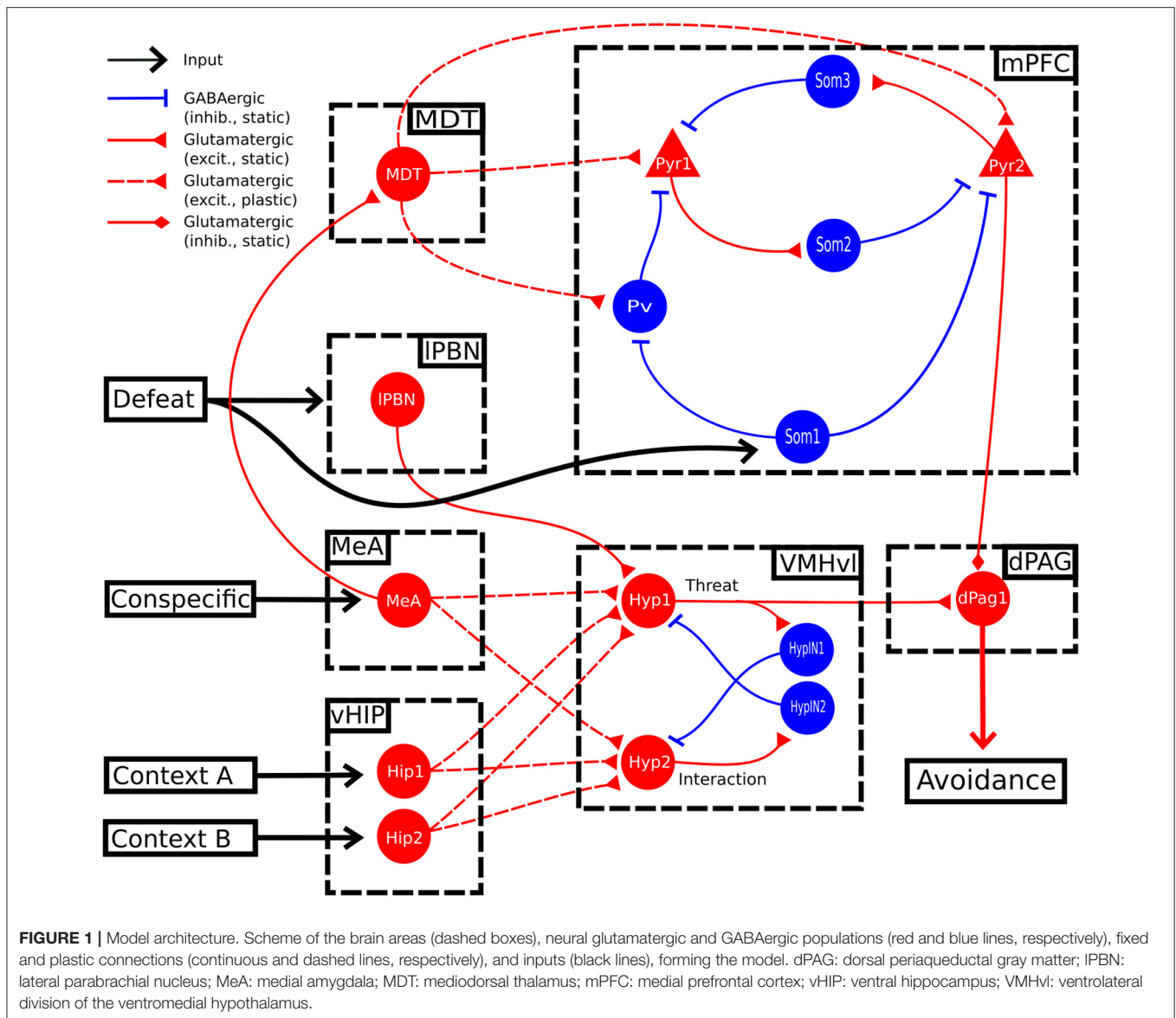
in our model. Knowledge gaps in the literature were filled by formulating specific hypotheses, that were subsequently validated with the reproduction of the target experiments.

The model receives three inputs, representing the conspecific, the spatial context, and the social defeat. The first input activates the MeA (Li et al., 2017; Nordman et al., 2020) and is related to the detection of sensory cues signaling the presence of a conspecific; the second input activates the vHIP that encodes spatial information about two possible alternative contexts through two neural populations, Hip1 and Hip2. Both the conspecific and contextual input reach two populations of the VMHvl (see below; Sakurai et al. 2016; Chang and Gean 2019; Wang et al. 2019; Falkner et al. 2020; Krzywkowski et al. 2020). The third input encodes the social defeat and activates the lateral parabrachial nucleus (IPBN), which in turn conveys excitatory pain-related signals to the VMHvl (Chiang et al., 2020). The VMHvl sends excitatory connections to the dPAG excitatory neural unit dPag1, which triggers a social avoidance behavioral output (Silva et al., 2013; Deng et al., 2016).

In the model, there are two units in the VMHvl, Hyp1 and Hyp2, whose activity is antagonistic. We modeled them on the basis of the results from Krzywkowski et al. (2020) who showed that, during the aversive social encounter, a population of VMHvl neurons increase its activity while another one decreases it (the authors called them “Defeat+” and “Defeat-.”) In the construction of the VMHvl connectivity, we hypothesized that the two excitatory neural populations within the VMHvl exhibit mutual lateral inhibition through local GABAergic interneurons (HypIN1 and HypIN2), effectively implementing a *winner-takes-all* mechanism corresponding to a competition to control behavior between the two populations. The presence of inhibitory interneurons within VMHvl has been proposed in previous studies (Choi et al., 2005; Kim et al., 2019; Lo et al., 2019).

In addition to VMHvl, the projections from the MeA reach the MDT (Krettek and Price, 1977; Canteras et al., 1995; Mitchell and Chakraborty, 2013). This relays multi-modal cognitive/emotional information to the cortex through plastic connections (Kuroda et al., 1998; Delevich et al., 2015; Franklin et al., 2017; Collins et al., 2018; Nelson et al., 2019).

In the construction of the mPFC, we followed the standard approach of the top-down models (John et al., 2013; Moustafa et al., 2013; Li et al., 2016; Oliva et al., 2018; Bennett et al., 2019; Mattera et al., 2020), abstracting over the layered circuits and intra-cortical connectivity of the cortex. In particular, since ours is a system-level model considering several areas, it was not possible to include in the model all the known connections linking those areas. We thus followed a parsimony principle and included in the model only the connections that were functionally relevant to actually capture the target experiments. Since, however, the necessity of such connections was still an open problem, we also performed a sensitivity analysis (see below) to highlight the importance of the considered connections. Analogously to what stated in Mannella et al. (2016), this approach has the following advantages: (a) the production of hypotheses about the identity and connectivity of the neural populations of the mPFC relevant for the targeted phenomenon, here social avoidance; (b) the identification of the knowledge gaps



regarding the mechanisms of recruitment of those populations; and (c) the development of predictions based on them that can be empirically tested.

The authors of Xu et al. (2019) reported that sub-chronic social defeat leads to an increase of the firing of the mPFC pyramidal neurons, and that a stimulation of parvalbumin positive neurons (parvalbumin+), performed through the inhibition of the upstream somatostatin positive interneurons (somatostatin+), reduces their activity and reverts conditioning. We thus implemented in the model an excitatory neural population, Pyr1, targeted by a double-inhibition circuit formed by Pv and Som1, representing, respectively, parvalbumin+ and somatostatin+ neurons (Figure 1). Interestingly, auditory fear conditioning relies on a mPFC disinhibitory circuit analogous to the one described for social fear conditioning (Courtin et al., 2014; Cummings

and Clem, 2020). In particular, it has been shown that the acquisition of auditory fear conditioning requires the transient inhibition of the parvalbumin+ interneurons (Courtin et al., 2014). It has been suggested that a similar recruitment of a disinhibitory microcircuit, through the activation of somatostatin+ interneurons upstream of the parvalbumin+ interneurons, could establish social fear in the mPFC (Wang et al., 2020). On this basis, in the model we hypothesized that sub-chronic social defeat recruits both the nociceptive center IPBN and the unit Som1 in the mPFC causing the disinhibition of Pyr1.

A model where social avoidance is driven, at the mPFC level, only by the activation of a single population of pyramidal neurons cannot explain the findings of Franklin et al. (2017), which found that a decrease in the layer 5 of the mPFC output to the dPAG leads to social avoidance. We thus hypothesized

that another population of pyramidal neurons, besides Pyr1, is involved in sub-chronic social defeat. In the model, this population Pyr2 should inhibit social avoidance in an opposite way with respect to the Pyr1 population described by Xu et al. (2019). Supporting this claim, and in accordance to Franklin et al. (2017), a recent experimental study showed that the administration of the psychedelic lysergic acid diethylamide (LSD), which enhances excitatory burst firing, promotes social behavior in mice. Interestingly, the pro-social effect of LSD is prevented by the optogenetic inhibition of the mPFC pyramidal neurons (De Gregorio et al., 2021). The two units Pyr1 and Pyr2 are reciprocally connected by Som2 and Som3, representing two populations of putative somatostatin+ interneurons driving lateral inhibition (Kapfer et al., 2007; Silberberg and Markram, 2007; Riedemann, 2019).

The second hypothesis concerns the mPFC-dPAG projections. As Franklin et al. (2017) previously observed, prefrontal projections from pyramidal neurons located in layer 5 exert inhibitory control over the activity of the dPAG neurons. However, it has also been established that pyramidal cells projecting to the dPAG are glutamatergic and therefore the mechanism through which they exert the inhibitory control remains to be explained. In *ex vivo* circuit-mapping experiments, optogenetic-induced stimulation of pyramidal prefrontal projections to the dPAG induces short-latency excitatory currents in some Vglut2+ cells (13%), but not in the Vgat+ cell (Franklin et al., 2017). This suggests the absence of a feedforward GABAergic inhibition and the possibility of an indirect, long-latency suppression of incoming hypothalamic inputs driven by metabotropic glutamatergic receptors. For this reason, we simulated the prefrontal top-down control of Pyr2 on the dPag1 with a direct inhibition, thus abstracting over the specific mechanism supporting the process.

In addition to this “basic model,” we explored the possibility of an alternative model (Figure 3) where we connected another area, indicated as lateral septum (LS), to the threat unit Hyp1 of the VMHvl. The possibility of this interaction is supported by (a) anatomical tracing studies (Risold and Swanson, 1997) and (b) functional evidence showing that the optogenetic stimulation of the LS projections reaching VMHvl induces social investigation (Wong et al., 2016).

The weights of the connections of the two models (listed in Supplementary Table S2, together with the literature supporting their existence) were manually adjusted until we found a configuration sufficient to reproduce all the target experiments (Mattera et al., 2020). To test the robustness of the results to the modification of the parameters, we performed a sensitivity analysis (see Supplementary Material and Section 3.5). Given the absence of noise and of weight randomization the model is completely deterministic, so it was not needed to run the simulation multiple times with different seeds of the random number generator.

2.3. Synaptic Plasticity

Some of the connections of the models are fixed while others are plastic. The weights of the plastic connections are updated according to a simplified Bienenstock–Cooper–Munro (BCM)

learning rule (Bienenstock et al., 1982):

$$\Delta W = \alpha \cdot (F_{post} - \theta) \cdot F_{pre} \quad (3)$$

where α is the learning rate, θ is a threshold (Supplementary Table S1), and F_{post} and F_{pre} are the firing rates of the post- and pre-synaptic units. The plastic weights were clipped within a (W_{max} , W_{min}) range (Supplementary Table S1).

To decide which connections had to be plastic, we followed this strategy. First, we considered to be plastic all the connections that were found to be plastic in the experimental literature on the social behavior of mice (Supplementary Table S2). The rest were considered fixed. Finally, we ran the simulations to fit the experimental data. We observed that, to reproduce the experiment of Krzykowski et al. (2020), the connection between vHIP and VMHvl had to be made plastic. This corresponds to a principle of parsimony aimed at simplifying the model, staying as much as possible grounded on literature, and minimizing the hypothesis.

