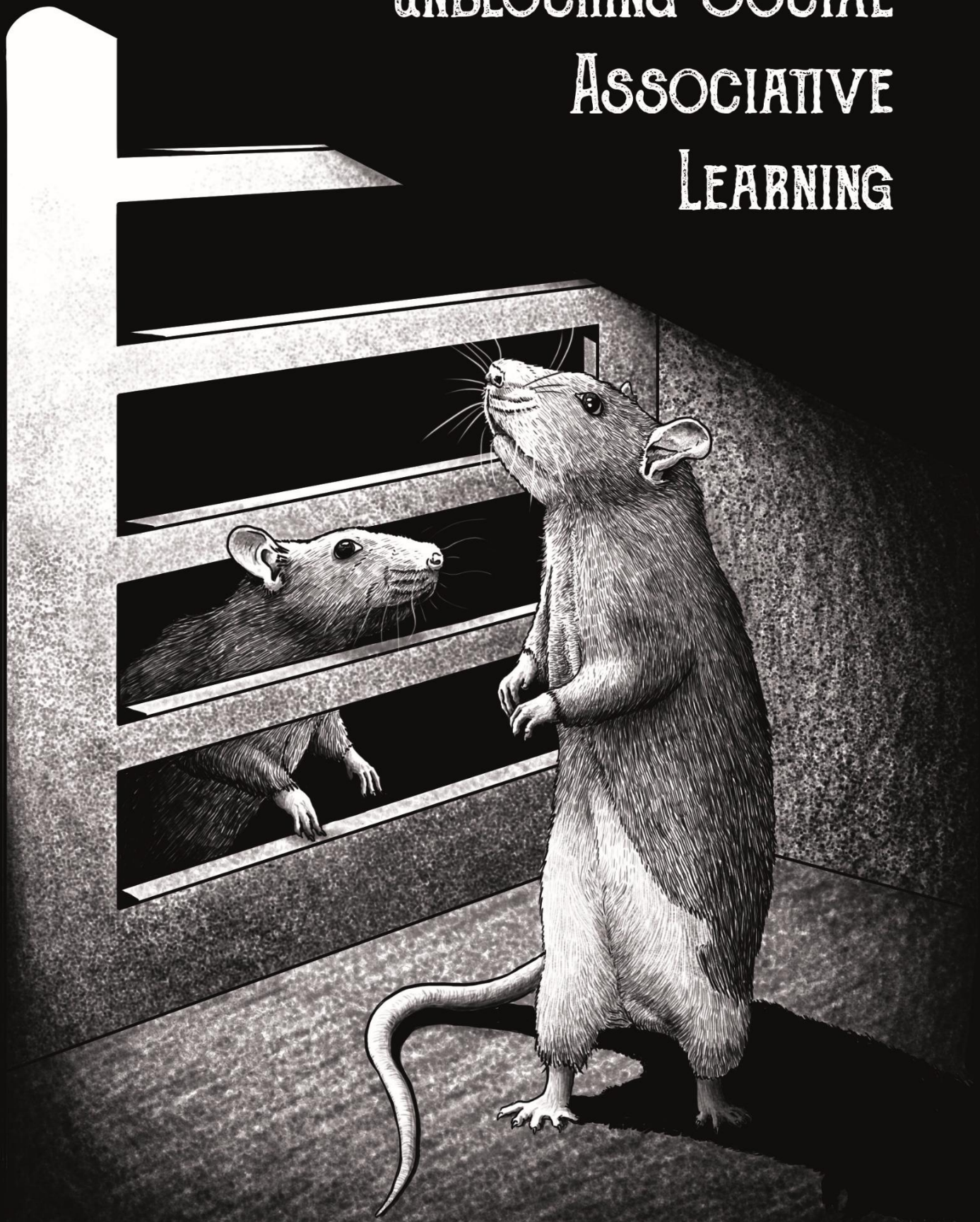
Unblocking Social Associative Learning



Sander van Gurp

Unblocking social associative learning

Sander van Gurp

Unblocking social associative learning

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences

at the Heinrich Heine University Düsseldorf

presented by

Sander van Gurp

from Sittard (Netherlands)

Sittard, April, 2021

From the Comparative Psychology, Institute of Experimental Psychology

at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at

Heinrich Heine University Düsseldorf

Supervisor: Associate Prof. Dr. Marijn van Wingerden

Co-supervisor: Prof. Dr. Tobias Kalenscher

Date of the oral examination:

24.09.2021

"All we have to decide is what to do with the time that is given us."

J.R.R Tolkien, The Fellowship of the Ring.

"Two things fill the mind with ever new and increasing admiration and awe, the more often and steadily we reflect upon them: the starry heavens above me and the moral law within me.

Immanuel Kant, Critique of Practical Reason

Table of contents

1. Introduction

1.1 Preface. People are motivated to seek social contact and maintain social relationships. The size of this social support system is a buffer against stress and is negatively associated with mortality. In addition, Hill (1987) constructed a scale to index the key factors important for measuring the positive aspect relative to social motivation. Four factors were identified: social comparison, emotional support, positive stimulation, and attention. In this thesis, I am primarily interested in positive stimulation, which Hill defined as the ability of affiliation to provide enjoyable affective and cognitive stimulation. Two items in the scale that measure positive stimulation are: "Just being around others and finding out about them is one of the most interesting things I can think of doing" and "The main thing I like about being around other people is the warm glow I get from contact with them." Examples of positive stimulation range from taking a simple walk in the park with a friend or going to the cinema with a loved one to enjoying a family dinner over the holidays. These experiences provide enjoyable affective and cognitive stimulation and are therefore considered valuable.

In the animal world, social motivation and interaction are also important. When young mice and rats are isolated from their parents, they begin vocalizing for their mothers because they miss the nest's warmth and sensory features. In response to this, the mother picks them up and returns them to the nest (Hofer, 1996). During adolescence, rats participate in social play, also referred to as rough and tumble play; the play partners take turns tumbling each other over onto their backs and consequently stand on top and pin the other down (Panksepp, 1981). Furthermore, monkeys negatively react when they see another monkey receive a more attractive reward for the same amount of effort and reject such offers (Brosnan & de Waal, 2003). Tamir and Hughes (2018) described ways to break down such complicated social behaviors into components. There is an immediate cause (e.g., the pups' calls and the presence of a conspecific rat in the previous example). In addition, there is the expression of the behavior (e.g., the pups' retrieval or the rough and tumble play). Performing such behaviors, finally, leads to positive affective enjoyment (e.g., the pup being reunited with the mother in the nest or the positive feelings associated with finishing a social play activity). What most of these behaviors have in common is that the organism (human, monkey, or rat) must be able to infer the needs and feelings of others to take part in the specific social behavior. This ability is called empathy and it contains a complex system of functions, such as sharing the emotional state of others, being influenced by it, and understanding others' specific needs (de Waal & Preston, 2017). A good example of easily identifying someone else's emotional state is by observing the other person expressing joy while eating their favorite ice cream. If they in turn observe that you picked your favorite ice cream and you enjoy eating it, then there is a

mutual match in sharing each other's positive state, further enhancing the value of the experience. Experiencing the rewarding outcome of sharing the ice cream might lead you and your friend to think about it days later, conjuring up specific sensory details like the taste. Importantly, the context in which you gained this rewarding outcome is also remembered; perhaps it occurred on the Rhein boulevard in Düsseldorf and you might recall the exact spot where you walked to afterward and on which bench you both sat. The mutual expression of enjoyment and positive valuation of social and contextual information shared during social interactions has connected the environmental context with the social reward. If you eat ice cream at another place on a different occasion, and the two of you enjoy the ice cream there less, you might remember that place and its features less vividly. Hypothetically, you visit both places an equal amount of times throughout the summer. Someone then gives you the choice of revisiting one of the two environments. The context of the place with the most rewarding shared ice cream would lead you to think back on that experience and the shared enjoyment you had, thereby motivating you to want to return there rather than the environment of the less liked ice cream.

I assume that many people (including myself) have had such experiences where the social experience becomes associated with the context that predicts them and the episodic social memories associated with them. The ice cream example can be interchanged with any other shared activity; for example, for me, it was roller-skating with a group of friends or swimming with a close friend. This thesis aims to advance knowledge about how the learning of such cues predicts positive social rewards and how the brain supports such learning. In the remainder of this introduction, three subjects are discussed: (1) the current state of knowledge regarding social reinforcement (see definitions; Table 1) learning and the introduction and description of a task that can measure social associative learning, (2) the potential role of the orbito-frontal cortex (OFC) in social reinforcement learning in that task, and (3) the role of ultrasonic vocalizations in a social – nonsocial spatial learning task.

1.2 Social reinforcement learning. As we saw in the general introduction, mutual rewards and social interaction are essential for our well-being and drive learning about the features in the environment that are significant predictors for acquiring social knowledge. This thesis aims to enhance knowledge about social reinforcement learning and introduce new behavioral paradigms that aid in studying the valuation of features in the environment that predict mutually rewarding outcomes and social interaction. In the following introductory paragraphs, I first discuss how social reinforcement learning works by looking at its different aspects: attention, incentive salience, Pavlovian (social) conditioning, and dopamine as a driving force in reinforcement learning (definitions, see Table 1). These are all aspects of social reinforcement learning that are all directly

active during mutual reward learning and therefore help the reader understand why and how mutual reward processing occurs. After discussing these more basic features and vicarious learning about other's outcomes, I discuss unblocking as a learning paradigm that is optimally suited to investigate a less understood component of vicarious learning in rats: appetitive other-regarding stimulus-outcome learning. The role of the OFC in social and non-social reinforcement learning is specified, and a potential role for this brain structure in the social unblocking task is described. Finally, the role of ultrasonic vocalizations as a communicative and potential social reinforcement signal is discussed together with the introduction of a task that enables the direct comparison of cues that predict social versus non-social rewards, which directly helps with understanding and quantifying the valuation of social interaction as a rewarding outcome.

1.3 Social stimuli drive social attention. People prefer social stimuli (definition; see Table 1) over non-social stimuli, as they orient more to and engage more with the social stimuli (Chakrabarti, Haffey, Canzano, Taylor, & Mcsorley, 2017). Furthermore, social rewards drive attentional orienting to stimuli and spatial locations (B. A. Anderson, 2016; Hayward, Pereira, Otto, & Ristic, 2018). In an experiment by Hayward et al. (2018), participants learned to click at spatial targets. Afterward, they went through a social reward modulation phase, in which they received positive social feedback from an experimenter. After this phase, when they were asked to win points for the experimenter, their reaction times for detecting spatial locations associated with winning those points were faster than those previously shown in a baseline period. Hayward et al. (2018) concluded that the social reward feedback had driven enhanced attentional orienting for the rewarded spatial location. Next to an enhanced attentional bias for social stimuli, people also enjoy direct social interaction. For example, when performing a ball-tossing game, a higher frequency of tosses to other players led to more enjoyment (Kawamichi et al., 2016). There is thus a positive value inherent in social interactions, which they describe as "emerging from increased feelings of social connection". These are rewarding feelings that motivate a person to engage in in social behaviors (Tamir & Hughes, 2018), as in the ball game, or capture and drive attention toward social stimuli.

The motivational value associated with social stimuli is also present throughout the animal kingdom. Multiple studies have shown that monkeys perform motor actions to access social stimuli, establishing their rewarding properties (e.g., Anderson, 1998). Furthermore, Deaner et al. (2005) found that monkeys value social information according to its social utility (i.e., images of other monkeys with a specific social status or reproductive potential) and that this utility was reflected in viewing time and the orienting value to those biologically significant cues. Studies demonstrated that rats were more attracted to rats that could move versus stuffed or anesthetized

rats (Latane, Joy, Meltzer, Lubell, & Cappell, 1972). Moreover, Latané and Steele (1975) found that the social interaction of rats in an open field increased from 30% of the time in the first five minutes to 67% in the last one and a half hours of a seven and a half hour period. This interaction was further increased by social isolation. The interactions in the first period involved mostly active social play activities, and the later period involved a more passive huddling. Their results showed that interest in social contact is long-lasting and that the expression patterns in rats are sophisticated and involve multiple behavioral repertoires. Social stimuli, thus, drive attention and influence behavior in such a way that they modulate an immediate orienting and engaging response size that is dependent on the value of the social stimuli. Capozzi and Ristic (2018) proposed a theory that stated that the three core processes necessary for this selective response toward social stimuli are perception, interpretation, and evaluation. The perceptual process ensures that attention is directed at the relevant stimuli, the evaluative process determines the social meaning, and the interpretative process links attention with the social stimulus's meaning. Social attention is thus crucial for the selection and processing of social information and appears to be essential for engaging in rewarding social behaviors.

A study by Dawson et al. (2004) highlights the importance of social attention. They studied the failure of children with autism spectrum disorder to orient to social stimuli and their impairment in coordinating attention between social partners and objects of interest (joint attention) by sharing or following their gaze. They found that children with autism spectrum disorder, but not children with a pervasive developmental disorder or typical development, had a deficit in orienting to social stimuli, such as an examiner calling the child's name or snapping their finger while looking at the child. The children with autism spectrum disorder furthermore made fewer attempts to initiate joint attention and were less likely to respond to an examiner that tried to engage in joint attention. This study is significant since it can help examiners identify deficits and start interventions early enough to help children learn to increase their motivation for attending to social stimuli. Besides orienting to social stimuli and engaging in social interaction, it is also crucial to interpret and evaluate social rewards, as Capozzi and Ristic (2018) stated. This process of social valuation will be discussed in the next section.

1.4 Social valuation, feeding motivation, and incentive salience. Why is it that social stimuli draw your attention and drive your approach and interaction? On the one hand, there are inherently rewarding social experiences like the social ball game (Kawamichi et al., 2016) discussed earlier or, for example, looking at social stimuli such as beautiful faces (Aharon et al., 2001). Social reward processing in humans, on the other hand, has been found to involve more complex social cognition like social status and social cooperation (Albrecht, von Essen, Fliessbach, & Falk, 2013;

Rand & Nowak, 2013). Moreover, it has become possible to investigate real-life social interactions in humans because of novel measurement technologies, such as the two-person simultaneous fMRI (definition, see Table 1). These scans enable tracking brain functions by looking at the BOLD signal (definition, see Table 1) during real-time interactions between two persons, for example, during turn-taking in conversations (Hari, Henriksson, Malinen, & Parkkonen, 2015). However, this research on human social reward valuation through real-life social interactions and more complex social cognition is beyond this thesis's scope. Instead, this thesis focuses on more controlled behavioral tasks used to quantify social reward in animal interaction using the following principles.

1. We use "animal species with a behavioral or psychological repertoire similar to humans so that the results of experiments with these animal models may throw light on seemingly related behavior in human beings" (Carroll and Overmier, 2001, as cited in Lickliter, 2004).
2. The relevance of animal research in this thesis and throughout the literature is defined as the meaningfulness and usefulness of the research for scientific purposes (van der Staay, 2006).

These principles lead to the following overall purpose and goal of this thesis:

- To identify the behavioral repertoires and understand the underlying brain substrates underlying valuation of social reward and its functional relevance to ongoing behavior in the rat.
- Reaching this goal fulfills the criteria of meaningfulness and usefulness of the thesis and the animal research it entails.

Trezza et al. (2011) have proposed that the behavioral repertoire associated with social valuation in rats consists of three processes: hedonic "liking," motivational "wanting," and cognitive "learning." These processes can also be found in the non-social valuation of food and drug reward (Berridge & Robinson, 2003), where liking is an objective pleasure reaction and includes, for example, the facial expression (tongue protrusion) induced by the hedonic impact of taste. Wanting or "incentive salience" is the transformation of sensory information about the reward into a wanted target of motivation. Cognitive learning is the important link connecting the two. It, therefore, supports the ability to choose the likable outcome that one (currently) wants (Trezza et al., 2011). Day et al. (1998) have provided a useful model system for this three-stage process of feeding motivation (i.e., the desire to feed). They established that feeding motivation is dependent on endogenous factors such as food deprivation and exogenous factors such as the availability of food items. Consequently, feeding motivation leads to the consumption of food through the act

of eating, which provides direct feedback, increasing the motivation (i.e., because of the hedonic aspects of eating) to either eat more or eat less (through satiation). Animals are furthermore motivated to assess whether a food is nutritionally beneficial or harmful, so that they can modify their feeding patterns if foods are low quality or have toxic properties. This framework for motivated food consumption contains components similar to the incentive salience theory: feeding motivation is a "wanting" for food and the consumption of food includes a hedonic "liking". Feedback terms are necessary to bridge the "wanting" with the sensory aspects of the hedonic "liking" components. This feeding motivation scheme can be easily transferred to represent social valuation. Social incentive salience, either due to social deprivation or the availability of social reward in the environment, activates a "wanting" for social reward (Figure 1). What follows is social exploration and upon encountering the social stimulus, attention is captured, and social interaction is initiated and "consumed", as when encountering and eating food. This interaction in turn elicits feedback, which when positive (i.e., joyful play) leads to more interaction

until a certain level of social satiation is reached, reducing the social motivation and possibly terminating the social interaction. The connection between the knowledge regarding the internal

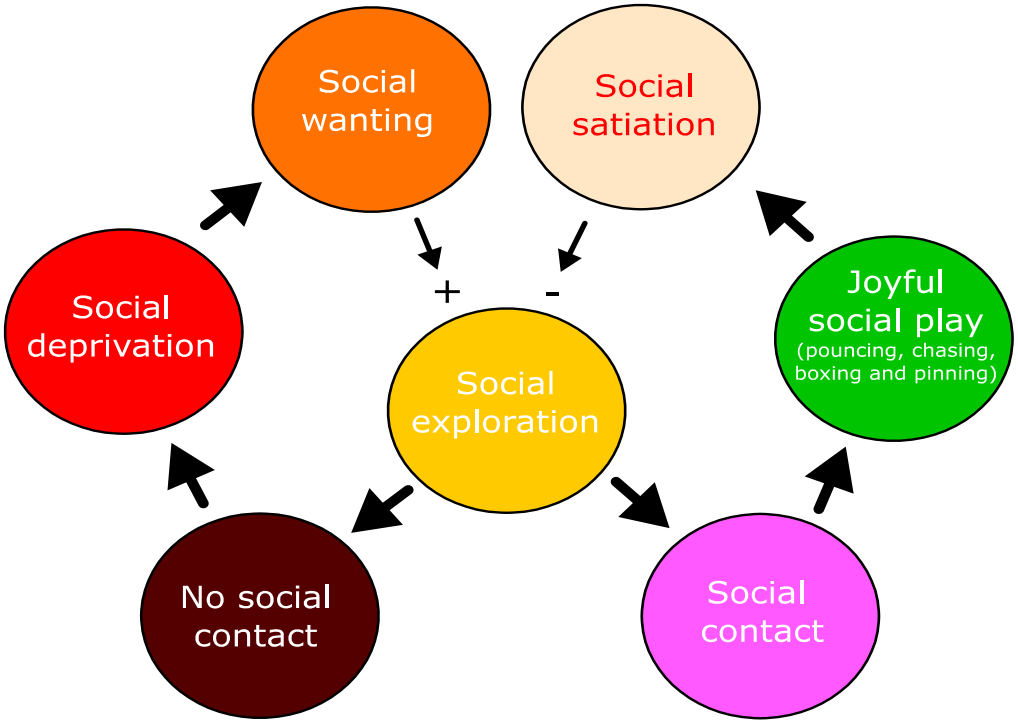


Figure 1. Social incentive salience cycle. If the social play is joyful, the rat is satiated and does not need more social play. If, however, the rat is not able to find a social play partner, the incentive salience system triggers a wanting for social play that leads to renewed social exploration, which upon contact can lead to social interaction and subsequent social play.

state of social need (level of social deprivation) and the feedback regarding the acquired value of social interaction defines the level of social learning. Here, I established a general framework for understanding the behavioral cycle that entails the social valuation process. What is important to note is that this framework has a crucial factor missing: the dependence of learning a particular (social) experience on the environment in which it occurred. After a successful cycle, when social satiation diminishes, it is critical to know how one found the social context and how much enjoyment it actually brought. One can make a more optimal decision by better remembering the social behavioral repertoire's costs (i.e., the energy necessary for social exploration and interaction) and benefits (amount of social joy). The following section focuses on the behavioral learning theory that aids the social valuation process by providing the associative connection between the environment that can make it easier for the searching animal to find the social stimulus and set in motion the approach and start the interaction.

1.5 (Social) Pavlovian conditioning and associative learning. Appetitive learning includes Pavlovian conditioning, whereby a cue, the conditioned stimulus (CS), in the animal's immediate environment, comes to serve as a signal that predicts the unconditioned stimulus (US; Bolles & Fanselow, 1980), which could be food pellets or social stimuli. A modern definition of Pavlovian conditioning (Fanselow & Wassum, 2016) is "the process whereby experience with a conditional relationship between stimuli bestows these stimuli with the ability to promote adaptive behavior patterns that did not occur before the experience." This modern definition was already present in how Pavlov himself described the process in 1927. This definition stems from his observation that dogs already began salivating at the sight or sound of the person bringing food. He stated that "the great advantage to the organism to react to the former stimuli is evident, for it is in the virtue of their action that food finding its way in the mouth immediately encounters plenty of moistening saliva" (Pavlov, 2010). In rats, Pavlovian conditioning similarly occurs when light or sound cues signal the availability of a food pellet reward.

The appetitive Pavlovian learning paradigm called "autoshaping" can be used to illustrate such learning. Cleland and Davey (1983) used this autoshaping procedure when they deprived rats of food and then exposed them to an experimental chamber. The rats could find food pellets in the middle of the chamber in a food tray on a variable interval schedule. On average, a pellet was delivered every 60 seconds. This is called "magazine training," and the rat learns that a spatial location signals food availability. Here, the physiological state of hunger drives a feeding motivation for food that subsequently drives exploratory behavior (Day et al., 1998), in addition to a normal intrinsic exploration mode aimed at exploring novel environments and objects (Hughes, 1997). These different exploratory modes drive the rat to explore and potentially learn

that the environment contains pellets at unexpected moments in time. The animal progresses through multiple stages within this simple first step: an affective hunger state, which triggers an exploratory action that may lead to a reward.

Consequently, reward consumption drives feedback about the nature of the food consumed, leading to the formation of an action-outcome association. If the rat explores this environment, then on some occasions, it can find food and satisfy its wanting. Afterward, Cleland and Davey presented a light for 10 seconds in the two ends of the rectangular maze signaling the delivery of pellets. Here, the conditioned stimulus (CS) presentation leads to the unconditioned stimulus (US) pellet, resulting in the rat approaching the light at the end of the chamber during the cue's presentation. The rats then learned that one CS (CS+) resulted in a food reward and another CS (CS-) resulted in the absence of a food reward. For a visual CS+, the rats approached the light and the food cup, whereas for an auditory CS+, the rats only approached the food cup. No approach was visible for both auditory and visual CS- and the rats did not spend time in the food cup. Cleland and Davey's (1983) results, therefore, conform to Fanselow and Wassum's (2016) definition of Pavlovian conditioning as the rats have learned to associate the CS with its outcome value (food or no food). Thus, the conditioned stimulus promotes an advantageous adaptive behavior that did not occur before, during the initial magazine training. A study by Timberlake and Grant (1975) found that a social CS (the presentation of a restrained rat) predictive of a food US elicited social behaviors such as approach, sniffing, and social contact with the restrained rat. This response was more significant than in the control conditions—rat present, but no food delivery—or when there was food delivery, but a wooden block served as the CS. These results are among the first to show that a social CS can predict a food outcome and elicit social behaviors in rats.

In addition to the CS itself being social in this case, a series of studies have shown that rats can learn to associate a chamber (CS+: neutral cues) on a maze with rewarding partner social play (US). The rats developed a preference for the arm associated with social play over an arm with a non-playful rat (socially conditioned place preference—definition see Table 1; Calcagnetti and Schechter, 1992). Likewise, in honeybees, it was found that naïve honeybees developed a preference for rewarding flowers in which a conspecific (honeybee belonging to the same species) was present (Leadbeater & Chittka, 2009). Furthermore, Al-imari and Gerlai (2008) found that zebrafish approached a cue card that had previously been presented together with the presentation of a conspecific zebrafish over a card that was not. These simple social Pavlovian conditioning schemes involved the social learning components of liking the social stimulus (US), and by repeatedly pairing the US with a neutral cue (CS) confers the cue with a wanting for the social

outcome (place preference). This type of simple social learning works by transferring the value of the conspecific's presence to a neutral cue or spatial location. This learning is thought to be dependent and an extension of the basic idea of stimulus enhancement formulated by Heyes (1994): "observation of a demonstrator (or its products) exposes the observer to a single stimulus (rather than a relationship between events) at t1 and single stimulus exposure effects a change in the observer detected, in any behavior, at t2". Thus, animals can clearly learn to associate novel

Table 1. Definitions.

Social reinforcement. A positive interpersonal stimulus, such as vocalization, smell, touch, or another social outcome, that increases the frequency of the behavior that immediately precedes it. Also called social reward (APA dictionary of psychology, n.d.).

Social stimuli: Any agent, event, or situation with social significance, particularly an individual or group, that elicits a response relevant to interpersonal relationships (APA dictionary of psychology, n.d.).

Incentive salience: Incentive salience is a motivational process that transforms sensory information about rewards and the cues that predict them (sights, sounds, and smells) into attractive and desired incentives (Berridge & Robinson, 2003).

Pavlovian Conditioning: The process whereby experience with a conditional relationship between stimuli bestows these stimuli with the ability to promote adaptive behavior patterns that did not occur before the experience (Fanselow & Wassum, 2016)

Conditioned place preference: A process in which experience with certain stimuli is rendered as reinforcing the place where that experience occurred. (APA dictionary of psychology, n.d.)

fMRI: Functional magnetic resonance imaging (fMRI) is a technique that measures brain activity by detecting changes associated with blood flow (Wikipedia contributors, 2021).

BOLD signal: The blood-oxygen-level-dependent (BOLD) signal detected in the fMRI reflects a change in deoxyhemoglobin, which is driven by changes in blood flow and blood oxygenation. The BOLD signal is directly coupled to the underlying neuronal activity (Hillman, 2014).

Fiber-photometry: Genetically-encoded calcium indicators (GECIs) change their fluorescence signal based on whether or not they are bound to calcium. When a neuron fires an action potential, the internal calcium concentration increases, increasing the ability of the GECIs to fluorescence. With fiber-photometry, excitation light is directed into the brain, and this fluorescence signal is collected (Martianova, Aronson, & Proulx, 2019).

Optogenetics: Optogenetics is a technique that uses light to control neural activity with high temporal resolution (Deisseroth et al., 2006).

Extracellular electrophysiology: Extracellular electrode recordings are used to monitor neuronal activity from outside the cell. It can measure patterns of action potentials within many areas of the peripheral and central nervous systems. The neuronal activity is recorded once the electrode is positioned close to a neuron (Ellenbroek et al., 2010).

stimuli with their (social) outcomes through Pavlovian conditioning when exploring their environments. These valuable stimuli promote advantageous behavior directed at fulfilling the "wanting" need with the "liked" social interaction by enhancing the likelihood of finishing the behavioral repertoire of the social incentive salience cycle (Figure 1). The following section highlights one of the most fundamental findings in neuroscience: the discovery of how dopamine neurons activity drives the formation and strength of such CS to appetitive or (social) US association during Pavlovian conditioning.

1.6 Dopamine as a social incentive-teaching signal. Section 1.5 discussed how rats could learn to make social and food reward predictions, namely, to learn that a CS+ predicts the value of both social and food outcomes. In contrast, a CS- predicts the absence of an outcome. The conditioned stimulus serves as a reward predictor and initiates, in our case, social or food approach behavior upon presentation. Importantly, the animal learns to maximize rewards and minimize punishment (absence of reward). It, therefore, obtains what is desired and avoids what is not wanted. This is the fundamental goal of reinforcement learning (Maia, 2009). The prediction error is the key teaching signal used to drive learning and updating. This prediction error is elicited when the animal experiences a reward (US: the social or food outcome in our case) that they did not (fully) anticipate in the presence of a neutral cue (CS) and is thought to originate because of a discrepancy between what is expected and what actually occurs (Nasser, Calu, Schoenbaum, & Sharpe, 2017). It can drive learning about neutral cues in positive and negative directions. Dopamine neurons in the ventral tegmental area (VTA) are activated by rewards (no prediction, reward occurs). This enhanced activity transfers during learning from when the primary reward is presented to the moment that the reward predicting stimulus is presented. Importantly, the neurons are not activated by the reward when the reward is fully learned (reward predicted, reward occurs). Finally, the neurons can show a dip in activity when a reward is omitted (reward predicted, no reward

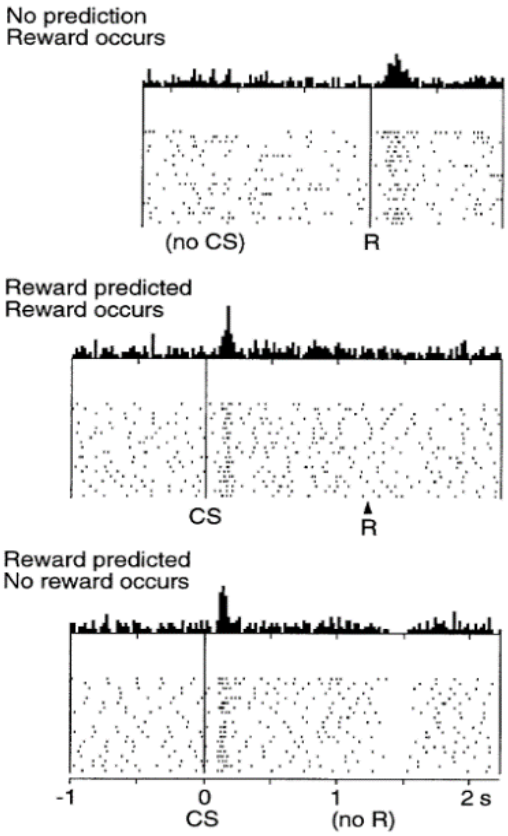


Figure 2. Bidirectional prediction error. Reprinted from Schultz, Dayan, & Montague (1997).

occurs). These features represent the bidirectional prediction error (see Figure 2) that drives a form of learning that attributes motivational value to neutral cues (W. Schultz, Dayan, & Montague, 1997). The dopamine response correlates the hedonic outcome with its preceding neutral stimulus. After learning, the CS then activates the dopamine neurons, which activate the incentive motivational properties that drive an appetitive approach behavior (Berridge & Robinson, 1998). These VTA dopamine neurons project to many brain regions, including the nucleus accumbens (NAcc; Phillipson & Griffiths, 1985). Knutson et al. (2001) found that cues predicting increases in reward magnitudes in people cause increased self-reported happiness ratings. The rating increase is associated with increased NAcc activity visible in the BOLD signal (definition; see Table 1) contrast score. Reward predictive cues that trigger an incentive response to gain a reward are impaired in rats when injecting D1 and D2 dopamine antagonists into the NAcc (Yun, Wakabayashi, Fields, & Nicola, 2004). In the same study, it was found that injecting the GABA B agonist baclofen into rats' VTA decreased the firing of (dopamine) neurons in NAcc, causing a decreased incentive response. Yun, Wakabayashi, Fields, & Nicola did not differentiate between dopamine or other neurons, but earlier studies had already shown that injecting baclofen in the VTA reduced dopamine release in the NAcc (Westerink, Kwint, & DeVries, 1996).

Similarly, NAcc activity correlates with people's incentive motivation for gaining social rewards. For example, people who see another person (that they empathize with) win a game experience increasing ventral striatum fMRI activity more than seeing a non-relatable person win (Mobbs et al., 2009). Moreover, in a study by Jones et al. (2011), it was found that human participants who often received positive social feedback for specific cues developed faster reaction times and higher ratings of likability for these cues than participants who received positive social feedback only rarely. The researchers then calculated the prediction error signal (Delta; actual social feedback—predicted social feedback) and found that ventral striatum BOLD signal activity correlated with this signal. In another study by Fliessbach et al. (2007), two people estimated the number of dots on a screen. When both were correct (1:1), either they were both rewarded with the same payment or the other participant received more than the self (1:2), and vice versa (2:1). Fliessbach et al. found that the BOLD signal in the striatum was the highest for the self > other payment, which was higher than equal payments, which was again higher than when more payment was given to the other. This study, therefore, has shown a direct social relative income effect on the ventral striatum. In rats, there is some evidence that reward presentation to others influences the neurons dopamine response in the ventral striatum (Figure 3A). When an actor rat observed a partner rat receive food, this triggered a greater release of dopamine in the NAcc than when the reward was delivered to an empty box. This observation of rewards delivered to others was associated with an

increased positive 50 kHz vocalizations (Kashtelyan, Lichtenberg, Chen, Cheer, & Roesch, 2014). Further clear evidence that the VTA to NAcc circuit is essential for driving social associative learning stems from Gunaydin et al.'s (2014) work that used fiber-photometry (definition; see Table 1) to record dopamine neurons. They found that dopamine neurons encode (Figure 3B) and predict social interaction. Furthermore, the optogenetic (definition; see Table 1) activation of VTA-dopamine neurons can drive interaction; more specifically, only activating specific VTA to NAcc projection neurons increases social interaction. These findings provide evidence that the dopamine circuit from the VTA to the NAcc is crucial in driving cue-activated incentive motivations directed at obtaining positive non-social and social rewards. Section 1.7 will show that next to the fact that rewards delivered to others are "wanted," the behavior of the other can also change the behavior of the self.

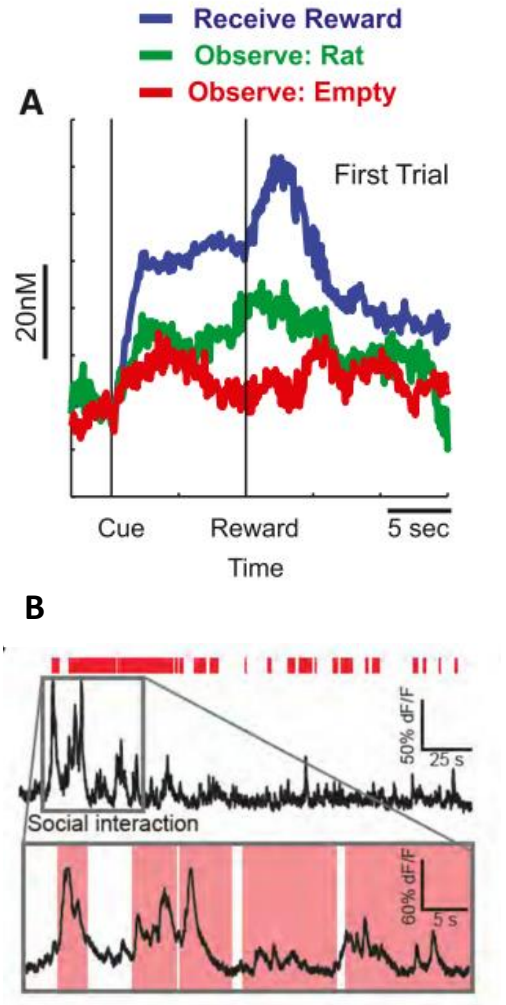


Figure 3. Dopamine and social reward. (A) Observing a partner rat receive food (green) triggers dopamine release. (B) Dopamine Fiber-photometry signal is enhanced during social interaction. (A) reprinted from Kashtelyan et al., 2014. (B) reprinted from Gunaydin et al., 2014.

1.7 Vicarious reinforcement: learning from the behavior of others. A rat can also show "a change in its behavior as a function of witnessing the consequences accompanying the performances of others," which Albert Bandura defined as vicarious reinforcement, causing the observer's behavior to either increase or decrease (Bandura, 1971). One example of this is a study in which a conspecific rat signals a change in the type of food that could be found in the immediate environment. Galef (2001) extensively investigated how these kinds of social signals influenced the food consumption of conspecifics; his main discovery was that an observer rat who interacted with a demonstrator rat developed a preference for the food (i.e., cinnamon) that the demonstrator had recently eaten in a separate chamber. Galef found strong evidence that the transfer of knowledge was based on olfactory information since when they made the observer anosmic (unable to smell; done by rinsing the nasal passages with a zinc sulfate

solution), the preference was absent. Conversely, even when Galef anesthetized the demonstrator rat and taped it to a Petri dish, the observer still had an enhanced preference for the partner's diet. Other important work involving observational learning, that is, the acquisition of information, skills, or behavior through watching the performance of others (APA dictionary of psychology, n.d.), comes from work by Zentall and Levine (1972). They found (see Figure 4) that water-deprived rats learned to press a bar for a water reward faster when they observed another rat pressing a bar and showing a consummatory drinking response as well (OB). Observation of both the action of pressing the bar itself and the drinking response was crucial because when the rat only observed the other rat drinking (OD), its response was similar to when another rat present did not drink or press a lever (OD OE). More evidence that rats can learn an action-outcome association from another rat comes from a study by Heyes and Dawson (1990). They showed that an observer rat, who observed another demonstrator rat perform a directional push on a joystick, pushed in a direction similar to the demonstrator when placed in the same position as the demonstrator rat who demonstrated the movement. In a follow-up control study, the researchers showed that the demonstrator rat must perform the directional push since when the joystick moved automatically without the demonstrator rat present, the observer rat had no enhanced directional response (Heyes, Jaldow, Nokes, & Dawson, 1994). Therefore, observing the others' action-outcome behavior enhanced learning, rather than just the mere presence of the other rat or the observation of the rewarded stimulus itself (rewarded joystick, CS+) signaling the reinforcement.

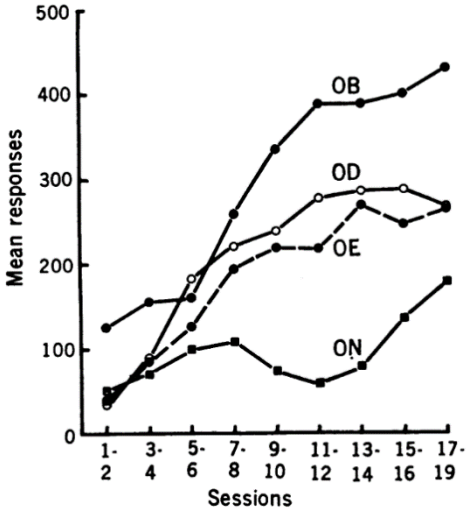


Figure 4. Observational learning. Learning curve for rats that press a bar that leads to a water reward (OB = observe drinking and pressing, OD = Observe drinking, OE = observe empty, and ON = No observation). Reprinted from Zentall and Levine, 1972)

A change caused by observing action-outcome associations alone is not the only driver of vicarious learning. Rewards given to a significant other that are predicted by neutral cues also drive vicarious learning. Chang et al. (2011), for example, found that monkeys (rhesus macaques), who first learned that specific cues predicted either a self-reward, both-reward, a reward for the other, or no reward for both, subsequently preferred the cue that predicted a reward for the other over the cue that predicted a reward for no one. Crucially, this preference was absent when a juice collection bottle replaced the monkey in the other reward condition.

Chang et al. further confirmed that the monkey looked more frequently at the other monkey after choosing to reward the other in the other/none condition, indicating that social attention plays an important role in vicarious reinforcement learning.

Interestingly, in another study by Azzi, Sirigu, & Duhamel. (2012), it was found that cues, which predicted mutual reward delivery, were less valuable for rhesus macaques than cues predicting a self-reward only. However, they also found that the macaques showed a preference for receiving a reward together with one macaque, but not with another macaque, potentially because the social status of the less preferred macaque influenced the preference. In an earlier study, it was found that unrelated chimpanzees do not prefer consuming a both-reward together over self-rewards (Silk et al., 2005). Horner et al. (2011) hypothesized that this absence was caused by other factors such as the task's complexity and the lack of communication between the chimpanzees. Horner et al. used a simpler design in their prosocial choice task where the chimpanzees could directly interact with each other. The chimpanzees could choose between a token that predicted a mutual both-reward or an own-reward. A human would hold up two hands for the both-reward, and consequently, a wrapped food reward was presented to the actor, and directly afterward, another food reward was given to the partner chimpanzee. Only one hand was presented for the own-reward, and only the actor chimpanzee was given a reward. With this setup, Horner et al. indeed showed that chimpanzees have prosocial tendencies, as they preferred that the food reward be given to both the actor and partner chimpanzee over only an actor reward. Overall, these experiments demonstrate that associative learning of vicarious cue-outcome associations was present in monkeys as the cues induced choice behavior that signaled either an increase or decrease in preference for cues that predict social outcomes to others. In rats, prosocial choice was investigated in two different tasks incorporating prosocial choice in an observational action-outcome learning paradigm (Márquez, Rennie, Costa, & Moita, 2015) versus a more cognitive spatial action-outcome learning paradigm (Hernandez-Lallement, van Wingerden, Marx, Srejic, & Kalenscher, 2015). In Márquez et al.'s (2015) task, a recipient rat showed a nose poke, indicating that it was seeking food, which the actor observed. Afterward, the actor rat decided to either feed him- or herself (self-reward) or be rewarded together with the other rat (both-reward). Rats showed a clear prosocial preference, preferring a both-reward (see Figure 5A) over a self-reward. When the rats could not observe the food-seeking behavior or when the recipient learned that food-seeking indication on a specific side would lead both sides to receive a both-reward, the preference was absent. In the task by Hernandez-Lallement et al. (2015), the prosocial choice was indicated by choices in a double-T maze, where the choice to go left or right indicated either a both-reward or a self-reward. After the actor entered their chosen arm, the partner also entered their arm on

the other side of the mirrored maze (see Figure 5B), and they were both rewarded. In contrast, for the self-reward option, only the actor was rewarded, not the partner.