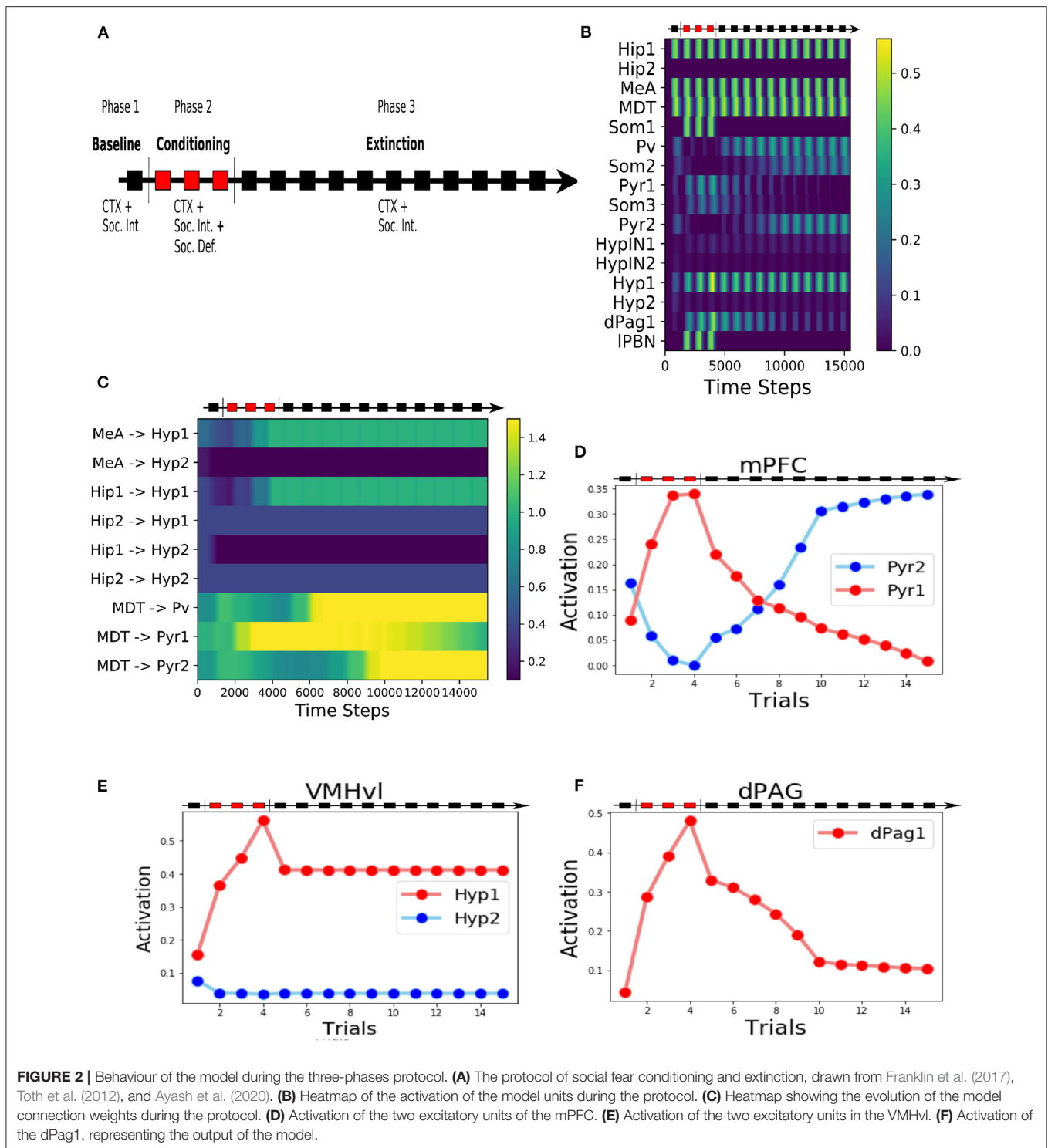
2.4. Simulation Protocol

In the model, the activation of the unit dPag1 represents social avoidance (Silva et al., 2013). To measure the baseline level of avoidance (Figure 2A), the model underwent a single trial (duration of 1 trial = 500 timesteps) of social interaction involving the presence of input to MeA (conspecific) and Hip1 (context). This was followed by three trials [as in Franklin et al. (2017)] of social conditioning (inputs to MeA and Hip1; defeat input to IPBN and Som1). Finally, these conditioning trials were followed by 11 trials [as in Toth et al. (2012)] of extinction (inputs to MeA and Hip1). This protocol is based on the foot-shock social conditioning and the sub-chronic defeat and aims to capture the common mechanisms underlying them. We cannot exclude that these 2 paradigms could induce plasticity at different sites. However, these 2 paradigms induce the same symptoms, measured in the three chamber apparatus (containing in one chamber the unfamiliar mouse to interact with) or in an open field with a cage containing the unfamiliar mouse (Toth et al., 2012; Franklin et al., 2017; Xu et al., 2019; Krzykowski et al., 2020) and are considered the best available to recapitulate the symptoms of SAD (Wang et al., 2020). Other protocols, such as the chronic defeat (Huhman, 2006) or the juvenile isolation (Yamamuro et al., 2020), were not taken into account because of the non-specific symptoms (that could indicate different fear pathways) and the difficulty to figure which input units to activate into the model in order to reproduce them. For the rest of this article, we will use the terms “sub-chronic social defeat” and “social fear conditioning” in an interchangeable way to refer to our social fear protocol.

3. RESULTS

3.1. The Overall Behavior of the Model

During the baseline (phase 1 of the protocol, Figure 2A), when the model was exposed only to the context (represented by the activation of the unit Hip1) and the conspecific (represented by



the activation of the MeA), the hypothalamic unit Hyp1 (the threat unit, see **Figures 1, 2B,E**) and the mPFC excitatory units Pyr1 and Pyr2 were mildly activated (**Figures 2B,D**). The dPag1 unit remained inactive (**Figure 2F**).

During the three trials of conditioning (Phase 2 of the protocol, **Figure 2A**) we activated the units Som1 in mPFC and IPBN, concomitantly with Hip1 and MeA. This induced

a reorganization of the weights at the cortical and subcortical levels (**Figure 2C**). In particular, at the subcortical level, the connection between MeA and the threat unit Hyp1 underwent LTP, while the connection between MeA and the interaction unit Hyp2 was depotentiated. At the cortical level, during conditioning Som1 deactivated Pyr2 and activated Pyr1, causing LTP between MDT and Pyr1 and LTD between MDT and Pyr2

(Figure 2C). Overall, the increase in Hyp1 activity in VMHvl and the decrease in Pyr2 activity in mPFC recruited the dPag1 unit (Figure 2F). The repeated presentation of the conditioned stimuli (consppecific and context) without the defeat stimuli (IPBN and Som1 activation) slowly extinguished the activation of the dPAG (phase 3 of the protocol, Figure 2A). In this model, the extinction was driven by potentiation/depotentialization in mPFC, while the connections in the VMHvl remained stable (Figure 2C).

In classical fear conditioning, three classes of neurons have been described on the basis of their responsivity to the conditioned stimulus: the fear neurons are active after conditioning but not after extinction, the extinction neurons behave in the opposite way, and the persistent neurons are turned on after conditioning and remain responsive to the stimulus even after extinction (Repa et al., 2001; Milad and Quirk, 2002; Herry et al., 2008; Santini et al., 2008; Amamo et al., 2010; An et al., 2012; Trouche et al., 2013). We looked for such populations in our model (Figure 2B). The unit Pyr1 in mPFC, that inhibits the activation of dPag1, behaved as a fear unit, while Pyr2 as an extinction unit (Figures 2B,D). We thus identified Pyr1 as a fear-ON population and Pyr2 as a fear-OFF population. Persistent neurons appeared in the VMHvl, represented by the unit Hyp1 (Figures 2B,E).

3.2. Alternative Model

We explored in a second model (Figure 3) the possibility that the VMHvl is also involved in the extinction.

In this model, the Hyp1 unit must receive inhibitory input from the lateral septum (LS), representing a biasing signal elicited by interaction with a conspecific (Figure 3). As expected, during the conditioning trials the excitatory connections reaching Hyp1 were subjected to LTP (Figure 4A), increasing their activity and promoting the downstream activation of dPag1 (Figures 4B,C). As the extinction trial progressed, we observed the depotentialization of the connections reaching Hyp1 and the potentiation of those reaching Hyp2, indicating that the extinction progress is occurring also inside the VMHvl (Figures 4A,B). As a result, the activity of dPag1 diminishes during extinction (Figure 4C).

We modulated the output of the LS to the VMHvl, in order to simulate an experiment of LS inhibition during extinction. We observed that, compared to the control condition (Supplementary Figures S1A, S1B), a reduction of LS output of 30% slows down extinction (Supplementary Figures S1C, S1D), and a reduction of 38% completely abolishes it (Supplementary Figures S1E, S1F).

3.3. Reproduction of Key Target Experiments

Given that the alternative models exhibited a behavior very similar to the main model, except for the presence or absence of the persistent neurons in the VMHvl, we continued all the following simulations using the latter model (Figure 1). To validate the model, we verified if it could qualitatively fit some relevant experiments from literature (see

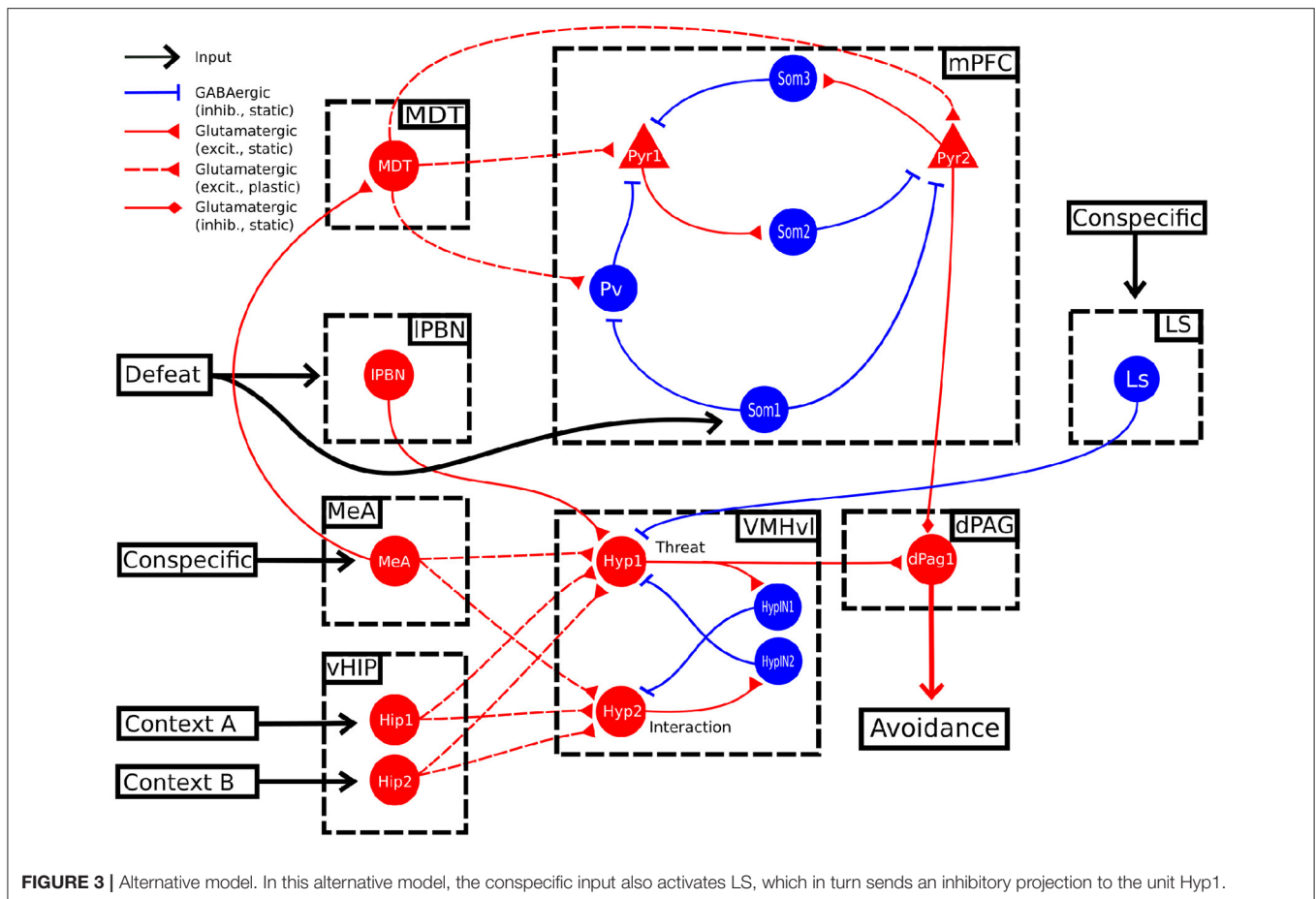
Supplementary Methods for the detailed protocol of experiment reproduction and **Supplementary Table S3**).