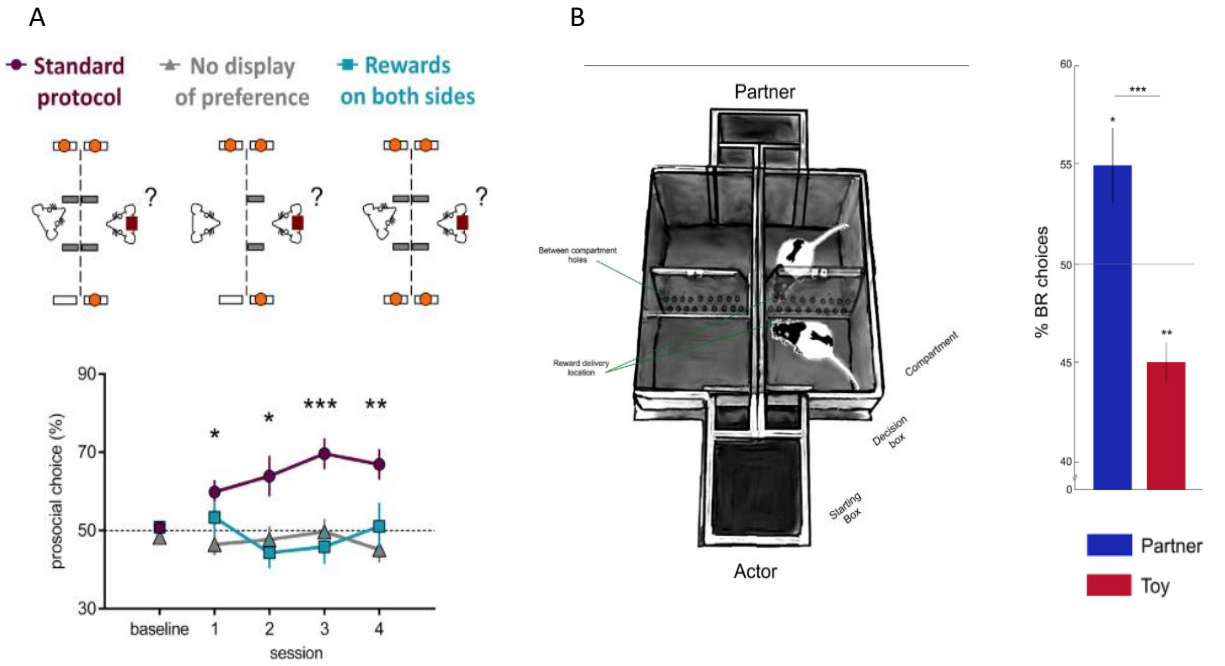


Figure 5. Pro-social choice. (A) In the study by Márquez et al. (2015), prosocial choices were observed when the recipient rat indicated that it was seeking food. (B) In the pro-social choice task by Hernandez-Lallement et al. (2015b), prosocial choices were observed when a partner was present and receiving food pellets together with the actor, but not when a toy was present. Reprinted from Hernandez-Lallement et al., 2015b and Márquez et al., 2015.

In this paradigm, rats also preferred to both be rewarded over self-rewarded, while the preference was absent when only a toy rat was present during the choice. In the first task, preferences for prosocial outcomes were more significant, most likely because it is easier for rats to learn social outcomes using observational action-outcomes association versus a more cognitive spatial action-outcome association in which learning is partly dependent on how well spatial learning can aid decision making. Overall, these two tasks provide evidence that rats demonstrate prosocial tendencies in which they care for the wellbeing of the other. Furthermore, Avital et al. (2016) developed a social cooperation task that extended the above findings by showing that rats also perform joint actions to achieve mutual benefit by learning to coordinate their movement through a social maze. These paradigms all establish that rats can learn from others in various ways and seem to prefer mutual rewards. Prosocial choice-related action-outcome associations were studied in these paradigms, where the social outcome was coupled with the prosocial or cooperative action

itself. In rats, unlike the rhesus macaque, it is currently unknown how vicarious stimulus-outcome associations are formed. While simpler forms of social associative learning discussed earlier showed that rats do indeed make Pavlovian social stimulus-outcomes associations, it is unclear whether rats would also prefer cues that predict mutual rewards over cues that predict self-reward. Section 1.8 discusses why a famous task from the reinforcement-learning field is optimally suited to investigate this question.

1.8 Introducing a social reinforcement-learning task. Previous sections have highlighted how social rewards drive attention and how learning about social rewards involves an incentive motivation framework as social rewards contain both a wanting and a liking component. Rats can furthermore learn to associate cues with their social rewards, and this reward learning is driven by a VTA to NAcc circuit that enables neutral cues to acquire social motivation similarly to non-social cues. Moreover, rats can learn vicariously from others by observing actions that lead to specific outcomes and prefer outcomes resulting in a mutual both-reward rather than outcomes resulting in an own-reward. However, it is currently unknown how the more complex vicarious associative reinforcement learning works in rats and which circuitry is involved in attributing vicarious social reward to neutral cues. I, therefore, developed a novel task in which rats learn to associate vicarious rewards with neutral stimuli based on the blocking and unblocking effect. To understand the task, I will first explain the blocking effect and subsequent discoveries of the unblocking effect.

The blocking effect was discovered by Kamin (1969) and involved three experimental steps. First, an animal was conditioned to a simple CS, consisting of element A. Second, the animal was conditioned to respond to a compound consisting of A presented and B (compound phase). Third, the animal's response to B alone was tested (probe trial). The question was: will the animal respond to element B? In his task, Kamin found strong evidence that element B would be completely blocked from conditioning if A were conditioned. Meanwhile, B was conditioned when A was not fully predicting the outcome before compound training. In this specific case, animals first learned to press a lever steadily. Afterward, they were presented with a three-minute noise CS ending with a 0.5 millisecond 1mA shock (US) during the lever presses. After finishing 16 trials during which the tone CS–shock association was learned, the rats would be presented with a compound consisting of the same noise plus an added light cue, again ending with a shock for eight trials. Finally, the light was presented alone without a shock in the probe trial. Kamin found that the rats started to fully inhibit their bar pressing already on the fifth trial during initial conditioning and compound conditioning when presented with noise A. This inhibition, however, did not occur when the rats were presented with the light in the probe trial; the light was blocked from conditioning.

This empirical finding was very influential in the animal conditioning field and led many to theorize why the animal did not learn about the second stimulus. Mackintosh (1975), for example, theorized that the A element was more informative and salient than the B element, and therefore the B element was swiftly losing its salience over the trials due to a rapid decline in attention. In contrast, Wagner & Rescorla (1972) theorized that if the associative strength of the CS (here A) that predicts the US was high, then the added compound stimulus reinforcement strength would be low and vice versa. Dickinson et al. (1976) challenged this theory with an experiment, in which they found that after an element A predicted a double shock, and the following compound predicted only one shock, presenting the element B was less effective in blocking. They concluded, therefore, that unblocking had occurred. Holland (1984), however, interpreted this finding as an extension of the Rescorla and Wagner model, hypothesizing that even if the reinforcement strength of a stimulus A is high, if there is a discrepancy between a new US B and the old highly reinforced US A, conditioned reinforcement (positive or negative) should be a function of the difference between the two USs. Holland's (1984) experiment formally tested this idea (Figure 6A). First, the rats received four presentations of a house light for 10 seconds. At the end of the presentation, one pellet was delivered (Phase 1, A+). The experimental group subsequently received a compounded stimulus X and then received one pellet followed by two more pellets 5 seconds later (Up: Phase 2, AX++). Finally, the rats were tested in the probe trial on the X cue alone without a reward (Test, X-). Importantly, in the control condition (Up/C), there was no upshift. What Holland found was that the rats made more head jerks (short fast movement directed at the food cup; Figure 6B), but also startled more (rapid jump or change in position) and made more magazine responses (standing motionless with the head in the food cup) for the upshift cue than the upshift control cue (Up/C). These results indicated that unblocking had occurred as the head jerk was found to be the main conditioned measure for auditory unblocked cues (Holland, 1977). Holland

A

Design of Experiment 1

Group	Phase 1	Phase 2	Test
Up	A+, B++	AX++, B++	X−
Up/C	A+, B++	A++, BX++	X−
Down	A+, B++	A+, BX+	X−
Down/C	A+, B++	AX+, B+	X−

B

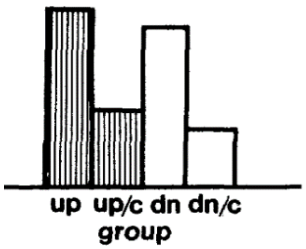


Figure 6. Appetitive unblocking. (A) Design of the appetitive unblocking paradigm. (B) Head startle during test trials. Reprinted from Holland, 1984.

(1984) concluded that the unblocking procedure facilitates the formation of an association between the US (added pellets) and the added cues.

Previous sections have shown that a social US has all the components necessary to accommodate the formation of a social US with added cues during unblocking. Social USs drive social attention, fit into the framework of incentive salience, and drive conditioning in simple social Pavlovian conditioning. The dopaminergic reward circuitry drives these social stimulus-outcome associations. Finally, rats prefer actions that lead to a both-reward over an own-reward. It thus stands to reason that social rewards can replace own-rewards in unblocking and drive the social unblocking effect. In Chapter 2, a social reinforcement task will be introduced that investigates whether mutual reward outcomes (i.e., reward outcomes to both self and another rat; Social Up) indeed form an association with an added CS during compound training that is higher than the association with an own-reward outcome (i.e., the actor is rewarded while the partner is not; Social Up/C). While it is clear from studies that the dopaminergic system is critical in forming associations in appetitive unblocking (Keiflin, Pribut, Shah, & Janak, 2019), Section 1.9 dives into the role of the OFC in unblocking and its possible role in social unblocking.

1.9 The neural circuit that drives (social) unblocking: a specific role for the orbitofrontal cortex? To understand the OFC's role in social learning, we start by addressing the two brain structures heavily involved in learning that a CS becomes associated with social and food outcomes. The CS that comes to predict the US acquires the incentive salient properties of the US that is driving approach and consumption. In addition to transferring the positive value to the CS via dopamine activity in appetitive learning, the characteristics of the US are also important. The basolateral amygdala (BLA) is prominently involved in connecting the US qualities with the CS that predict them. The BLA is necessary for a CS to gain access to the subjective value representing the outcome's incoming sensory properties, for example, during licking or chewing, while the central part is more relevant for conditioned motivational influences on behavior (Balleine & Killcross, 2006).

McDannald et al. (2012) have specified that learning can be based on model-free and model-based systems. They propose that model-free learning indicates that the (total) value is represented in a common currency devoid of specific references to the specific form and features of the reward it predicts, for example, when a neutral cue predicts a value of zero and subsequently is coupled with a food reward (one food pellet) until the cue fully predicts the reward. With model-based changes in value, McDannald et al. mean behavior that involves making a cognitive model of all stimuli and events in the environment and using that knowledge to predict future value. A change in a

reward's incoming sensory properties, such as during the unblocking procedure, can sometimes indicate a change in the value, which has both model-free and model-based properties, or a more pure model-based change. The full model-based change in identity features involves substituting two equally preferred rewards with different flavors, such as changing a grape-flavored pellet reward to a banana-flavored one. A combination of model-free and model-based learning occurs when one pellet reward changes to three pellets; here, the model-free-based change indicates an increase in the common currency of the reward (one to three). At the same time, the three pellets' actual experienced outcome features have a distinct model-based representation that indicates a sensory perception of chewing more pellets. The new learning in the form of unblocking can then occur after a new stimulus is compounded on a fully learned stimulus and predicts those changes in outcome features (Rescorla, 1999) or value (Holland, 1984) of the US.

The brain's circuitry that drives different aspects of unblocking involves the basolateral and central amygdala, ventral tegmental area, and the OFC. Lesioning the BLA impairs unblocking based upon a change in the identity, as when rats fully learned that a visual cue predicted a one-pellet reward. When they subsequently added a compound cue signaling a sucrose water reward, they showed no unblocking, while control rats with a sham lesion did. However, when the compound cue signaled a change to only more pellets, unblocking in the lesioned rats was not impaired (Chang, McDannald, Wheeler, & Holland, 2012). Conversely, the central amygdala is only involved when the change during unblocking is a downshift, not an upshift (Holland & Gallagher, 1993b). In two recent studies, unblocking was investigated using transgenic rats, where Cre-recombinase was expressed under the control of the tyrosine hydroxylase promoter (cf. Witten et al., 2011). Using this genetic technique, researchers can target the VTA dopamine neurons specifically and activate only them using optogenetic stimulation via chronic optical fiber implants (optogenetics; definition, see Table 1). Keiflin et al. (2019) and Steinberg et al. (2013) found that if they replaced the normal increase with more food pellets (going from one to three) during unblocking with optogenetic activation of the VTA cells, they could mimic the unblocking in those rats that expressed the Cre-recombinase (Cre+), but not in their littermates (Cre-) or when they performed the stimulation in the Cre+ animals during the intertrial-interval. Keiflin et al. (2019)concluded that the value associated with the sucrose's features (its identity) was unblocked since when they devalued the sucrose after optogenetic stimulation already occurred, the unblocking was eliminated. The OFC processes information such as the specific expected outcome's size and identity and the relative preference for outcomes. Furthermore, it facilitates updating these predicted outcome components when contingencies change (such as devaluation; Schoenbaum, Roesch, Stalnaker, & Takahashi, 2011).

McDannald et al. (2011) found that the lateral part of the OFC is necessary to update changes in identity features (going from a grape to a banana flavor) but not changes in value (going from one to three pellets) during unblocking (Figure 7A, B, C, and D). During the probe trial, responses to the Y cue, which signaled a change in features during compound conditioning, diminished in lateral OFC-lesioned rats but not in control animals (Figure 7). Other studies using extracellular electrophysiology (definition, see Table 1) have found that the firing rate of posterior-lateral OFC neurons increased in the first day of unblocking for olfactory cues that predicted a "valueless" change in flavor identity (McDannald et al., 2014) and for cues that predicted a pure upshift or downshift in the reward value (smaller or larger milk drops, not more or fewer drops; Lopatina et al., 2015). Conversely, the medial OFC showed mainly downshift-related changes in firing rates (Lopatina et al., 2016). It is, therefore, clear that the OFC anatomy (Izquierdo, 2018, see Figure 7E) partially defines how the reward's identity (taste) needs to be updated in a positive (lateral OFC) or negative (medial and lateral OFC) direction. Interaction between the BLA and the lateral OFC could very likely be essential for assigning enhanced positive value to the correct new stimulus identity. Another paradigm where this connection is also essential is the specific

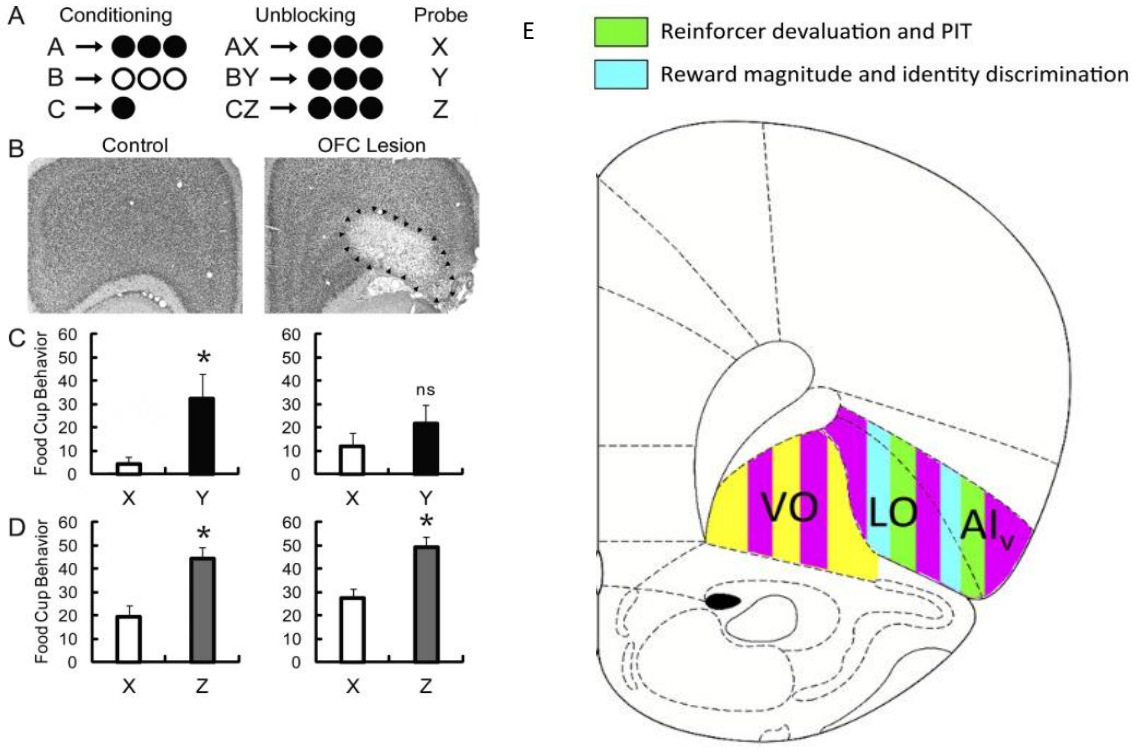


Figure 7. Identity unblocking in the rats' lateral OFC. (A) Appetitive unblocking task design. (B) Lateral OFC lesion (C) Lateral OFC lesion impairs identity unblocking (D) Lateral OFC lesion does not impair value unblocking. (E) Anatomical position of the posterior-lateral OFC and its function. (A, B, C, and D) Reprinted from McDannald et al., (2011) and (E) Reprinted from Izquierdo, 2018.

Pavlovian-to-instrumental transfer (specific PIT), where a rat learns to associate one tone with a sucrose reward and a noise with a pellet reward outcome. Afterward, the rat learns that a left lever press would result in a pellet reward and a right lever press with a sucrose reward. In a final session, researchers play back the sounds during the lever press, which typically results in an enhanced response for the left lever if the noise is played, but not when the tone is played and vice versa (an enhanced response for the right lever is shown when the tone is played). It is thus critical that the stimulus-outcome matches the action-outcome for specific PIT to work. A lateral OFC lesion impaired the typically observed enhanced response of the specific Pavlovian cue on its matched instrumental response (Ostlund & Balleine, 2007).

Furthermore, when inactivating BLA terminals in the lateral OFC, this specific PIT, as mentioned above, is impaired, possibly due to an impairment of the cue to trigger increased motivation (Lichtenberg et al., 2017). This finding is a further refinement of an earlier finding from Burke et al. (2008), which showed that lateral OFC-lesioned rats, in addition to causing an impairment identity by unblocking added cues that predicted a change in the flavor of the outcome, did not show an enhanced lever press response for the unblocked cue over the blocked, while the rats in the control group did.

We can conclude from these experiments that a circuit including the BLA, VTA, and lateral OFC plays a crucial role in specific model-based unblocking. While the VTA signals both value and identity signals, the BLA and the OFC are more important for making a cognitive model of the identity features and their associated predicted reward outcomes. In Chapter 2, we found that a vicarious reward unblocks associative learning regarding novel cues. In the task, rats learned to associate a novel compound cue during unblocking with three conditions: 1) a both-reward condition where both an actor and a partner were rewarded, 2) an own-reward where the actor rats were rewarded, and 3) a no-reward where neither rat was rewarded. A probe test found that the cue predicting a both-reward showed a higher conditioned response than the cue predicting an own-reward or no-Reward. This social unblocking effect, where the both-reward has a higher value for the rat than the own-reward, could likely be orchestrated through the above described lateral OFC to BLA circuitry. BLA lesions impair prosocial choices in the prosocial choice task where a both-reward is preferred over an own-reward (Hernandez-Lallement, van Wingerden, Schäble, & Kalenscher, 2016b). The BLA is reciprocally connected to the lateral OFC (Barreiros, Panayi, & Walton, 2021; Lichtenberg et al., 2017). Mutual rewarding outcomes are potentially learned through observing the specific outcomes of others since presenting a toy rat during mutual reward presentation prevents learning (Hernandez-Lallement et al., 2015). This information from the BLA could well be integrated, and lateral OFC could likely provide a cognitive model for the

social reward outcomes. If this mutual versus self-reward outcome is coded as two specific actions like during the specific PIT mentioned above, mutual rewards and the positive social reward and feedback related to the identity model (mutual self and other vs. self-only reward) of the reward could enhance the motor response. According to the social reinforcement learning hypothesis dictated by Hernandez-Lallement et al. (2016b), communicative feedback during mutual reward learning could signal the outcome's identity and drive the formation of these new social actions and stimulus-outcome associations.

A role for the OFC in encoding such social reward signals comes mainly from a study by Jones et al. (2011). They found that people who received continuous positive social feedback from peers during the presentation of stimuli displayed faster reaction times to those stimuli than to stimuli that predicted positive reinforcement from peers less often. The researchers then calculated a prediction error from the difference between the experienced positive or negative feedback and the expected outcome for each peer. They found that activity in the lateral OFC was positively associated with the social prediction error. The rostral cingulate cortex was furthermore found to code the feedback's expected value. This finding is further established by Chang et al. (2013), who found that cues predicting a reward for another monkey correlated with increased activity in the anterior cingulate, while lateral OFC activity was mainly correlated to self-reward coding (Watson & Platt, 2012). Another paper, however, has found increased activity in the lateral OFC when rewards are mutual, but only if the monkey is the preferred partner (Azzi et al., 2012). It becomes clear that the lateral OFC has a more prominent role when the value of the social feedback from others is defined by the correct assignment of predictive value, taking into account the outcomes' social cognitive model. In the above-described paper, the social cognitive model depended on assigning the accurate chance to cues that predict positive social reinforcement by peers. In tasks with monkeys, the cognitive model must include the identity of the rewarded monkey (who that monkey is in the hierarchy could influence future food-sharing behavior). We can conclude that a mutual reward identity signal is likely mediated by the BLA and lateral OFC and is hypothesized to be necessary for driving model-based prediction errors if a social cognitive model that includes positive social feedback or a social reference frame is involved in the learning.

Another hypothesis is that these model-based predictive signals do not drive social unblocking, but rather the OFC cortex is only necessary when the model-based social predictive value must be rapidly updated. In identity-unblocking, a different flavor is still desired even though the value is the same as the other flavor (i.e., the rat shows a conditioned approach to the cue representing the different flavor). This identity unblocking has much in common with reversal learning where stimulus-reward contingencies are changed (a rule switch). It was recently discovered that in rats,

long-range projections of lateral OFC neurons to the somatosensory cortex (S1) were important for updating the sensory representation of stimuli to signify the correct value of either a no-go cue or a go-cue assigned to it (Banerjee et al., 2020). If the identity value is the same for differently flavored foods, but the features are different, remapping the correct stimulus-outcome associations must still occur. The OFC plays an essential role in this remapping, as has been found in extensive research on its involvement in reversal learning (Schoenbaum, Roesch, Stalnaker, & Takahashi, 2011).

Another experiment by Walton et al. (2010) shows that this stimulus-outcome association remapping is related to assigning the proper credit to the right stimulus. The correct credit assignment refers to the consequence of a particular choice among several alternatives. Walton et al. found that monkeys with a lateral OFC lesion have problems updating the value of their choice for specific stimuli when the environment signaled a rapid value change of that specific stimuli. The lateral OFC-lesioned monkeys showed this deficit because their responses were based more on integrating the value of recent choices than on the actual consequence of each choice. This response deficit indicates an unstable stimulus-outcome mapping due to the incorrect assignment of credit to each cue in the lateral OFC-lesioned monkeys. The integration of a self-reward with a no-reward or reward to another rat may be impaired when lesioning the lateral OFC. It could likely be due to the inability to assign a self-initiated approach value, the value associated with another rat approaching or not approaching a food cup, when reward presentations rapidly change during social unblocking. Chapter 2 will show whether the posterior-lateral OFC has a role in unblocking cues by representing more model-based prediction errors involving a social cognitive model that includes positive social feedback or a social reference frame indicating mutual reward, or whether the posterior-lateral OFC is involved in correctly assigning credit to the different social stimulus-outcome associations.

1.10 Directly comparing social versus non-social rewards and the role of ultrasonic vocalizations. While the unblocking paradigm is a good method to isolate new social learning about novel stimulus-outcome associations and how they are formed, it is a paradigm where the rat does not choose to turn on the light to move to a place where the two rats can both be rewarded. Every stimulus that occurs is a surprise, and learning is driven by a motivational drive initiated after learning that occurs as soon as the cue is triggered. Conversely, motivation is often self-initiated out of a desire to fulfill a social need for interaction. We have previously discussed that when searching for such social outcomes, both social and non-social stimuli found in the environment can result in those stimuli and places becoming conditioned. A conditioned cue predicting sucrose water, for example, can lead to an increased motivation for feeding when the

cue is presented (Weingarten, 1984). In contrast, a place that predicts social contact leads to social approach behavior (Nadler et al., 2004), indicating the formation of a social place preference (Van Den Berg et al., 1999).

However, there is a lack of precise scientific knowledge regarding preference for social rewards relative to food rewards. Padoa-Schioppa and Assad (2006) have found that one reward can be quantified in units of the other: monkeys choose rewards by assigning a general value to stimuli that predict food rewards and decide between different, varying quantities of both rewards. Theoretically, animals should express similar choice preferences when confronted with a choice between food and social contact, leading to the intriguing possibility that one can identify an indifference point where the decision-maker chooses the different options with equal probability, similar to what has been found in delay discounting (Richards, Mitchell, de Wit, & Seiden, 1997). That rats can indeed indicate a preference for food with different hedonic levels comes from studies into the licking of sucrose water with varying concentration. Spector et al. (1998) found that rats increased their average licks per minute, going from a low sucrose concentration to a higher one (10%). Fonseca et al. (2018) have shown that next to changing their lick rate based upon sucrose percentage, rats can also discriminate between a high (18%) versus low (3%) amount of sucrose. Furthermore, when given the choice of other concentrations, rats showed an increased preference of 18% > 11.75% > 7.5% > 4.75% > 3%, implicating that the rats use the sucrose intensity to solve the discrimination problem. This increased preference is caused by a positive hedonic response to the sweet taste, which is mediated by the activity of the brain's pleasure circuitry and is accompanied by a behavioral response of facial relaxation and smiling (Berridge, Robinson, & Aldridge, 2009). This preference for sweet food is pronounced in young adults but becomes less significant in adults (Desor & Beauchamp, 1987). Furthermore, this inclination is strong enough for young adult rats to choose a 0.2% saccharine reward, when given the choice of selecting to press a lever associated with a 0.2% saccharine reward vs. a lever associated with an infusion of cocaine or heroin (Madsen & Ahmed, 2015).

Evidence that the social context of food consumption can influence food choice comes from Birch et al. (1980), who found that children preferred food more when the food presentation was given contingent on positive social feedback than when food was found only at a specific place. Another study found that perceived loneliness was associated with increased sucrose beverage consumption (Henriksen, Torsheim, & Thuen, 2014). In monkeys, the influence of social stimuli on food preference can be investigated by asking a monkey how much tasty juice rewards they are willing to forego to watch rewarding pictures of conspecifics (Deaner et al., 2005). With rats, however, an indication of preference can only be deduced by looking at specific movement patterns, choice

allocations, and ultrasonic vocalizations. 50 kHz vocalizations are a call type defined by a wide range of different sounds modulated in frequency, pitch, and shape and made visible on spectrograms. They mainly indicate positive social behaviors such as play, tickling, and sexual behaviors (Brudzynski, 2013; Brian Knutson, Burgdorf, & Panksepp, 2002). Importantly, when one plays back these kinds of calls to rats, it leads to an approach behavior (Wöhr & Schwarting, 2007) of the rat, and simultaneously, dopamine is released in the NAcc (I. Willuhn et al., 2014). These two elements make 50 KHz calls a good indicator of incentive salience and possibly a signal that could contribute to establishing social CS-US associations. As it is currently unclear how social interaction in rats influences the preferences for sucrose water and what role 50 KHz calls play, Chapter 3 introduces a task that directly investigates how rats choose between sucrose on different levels of sweetness and the concurrent availability on all levels of a social stimulus. This task allowed us to establish an indifference point and quantify the value of the social stimulus in units of the sucrose level. We also established the role of 50 KHz calls, including their subtypes (Wright, Gourdon, & Clarke, 2010), which further elucidates the importance of the subjective valuation of a juvenile when food is present.

1.11 Thesis goal and chapters overview. My goal is to identify the behavioral repertoires and understand the underlying brain substrates underlying valuation of social reward and its functional relevance to ongoing behavior in the rat. To answer these questions I have worked on establishing two new paradigms and performed neuroscientific studies. In the **second chapter**, I have established a task in which the *vicarious* associative value of cues that predict mutual reward versus self-reward are measured. In the **third chapter**, I find that the posterior-lateral Orbitofrontal cortex is not necessary for the formation of the stimulus-outcome association that predict mutual reward, but is necessary for the accurate formation of stimulus-outcome association that predict no additional reward to a partner. In the **fourth chapter**, I have helped establish a task that investigates the influence of different levels of an appetitive sucrose reward on the valuation of social interaction and the quantification of an important role for 50 KHz vocalizations and its subtypes in driving preferences during this task.

2. Vicarious reward unblocks associative learning about novel cues in male rats.

Abstract. Many species, including rats, are sensitive to social signals and their valuation is important in social learning. Here, we introduce a task that investigates if mutual reward delivery in male rats can drive associative learning. We found that when actor rats have fully learned a stimulus-self reward association, adding a cue that predicted additional reward to a partner unblocked associative learning about this cue. In contrast, additional cues that did not predict partner reward remained blocked from acquiring positive associative value. Importantly, this social unblocking effect was still present when controlling for secondary reinforcement but absent when social information exchange was impeded, when mutual reward outcomes were disadvantageously unequal to the actor or when the added cue predicted reward delivery to an empty chamber. Taken together, these results suggest that mutual rewards can drive associative learning in rats and is dependent on vicariously experienced social and food related cues.

2.1 Introduction.

Humans and other animals have developed a capacity for mutual cooperative behavior (Nowak, 2006; Rand & Nowak, 2013; Rilling et al., 2002; Suchak, Eppley, Campbell, & de Waal, 2014), a preference for prosocial outcomes to familiar partners (Hernandez-Lallement et al., 2015; Horner et al., 2011; Márquez et al., 2015) and helping behavior towards others in need (Ben-Ami Bartal, Decety, & Mason, 2011; Fehr & Rockenbach, 2004). These behaviors are sometimes costly, prompting questions why actor engage in them (F. B. de Waal & Suchak, 2010; Hamilton, 1963; Stevens, Cushman, & Hauser, 2005; Trivers, 1971). Some researchers have focused on putative future reciprocation (Taborsky, Frommen, & Riehl, 2016) as a potential driver, while others have highlighted that acting generously could generate self-reward internally (Harbaugh, Mayr, & Burghart, 2007; Park et al., 2017) or through positive social reward signals such as friendly faces in humans (Spreckelmeyer et al., 2009). Indeed, the capacity to identify positive, rewarding outcomes delivered to others is a fundamental aspect of social observational learning (Thomas R. Zentall, 2012). Underlying some of these suggestions is the assumption that rewarding outcomes to a social partner could also represent value to oneself and thus drive a proximate reward/learning mechanism (Hernandez-Lallement et al., 2016a; Ruff & Fehr, 2014).

** Chapter 2 is published in eLife: van Gurp S, Hoog J, Kalenscher T, van Wingerden M. 2020. Vicarious reward unblocks associative learning about novel cues in male rats. Elife 9:1–25. doi:10.7554/elife.60755*

By this logic, animals, including humans, choose pro-social outcomes, cooperate or act altruistically because these actions result in vicarious reward, experienced through sensitivity to the behavioral and/or affective state of the partner (F. B. M. de Waal & Preston, 2017; Prochazkova & Kret, 2017), in addition to putative anticipated future reciprocal reward (Taborsky et al., 2016). One important aspect of social learning is identifying the features of the environment that predict (vicariously) rewarding outcomes, and learning the (instrumental) action sequence, appropriate to the context, for acquiring these vicariously rewarding outcomes. There is evidence that cues that predict social reward can become valuable as humans learn to respond faster to stimuli that become associated with positive social reinforcement (Jones et al., 2011) and monkeys preferred stimuli that predicted a reward delivery to a conspecific more than the stimuli that predicted no reward delivery (Chang et al., 2011). In rats, it was found that observing another rat being rewarded is (vicariously) rewarding by itself as it is accompanied by 50 kHz vocalisations, indicative of a positive appetitive state (Burgdorf, Panksepp, & Moskal, 2011b; Panksepp, 2007), and dopamine release in the NAcc of the observer rat (Kashtelyan et al., 2014). Indeed, playback of 50 kHz leads to both an approach response (Wöhr & Schwarting, 2007) and results in dopamine release in the Nucleus Accumbens NAcc (Ingo Willuhn et al., 2014). We therefore hypothesised that vicarious reward, associated with rewards delivered to others, could also reinforce Pavlovian associative learning about novel cues, as has been found in the appetitive domain (Berridge, 2012; Wolfram Schultz, 2016). To investigate our hypothesis, we use a well-established behavioral paradigm in associative learning called blocking. Kamin (1969) found, in simple stimulus-outcome association tasks, that if new stimuli are added to a stimulus that already fully predicts a reward, associative learning about those additional stimuli will be blocked. Reinforcement learning about additional stimuli can become unblocked, however, by an increase in reward value or a change in reward identity contingent on the presentation of the new stimuli. This change in value is then thought to be associated to these new stimuli and thus alters their incentive value (Holland, 1984). We hypothesise that rewarding social outcomes, such as sugar pellet deliveries to a partner rat, will also be capable to unblock learning about novel stimuli added in compound, indicative of an increased, partially vicarious value of mutual rewards relative to own-rewards. We tested this hypothesis by adopting a task from McDannald et al. (2011) where unblocking is operationalised by adding additional pellet deliveries conditional on a second cue presented in compound with a learned cue that already fully predicted reward. We modified this task in such a way that the second cue is now followed by a food reward delivery to a partner rat, rather than increasing one's own reward. In addition, a third control cue added in compound to the learned cue (on different trials) was not followed by food reward delivery to a partner rat. Concretely, we thus hypothesized that associative

learning about the second stimulus would become unblocked through a vicarious experience of the partner reward exclusively during mutual reward outcomes. In contrast, the third cue should remain blocked from acquiring associative value due to absence of a reward outcome for the partner. We indeed found, when tested in extinction, that the unblocked cue had acquired more associative value, as indexed by conditioned responding at the food trough, in comparison to the blocked cue. Importantly, this effect was still present when controlling for potential effect of secondary reinforcement associated with increased pellet deliveries. Crucially, this difference was absent when 1) social information exchange was impeded 2) when the partner rat was absent during mutual reward delivery 3) and when the unblocking cue was associated solely with partner reward but not actor reward, presenting a disadvantageous unequal reward distribution to the actor rat.

We thus conclude that mutual, equal reward delivery can trigger a positive vicarious reward experience that supports unblocking of associative learning about novel cues. This opens up possibilities to investigate behavioral aspects of the social-value driven reinforcement learning and its associated neural basis, processes that might be disturbed in psychiatric disorders marked by impaired reinforcement learning and/or social behavior such as autism (Kohls, Chevallier, Troiani, & Schultz, 2012) and schizophrenia (Fulford, Campellone, & Gard, 2018).

2.2 Methods.

Subjects. 88 male Long Evans rats were housed in pairs of two and kept under an inverted 12:12 h light dark cycle, in a temperature (20 ± 2 ºC) and humidity-controlled (approx. 60%) colony room. All rats had ad libitum access to food, except during the testing period. During behavioral testing, the rats where food restricted (20 grams on weekdays and 22 grams in the weekend) and maintained on a body weight of about 90% of their free-feeding weight. All testing was performed in accordance with the German Welfare Act and was approved by the local authority LANUV (Landesamt für Natur-, Umwelt und Verbraucherschutz North Rhine-Westphalia, Germany).

Apparatus. Testing was conducted in 4 customised PhenoTyper (Noldus Information Technology) behavioral testing boxes (Fig. 1A) of 45 by 45 by 55 cm, supplemented with operant devices (Med Associates) and placed inside a custom-made sound- and lightproof ventilated box. The boxes where modified by adding a custom-made Plexiglas separation wall (Fig. 1A, left panel), which divided the box into two compartments, to allow the training of a pair of rats at the same time. The separation wall was equipped with a sliding door (dimensions: 20 by 20 cm, located at 7

cm from the left side of the Skinner box) and 4 rectangular interaction windows (Fig. 1A, left panel; size: 10 by 1.5 cm) that were positioned exactly in between the door and the wall holding the stimulation devices used for conditioning. Both compartments of the box contained a food trough (Med Associates, ENV-254-CB) positioned in the middle on the right side. The food troughs were adapted in such a way that the detection photobeams were positioned at the entry point of the food trough. The food trough was connected to an automated pellet dispenser (PTPD-0010, Noldus Information Technology) that delivered sucrose pellets (20 mg dustless precision pellets, Bio-Serv, Germany). Operant devices were positioned on the right side of the box at the level of the separation wall: an LED Stimulus Light (Med Associates, ENV-211m) with green cover was positioned 10 cm above the ground and a house light (Med Associates, ENV-215m) 28 cm above the ground. A speaker (Med Associates, ENV-224am) was positioned at 20 cm above the ground for the playback of auditory stimuli (Fig. 1A, right panel). Auditory stimuli were played back at a loudness of 75 dB measured with a hand-held analyser (type 2250-S from Brüel and Kjaer) right in front of the speaker. In the top cover of the Skinner boxes, a camera (Basler, acA1300-60gc, GigE) was positioned to obtain videos of the behavioral experiment at 25 fps. Analyses of the recorded videos was performed with EthoVision XT 11.5 (Noldus Information Technology). Finally, a USV-microphone was positioned next to the camera for recording ultrasonic vocalisations using Ultra Vox XT (Noldus Information Technology).

Pavlovian discrimination task. Before the start of behavioral training, rats were put on food restriction to reduce their weight to 90% of their free-feeding weight. Within a pair of cage mates, one rat was assigned at random as the actor animal, and the other as the partner animals. As a first step, they were habituated to their pre-determined training side of the customised PhenoTyper for 3 days (15 minutes per day). During this period, they could retrieve 6 pellets that were put along the edges of their respective side of the box. Subsequently, the discrimination learning phase started. Here, the pairs of rat cage mates were divided into two groups; one group of rat pairs would learn a visual discrimination problem, and the other an auditory discrimination problem (Fig 1B, left panel). The visual stimuli to be discriminated consisted of a houselights flashing at 1 Hz (0.1sec on, 0.9 sec off) and a steady green light; the auditory stimuli were made up by a 4.0 kHz clicker (0.1sec on, 0.9 sec off) and a 1.5 kHz (75 dB) steady tone (see Figure 1 – figure supplement 1 for overview of stimulus contingencies). The different groups (Auditory vs Visual) were each trained alone either in the upper or lower compartment of the Skinner box, and the side assignments between actor and partner rats were counterbalanced between experiments (Fig. 1A).

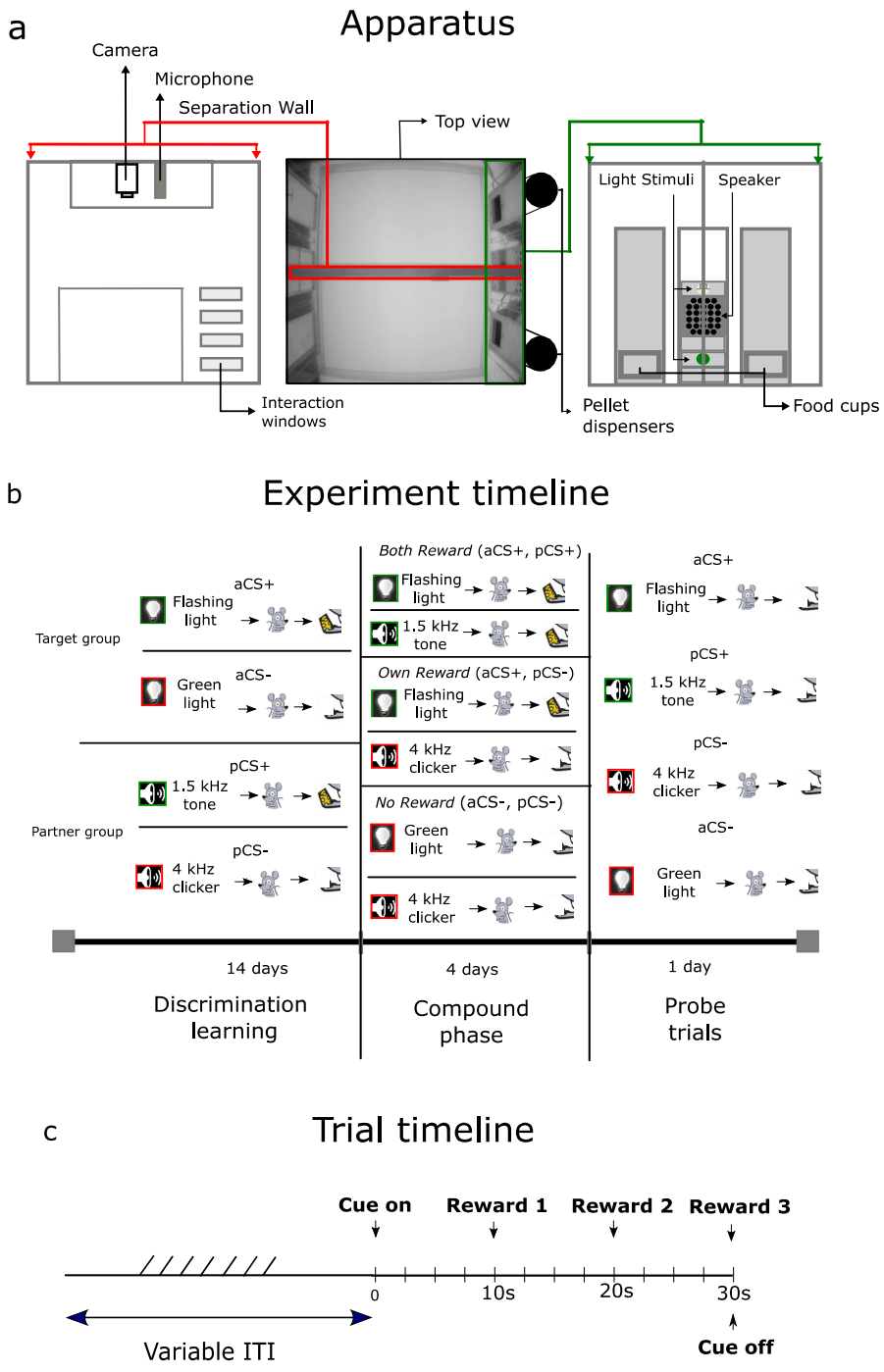


Figure 1. *Behavioural apparatus, Experimental timeline and Trial timeline.* (a) The PhenoTyper consisting of lower and upper compartment in which behavioural training took place is displayed in the middle. On the left the custom made separation wall is shown with interaction windows, camera and microphone. On the right, the right side of the PhenoTyper is displayed with the used operant devices and in both sides of the box the food cup. (b) An example experimental time line is displayed. Actor rats learn to discriminate two visual cues in upper compartment while at a different time partner rats learned to discriminate two auditory cues in the lower compartment. In the compound phase actor and partner rat are either both rewarded (BR, aCS+/pCS+), actor rat is rewarded while the partner is not rewarded (OR, aCS+/pCS-) or both actor and partner rat are not rewarded (NR, aCS- pCS-). In the probe trials all learned cues are presented to the actor and at a different time to the partner rat without reward. (c) Here, a timeline is shown with the different components that make up a single trial throughout the discrimination learning, compound phase and probe trials.

Each rat received 14 days of discrimination training. One daily session consisted of 40 trials, of which 20 trials were aCS+ and 20 aCS-. The order of aCS+ and aCS- trials was pseudo-randomized, with no more than three trials of one kind occurring in a row. Stimuli were presented for 30s and at every 10s (+ 0.1 to 0.4 sec jitter), a pellet was delivered (Fig. 1C). We trained a total of N= 20 actor rat and 20 partner rats on the discrimination problem in the experimental group, 16 actors and 16 partners in the control group 1 and 16 animals (all considered actors) in control group 2 (Unequal Outcomes). The experimental group was divided in subgroup 1A (Social-Appetitive subgroup) and 1B (Social-Only subgroup) and in subgroup 1A (Inserted Wall) and 1B (No Partner Present; See Figure 1 - figure supplement 2 for overview). The Social-Only subgroup consisted of 12 actors and 12 partners of the N=40 experimental group. Here, a second pellet dispenser was already placed outside of the behavioral box during the entire discrimination learning phase, at the opposite side of where the current rat was trained, delivering pellets outside of the box. This additional dispenser placement ensured that the sound of additional pellet drops was similar to the compound conditioning phase (see below). Providing a uniform pellet delivery sound associated with self-reward pellet delivery throughout the experiment prevented any difference in pellet delivery related sounds as a source of secondary reinforcement from influencing the conditioning to added cues in the compound phase. In the additional Social-Appetitive subgroup, the second pellet dispenser was not active during the discrimination phase (Figure 1 - figure supplement 2). This gave us the opportunity to make direct comparison within the experimental group to investigate potential effects of secondary reinforcement (Figure 1 - figure supplement 2). The ITI in both experimental and control group was made up of a fixed 30s window supplemented with a randomized time window ranging from 5 to 100 seconds with steps of 5 ms, uniformly distributed. The ITIs were thus fully randomized, resulting in a total variable ITI with a mean of 80 ms. Ultrasonic vocalisations where recorded from 10s before cue onset to 20s after cue offset, for a total duration of 60s per trial. After completion of the discrimination phase, rats progressed to the compound conditioning stage.

Compound conditioning. After discrimination training was completed, rats in the visual discrimination group received 1 day of pre-exposure to the two novel auditory stimuli while the rats in the auditory discrimination group received 1 day of pre-exposure to the two novel visual stimuli. The pre-exposure session consisted of one session with 6 trials. The stimuli were presented in a randomized order with ITIs of 15, 30 45, 60, 75 and 90s. Pre-exposure was done to minimize an influence of novelty induced enhancement on the conditioning of added compound stimuli (Holland & Gallagher, 1993a) and enhance the discriminability of these added stimuli (Honey & Hall, 1989) for each group. This would strengthen the evidence that any observed blocking or

unblocking would be related to task conditions, rather than novelty. In the compound phase, three different conditions were used (Fig 1B, middle panel). In Both Reward (BR) trials, both the (respective visual or auditory) CS+ of the Actor group (aCS+) and the Partner group (pCS+) were simultaneously displayed and both rats were rewarded with 3 pellets. In the Own Reward (OR) trials, the respective aCS+ was simultaneously displayed with the aCS- of the Partner group and only the Actor group was rewarded. In the NR trials, the respective aCS- was simultaneously displayed with the aCS- of the Partner group and neither Actor nor Partner were rewarded. A compound conditioning session consisted of 20 trials per condition. The Conditions BR, OR and NR were pseudo-randomized with every condition not being repeated more than 3 times in a row. ITI randomization, stimulus presentation and reward delivery were implemented as in the discrimination phase.

Probe trials. During probe trials, all rats were tested in isolation for one extinction session in their assigned box compartment. All stimuli were now presented in isolation, both the aCS+ and aCS- learned in the Pavlovian discrimination task as well as the two novel stimuli pCS+ (Both reward CS+) and pCS- (Own Reward CS-) added in the compound phase, for which learning was hypothesised to become unblocked and blocked, respectively. Rats in both groups went through 10 trials for each of these 4 stimuli, presented in isolation and without reward delivery (Fig. 1B, right panel). The 4 stimuli were pseudo-randomized with every condition not being repeated more than 3 times in a row.

Control experiments: Inserted Wall (1A). In the Inserted Wall control experiment, 8 Actor rats and 8 Partner rats went through the same three experimental conditions. The only difference here is that during the compound phase, the wall that separated the Skinner box compartments was rendered opaque by adding an additional black wall, to block contact between the Actor and Partner rats. We hypothesised that if visual, and/or auditory and/or olfactory contact between the rats facilitated the social information transmission that helps to unblock reinforcement learning of compound cues, then obstructing these transmission cues should impair unblocking. In the Inserted Wall control group, we chose to also implement the 1-pellet dispenser condition (see Figure 1 - figure supplement 2), to match our results to the condition where secondary reinforcement might still play a role. If differences between the Inserted Wall control and the Social-Appetitive experimental conditions would still emerge, this would strengthen the interpretation that social unblocking was driven primarily by vicarious reward, and not by secondary reinforcement learning, as the putative reinforcing effect of an additional pellet drop during the compound phase was present in both the Social-Appetitive experimental and Inserted Wall control conditions.

In the No Partner Present control experiment, 8 actors went through the same experimental conditions. Here, the only difference was that during the compound phase the partner rat was not present. Instead, pellet dispensers dropped pellets in a custom-made 3d printed plastic food cup including the metal parts which were used in the original food cup for catching the pellet. This made sure that the sound of pellet delivery was similar as in the experimental group. Pellets furthermore fell through the custom-made food cup in a small cup underneath as to avoid the pellets to stack up in view of the actor rat. Finally, here it is important to note that secondary reinforcement of the additional pellet dispenser activity itself was controlled for by delivering pellets outside of the box during discrimination learning, as in the Social-Only experimental subgroup (see Figure 1 - figure supplement 2). This would ensure that only the pellet delivery related sound of falling in the food cup (of the empty partner side) and not sounds made by the pellet dispenser itself would influence associative learning. This control condition was used to assess if visual and auditory observation of pellet delivery in the food cup could unblock learning by itself. Finally, in control experiment 2: Unequal Outcomes, actor and partner rats went through the same stages of conditioning only now during compound conditioning the BR condition became a Partner reward condition while the OR remained the same. With this symmetric implementation, actor rats' OR is partner rats' PR and vice versa and both groups of rats can be treated as actors, doubling the sampling size for one experiment. This control experiment was used to assess if disadvantageous unequal reward outcomes to partner rats would unblock learning.

Statistical data analyses. Entries into the food trough were recorded as photobeam breaks. Raw data were processed in EthoVision XT 11.5 (Noldus Information Technology) to extract our dependent variables: time spent in the food trough and number of entries in the trough (food cup rate). Food cup directed behavior in the form of time spent in the food trough and latency to entry were analysed per trial and per condition for all stages of learning; further analysis and graph preparation was performed using custom-made scripts in MATLAB (version 2014b, MathWorks). All statistics was performed using SPSS (IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp). To assess the strength of learning during discrimination and compound conditioning, only the first 10s of the cue period was analysed to avoid the influence of reward delivery/omission feedback (M. A. McDannald et al., 2011) and the time spent in the food cup was used as measure for conditioned responding . In the probe trials however reward was absent, therefore here we analysed both 10s and 30s period. Previously (Burke et al., 2008; McDannald et al., 2011) used the percentage of time spent at the food trough and the food cup rate to assess value unblocking and identity unblocking ,respectively as measures for conditioned responding. As social unblocking is thought to mainly reflect value unblocking, we

report the percentage of time spent as our main outcome parameter. For completion, we also report food cup rate to assess identity unblocking. Discrimination learning performance was quantified by averaging responding to the cues over the last 4 days of training and comparing the mean between aCS+ and aCS- and difference scores of CS+ - CS- for contrasting cue modalities using paired sample t-tests. Performance in the compound phase was quantified using a 2 factor repeated measures ANOVA on the mean response rate per day across conditions (BR, OR and NR) and post hoc tests were performed to assess the significance of any differences between conditions, corrected for multiple comparisons. Performance in the probe trials was assessed by averaging responding of the actor rats time spent in the food cup and food cup rate over 5 bins (2 trials per bin) and running a two factor repeated measures ANOVA over these bins and the 4 stimuli types (aCS+, pCS+ (unblocked), pCS- (blocked) and aCS-) separately for experimental and control experiments. Differences between conditions and bins were assessed with post-hoc tests, again corrected for multiple comparisons. A putative difference between latencies to entry was analysed in a two-step process, as latencies were not normally distributed. Using a bootstrap procedure (N=5000 iterations), per experimental condition we sampled N probe-trial latencies to entry (with N resampled with replacement, equal to the number of trials with valid entries excluding non-entries and latencies <0.040s) for the unblocked cue and the blocked cue throughout all probe trials, and stored (per iteration) the difference in mean latency for these samples. This generated an N=5000 bootstrap population of mean latency differences per experiment, with all of these distributions following a normal-like distribution (see Fig. 3 – figure supplement 1). Using a Z-test, we assessed 1) whether each distribution was significantly different from 0 (suggesting a significant difference in latency to enter between trial types) and 2) whether this latency difference was significantly different between experimental conditions. For the direct comparison between Experiment and control group in the probe trials we performed a mixed repeated measures ANOVA with factors trials (trial 1 to 6) and stimuli (aCS+, pCS+ (unblocked), pCS- (blocked) and aCS-) and experiment group as a between-subjects factor (Experiment, Control). Differences between conditions between experiments were assessed with post-hoc tests, corrected for multiple comparisons. Finally, to further in depth look at the difference between experiment and control group difference scores were calculated for every available contrast (aCS+/aCS-, pCS+/pCS-, pCS+/aCS- and pCS-/aCS) and for these contrasts a two factor repeated measures ANOVA was calculated. For all RM-ANOVA's, Mauchly's test of sphericity was performed and, when significant, the Greenhouse-Geisser correction was applied.

2.3 Results.

All groups of actor and partner rats were initially trained separately on a Pavlovian discrimination problem. Subsequently, the rats went through a social learning phase were actor rats could learn to associate additional compounded cues with different reward outcomes delivered to the partner rat (social unblocking). Finally, we tested the associative strength of all cues, each presented in isolation, in a probe phase without a reward. In the Inserted Wall control experiment, we impeded the exchange of visual information by implementing an opaque wall and in the No Partner Present control experiment, we implemented the social learning phase without a partner rat present. Finally, in the Unequal Outcomes control experiment, we implemented the social learning phase with unequal, disadvantageous reward outcomes (see Method section for further details). We illustrate the actor rats' conditioned responses with the time spent in the food cup, the food cup rate and their latency to entry as dependent variables. We subdivide the result section in two parts. We demonstrate that cues that predicts no additional reward for the partner are blocked in both the experimental group and all control groups. We then show that, generally, we find vicarious unblocking for food cup occupancy in experimental group 1 (combined Social-Appetitive and Social-Only subgroups) but not in control group 1 (combined Inserted Wall and No Partner Present subgroups) and control group 2 (Unequal Outcomes). We furthermore examine the pattern of unblocking over time and in addition, we investigate potential identity unblocking by looking at the food cup rate. The second part of the results section present several control experiments that show that the vicarious unblocking response is still present when controlling for secondary reinforcement of pellet dispenser sounds but not when a wall was placed between partners to prevent social information exchange (control group 1, Inserted Wall subgroup). Likewise, unblocking was diminished when there was no partner present during the social learning phase (control group 1, No Partner Present subgroup).

Discrimination learning. Actor rats (N=20) were trained on a visual or auditory discrimination task with counterbalanced exemplars as aCS+ and aCS- stimuli (see Figure 1 – figure supplement 1). All actor rats developed a conditioned response to their own aCS+, resulting in an increase with learning in time spent in the food trough on aCS+ trials in anticipation of reward, independent of cue modality (see below). Concurrently, they learned to expect no reward during aCS- presentations, as witnessed by a steady decrease in time spent in the food trough on aCS- trials (Fig. 2a, c, e and an example trial in Video 1). A paired samples *t*-test examining the mean responding over the last 4 days of conditioning was performed. We found a significant difference in time spent in the food trough between the aCS+ and aCS- of the experimental group (M = 58.76, SD=12.86; M=21.19, SD= 13,21; $t(19) = 12.116$, $p < 0.001$), control group 1 (M = 54.649,

SD=14,604; M=15.61, SD= 7.86; $t(19) = 13.472$, $p < 0.001$) and control group 2 (M = 53.82, SD=18.06; M=17.66, SD= 9.02; $t(19) = 7.57$, $p < 0.001$). We performed a two-way ANOVA to assess whether discrimination ability was similar in the experimental conditions and for the different stimulus types (auditory or visual) using the difference scores (aCS+/aCS-) on the last 4 days of training. There was no significant difference between groups ($F(2, 46) = 0.141$, $P = 0.869$), no difference between auditory and visual discrimination learning ($F_{(1, 46)} = 0.076$, $P = 0.785$) and finally no interaction between experiment and stimulus type ($F(2, 46) = 0.297$, $P = 0.745$).

Social learning. In this phase, rats were trained together. The aCS+/aCS- of the actor and pCS+/pCS- of the partner were combined in three compound combinations with the following reward outcomes: Both Reward (BR: aCS+/pCS+), Own Reward (OR: aCS+/pCS-) and No Reward (NR: aCS-/pCS-; Fig. 1B). In the main experiments, we chose to omit the "Partner Reward" condition where the target rats would *not* receive reward, while the partner rats would (PR: aCS-/pCS+; but see Unequal Outcomes control), to avoid a potential reward/value conflict due to disadvantageous inequity aversion (Fehr and Schmidt, 1999), which has been reported in rats as well (Oberliessen et al., 2016). Rats' conditioned responses to these compound cues are shown (Fig. 2b, d, f and an example trial in video 2 and 3) and a direct comparison of these responses to the original aCS+ and aCS- cues was made, both indexed by time spent in the food cup and the food cup rate. In the subsequent analysis, only the behavior of the Actor rats is reported. We applied a mixed repeated measures ANOVA design with the three compound trial types (BR (Partner reward (PR) for control group 2), OR and NR) and day 1 to 4 as within subject factors, and with group (experimental vs control 1 vs control 2) as between subject factor.Time spent in the food cup during the first 10 seconds after the cue onset was chosen as the dependent variable. We found a significant main effect of trial type ($F_{(1.568, 76.845)} = 161.520$, $p < 0.001$, $\eta_p^2 = 0.767$) and importantly found an interaction effect of Experiment * trial type ($F_{(3.137, 76.845)} = 28.243$, $p < 0.001$, $\eta_p^2 = 0.537$), reflecting the difference in experiments that employed BR vs PR trials; and no effect of day ($F_{(2.223, 108.195)} = 0.017$, $p< 0.001$, $\eta_p^2 = 0.997$). Post-hoc comparison revealed that actors' responding to the BR cue did not differ significantly from the OR cue in experimental group 1 (Mean Difference = 0.490, Std Error = 2.367, p = 1.00) and control group 1 (Mean Difference = 1.603, Std Error = 2.646, p = 1.00), while in control group 2 responding is smaller in PR (Partner Reward) than OR trials (Mean Difference = -38.784, Std Error = 2.646, $p < 0.001$). BR responding was furthermore significantly higher than NR in Experimental group 1 (Mean Difference = 32.984.784, Std Error = 2.848, $p < 0.001$) and control group 1 (Mean Difference = 32.684, Std Error = 3.185, $p < 0.001$), while in control group 2, PR responding was not significantly different from NR (Mean Difference = -1.833, Std Error = 3.185, p = 1.00), arguing

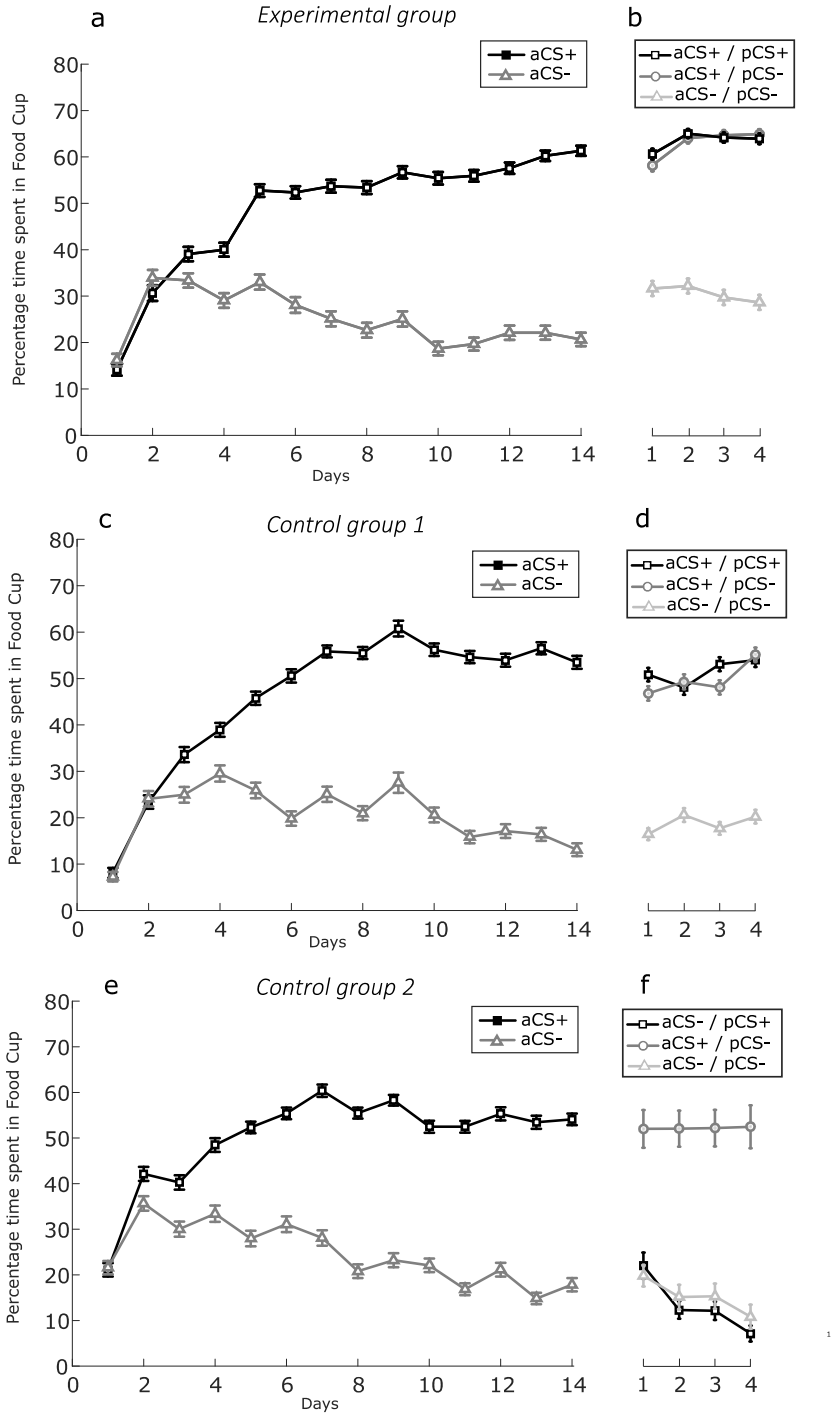


Figure 2. ***Conditioning per experimental phase.*** *Experimental group (combined Social-Appetitive and Social-Only subgroups).* (a) Percentage of time spent in food cup for discrimination learning between aCS+ and aCS- over days. (b) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days. *Control group 1 (combined Inserted Wall and No Partner Present subgroups).* (c) Percentage of time spent in food cup for discrimination learning between aCS+ and aCS- over days. (d) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days. *Control group 2 (Unequal Outcomes).* (e) Percentage of time spent in food cup for discrimination learning between aCS+ and aCS- over days. (f) Percentage of time spent in food cup for the compounds PR (aCS-, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days. Error bars indicate SEM.

against social facilitation of conditioned responding as a social learning mechanism. Finally, OR responding was significantly higher than NR responding in all groups (experimental group 1: Mean Difference = 32.495, Std Error = 3.704, $p < 0.001$; control group 1 Mean Difference = 31.081, Std Error = 4.141, $p < 0.001$; control group 2 Mean Difference = 36.901, Std Error = 4.141, $p < 0.001$). We furthermore assessed if the average compound phase food cup responses over 4 days changed in comparison to the last 4 days of discrimination learning, due to addition of the pCS+ and pCS- cues. Next to that, we assessed if there were any between-group and within-condition differences in the compound phase food cup responses to BR, OR and NR cues. We first ran a mixed repeated measures ANOVA analysis, with three difference scores (aCS+/aCS-, BR/NR and OR/NR) as within subject factors and Group (experimental vs control 1 vs control 2) and Stimulus Type (Auditory/ visual) as between subject factors (Figure 2 – figure supplement 1a, b and c). We found a significant main effect of trial type ($F_{(1.435, 66.017)} = 35.071$, $p < 0.001$, $\eta_p^2 = 0.433$), we furthermore found an interaction effect of trial type * Group ($F_{(2.870, 66.017)} = 22.188$, $p < 0.001$, $\eta_p^2 = 0.491$), an interaction effect of trial type * Stimulus type ($F_{(1.435, 66.017)} = 4.286$, $p = 0.029$, $\eta_p^2 = 0.085$) but no effect of Experiment * trial type * stimulus type ($F_{(2.870, 66.017)} = 1.577$, $p = 0.187$, $\eta_p^2 = 0.064$). Post-hoc comparison for the Trial type * Stimulus type interaction found that rats, over all experiments, have a significantly smaller contrast score for visual than auditory cues in the OR/NR contrast (Mean Difference = -10.862, Std Error = 4.141, $p < 0.001$), near significantly smaller in the BR/NR (Mean Difference = -6.409, Std Error = 3,432, $p = 0.068$) but not smaller for the aCS+/aCS- (Mean Difference = +1.191, Std Error = 4.331, $p = 0.785$). There are furthermore no differences within-experiments between aCS+/aCS- contrast scores and BR/NR or OR/NR contrast scores in all experimental groups, arguing against a putative effect on associative learning of the compound cues due to more vigorous responding in the compound phase. The only expected differences observed here is that the PR/NR contrast in control experiment 2 is smaller than the OR/NR (Mean Difference = 38.784, Std Error = 2.532, $p < 0.001$) and aCS+ /aCS- contrast (Mean Difference = 38.044, Std Error = 3.676, $p < 0.001$), because of the altered reward contingencies. Finally, most importantly, we do not find any differences between contrasts between different experimental groups with the only exception again for the PR/NR contrast which is lower for the control group 2 (PR) compared Experimental group 1 (BR) and control group 1 (BR). We find that the observed effect of stimulus type is mostly captured by a slight shift in conditioned responding to the NR compound in comparison to the aCS- responses during discrimination learning. When running a mixed repeated measures ANOVA design, with the aCS- and NR as within subject factors and group (experimental vs control 1 vs control 2) and stimulus type (Auditory/ visual) we find a triple interaction effect (Figure 2 – figure

supplement 1d, e and f; $F(2,46) = 5.247$, $p = 0.009$, $\eta_p^2 = 0.186$). It becomes clear that the rats show significantly more food cup responses to the visual cues in comparisons to the auditory cues in the NR (Mean Difference = 18.966, Std Error = 3.061, $p < 0.001$) compared to the aCS- in experimental group 1 but not control group 1 and 2. Conditioned responses to visual cues are furthermore higher in NR over aCS- in both experimental group 1 (Mean Difference = 25.247, Std Error = 4.412, $p < 0.001$) and near significantly higher in control group 1 (Mean Difference = 9.00, Std Error = 4.697, $p = 0.062$) but not control group 2.

These results indicate that adding an additional cue predicting a BR or OR outcome does not change the conditioned response in comparison to NR during discrimination learning in the experimental and control groups. Importantly, no differences were observed between rewarded conditions indicating that partner presence does not influence food cup responses by itself. The only difference we notice is that adding a visual cue to an auditory cue leads to increased food cup response in the NR condition experimental group 1 and control group 1 but not in control group 2 compared to the aCS-. This could indicate a deficit in inhibitory action control of a learned auditory CS- because of partner presence, or reflect some difference in stimulus efficacy or asymmetrical processing interacting with social partner presence that cannot be entirely interpreted.

Probing vicarious associative learning. In the probe trials, we aimed to show the effect of associative learning driven by self and vicarious reward. In an extinction setting, rats were individually exposed to the cues in isolation (i.e. one at a time), omitting reward. The learned associative value of each cue was indexed by the time spent in the food cup, the food cup rate and the latency to entry over 10 extinction trials per cue (the presentation order of cues was intermixed). We show the percentage conditioned responding of the actor rats to the first 10 seconds of 10 presentations each of the aCS+ and aCS-, the pCS+ (unblocked) cue (Figure 3 a, b, c and an example trial in Video 4 and 5) associated with an added reward to the partner (BR) and the pCS- (blocked) cue associated with no additional reward to the partner (OR) and no reward to self (NR). We binned responses in groups of 2 trials. Summary statistics for the Figure 3 comparisons (F-stats, p-values, effect sizes), including the time spent in the food cup for the 30 seconds after cue onset (Figure 3d, e and f), can be found in Figure 3 –Source Data 1.

Vicarious reward unblocks associative learning. A two-factor repeated measures ANOVA with stimulus type and bin as factors and the time spent in the food trough as the dependent variable was performed for the experimental group (combined Social-Appetitive and Social-Only subgroups). To sum up, we found a significant probe phase main effect of cue type on time spent

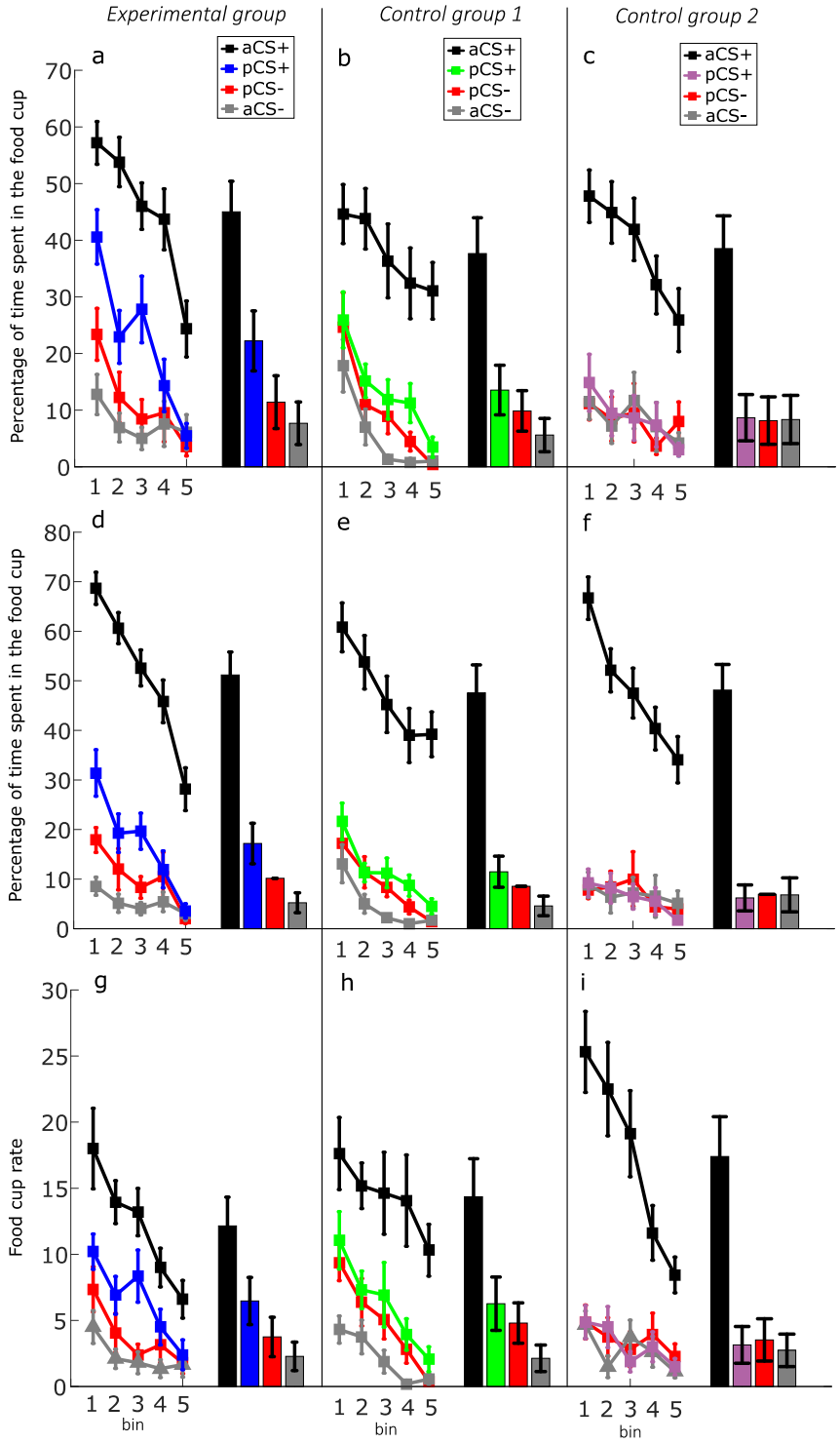


Figure 3. ***Food cup response during the probe trials.*** *Experimental group 1 (combined Social-Appetitive and Social-Only subgroups)*: (a) Percentage of time spent in the food cup in the 10 second period after cue onset during extinction (d) Percentage of time spent in the food cup in the 30 second period after cue onset (g) Food cup rate per minute in the 10 second period after cue onset. *Control group 1 (combined Inserted Wall and No Partner Present subgroups)*: (b) Percentage of time spent in the food cup in the 10 second period during extinction after cue onset (e) Percentage of time spent in the food cup in the 30 second period after cue onset (h) Food cup rate per minute in the 10 second period after cue onset. *Control group 2 (Unequal Outcomes)*: (c) Percentage of time spent in the food cup in the 10 second period after cue onset during extinction (f) Percentage of time spent in the food cup in the 30 second period after cue onset (i) Food cup rate per minute in the 10 second period after cue onset. Extinction in all groups is depicted in 5 bins of 2 trials per bin. Bar plots indicate averaged time spent in the food cup over 10 probe trials between aCS+, pCS+(unblocked), pCS- (blocked) and aCS-. Error bars indicate SEM.

in food trough ($F_{(3, 57)} = 83.180$, $p < 0.01$, $\eta_p^2 = 0.814$). As expected, time spent in the food trough was higher for aCS+ than aCS- trials (Mean Difference = 37.322, Std Error = 2.958, $p < 0.001$; Cohen's d = 1.92, Fig. 3a). Critically, pairwise comparisons revealed that the actors spent more time in the food trough for pCS+ (unblocked) cues than for pCS- (blocked) cues (Mean Difference = 10.804, Std Error = 2.592, $p = 0.003$, Cohen's d = 0.60; Fig. 3a). Furthermore, we also found a significantly higher responding to the pCS+ cue compared to the aCS- cue (Mean Difference = 14.544, Std Error = 2.257, $p < 0.001$, Cohen's d = 0.88) while responding to the pCS- cue was not significantly different from the aCS- cue (Mean Difference = 3.740, Std Error = 1.190, $p = 0.390$, Cohen's d = 0.25), suggesting that the blocked cue is treated similarly to the aCS-, in line with learning theory. Additionally, we found a main effect of bin number on time spent in food trough ($F_{(4, 76)} = 18.678$, $p < 0.01$, $\eta_p^2 = 0.496$), reflecting the extinction process, and finally, we found an interaction between cue type and bin number on the time spent in food trough ($F_{(12,228)} = 2.930$, $p = 0.001$, $\eta_p^2 = 0.137$). Simple effects analysis revealed that the food cup response for the pCS+ was significantly higher than pCS- for bin 1 (Mean Difference = 17.200, Std Error = 5.461, $p < 0.032$), bin 3 (Mean Difference = 19.360, Std Error = 0.007, $p = 0.007$) but that this difference disappeared from bin 4 (Mean Difference = 4.830, Std Error = 7.076, $p = 1.00$).

Blocking the Vicarious experience of reward impairs associative learning. Here, a similar two-factor repeated measures ANOVA with stimulus type and bin as factors and the time spent in the food trough as the dependent variable was performed for the control group 1 (Combined Inserted Wall and No Partner Present subgroups). We also found a significant main effect of cue type on time spent in food trough ($F_{(1.372, 20.585)} = 31.215$, $p < 0.001$, $\eta_p^2 = 0.675$). Here, we find again that the time spent in food trough was higher for aCS+ than aCS- trials (Mean Difference = 32.043, Std Error = 2.958, $p < 0.001$; Cohen's d = 1.91, Fig. 3b). Crucially, pairwise comparisons revealed that the time actors spent in the food trough did not differ for pCS+ cues compared to pCS- cues (Mean Difference = 3.687, Std Error = 2.146, $p = 0.637$, Cohen's d = 0.30; Fig. 3b). We did find a significantly higher responding to the pCS+ cue compared to the aCS- cue (Mean Difference = 7.948, Std Error = 1.862, $p = 0.004$, Cohen's d = 0.75), potentially reflecting some non-social appetitive value related to second-order conditioning, while responding to the pCS- cue was not significantly different from the aCS- cue (Mean Difference = 4.260, Std Error = 1.634, $p = 0.119$, Cohen's d = 0.44; Fig. 3b). Additionally, we again found a main effect of bin number on time spent in food trough ($F_{(1.789, 26.834)} = 13.270$, $p < 0.001$, $\eta_p^2 = 0.469$), but no interaction between condition type and bin number on the time spent in food trough ($F_{(6.029, 20.585)} = 0.835$, $p = 0.547$, $\eta_p^2 = 0.137$).

Unequal outcomes prevent associative learning. Here, a similar two-factor repeated measures ANOVA with stimulus type and bin as factors and the time spent in the food trough as the dependent variable was performed for the Unequal Outcomes control group. The two-factor repeated measures ANOVA revealed a significant main effect of probe trial type on time spent in food trough ($F_{(1.732, 25.980)} = 46.215$, $p < 0.001$, $\eta_p^2 = 0.755$). Again, responding was higher for aCS+ than aCS- trials (Mean Difference = 30.162, Std Error = 3.754, $p < 0.001$; Cohen's d = 1.76, Fig. 3c). Crucially, responding to pCS+ cues did not differ from pCS- cues across the 5 bins of extinction (Mean Difference = 0.513, Std Error = 2.294, p = 1.00, Cohen's d = 0.04; Fig. 3c). No differences were found for pCS+ cue compared to the aCS- cue (Mean Difference = 0.318, Std Error = 1.622, p = 1.00, Cohen's d =0.02) or the pCS- cue versus the aCS- cue (Mean Difference = -0.195, Std Error = 2.100, p = 1.00, Cohen's d = -0.01; Fig. 3c). As expected, we found a main extinction effect of bin number on time spent in food trough ($F_{(4, 60)} = 8.291$, $p < 0.001$, $\eta_p^2 = 0.356$), but found no interaction between condition type and bin number on the time spent in food trough ($F_{(4.528, 67.915)} = 1.671$, $p = 0.160$, $\eta_p^2 = 0.100$).

Socially unblocked cues associated with faster food cup entry than control group cues. Latency scores during the probe trials for the experimental group (combined Social-Appetitive and Social-Only subgroups), Control group 1 (Combined Inserted Wall and No Partner Present subgroups) and the Control group 2 (Unequal Outcomes) are shown in Figure 3 – figure supplement 1 (a, b and c). We ran a bootstrapped analysis of latency differences with N=5000 iterations per experiment, drawing with replacement from the pCS+ and pCS- trials (according to their N) per iteration and storing the difference in mean latency between these trial type. From these distributions (Figure 3 – figure supplement 1d) of mean latencies, we assessed whether these distributions differed from zero and whether they differed between experimental groups with Z-tests. We found that the latency difference scores differed from zero in the experiment group (Z = -2.79, p = 0.003) but not control group 1 (Z = -0.90, p = 0.18) and control group 2 (Z = 2.08, p = 0.98). We furthermore found that the latency difference is bigger in the experiment group compared to control group 1 (Z = -1.712, p = 0.043, one-sided) and control group 2 (Z = -4.72, $p < 0.001$). Finally, control group 1 has larger latency differences than control group 2 (Z = -3.16, $p < 0.001$). We conclude from these results that the rats in the experimental groups showed shorter latencies on pCS+ than on pCS- trials, but this was not the case for rats in control group 1 and control group 2. Importantly, this difference is significantly larger in the experimental group than in both control groups, further supporting the interpretation that pCS+ cues acquire associative value in the experimental group, supporting social unblocking.

Taken together, these results show that the actor rats exhibited more food cup directed behavior for the pCS+ cue than both the aCS- and pCS- cue over 10 trials of extinction in the experimental condition only. This means that when actor rats have fully learned a stimulus-reward association producing reward for themselves, adding a cue that predicted an additional reward delivery to a partner rat unblocked associative *social* learning (pCS+ > pCS-) about this cue, putatively due to a vicarious reward experience. In contrast, rats did not spend more time in the food cup for the pCS- cue compared to the aCS-, suggesting that additional cues that did *not* predict vicarious reward remained blocked from acquiring associative value. Contrary to the findings for the experimental group, the rats in control group 1 and 2 did not show such conditioned responding, indicative of acquired value for the unblocked pCS+ cue, over 10 trials of extinction. This suggests that acquiring associative *social* value in this unblocking experiment requires social information exchange (control group: Inserted Wall) and/ or the presence of a partner (control group: No Partner Present). Interestingly, disadvantageous unequal reward distributions putatively modulated the vicarious reward experience, impeding the unblocking effect. Our results reflect that cues related to mere reward delivery have to be controlled for, as witnessed by the pCS+ over aCS- difference even in the control experiments, highlighting the need for an active blocking control cue (pCS-) as implemented here.

Strength of the social unblocking effect over trials. We conclude from the simple effects analyses on the interaction effects in the experimental group that the associative value of unblocked novel cues can be measured for approximately 6 trials in extinction and will use this analysis window going forward. First, we extended the previous analyses with a mixed repeated measures ANOVA design with trial type (aCS+, pCS+ (unblocked), pCS- (blocked), aCS-) and trial 1- 6 (bin 1-3) as within subject factors and group (experimental group 1 vs control group 1 (a and b) vs control group 2) as between subject factor. Performing this analysis for both the 10 seconds and 30 seconds period, we found that rats exhibited more food cup directed behavior for the pCS+ cue than both the pCS- cue over 6 trials of extinction in the experimental group 1 and not control group 1 and 2 (see Figure 3 – Source Data 1).