First, we reproduced the data obtained by Silva et al. (2013) who found that pharmacogenetic inhibition of the VMHvl results in a significant decrease of the time spent in defensive postures during the exposure to an aggressive conspecific. In particular, we verified that blockage of the activity of all populations in VMHvl after social fear conditioning reduces the dPAG activation induced by the conspecific exposure (Figure 5A).

Franklin et al. (2017) demonstrated that the selective inhibition of the prefrontal projections from the layer 5 to the dPAG mimicked social avoidance in mice that had not experienced previous defeat. In the model, when the system receives the inputs representing the conspecific and the context, if the output of Pyr2 is prevented to reach dPAG we observe a substantial activity in the dPag1 population (Figure 5B). Indeed, the dPag1 population is under the control of both the excitatory input from Hyp1 and the inhibitory input from Pyr2 (Figure 1). In the absence of Pyr2, the excitatory input prevails and dPag1 becomes active (Supplementary Figure S2).

The experiments of Xu et al. (2019) provide insights about the functions of some populations of mPFC neurons in social fear conditioning. The model reproduces two key experiments of the authors, namely the stimulation of parvalbumin+ interneurons and the inhibition of the somatostatin+ interneurons after the three trials of social defeat. We observed that the activation of the Pv unit or the inhibition of the Som1, Som2, and Som3 units reduced the dPag1 activation induced by the defeat protocol (Figure 5C). We also observed a phenomenon that has not been investigated by Xu et al. (2019): the manipulation of the mPFC interneurons during the first extinction trial accelerated the subsequent acquisition of fear extinction, which thus was completed three trials before the control (Figure 5C). In particular, both the Pv stimulation and the inhibition of the Som units facilitated the potentiation of the connection between MDT and Pyr2 and the depotentialization of the connection between MDT and Pyr1 (Supplementary Figures S3A, S3B). This led to a more powerful activation of the Pyr2 fear-OFF population, compared to the non manipulated control, and to a consequent deactivation of the Pyr1 fear-ON population (Figures 5D,E).

Krzywkowski et al. (2020) studied the activation of different VMHvl populations during the aversive social encounter and when the mice were re-exposed to the same context without the opponent. Their Defeat+ population corresponds to our threat unit Hyp1, which is active during, but not before, conditioning (Figure 5F; see also Figure 2B). Re-exposure to the same context of the defeat, implemented through the activation of the Hip1 input without the conspecific or defeat inputs (MeA, IPBN, Som1 populations remained silent), recruited Hyp1 similarly to what done by the defeat. Conversely, exposure to a different context (Hip2) did not induce any Hyp1 activity (Figure 5F). This reproduces the findings of Krzywkowski et al. (2020) and implies that the connection between Hip1 and Hyp1 populations encodes the memory related to the contextual social fear conditioning. Indeed, we observed that in context 1 (Hip1 active), the activation of Hyp1 is able to drive dPag1 activation, even in the absence of the conspecific (Figure 5G).



3.4. Predictions

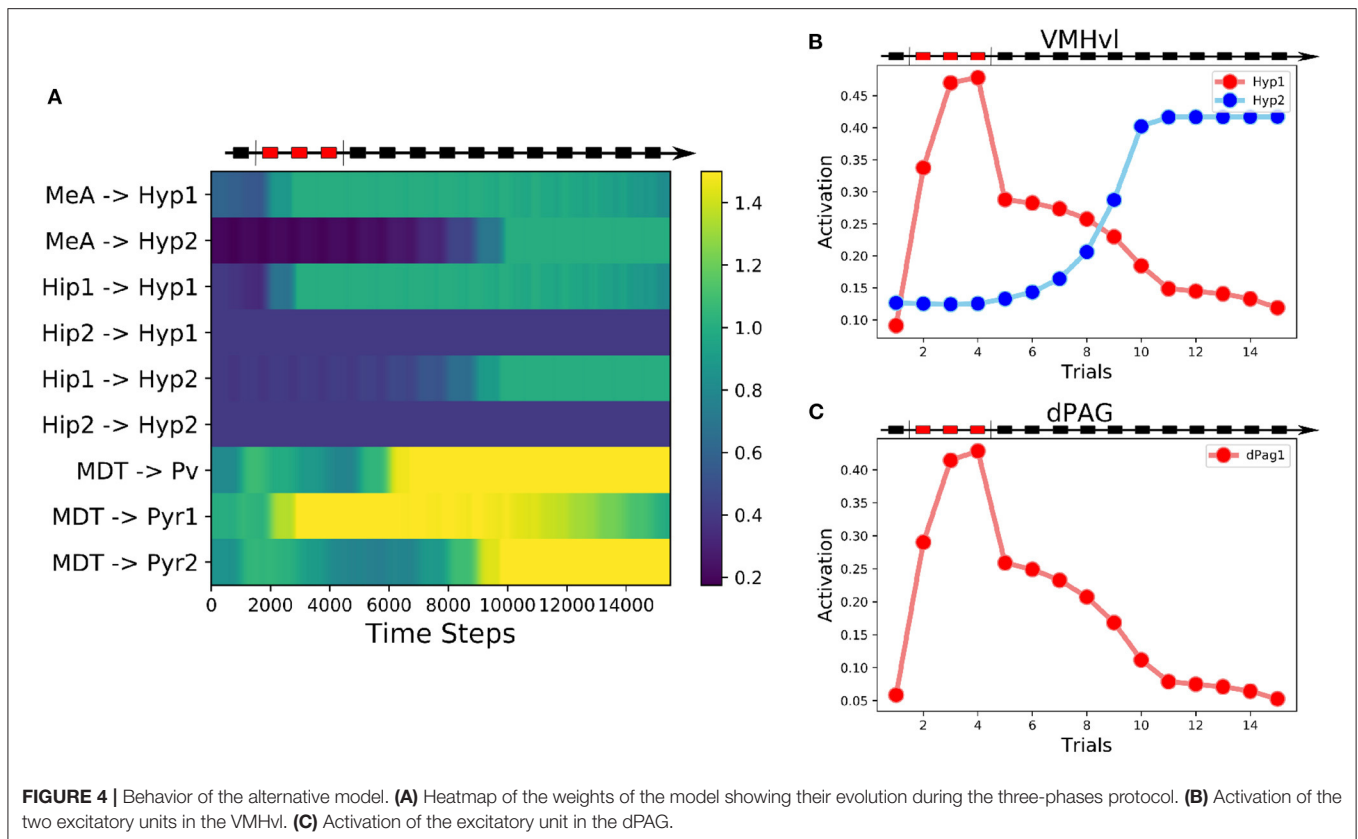
The manipulations of the model that simulated plausible but never attempted *in vivo* experiments allowed us to make four specific testable predictions. First, we blocked synaptic plasticity inside the VMHvl area during the conditioning trials, in order to simulate an artificial impairment of LTP. As shown in **Figures 6A,B**, during the conditioning trials the Hyp1 and dPag1 units increase their excitation (although to a lesser extent than controls). However, their activation drastically decreases during the first trial of extinction (trial 5) and remains low in the last trials, thus indicating that the plasticity in VMHvl is mandatory for the acquisition of social fear memory.

The second prediction regards the role of the IPBN. We blocked its activity during the conditioning trials to simulate a pharmacological or optogenetic inhibition. Without the upstream input from IPBN, the hypothalamic fear population Hyp1 remains silent during the defeat and the extinction phases (**Figure 6C**). As a consequence, the connections between the input areas (MeA and Hip) and the VMHvl do not undergo synaptic plasticity (**Supplementary Figure S4A**) and dPag1, despite being slightly stimulated during the conditioning trials due to the weakening of the prefrontal inhibition (**Supplementary Figure S4B**), drastically decreases its levels of

activation during extinction, indicating that the conditioning process is strongly impaired (**Figure 6D**).

The third prediction concerns what happens if the prefrontal Pyr2 population projecting to dPAG1 is stimulated after social conditioning. This corresponds to an experiment of pharmacological or optogenetic stimulation, executed during the first trial of extinction, of prefrontal neurons appropriately isolated from other PFC pyramidal neurons through retrograde labeling of the dPAG neurons. As shown in **Figure 6E**, the Pyr2 stimulation during the first extinction trial induces a rapid decrease of the levels of dPag1 activation compared to the control, which results in a one-trial extinction. This is due to the fact that in the mPFC (**Figure 6F** and **Supplementary Figure S5**) the strong activation of Pyr2 in the first extinction trial shuts down the competing Pyr1 population. The effect is an enhancement of LTP of the connection MDT-Pyr2 and the LTD of the connection MDT-Pyr1. This rapid reorganization in the mPFC circuit results in the dPag1 inhibition observed in **Figure 6E**.

The last prediction is linked to the results of Herry and Garcia (2002), who showed in *in vivo* experiments with mice undergoing auditory fear conditioning that the manipulation of the MDT-mPFC connections influences extinction. In particular, we investigated whether a similar outcome could be predicted for social fear conditioning. We, thus, exposed



the model to a protocol that mimics the effect of either high-frequency stimulation (HFS) or low-frequency stimulation (LFS) of the MDT, administered after conditioning (see **Supplementary Methods** for the details), to observe the effect on extinction. As shown in **Figure 6G**, HFS and LFS, respectively, accelerate or completely impair extinction. The cause of this is that HFS induces the potentiation of the MDT-Pyr2 connection. The simultaneous potentiation of the MDT-Pv connection (see **Supplementary Materials**), combined with the competition between Pyr1 and Pyr2, shuts down Pyr1 (**Figure 6H**) and this leads to a depotentiation of MDT-Pyr1. As a result, extinction is accelerated. LFS has the opposite effect: it induces a depotentiation of the MDT-Pyr2 and the MDT-Pv connections and, because of the resulting reduced inhibition on Pyr1, a potentiation of the MDT-Pyr1 connection (**Supplementary Figure S7D**). This results in an increase of the Pyr1 activity compared to the control, and in a full inhibition of the Pyr2 activity (**Figure 6I**). As a result, extinction is impaired.

3.5. Sensitivity Analysis

In order to examine the robustness of the model, we tested how changing the weights impacts on the reproduction of the conditioning, extinction and of the key target experiments (see **Supplementary Methods**, **Supplementary Figure S6** and **Supplementary Table S4**). The model is very robust to the perturbation of almost all parameters. Notable exceptions are the weight of the connections between MDT and Pyr1 and between

IPBN to Hyp1, that can be changed, respectively, in a range of 30 and 55%. This is reasonable, being these, respectively, the connection that conveys the defeat to the VMHvl and the connection mainly responsible for the reorganization of the mPFC after the conditioning. Finally, we observed that the HypIN1 neuron and the connection from Som1 to Pyr2 in mPFC can be eliminated from the model without affecting the outcome of the simulations.

4. DISCUSSION

4.1. Contribution of the Model

In the field of Pavlovian cued and contextual fear conditioning, theoretical models have played an important role to integrate information and offer new perspectives on the experimental findings. This effort has produced novel predictions, has highlighted the knowledge gaps, and has driven further experimental research that has been used to build more sophisticated and biologically grounded models (Burgos and Murillo-Rodríguez, 2007; Mannella et al., 2008; Li et al., 2009, 2016; Krasne et al., 2011; Anastasio, 2013; John et al., 2013; Moustafa et al., 2013; Carrere and Alexandre, 2015; Kim et al., 2016; Nair et al., 2016; Oliva et al., 2018; Bennett et al., 2019; Mattera et al., 2020). We believe that social fear research could also benefit from this approach. Here, we have thus proposed a computational model of social avoidance that, to the best of our knowledge, is the first to address this complex

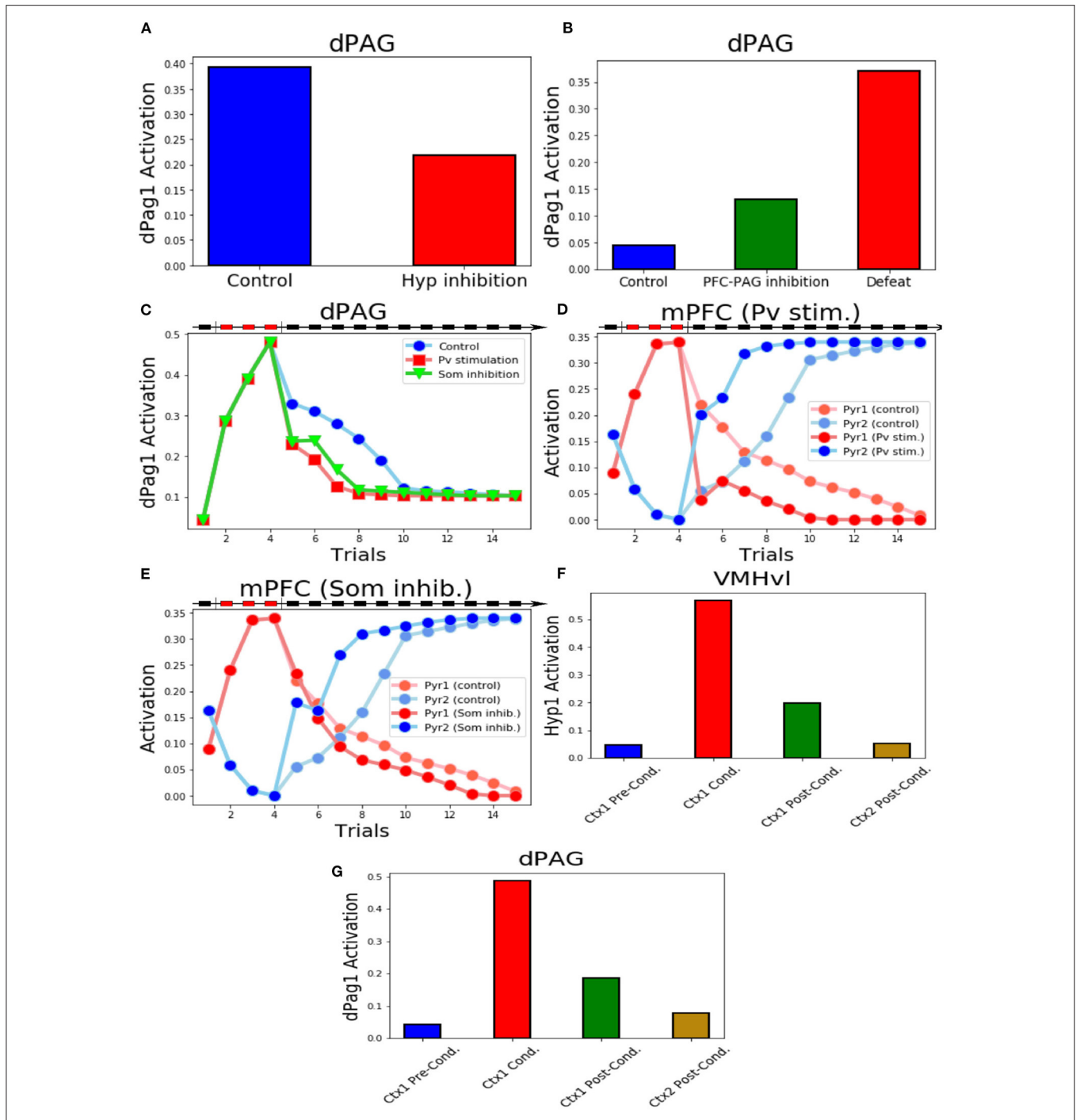


FIGURE 5 | Key experiments reproduced by the model. **(A)** Inhibition of the excitatory units in VMHvl reduces the activity of dPag1. **(B)** Inhibition of the output of Pyr2 induces the activation of dPag1 in the system exposed to the conspecific and context input, even if it was never applied a protocol of conditioning. We show, for comparison, the dPag1 activation in a system subjected to the defeat protocol, where the unit Pyr2 is not manipulated. **(C)** The activation of Pv or the inhibition of Som1, Som2, and Som3 during the first trial of extinction, respectively, slow down or faster the extinction process. **(D)** Behaviour of the excitatory units of the mPFC when Pv is stimulated during the first trial of extinction and respective controls **(E)** Behaviour of the excitatory units of the mPFC when Som1, Som2 and Som3 are inhibited during the first trial of extinction and respective controls. **(F,G)** Activation of the VMHvl unit Hyp1. **(F)** and the dPag1 **(G)** in context 1 before, during and after conditioning and in context 2 after conditioning.

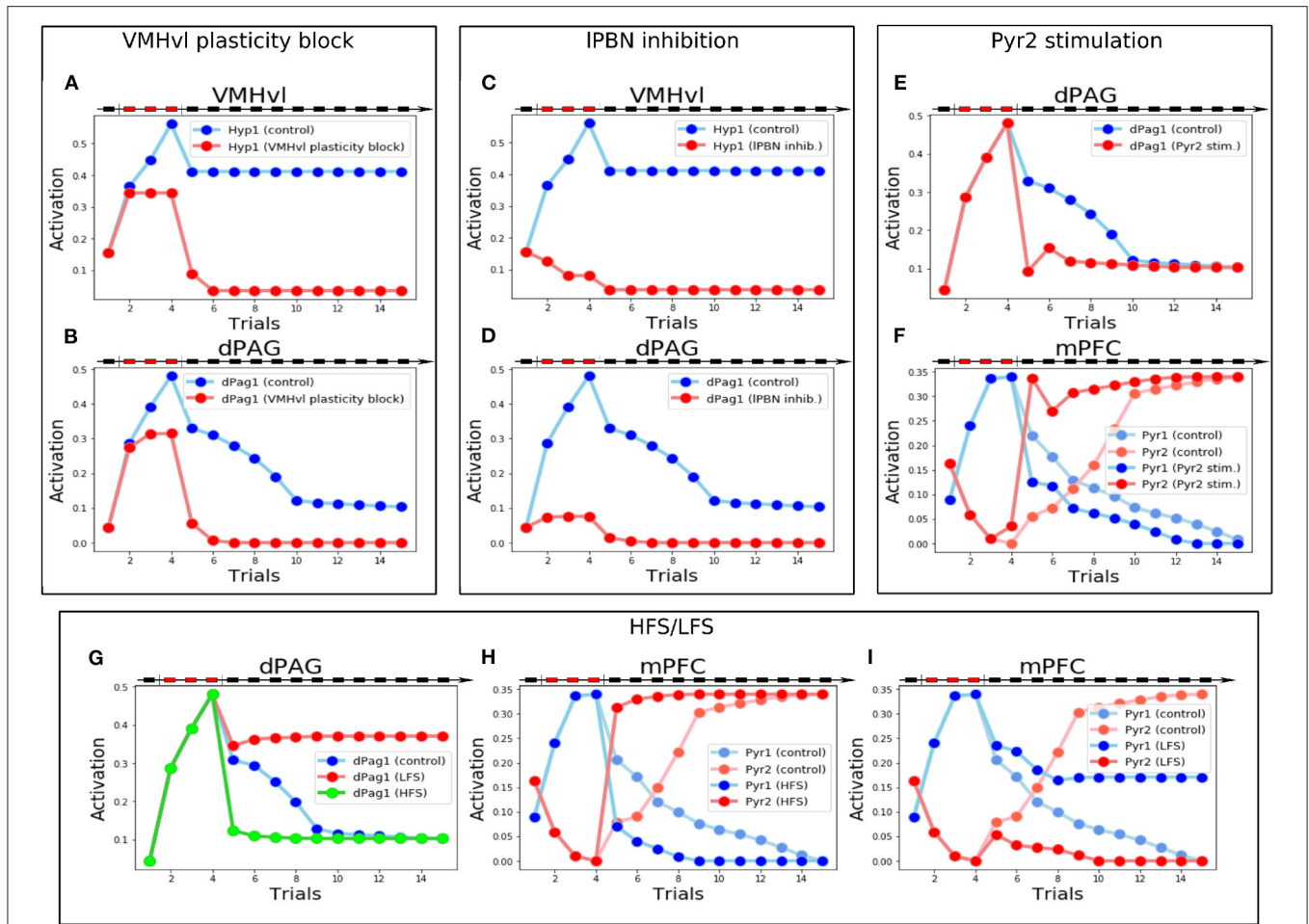


FIGURE 6 | Manipulations of the model that generate testable predictions. **(A,B)** Activity of Hyp1 **(A)** and dPag1 **(B)**: with respect to the control, when the VMHvl plasticity is blocked during conditioning the memory of conditioning is fully lost. **(C,D)** Activity of Hyp1 **(C)** and dPag1 **(D)**: with respect to the control, when the IPBN activity is blocked during conditioning Hyp1 remains silent during the defeat and the extinction phases and dPag1 does not exhibit conditioning. **(E,F)** Stimulation of Pyr2 during the first trial of extinction: compared to the control, this changes the plasticity and activations in the mPFC **(E)** resulting in a fast deactivation of dPAG **(E)** and extinction. **(G–I)** Effects of MDT stimulation through a protocol mimicking HFS or LFS: with respect to the control, the dPAG1 undergoes, respectively, a rapid drop or a lack of the decrease exhibited by the control **(G)**; in the case of HFS, after conditioning the mPFC Pyr1 activity decreases faster and Pyr2 remains inhibited **(H)**; in the case of LFS, the mPFC Pyr1 does not decrease completely with extinction, and Pyr2 activity increases faster than the control **(I)**.

phenomenon. A strong advantage of the model is that it ties together fragmented information regarding different brain areas involved in conditioning and extinction of social avoidance. In particular, social avoidance is mainly investigated with a focus on the VMHvl (Silva et al., 2013; Sakurai et al., 2016; Wang et al., 2019; Krzywkowski et al., 2020) or the mPFC (Franklin et al., 2017; Xu et al., 2019). However, how these structures interact and contribute to the phenomenon is a problem that has not yet been addressed experimentally.

In the proposed model, the threat neural population Hyp1 in VMHvl represents the Defeat+ population of Krzywkowski et al. (2020), shown to be activated by aversive social encounter with a conspecific. The activation of this population drives the firing of the dPAG neural population promoting the behavioral manifestations of fear (Figures 1, 2E,F). In agreement with this, it has been shown that the optogenetic activation of social fear

neurons in VMHvl, marked by cFos expression after defeat, is sufficient to induce fear in mice (Sakurai et al., 2016). On the other hand, if these neurons are inhibited during an encounter with an aggressive conspecific, the defensive behavior is reduced (Silva et al. 2013; Figure 5A). The model proposes that the mPFC acts as a brake on subcortical structures, in particular by dampening the dPAG activation and thus preventing social avoidance expression. As shown by Franklin et al. (2017), and reproduced by the model (Figure 5B), the silencing of the connection between the mPFC and the dPAG induces fear in the undefeated mice. This fits with the finding for which a lesion of the mPFC reduces social interaction (Murray et al., 2015). In the model depicted in Figures 1, 3, the output unit dPag1 receives two distinct types of signals: the excitatory Hyp1 signal from the VMHvl and the inhibitory Pyr2 signal from the mPFC. In the baseline condition, before conditioning the presence of

a conspecific activates only loosely the Hyp1 unit (**Figure 2B**), and the concomitant activation of Pyr2 is enough to turn off dPag1. In order to elicit social avoidance, the Hyp1 population must be more strongly activated by a conspecific [as artificially done by Sakurai et al. (2016)] or the Pyr2 population must be less responsive [as artificially done by Murray et al. (2015), Franklin et al. (2017)]. Social defeat triggers a reorganization of the synaptic weights in VMHvl and in the mPFC such that both conditions are met (**Figure 2C**). During extinction, the relative levels of the mPFC Pyr2, but not Hyp1, return to the naive condition the main model (**Figure 2**). Instead, the excitation of both Hyp1 and Pyr2 are restored in the alternative model (**Figure 4A**).

On the basis of the overall picture emerging from these results, the model suggests that a possible rescue strategy for the social avoidance disease (SAD) is to restore, or potentiate, the brake function of the mPFC on the subcortical circuit formed by the threat population Hyp1 of the VMHvl and the dPAG. In support of this view, imaging studies have shown an altered activation of the mPFC in SAD patients (Labuschagne et al., 2012; Stein, 2015). Moreover, remarkably in SAD the functional connectivity between the mPFC and subcortical structures is significantly decreased (Prater et al., 2013; Gold et al., 2016). This pathological condition is reproduced in the murine model of sub-chronic social defeat, where the local field potential coherence between the mPFC and dPAG is reduced (Franklin et al., 2017). Conversely, symptom improvement after cognitive-behavioral therapy (CBT) is predicted by an increase of mPFC-subcortical connectivity (Klumpp et al., 2014).