To directly contrast the unblocking effect between the experimental and control conditions, we calculated difference scores (Figure 4) for the direct comparison of the aCS+/aCS-, pCS+/pCS-, pCS+/aCS- and pCS-/aCS- trial types, and tested for difference in these contrasts between experimental groups. Summary statistics for the Figure 4 comparisons (F-stats, p-values, effect sizes) can be found in Figure 4 – Source data 1.We would expect no difference in the initial discrimination learning contrast, aCS+/aCS- tested in extinction, between groups. Indeed, in a two way repeated measures ANOVA, we found no significant main (within subject) effect of trial

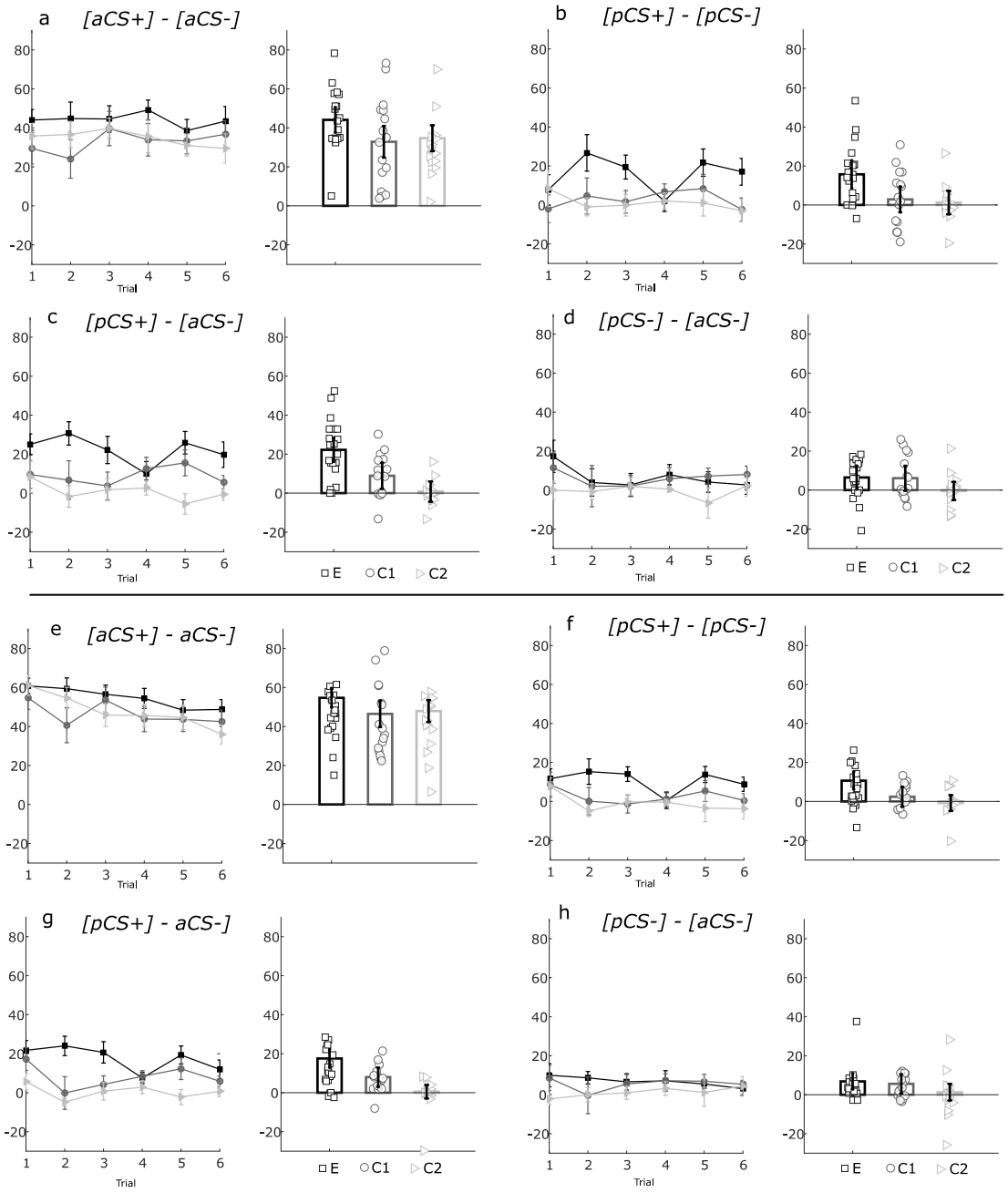


Figure 4. *Difference scores of the actor rat during extinction for 10s and 30s data for experimental group (combined Social-Appetitive and Social-Only subgroups: E, black squares), Control group 1 (combined Inserted Wall and No Partner Present subgroups: C1, grey circles) and Control group 2 (Unequal Outcomes: C2, grey triangles).* (a,e) Difference scores of the percentage of time spent in food cup in the 10 second period (a) and 30 seconds (e) after cue onset over 6 trials for the [aCS+]- [aCS-] difference scores. Bar plots show the average over trials of the [aCS+]- [aCS-] difference scores with dots showing the mean per rat. (b,f) Difference of the percentage of time spent in food cup in the 10 second (b) and 30 seconds (f) period after cue onset over 6 trials for the [pCS+]- [pCS-] difference scores. Bar plots show the average over 6 trials of the [pCS+]- [pCS-] difference scores with dots showing the mean per rat. (c,g) Difference of the percentage of time spent in the food cup in the 10 second (c) and 30 (g) second period after cue onset over 6 trials for the [pCS+]- [aCS-] difference scores. Bar plots show the average over 6 trials of the [pCs+]- [aCS-] difference scores with dots showing the mean per rat. (d,h) Difference of the percentage of time spent in the food cup in the 10 second (d) and 30 second period after cue onset over 6 trials for the [pCS-]- [aCS-] difference scores. Bar plots show average over 6 trials of the [pCs-]- [aCS-] difference scores with dots showing the mean per rat. Error bars indicate SEM.

number on this contrast ($F_{(4.016, 8.032)}$ = 5.96, p = 0.666) and no interaction effect ($F_{(10,130)}$ = 0.460, p= 0.914). Importantly, we did not find evidence for a between-subjects effect of group ($F_{(2, 49)}$ = 1.859, p =0.167, η_p^2 = 0.071; Fig. 4a). The aCS+/aCS- contrast score of the experimental group was not higher than control group 1 (mean difference = 11.242, std error = 6.363, p = 0.251) and control group 2 (mean difference = 9.400, std error = 6.363, p = 0.438) . When directly comparing the unblocked/blocked (pCS+/pCS-) contrast between groups, we *did* expect to find a between subjects group effect. Indeed, when we examined the difference scores of the pCS+/pCS- contrast with a two way repeated measures ANOVA with group (Experimental, control) and trial (1-6) as factors, we found a significant main (between subject) effect of group ($F_{(2, 49)}$ = 6.397, p = 0.003, η_p^2 = 0.207; Fig. 4b). This analysis revealed that the percent difference in responding between pCS+ and pCS- cues was higher for the experimental group than the control group 1 (mean difference = 12.903, std error = 4.531, p = 0.019) and control group 2 (mean difference = 14.520, std error = 4.531, p = 0.007). We found no within-subject effect of trial number $F_{(4.111, 8.222)}$ = 0.627, p = 0.680 η_p^2 = 0.013), suggesting that the difference is relatively stable across trials, and no interaction ($F_{(5,130)}$ = 1.090, p= 0.370, η_p^2 = 0.043). In the contrast analysis for the 30-seconds period, we also found a significant main (between subject) effect of group for the pCS+/pCS- contrast ($F_{(2, 49)}$ = 5.976, p = 0.005, η_p^2 = 0.196; Fig. 4f), revealing that the percent difference in responding between pCS+ and pCS- cues was significantly higher (one-sided) for the experimental group than control group 1 (mean difference = 8.278, std error = 3.478, p = 0.064) and significantly higher than control group 2 (mean difference = 11.493, std error = 3.478, p = 0.005). As expected, also for the 30s period, the aCS+/aCS- revealed no main (between subject) effect and pCS-/aCS- can be found in Table S1. We conclude from these results that the pCS+, predicting partner reward in a BR compound, became unblocked and acquired associative value in the experimental group, but not in control groups 1 and 2, as witnessed by a significantly larger unblocked vs. blocked contrast in the experimental vs. control of group ($F_{(2, 49)}$ = 1.251, p = 0.295, η_p^2 = 0.049; Fig. 4e). Results from the full contrast analyses for the pCS+/aCS- groups for both the first 10 seconds after cue onset and the whole 30 seconds period. We attribute this differential social unblocking effect to a putative difference in experienced vicarious reward. The control conditions, impeding social information exchange (control group; Inserted Wall) and/or the absence of the partner rat (control group; No Partner Present) presumably attenuated vicarious reward experience. In addition, disadvantageous unequal reward distributions did not lead to unblocking, suggesting that such distributions do not reflect vicarious reward experiences for our rats, in line with previous behavioral evidence of inequity aversion.

Probing vicarious reward identity. Burke et al. (2008) found that changing the sensory identity (flavour) of an outcome associated with an added cue in a compound also unblocked this cue and that this identity unblocking was captured by scoring the food cup rate, e.g. the frequency or number of entries into the food cup irrespective of the total duration of visits. In our paradigm, social unblocking could also be interpreted as a reward identity switch in that the additional partner outcome changes the sensory aspects of the reward by virtue of the partner receiving and eating the rewards. Food cup rate in the probe trials, next to food cup occupancy, could therefore potentially reflect model-based reward identity unblocking and therefore could provide insight in the influence of sensory features of social unblocking. Alternatively, if the additional partner reward is interpreted solely as a change in value, but not identity, we would hypothesize that food cup rate would not be affected.

Vicarious reward unblocks food cup rate in Experimental but not in Control group 1 and 2. To further explore the food cup rate as a measure of the potential identity unblocking effect in the experimental vs. control condition, we applied the same a mixed repeated measures ANOVA design with trial type (aCS+, pCS+ (unblocked), pCS- (blocked), aCS-) and bin 1-3 as within subject factors and group (experimental (N=20) vs control 1 (N= 16 vs control 2 (N=16)) as between subject factor with food cup rate per minute for the first 10 second after the cue onset as dependent variable. We found a significant main effect of trial type ($F_{(1.761, 86.275)} = 78.460$, $p < 0.001$, $\eta_p^2 = 0.616$) and an effect of bin ($F_{(1.953, 5.713)} = 21.9968$, $p< 0.001$, $\eta_p^2 = 0.310$). We also find an interaction effect of Experiment * trial type ($F_{(3.52, 86.275)} = 5.033$, $p= 0.002$, $\eta_p^2 = 0.170$). Post-hoc comparison revealed here as well that the food cup rate was significantly higher for pCS+ cue in comparison to the pCS- cue in the experimental group (mean difference = 3.883, std error = 0.1.207, p = 0.014; Fig. 3g) but not in control group 1 (mean difference = 1.500, std error = 1.350, p = 1.00; Fig. 3h) and not in control group 2 (mean difference = -0.062, std error = 1.350, p = 1.00; fig 3i). Interestingly, we furthermore find that the food cup rate for the pCS+ cue is significantly higher than the aCS- in both the experimental group (mean difference = 5.683, std error = 1.074, p < 0.001; Fig. 3g) and control group 1 (mean difference = 5.125, std error = 1.200, p = 0.001; Fig. 3h) but not in control group 2 (mean difference = 0.438, std error = 1.200, p = 1.00; Fig. 3i). These results could potentially indicate that the actor rats had a clear idea that the identity of the US food rewards in BR trials had changed, even though in the Inserted Wall control group where information exchange is impeded and/or in the No Partner Present control group where the partner is absent. We can conclude that identity unblocking as measured by the number of entries into the food cup is also present in this task and that the sensory aspects of the additional

partner presence are necessary to associate the novel cue with positive social associative value (BR > OR).

Probing associative learning in the partner rat. During the compound phase in the experimental group and the control group Inserted Wall, the partner rat learns to associate another set of outcomes to the compound cues (Figure 5b, e) : Both Reward (BR: pCS+, aCS+), Actor Reward (OR: pCS- aCS+) and No Reward (NR: pCS-, aCS-) after going through Discrimination Learning (Figure 5a, d). For the partner, learning about the aCS+ cue is thus confounded by being paired with two qualitatively different outcomes: from the perspective of the partner, it represented both a mutual reward outcome and an unequal disadvantageous reward outcome. It is thus likely that the aCS+ cue value would be increased due to the BR associated value but at the same time decreased due to the disadvantageous unequal outcome on OR trials. We tested whether the aCS+, associated with these multiple types of partner-own-reward outcomes, still showed evidence of unblocking by performing a two-factor repeated measures ANOVA with stimulus type and trial 1 to 6 as factors and the time spent in the food trough as the dependent variable. For the experimental group 1, we found a significant main effect of probe trial condition on time spent in food trough ($F_{(1.551, 29.472)} = 79.840$, $p < 0.001$, $\eta_p^2 = 0.808$). As expected, time spent in food trough was higher for pCS+ than pCS- trials (Mean Difference = 39.00, Std Error = 3.774, $p < 0.001$, Fig 5c), reflecting the partner's initial discrimination learning. Critically however, pairwise comparisons revealed that the partners time spent in food trough was *not* higher for aCS+ cues compared to aCS- cues across 6 trials (Mean Difference = 3.530, Std Error = 1.691, $p = 0.303$; Fig 5c). Furthermore, to see whether the absence of social information exchange (Inserted Wall control group) would influence partner learning we also performed a two-factor repeated measures ANOVA on the data of Inserted Wall control group. We found a significant main effect of probe trial condition on time spent in food trough ($F_{(3, 21)} = 23.490$, $p < 0.001$, $\eta_p^2 = 0.770$). As expected, time spent in food trough was again higher for pCS+ than pCS- trials (Mean Difference = 43.525, Std Error = 6.56, $p = 0.002$; Fig 5f) here though there was a trend towards higher responding for for aCS+ cues than for aCS- cues (Mean Difference = 9.958, Std Error = 2.855, $p = 0.061$; Fig 5f). We conclude from these results that the compounded cue aCS+ has not become unblocked for the partner rat in the experimental group, however for the control group 1 we observe a clear trend indicative of unblocking. This potential unblocking could be influenced by two factors. First, secondary reinforcement of actor reward delivery (in aCS+ containing compounds) without observation of the actual reward delivery to the actor could have inhibited the attenuating effect of disadvantageous inequity aversion on unblocking and lead to this trend towards unblocking. Next to that, attentional-based unblocking

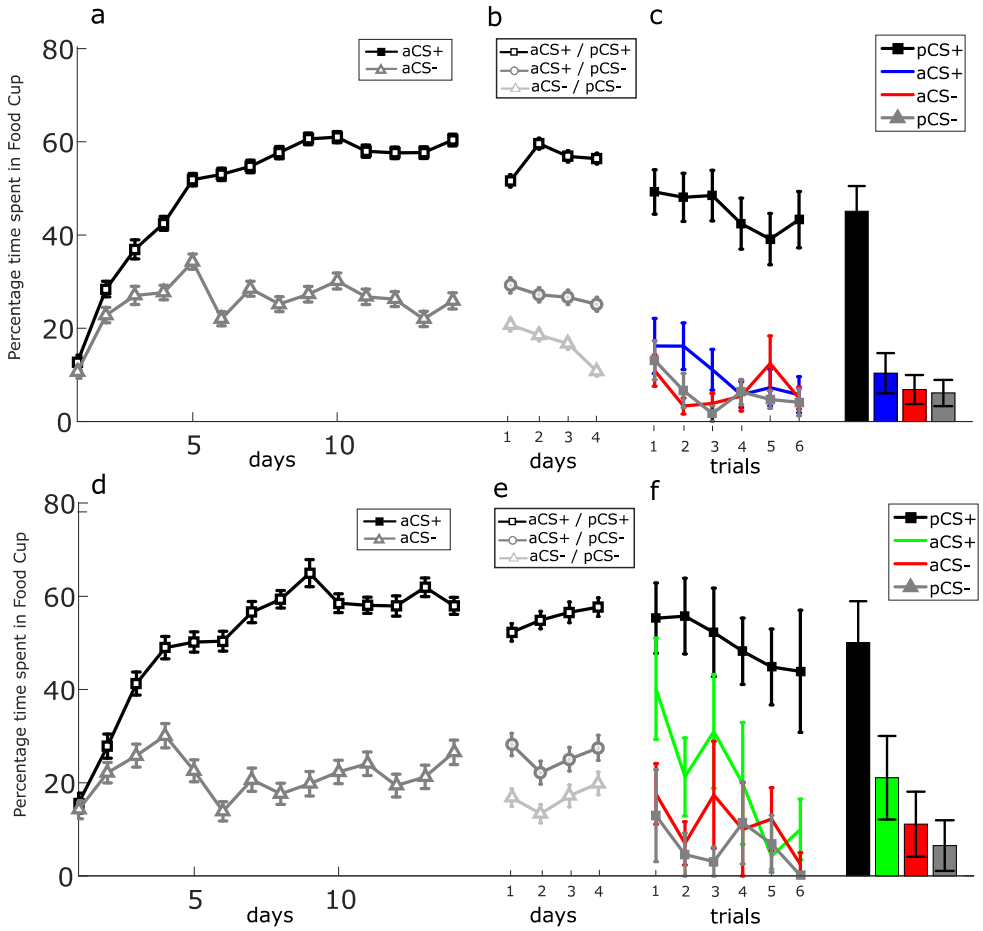


Figure 5. *Food cup responses of the partner rat during learning. Experimental group 1 (combined Social-Appetitive and Social-Only subgroups)*: (a) Percentage of time spent in the food cup during discrimination learning. (b) Percentage of time spent in the food cup during social learning. (c) Percentage of time spent in the food cup during the first 6 trials of the probe phase. *Control group 1a (Inserted Wall)*: (d) Percentage of time spent in the food cup during discrimination learning. (e) Percentage of time spent in the food cup during social learning. (f) Percentage of time spent in the food cup during the first 6 trials of the probe phase. Bar plots indicate averaged time spent in the food cup over 6 probe trials between aCS+, pCS+ (unblocked), pCS- (blocked) and aCS-. Error bars indicate SEM.

(Haselgrove, Tam, & Jones, 2013) could play a role for the partner rat when it learns that one added cue predicts both reward and the omission of reward. This attentional unblocking effect would also be stronger if evidence of actor presence/reward would be blocked. A direct test of unblocking where both rats experienced disadvantageous unequal rewards was implemented as Unequal Outcomes. We did not include a version of the Unequal Outcomes control experiment where we also impeded social information transfer, but would speculate that, in that case, unblocking would remain supressed as well.

The social unblocking effect persists when controlling for secondary reinforcement, but not when the partner is not present. Besides the vicarious experience of reward, other confounding factors could have contributed to learning/unblocking in our paradigm. Most

notably, sources of secondary reinforcement should be excluded as potential drivers of learning. During discrimination learning, the actor rat is conditioned to receive pellets contingent on its aCS+. Afterwards, in the compound phase, the rat is presented with an auditory–visual compound. Instead of one pellet drop (self-reward), now, on some trials, two pellets drop simultaneously (Both Reward trials). It is possible that the additional pellet delivery sound acted as a third CS+ in the compound, in addition to the aCS+ and pCS+. Because the sound of the pellet dispenser is already associated with the aCS+ of the actor rat, it is possible that the appetitive value increased with the intensity of this cue (two pellets dropping instead of one), thus enhancing the total value of cue configuration, leading to unblocking of the pCS+. To control for this possible source of secondary reinforcement, in a subgroup of rats (experimental group; Social-Only), we added a pellet dispenser aimed outside the box (placed at the same location as in the compound phase) already during the discrimination phase, providing the same acoustic features of pellet delivery to the target rat, without presenting additional reward (pellets were collected outside of the box). Next to that, our control group 1 consisted of a subgroup 1A where a wall impeded social contact and a subgroup 1B where there was no partner present. It is clear that these two groups are similar in that the actor rat does not observe food delivery to the partner. However, in the impeded wall condition, US conditioning could still occur due to pellet dispenser sounds (not controlled for; see Figure 1 - figure supplement 1) and partner rat related sounds while in the no partner present condition conditioning might still be caused by the observation of food delivery in the other compartment but not pellet dispenser related sounds (controlled for; see Figure 1 - figure supplement 1).

We found a significant main effect ($F_{(1, 32)} = 17.964$, $p < 0.001$, $\eta_p^2 = 0.360$) of trial type, an interaction of Group * trial type ($F_{(3,32)} = 4.559$, $= 0.009$, $\eta_p^2 = 0.299$; Fig. 6a) and an effect of trial number ($F_{(5,160)} = 7.286$, $p < 0.001$, $\eta_p^2 = 0.186$). Post hoc comparison reveals that responding to the pCS+ cue in extinction differs significantly from the pCS- cue in the Social-Appetitive group (one pellet added; mean difference = 24.808, std error = 4.578, $p < 0.001$), Social-Only group (no new pellets added; mean difference = 9.717, std error = 4.047, $p = 0.022$) but not the Inserted Wall group (mean difference = -0.175, std error = 4.956, $p = 0.972$,) or the No Partner Present group (mean difference = 5.525, std error = 4.578, $p = 0.273$). We then also compared responding during the full 30 seconds, which would equate to the addition of 3 extra pellets. We performed a similar mixed repeated measures ANOVA design with trial type (pCS+, pCS-) and trial (1 - 6) as within subject factors and 4 group (experimental group; Social-Appetitive: N=8; one pellet added vs. experimental group; Social-Only: N=12; no new pellets added vs. Inserted Wall control: N= 8; one pellet added vs No Partner Present control N= 8; one pellet added) factors as between subject factor. We also found a significant main effect ($F_{(1, 32)} = 16.682$, $p < 0.001$, $\eta_p^2 = 0.343$), an

interaction effect of Group * trial type ($F_{(3,32)} = 4.211$, $p = 0.013$, $\eta_p^2 = 0.283$; Fig. 6b) and an effect of trial ($F_{(3.406, 108.98} = 6.084$, $p < 0.001$, $\eta_p^2 = 0.160$). Post hoc comparison reveals that the pCS+ cue differs significantly from the pCS- cue in Social-Appetitive group (one pellet added; mean difference = 16.844, std error = 3.623, $p < 0.001$), in the Social-Only group (no new pellets added; mean difference = 6.622, std error = 2.58, $p = 0.032$) but not the Inserted Wall control (mean difference = -1.239, std error = 3.623, $p = 0.971$) and No Partner Present control (mean difference = 6.106, std error = 3.623, $p = 0.102$).

Finally, to zoom in on the temporal dynamics of pCS+ vs. pCS- responding in these four experiments, we created time-resolved cumulative occupancy plots. We split the 10 seconds before and after cue onset in 5 bins of 2 seconds each and averaged responding during these bins over the 6 trials which we found have shown the effect. For an additional statistical analysis, we then looked at the cumulative responding over these 5 post-cue onset bins. We performed a mixed repeated measures ANOVA design with trial type (pCS+, pCS-) and time bins as within subject factors and the four groups as between subject factor. We find a significant main effect ($F_{(1,32)} = 12.601$, $p = 0.001$, $\eta_p^2 = 0.283$), an interaction effect of Group * trial type ($F_{(3,32)} = 3.780$, $p = 0.020$, $\eta_p^2 = 0.262$;) and an effect of Group * trial type * bin number ($F_{(3.674, 39.674} = 6.084$, $p = 0.025$, $\eta_p^2 = 0.233$; Fig. 6 c,d,e,f). Post hoc comparison revealed that cumulative response during the pCS+ cue differs significantly from the pCS- cue in Social-Appetitive subgroup from bin 2 (mean difference = 0.493, std error = 0.185, $p = 0.012$) and onwards on bin 3 (mean difference = 1.091, std error = 0.278, $p < 0.001$), 4 (mean difference = 1.601, std error = 0.385, $p < 0.001$) and 5 (mean difference = 2.427, std error = 0.495, $p < 0.001$) after cue onset . In the Social-Only subgroup we find significant differences in bin 4 (mean difference = 0.804, std error = 0.314, $p = 0.015$) and 5 (mean difference = 0.972, std error = 0.404, $p = 0.022$). However, no significant temporal bins were found in the Inserted Wall and No Partner Present control groups.

While descriptively, the magnitude of the *social* unblocking effect is larger when not controlling for additional pellet drops (Social-Appetitive subgroup) than when such a control is implemented (Social-Only subgroup), we conclude that the *social* unblocking effect still exist when explicitly controlling for additional pellets falling in the compound phase for the first 10 seconds and 30 seconds period after cue onset but not when social information exchange is impeded and finally also not when no partner is present during the compound phase.

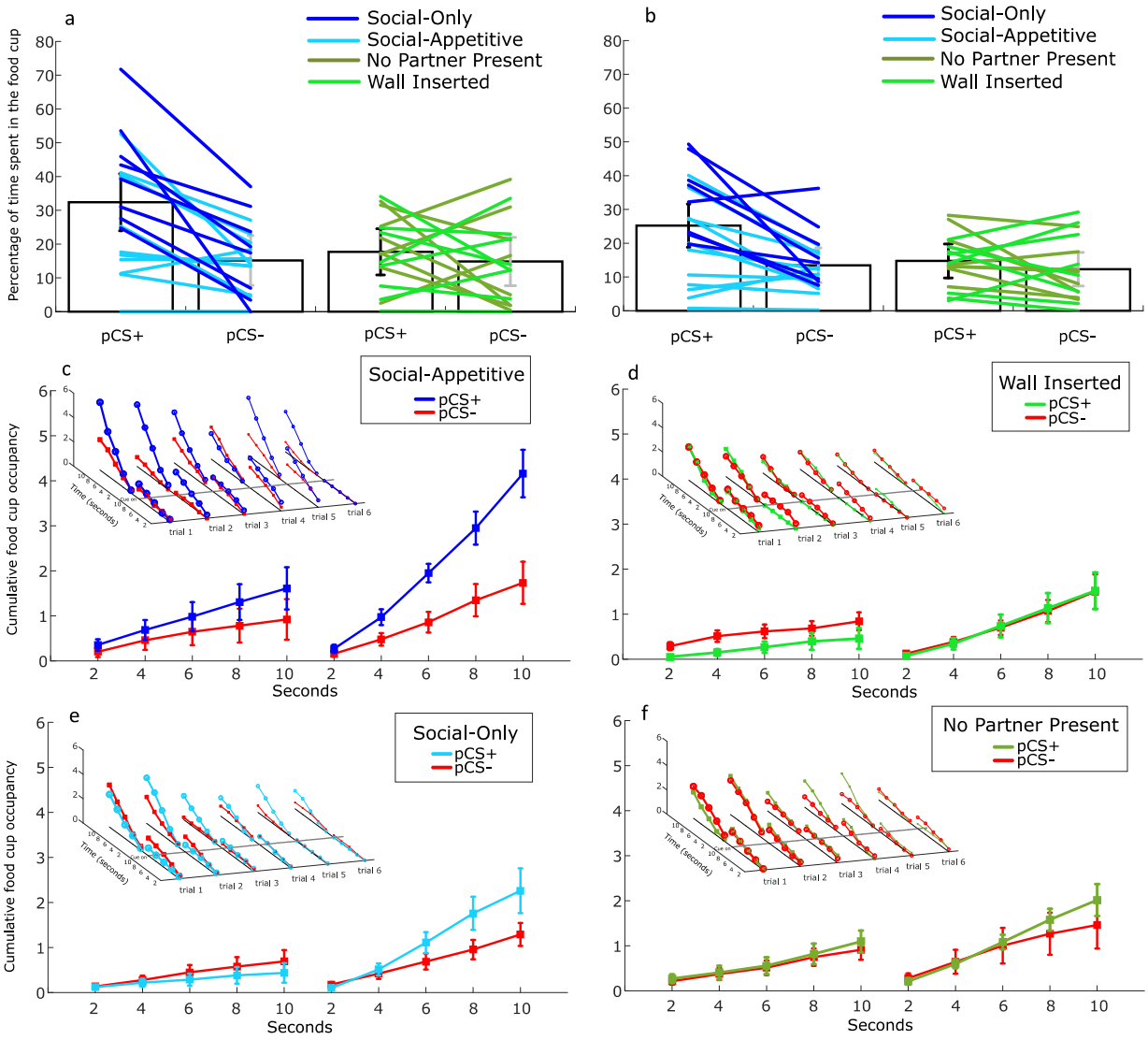


Figure 6. *Effect of secondary reinforcement, impeding wall and partner absence on food cup occupancy in the first 10 seconds after cue onset.* (a, b) Mean percentage of spent in the food cup, for the 10 second period (a) and 30 second period (b) after cue onset for pCS+ vs pCS- for trial 1 to 6 for the experimental group 1 (Social-appetitive and Social-only) and control group 1 (Wall impeded and No partner present) (c) Average cumulative food cup occupancy over 6 trials for the 10 seconds pre and post cue on for the Social-Appetitive group. (d) Average cumulative food cup occupancy over 6 trials for the 10 seconds pre and post cue on for the Wall Inserted group. (e) Average cumulative food cup occupancy over 6 trials for the 10 seconds pre and post cue on for the Social-Only group. (f) Average cumulative food cup occupancy over 6 trials for the 10 seconds pre and post cue on for the No Partner Present group. 3d plot depicts cumulative food cup occupancy per trial per group. Error bars indicate SEM.

2.4 Discussion.

Summary. Social valuation is crucial in forming and maintaining social relationships and, presumably, in experiencing the pleasurable and reinforcing aspects of social interaction. However, it remained unclear whether vicarious reward value, which we define here as value derived from social signals associated with reward delivery to another (Ruff & Fehr, 2014), could drive learning

just as self-experienced value. If this was the case, then vicariously experienced reward should be able to reinforce behavior in a formal Pavlovian learning paradigm. Here we addressed this question by introducing a novel social unblocking task. We find that vicarious reward experience, operationalized in this task as rewards delivered to social partners (cagemates), can indeed drive learning about novel stimuli. After having fully learned that a specific CS+ cue predicted a self-reward, learning about a second cue delivered in compound with this CS+ was blocked, as predicted by learning theory (Fanselow & Wassum, 2016; Rescorla & Wagner, 1972), when no additional self or other reward was contingent on this cue. Blocking was found when comparing the food cup response in extinction between trial types, here specifically for the pCS- (CS- cue predicting no partner reward) compared to the aCS- (CS- cue predicting no actor reward). Learning was unblocked, however, by providing an additional reward delivered to the partner simultaneously with the fully predicted self-reward as witnessed by a higher food cup response of actor on the pCS+ (CS+ cue predicting partner reward) compared directly to the pCS-. The social nature of the positive vicarious reward experience was specifically assessed in three control experiments: 1) Preventing the exchange of social information in the compound learning phase impeded the social unblocking effect (pCS+ ≈ pCS-). 2) Partner rat absence during mutual reward showed unblocking for the novel stimuli (pCS+ > aCS-) but crucially, not the social unblocking effect (pCS+ ≈ pCS-). 3) When the partner was rewarded and the actor not (aCS-, pCS+) and when the actor was rewarded but not the partner (aCS+, pCS-) we found no evidence either unblocking (pCS+ ≈ aCS-) or social unblocking (pCS+ ≈ pCS-). These results suggest that vicarious reward experience can indeed drive learning processes, in line with formal behavioral learning theory and that specific social aspects of the environment such as partner presence and partner visibility are necessary for observing a social unblocking effect (pCS+ > pCS-).

Learning Theory Our results extend previous work by Peter Holland (Holland, 1984) and Geoffrey Schoenbaum (Lopatina et al., 2015; McDannald et al., 2011) on unblocking in appetitive Pavlovian conditioning. These authors found that rats, after learning that distinct cues have specific food outcomes, can show unblocking of learning for cues added in compound, when self-rewards were altered by increasing reward value (e.g. an upward shift from 1 to 3 pellets) or a change in reward identity (same reward type but a shift in reward features such as flavour). In contrast, learning was blocked when no such reward change occurred (e.g. same reward amount or same identity). According to reinforcement learning theory, the upshift or change in identity led to a discrepancy between the expected reward (1 pellet) and the received reward (3 pellets), thus producing a reward prediction error. The theory states that if the added cue reliably predicts the increase in reward/ identity outcome, it will acquire the value inherent in the reward itself (Sutton

& Barto, 1981). The main indicator of learning about the value of a (novel) cue in the unblocking paradigm is an increase in the time spent in the food cup in the probe (extinction) phase (Lopatina et al., 2015; McDannald et al., 2011). Indeed, we observed a higher time spent in the food cup for the cue predicting mutual rewards to the actor rat and its conspecific than to the cue that predicts own-rewards. Taking into account the results of the control experiments, we conclude that the observed enhanced food cup response i.e. the social unblocking effect could be driven by an upshift-related vicarious reward prediction error. The observed social unblocking effect adds to the emerging literature showing that animals attach value to rewards delivered to conspecifics (Hernandez-Lallement et al., 2016a; Kashtelyan et al., 2014) and learn about cues that predict rewards delivered to others. The social reinforcement learning hypothesis (Hernandez-Lallement et al., 2016a) proposes that integration of social signals expressed by partners can aid in making appropriate decisions in a social context. Evidence for this hypothesis comes from the prosocial choice task (PCT) in which it was found that, rewards delivered to oneself and to a partner are preferred over a reward delivery only to the actor himself in both monkeys (Horner et al., 2011) and rats (Hernandez-Lallement et al., 2015; Márquez et al., 2015). In rats, it was found that this effect was modulated by the behavior displayed by the other rat (Márquez et al., 2015) and that this effect was impaired when the partner was replaced by a toy (Hernandez-Lallement et al., 2015) or when the display of the partner's preference was impeded. In Monkeys it was furthermore found that cues that are associated with reward delivery to another monkey were preferred over cues that were associated with juice delivery to a chair with no monkey in it and this preference was absent, in the non-social condition, when there was only a juice bottle present (Chang et al., 2011). We found that social learning occurs when additional reward was delivered to a visible partner but not when preventing the exchange of social information by an opaque wall or when the partner was absent during social learning, providing further evidence that social signals are indeed necessary for learning in social, other regarding, paradigms. A recent study furthermore, shows that macaques increase licking frequency in line with a higher probability of self-reward but decrease their anticipatory licking with increased probability of reward delivery to another monkey (Noritake, Ninomiya, & Isoda, 2018). The authors interpret this decrease of anticipatory licking as an indicator of the negative affect associated with unequal disadvantageous reward pay-outs. Both monkeys and rats have been found to have a distaste for these unequal pay-outs (Brosnan & de Waal, 2003; Oberliessen et al., 2016). Here, we provide similar evidence that, in rats tested in our social unblocking paradigm, disadvantageous unequal reward outcomes do not support unblocking of cues that predict reward to the other rat but not oneself. Our finding extend the results of Rescorla (1999) who found that changing the outcome of an aCS- cue by adding a cue in

compound that predicts self-reward leads to unblocking of that added cue. This contrasts with the lack of unblocking found here, indicating that the observation of reward delivery to the social partner does not have similar reinforcement properties as adding self-reward, possibly due to the negative affect associated with disadvantageous unequal reward outcomes.

Further research is necessary to see whether cues associated with vicarious reward or social reinforcement can also act as a conditioned reinforcer for instrumental responses of rats, as has been found humans (Lehner et al., 2017), in a similar way as has been found for appetitive cues (Burke et al., 2008; Kruse, Overmier, Konz, & Rokke, 1983; R. A. Rescorla, 1994). Finally, it is important to investigate if cues predicting vicarious rewards can guide rats' choices in a social setting. It has been found that rats choose a reward arm in a T-maze that leads to play behavior more than an arm leading to a social encounter where play was absent (Humphreys & Einon, 1981). Furthermore, social play can induce a social place preference (Calcagnetti & Schechter, 1992) and rats are willing to lever press for social play reinforcement (Achterberg et al., 2016). Our task indicates that the unblocked cue has attained the rewarding properties of social reward and it is therefore likely that a presented unblocked cue would be preferred over a blocked cue when tested in a two-alternative forced choice task. We finally expect that our social unblocking effect depends on the successful transmission of social signals between the actor and partner rat (Nicol, 1995; Hernandez-Lallement et al., 2016) and that different signal modalities (auditory, visual, olfactory) might contribute and combine in additive or interactive fashion.

Conclusion. Overall, these data provide evidence that vicarious reward experience can drive associative learning in rats, and that the transmission of social cues between rats is necessary for this learning. Further experiments should be conducted to reveal which mode(s) of social information processing are necessary and sufficient to drive unblocking through social value. Overall, our novel behavioral paradigm could be used to further explore how rats learn about value in social contexts and is well suited to probe the neural circuits involved in social reinforcement learning.

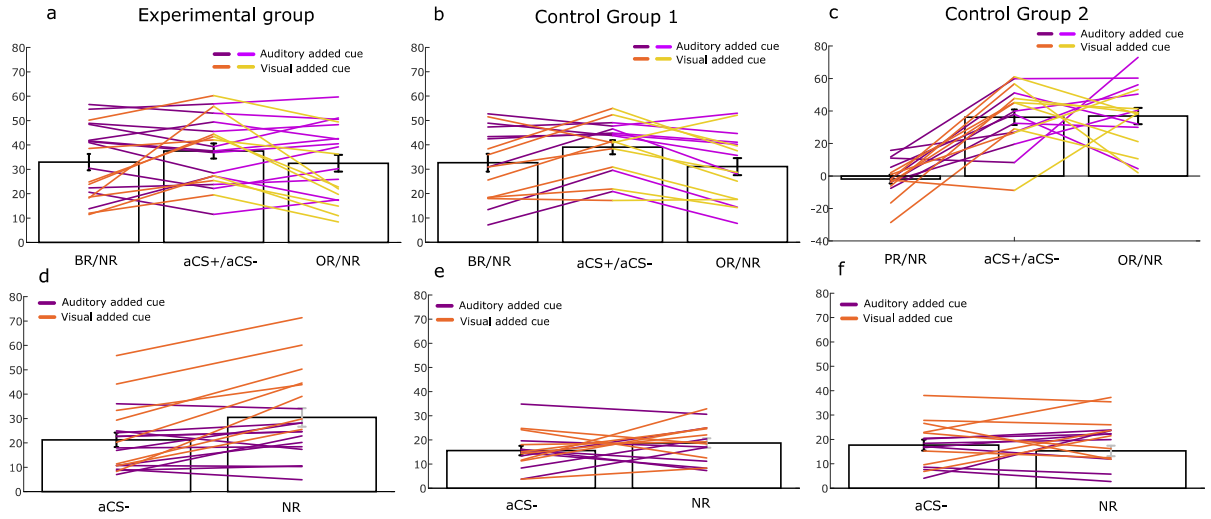
2.5 Supplemental Figures. Vicarious reward unblocks associative learning about novel cues in male rats.

Group names	Associative Learning	Pellet Dispensers active during Discrimination learning	Pellet Dispensers active during Social learning	Group size
Experimental group **Social-Appetitive (1A)**	BR, OR, NR No Control for secondary reinforcement of Pellet dispenser related cues	1 pellet dispensers	2 pellet dispensers	8 actors and 8 partners
Experimental group **Social-Only (1B)**	BR, OR, NR Controlled for secondary reinforcement of Pellet dispenser related cues	2 pellet dispensers	2 pellet dispensers	12 actors and 12 partners
Control group 1 **Inserted Wall (1A)**	BR, OR, NR Wall impeding social information transfer	1 pellet dispensers	2 pellet dispensers	8 actors and 8 partners
Control group 1 **No Partner Present (1B)**	BR, OR, NR Partner absent	2 pellet dispensers	2 pellet dispensers	8 actors and 8 partners
Control group 2 **Unequal Outcomes**	PR, OR, NR Outcomes Unequal	2 pellet dispensers	2 pellet dispensers	16 actors

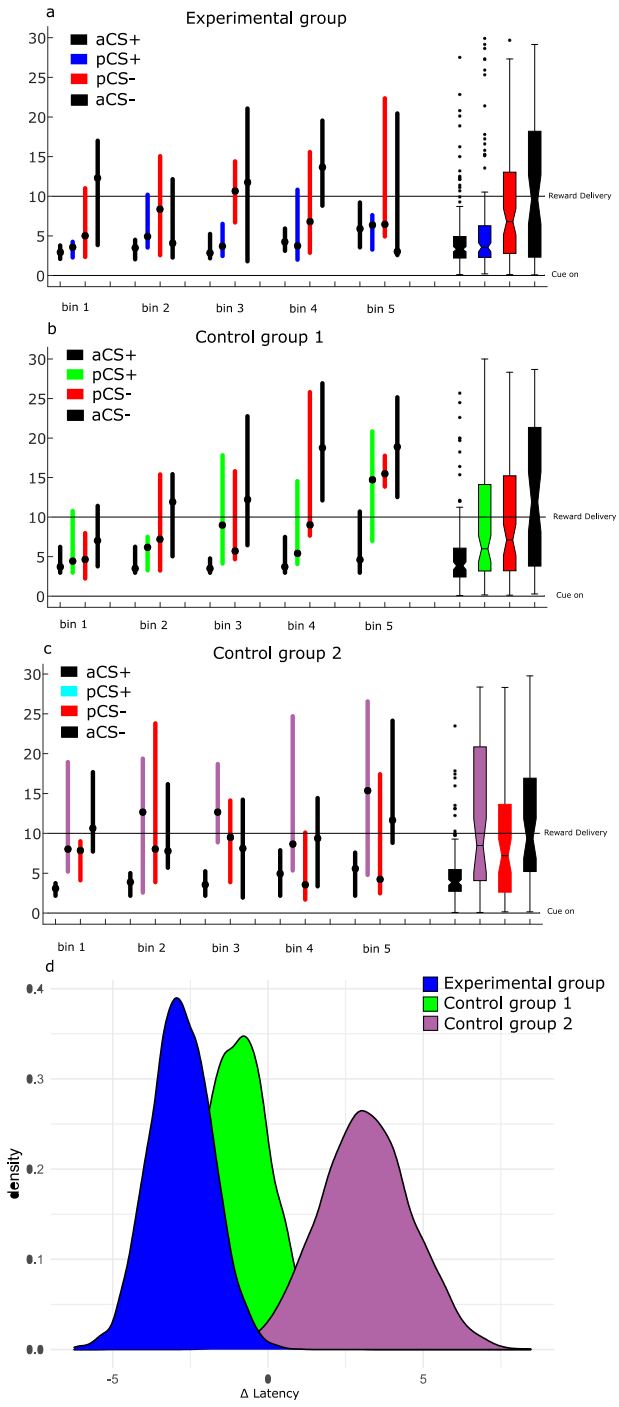
2. figure 1 - figure supplement 1.