From a theoretical perspective, what is the need for a social animal to have both an “accelerator” (VMHvl) and a “brake” (mPFC) of the social avoidance? The prefrontal cortex is an area that underlies executive functions (working memory, inhibition, flexibility) to support goal-directed behavior (Uylings et al., 2003; Kesner and Churchwell, 2011; Diamond, 2013). It has been recently highlighted that the prefrontal regulation of the basolateral amygdala in auditory and contextual fear conditioning should be considered in this framework (Gonzalez and Fanselow, 2020). In this view, it seems plausible that, when some particular goals are pursued, an animal could get advantaged to override a subcortical impulse to flight through the inhibitory cortical control of sub-cortical areas.

Besides executive functions, another factor that could influence social avoidance is social rank. Notably, social rank is encoded in the mPFC and shares some pathways with those of social fear. In particular, the strength of the connection between the MDT and the mPFC, and the excitability in layer 5 pyramidal neurons, determine the animal hierarchical status (Wang et al., 2011; Zhou et al., 2017). As expected from this, it has been observed that the animal rank predicts the susceptibility to a social defeat paradigm (Larrieu et al., 2017).

4.2. Anatomical and Functional Considerations

In the construction of the model mPFC, we derived the neural populations and the connections between them from

both experimental findings on the activation/deactivation of different neurons during social conditioning, and from some hypotheses formulated on the basis of the analysis of the literature (see **Supplementary Table S1**). Here we thus speculate on the possible localization of these populations in the mPFC anatomy.

We propose that the brain correspondent of the model neural population Pyr2 is located in layer 5, having been shown that the descending projections from mPFC to dPAG originate from it (Franklin et al., 2017). The mPFC lacks a thalamorecipient layer and the MDT input reaches both layer 5 and layer 2/3 (Collins et al., 2018). It is thus possible MDT reaches Pyr2 through a direct connection. The location of the Pyr1 population is more difficult to infer. These neurons are specifically activated by social fear conditioning, as shown with c-Fos labeling and calcium imaging (Xu et al., 2019). Remarkably, Xu et al. (2019) reported a significant activation of c-Fos in prelimbic, but not infralimbic cortex. On the other hand, Hinwood et al. (2011) previously showed that a protocol of repeated social defeat recruits Δ FosB positive neurons in layer 2/3 of the infralimbic cortex. Although the protocols are different, this suggests the possibility that the neurons described by Xu et al. (2019) belong to the higher layers of the mPFC. Unfortunately, there is paucity of experimental data on the involvement of prelimbic and infralimbic cortex in social avoidance after a sub-chronic social defeat or a social fear conditioning protocol. Future work could aim to investigate if the dynamics of the neural populations reproduced in the mPFC of our model could be related to the behaviour of prelimbic and infralimbic populations.

Following the aforementioned experimental findings, the parvalbumin+ population (represented in our model by the unit Pv, **Figure 1**), that after social fear conditioning becomes hyperactive and depresses the pyramidal neurons firing during conspecific encounters (Xu et al., 2019), is the one supporting feedforward inhibition in layer 2/3 after MDT stimulation (Delevich et al., 2015). We could also try to locate the interneuronal somatostatin+ populations represented in our model by Som2 and Som3. Som2, which in our model is activated by Pyr1 and inhibits Pyr2, could represent a translaminar inhibitory circuit from layer 2/3 to layer 5. Optogenetic mapping suggests that somatostatin+ interneurons that receive from layer 2/3 and project to layer 5 are located within layer 5 (Naka and Adesnik, 2016). On the other hand, cases have been described of somatostatin+ cells from layer 5 that project back to pyramidal neurons in layer 2/3 (Kapfer et al., 2007; Nigro et al., 2018), as could be for the population represented by our Som3 unit.

One of the main mechanisms of social avoidance that still lacks an explanation supported by empirical evidence is the “defeat signal” that triggers the functional reorganization of the mPFC. We observed (see Materials and Methods) that the mPFC disinhibitory circuit comprising somatostatin+ and parvalbumin+ interneurons upstream of the pyramidal neurons is shared by auditory and social fear conditioning (Courtin et al., 2014; Xu et al., 2019; Cummings and Clem, 2020). A similar disinhibitory mechanism is present in the auditory cortex, where layer 1 interneurons negatively regulate layer 2/3 parvalbumin+ interneurons, which in turn inhibit excitatory layer 2/3 neurons (Letzkus et al., 2011). Auditory fear conditioning requires that

the auditory cortex is disinhibited through the activation of layer 1 interneurons at the moment of fear learning. The key element that recruits layer 1 interneurons is the nicotine acetylcholine (ACh) receptors (Letzkus et al., 2011). It has been shown that pairing a tone with the stimulation of the nucleus basalis, that releases ACh to the cortex, causes an unbalanced excitation/inhibition through the induction of synaptic plasticity (Froemke et al., 2007). Thus, in certain conditions, ACh acts as a negative reward signal that remodels cortical synapses, as also observed using a go/no-go paradigm (Hangya et al., 2015). In the mPFC, ACh from the nucleus basalis mainly targets somatostatin+ interneurons (Sun et al., 2019). On the basis of this evidence, and in line with an idea recently discussed in the literature (Wang et al., 2020), we suggest that the defeat signal recruiting somatostatin+ interneurons in mPFC (Som 1 in our model, see **Figure 1**) during social fear conditioning could be represented by ACh. We propose that a possible experiment usable to investigate a coincident activation of the central cholinergic neurons during the social fear protocol could be the *in vivo* electrophysiological recordings in the nucleus basalis, paired with an identification of the ACh-releasing neurons through optogenetic tagging (Lima et al., 2009; Hangya et al., 2015). If an activation of cholinergic neurons is recorded, a direct way to demonstrate the role of ACh in the mPFC in fear conditioning could be the *in loco* administration of antagonists of ACh receptors (Gu et al., 2020). Finally, cell specific knockout of the receptors (Hernandez et al., 2014; Gu et al., 2020) could be used to locate the subtype of neurons on which ACh acts to produce the fear conditioning.

In the alternative model (**Figure 3**) we proposed the existence of an inhibitory input on the threat neurons in the VMHvl. When this modification is inserted in the model, the persistent population in VMHvl is substituted by an extinction population. We speculate that this possible inhibitory afferent is represented by the LS, a key area involved in social interaction and social fear pathways. Indeed, LS sends inhibitory projections to the VMHvl to control aggressive behaviors against conspecifics and, when optogenetically stimulated, promotes social investigation (Wong et al., 2016). Moreover, oxytocin (OXT), which has pro-social effects and has been proposed as a therapy for SAD (Guastella et al., 2009; Labuschagne et al., 2010), when directly administered to the LS of mice facilitates social fear extinction (Zoicas et al., 2014; Menon et al., 2018). This suggests the idea that neural projections from the main oxytocinergic hypothalamic nuclei, such as the paraventricular nucleus and particularly the supraoptic nucleus, are important upstream regulators of the levels of inhibition provided by the LS to the VMHvl. It is worth noting that social fear conditioning induces a reduction of the levels of oxytocin release inside the LS, with a mechanism still to be explored (Zoicas et al., 2014). This is possibly due to the alterations induced by social fear conditioning inside the main oxytocinergic nuclei. Indeed, it has been observed that stress has a detrimental impact on the PVT, inducing a reduction of the expression of serotonin 5-HT1A receptors (Florez et al., 2017). Future studies are required to elucidate the involvement of this pathway on social fear conditioning and extinction, especially in relation to the

TABLE 1 | Testable predictions of the model.

Prediction	Possible test strategy
Presence of fear and extinction neurons in mPFC; presence of persistent neurons in VMHvl	<i>In vivo</i> recordings in mPFC and VMHvl during a conspecific exposure, done before conditioning, after conditioning and after extinction
Blocking the plasticity in VMHvl does not alter the social fear responses but impairs the formation of the fear memory	Stereotactic injection of AP5 or other plasticity inhibitors in VMHvl before social fear conditioning
IPBN nucleus conveys the nociceptive information necessary to the social conditioning to the VMHvl	Stereotactic injection of muscimol in IPBN before conditioning
Layer 5 mPFC pyramidal neurons stimulation during a social encounter boosts the extinction of social fear	Administration of drugs that enhance the excitability of layer 5 neurons, like LSD, before extinction
HFS and LFS of MDT, respectively, increase or completely abolish extinction	<i>In vivo</i> HFS of LFS of MDT with electrodes before extinction

presence or absence of persistent neurons in VMHvl. As we observed in **Supplementary Figure S1**, in the alternative model the inhibition of LS slows down extinction and, if sufficiently high, inhibits it. This experiment, once performed in mice, would allow the discrimination which model is more accurate in the description of murine social avoidance pathways.

4.3. Predictions of the Model

Our model produced various predictions (**Table 1**), some derived from specific manipulations of the model aimed to simulate possible new experiments and others deduced from the reproduction of data already present in the literature.

We observed that, similarly to what happens in auditory and contextual fear conditioning, the model exhibits the three classes of fear, extinction, and persistent neural populations. In particular, on the basis of our simulations, we expect to find fear and extinction neurons in the mPFC, corresponding to, respectively, the neural populations Pyr1 and Pyr2 (**Figure 2B**). Pyr1 corresponds to the population of neurons activated by the conspecific exposure in the socially conditioned mice (Xu et al., 2019). Unfortunately, the authors of the research did not investigate the effect of the extinction on the firing of this population, and it is thus not possible to label them as fear or persistent neurons. On the other hand, the simulations predict that VMHvl contains persistent neurons (**Figure 2B**), unless a biasing signal inhibits the threat population during social encounters, such as LS inhibition in the alternative model (**Figure 2**). Persistent neurons have an important implication in the context of SAD. Indeed, it has been observed that the existence of these neurons in the amygdala after auditory fear conditioning prevents the return to a pre-conditioning state even after extinction. For this reason, the fear memory can be easily reinstated (Maren and Holmes, 2016). Moreover, if the subcortical areas, but not the mPFC, contain persistent neurons, this would imply that the extinction is entirely entrusted to the cortex, that thanks to its flexibility is able to return to the naive condition.

After blocking the plasticity of the VMHvl during the conditioning procedure the model is still able to acutely evoke the avoidance response. However, the acquisition of social avoidance is drastically impaired and the model fails to elicit the fear response in the following trials (**Figures 6A,B**). This result has analogies with the results of the experiments of Silva et al. (2016b), where it was seen that silencing the dorsomedial part of the VMH is sufficient to impair the acquisition of predatory fear memory. We speculate that the same process occurs inside the VMHvl during conditioning and that the plasticity taking place in this structure is fundamental for the integration of contextual, nociceptive, and social information, and for the acquisition of social avoidance.

An important afferent of the model is the IPBN, which delivers pain signals to different subcortical structures to drive appropriate responses (Chiang et al., 2019, 2020). In particular, the activation of the projections to the VMHvl induces escape (Chiang et al., 2020). Moreover, IPBN input to the amygdala is necessary and sufficient to induce fear memories (Sato et al., 2015; Chiang et al., 2020). We hypothesized that the unconditioned signal necessary for the conditioning of the subcortical compartment arises from this area. For this reason, we predict that its inhibition would result in a failed conditioning (**Figures 6C,D**).

When we stimulated the unit Pyr2 during the first trial of extinction, we observed a fast one-shot abolishment of social avoidance (**Figure 6E**). This is especially interesting in the context of a possible pharmacological treatment of SAD. In the light of the aforementioned speculations regarding the localization of the Pyr2 population in the mPFC, we suggest that a therapy of stimulation of the layer 5 pyramidal neurons projecting to the dPAG during the exposure to the fearful stimuli could induce a fast therapeutic recovery. In this perspective, a possible pharmacologic agent would be LSD, an agonist of 5-HT_{2A} receptors (Nichols, 2004). Indeed the activation of these receptors enhances glutamatergic currents in the apical dendrites of layer 5 pyramidal neurons in mPFC (Aghajanian and Marek, 1997, 1999). In addition, LSD has been recently shown to promote social behavior in rodents through the activation of the prefrontal excitatory neurons. Optogenetic inhibition of the mPFC, on the other hand, not only blocks the pro-social effect of LSD, but also induces social avoidance (De Gregorio et al., 2021; this is in line with the simulation shown in **Figure 2B**, where we blocked the output of the mPFC). It is worth noting that an LSD-induced enhancement of sociability has also been reported in humans (Dolder et al., 2016; Duerler et al., 2020). Remarkably, LSD has never been investigated in defeated mice but our simulations and our theoretical framework support the hypothesis of a beneficial effect on social avoidance.