Cues used *Flashing light (FL), Green light (GL),4 KhZ tone (4T), 1.5 KhZ tone (1.5T), 4 KhZ clicker (4C)*	**Discrimination learning**	**Social learning**	**Probe trials**
Experimental group 1a	*Visual (N = 4):* aCS+ = FL / aCS- = GL *Auditory (N = 4):* CS+ = 4T / aCS- = 1.5T	*Auditory unblocking (N = 4):* aCS+/pCS+ = FL + 4T aCS+/pCS- = FL + 1.5T aCS-/pCS- = GL + 1.5T *Visual unblocking (N = 4):* aCS+/pCS+ = 4T + FL aCS+/pCS- = 4T + GL aCS-/pCS- = 1.5T + GL	aCS+ = FL pCS+ = 4T pCS- = 1.5T aCS- = GL aCS+ = 4T pCS+ = FL pCS- = GL aCS- = 1.5T
Experimental group 1b	*Auditory (N = 4):* aCS+ = 4T / aCS- = 4C *Visual (N = 4):* aCS+ = FL / aCS- = GL *Visual (N = 4):* aCS+ = GL / aCS- = FL	*Visual unblocking (N = 4):* aCS+/pCS+ = 1.5T + FL aCS+/pCS- = 1.5T + GL aCS-/pCS+ = 4C + GL *Auditory unblocking (N = 4):* aCS+/pCS+ = FL + 1.5T aCS+/pCS- = FL + 4C aCS-/pCS- = GL + 4C *Auditory (N = 4):* aCS+/pCS+ = GL + 4C aCS+/pCS- = GL + 1.5T aCS-/pCS- = FL + 1.5T	aCS+ = 1.5T pCS+ = FL pCS- = GL aCS- = 4C aCS+ = FL pCS+ = 1.5T pCS- = 4C aCS- = GL aCS+ = GL pCS+ = 4C pCS- = 1.5T aCS- = FL
Control group 1	*Auditory (N= 8):* aCS+ = 1.5T / aCS- = 4C *Visual (N= 8) :* aCS+ = FL / aCS- = GL	*Visual unblocking* (N = 8): aCS+/pCS+ = 1.5T + FL aCS+/pCS- = 1.5T + GL aCS-/pCS- = 4C + GL *Auditory unblocking (N = 8):* aCS+/pCS+ = FL + 1.5T aCS+/pCS- = FL + 4C aCS-/pCS- = GL + 4C	aCS+ = 1.5T pCS+ = FL pCS- = GL aCS- = 4C aCS+ = FL pCS+ = 1.5T pCS- = 4C aCS- = GL
Control group 2	*Auditory (N= 8):* aCS+ = 1.5T / aCS- = 4C *Visual (N= 8) :* aCS+ = FL / aCS- = GL	*Visual unblocking* (N = 8): aCS-/pCS+ = 4C + FL aCS+/pCS+ = 1.5T + GL aCS-/pCS- = 4C + GL *Auditory unblocking (N = 8):* aCS-/pCS+ = GL + 1.5T aCS+/pCS- = FL + 4C aCS-/pCS- = GL + 4C	aCS+ = 1.5T pCS+ = FL pCS- = GL aCS- = 4C aCS+ = FL pCS+ = 1.5T pCS- = 4C aCS- = GL

2. figure 1 - figure supplement 2.



2. figure 2 - figure supplement 2.



2. figure 3 - figure supplement 1. Latency to entry for the different groups. (**a**) Experimental group (combined social-appetitive and social-only subgroups), (**b**) control group (combined inserted wall and no partner present subgroups) 1 (**c**) control group 2 (unequal outcomes). In all groups, the median latency to enter the food cup during the probe trials is shown over five bins (two trials per bin). Black dots indicate the median and bars display the interquartile range. Notched boxplot on the right display the distribution per condition over all bins. (**d**) Bootstrap distributions of mean latency differences between trial types pCS+ and pCS-.

Figure 3 - Source Data 1. Probing vicarious associative learning				
Experimental group	**F value** (main effect)	**P value**	η_p^2	
RMA -- WS; aCS+,aCS-, pCS+, pCS-, bin 1- 5	$_{(3, 57)} = 83.180$	**< 0.001**	$\eta_p^2 = 0.814$	
Post hoc comparisons for main effect	**Mean difference**	**Std error**	**P value**	**Cohen's D** (confidence level = 0.95)
aCS+ | aCS-	37.406	2.964	**< 0.001**	2.23
pCS+ | aCS-	14.544	2.257	**< 0.001**	0.88
pCS+ | pCS-	10.804	2.592	**0.003**	0.6
pCS- | aCS-	3.74	1.19	0.39	0.25
Control Group 1	**F value** (main effect)	**P value**	η_p^2	
RM anova *RMA* -- WS; aCS+,aCS-, pCS+, pCS-, bin 1- 5)	$F_{(1.372, 20.585)} = 31.215$	**< 0.001**	0.675	
Post hoc comparisons for main effect	**Mean difference**	**Std error**	**P value**	**Cohen's D**
aCS+ | aCS-	32.043	2.958	**0.021**	1.91
pCS+ | aCS-	7.948	1.862	**0.004**	0.75
pCS+ | pCS-	3.687	2.146	0.637	0.3
pCS- | aCS-	4.26	1.634	0.119	0.44
Control Group 2	**F value** (main effect)	**P value**	η_p^2	
RMA -- WS; aCS+,aCS-, pCS+, pCS-, bin 1- 5)	$(F_{(1.732, 25.980)} = 46.215$	**< 0.001**	0.755	
Post hoc comparisons for main effect	**Mean difference**	**Std error**	**P value**	**Cohen's D**
aCS+ | aCS-	30.162	3.754	**< 0.001**	1.76
pCS+ | aCS-	0.318	2.294	1	0.04
pCS+ | pCS-	0.513	2.662	1	0.02
pCS- | aCS-	-0.195	2.1	1	-0.01
Experimental vs Control 1 & 2	**F value**	**P value**	η_p^2	
Group comparison	(interaction effect, *10 seconds / 30 seconds period after cue onset)*	*(10 seconds / 30 seconds period after cue onset)*	*(10 seconds / 30 seconds period after cue onset)*	
Mixed–RMA -- WS; aCS+,aCS-, pCS+, pCS-; trial 1 - 6 & BS; Experimental vs control 1 & 2	$F_{(4.300, 05.354)} = 3.713$ / $F_{(4.140, 101.422)} = 2.748$	**0.002 / 0.031**	0.132 / 0.101	
Post hoc comparisons for interaction effect	**Mean difference**	**Std error**	**P value**	
Experimental group	10 seconds/ 30 seconds			
aCS+ | aCS-	44.10 / 54.727	4.242 / 3.760	**< 0.001 / < 0.001**	
pCS+ | pCS-	15.753 / 10.711	3.021 / 2.319	**< 0.001 / < 0.001**	
Control group 1				
aCS+ | aCS-	32.851 / 46.504	5.556 / 4.204	**< 0.001 / < 0.001**	
pCS+ | pCS-	2.850 / 2.433	3.377 / 2.593	1.00 / 1.00	
Control group 2				
aCS+ | aCS-	34.700 / 47.969	5.556 / 4.204	**< 0.001 / < 0.001**	
pCS+ | pCS-	1.233 / -0.782	4.743 / 2.593	1.00 / 1.00	

	Figure 4 - Source Data 1. Strength of the social unblocking effect over trials			
Experimental vs Control 1 & 2 - Difference scores				
(Mixed–RMA -- WS; aCS+/aCS-, pCS+/pCS-, pCS+/aCS- and pCS-/aCS- contrasts, trial 1 – 6; BS; Experimental vs control 1)	**F value, P value, η_p^2** *(Between subjects effect)*	**Experimental – control 1:** *Mean difference, std error, p-value*	**Experimental – control 2:** *Mean difference, std error, p-value*	**Control 1- Control 2:** *Mean difference, std error, p-value*
10 seconds period after cue onset				
aCS+/aCS-	$F_{(2, 49)} = 1.859$, 0.167, 0.071	11.242, 6.363, 0.251	9.400 , 6.363, 0.438	-1.842, 6.707, 1.00
pCS+/aCS-	$F_{(2, 49)} = 6.397$, **0.003**, 0.207	12.903, 4.531, **0.019**	14.520, 4.531, **0.007**	1.617, 4.776, 1.00
pCS+/pCS-	$F_{(2, 49)} = 15.554$, **< 0.001**, 0.388	13.250, 3.908, **0.004**	21.392, 3.908, **0.001**	8.142, 4.119, 0.161
pCS-/aCS-	$F_{(2, 49)} = 2.429$, 0.099, 0.090	0.4347, 3.406, 1.00	6.872, 3.406, 1.00	6.525, 3.590, 0.226
30 seconds period after cue onset				
aCS+/aCS-	$F_{(2, 49)} = 1.251$, 0.295, 0.049	8.223, 5.640, 0.454	6.757, 5.640, 0.710	-1.465, 5.945, 1.00
pCS+/aCS-	$F_{(2, 49)} = 5.976$, **0.005**, 0.196	8.278, 3.478, 0.064	11.493, 3.478, **0.005**	3.215, 3.667, 1.00
pCS+/pCS-	$F_{(2, 49)} = 12.881$, **< 0.001**, 0.345	9.654, 3.388, **0.021**	17.054, 3.388, **< 0.001**	7.490, 3,572, 0.123
pCS-/aCS-	$F_{(2, 49)} = 1.467$, 0.241	1.286, 3.337, 1.00	5.561, 3,337, 0.306	4.275, 3.518, 0.690

Videos can be found at: https://elifesciences.org/articles/60755/figures#content

3. Posterior-lateral Orbitofrontal cortex lesion leaves social and appetitive unblocking intact but impairs the blocking effect related to partner no-reward.

Abstract. The lateral Orbitofrontal cortex in rats supports cue-based outcome predictions and has been found to be necessary for coding the identity of reward during appetitive unblocking. A more posterior part of the lateral Orbitofrontal cortex (PLO) has been found to contain the cells that code the identity of upshifts/downshifts in appetitive unblocking. We hypothesised that during social (un)blocking, a procedure where a change in partner reward, rather than self-reward, is used to drive unblocking, PLO could also support the processing of partner related social-identity signals during mutual reward versus own-reward. PLO would be well positioned to link novel cues to social reward information and thus enable a vicarious associative learning process to unblock the value of these cues. We found, however, that PLO lesions did not impair social unblocking but replicate earlier findings that appetitive non-social unblocking remained intact in PLO lesioned rats. In contrast, we found moderate evidence that PLO lesioned actor rats showed impairments in blocking for novel cues associated with partner no-reward.

3.1 Introduction

Rats coordinate mutual behaviors (Tan & Hackenberg, 2016), reciprocally share food rewards (Rutte & Taborsky, 2008), forage together (Weiss, Dorfman, Ram, Zadicario, & Eilam, 2017) and make prosocial choices that depend on the behavior of the partner (Hernandez-Lallement et al., 2015; Márquez et al., 2015). It is vital that organisms learn to predict which cues in an environment signal such social experiences, and which cues are socially irrelevant. In learning theory, when an agent learns that a stimulus is fully predictive of a reward, and subsequently an additional stimulus is added while keeping the reward the same, learning about this additional stimulus is blocked (Kamin, 1969). Learning about additional predictive stimuli can become unblocked, however, when, for example the amount of reward is increased (upshift) compared to the amount predicted by the originally learned stimulus (Holland, 1984). Another way to unblock learning is to change the sensory identity of rewards, by changing a food pellet to an liquid sucrose reward (identity unblocking; Rescorla, 1999).

Chapter 3 is submitted and under review in eNeuro: van Gurp S, Seidisarouei M, Kalenscher T, van Wingerden M. 2020. Posterior-lateral Orbitofrontal cortex lesion leaves social and appetitive unblocking intact but impairs the blocking effect related to partner no-reward. eNeuro.

McDannald et al. (2011) found that lesioning the lateral Orbitofrontal cortex impaired appetive unblocking driven by identity (change in flavour) but not value upshift signals. In later studies they found that a specific posterior region of the lateral orbitofrontal cortex (PLO) contains cells that fire to novel odors that predict such a 'valueless' change in identity (McDannald et al., 2014) and direct upshifts in value (Lopatina et al., 2015) during unblocking.

In a recently introduced social unblocking paradigm, the increase in reward was related to rewards delivered to a partner rat, receiving a food pellet reward together with an actor, instead of changes in self-reward (van Gurp, Hoog, Kalenscher, & van Wingerden, 2020). It was found that learning about cues that predicted food to the partner would become unblocked, when they were associated with reward to both an actor and a partner (Both-Reward; BR) but remained blocked when the actor was rewarded but not the partner (Own-reward; OR). This social unblocking effect was hypothesed to be driven by a vicarious other-regarding learning process through a change in social-identity signals related to the partner's behavior that indicated the value of the mutual reward. Actor rats could observe the partner rat approach the food cup and consume the pellet rewards during Both-reward but not Own-reward trials (as preventing visual observation of the partner, and partner absence both impaired the social unblocking effect) and, furthermore, the actor could learn about the presence of food via the breath of the partner through openings in the cage divider (cf. Galef, 2001) after the Both Reward but not Own-reward (as preventing social interaction impaired the social unblocking effect). In social learning tasks in primates, there is mixed evidence that the orbitofrontal cortex is involved in making social value computations, as it contains subpopulations of cells encoding juice rewards delivered to a monkey itself, but not juice rewards delivered to another partner monkey (Chang et al., 2013) but also was shown to code a signal for cue specific identity predictions for both-reward versus self-reward outcome and contain cells encoding the value of both-reward, but only when the decisions involved the preferred social partner (Azzi et al., 2012). Even though lesioning the orbitofrontal cortex does not degrade normal social interaction between adult rats (Rudebeck et al., 2007), this part of the brain in rodents has been found to contain cells that increase activity during social approach and interaction and, when activated, inhibit consummatory behavior in mice (Jennings et al., 2019). It is thus likely that the rat posterior-lateral Orbitofrontal cortex (PLO) encodes partner related social-identity signals that characterise Both-reward and/or Own-reward during social unblocking. We therefore hypothesised that the social unblocking effect depended on PLO integrity and would be impaired by lesioning the PLO. When lesioning the PLO, we found however that neither the social (Partner-Reward > Partner No-Reward) nor the non-social, appetitive (Upshift > No-Shift) unblocking effect was affected when tested in extinction, compared to a sham lesion and a behavioral control

group. This provides evidence that the PLO, at least in this task configuration, is not required for the calculation of social-identity signals and the expression of the social unblocking effect. We found, however, that lOFC lesioned rats showed impaired blocking to the cue that predicted no additional partner reward (and thus should remain blocked from acquiring associative value). This surprising effect is discussed as a potential deficit in correct credit assignment to the blocked cue, possibly because some Both-reward or Own-Reward value transferred to the partner No-Reward added compound cue (pCS-) in lesioned animals.

3.2 Methods.

Subjects and housing. 54 male adult Long Evans Rats were housed in pairs of two per cage and kept under an inverted 12:12 h light-dark cycle, in a temperature (20 ± 2 ºC) and humidity-controlled (approx. 60%) colony room. During behavioral testing, the rats where food-restricted (20 grams on weekdays and 22 grams on the weekend) and maintained a bodyweight of about 90% of their free-feeding weight. All testing was performed in accordance with the German Welfare Act and was approved by the local authority LANUV (Landesamt für Natur-, Umwelt und Verbraucherschutz North Rhine-Westphalia, Germany, AZ 84-02.04.2016.A522).

Setup. The task setup (see Figure 1a) consisted of a customized PhenoTyper (Noldus Information Technology) that was divided in two compartments, in one compartment the actor was conditioned and in the other the partner. The compartments were separated by a transparent wall. In each compartment there was a food cup present (Med Associates, ENV-254-CB) were the rats could obtain food pellets (20 mg dustless precision pellets, Bio-Serv, Germany). Two light sources were used, one LED Stimulus Light (Med Associates, ENV-211m) with a green cover and a house light (Med Associates, ENV-215m) and sounds were played from a speaker (Med Associates, ENV-224am) at 75 dB.

The social unblocking effect: experiment timeline.

The social unblocking procedure (see figure 1b) consists of a discrimination learning phase, a compound phase (referred to as: Social learning phase) and probe trials similar to a standard appetitive unblocking paradigm. In the discrimination learning phase, both actor and partner rats separately, in their assigned compartments, learned a visual or auditory discrimination task where they discriminated between an aCS+ (conditioned stimulus leading to a reward; "A" stand for actor) from an aCS- (conditioned stimulus leading to the absence of reward). Afterwards, during the compound phase, the visual stimulus is added into a compound with the auditory stimulus or vice versa. If the visual stimulus is learned by the actor than the compound phase contains of three

conditions: Both-reward (BR; actor aCS+ visual, partner pCS+ auditory), Own-reward (OR; actor aCS+ visual, partner aCS- auditory) or No-reward (NR; actor aCS- visual, partner aCS- auditory). All learned reward associations for Actor and Partner remain the same. Importantly, from the actor's perspective, the total reward delivered in the OR condition and NR conditions is not different from the total reward predicted by the aCS+ or aCS-, respectively. However, the BR conditions adds a partner reward that was not predicted by the aCS+, but is associated to the compound pCS+ stimulus. The social learning phase last for 4 days and the rats are trained

Figure 1. Experiment overview and trial timeline. (a) Task setup topview with two compartment visible and on the left the wall seperating the compartments and on the right the cue lights and food cups (b)Experiment timeline: Firstly, Actor rats learn to discriminate two visual cues while, at a different timepoint, partner rats learned to discriminate two auditory cues. Consequently actor rats receive either Sham or Lesion surgery of the PLO after which they recover for 10 days. After this the rats are retrained on the discrimination learning for 5 days. Then in the social learning phase actor and partner rat are either both rewarded (BR, aCS+/pCS+), actor rat is rewarded while the partner is not rewarded (OR, aCS+/pCS-) or neither actor nor partner rat was rewarded (NR, aCS/pCS-). In the probe trials, all compound cues are presented to the actor and at a different time to the partner rat without reward. After the social probe trial actor rats received one day of a reminder trials. Actor rats then went through 3 conditions: Upshift (aCS+,UCS), Same reward positive (aCS+,sCS1) and Same reward negative (aCS+,sCS2). Finally, the learned cues (aCS+, UCS, sCS1 and aCS-) were shown in an exinction trial. (c) The trial timeline for social unblocking is shown. Each trial a cue light is presented for 30 seconds with pellet drops at 10, 20 and 30 seconds. (d) The trial timeline for appetitive unblocking is shown. Each trial a cue light is presented for 30 seconds with two pellet drops at 10, 20and 30 seconds

together, receiving 20 pseudo-randomized trials per condition. Afterwards in the probe trials the cues (aCS+, pCS+, pCS-, aCS-) are presented without a reward to both actor and partner separately. The probe trials consisted of 10 trials per cue, again presented in a pseudo-randomized manner. It was found that, during the first six probe trials, the actor rats spent more time in the food cup for the partners pCS+ than the pCS- (van Gurp et al., 2020). This effect (pCS+ > pCS-) was termed the social unblocking effect and was indicative that social learning had taken place. Somehow, the value associated with the partner rat had become associated to the pCS+, leading to approach in extinction. The social nature of the vicarious social reward was established by confirming that both preventing social information exchange and partner rat absence impaired the social unblocking effect. Here, we used the same three learning stages of the social unblocking task (discrimination learning, social learning phase and probe trial), but lesioned the posterior-lateral part of the orbitofrontal cortex directly after the discrimination learning phase to investigate if this region is necessary for the social unblocking effect to occur. A post-surgery retraining period on the discrimination period was added to confirm that basic discrimination was still intact. Afterwards, to be able to directly compare the PLO lesion effect on social versus appetitive unblocking, we added an appetitive compound phase with its own probe phase. In the appetitive compound phase the actor was trained alone with three conditions. In the Upshift condition (aCS+, upCS), the added stimulus (upCS) signalled going from three to six pellets, while the Same reward positive condition (aCS+, sCS1) signalled no change; three pellets remained three pellets and finally in the Same reward negative condition (aCS-, sCS2) signalled absence of pellets for the added sCS2 cue. Here, during the probe trials the a CS+, upCS, sCS1 and the aCS- were shown without reward. An appetive (upshift) unblocking effect would be visible if the time in spent in the food cup would be higher for the upCS than the sCS1.

Social unblocking: detailed procedure.

Discrimination learning. Actor and partner rats were randomly assigned in their home cage, and both actors and partners were habituated to their assigned compartment of the PhenoTyper for 3 days. Afterward, both actor and partner went through 14 days of discrimination learning in their respective compartment. Importantly, here we decided to simplify interpretation of the unblocking effect by focussing on auditory unblocking (BR example: aCS+ visual; pCS+ auditory) only, excluding visual unblocking (BR example: aCS+ auditory; pCS+ visual). This gave us the possibility to directly compare, within the auditory stimulus modality, social and appetitive unblocking (Figure 1b). Actor rats always learned to discriminate between a flashing light (flashing at 1 Hz 0.1 sec on, 0.9 sec off; aCS+) and a steady green light (aCS-). Partner rats, on the other hand, always learned to discriminate between a 1.5 kHz (75 dB; aCS+) steady tone a 4.0 kHz clicker (0.1sec on, 0.9 sec

off; aCS-). During the discrimination-learning all rats went through 20 trials a day per condition and the order of aCS+ and aCS- trials was pseudo-randomized. Stimuli were presented for 30s and for the actors aCS+ a pellet was dropped at every 10s, 20s and 30s (+ 0.1 to 0.4 sec jitter) after cue onset (Figure 1c). Finally, we added a pellet dispenser delivering pellets outside of the box during discrimination learning, which we previously have shown to act as a suitable control for secondary reinforcement related to pellet dispenser reward delivery cues during compound conditioning.

Surgery. When discrimination learning was ready on day 14, rats underwent surgery. First, they were anesthetized by ways of isofluorane inhalation (5% for induction and 3% for maintenance), and positioned on a stereotaxic frame (David Kopf Instruments, USA). On each hemisphere one hole was drilled at the following coordinates (anteroposterior (AP) - 3.0, mediolateral (ML) - 3.2 mm, and dorsoventral (DV) - 5.0 mm) targeting the posterior-lateral Orbitofrontal cortex (PLO) previously found to be related to value upshifts in unblocking (Lopatina et al., 2015) and learning driven by unexpected outcomes (Takahashi et al., 2009). Actor rats (N = 14) received bilateral infusions using a 0.3 mm injection needle (PlasticsOne) connected via polyethylene tubing to a 10 µl Hamilton syringe within a microinfusion pump (Harvard apparatus). Infusions were made using 0.2 µl of 0.09 M quinolinic acid dissolved in 0.1 M phosphate buffer solution (PBS, pH value 7.4) at an infusion rate of 1 µl/min, after which the needle was left in place for three minutes allowing the substance to diffuse away from the injection site. Sham surgeries (n= 13) were made by infusing vehicle solutions (0.1 M PBS, pH value 7.4) according to the same protocol. After completing the surgery, animals received analgesic injections (Carprofen; 5 mg/ml) for three consecutive days and recovered for a period of 10 days. Both the actors of the lesion (N = 14) and sham group (N=13) and their partners (N = 27) were trained for 5 days of days after surgery and then went on to the compound phase.

Social Unblocking. During the social learning phase the actor and partner rat were trained together, in their respective compartments, receiving the conditions: BR (aCS+, pCS-), OR (aCS+, pCS-) and NR (aCS-, pCS-) to induce social unblocking/blocking (see paragraph the social unblocking effect and Figure 1b). Stimuli were again presented for 30s and for both the actors aCS+ and partners pCS+ a pellet was dropped in their respective compartments, at every 10s, 20s and 30s (+ 0.1 to 0.4 sec jitter) after cue onset (Figure 1c). .During probe trials, rats were tested in isolation receiving all stimuli (aCS+, aCS-, pCS+, pCS-) alone without reward.

Appetitive Unblocking. After social unblocking, only the actor rats continued and got one day of conditioning on the discrimination problem. Subsequently, actor rats went on to go through 4 days of appetitive unblocking/blocking (Figure 1a). Here, the conditions, as described in social

unblocking effect introduction, were as follows: Upshift (aCS+, upCS), Same reward positive (aCS+, sCS1), Same reward negative (aCS-, sCS2). Stimuli were presented for the 30s, and for the aCS+ /upCS a pellet was dropped (+ 0.1 to 0.4-sec jitter) at 10 and 11 seconds, 20 and 21 seconds, and finally at 30 and 31 seconds after cue onset (Figure 1c). During Probe trials, the rats were again tested in isolation receiving the stimuli (aCS+, upCS, sCS1, aCS-) alone without reward. The sCS2 cue was not tested during probe trials due to maintain the same number of items tested in extinction.

Statistical data analyses. Photo beam breaks in the food cup were extracted from EthoVision XT 11.5 (Noldus Information Technology), and from this, we calculated the time spent in the food trough. Graphs were made using Matlab (version 2014b, MathWorks) and all statistics were done using JASP (JASP Team (2020), JASP (Version 0.13.1). For discrimination learning and the social learning phase we calculated the percentage time spent in the food cup for the first 10 seconds after cue onset. In the lesion actor group, the discrimination and retraining data of two rats were lost because of a setup error, and in both the sham and lesion group, the data for 8 partner rats was lost due to a broken experiment file. For both appetitive and social probe trial analyses, we applied baseline correction in calculating cue-based activity. This baseline-corrected score was calculated by subtracting a percentage of time spent in the food cup during the baseline period of 30 seconds in the food cup from the percentage of time spent in the food cup in the 10 seconds after cue onset. Unblocking was assessed by the difference in baseline-corrected response between to conditions (unblocking difference score). We used mixed repeated measures ANOVA and t-tests for the statistical analyses or non-parametric equivalents (Wilcoxon signed rank test, or Mann Whitney U test) when appropriate. For further qualification of evidence, we added Bayesian independent and dependent t-tests. Post hoc comparisons were bonferonni corrected. For all the repeated measures ANOVAs a greenhouse geisser corrections were used if the assumption of sphericity was violated.

3.3 Results.

Histology results. Lesion targeted the posterior lateral part of the orbitofrontal cortex where Lopatina et al. (2015) found neurons acquired responses to upshift, downshift cues during unblocking, which are likely to be required for identity unblocking (McDannald et al., 2011) and roughly appears similar to a recent paper in which posterior and anterior lesion were differentiated (Panayi & Killcross, 2018). Histological assessment was performed by S.G. and confirmed by one additional person, who was blind to the experimental groups. Excitotoxic lesion damage (N=11) was small but specific and included five actor rats in which the lesion was mostly located from

anterior 3.0 to anterior 3.4 (Figure 2c) and six actor rats where the lesion was located anterior from 3.2 to 3.7 (Figure 2a). An example of the lesion at 3.2 is shown in figure 2b and an example of the lesion at 3.7 is shown in figure 2d. Three actor rats from the lesion group were excluded due to lesion location outside the PLO. Two actor rats from the Sham group were excluded due to signs of a possible lesion. In all other rats from the sham group, no lesion was apparent (N=11). All partner rats of the excluded actor rats were also excluded. The final groups included in the data-analyses below for all three phases were: actor-lesion (N = 11), actor-sham = 11, partner-lesion (N = 11), partner-sham = 11).

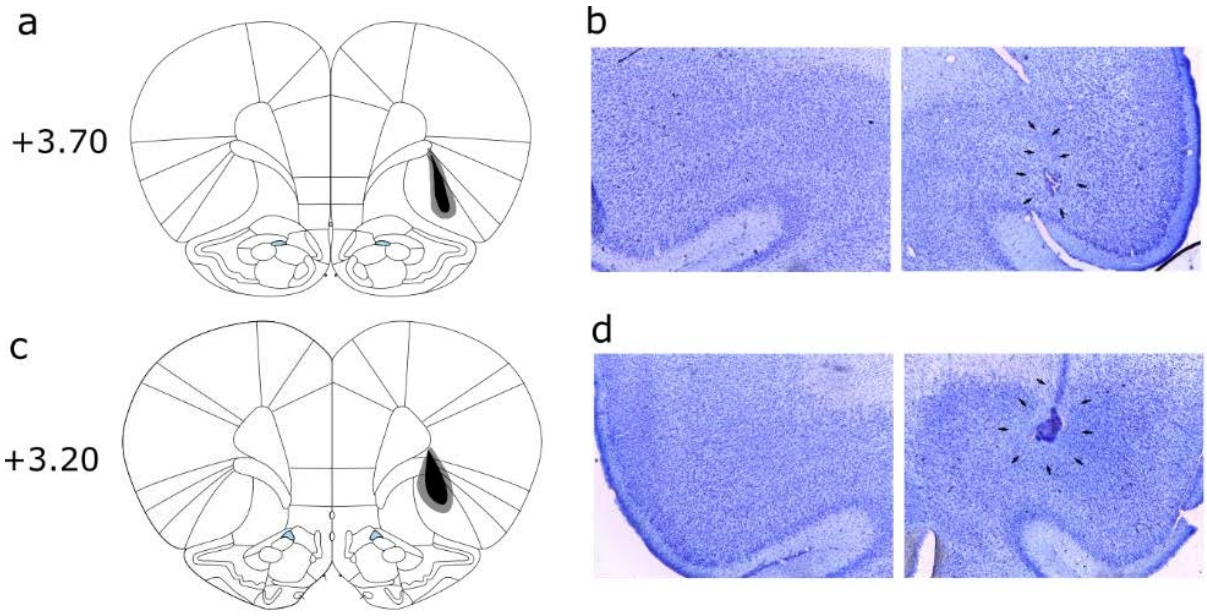


Figure 2. Histology. (a, c) Mininum (blac) and maximum (light grey) posterior-lateral OFC lesions are showns at bregma +3.2 and +3.7 AP. (adapted from Paxinos and Watson, 2007) (b) Example of sham (left) and lesion (right) at bregma + 3.7 (d) Example of sham (left) and lesion (right) at bregma + 3.2

Both Sham and Lesion rats show similar discrimination and social learning. Actor rats of both groups were trained for 14 days on a visual discrimination problem (Figure 2a, c) while their partner rats trained on an auditory discrimination problem (Supplemental Figure 2b, d). Afterward, the actor and partner rats went through a social compound phase, and finally, they went through 10 probe trials (Figure 2a). Here, we first describe the actor rats performance throughout discrimination and social learning phase and check if there are any differences between the Lesion and Sham group. During discrimination learning, Actor rats of both groups acquired a conditioned response to the aCS+ while diminishing their responding to the aCS-. Average responding during the last 4 days was larger for aCS+ over aCS- in the Sham group ($T(10) = 8.045$, $p < 0.001$; Figure 2a) and Lesion group ($T(8) = 12.683$, $p < 0.001$; Figure 2c). To directly compare the discrimination learning, we use a difference score of aCS+/aCS-. A Mann-Whitney U test showed no significant

differences in responding on the last 4 days of training between groups in the strength of discrimination learning (U (N-Lesion = 9, N-Sham = 11) = 66, p = 0.230). After the 14 days of training, actor rats received lateral Orbitofrontal cortex lesions, recovered for 10 days, and were subsequently retrained for 5 days. A paired sample t-test revealed that Discrimination learning strength was not significantly affected by the lesion (Figure 2a, c) when comparing the difference scores (pre versus post-lesion; (T (19) = -1.092, p = 0.856). We also found no significant differences in discrimination learning when directly comparing the difference score post-lesion (T(18) = 0.778, p = 0.477). After retraining, actor rats went on to the social learning phase (Figure 2b, d) were they were paired with their partners and trained in the conditions BR (aCS+, pCS+), OR (aCS+, pCS-), and NR, (aCS-, pCS-). For the analyses, we calculated the mean over 4 days of responding and performed a mixed ANOVA on the actor rats percentage of time spent in the food cup during the BR, OR, and NR with Conditions (BR, OR, and NR) as a within-subject factor and Group (Sham, Lesion) as a between between-subject factor and two separate repeated measure Anova for Sham and Lesion group separately. We find that there are no differences between the groups during the social learning phase (for statistics see: Table S1). Finally, we checked whether the act of just adding a novel cue in the social learning phase changed responding between groups. In addition to the post-lesion discrimination difference score (aCS+/aCS-) we calculated a BR/NR and OR/NR difference score. We performed a repeated-measures ANOVA with conditions (aCS+/aCS-; BR/NR) and the between-subject factor group (Sham/Lesion). We find no significant interaction ($F(1,17) = 0.082$, $p = 0.778$, $\eta_p^2 = 0.002$) indicating that adding a cue that predicts mutual reward did not differentially change responding towards the food cup between groups. We then performed a repeated-measures ANOVA with conditions (aCS+/aCS-; OR/NR) and the between-subject factor group (Sham/Lesion). We find no interaction ($F(1,17) = 0.131$, $p = 0.722$, $\eta_p^2 = 0.003$) indicating that adding a cue that predicts own reward and partner no-reward did not differentially change responding towards the food cup in between groups. Overall, these results indicate that lOFC lesions did not affect food cup responding during discrimination or social learning.

Both Sham and Posterior-lateral OFC lesioned rats show the social unblocking effect. After the actor rats of both the Lesion and sham group went through the social learning phase, we presented the aCS+, pCS+, pCS- and aCS- in 10 extinction trials without food reward (See Sham group in Figure 4c, for Lesion group in figure 4f and for the behavioral data group in figure 4i). Because only an auditory steady tone cue was used for unblocking during this experiment, we include the auditory data from the van Gurp et al. (2020) dataset (Behavioral data group; N = 8) to make a direct comparison of the strength of social unblocking effect in this experiment.

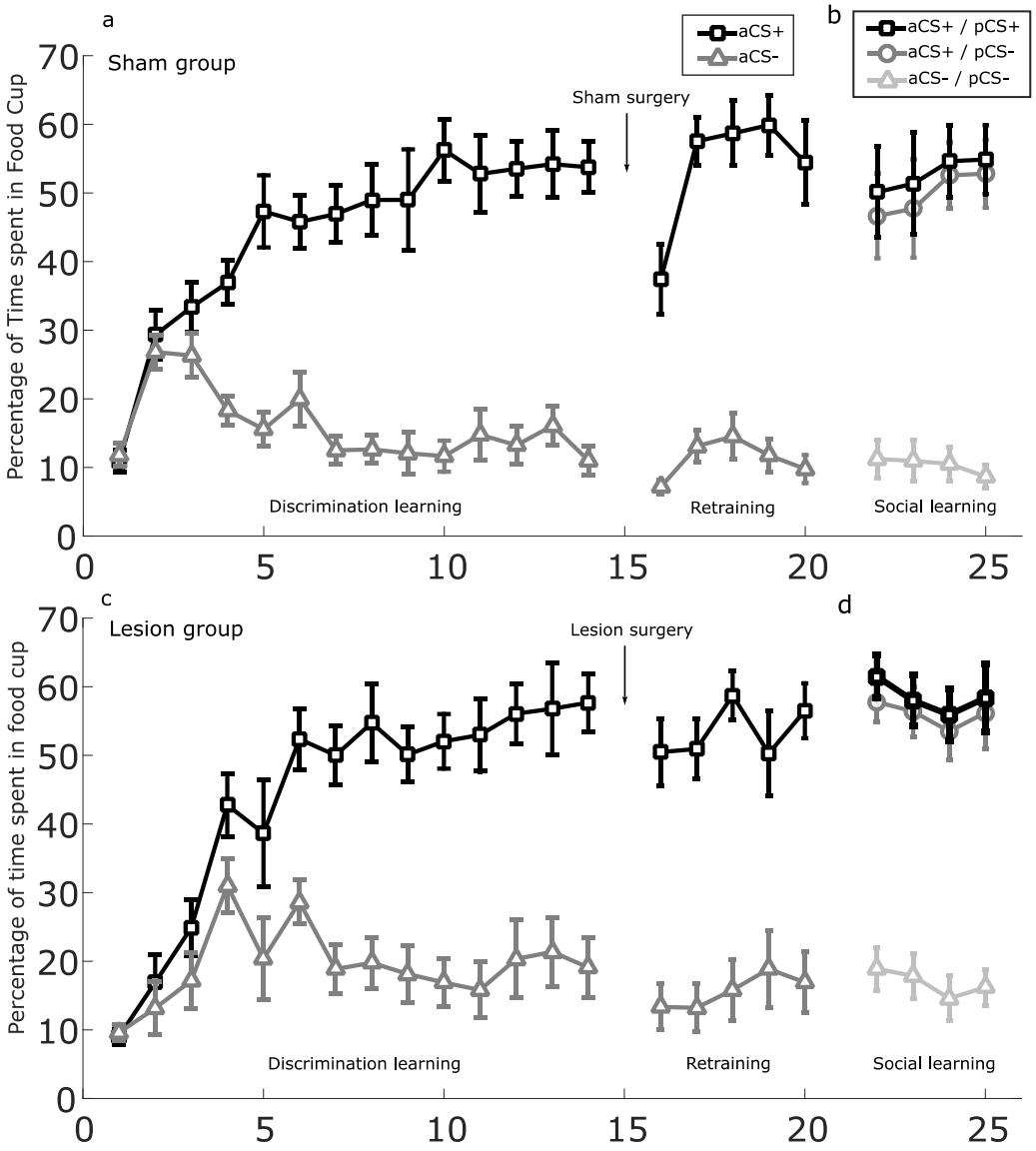


Figure 2. Conditioning per experimental phase. *Sham group.* (a) Percentage of time spent in food cup for discrimination learning between aCS+ and aCS- over days. (b) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days. *Lesion group.* (c) Percentage of time spent in food cup for discrimination learning between aCS+ and aCS- over days. (d) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days.

In this dataset, it was observed that the social unblocking effect lasted for the first 6 trials during the probe trials, with the strongest effects on trial 1. We, therefore, focused here on responding on the first 6 trials and, more specifically, on trial 1. For all conditions, we calculated a baseline-normalized difference score (i.e., the percentage of time spent in a 30 seconds baseline response subtracted from the percentage of time spent in the 10 seconds post cue onset; see method section for details) to ensure the data's reliability. If social unblocking is dependent on a social-identity signal dependent coding of mutual rewards, we expect lOFC lesions to impair social unblocking

(cf. McDannald, 2011). To see if lOFC lesions affect the social unblocking effect we compared the sham lesion group (N = 11) to the lOFC lesion group (N = 11) and the Behavioral data group.

Sham actor rats and behavioral data actor rats, but not lesioned actor rats show unblocking on trial one. We first checked for clear difference between groups (Sham, Behavioral data, Lesion) on the social unblocking effect (pCS+ < pCS-) by performed a two factor repeated measures ANOVA with Condition (aCS+, pCS+, pCS- and aCS-) and trial (trial 1 to 6) as factors and Group (Sham, Behavioral data, Lesion) as a between-subjects factor. While this analyses (for statistics see: Table S1) showed no significant interaction effects between group and condition or group, condition and trial, we decided for a more direct comparison. As the unblocking effect is present when the pCS+ > pCS-, we directly compared the pCS+ and pCS- difference score on trial 1 and afterwards over 6 trials of responding. Here, we expected to observe unblocking in the sham but not lesion group if social-identity signals were necessary for social unblocking. Paired T-test revealed that difference score in the Sham group was higher for pCS+ (M = 22.642, SD = 34.675) over pCS- (M = 22.642, SD = 34.675) in trial 1 (T (10) = 2.251, p = 0.024, BF_{+0} = 3.402, with median posterior delta = 0.571, 95% CI = 0.075 - 1.218; Figure 4a). Paired T-test furthermore revealed that difference score in the lesion group was not higher for pCS+ (M = 21.139, SD = 30.720) over pCS- (M = 11.976, SD = 18.231) in trial 1 (T (10) = 0.903, p = 0.194, BF_{+0} = 0.660, with median posterior delta = 0.291, 95% CI = 0.018-0.808; Figure 4d). Finally, a Paired T-test revealed that difference score in the Behavioral data group was higher for pCS+ (M = 31.150 SD = 30.720) over pCS- (M = -11.067, SD = 23.736) in trial 1 (T (10) = 3.350, p = 0.006, BF_{+0} = 11.021, with median posterior delta = 0.955, 95% CI = 0.180-1.913; Figure 4g). These results show that in trial 1 of the probe trials boths Sham and behavioral data actor rats show the social unblocking effect (pCS+ > pCS-) but not the PLO lesioned actor rats.