The last prediction regards how an MDT stimulation inducing plasticity to the connections with the mPFC influence extinction. We were interested in this simulation for 2 reasons. First, this manipulation determines the success of extinction to auditory conditioning (Herry and Garcia, 2002). We reasoned that a similar outcome would suggest that auditory and social fear share some pathways or mechanisms at the level of mPFC. Second, the strength of the MDT-mPFC synapse determines the social

rank of the rodents and influences the winning individual in the tube test (Zhou et al., 2017). The outcome of our simulations (**Figures 6G–I**) was very similar to those obtained by Herry and Garcia (2002), where HFS and LFS, respectively, increase or completely abolish extinction. Unfortunately, the influence of rank on social fear is still poorly investigated (Larrieu et al., 2017), but our results raise the possibility that an increase in the social rank scale would have a positive effect on the social fear extinction.

4.4. Limitations of the Model

We now discuss some limitations of the model that should be addressed in future work. The first limitation regards the connectivity of the mPFC with other structures beyond dPAG. For example, layer 2/3 excitatory neurons (possibly our Pyr1 population) send projections to the Nucleus Accumbens (NAcc; Franklin et al. 2017) and this structure has been shown to be involved in the expression, but not the acquisition, of social fear (Lockett et al., 2012). This suggests that NAcc and its downstream areas are involved in the fear-promoting pathway of the Pyr1 population. Besides NAcc, it has been recently shown that the pathway connecting the mPFC to the paraventricular nucleus of the thalamus (PVT) is also part of the sociability circuit and that the inhibition of the mPFC to PVT pathway leads to a reduced social preference (Yamamuro et al., 2020). Here, we focused on a minimal circuit able to explain the principal experimental findings and to produce predictions without constructing an overly complex architecture. For the same reason, we have not taken into account other neural structures involved in parallel or complementary circuits of social fear behavior such as the bed nucleus of the stria terminalis (Clauss et al., 2019) and the dorsal pre-mammillary nucleus (Motta et al., 2009). These structures were not necessary to explain the experiments targeted here but they should be taken into account in future models.

Another limitation of this work regards the kind of modeling approach employed in this work. In particular, the firing rate neural model is capable to simulate the long-range projections and the excitatory/inhibitory relationship between the different brain areas, but says little about their internal computation and cannot reproduce some complex phenomena. For instance, a recent work of the group of Xu (Liu et al., 2020) revealed unexpected complexity in the social control exerted by the mPFC. In particular, the authors observed that activation of parvalbumin+, but also somatostatin+ interneurons at the gamma frequency, results in a pro-social effect. On the other hand, a nonrhythmic activation of somatostatin+ interneurons has the opposite outcome. Investigating these phenomena would give a more complete understanding of the mPFC control of social fear, but would also require spiking neural networks to simulate brain oscillations (e.g., Wei et al. 2016; Capone et al. 2019).

Notwithstanding these needed developments, as happened for the classical fear conditioning studies centered on the amygdala, the results presented here should have demonstrated the utility of studying social avoidance through top-down models. These models are indeed capable of systematising and conceptualising the experimental findings on the involved

neural areas and plasticity mechanisms and also of paving the way to additional studies using more detailed bottom-up models, as it is done in the study of auditory fear conditioning (Nair et al., 2016).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

VA, AM, and GB: conceived the model, conducted the simulation, analyzed the results, and wrote and reviewed this

article. All authors contributed to the article and approved the submitted version.

FUNDING

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 713010 (GOAL-Robots-Goal-based Open-ended Autonomous Learning Robots).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnsys.2022.841085/full#supplementary-material>

REFERENCES

- Aghajanian, G., and Marek, G. (1997). Serotonin induces excitatory postsynaptic potentials in apical dendrites of neocortical pyramidal cells. *Neuropharmacology* 36, 589–599.
- Aghajanian, G. K., and Marek, G. J. (1999). Serotonin, via 5-HT_{2A} receptors, increases epsps in layer V pyramidal cells of prefrontal cortex by an asynchronous mode of glutamate release. *Brain Res.* 825, 161–171.
- Amano, T., Unal, C. T., and Paré, D. (2010). Synaptic correlates of fear extinction in the amygdala. *Nat. Neurosci.* 13, 489. doi: 10.1038/nn.2499
- An, B., Hong, I., and Choi, S. (2012). Long-term neural correlates of reversible fear learning in the lateral amygdala. *J. Neurosci.* 32, 16845–16856. doi: 10.1523/JNEUROSCI.3017-12.2012
- Anastasio, T. (2013). Computational search for hypotheses concerning the endocannabinoid contribution to the extinction of fear conditioning. *Front. Comput. Neurosci.* 7, 74. doi: 10.3389/fncom.2013.00074
- Arnold Anteraper, S., Triantafyllou, C., Sawyer, A. T., Hofmann, S. G., Gabrieli, J. D., and Whitfield-Gabrieli, S. (2014). Hyper-connectivity of subcortical resting-state networks in social anxiety disorder. *Brain Connect.* 4, 81–90. doi: 10.1089/brain.2013.0180
- Ayash, S., Schmitt, U., and Müller, M. B. (2020). Chronic social defeat-induced social avoidance as a proxy of stress resilience in mice involves conditioned learning. *J. Psychiatric Res.* 120, 64–71. doi: 10.1016/j.jpsychires.2019.10.001
- Bennett, M. R., Farnell, L., Gibson, W. G., and Lagopoulos, J. (2019). A model of amygdala function following plastic changes at specific synapses during extinction. *Neurobiol. Stress* 10, 100159. doi: 10.1016/j.ynstr.2019.100159
- Berkowitz, R. L., Coplan, J. D., Reddy, D. P., and Gorman, J. M. (2007). The human dimension: how the prefrontal cortex modulates the subcortical fear response. *Rev. Neurosci.* 18, 191–208. doi: 10.1515/revneuro.2007.18.3-4.191
- Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* 2, 32–48.
- Biro, L., Toth, M., Sipos, E., Bruzsik, B., Tulogdi, A., Bendahan, S., et al. (2017). Structural and functional alterations in the prefrontal cortex after post-weaning social isolation: relationship with species-typical and deviant aggression. *Brain Struct. Funct.* 222, 1861–1875. doi: 10.1007/s00429-016-1312-z
- Burgos, J. E., and Murillo-Rodríguez, E. (2007). Neural-network simulations of two context-dependence phenomena. *Behav. Process.* 75, 242–249. doi: 10.1016/j.beproc.2007.02.003
- Canteras, N., Simerly, R., and Swanson, L. (1995). Organization of projections from the medial nucleus of the amygdala: a phal study in the rat. *J. Comparat. Neurol.* 360, 213–245.
- Capone, C., Pastorelli, E., Golosio, B., and Paolucci, P. S. (2019). Sleep-like slow oscillations improve visual classification through synaptic homeostasis and memory association in a thalamo-cortical model. *Sci. Rep.* 9, 1–11. doi: 10.1038/s41598-019-45525-0
- Carrere, M., and Alexandre, F. (2015). A pavlovian model of the amygdala and its influence within the medial temporal lobe. *Front. Syst. Neurosci.* 9, 41. doi: 10.3389/fnsys.2015.00041
- Chang, C.-H., and Gean, P.-W. (2019). The ventral hippocampus controls stress-provoked impulsive aggression through the ventromedial hypothalamus in post-weaning social isolation mice. *Cell Rep.* 28, 1195–1205. doi: 10.1016/j.celrep.2019.07.005
- Chen, P., and Hong, W. (2018). Neural circuit mechanisms of social behavior. *Neuron* 98, 16–30.
- Chiang, M. C., Bowen, A., Schier, L. A., Tupone, D., Uddin, O., and Heinricher, M. M. (2019). Parabrachial complex: a hub for pain and aversion. *J. Neurosci.* 39, 8225–8230. doi: 10.1523/JNEUROSCI.1162-19.2019
- Chiang, M. C., Nguyen, E. K., Canto-Bustos, M., Papale, A. E., Oswald, A.-M. M., and Ross, S. E. (2020). Divergent neural pathways emanating from the lateral parabrachial nucleus mediate distinct components of the pain response. *Neuron* 106, 927–939.e5. doi: 10.1016/j.neuron.2020.03.014
- Choi, G. B., Dong, H.-W., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Swanson, L. W., et al. (2005). Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. *Neuron* 46, 647–660. doi: 10.1016/j.neuron.2005.04.011
- Christoffel, D. J., Golden, S. A., Walsh, J. J., Guise, K. G., Heshmati, M., Friedman, A. K., et al. (2015). Excitatory transmission at thalamo-striatal synapses mediates susceptibility to social stress. *Nat. Neurosci.* 18, 962–964. doi: 10.1038/nn.4034
- Clauss, J. A., Avery, S. N., Benningfield, M. M., and Blackford, J. U. (2019). Social anxiety is associated with bnst response to unpredictability. *Depress. Anxiety* 36, 666–675. doi: 10.1002/da.22891
- Cohen, S., Janicki-Deverts, D., and Miller, G. E. (2007). Psychological stress and disease. *JAMA* 298, 1685–1687. doi: 10.1001/jama.298.14.1685
- Collins, D. P., Anastasiades, P. G., Marlin, J. J., and Carter, A. G. (2018). Reciprocal circuits linking the prefrontal cortex with dorsal and ventral thalamic nuclei. *Neuron* 98, 366–379. doi: 10.1016/j.neuron.2018.03.024
- Courtin, J., Chaudun, F., Rozeske, R. R., Karalis, N., Gonzalez-Campo, C., Wurtz, H., et al. (2014). Prefrontal parvalbumin interneurons shape neuronal activity to drive fear expression. *Nature* 505, 92–96. doi: 10.1038/nature12755
- Cummings, K. A., and Clem, R. L. (2020). Prefrontal somatostatin interneurons encode fear memory. *Nat. Neurosci.* 23, 61–74. doi: 10.1038/s41593-019-0552-7
- De Gregorio, D., Popic, J., Enns, J. P., Inserra, A., Skalecka, A., Markopoulos, A., et al. (2021). Lysergic acid diethylamide (LSD) promotes social behavior through mTORC1 in the excitatory neurotransmission. *Proc. Natl. Acad. Sci. U.S.A.* 118, e2020705118. doi: 10.1073/pnas.2020705118
- Delevich, K., Tucciaroni, J., Huang, Z. J., and Li, B. (2015). The mediodorsal thalamus drives feedforward inhibition in the anterior cingulate cortex via parvalbumin interneurons. *J. Neurosci.* 35, 5743–5753. doi: 10.1523/JNEUROSCI.4565-14.2015