Lesioned and Behavioral Data actor rats, but not Sham actor rats, show unblocking during the first six trials. We then looked at the first six trials of the probe trials to see if the social unblocking effect was present. We performed a repeated measures design with the factors condition (pCS+ and pCS-) and trial (1 to 6) and the between-subjects factor Group (Sham, Lesion, Behavioral Data). We find an effect of condition F (1,27) = 26.569 , p < 0.001, η_p^2 = 0.496) and a near significant interaction effect of group * Condition F (2,27) = 2.712 , p = 0.084, η_p^2 = 0.167). Post hoc comparison revealed that the difference score was higher for pCS+ over pCS- over 6 trials in the Lesion (mean difference = 16.790, se = 3.628, p = 0.004 **; Figure 4e) and in the Behavioral Data group (mean difference = 14.619, se = 4.197, p = 0.026 *; Figure 4h) but not the Sham Group (mean difference = 4.33, se = 3.579, p = 1.00; Figure 4b). These data provide evidence that the sham group showed unblocking on the first trial but not during the first

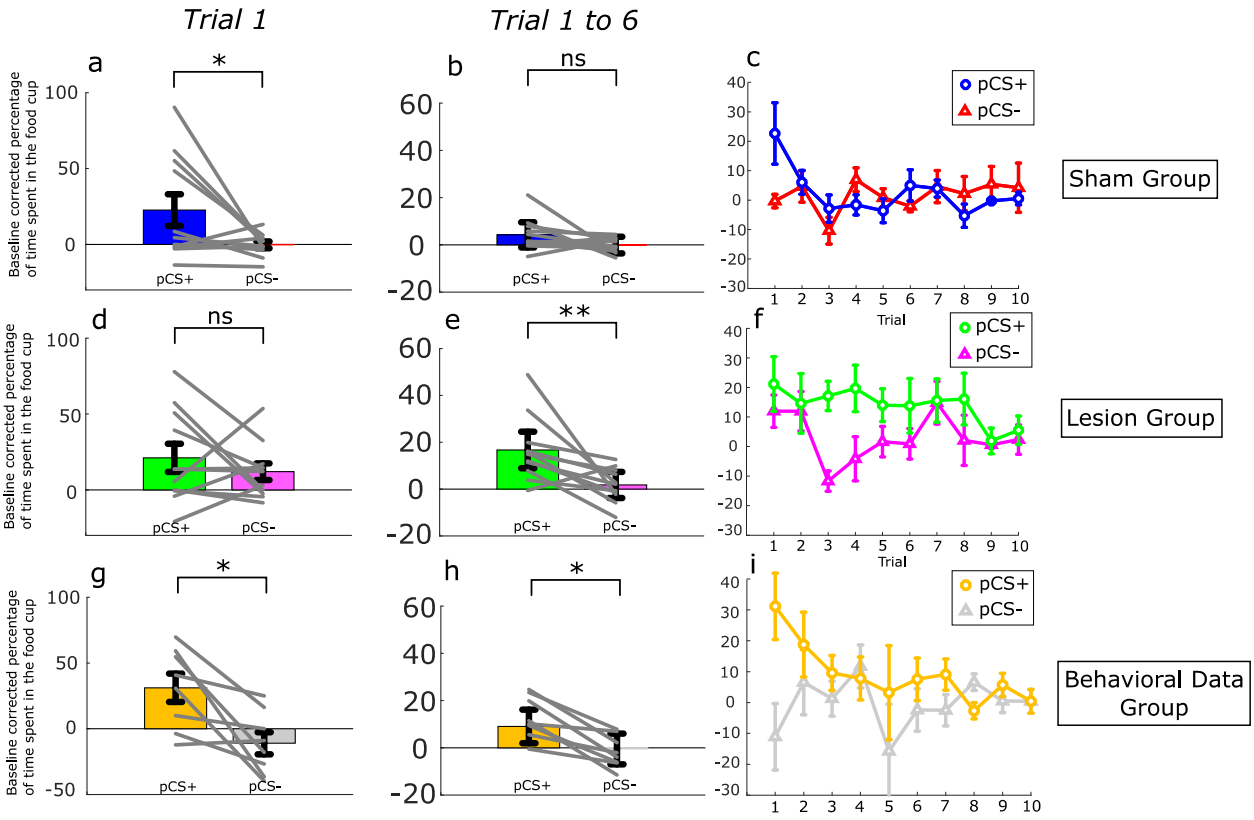


Figure 4. Food cup response of the actor rat during social probe trials. (a) Percentage of time spent in the food cup of the Sham group for the pCS+ and pCS- cues on trial 1. (b) Percentage of time spent in the food cup of the Sham group for the pCS+ and pCS- cues averaged over the first 6 trials. (c) Sham group's Percentage of Time spent in the food cup of the pCS+ and pCS- during extinction. (d) Percentage of time spent in the food cup of the Lesion group for the pCS+ and pCS- cues on trial 1. (e) Percentage of time spent in the food cup of the Lesion group for the pCS+ and pCS- cues averaged over the first 6 trials. (f) Lesion group's percentage of time spent in the food cup of the pCS+ and pCS- during extinction. (g) Percentage of time spent spent in the food cup of the Behavioral data group for the pCS+ and pCS- cues on trial 1. (h) Percentage of time spent spent in the food cup of the Behavioral data for the pCS+ and pCS- cues averaged over the first 6 trials. (i) Behavioral data group's Percentage of time spent in the food cup of the pCS+ and pCS- during extinction. All percentage of time spent in the food cup was baseline corrected.

six trials while the lesion group showed social unblocking during the first six trials but not on the first trial. It is, therefore, clear that all groups showed characteristics of social unblocking with differing time courses. We, therefore conclude here that, even though not in a similar manner, the social unblocking effect (pCS+ > pCS-) is present in both Sham and PLO lesioned rats.

Posterior lateral OFC lesions impair the blocking effect for partner no-reward. In PLO lesioned animals, but not in animals of the sham and the behavioral groups, we found that the social unblocking effect is impaired in trial 1. This could be due to 1) no increase in responding to the pCS+ and to pCS- or 2) a similar increase in responding to pCS+ and pCS-. After learning

that the aCS+ fully predicts a reward, we expected that the pCS- would be blocked from acquiring value, as it is coupled to cues that predict no additional reward delivered to either the actor or the partner (Own-reward; aCS+, pCS- and No-reward; aCS-, pCS-). We found that the absence of the unblocking effect is mainly due to a higher score on the pCS- cue on trial 1, indicating a decreased blocking effect (figure 5b). To see how strong this evidence is, we performed an independent t-test on the pCS- response with sham and lesion as between-subject factors. We found moderate evidence that indeed the response to the blocked cue is higher for the Lesion group than the Sham group in trial 1 (T (20) = 2.070, p = 0.026, Cohen's D = -0.883, BF_{+0} = 3.13, with median posterior delta = -0.678, 95% CI = -1.544 - -0.074; Figure 5c). We then found similar moderate evidence that the blocked cue response was higher for the lesion group compared to the Behavioral data as well (T(16) = 2.278, p = 0.018, Cohen's D = 1.080, BF_{+0} = 4.052, with median posterior delta = 0.805, 95% CI = 0.093 - 1.816; Figure 5c). These results provide evidence that blocking of partner no-reward related cues was impaired in trial 1 in the Lesion group, but not Sham and the Behavioral Data group. To see if this blocking impairment for the pCS- is not due to a general enhancement of responding to the NR compound (aCS-/pCS-) during social learning, we zoomed in to get a detailed look at responding to the NR second to second responses in the food cup. This was not based on the difference score but included the baseline and cue on period (see Figure 5a; 2 seconds per bin), averaged over the 4 days of social learning. We performed a repeated-measures ANOVA with the time spent in the food cup in the NR condition as DV and two factors as IVs: time bin as a within-subjects factor (20 bins of 2 seconds) and group (sham, lesion) as the between-subjects factor.

We found a interaction effect between time bin and Group F(1.725, 34.507) = 4.447, p = 0.024, η_p^2 = 0.182). Post-hoc comparisons showed that the time spent in the food cup does not differ between sham and lesion group for baseline bins 1-5 (all U (N-Lesion = 11, N-Sham = 11) = 36 or lower, all $p \geq 0.058$). At cue onset however, the NR response becomes significantly higher in the lesion over the sham group on bin 1 (U (N-Lesion = 11, N-Sham = 11) = 29, p = 0.020) and bin 2 (U (N-Lesion = 11, N-Sham = 11) = 29, p = 0.020) and nearly significant at bin 3 (U (N-Lesion = 11, N-Sham = 11) = 37, p = 0.66; Figure 5a). These results suggest that the observed lack of blocking (i.e., responding to the pCS- cue) can therefore be explained due to a very specific higher response rate to the NR cue during social conditioning in the lesion group compared to the sham group, potentially due to a spillover over of value from the OR or BR cue that is attached to the pCS- cue. We conclude from these results that PLO integrity is required to implement "social" blocking of a cue that predicts no additional partner reward.

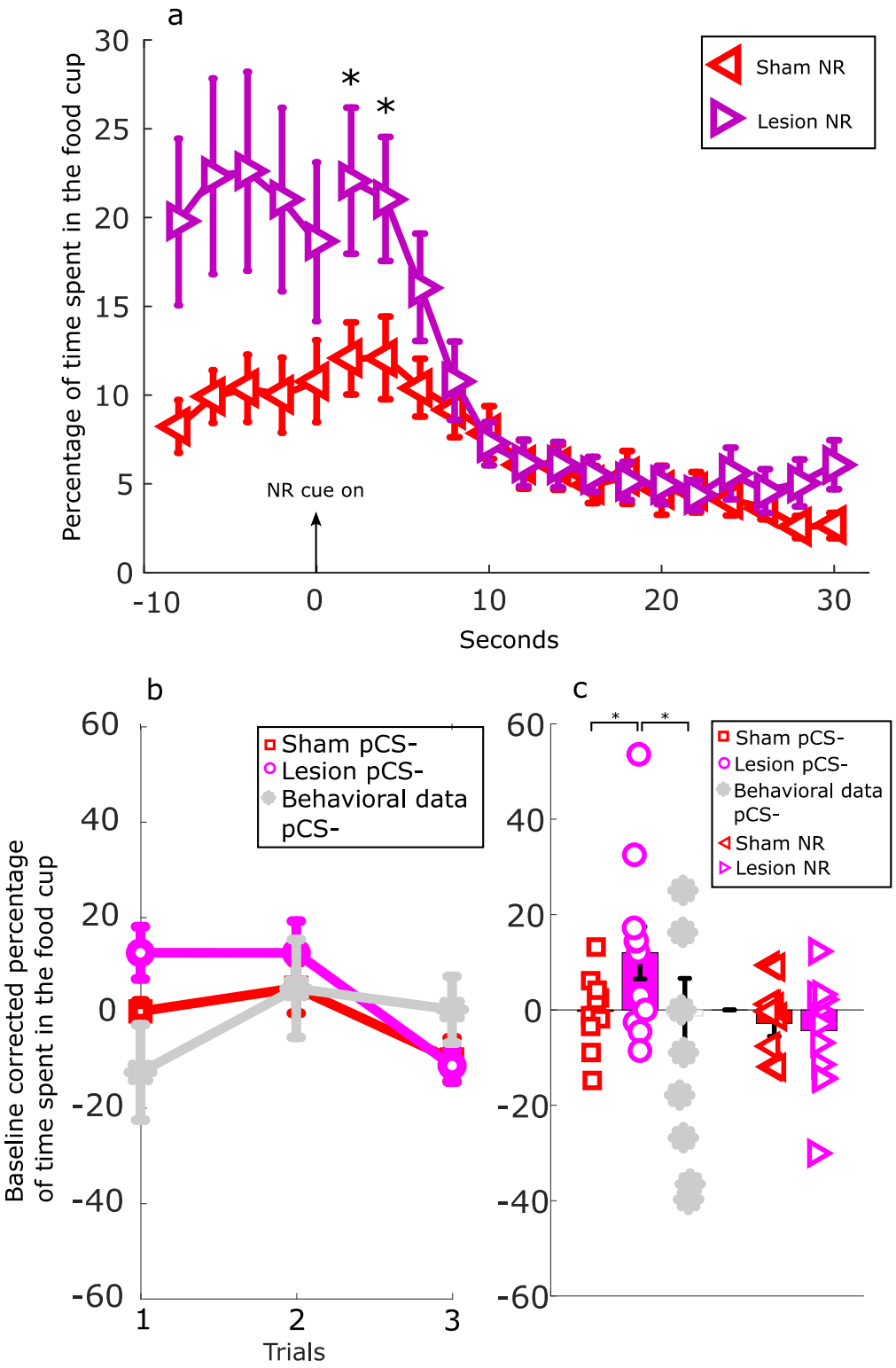


Figure 5. Social blocking. (a) Percentage of time spent in the food cup averaged over 4 days of NR (aCS-/pCS-). Plotted is the baseline period of 10 seconds and and the cue on period of 30 seconds. The time x-axes is defined by 2 seconds bins. The NR compound cue goes on at 10 seconds.(b) Baseline corrected time spent in the food cup during the first 3 trials of the blocked pCS- cue (including a baseline of 30 seconds). (c) Bar plot shows baseline corrected percentage of time spent in the food cup during trial 1 of the pCS- in Sham, Lesion and Behavioral data group. On the right the baseline corrected time spent in food cup for the NR response on the last day of compound training is shown. Scatter points display indivividual rats.

No evidence for social unblocking in the partner rat cue that predicts both Mutual and actor reward only in the Sham group. After discrimination learning of 19 days, partner rats went through 4 days of compound training with the conditions BR (pCS+, aCS+, rewarded), OR (pCS-, aCS+, no reward) and NR (pCS-, aCS-, no reward). For partners rats, the average responding during the last 4 days of discrimination learning was larger for pCS+ over pCS- in the Sham group ($T(10) = 19.337$, $p < 0.001$; Figure S1a) and Lesion group ($T(8) = 6.815$, $p < 0.001$; Figure S1c). For the analyses, we calculated the mean time spent in the food cup over 4 days of responding and performed a repeated-measures ANOVA to compare responding on the mean BR, OR (partner no-reward, actor reward) and NR per group. To see if there were any significant differences between groups, we directly compared the percentage of time spent in the food cup on the BR, OR and NR. We performed a repeated-measures ANOVA with Conditions (BR, OR, and NR) and as a between-subject factor group (Sham and Lesion). We find no interaction effect ($F(1.284, 25.671) = 0.546$, $p = 0.509$, $\eta_p^2 = 0.003$) indicating that there were no significant differences between the Sham and lesion Group partner rats during social learning. Both groups of partner thus showed similar responses during the social phase. In the Sham group we find an effect of

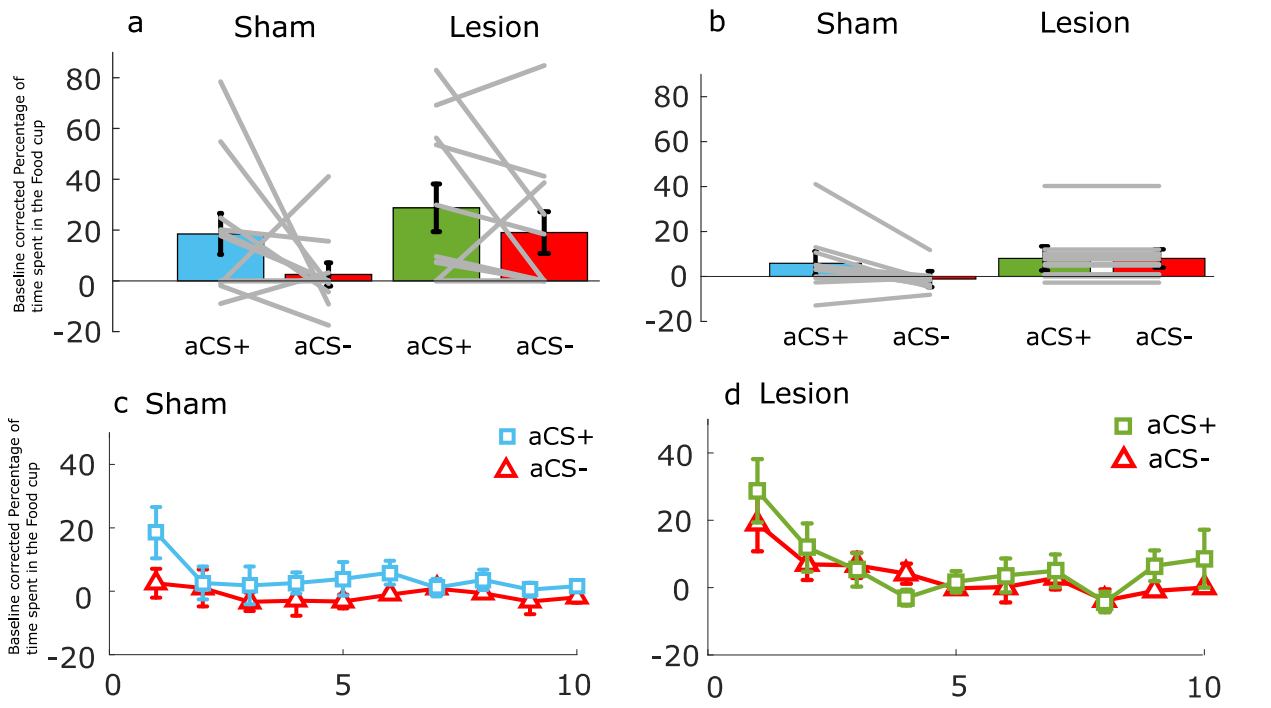


Figure 6. Food cup response of the partner rat during compound phase and probe trials. (a) Percentage time spent in the food cup of the Sham and lesion group for the aCS+ and aCS- cues on trial 1. (b) Percentage time spent in the food cup of the Sham and lesion group for the aCS+ and aCS- cues average over the first 6 trials. (c) Food cup response of the partner of the sham group during the probe trials. (d) Food cup response of the partner of the Lesion group during probe trials. All percentage of time spent in the food cup was baseline corrected.

condition (F (1.212, 12.116) = 81.243, p < 0.001, η_p^2 = 0.890). The percentage of time spent in the food cup in the BR was, as expected, significantly higher than OR responding (mean difference = 29.151, se = 3.352, p < 0.001). The partner rats spent more time in the food cup in the BR condition than the NR (BR: mean difference = 38.192, se = 3.979, p < 0.001). Importantly they also spent more time in the food cup in the OR (partner no-reward, actor reward) than the NR (mean difference = 9.041, se = 1.534, p < 0.001). In the lesion group we similarly find an effect of condition (F(1.326, 27.839) = 104.327, p < 0.001, η_p^2 = 0.832) the percentage of time spent food cup. The percentage of time spent food cup in the BR was, as expected, significantly higher than OR responding (26.508, se = 2.871, p < 0.001). Percentage of time spent food cup in the BR was higher than the NR (BR: mean difference = 37.763, se = 3.324, p < 0.001). Here, similar as in the sham group we find a higher Percentage of time spent food cup in the in the OR (partner no-reward, actor reward) than the NR (neither reward; mean difference = 11.255, se = 1.526, p < 0.001). We can thus conclude that partner rats in both groups are influenced by unequal reward outcomes, similarly, as it was found in van Gurp et al. (2020), leading the partner rat to lose the learned pCS- inhibition and respond more to food cup, even in the absence of a partner reward on OR trials. We previously found that the partner rats do not show unblocking to the aCS+ cue (that predicts actor reward in both the BR and OR conditions). To see whether partner rats unblocking is absent in this experiment as well and to investigate whether there are any significant differences between Sham and Lesion group, we ran a two factor repeated measures ANOVA with Condition (pCS+, aCS+, aCS- and pCS-) and trial (trial 1 to 6) as factors and Group (Sham, Lesion) as a between-subjects factor (figure 6b, c, d). The time spent in the food trough for the 10 seconds after cue onset was the dependent variable. We found no interaction effect of Condition * group F (2.312, 46.238 = 0.762, p = 0.490, η_p^2 = 0.003) and no interaction effect of condition * trial * group F (6.119, 122.381 = 0.489, p = 0.819, η_p^2 = 0.006). To check in more detail this result this we again performed a T-test comparing the percentage of time spent in the food cup during aCS+ with the aCS- in both Sham and lesion partner groups. In the Sham partner group we found no significant difference in the percentage of time spent in the food cup for the aCS+ (M = 18.485, SD = 26.783) over pCS- (M = 2.521, SD = 15.101) in trial 1 (T (10) = 1.549, p = 0.076, BF_{+0} = 1.372, with median posterior delta = 0.411, 95% CI = 0.035 0.92; Figure 6a). We furthermore found no significant difference in the lesion group in the percentage of time spent in the food cup for the aCS+ (M = 28.752, SD = 31.155) over pCS- (M = 19.108, SD = 27.252) in trial 1 (T (10) = 1.173, p = 0.134, BF_{+0} = 0.883, with median posterior delta = 0.338, 95% CI = 0.023 0.881; Figure 6a). We conclude from these results that there is no evidence that cues that predict

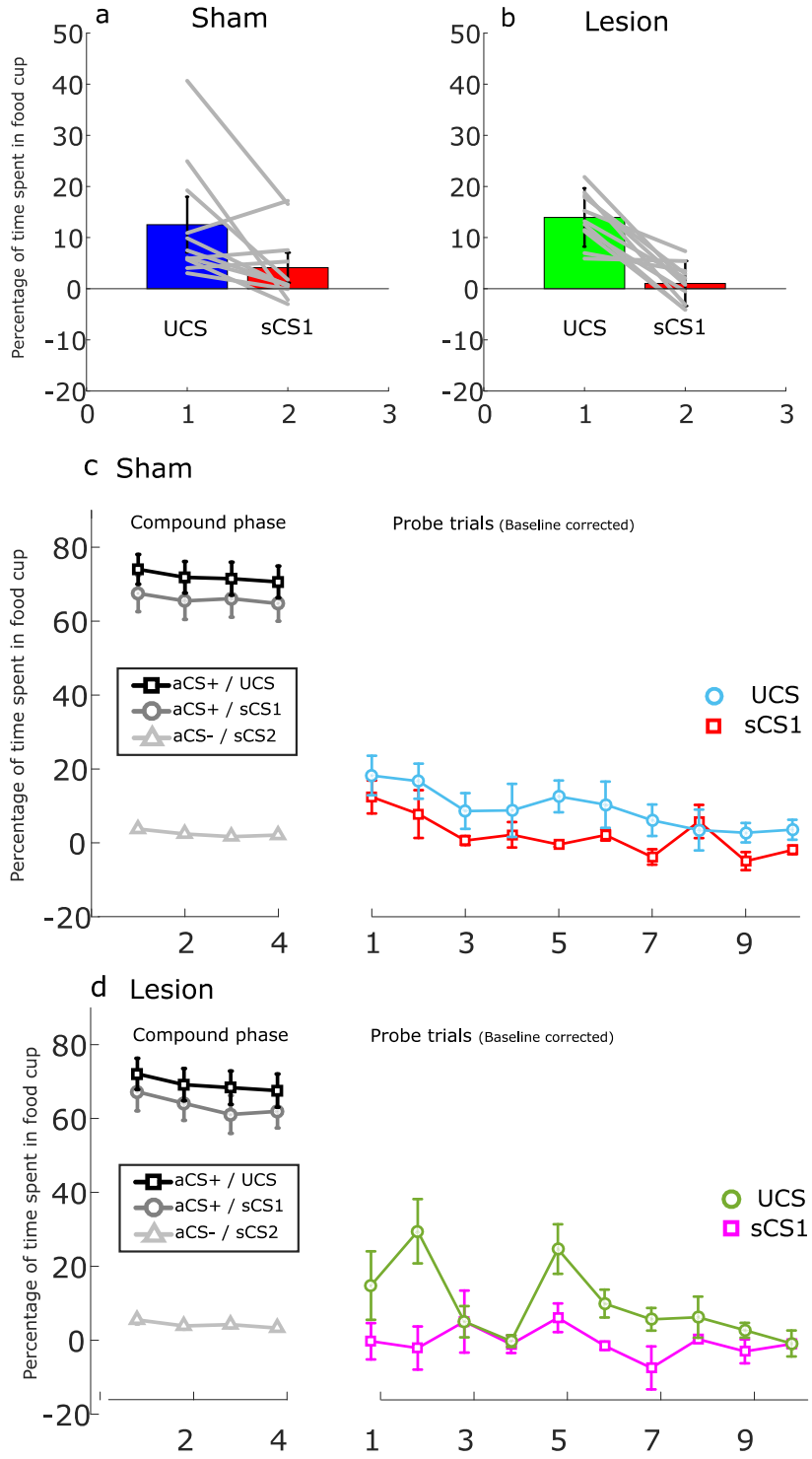


Figure 7. Food cup response of the actor rat during appetive unblocking. (a) Percentage of time spent in the food cup of the Sham group during the probe trials, for the UCS and sCS1 cues averaged over the first 6 trials. Single lines indicates individual rats. (b) Percentage of time spent in the food cup of the Lesion group during the probe trials, for the UCS and sCS1 cues averaged over the first 6 trials. (c) Sham group's percentage oftime spent in the food cup of the pCS+ and pCS- during the appetitive compound phase and the percentage of time spent in the food cup during the appetitive probe trials. (d) Lesion group's percentage of time spent in the food cup of the UCS and sCS1 during the appetitive compound phase and the percentage of time spent in the food cup during the appetitive probe trials. Probe trials, but not compound phase percentage of time spent in the food cup was baseline corrected.

additional actor reward (in BR and OR trials) cause social unblocking (aCS+ > aCS-) in either the lOFC lesion or Sham partner groups.

PLO lesions leave appetitive unblocking intact. After social unblocking, the actor rats went through a non-social appetitive-unblocking control phase. Here, we expected to replicate findings by McDannald et al. (2011) who found that appetitive unblocking is unaffected by lateral Orbitofrontal cortex lesion. Briefly, we indeed replicated appetitive unblocking in both groups, and no statistical test of interactions between responses in the upshift or probe phase and group membership was significant. In both groups we found that, during the appetitive compound phase (see figure 7c, d) rats spent more time in the food cup for a cue that predicted Upshift (+ 3 pellets; aCS+, upCS) over a cue that predicted Same reward positive (no added pellets; aCS+, sCS1). Importantly, during the probe trials, the baseline corrected time spent in the food cup (over a 6 trial period) in the upshift cue (upCS; + 3 pellets) was higher than the cue that predicted no added reward (sCS1; + 0 pellets) in both Sham (see figure 7a, c) and the Lesion group (see figure 7b, d). These results (for statistics see: Table S2) replicate earlier findings that the lateral OFC is not necessary for appetitive unblocking (McDannald et al., 2011). We furthermore find that an upshift from 3 to 6 pellets delivery also induces appetitive unblocking, similarly as has been found with a 1 to 3 pellets upshift (Holland, 1984; Keiflin et al., 2019; M. A. McDannald et al., 2011) and can therefore be used for the investigating of upshift learning during unblocking.

3.4 Discussion.

PLO lesions leave social unblocking intact. We find here that bilateral posterior-lateral Orbitofrontal cortex lesioned animals do not show impairments on both social and appetitive unblocking. This result replicates findings of McDannald et al. (2011) who found that appetitive unblocking is unaffected by lateral orbitofrontal cortex lesions and extends it by showing that a more specific posterior part of the lateral orbitofrontal cortex (PLO) that has been shown to code upshifts in appetitive self-rewards (Lopatina et al., 2015), does neither affect social nor appetitive unblocking. It thus seems that PLO, an area shown to be crucial in supporting model-based identity unblocking, is not required to support other-reward related social unblocking driven by social-identity signals. We hypothesized that the process that links social-identity signals related to mutual reward to novel cues introduced in a compound learning phase would depend on PLO integrity. Such identity signals could be useful to support a "self-reward" from an "other-reward" value distinction. Vicarious other reward signals are thought to consist of visually observing the partner rats' approach the food cup, observes the pellet being delivered to the partner and afterwards hearing the comsummatory behavior (van Gurp et al., 2020) and/or the possible the

transfer of olfactory food-related information on the breath of the partner rat to actor rat after comsummatory behavior (Galef, Iliffe, & Whiskin, 1994) and/or ultrasonic vocalisations associated with appetitive states. We thus conclude here that these social-identity reward signals are most likely not processed in the PLO. Alternatively, this kind of model-based social-identity reward information could be processed in the ACC (Carrillo et al., 2019; Schneider, Sciarillo, Nudelman, Cheer, & Roesch, 2020) and ACC activity could contribute to vicarious learning of the association between such social-identity reward signals that and the cues that predict them as has been found in monkeys (Chang et al., 2013). An even simpler model would rely on model-free integration of self- and other reward as a driver of unblocking that would not require OFC involvement at all, consistent with the common currency hypothesis (Ruff & Fehr, 2014). The BLA could furthermore be involved, as socially transmitted food preferences are impaired when this structure is temporarily lesioned (Wang, Fontanini, & Katz, 2006) and therefore it could support the processing of the transfer of olfactory food-related information on the breath of the partner rat to actor rat. The processing of other social reward signals such as positive 50 kHz and negative 22 kHz ultrasonic vocalisations (Kashtelyan et al., 2014) could also drive the social unblocking as a proxy for social value. These signals are however most likely not processed in the PLO, at least not directly, as these vocalisations are more likely to activate neuronal populations in the VTA (Gunaydin et al., 2014) and, via the NAcc, drive the formation of social reward learning as has been found during appetitive unblocking (Keiflin et al., 2019).

Varying social unblocking effect strength. It is unclear why the Sham group only shows the social unblocking effect on trial 1 but less pronounced during the first 6 trials as observed in the behavioral control group. Firstly, it is possible that the fast extinction of social unblocking in the sham group is a potential false positive, and the real social unblocking effect for auditory cues lasts longer, such as observed in the data of the Behavioral group. The second explanation is that we expect that the social unblocking effect in the control group is smaller due to controlling for secondary reinforcement of the added pellet dispenser during discrimination learning (as found in van Gurp et al. (2020)) and that, therefore the PLO lesion rats should show smaller levels of social unblocking as the pellet dispenser was also added during discrimination learning in the lesion group. The behavioral data control group, however, consisted of both rats in which secondary reinforcement was controlled for (van Gurp et al. (2020); N = 4; Social-Only group) and where it was not controlled for (N=4; Social-appetitive group). We, therefore, have a too small comparison group to weigh evidence to support a conclusion stating that social unblocking is too strong in the lesion group or too small in the Sham group.

PLO lesions impair social blocking of cues associated with partner no-reward. Besides the absence of involvement of the PLO in social and appetitive unblocking, we found evidence that PLO-lesioned rats show enhanced responding to partners pCS- cue that predicted both Own-reward (OR; aCS+/pCS-) and No-Reward (NR; aCS+/pCS-) in the first trial of the probe trials. Blocking theory predicts that, because the aCS+ and aCS- have been fully learned, learning about the pCS- should remain blocked because, in both the OR and NR compounds, no change in total reward value is induced. During this task, however, the task contingencies are more detailed and involve different interleaved probabilities of cue- reward outcomes. Although the aCS+ has been fully learned, the pCS- is coupled to reward in only 50 percent of the trials. Although the value of the compounds do not change, animals have to keep track of the two contexts in which the pCS- appears. Taking the reward context into account, the predicted value associated with the pCS- cue could be changed when an NR (aCS-/pCS-) follows OR (aCS+/pCS-), leading to a negative prediction error for the pCS- (going from reward to no-reward) and vice versa when OR (aCS+/pCS-) follows NR (aCS-/pCS-) leading to a positive prediction error. In normal blocking/unblocking learning, this kind of prediction should, however, not occur as any update to the pCS- cue would be blocked due to comparison with the learned values of aCS+ and aCS-. We find, however, that lesioned rats do show enhanced responding to the NR (aCS-/pCS-) cue in the cue on period. We interpret this effect in line with Walton et al. (2010) work on the OFC's role in credit assignment and the spread of effect, raising the possibility that the value associated to pCS- because of the spillover of positive social reinforcement, from previous trials or within the compound by failing to account for the complete reward prediction carried by aCS+. Credit assignment is defined as the ability to assign reward outcomes to the right previous choices (Walton et al., 2010). Crucially, they refer to the spread of effect phenomenon first observed by (Thorndike, 1933), that is, rewards do not only reinforce the choices that lead to them but also other choices which happened in the recent past or closely following the choice. Walton et al. (2010) remarkably found that, in monkeys with a lesion to the lateral orbitofrontal cortex, the choices for a specific stimulus-outcome, were strongly influenced by its most recent choices for another stimulus-outcome and not like normal animals, by the appropriate weighing of the specific contingent stimulus-outcome. It is thus possible that the impaired blocking effect observed in our lesion group is caused by an enhanced responding to the NR because of its recent history of reinforcement to the OR (aCS+/pCS-) positive value, impairing the normally observed blocking effect. The responding to the blocked cues observed here is normally observed in the control condition during blocking in both the social and non-social domains. In humans, when a cue Y is added to a cue B that predict no reward and then when this compound is reinforced, it leads the Y cue to acquire

conditioned responses during extinction for both social and non-social cues (Seid-Fatemi & Tobler, 2015; Philippe N Tobler, O'doherty, Dolan, & Schultz, 2006). In the blocking experiment of Tobler (2006), the level of learning of the Y cue was correlated to the amount of blocking a participant showed and that this learning was related to lateral orbitofrontal cortex activity. In our experiment, it is possible that incorrect credit assignment in lesioned animals leads to a spread of associative value to the pCS- cue (50% associated with own reward), and thus (partial) unblocking, similar to the Y cue that becomes unblocked in the blocking experiment of Tobler (2006). This blocking or credit assignment impairment could harm social foraging behavior, when actor animals expend energy approaching cues that were co-present with reward, but do not predict it.

To investigate the precise role for the PLO in social vicarious learning, a task would be needed in which animals update the cues predicting social reward offline, for example by a procedure that devalues the cue that predicts reward to others. This could be achieved by satiating the partner and/or actor rat before learning. The decreased motivation of the actor rat itself to participate in social interaction and/or reward learning, could likely be captured by diminished processing of the social reward-identity signal, as the social outcomes or the actor's interest in the social outcome was devalued, leading to decreased social unblocking. We hypothesise that lesions of the lateral OFC would impair such social downshift valuations based on offline cue updating that requires an internal model of self- and other-related social motivation representations.

3.5 Supplemental Materials.

Figure S1

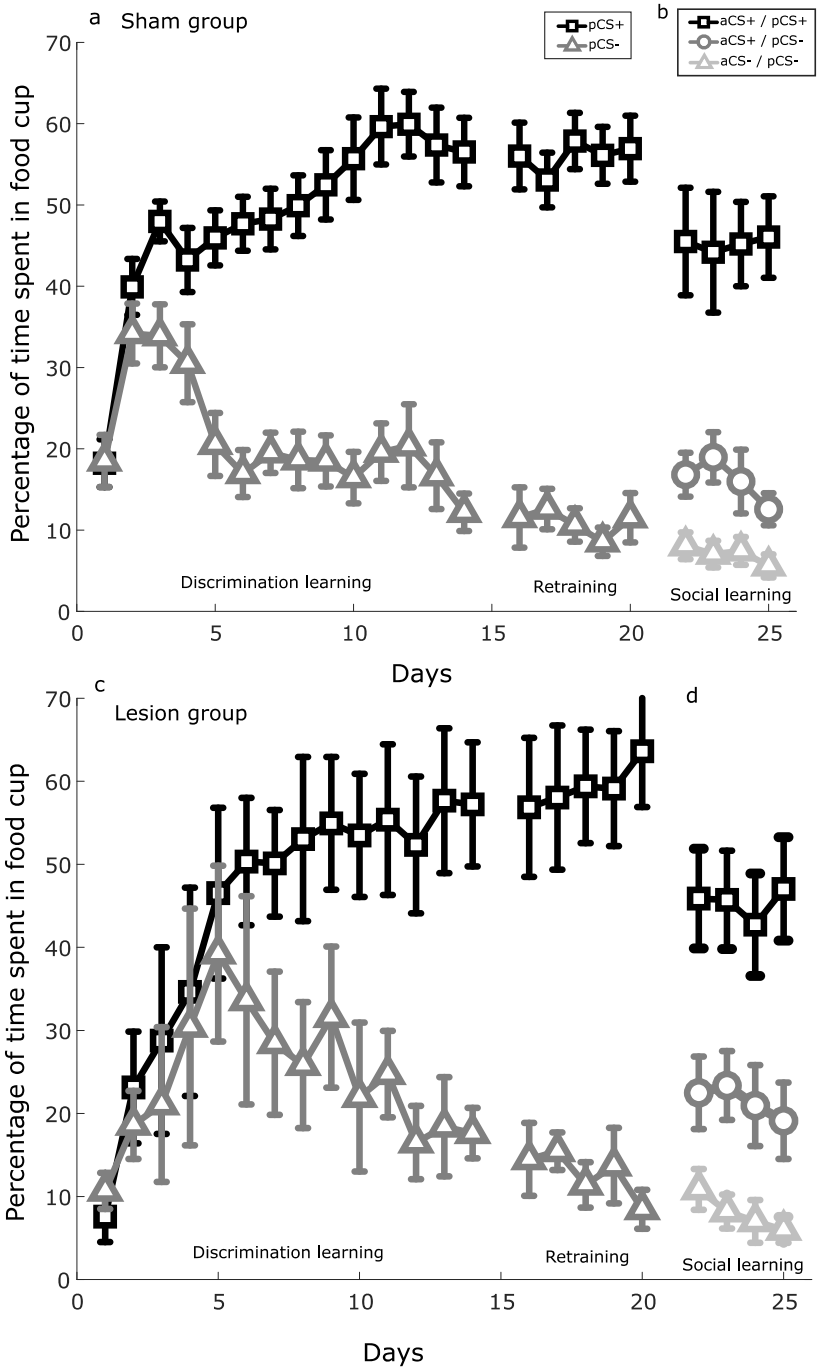


Figure S1. Conditioning per experimental phase of the partner rats. *Sham partner group.* (a) Percentage of time spent in food cup for discrimination learning between pCS+ and pCS- over days. (b) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days. *Lesion partner group.* (c) Percentage of time spent in food cup for discrimination learning between pCS+ and pCS- over days. (d) Percentage of time spent in food cup for the compounds BR (aCS+, pCS+), OR (aCS+, pCS-) and NR (aCS-,pCS-) over days.

Table S1

Table S1. Social learning Phase and general probe trials statistics			
Social learning Phase **Mixed repeated measures ANOVA**	**F value** **(Interaction Condition * Group)**	**P value**	**η_p^2**
Conditions (BR, OR, and NR) and Group (Sham, Lesion)	**F (1.069, 21.379) = 0.015**	**p = 0.916**	**η_p^2 < 0.001**
Sham: Repeated measures ANOVA	**F value** **(main effect)**	**P value**	**η_p^2**
conditions (BR, OR, and NR)	**F (1.108, 11.076) = 121.594**	**p < 0.001**	**η_p^2 = 0.924**
Post hoc comparisons for main effect	**Mean difference**	**Std error**	**P value**
BR | OR	**2.820**	**1.123**	**0.092**
BR | NR	**42.407**	**3.899**	**p < 0.001**
OR | NR	**39.587**	**3.360**	**p < 0.001**
Lesion: Repeated measures ANOVA	**F value** **(main effect)**	**P value**	**η_p^2**
conditions (BR, OR, and NR)	**F(1.103, 10.336) = 71.335**	**p < 0.001**	**η_p^2 = 0.877**
Post hoc comparisons for main effect	**Mean difference**	**Std error**	**P value**
BR | OR	**2.471**	**1.123**	**p = 0.018**
BR | NR	**41.552**	**4.770**	**p < 0.001**
OR | NR	**39.082**	**4.736**	**p < 0.001**
Social unblocking ***probe trials*** **Mixed**	**F value** **(main effect)**	**P value**	**η_p^2**

repeated measures ANOVA			
Group: Sham, Lesion, Behavioral group, Condition: aCS+, pCS+, pCS- and aCS- and trial: 1 to 6	**F value (main effect)**	**P value**	**η_p^2**
Condition * group	**F (4.407, 59.497) = 1.088**	**p = 0.373**	**η_p^2 = 0.075**
condition * trial * group	**F (9.644, 260.393) = 1.084**	**p = 0.368**	**η_p^2 = 0.074**

Table S2

Table S2: Appetitive unblocking statistics			
<u>Compound phase</u> ***SHAM: repeated measures ANOVA***	***F value (Interaction Condition * Group)***	***P value***	***η_p^2***
Conditions: Upshift (+ 3 pellets; aCS+, upCS), Same reward positive (no added pellets; aCS+, sCS1) and Same reward negative (no added pellets; aCS-, sCS2) and days (day 1 to 4)	**F(1.135, 11.346) = 182.693**	**p < 0.001**	**η_p^2 = 0.917**
Post hoc comparisons for main effect	***Mean difference***	***Std error***	***P value***
Upshift | Same reward positive	**6.432**	**3.126**	**p < 0.001**
Lesion: repeated measures ANOVA	***F value (Interaction Condition * Group)***	***P value***	***η_p^2***

Conditions: Upshift (+ 3 pellets; aCS+, upCS), Same reward positive (no added pellets; aCS+, sCS1) and Same reward negative (no added pellets; aCS-, sCS2) and days (day 1 to 4)	**F(1.032, 10.322) = 65.333**	**p < 0.001**	**η_p^2 = 0.804**
Post hoc comparisons for main effect	***Mean difference***	***Std error***	***P value***
Upshift | Same reward positive	**5.093**	**5.391**	**p < 0.001**
Unblocking phase ***Mixed repeated measures ANOVA***	***F value*** ***(Interaction Condition * Group/ condition * trial * group)***	***P value***	***η_p^2***
Condition (aCS+, upCS+, sCS- and aCS-), trial (trial 1 to 6) and Group (Sham, Lesion	**F (2.541, 50.828) = 0.624 / F (6.297, 125.946) = 1.089**	**p = 0.577 / p = 0.373**	**η_p^2 < 0.001 / η_p^2 = 0.373**
Sham: repeated measures ANOVA	***F value*** ***(Main effect)***	***P value***	***η_p^2***
Conditions: upCS and sCS1 and trial (trial 1 to 6)	**F(1,10) = 6.908**	**p = 0.025**	**η_p^2 = 0.091**
Post hoc comparisons for main effect	***Mean difference***	***Std error***	***P value***
upCS | sCS1	**8.422**	**3.204**	**0.025**
Lesion: repeated measures ANOVA	***F value*** ***(Main effect)***	***P value***	***η_p^2***
Conditions: upCS and sCS1 and trial (trial 1 to 6)	**F(1,10) = 40.059**	**p < 0.001**	**η_p^2 = 0.099**
Post hoc comparisons for main effect	***Mean difference***	***Std error***	***P value***

upCS | sCS1	**12.913**	**2.040**	**$p < 0.001$**

4. Distinct profiles of 50 kHz vocalizations differentiate between social versus non-social reward approach and consumption.

Abstract

Social animals tend to possess an elaborate vocal communication repertoire, and rats are no exception. Rats utilize ultrasonic vocalizations (USVs) to relay information about a wide range of socially relevant cues, as well as information regarding the valence of the behavior and/or surrounding environment. Both quantitative and qualitative acoustic properties of these USVs are thought to communicate context-specific information to conspecifics. Rat USVs have been broadly categorized into 22 and 50 kHz call categories, which can be further classified into subtypes based on their sonographic features. Recent research indicates that the 50 kHz calls and their various subtype profiles may be related to the processing of social and nonsocial rewards. However, only a handful of studies have investigated USV elicitation in the context of both social and nonsocial rewards. Here, we employ a novel behavioral paradigm, the social-sucrose preference test, that allowed us to measure rats' vocal responses to both nonsocial (i.e., 2, 5, and 10% sucrose) and social reward (interact with a Juvenile rat), presented concurrently. We analysed adult male Long-Evans rats' vocal responses towards social and nonsocial rewards, with a specific focus on 50 kHz calls and their 14 subtypes. We demonstrate that rats' preferences and their vocal responses towards a social reward were both influenced by the concentration of the nonsocial reward in the maze. Furthermore, we report that the proportion of individual subtypes of 50 kHz calls, as well as the total USV counts, showed variation across different types of rewards as well. Our findings provide a thorough overview of rat vocal responses towards nonsocial and social rewards, and are a clear depiction of the variability in the rat vocalization repertoire, establishing the role of call subtypes as key players driving context-specific vocal responses of rats.

4.1 Introduction.

Rats are social animals (Whishaw & Kolb, 2009) that form relatively large and tightly organized groups. As nocturnal animals, many rodent species rely on complex vocalizations for

** Chapter 4 is submitted and under review in Frontiers in Behavioral Neuroscience. Seidisarouei M, van Gurp S, Pranic NM, Noguer Calabús I, Kalenscher T, van Wingerden M. 2020. Distinct profiles of 50 kHz vocalizations differentiate between social versus non-social reward approach and consumption. Frontiers in Behavioral Neuroscience, section Motivation and Reward.*

communication and social coordination. The extent of their vocalization vocabulary depends on their social structure and inter-individual interactions (for a review, see Brudzynski, 2014).

Among rodents, rats, in particular, have developed an elaborate system of ultrasonic communication which has been suggested to have adaptive significance by signalling socially relevant information: ultrasonic vocalizations (USVs) emitted by rats have been implied to play a role in warning conspecifics (Brudzynski, 2013; Litvin, Blanchard, & Blanchard, 2007), as well as acting as indices of rats' affective states (Brudzynski, 2013; Brian Knutson et al., 2002) and social motivation (Mulvihill & Brudzynski, 2018b). Additionally, Himmler et al. (2014) have demonstrated the function of rat USVs in facilitating and maintaining play behavior, pointing to their social communicative value. Thus, it has been suggested that the wide range of calls emitted by rats serve a multitude of context-dependent functions.

The USVs emitted by pups, adolescent and adult rats can be divided to three major sub-groups: (i) 22- kilohertz (kHz) alarm calls (Litvin et al., 2007) produced in response to an aversive circumstance (Wöhr & Schwarting, 2013), (ii) 50-kHz USVs that signal appetitive and rewarding states (Panksepp & Burgdorf, 2000) and (iii) 40 kHz vocalizations produced by socially isolated pups (Wöhr, Houx, Schwarting, & Spruijt, 2008). The acoustic features of the 50 kHz calls differ substantially from 22 kHz USVs (Brudzynski & Holland, 2005; Brudzynski & Pniak, 2002; Thompson, Leonard, & Brudzynski, 2006), allowing distinct and clear-cut classifications. Specifically, 50 kHz USVs have a concise call duration between 30 – 40 ms, a bandwidth of 5-7 kHz, and a peak frequency remaining within 45-55 kHz, although the calls can reach 70 kHz or higher.

The 22 and 50 kHz call categories emitted by rats thus represent general qualitative information regarding the valence of the environment or behavior, but these call categories can be further organized into subtypes of vocalizations (Brudzynski, 2015; Himmler et al., 2014; Wright et al., 2010) that differ in sonographic features. For instance, 50 kHz USVs can be classified into Flat and frequency-modulated (FM) subtypes based on the bandwidth of frequencies they extend over in spectrograms (Burgdorf & Panksepp, 2006; Wöhr et al., 2008). Several lines of evidence demonstrate that rats emit Flat- and FM-50 kHz USVs in different situations, suggesting that these subgroups of 50 kHz USVs may have distinct and disparate communicative roles of behavioral significance. Flat calls, for instance, have been suggested to be involved in (initiating) social contact (Burgdorf, Panksepp, and Moskal, 2011) and social coordination (Wöhr & Schwarting, 2008). FM 50 kHz USVs, on the other hand, are more commonly emitted during rewarding situations or high positive emotional arousal (Burgdorf et al., 2011a). The FM subgroup of 50 kHz USVs has been

further grouped into subtypes based on the extent of their frequency modulation and the shape they assume in the spectrogram (Brudzynski & Zeskind, 2018). In the most comprehensive classification, the 50 kHz USVs were categorized into 14 distinct subtypes (Wright, Gourdon, and Clarke, 2010). This categorization, however, is not one without controversy. Coffey, Marx, and Neumaier (2019), for instance, have recently utilized the DeepSqueak software to classify USVs using unsupervised machine learning techniques into 18 separate clusters instead of 14 subtypes. In addition, the behavioral relevance of these various call subtypes remains largely unknown.

Because of their association with appetitive situations, 50kHz calls could potentially also be utilized in quantifying the value that individual rats attribute to a reward (Garcia, McCowan, & Cain, 2015) as well as to the expectation of a reward (Binkley, Webber, Powers, & Cromwell, 2014). Calls emitted in the presence of nonsocial and social rewards have been investigated thoroughly in the literature. Cues for nutritional reward have been shown to elicit 50 kHz responses from rats (Brenes & Schwarting, 2014), and a preference for sweet pellets over regular pellets is associated with an increase in the frequency of 50 kHz vocalizations (Mateus-Pinheiro et al., 2014). Nevertheless, Schwarting, Jegan, and Wöhr (2007) found no difference between the 50 kHz calls produced by food-deprived animals and the ones exposed to ad-libitum feeding, when they were alone in the home cage. In another intricate design, Browning et al. (2011) have demonstrated that rats trained for cocaine and sucrose self-administration showed more 50 kHz calls during the reward self-administration and reinstatement phase (after a period of extinction training), compared to naïve controls who were not rewarded.

Juvenile, adolescent, and adult rats have been shown to emit 50 kHz calls during interactions with their conspecifics, such as rough and tumble play (Brian Knutson, Burgdorf, & Panksepp, 1999) and mating (White, Cagiano, Moises, & Barfield, 1990). Female rats also produce 50 kHz calls when encountering a social partner (Börner, Hjemdahl, Götz, & Brown, 2016). The calls emitted by adult rats can thus give clues about their social behavior (but see, Manduca et al., 2014). It has been shown that rats emit more 50 kHz calls when exposed to another conspecific (Brudzynski & Pniak, 2002), and display a preference for rats producing more 50 kHz calls (Panksepp, Gordon, & Burgdorf, 2002). In contrast, rats selectively bred to emit lower rates of 50 kHz calls spent less time with conspecifics in a social interaction test than the randomly bred line (Burgdorf et al., 2009). Similarly, playful experiences are significantly less frequent in pairs of devocalized rats than in their vocalizing counterparts, emphasizing the role of these 50 kHz calls in maintaining play behavior (Himmler et al., 2014).

Łopuch and Popik (2011) and Kalenscher et al. (2020) have also argued that the cooperative behavior of rats positively correlates with the 50 kHz vocalizations they produce, as 50 kHz USVs may act as social vicarious reward signals (Hernandez-Lallement et al., 2016a; Kalenscher et al., 2020; van Gurp et al., 2020). Neural processing of USVs has been implicated in the amygdala, with opposing coding schemes for 22 vs. 50 kHz USVs (Parsana, Li, & Brown, 2012), and indeed, lesions of the BLA impair the social approach that is usually observed to 50 kHz USV playback (Schönfeld, Zech, Schäble, Wöhr, & Kalenscher, 2020; Seffer, Schwarting, & Wöhr, 2014; Wöhr & Schwarting, 2007).

In short, both qualitative and quantitative differences in 50 kHz USV production have been found across a range of social and nonsocial rewarding situations. Only a handful of studies in the literature, however, have investigated USV production in the context of concurrent social and nonsocial rewards. Utilizing selective breeding procedures, (Burgdorf et al., 2009) have demonstrated that rats bred to emit higher rates of 50 kHz calls were more likely to prefer a sucrose solution to tap water than randomly bred rats. Willey and Spear (2013) analysed the calls and approach behavior towards both food-related and social stimuli in rats exposed to varying degrees of social deprivation. The time animals spent investigating the social stimulus within the apparatus positively correlated with the frequency-modulated (FM) calls they emitted. However, these authors did not find a relationship between animals' responses to food stimuli and their USV production. In a novel design, Mulvihill and Brudzynski (2018b) analysed the USVs produced by male rats separately allowed to freely explore a female, a littermate, as well as two nonsocial conditions, namely Fruit Loop rewards and 2% ethanol solution. Their results indicated that out of the four groups, only rats exposed to a cycling female produced a higher proportion of calls than the baseline. Mulvihill and Brudzynski (2018b) also demonstrate significant differences between the types of calls made in nonsocial versus social conditions. Specifically, rats exposed to nonsocial stimuli produced more flat calls than non-trill FM calls, whereas the non-trill FM subtype dominated the 50 kHz calls in the social contexts.

Thus, in summary, there is growing evidence that 50 kHz USVs, and the 50 kHz subtypes, are related to the subjective experience of social vs. nonsocial rewards, which could be related to reward processing traits (such as sucrose preferences), to individual communicative traits, or a combination of these factors. If there indeed is a structure to the type of vocalizations emitted in social and nonsocial situations, akin to a selective "vocabulary" for different behavioral contexts, it should be possible to distinguish these contexts when presented in direct competition, based on the vocalization patterns that are recorded.

To study this question, we employed a novel behavioral paradigm, the social-sucrose preference test. It is conducted on an XCST (X-shape chambered sociability test) maze. The XCST maze is a modified version of a radial arm maze previously utilized by Schönfeld et al. (2020) that can be used to contrast behavioral responses to both a social reward (Juvenile conspecific in an open-bar sociability cage) and varying levels of nonsocial reward (sucrose solutions) in different arms of the apparatus while recording the USVs emitted by the animals. Thus, we systematically investigated how the occurrence of the 14 subtypes of rat USVs was related to rats' choice behavior in the trade-off between social and nonsocial rewards.

4.2 Methods.

Subjects. The experiment was conducted according to the European Union Directive 2010/63/EU for animal experimentation and was approved by the local authority (Landesamt für Natur, Umwelt und Verbraucherschutz North-Rhine Westphalia, Germany). Fifteen male Long-Evans rats (*Charles River, Italy*) in total were obtained in a batch of 12 actor rats (PND 40, M_{weight} = 320 g, at the starting day of the experiment) and 3 Juvenile rats (PND 28, M_{weight} = 290 g, at the starting day of the *Social-Sucrose Preference Test (SSPT)*, serving as social stimulus/reward. Experimental rats were housed in groups of N=3 rats in standard Type IV Macrolon cages under a reversed 12:12 h light-dark cycle. The housing room was kept at a constant temperature of 22°C and a humidity of 60%. Throughout the experiment, all rats received standard laboratory rodent food, *ad libitum*, except for the *Sucrose Discrimination Test* (SDT) phase in which all actors were limited in their food intake (food per rat per day: 22g on weekdays and 25g on weekends).

Behavioral Task Setup. We used an eight-arm radial maze as previously adapted by Schönfeld et al. (2020), detached four arms to arrive at a cross/plus-maze setup (Fig 1.A). The maze consisted of a central platform (36 cm diameter) and four arms (14 cm wide and 60 cm long) that extended from the central platform in an octagon-shaped pattern. Each of the four arms was consistently associated with one single reward type: 3 arms with three different levels of a sucrose solution reward (see Fig 1.A) and one arm with a social stimulus. To circumvent any spatial bias, we divided our subjects into two groups (A and B, per group = 6) with a different allocation of reward positions for each group. Notably, during any test day in the experiment, only 2 out of 4 arms were open at a time to provide a head-to-head preference test between two rewards. On the arm of the maze assigned to the social reward, an unfamiliar Juvenile rat could be placed in a fixed cylindrical restrainer built from metal bars and compact plastic for its floor and ceiling (Height: 25.5 cm, Diameter: 17 cm, Ugo Basile Sociability Cage). The restrainer was fixed on the maze at the end of the Juvenile's arm, and the Juvenile could move around in this restrainer, and social contact

through the openings between the bars was possible. On the arms allocated to nonsocial reward (i.e., different sucrose concentrations 2%, 5%, and 10%), sucrose solution was provided to the actor animals in a cube plastic dish (8 x 8 cm) mounted at the end of each arm. Additionally, in order to facilitate spatial learning of the reward conditions in each arm over days, we included sandpapers in the entrance of each arm that the rats' whiskers touch when entering the arms. The sandpapers had varying grades (Group A: 2% [P800], 5% [P400], 10% [P150], and Juvenile [P1200], Group B: 2% [P150], 5% [P1200], 10% [P800], and Juvenile [P400]), following the findings of Guic-Robles, Valdivieso, & Guajardo (1989). These authors have demonstrated that rats' whiskers can discriminate between sandpapers with 200 grains/cm2 and 25 grains/cm2. To record the ultrasonic vocalizations (USVs), four ultrasonic microphones (Condenser Microphone CM16/CMPA, Avisoft Bioacoustics, Glienecke, Germany) were positioned via a microphone stand to approximately 20 cm on the right side of each reward dish and the restrainer (See Fig. 1A).

Social-Sucrose Preference Test design (SSPT). Behavioral testing on the SSPT included three phases (see Fig. 1B). In the first habituation phase, all four arms were open and unbaited, and each actor rat explored the maze for 10 minutes. This phase aimed to find out whether animals were inherently biased towards selecting one specific arm or sandpaper (see Fig. 2A-B). The second phase of training was the Sucrose Discrimination Test (SDT), which was implemented to verify that the actors could indeed distinguish among the three selected sucrose concentrations (2%, 5%, and 10%). Food deprived animals were tested on the SDT phase over nine days in three repetitions of three different conditions. In each condition, only two arms were open, and rats chose to allocate their time between rewards on the maze in the following order of conditions: *2% vs. 5%*, *2% vs. 10%*, and *5% vs. 10%*. Notably, each animal was tested in only one condition each day. Each test trial took 10 minutes; during this time, actors could move freely in the two open arms and drink up to 20 ml sucrose solution per plastic dish at the end of each arm. Both dishes were filled with fresh sucrose solution for each new trial/actor. After passing the SDT phase (Fig. 2C), the experiment was continued to the SSPT phase. In this phase, over each trial with a duration of 10 minutes, the actor animal could similarly move freely between two open arms: either to explore the arm baited with sucrose, or to investigate the Juvenile rat in the restrainer at the end of the Juvenile arm. Animals were tested once per day, in three conditions (*Juvenile vs. 2%, Juvenile vs. 5%, and Juvenile vs. 10%*) spread out over the three SSPT testing

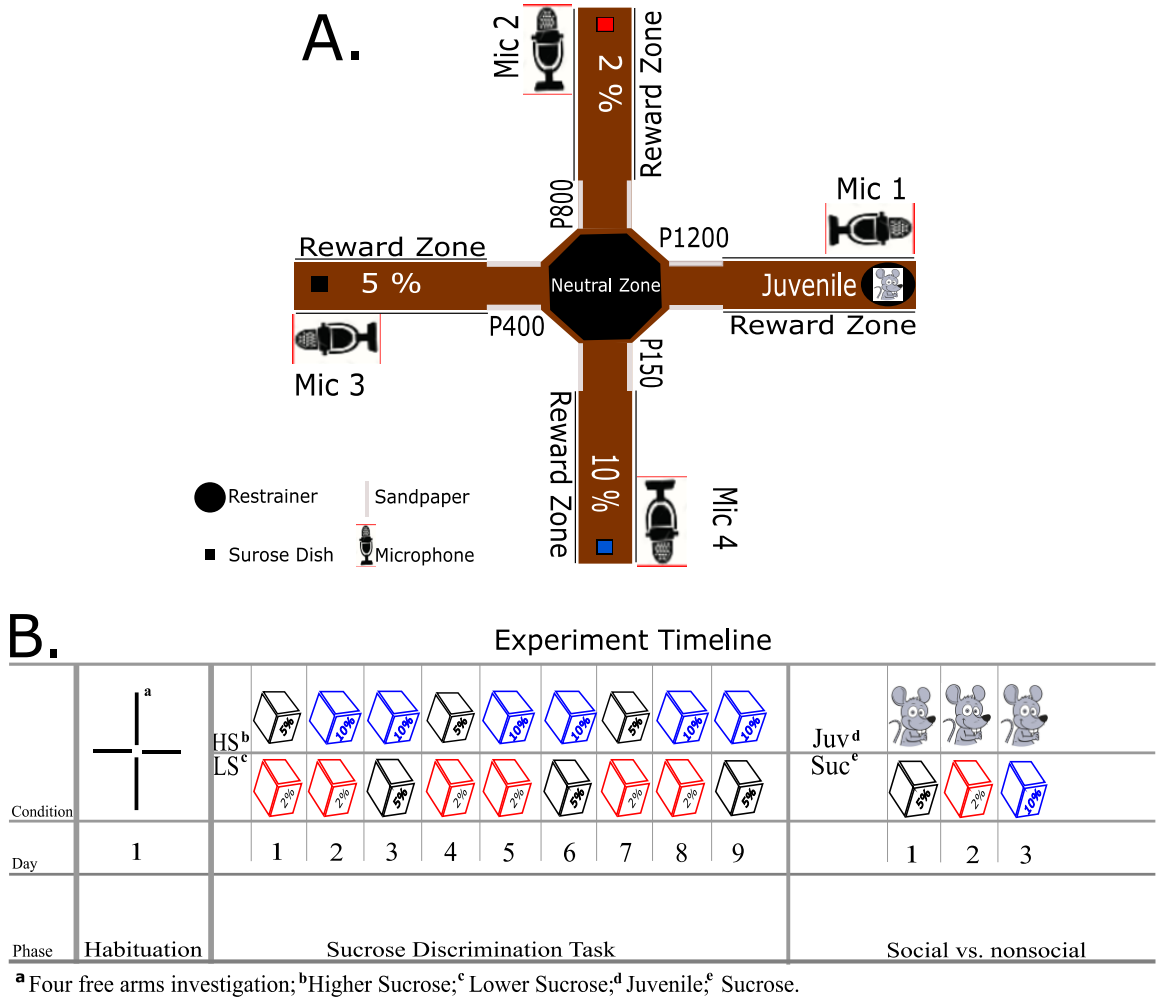


Figure 1. A schematic diagram of XCST setup with non-social reward positions, the restrainer for the juvenile social reward, sandpaper positions, and microphones. Every arm was assigned to a specific reward throughout the experiment. B shows the experiment timeline for different phases, days, and conditions–the cubes represent the sucrose in different concentrations.

days (see Fig. 1B). To keep baseline motivation equal for both types of reward (social vs. nonsocial), food deprivation was stopped after the final SDT test day, and animals were allowed to recover weight over two days before starting the SSPT. For the remainder of the experiment, animals were kept *ad libitum*. Rats usually spend more time exploring novel conspecifics than familiar ones (C. J.W. Smith, Wilkins, Mogavero, & Veenema, 2015; Caroline J.W. Smith, Mogavero, Tulimieri, & Veenema, 2017), suggesting that the value of social interaction dynamically decreases over days with increasing familiarity with the conspecific. To keep the novelty, and, hence, the value of investigation of the social stimulus similar across testing sessions, three different Juvenile rats were used in all three conditions of SSPT for each actor. The order of the identities of these Juveniles was counterbalanced across actor rats to exclude identity effects. All

USVs from all trials over the two phases (SDT and SSPT) were recorded for the full 10-minute trial duration, with the sampling rate set at 250 kHz.

Behavioral analysis: Video-tracking. For the recorded videos from all sessions, Ethovision (EthoVision XT version 11.5, Noldus) was used to track the animals' position. Tracking settings were optimized separately for each different phase of the study (Habituation, SDT, SSPT). In the habituation phase, each arm was divided into two zones (Sandpaper zone and Reward zone) to check for any inherent bias for the different reward zones and sandpaper zones. For the SDT and SSPT phases, we used the time that the animals spent in the reward zones (see reward zones; Fig. 1A). The time spent in the neutral zone was excluded from the analysis.

Ultrasonic Vocalization Recording, Labeling Procedure, and Synchronisation. Acoustic analysis of the USVs was executed using Avisoft SASLab Pro (Version 5.2, Avisoft Bioacoustics, Berlin, Germany). Spectrograms were generated with a fast Fourier transform (FFT)-length of 512 points and an overlap of 75% (Flat Top window, 100% frame size). Correspondingly, spectrograms had a frequency resolution of 390 Hz and a time resolution of 0.64 ms. In the setup, we recorded the USVs through 4 microphones, providing a four-channel spectrogram recording. The amplitude of the ultrasonic vocalizations differed depending on the distance between the animal and the different microphones (Supplementary material Appendix 1, Fig. A). The microphone channel that recorded the largest amplitude was selected for labeling for each USV in the spectrograms. This channel differed between the conditions and minutes of the trial. The labeling phase was conducted by two trained, independent labelers who labeled and classified each USV based on its sonographic features (as in Wright et al., 2010). Notably, in the SSPT phases, calls could be emitted by both the actor animal and the Juvenile social stimulus. In these analyses, we did not attempt to tease apart the source of these vocalizations, but instead rely on within-subject comparisons of actor animals to quantify differences.

The labeling phase consisted of two steps: calibration and final labelling. During the first step, two labelers became familiar (under the supervision of the expert labeler) with sonographic features of each of the 50 kHz USV subtypes (and 22KHz) according to the classification suggested by Wright et al. (2010; for an overview of the different USV subtypes considered in this study, Fig. 3 G). They initially labeled USVs together to reach a consensus labelling scheme. After this calibration step, they separately labeled the same 400 USVs and, subsequently, compared their labeling match. In total, inter-rater reliability was high (Cohen's kappa=0.95), such that 94.3% of 50 kHz USV's subtypes were labelled with the same category by both labellers. Due to technical problems, the USV files of the condition *2% vs. 5%* and some animals (1,10,11,12) from the SSPT task were lost.

Therefore, for all USV related statistical analyses, we only applied the USVs from 8 animals for both tasks. Thirty-two trials from SDTs' phase, including two days (2 and 3) for conditions (*2% vs. 10%* and *5% vs. 10%*), were labeled. For the USVs from the SSPT phase, the recordings from all three test days (N=24 recordings in total) were labeled. Both labelers tagged half of all USVs from the same conditions (every odd minute of each trial).

USV call production definition and Behavior-USV Synchronisation. When labels were assigned in Avisoft, we exported the raw data to generate a time series of vocalisation labels with a temporal resolution of 25 Hz, synchronised to the video stream and position data. Thus, each 0.040 ms sample had a one-hot encoded binary label, corresponding to the presence/absence of each of the 50 kHz subtypes, 22 kHz or background/noise. We first looked at the summed frames spent vocalizing, including all rats, to establish inclusion/exclusion criteria. We found that rats emitted vocalizations in a total of N=7252 call frames (290 seconds, combined SDT, and SSPT – 2.4% of total recorded frames). The 22-kHz USVs accounted for 23.3% of all samples with USVs, counted in ms spent vocalizing (Fig. 3D). This high proportion of 22 kHz frames is mainly caused by the naturally longer length of a 22kHz USV compared to the length of a 50 kHz call. As the main goal of this experiment only covers the 50 kHz calls, no further analysis was conducted on the 22 kHz calls (Fig. 3E shows proportions excluding 22 kHz calls). Fig. 3F shows the inter-individual variation in USV production, warranting a within-subjects approach that includes normalization to correct for these inter-individual differences in calculating group contrasts (see below). During the labeling phase, 3.9% of all call frames could not be clearly labelled in any of the 14 categories of 50 kHz subtypes. These USVs with varying sonographic features, were called Unclear (un, Fig. 3D, and Supplementary material Appendix 1, Fig B) and excluded from USVs within-between analyses. After labeling all 50 kHz USVs, six subtypes (Step-Down, Inverted-U, Step-Up, Multi-Step, Downward Ramp, and Upward Ramp) were excluded because of their small incidence (<2% of all call frames [an arbitrary cut-off], Fig. 3E). The selected call subtypes were thus: Trill, Flat, Complex, Composite, short, Flat-Trill-combination, Split and Trill-with-Jump (marked "S" in Fig. 3G)

Statistical Analyses. To rule out any spatial biases for or against some arms over others in the maze, independent of the reward contingencies, we applied independent samples t-tests to check for differences in time spent in each reward zone between groups A and B. To check for spatial bias related to any inherent preference for the different reward zones and sandpapers, we performed a repeated-measures ANOVA to assess the effect of sandpaper type and reward zones as independent variables (IVs) on the time animals spent in each reward and sandpaper zone during habituation to the maze, when rewards were not yet introduced. To find out whether rats

discriminated between different sucrose levels in the SDT, first, we calculated the SDT sucrose solution preference score for each day/condition in the SDT as a percentage of time spent with the higher sucrose (Fig.2C).

$$\text{SDT sucrose preference score} = \frac{\text{Time spent in high sucrose reward zone}}{(\text{Time spent in high sucrose reward zone} + \text{Time spent in low sucrose reward zone})} * 100$$

With these sucrose preference scores, we conducted a two-way repeated-measures ANOVA with the condition (three levels: *2% vs. 5%*, *2% vs. 10%*, and *5% vs. 10%*) and task repetition day (three levels: day 1, day 2, day 3) as (IVs) and % time spent in the higher sucrose zone as dependent variable (DV) (Fig.2D).

Similarly, for the SSPT task, first, we calculated a Juvenile preference score,

$$\text{SSPT Juvenile preference score} = \frac{\text{Time spent in the Social Reward zone}}{(\text{Time spent in the Social Reward zone} + \text{Time spent in the Nonsocial reward zones})} * 100$$

and used this Juvenile preference score to run a repeated-measures ANOVA to detect any differences in Juvenile preference as a function of sucrose concentration (*Juvenile vs. 2%*, *Juvenile vs. 5%*, and *Juvenile vs. 10%*). To find out if animals preferred a particular reward type over the other in each condition, we analysed their preference by applying a paired samples t-test. Finally, regarding the design of the maze, animals could also spend their time in the Neutral zone, as SSPT Juvenile preference score only considered the percentage of the time animals spent in reward zones, in order to know whether animals spent different time for a particular reward (either Social or Nonsocial) over the three conditions, we conducted a two-way repeated-measures ANOVAs with Conditions and Zone as IV and absolute time spent per reward zone as DV, and performed post-hoc paired-sample t-tests to compare the absolute time spent between zones per condition. For all statistical analyses, the significance level was $p < 0.05$ and all the post-hoc tests p-values were Bonferroni-corrected for multiple comparisons.

Vocalization Analyses. Our initial analysis focused on a Combined USV vocalization score (CVS), including all 15 subtypes (Including un and excluding only the 22 kHz) per session to look for overall differences in vocalization rates between conditions. Here, we first summed up all frames the rats vocalized for each of the 15 subtypes in a certain zone and then divided that score

by the time the animal spent in that zone, thus normalizing the vocalization time to the occupation time per zone, creating a normalized vocalisation rate. As inter-individual differences resulted in a skewed distribution of normalized vocalistion rates, we performed a log transformation on these combined vocalisation scores to reduced skewness and facilitate visualization. To investigate if the number of vocalizations differed depending on the reward type (social vs. nonsocial) or sucrose concentration, we applied a two-way repeated-measures ANOVA for each task (SDT and SSPT) separately. Here, we considered the condition with two levels for SDT (conditions: 2vs10% and 5v10%), three levels for SSPT (*Juvenile vs. 2%*, *Juvenile vs. 5%* and *Juvenile vs. 10%*), and two levels reward zone (SDT: higher/lower sucrose and SSPT: Juvenile/Sucrose) as IVs, and the log of the CVS of each task as DV.

To zoom in to differences between subtypes, we performed a similar analysis pipeline per subtype: after excluding the 22 kHz, unclear and infrequent call subtypes, for the remaining eight categories, we again normalized the subtype-specific vocalization rate to the spatial occupancy per zone to calculate a subtype vocalization score (SVS). This SVS was thus calculated by summing the number of frames the rat vocalized a specific subtype (1 frame = 0.040 ms) in a given zone and dividing it by the time the animal spent in that zone.

As a within-subjects normalization step, from these SVSs, we calculated a delta SVS score to show the differences in vocalisation rate between zones, for a given subtype. The delta SVS score was calculated as follows: i) SVS score in the low sucrose zone subtracted from the SVS score in the high sucrose zone for SDT and ii) SVS score in the nonsocial reward zone subtracted from the SVS score in the social reward zone for SSPT. We used this deltaSVS to compare normalized vocalisation rates between subtypes in a given condition (between-subtype analyses) and within a subtype, between conditions (within-subtype analyses).

In the between-subtype analysis, with these dSVS, we ran a Kruskal Wallis test per condition for the SDT and SSPT data, with the subtype as the IV and the dSVS score as DV for each condition.

In the within-subtype analysis, we performed a Wilcoxon Signed-Rank test for the SDT sessions, comparing the vocalization of the given subtype in two conditions [*2% vs. 5%* and *5% vs. 10%*]) and a Friedman test for each subtype across the three SSPT conditions (*Juvenile vs. 2%*, *Juvenile vs. 5%* and *Juvenile vs. 10%*). For all statistical analyses, the significance level was set at $p < 0.05$, and all the post-hoc tests p-values are Bonferroni- corrected for multiple comparisons.

Mixed Linear Model Analyses. To exploit the continuous range of sucrose solutions used in the SSPT, to look for a linear association between vocalisations and sucrose solution, we conducted

two mixed linear models, one on total calls (CVS), and one on subtype-specific SVS. Both models entered Animals as random effects, Conditions (2% vs. Juvenile, 5% vs. Juvenile and 10% vs. Juvenile) as fixed effects and CVS/dSVS as the dependent variable.

Software. All statistical analyses were carried out using SPSS Statistics (version 24; IBM, USA) and R 3.5.1 (R Core Team, 2018). We applied the following libraries in R: the tidyverse (Wickham, 2017), the haven psycho , the readxl (Wickham, Bryan, et al., 2019), the tidyr (Wickham, Averick, et al., 2019), the tibble (Wickham, Averick, et al., 2019), the sjplot (Lüdecke, 2017) the ggstatsplot (Patil,1. 2018) and the rockchalk (Johnson, 2019). Moreover, visualizations of some figures (Fig.2D and Supplementary material Appendix 1 (Fig. D)), were made using Jupyter Notebook (Kluyver et al., 2016) through the packages matplotlib (Hunter, 2007), pandas (McKinney, 2010) and seaborn (Waskom, 2018). Remaining figures were created by Inkscape (version 0.92.1, Inkscape project, 2020)

4.3 Results.

Behavior. A between-group comparison did not find evidence for a difference in spatial/reward preference based on the maze layout for groups A and B (Supplementary material Appendix 1, Fig. CA & CB). Similarly, an analysis of the habituation period did not find any evidence for a preference for a specific zone of reward ($F(3, 33) = 1.35$, $p > .05$; Fig. 2.A) or sandpaper zone ($F(3, 33) = 1.6$, $p > .05$; Fig. 2.B).

SDT. To determine whether actor animals could indeed discriminate between different sucrose concentrations (i.e., 2%, 5%, and 10%), we conducted a two-way repeated-measures ANOVA with task condition and task repetition day as within-subject factors and percentage of the higher sucrose reward as DV. We found no significant main effect of task condition, suggesting that animals did not significantly differ in their preference for the sweeter sucrose solution across sessions with different levels of sucrose concentrations. We did observe a significant main effect of day [$F(2, 22) = 15.2$, $p < 0.001$, $\eta_p^2 = .581$]. Post-hoc analysis revealed that animals preferred the higher-percentage sucrose solution significantly more in all conditions on day three (M=81.7, SE=2.7) compared to day two (M=69.1, SE=2.7, $p < .05$, $d = 4.6$) and day one (M=63.3, SE=2.7, $p < .001$, $d = 6.8$). The data thus showed that animals develop a clearer preference for the sweeter sucrose solution over days (Fig. 2C), probably as a consequence of learning. There was no significant interaction effect. *SSPT*. To assess whether animals expressed a significant preference between social and nonsocial rewards (with three different sucrose concentrations) in the social-sucrose preference test (SSPT), we conducted a one-way repeated-measures ANOVA on the percentage of time spent with the social reward (Juvenile zone). The results showed that

preferences for the Juvenile differed significantly between conditions [F (2, 22) = 52.2, p < 0.001, η_p^2 =.826]. Post-hoc tests revealed that the animals' preference for the Juvenile increased significantly from the condition *Juvenile vs. 10%* (juv. pref: M=19%, SD=10%) condition to the *Juvenile vs. 5%* (juv. pref: M=55%, SD= 15%, p <.001, *d* = 12.2) condition. There was a further, but non-significant increase in Juvenile preference when reducing the sucrose concentration to 2%; in this condition, Juvenile preference was also significantly higher than in the *Juvenile vs. 10%* condition (juv. pref: M=61%, SD= 13%, p <.001, *d* = 9.4). Three one-sample t-tests vs. indifference (50%) showed that animals preferred the social reward in Juvenile vs. 2% (M=61.5, SD=13, t(11)=3.06, p <.05), were indifferent between Juvenile vs. 5% (M=54.7, SD=15, t(11)=1.08, p >.05) and preferred the sucrose reward in Juvenile vs. 10% (M=80.6, SD=10, t(11)=9.9, p <.001). These results show clearly that animals indeed traded off interacting with a Juvenile to the consumption of sucrose and also that a preference for interacting with the Juvenile when sucrose levels were low (2%) could be reversed when confronted with a more preferred 10% sucrose solution (Fig. 2D). These between-condition differences could be due to a change in time (%) spent at the sucrose reward, the social reward, or both. To quantify this, we investigated if the absolute time that animals spent on each reward zone differed between different conditions. A repeated-measures ANOVA on the absolute time animals spent on social reward showed a significant effect of conditions [F (2, 22) = 33.2, p < 0.001, η_p^2 =.751]. Post-hoc tests revealed that the absolute time that animals spent in the Juvenile zone in the condition of *Juvenile vs. 10%* (M=97.7, SD=55) was significantly less than in the condition *Juvenile vs. 5%* (M=250, SD=76, p <.001, *d* = 2.2) and the condition *Juvenile vs. 2%* (M=259, SD=64, p <.001, *d* = 2.7). There was no significant difference between the condition *Juvenile vs. 2%* and *Juvenile vs. 5%*. A second repeated-measures ANOVA on the absolute time animals spent with nonsocial rewards also showed a significant effect of condition [F (2, 22) = 74.7, p < 0.001, η_p^2 =.872].

Here, post-hoc tests revealed that the absolute time that animals spent in the sucrose zone in the condition *Juvenile vs. 10%* (M=408, SD=19) was significantly more than the condition *Juvenile vs. 5%* (M=205, SD=20, p <.001, *d* = 10.4) and the condition *Juvenile vs. 2%* (M=159, SD=15, p <.001, *d* = 14.5). No significant difference was found between the conditions *Juvenile vs. 2%* and *Juvenile vs. 5%* (Fig.2E). As a follow-up analysis, a paired sample t-test per condition revealed that in Juvenile vs. 2%, Juvenile (M=259, SD=64) was significantly (p < 0.05, *d*= 1.7) preferred to sucrose (M=159, SD=53). In the condition *Juvenile vs. 5%*, animals were indifferent between the reward types (Juvenile: M=250, SD=76; 5% sucrose: M=205, SD=71). In contrast, in the condition *Juvenile vs. 10%,* the sucrose side (M=408, SD=67) was preferred significantly (p < 0.001, *d*= 5.6) over the Juvenile (M= 97, SD=55).

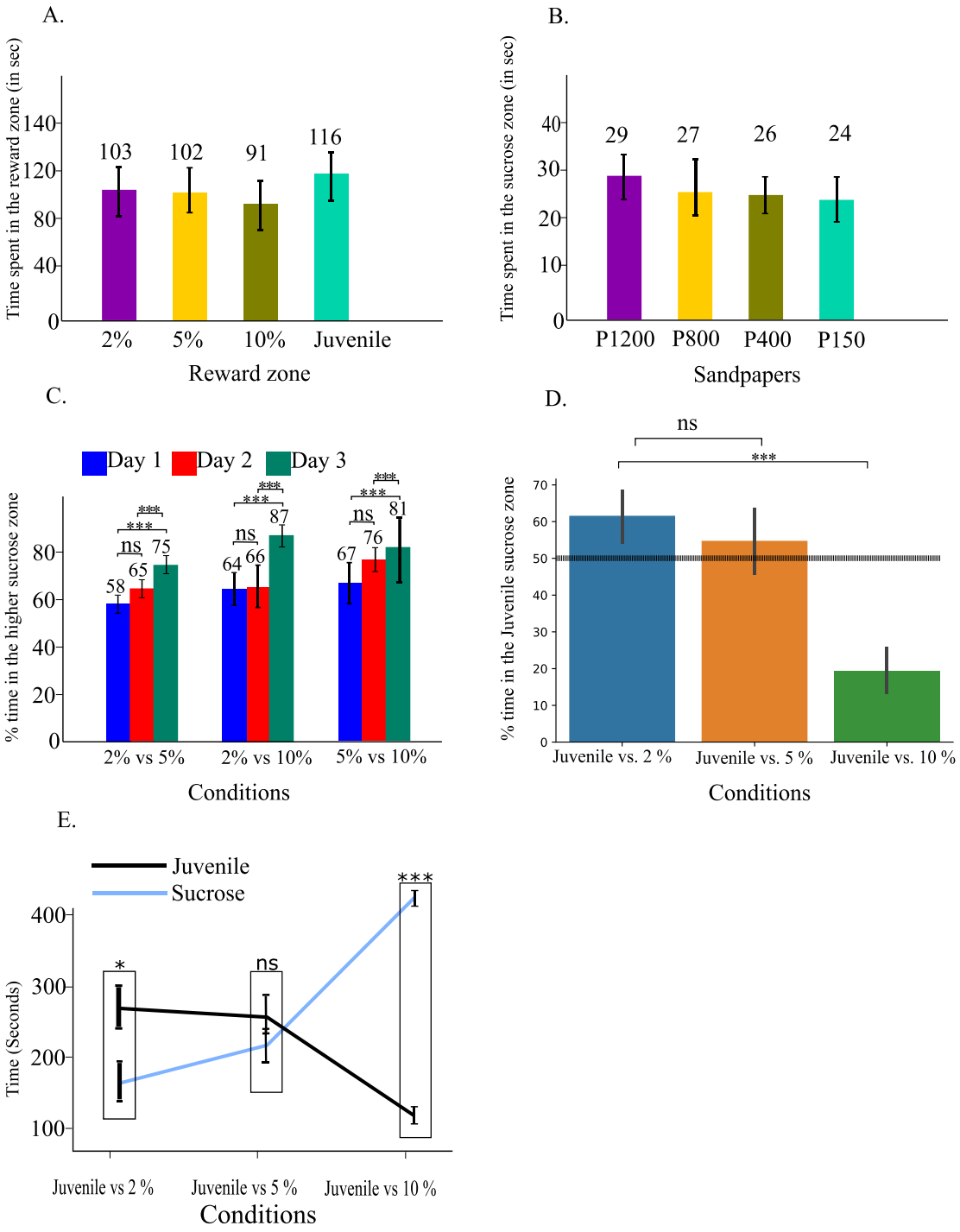


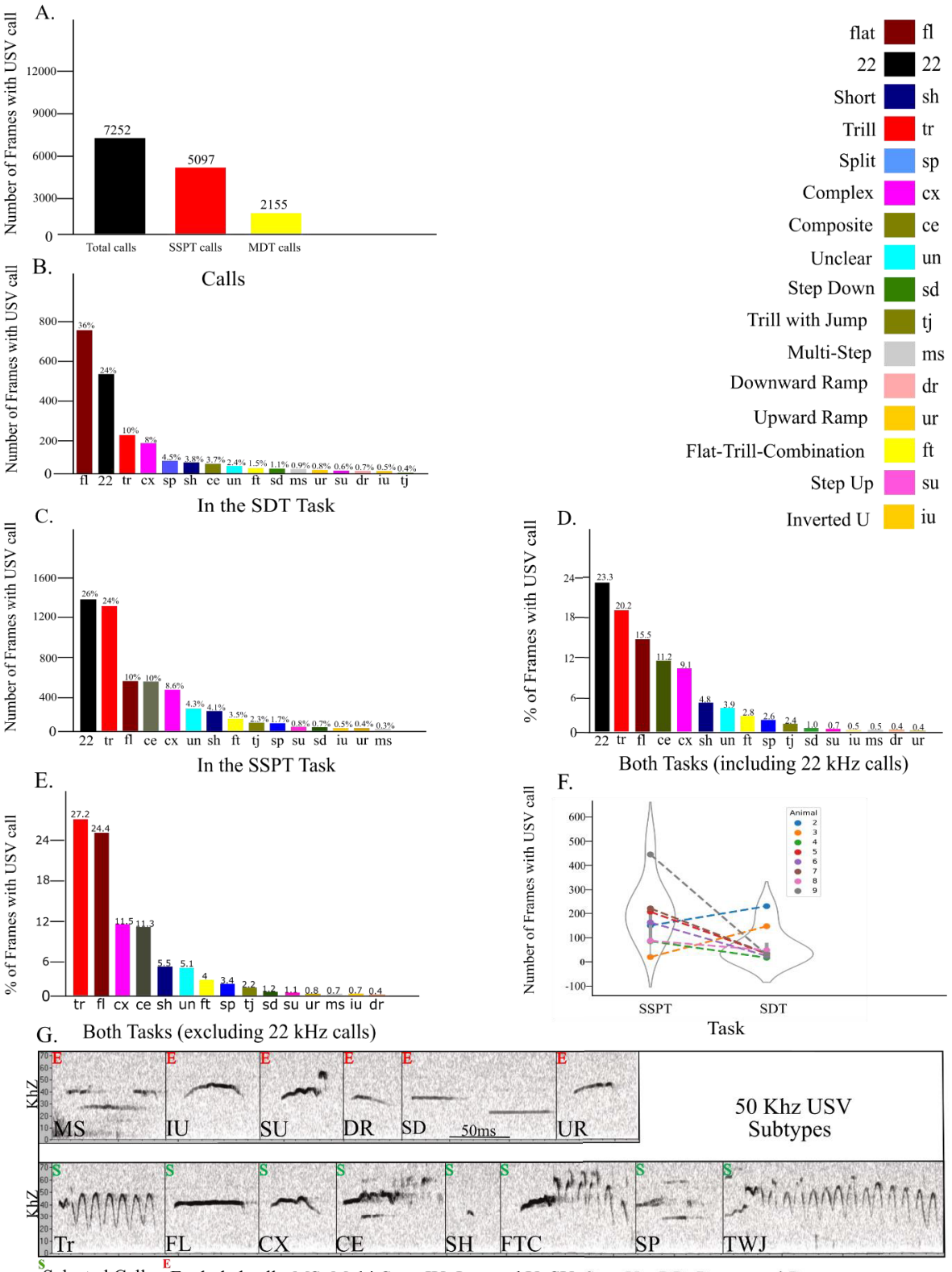
Figure 2. A time spent in each reward zone during the Habituation phase. B time spent in each sandpaper zone during the Habituation phase. C time spent in the higher sucrose zone in all three conditions of SDT. D time spent in the Juvenile zone in all three conditions of SSPT, the dashed line shows the 50% point. E absolute time spent in each reward zone for all three conditions. All error bars show the standard deviation. *p < .05, **p < .01, ***p < .001

Characterization of USV. As indicated in the methods section, the 50 kHz USVs produced by actor animals in the SSPT were labeled and further categorized into subtypes. Descriptive statistics were generated for each of the subtypes included in our analyses, along with within-condition and between-condition comparisons. Based on prevalence, we selected eight subtypes: Trill (Tr), Flat (Fl), Complex (Cx), Trill-with-Jump (Tj), Short (Sh), Flat-Trill-combination (Ft), Split (Sp), and Composite (Ce) for further analysis (Fig. 3G). Six subtypes (Step-Down, Step-Up, Upward Ramp, Multi-Step, Inverted-U, Downward Ramp) were excluded from analysis due to their limited occurrence (<2% of calling time, Fig.3E). From the selected subtypes, Tr (27.2%), Fl (24.4%), Cx (11.5%), and Ce (11.3%) were the most prevalent, while Sh (5.5%), Ft (4%), Sp (3.4%), and Tj (2.2%) were least prevalent in both tasks (Fig. 3E).