- Deng, H., Xiao, X., and Wang, Z. (2016). Periaqueductal gray neuronal activities underlie different aspects of defensive behaviors. *J. Neurosci.* 36, 7580–7588. doi: 10.1523/JNEUROSCI.4425-15.2016
- Diamond, A. (2013). Executive functions. *Ann. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750
- Diaz, V., and Lin, D. (2020). Neural circuits for coping with social defeat. *Curr. Opin. Neurobiol.* 60, 99–107. doi: 10.1016/j.conb.2019.11.016
- Dolder, P. C., Schmid, Y., Müller, F., Borgwardt, S., and Liechti, M. E. (2016). Lsd acutely impairs fear recognition and enhances emotional empathy and sociality. *Neuropsychopharmacology* 41, 2638–2646. doi: 10.1038/npp.2016.82
- Dos Santos, R. G., de Lima Osório, F., Martin-Santos, R., Zuardi, A. W., Hallak, J. E. C., and Crippa, J. A. S. (2019). Modulation of the endocannabinoid and oxytocinergic systems as a potential treatment approach for social anxiety disorder. *CNS Drugs* 33, 1031–1038. doi: 10.1007/s40263-019-00669-5
- Duerler, P., Schilbach, L., Stämpfli, P., Vollenweider, F. X., and Preller, K. H. (2020). Lsd-induced increases in social adaptation to opinions similar to one's own are associated with stimulation of serotonin receptors. *Sci. Rep.* 10, 1–11. doi: 10.1038/s41598-020-68899-y
- Duval, E. R., Javanbakht, A., and Liberzon, I. (2015). Neural circuits in anxiety and stress disorders: a focused review. *Therapeut. Clin. Risk Manag.* 11, 115. doi: 10.2147/TCRM.S48528
- Falkner, A. L., Wei, D., Song, A., Watsek, L. W., Chen, I., Chen, P., Feng, J. E., and Lin, D. (2020). Hierarchical representations of aggression in a hypothalamic-midbrain circuit. *Neuron* 106, 637–648.e6. doi: 10.1016/j.neuron.2020.02.014
- Florez, S., Solano, J., and Cardenas Parra, F. (2017). Effects of chronic social stress in both social interaction and expression of serotonin 5-HT_{1A} receptors in the paraventricular and the supraoptic nuclei of the hypothalamus. *Psychol. Neurosci.* 10, 252. doi: 10.1037/pne0000089
- Franklin, T. B., Silva, B. A., Perova, Z., Marrone, L., Masferrer, M. E., Zhan, Y., Kaplan, A., Greetham, L., Verrechia, V., Halman, A., et al. (2017). Prefrontal cortical control of a brainstem social behavior circuit. *Nat. Neurosci.* 20, 260. doi: 10.1038/nn.4470
- Froemke, R. C., Merzenich, M. M., and Schreiner, C. E. (2007). A synaptic memory trace for cortical receptive field plasticity. *Nature* 450, 425–429. doi: 10.1038/nature06289
- Gold, A. L., Shechner, T., Farber, M. J., Spiro, C. N., Leibenluft, E., Pine, D. S., et al. (2016). Amygdala–cortical connectivity: associations with anxiety, development, and threat. *Depress. Anxiety* 33, 917–926. doi: 10.1002/da.22470
- Goldin, P. R., Manber, T., Hakimi, S., Canli, T., and Gross, J. J. (2009). Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. *Arch. Gen. Psychiatry* 66, 170–180. doi: 10.1001/archgenpsychiatry.2008.525
- Gonzalez, S. T., and Fanselow, M. S. (2020). The role of the ventromedial prefrontal cortex and context in regulating fear learning and extinction. *Psychol. Neurosci.* 13, 459–472. doi: 10.1037/pne0000207
- Gross, C. T., and Canteras, N. S. (2012). The many paths to fear. *Nat. Rev. Neurosci.* 13, 651–658. doi: 10.1038/nrn3301
- Gu, Z., Smith, K. G., Alexander, G. M., Guerreiro, I., Dudek, S. M., Gutkin, B., et al. (2020). Hippocampal interneuronal $\alpha 7$ nAChRs modulate theta oscillations in freely moving mice. *Cell Rep.* 31, 107740. doi: 10.1016/j.celrep.2020.10.7740
- Guastella, A. J., Howard, A. L., Dadds, M. R., Mitchell, P., and Carson, D. S. (2009). A randomized controlled trial of intranasal oxytocin as an adjunct to exposure therapy for social anxiety disorder. *Psychoneuroendocrinology* 34, 917–923. doi: 10.1016/j.psyneuen.2009.01.005
- Hambrick, J. P., Weeks, J. W., Harb, G. C., and Heimberg, R. G. (2003). Cognitive-behavioral therapy for social anxiety disorder: supporting evidence and future directions. *CNS Spectr.* 8, 373–381. doi: 10.1017/s1092852900018630
- Hangya, B., Ranade, S. P., Lorenc, M., and Kepecs, A. (2015). Central cholinergic neurons are rapidly recruited by reinforcement feedback. *Cell* 162, 1155–1168. doi: 10.1016/j.cell.2015.07.057
- Hernandez, C. M., Cortez, I., Gu, Z., Colón-Sáez, J. O., Lamb, P. W., Wakamiya, M., et al. (2014). Research tool: validation of floxed $\alpha 7$ nicotinic acetylcholine receptor conditional knockout mice using in vitro and in vivo approaches. *J. Physiol.* 592, 3201–3214. doi: 10.1113/jphysiol.2014.272054
- Herry, C., Ciocchi, S., Senn, V., Demmou, L., Müller, C., and Lüthi, A. (2008). Switching on and off fear by distinct neuronal circuits. *Nature* 454, 600–606. doi: 10.1038/nature07166
- Herry, C., and Garcia, R. (2002). Prefrontal cortex long-term potentiation, but not long-term depression, is associated with the maintenance of extinction of learned fear in mice. *J. Neurosci.* 22, 577–583. doi: 10.1523/JNEUROSCI.22-02-00577.2002
- Hinwood, M., Tynan, R. J., Day, T. A., and Walker, F. R. (2011). Repeated social defeat selectively increases *delta*fosb expression and histone h3 acetylation in the infralimbic medial prefrontal cortex. *Cereb. Cortex* 21, 262–271. doi: 10.1093/cercor/bhq080
- Huhman, K. L. (2006). Social conflict models: can they inform us about human psychopathology? *Horm. Behav.* 50, 640–646. doi: 10.1016/j.yhbeh.2006.06.022
- John, Y. J., Bullock, D., Zikopoulos, B., and Barbas, H. (2013). Anatomy and computational modeling of networks underlying cognitive-emotional interaction. *Front. Hum. Neurosci.* 7, 101. doi: 10.3389/fnhum.2013.00101
- Kapfer, C., Glickfeld, L. L., Atallah, B. V., and Scanziani, M. (2007). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nat. Neurosci.* 10, 743–753. doi: 10.1038/nn1909
- Kelly, J. M., Jakubovski, E., and Bloch, M. H. (2014). Prognostic subgroups for remission and response in the coordinated anxiety learning and management (calm) trial. *J. Clin. Psychiatry* 76, 267–278. doi: 10.4088/JCP.13m08922
- Kesner, R. P., and Churchwell, J. C. (2011). An analysis of rat prefrontal cortex in mediating executive function. *Neurobiol. Learn. Mem.* 96, 417–431. doi: 10.1016/j.nlm.2011.07.002
- Kim, D., Samarth, P., Feng, F., Pare, D., and Nair, S. S. (2016). Synaptic competition in the lateral amygdala and the stimulus specificity of conditioned fear: a biophysical modeling study. *Brain Struct. Funct.* 221, 2163–2182. doi: 10.1007/s00429-015-1037-4
- Kim, D.-W., Yao, Z., Graybuck, L. T., Kim, T. K., Nguyen, T. N., Smith, K. A., et al. (2019). Multimodal analysis of cell types in a hypothalamic node controlling social behavior. *Cell* 179, 713–728. doi: 10.1016/j.cell.2019.09.020
- Klumpp, H., Keutmann, M. K., Fitzgerald, D. A., Shankman, S. A., and Phan, K. L. (2014). Resting state amygdala–prefrontal connectivity predicts symptom change after cognitive behavioral therapy in generalized social anxiety disorder. *Biol. Mood Anxiety Disor.* 4, 1–7. doi: 10.1186/s13587-014-0014-5
- Krasne, F. B., Fanselow, M., and Zelikowsky, M. (2011). Design of a neurally plausible model of fear learning. *Front. Behav. Neurosci.* 5, 41. doi: 10.3389/fnbeh.2011.00041
- Krettek, J., and Price, J. (1977). Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *J. Comparat. Neurol.* 172, 687–722.
- Krzywkowski, P., Penna, B., and Gross, C. T. (2020). Dynamic encoding of social threat and spatial context in the hypothalamus. *Elife* 9, e57148. doi: 10.7554/eLife.57148
- Kuroda, M., Yokofujita, J., and Murakami, K. (1998). An ultrastructural study of the neural circuit between the prefrontal cortex and the mediodorsal nucleus of the thalamus. *Progr. Neurobiol.* 54, 417–458.
- Labuschagne, I., Phan, K. L., Wood, A., Angstadt, M., Chua, P., Heinrichs, M., et al. (2010). Oxytocin attenuates amygdala reactivity to fear in generalized social anxiety disorder. *Neuropsychopharmacology* 35, 2403–2413. doi: 10.1038/npp.2010.123
- Labuschagne, I., Phan, K. L., Wood, A., Angstadt, M., Chua, P., Heinrichs, M., et al. (2012). Medial frontal hyperactivity to sad faces in generalized social anxiety disorder and modulation by oxytocin. *Int. J. Neuropsychopharmacol.* 15, 883–896. doi: 10.1017/S1461145711001489
- Larrieu, T., Cherix, A., Duque, A., Rodrigues, J., Lei, H., Gruetter, R., et al. (2017). Hierarchical status predicts behavioral vulnerability and nucleus accumbens metabolic profile following chronic social defeat stress. *Curr. Biol.* 27, 2202–2210. doi: 10.1016/j.cub.2017.06.027
- Leichsenring, F., and Leweke, F. (2017). Social anxiety disorder. *New England J. Med.* 376, 2255–2264. doi: 10.1056/NEJMcp1614701
- Letzkus, J. J., Wolff, S. B., Meyer, E. M., Tovote, P., Courtin, J., Herry, C., et al. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature* 480, 331–335. doi: 10.1038/nature10674
- Li, G., Nair, S. S., and Quirk, G. J. (2009). A biologically realistic network model of acquisition and extinction of conditioned fear associations in lateral amygdala neurons. *J. Neurophysiol.* 101, 1629–1646. doi: 10.1152/jn.90765.2008
- Li, Y., Mathis, A., Grewe, B. F., Osterhout, J. A., Ahanonu, B., Schnitzer, M. J., et al. (2017). Neuronal representation of social information in the medial amygdala of awake behaving mice. *Cell* 171, 1176–1190. doi: 10.1016/j.cell.2017.10.015

- Li, Y., Nakae, K., Ishii, S., and Naoki, H. (2016). Uncertainty-dependent extinction of fear memory in an amygdala-mpfc neural circuit model. *PLoS Comput. Biol.* 12, e1005099. doi: 10.1371/journal.pcbi.1005099
- Liang, B., Zhang, L., Barbera, G., Fang, W., Zhang, J., Chen, X., et al. (2018). Distinct and dynamic on and off neural ensembles in the prefrontal cortex code social exploration. *Neuron* 100, 700–714. doi: 10.1016/j.neuron.2018.08.043
- Lima, S. Q., Hromádka, T., Znamenskiy, P., and Zador, A. M. (2009). Pinp: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS ONE* 4, e6099. doi: 10.1371/journal.pone.0006099
- Lin, D., Boyle, M. P., Dollar, P., Lee, H., Lein, E., Perona, P., and Anderson, D. J. (2011). Functional identification of an aggression locus in the mouse hypothalamus. *Nature* 470, 221–226. doi: 10.1038/nature09736
- Liu, L., Xu, H., Wang, J., Li, J., Tian, Y., Zheng, J., et al. (2020). Cell type–differential modulation of prefrontal cortical gabaergic interneurons on low gamma rhythm and social interaction. *Sci. Adv.* 6, eaay4073. doi: 10.1126/sciadv.aay4073
- Lo, L., Yao, S., Kim, D.-W., Cetin, A., Harris, J., Zeng, H., et al. (2019). Connectional architecture of a mouse hypothalamic circuit node controlling social behavior. *Proc. Natl. Acad. Sci. U.S.A.* 116, 7503–7512. doi: 10.1073/pnas.1817503116
- Luckett, C., Norvelle, A., and Huhman, K. (2012). The role of the nucleus accumbens in the acquisition and expression of conditioned defeat. *Behav. Brain Res.* 227, 208–214. doi: 10.1016/j.bbr.2011.10.012
- Mannella, F., Miroli, M., and Baldassarre, G. (2016). Goal-directed behavior and instrumental devaluation: a neural system-level computational model. *Front. Behav. Neurosci.* 10, 181. doi: 10.3389/fnbeh.2016.00181
- Mannella, F., Zappacosta, S., Miroli, M., and Baldassarre, G. (2008). “A computational model of the amygdala nuclei’s role in second order conditioning,” in *International Conference on Simulation of Adaptive Behavior* (Osaka: Springer), 321–330.
- Marazziti, D., Abelli, M., Baroni, S., Carpita, B., Ramacciotti, C. E., and Dell’Osso, L. (2015). Neurobiological correlates of social anxiety disorder: an update. *CNS Spectr.* 20, 100–111. doi: 10.1017/S109285291400008X
- Maren, S., and Holmes, A. (2016). Stress and fear extinction. *Neuropsychopharmacology* 41, 58–79. doi: 10.1038/npp.2015.180
- Mattera, A., Pagani, M., and Baldassarre, G. (2020). A computational model integrating multiple phenomena on cued fear conditioning, extinction, and reinstatement. *Front. Syst. Neurosci.* 14, 569108. doi: 10.3389/fnsys.2020.569108
- Mayo-Wilson, E., Dias, S., Mavranzouli, I., Kew, K., Clark, D. M., Ades, A., and Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry* 1, 368–376. doi: 10.1016/S2215-0366(14)70329-3
- Menon, R., Grund, T., Zoicas, I., Althammer, F., Fiedler, D., Biermeier, V., et al. (2018). Oxytocin signaling in the lateral septum prevents social fear during lactation. *Curr. Biol.* 28, 1066–1078. doi: 10.1016/j.cub.2018.02.044
- Milad, M. R., and Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420, 70–74. doi: 10.1038/nature01138
- Mitchell, A. S., and Chakraborty, S. (2013). What does the mediodorsal thalamus do? *Front. Syst. Neurosci.* 7, 37. doi: 10.3389/fnsys.2013.00037
- Motta, S. C., Goto, M., Gouveia, F. V., Baldo, M. V., Canteras, N. S., and Swanson, L. W. (2009). Dissecting the brain’s fear system reveals the hypothalamus is critical for responding in subordinate conspecific intruders. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4870–4875. doi: 10.1073/pnas.0900939106
- Moustafa, A. A., Gilbertson, M. W., Orr, S. P., Herzallah, M. M., Servatius, R. J., and Myers, C. E. (2013). A model of amygdala–hippocampal–prefrontal interaction in fear conditioning and extinction in animals. *Brain Cogn.* 81, 29–43. doi: 10.1016/j.bandc.2012.10.005
- Murray, A. J., Woloszynowska-Fraser, M. U., Ansel-Bollepalli, L., Cole, K. L., Foggetti, A., Crouch, B., et al. (2015). Parvalbumin-positive interneurons of the prefrontal cortex support working memory and cognitive flexibility. *Sci. Rep.* 5, 1–14. doi: 10.1038/srep16778
- Nair, S. S., Paré, D., and Vicentic, A. (2016). Biologically based neural circuit modelling for the study of fear learning and extinction. *NPJ Sci. Learn.* 1, 1–7. doi: 10.1038/npjscilearn.2016.15
- Naka, A., and Adesnik, H. (2016). Inhibitory circuits in cortical layer 5. *Front. Neural Circuits* 10, 35. doi: 10.3389/fncir.2016.00035
- Nelson, A. C., Kapoor, V., Vaughn, E., Gnanasegaram, J. A., Rubinstein, N. D., Murthy, V. N., et al. (2019). Molecular and circuit architecture of social hierarchy. *bioRxiv* 838664.
- Nichols, D. E. (2004). Hallucinogens. *Pharmacol. Ther.* 101, 131–181. doi: 10.1016/j.pharmthera.2003.11.002
- Nigro, M. J., Hashikawa-Yamasaki, Y., and Rudy, B. (2018). Diversity and connectivity of layer 5 somatostatin-expressing interneurons in the mouse barrel cortex. *J. Neurosci.* 38, 1622–1633. doi: 10.1523/JNEUROSCI.2415-17.2017
- Nordman, J. C., Ma, X., Gu, Q., Potegal, M., Li, H., Kravitz, A. V., et al. (2020). Potentiation of divergent medial amygdala pathways drives experience-dependent aggression escalation. *J. Neurosci.* 40, 4858–4880. doi: 10.1523/JNEUROSCI.0370-20.2020
- Oliva, V., Cartoni, E., Latagliata, E. C., Puglisi-Allegra, S., and Baldassarre, G. (2018). Interplay of prefrontal cortex and amygdala during extinction of drug seeking. *Brain Struct. Funct.* 223, 1071–1089. doi: 10.1007/s00429-017-1533-9
- Pilling, S., Mayo-Wilson, E., Mavranzouli, I., Kew, K., Taylor, C., and Clark, D. M. (2013). Recognition, assessment and treatment of social anxiety disorder: summary of nice guidance. *BMJ* 346, bmj.f2541. doi: 10.1136/bmj.f2541
- Prater, K. E., Hosanagar, A., Klumpp, H., Angstadt, M., and Luan Phan, K. (2013). Aberrant amygdala–frontal cortex connectivity during perception of fearful faces and at rest in generalized social anxiety disorder. *Depress. Anxiety* 30, 234–241. doi: 10.1002/da.22014
- Rao, T. S., and Andrade, C. (2017). A possible role for ketamine in suicide prevention in emergency and mainstream psychiatry. *Indian J. Psychiatry* 59, 259. doi: 10.4103/psychiatry.IndianJPsychiatry_345_17
- Repa, J. C., Muller, J., Apergis, J., Desrochers, T. M., Zhou, Y., and LeDoux, J. E. (2001). Two different lateral amygdala cell populations contribute to the initiation and storage of memory. *Nat. Neurosci.* 4, 724–731. doi: 10.1038/89512
- Riedemann, T. (2019). Diversity and function of somatostatin-expressing interneurons in the cerebral cortex. *Int. J. Mol. Sci.* 20, 2952. doi: 10.3390/ijms20122952
- Risold, P., and Swanson, L. (1997). Connections of the rat lateral septal complex. *Brain Res. Rev.* 24, 115–195.
- Sakurai, K., Zhao, S., Takatoh, J., Rodriguez, E., Lu, J., Leavitt, A. D., et al. (2016). Capturing and manipulating activated neuronal ensembles with cane delineates a hypothalamic social-fear circuit. *Neuron* 92, 739–753. doi: 10.1016/j.neuron.2016.10.015
- Santini, E., Quirk, G. J., and Porter, J. T. (2008). Fear conditioning and extinction differentially modify the intrinsic excitability of infralimbic neurons. *J. Neurosci.* 28, 4028–4036. doi: 10.1523/JNEUROSCI.2623-07.2008
- Sato, M., Ito, M., Nagase, M., Sugimura, Y. K., Takahashi, Y., Watabe, A. M., et al. (2015). The lateral parabrachial nucleus is actively involved in the acquisition of fear memory in mice. *Mol. Brain* 8, 1–15. doi: 10.1186/s13041-015-0108-z
- Scaini, S., Belotti, R., Ogliari, A., and Battaglia, M. (2016). A comprehensive meta-analysis of cognitive-behavioral interventions for social anxiety disorder in children and adolescents. *J. Anxiety Disord.* 42, 105–112. doi: 10.1016/j.janxdis.2016.05.008
- Shen, B., and McNaughton, B. L. (1996). Modeling the spontaneous reactivation of experience-specific hippocampal cell assemblies during sleep. *Hippocampus* 6, 685–692.
- Silberberg, G., and Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by martinotti cells. *Neuron* 53, 735–746. doi: 10.1016/j.neuron.2007.02.012
- Silva, B. A., Gross, C. T., and Gräff, J. (2016a). The neural circuits of innate fear: detection, integration, action, and memorization. *Learn. Mem.* 23, 544–555. doi: 10.1101/lm.042812.116
- Silva, B. A., Mattucci, C., Krzykowski, P., Cuzzo, R., Carbonari, L., and Gross, C. T. (2016b). The ventromedial hypothalamus mediates predator fear memory. *Eur. J. Neurosci.* 43, 1431–1439. doi: 10.1111/ejn.13239
- Silva, B. A., Mattucci, C., Krzykowski, P., Murana, E., Illarionova, A., Grinevich, V., Canteras, N. S., Ragozzino, D., and Gross, C. T. (2013). Independent hypothalamic circuits for social and predator fear. *Nat. Neurosci.* 16, 1731–1733. doi: 10.1038/nn.3573
- Stein, D. J. (2015). Social anxiety disorder and the psychobiology of self-consciousness. *Front. Hum. Neurosci.* 9, 489. doi: 10.3389/fnhum.2015.00489
- Stein, M. B., and Stein, D. J. (2008). Social anxiety disorder. *Lancet* 371, 1115–1125. doi: 10.1016/S0140-6736(08)60488-2

- Sun, Q., Li, X., Ren, M., Zhao, M., Zhong, Q., Ren, Y., et al. (2019). A whole-brain map of long-range inputs to GABAergic interneurons in the mouse medial prefrontal cortex. *Nat. Neurosci.* 22, 1357–1370. doi: 10.1038/s41593-019-0429-9
- Toth, I., Neumann, I. D., and Slattery, D. A. (2012). Social fear conditioning: a novel and specific animal model to study social anxiety disorder. *Neuropsychopharmacology* 37, 1433–1443. doi: 10.1038/npp.2011.329
- Toth, I., Neumann, I. D., and Slattery, D. A. (2013). Social fear conditioning as an animal model of social anxiety disorder. *Curr. Protocols Neurosci.* 63, 9–42. doi: 10.1002/0471142301.ns0942s63
- Trouche, S., Sasaki, J. M., Tu, T., and Reijmers, L. G. (2013). Fear extinction causes target-specific remodeling of perisomatic inhibitory synapses. *Neuron* 80, 1054–1065. doi: 10.1016/j.neuron.2013.07.047
- Uylings, H. B., Groenewegen, H. J., and Kolb, B. (2003). Do rats have a prefrontal cortex? *Behav. Brain Res.* 146, 3–17. doi: 10.1016/j.bbr.2003.09.028
- Vitiello, B. (2009). Combined cognitive-behavioural therapy and pharmacotherapy for adolescent depression. *CNS Drugs* 23, 271–280. doi: 10.2165/00023210-200923040-00001
- Wang, F., Zhu, J., Zhu, H., Zhang, Q., Lin, Z., and Hu, H. (2011). Bidirectional control of social hierarchy by synaptic efficacy in medial prefrontal cortex. *Science* 334, 693–697. doi: 10.1126/science.1209951
- Wang, J., Tian, Y., Zeng, L.-H., and Xu, H. (2020). Prefrontal disinhibition in social fear: a vital action of somatostatin interneurons. *Front. Cell. Neurosci.* 14, 611732. doi: 10.3389/fncel.2020.611732
- Wang, L., Talwar, V., Osakada, T., Kuang, A., Guo, Z., Yamaguchi, T., and Lin, D. (2019). Hypothalamic control of conspecific self-defense. *Cell Rep.* 26, 1747–1758. doi: 10.1016/j.celrep.2019.01.078
- Wei, Y., Krishnan, G. P., and Bazhenov, M. (2016). Synaptic mechanisms of memory consolidation during sleep slow oscillations. *J. Neurosci.* 36, 4231–4247. doi: 10.1523/JNEUROSCI.3648-15.2016
- Wong, L. C., Wang, L., D'amour, J. A., Yumita, T., Chen, G., Yamaguchi, T., et al. (2016). Effective modulation of male aggression through lateral septum to medial hypothalamus projection. *Curr. Biol.* 26, 593–604. doi: 10.1016/j.cub.2015.12.065
- Xu, H., Liu, L., Tian, Y., Wang, J., Li, J., Zheng, J., et al. (2019). A disinhibitory microcircuit mediates conditioned social fear in the prefrontal cortex. *Neuron* 102, 668–682. doi: 10.1016/j.neuron.2019.02.026
- Yamamoto, K., Bicks, L. K., Leventhal, M. B., Kato, D., Im, S., Flanigan, M. E., et al. (2020). A prefrontal–paraventricular thalamus circuit requires juvenile social experience to regulate adult sociability in mice. *Nat. Neurosci.* 23, 1240–1252. doi: 10.1038/s41593-020-0695-6
- Zhou, T., Zhu, H., Fan, Z., Wang, F., Chen, Y., Liang, H., et al. (2017). History of winning remodels thalamo-pfc circuit to reinforce social dominance. *Science* 357(6347):162–168. doi: 10.1126/science.aak9726
- Zoicas, I., Slattery, D. A., and Neumann, I. D. (2014). Brain oxytocin in social fear conditioning and its extinction: involvement of the lateral septum. *Neuropsychopharmacology* 39, 3027–3035. doi: 10.1038/npp.2014.156

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Alfieri, Mattered and Baldassarre. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.