SDT. In total, throughout the SDT, 2155 call frames were found in which the rats were vocalizing, and from these eight selected subtypes, Fl (36%), Tr (10%), Cx (8%), Sp (4.5%), Sh (3.8%) and Ce (3.7%) were most prevalent while, ft (1.5%), and Tj (0.04%), were least prevalent in SDTs' conditions (Fig. 3B). SSPT. In total, in the SSPT, 7252 call frames were found in which the rats were vocalizing, and from these eight selected subtypes, Tr (24%), Fl (10%), Ce (10%), Cx (8.6%) were most prevalent while Sh (4.1%), Ft (3.5%), Tj (2.3%) and Sp (1.7%) were least prevalent (Fig. 3C) in SSPTs' conditions.

Analysis of Total USVs. To determine if the number of frames that the rat vocalised was affected by sucrose concentration or type of rewards in the different conditions, we conducted a two-way repeated-measures ANOVA on the Combined USV vocalization score (CVS; the number of frames vocalised relative to the time spent in the visited zone, see methods) with condition and reward zone as factors, separate for SDT and SSPT.

SDT. The SDT analyses found a significant effect of condition on the CVS [$F(1, 7) = 14.9$, $p < .01$, $\eta_p^2 = .680$]. The main effect showed that the CVS was significantly higher in the condition *2% vs. 10%* higher (*M*=.310, *SE*= .075) than in the condition *5% vs. 10%* (*M*=.128, *SE*= .054; Fig. 4A). The factor reward zone also had a significant effect on the CVS [$F(1, 7) = 14.3$, $p < .01$, $\eta_p^2 = .672$; Fig. 4B]. The main effect showed that the CVS was, surprisingly, higher ($p<.01$) in the lower sucrose concentration zone (*M*=.268, *SE*= .065) compared to the higher sucrose concentration zone (*M*=.171, *SE*= .058). There was also a significant interaction effect of conditions and reward zones ([$F(1, 7) = 5.9$, $p < .05$, $\eta_p^2 = .459$; Fig. 4E&F]). Post-hoc comparisons showed that CVS was higher in lower-reward zones only for the condition *2% vs. 10%*, in the zone



Selected Calls; Excluded calls. MS: Multi-Step, IU: Inverted U, SU: Step-Up, DR: Downward-Ramp
SD: Step Down, UR: Upward-Ramp,TWJ: Trill with Jump,Tr: Trill, FL: Flat, CX: Complex, CE: Composite
SH: Short, FTC: Flat Trill combination, SP: Split

Figure 3. A number of frames animals vocalized in both tasks and separately during each task. B number of call frames for each distinct subtype in the SDT task. Please note that the high number of frames spent vocalising 22 kHz reflects a relatively small number of long calls. C number of call frames of each subtype in the SSPT. D percentage of each subtype vocalized in both tasks, including 22 kHz. E percentage of each subtype vocalized in both tasks, excluding the 22 kHz. F number of call frames per animal per task averaged over all three conditions and G Examples of the fourteen 50 kHz USV Subtypes (labeled according to Wright, Gourdon and Clarke, 2010). Subtypes are marked with S (selected) or E (excluded, see text).

of lower sucrose concentration (*M*=.407, *SE*= .093) the animals had a higher CVS (p<.05) than the condition *5% vs. 10%* (*M*=.213, *SE*= .064, see Fig.4 A&B).

SSPT. For the SSPT we again performed a two-way within-subjects repeated-measures ANOVA. We found a significant effect of reward type ([F (1, 7) = 13.6, p <.01, η_p^2 =.658, Fig. 4D]). Post-hoc comparisons showed that the CVS was significantly higher in the Juvenile zone (M=.544, SE= .075) than in the sucrose zone (M=.313, SE= .067; p<.01). Furthermore, there was a significant interaction between condition and reward types [F (2, 14) = 5.1, p <.05, η_p^2 =.426, Fig. 4 G-I]. Post-hoc comparisons showed that animals' CVS in the Juvenile vs. 10% condition was significantly higher (p < .01) in the Juvenile zone (M=.685, SE= .121) compared to the sucrose zone (M=.297, SE=.064). No significant differences in CVS between reward zones were found for the Juvenile vs. 2% (p=.06) and Juvenile vs. 5% conditions [p =.07].

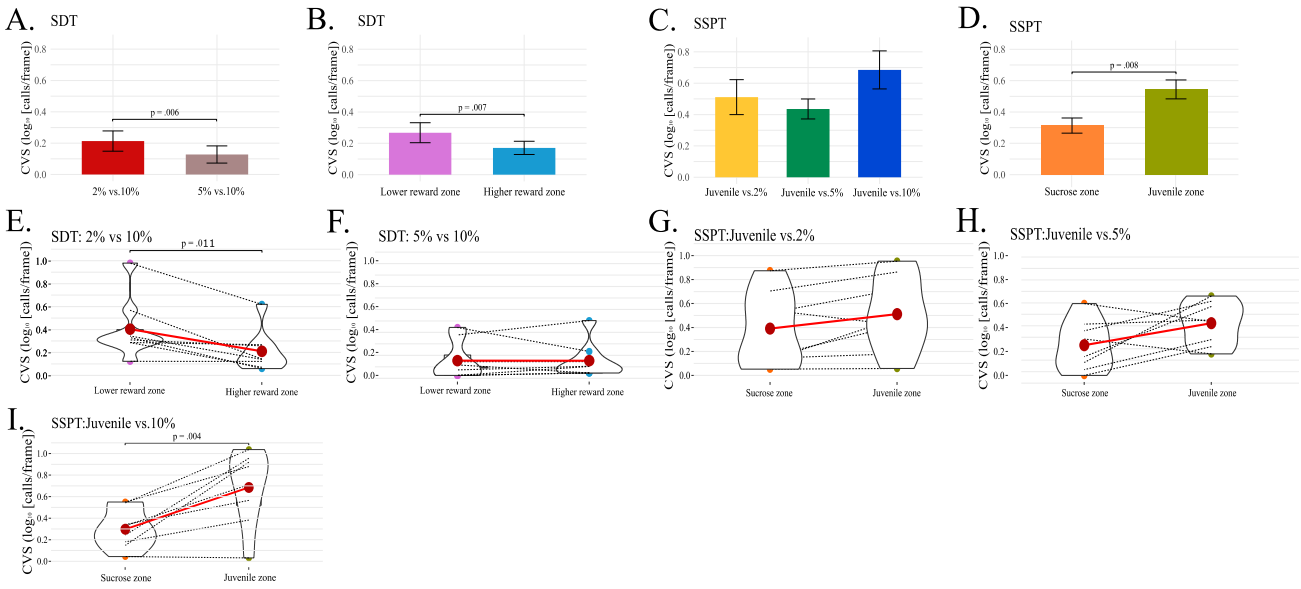


Figure 4. A, B, C, and D present animals' CVS in conditions (A, C) and reward zones (B, D) of the SDT and SSPT tasks, respectively. E and F show each animal's CVS in the two reward zones of the SDT task's two conditions. G, H, and I demonstrate each animal's CVS in the two reward zones of the SSPT task's three conditions. Error bars indicate Standard Error for A, B, D, and Standard Deviation for C.

These results already indicate an interesting finding: while behavioral preferences shifted towards the sucrose reward zone with higher sucrose concentration, the vocalisation rate showed the opposite trend, with increasing vocalisations recorded in the juvenile zone with increasing sucrose concentrations. We next investigated whether this pattern was present for specific subtypes, and if there were differences between subtypes.

Comparing USV subtypes between and within conditions.

Between-Subtypes Analyses. As one of the main questions of this study, we were interested in finding out if the different sucrose concentrations or different reward types were associated with a different vocalization palette across the 50 kHz USV subtypes. Here, we used the delta Subtype Vocalisation Score (dSVS; see methods), indexing the relative difference in vocalization rates between reward zones in a given session for these analyses, as it accounts for normalization of inter-individual differences in absolute call rates. **SDT.** We conducted a Kruskal Wallis test separately for each condition (*2% vs. 10%* and *5% vs. 10%*) by taking the eight subtypes observed in the SDT as a factor and their dSVS as the dependent variable (DV).

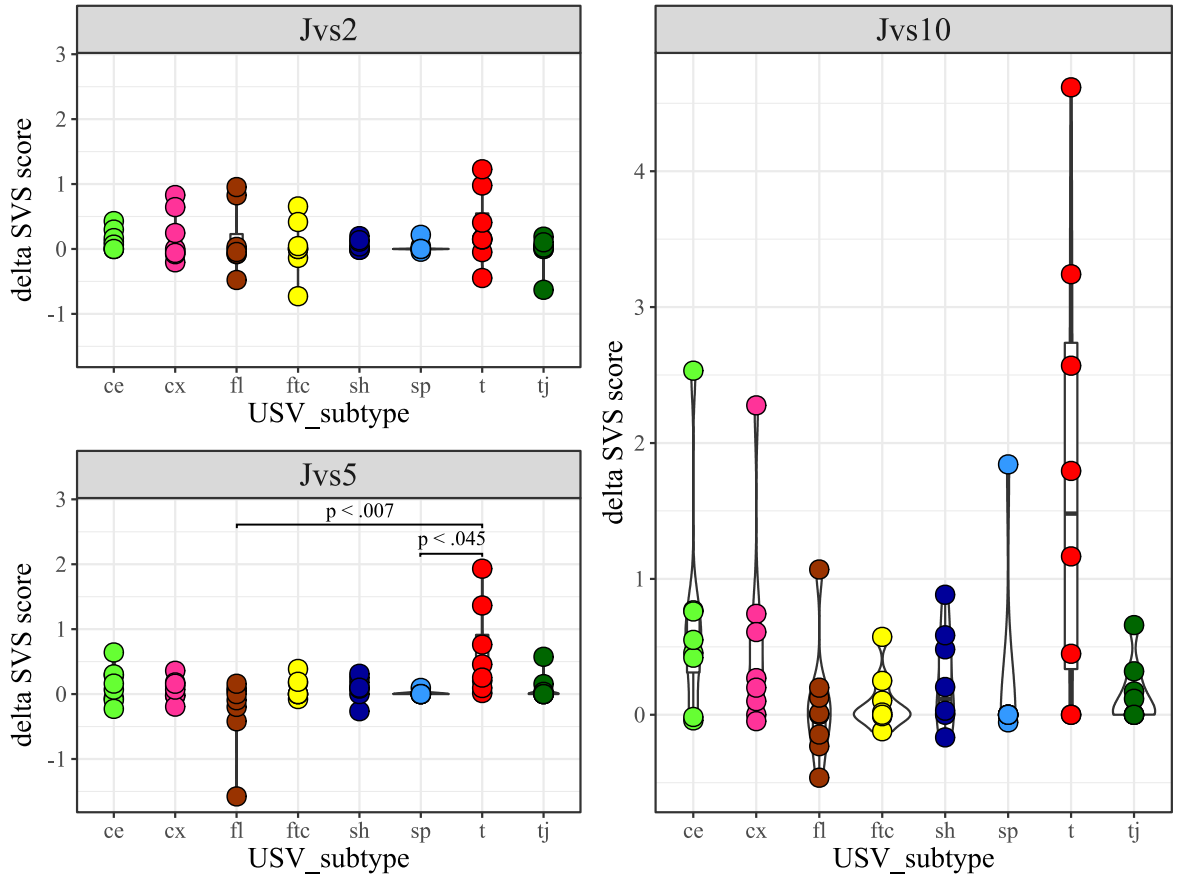


Figure 5. dSVS for each subtype, split by conditions in the SSPT.

We found no significant difference in the dSVS between subtypes for any condition (Supplementary material Appendix 1, Fig. D). *SSPT*. We similarly conducted a Kruskal Wallis test for each condition (*Juvenile vs. 2%*, *Juvenile vs. 5%*, and *Juvenile vs. 10%*). In the condition *Juvenile vs. 5%*, we found a significant difference ($H(7) = 16.6$, $p < .05$). Post- hoc pairwise comparisons showed a significant difference between dSVS of the subtypes Tr (median= 0.3) and Fl (median= -0.04), (Mann-Whitney U-test, $p < .01$) and dSVS of subtypes Tr and Sp (median= 0, $p < .05$; Fig.5)

Within subtype analyses.

SDT. this analysis was conducted to determine whether dSVS for a given subtype differed between conditions. The Wilcoxon Signed-Rank test results showed that the dSVS score of Tr was lower in condition *2% vs. 10%* (median=-0.4) than in condition *5% vs. 10%* (median= 0), $Z=2.1$, $p < .05$). There was no other significant difference within any subtypes between conditions (Fig. 6).

Mixed linear model analyses. For the within-subtype analysis of call rates in the SSPT, we exploited the continuous nature of the sucrose concentration in a mixed linear model, estimating the relationship between sucrose concentration (in %) and dSVS with individual animals modelled

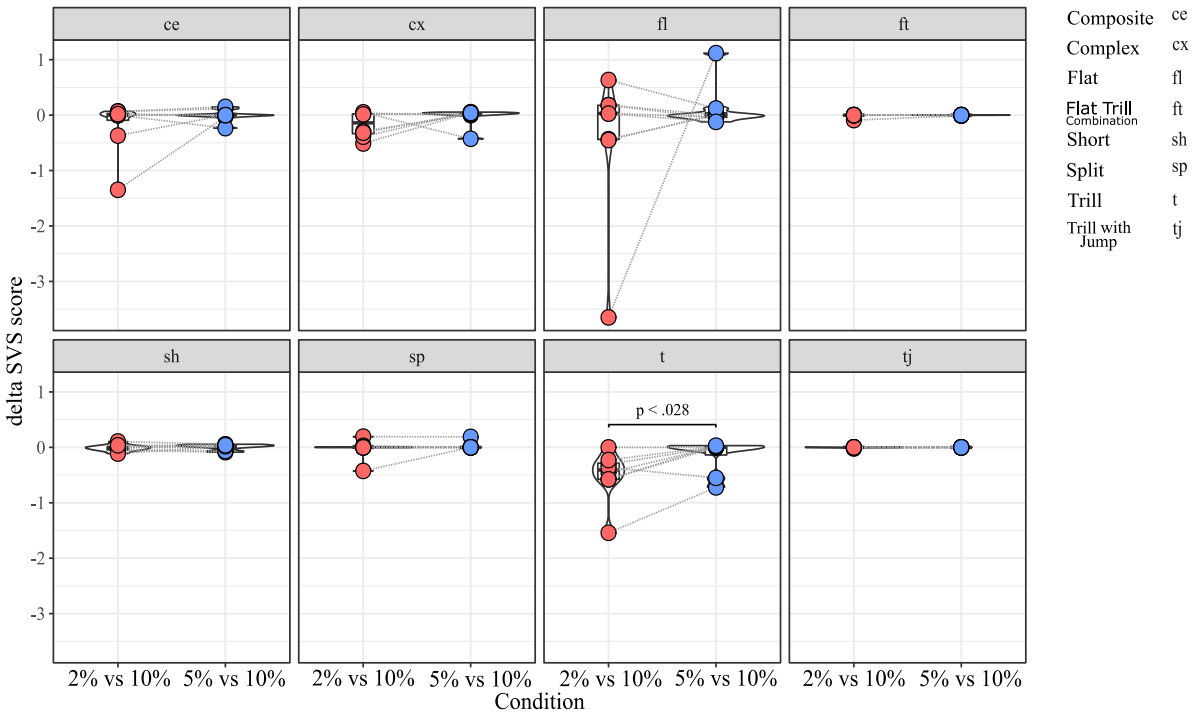


Figure 6. dSVS split by subtype between conditions

as random effects. We first modelled the total call rate (all calls combined) using Combined USV vocalization score (delta CVS; see methods). The mixed linear model showed a linear association between the delta CVS and the sucrose level (beta = 0.034, 95% CI [0.01 – 0.06], t (15) = 3.27, p <.01, R2 fixed effect = 0.208). This suggests that the difference in total vocalisation time in the Juvenile over the Sucrose zone significantly *increased* with higher levels of sucrose concentration (see Fig.7A and supplementary material Appendix 1, Table 1). We then modelled the sucrose concentration to delta SVS relationship in linear mixed models separately for each subtype. The models showed a significant association for the subtypes Tr (beta=0.18, 95% CI [0.05 – 0.031], p< .05) and Ce (beta=0.07, 95% CI [0.01– 0.013], p<.05). This means that, for these two subtypes, the difference in the number of frames vocalized in the Juvenile over the Sucrose zone significantly increased with higher levels of sucrose concentrations. (see Fig. 7B and supplementary material Appendix 1, Table 2.A & B for more individual model statistics).

4.4 Discussion.

Communication is essential for social animals, and rats are no exception. Rats utilize vocalizations in the ultrasonic range to communicate with their conspecifics. However, whether these vocalizations differ in response to different rewards when presented together, and whether vocalisations quantitatively index reward magnitude remained mostly unexplored.

Here, we presented a paradigm to test preferences for two different reward types head-to-head in distinct spatial locations on a four arm-maze. We simultaneously quantified social vs. nonsocial reward value through relative reward zone time allocation and reward type preference profiles by estimating slopes over three clearly discriminable (Fig. 2C) nonsocial reward values (sucrose concentrations). Rats, indeed, changed their time allocation over reward sites as a function of reward sucrose concentration (Fig. 2E) and even exhibited preference reversals, switching from preferring social interaction when it competed with 2% sucrose to preferring sucrose consumption when its concentration was upped to 10%. This change in behavioral preference and time allocation could be exploited to estimate the association between different 50 kHz USV subtypes and social vs. nonsocial reward, controlling for individual differences in overall vocalisation rate and variance in time spent at each reward site (Fig. 4). We found that, when controlling for occupancy and individual differences in this way, the overall difference in vocalisation rate between social and nonsocial reward sites (dCVS; normalized vocalization rate social minus nonsocial) increased from 2 to 5 to 10% sucrose conditions, as estimated with a linear model, suggesting that

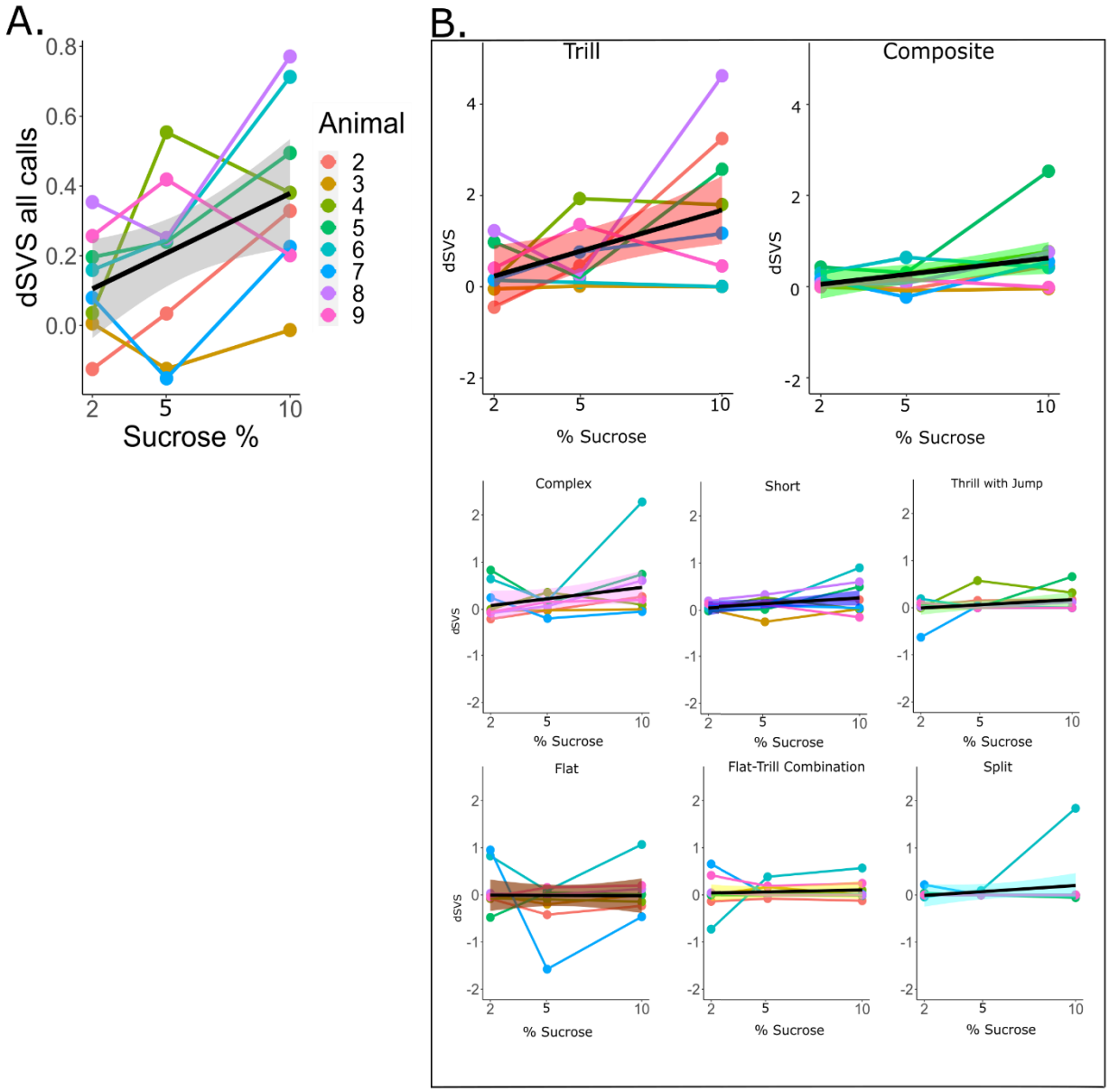


Figure 7. A. dCVS for all calls across different levels of sucrose in the SSPT. The black line (±standard error of the mean; grey shade) shows the estimated linear relationship between dSVS and sucrose concentrations, across all rats. Linked dots represent individual rats, modelled as a random effect. B. Each plot shows the change in dSVS of a certain subtype across three SSPT task conditions. Black lines represent the mean linear trends across all rats and (±standard error of the mean is represented by shade; coloured differently for each subtype.). The slopes for Trill and Composite subtypes are significant [supplementary material Appendix 1, Table 2.A]

animals vocalised *more* in the social zone even though the actor animals spent *less* time in the social side when the alternative was a high-sucrose solution. The vocalisation rate was not purely determined by appetitive sucrose consumption either, as witnessed by the dramatic reduction in call rate in the SDT conditions, even though animals exhibited comparable levels of sucrose consumption and behavioral preferences. As several studies already showed, 50 kHz USV calls are emitted during various appetitive states (Brudzynski & Zeskind, 2018), such as sucrose consumption and social play (Browning et al., 2011). Therefore, we hypothesized that, in the SDT task, more calls would be emitted in the *5% vs. 10%* condition than the *2% vs. 10%* condition (overall more sucrose) and that a higher percentage of calls would be scored in the higher sucrose zone in both conditions. Both hypotheses were *rejected*, however, as the rats vocalized significantly *more* in the *2% vs. 10%* condition, controlling for occupancy, and more calls we found in the *lower sucrose* zone in both conditions.

These findings, thus, rather support a view of USVs as a context-dependent communicative device aimed perhaps at establishing/inviting social contact compared to the alternative hypothesis that casts USVs as (static) epiphenomena of reward value linked to the consumption of social contact or nonsocial rewards. Many researchers have pointed to the associations between the various 50 kHz USV subtypes and certain types of overt behavior (Mulvihill & Brudzynski, 2018a, 2018b; Wöhr et al., 2008; Wright et al., 2010). When we zoomed in to the level of the various 50 kHz subtypes, we found that in our experiments, eight subtypes (Tr, Fl, Cx, Tj, Sh, Ft, Sp, and Ce) were vocalized much more prevalently than the other remaining subtypes identified by Wright et al. (Wright et al., 2010). We thus investigated whether the vocalisation rate of these subtypes could be used to discriminate between Social and nonsocial reward-related contexts.

When considering the SDT sessions, the Flat subtype was vocalized at a much higher rate compared to the remaining eight selected subtypes (Fig.3B). This parallels the findings of Mulvihill and Brudzynski (2018b), who reported that nonsocial conditions appeared to induce a greater proportion of flat calls as well as the findings of Wöhr and Schwarting (2013), who found an association of flat 50 kHz USVs and feeding behavior. Likewise, Wright et al. (2010) also found that flat calls were more prevalent in singly-tested rats than pair-tested rats. However, in our hands, the proportion of flat calls across high- and low-reward zones (dSVS) did not differ between flat calls and the other subtypes (Supplementary material Appendix 1 (Fig. D)) or across SDT conditions for flat calls (Fig. 6), arguing against a direct, parametric association between flat calls and hedonic state.

In contrast, similar to the findings of Wright, Gourdon, and Clarke (2010) and Brudzynski and Pniak (2002), demonstrating that animals generally vocalize more in the presence of conspecifics, in the SSPT, our subjects also vocalized more in the social reward zone than the nonsocial reward zone. Moreover, sucrose levels influenced this effect as conditions with a competing higher concentration of sucrose elicited higher vocalisation of 50 kHz USVs in the social zone (Fig. 7A). This result parallels the results of Mulvihill and Brudzynski (2018a), who demonstrate that social contexts in particular conditions induce call emission more robustly. In particular, the Trill and Composite subtypes drove this effect and were produced at increasing rates in the social zone when animals were deciding between visiting the Juvenile and increasing sucrose (Fig. 7B). This finding becomes particularly interesting when we consider that animals spent more time at the *nonsocial* zone at higher sucrose concentration conditions. What could explain this inverse relationship between behavioral preferences and differential USV production? We offer three putative explanations:

1) The sessions with higher sucrose concentrations induce an overall higher hedonic state that potentiates "chattiness" when the actor animal visits the Juvenile zone

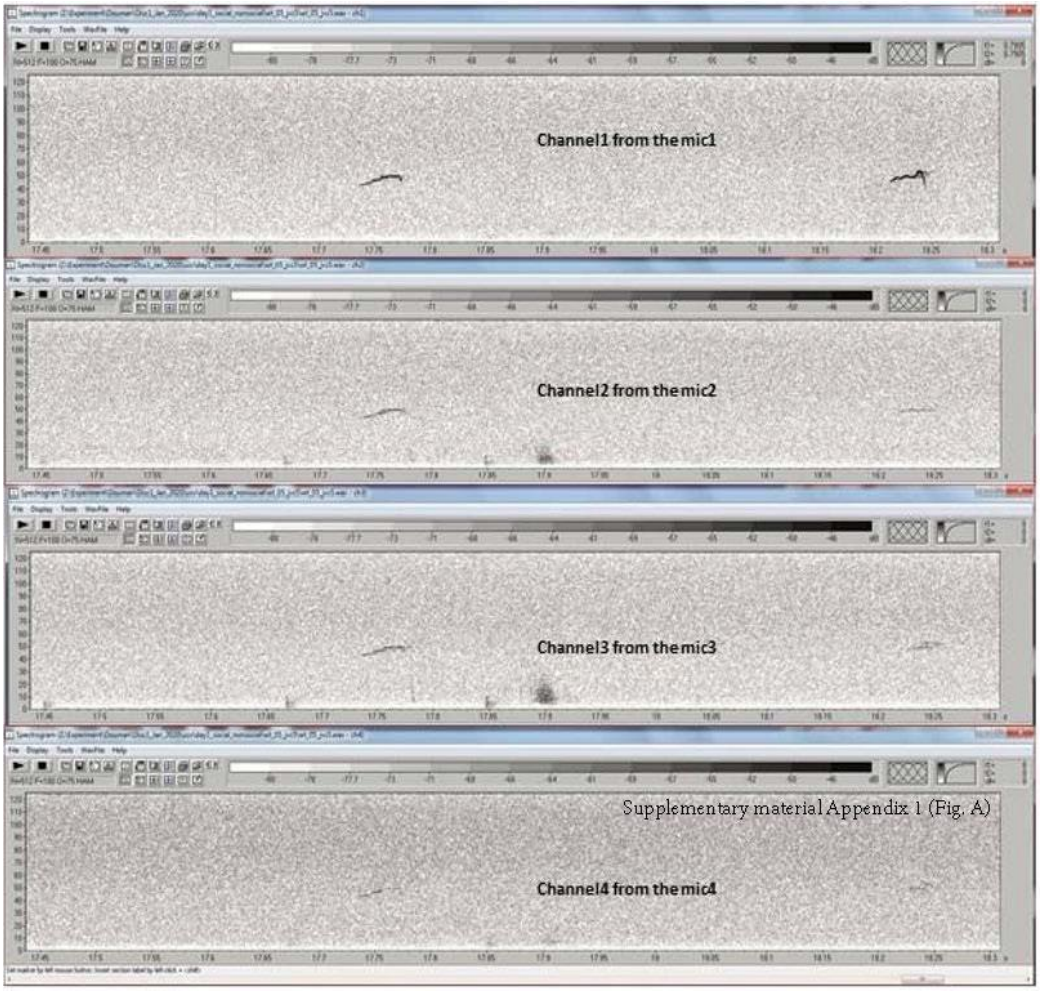
2) The higher sucrose content influences the breath of the actor, which in turn modulates the USV production when the animals are interacting

3) With increasing sucrose concentration, the actor animals shuttle more and faster between reward sites (anecdotal observations). If USV production decays exponentially with interaction time, shorter interactions yield a higher (normalized) call rate.

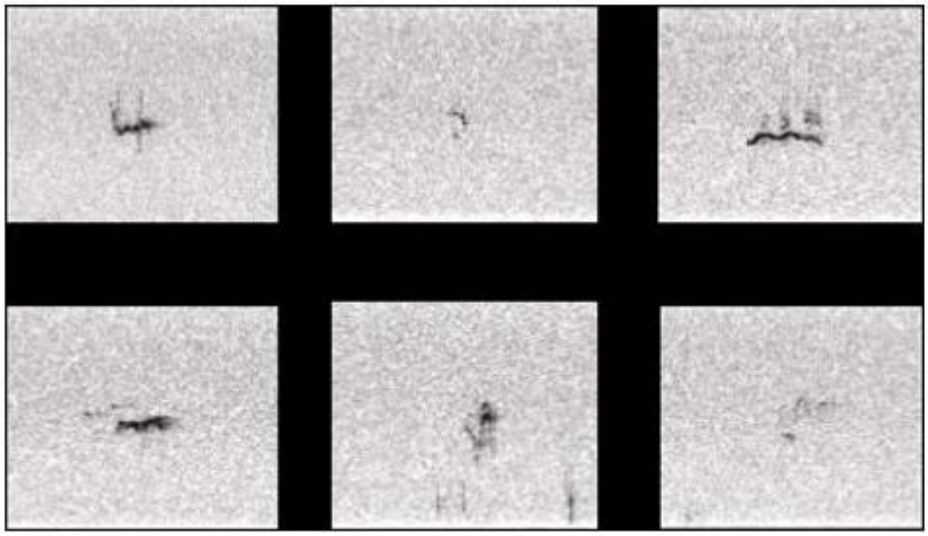
Limitations & Future directions. Adjudicating between these options will require further studies. One important limitation worth mentioning is that we utilized rats raised and tested in laboratory conditions. In a sense, our design is a drastically simplified version of what a rat might encounter in naturalistic settings. Studies such as ours aimed at elucidating the intricate patterns and subtypes of vocalizations in a micro-scale should be consolidated with field studies and naturalistic designs of rodent vocal behavior. Another important limitation of our study is that when the actor and Juvenile rat in the maze were in close proximity to each other, we were unable to determine with accurate precision whether the actor rat or the Juvenile was producing the calls that were recorded. Though several attempts have been made, using triangulation, microphone arrays (Heckman et al., 2017), or Onboard wireless EMG recordings of the larynx (Kelm-Nelson, Lenell, Johnson, & Ciucci, 2018) to arrive at precise disambiguation of the USV source, the current setup did not allow this objective to be met in our study.

Taken together, our study provides a first systematic overview of behavioral preferences and vocalization patterns recorded when rats are choosing between social and nonsocial rewards. The underlying behavioral and/or genetic traits and the neural correlations regulating the rats' specific preferences are yet to be explored. Recent studies utilizing a combination of cutting edge genetic techniques to pinpoint neural underpinnings of rodent vocal communication (Gao, Wei, Wang, & Xu, 2019; Kisko et al., 2018; Tschida et al., 2019) have illustrated the value of rodent models in elucidating the social behavior and pro-social 50-kHz ultrasonic communication as models of psychiatric illness as well as neural underpinnings of mammalian vocal communication. Our results again highlight the variance in rat vocalizations between individuals and within their repertoire. Not only did the total number of USVs differ depending on the type of and level of reward, but the specific subtypes themselves showed variation between conditions and rewards, and in some cases, were predictive of the level of reward. So what is the ultimate role of the different USV subtypes? We and others propose that these USV subtypes allow rats plasticity in their vocal behavior, enabling flexible communication to respond to the (social) cues from their surroundings in an appropriate manner. The conditional probability of one subtype following another is not random (Coffey et al., 2019), suggesting the possibility of syntax, or perhaps even turn-taking in an interacting rodent dyad. Such analyses could be combined with data-driven approaches to USV categorization that include frequency and/or amplitude information and machine learning in addition to expert-based pattern recognition of USV subtypes. Creating synthetic USV sentences that could outperform random sequences in eliciting approach behavior, now used as the gold standard (Seffer et al., 2014), would indicate the importance of subtypes in a USV call structure.

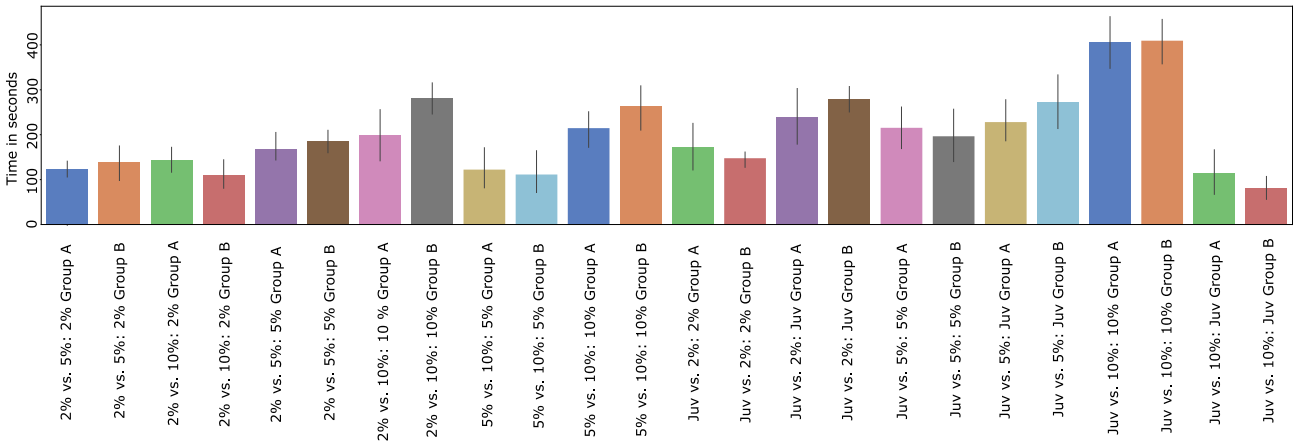
4.5 Supplemental Materials. Distinct profiles of 50 kHz vocalizations differentiate between social versus nonsocial reward approach and consumption.



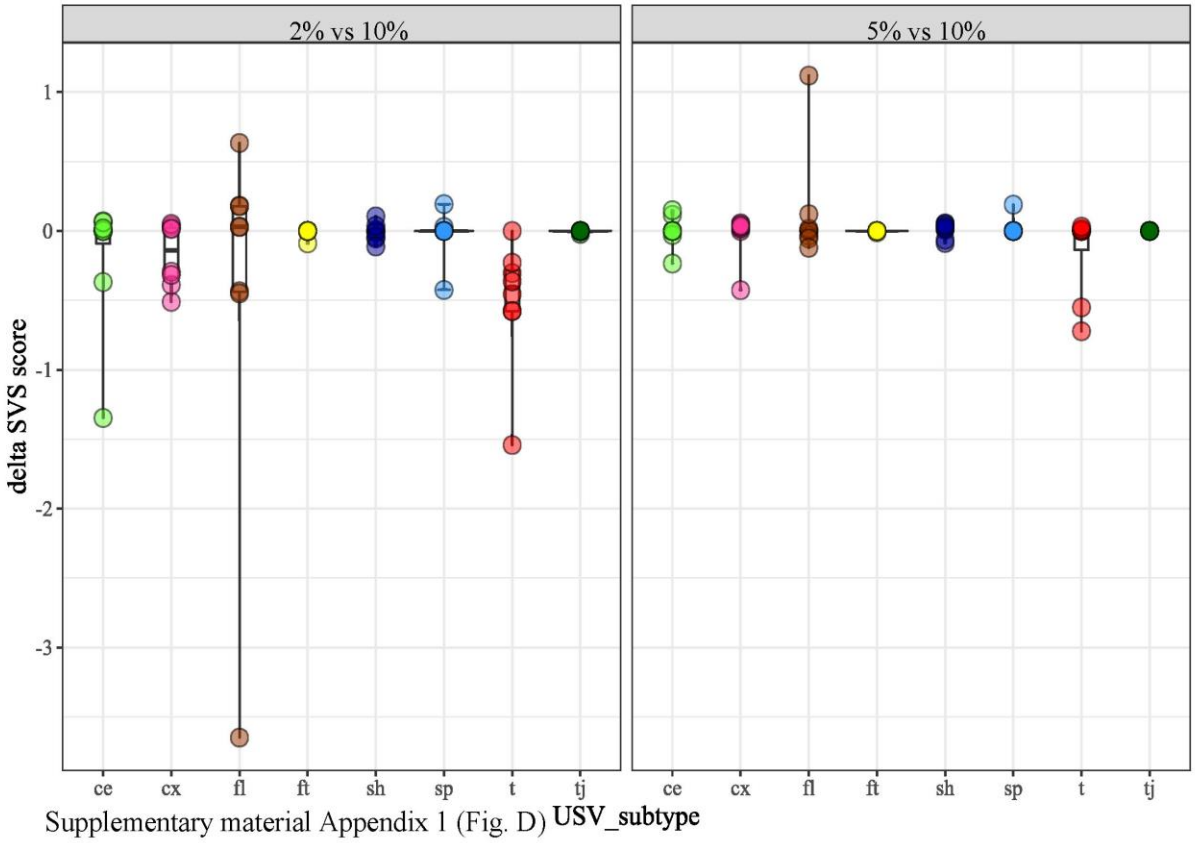
Supplementary material appendix 1. Figure A.



Supplementary material appendix 1. Figure B.



Supplementary material Appendix 1. Figure CA & CB.



Supplementary material Appendix 1 (Fig. D)

Supplementary material Appendix 1. Figure D

	log_delta_CVS		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.04	-0.15 – 0.22	0.675
sucrose_conc	0.03	0.01 – 0.06	**0.005**
Random Effects			
σ^2	0.03		
τ_{00} Animal	0.02		
ICC	0.44		
N Animal	8		
Observations	24		
Marginal R^2 / Conditional R^2	0.208 / 0.553		

Supplementary material Appendix 1. Table 1

	ce			t			cx			sh		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.09	-0.52 – 0.33	0.646	-0.13	-1.04 – 0.78	0.769	-0.02	-0.44 – 0.40	0.929	-0.00	-0.20 – 0.19	0.972
Condition	0.07	0.01 – 0.13	**0.018**	0.18	0.05 – 0.31	**0.011**	0.05	-0.01 – 0.10	0.085	0.02	-0.00 – 0.05	0.084
Random Effects												
σ^2	0.19			1.02			0.18			0.05		
τ00	0.06 Animal			0.11 Animal			0.07 Animal			0.01 Animal		
ICC	0.23			0.10			0.29			0.14		
N	8 Animal			8 Animal			8 Animal			8 Animal		
Observations	24			24			24			24		
Marginal R^2 / Conditional R^2	0.189 / 0.377			0.246 / 0.320			0.094 / 0.359			0.114 / 0.237		

Supplementary material Appendix 1. Table 2A

	fl			sp			tj			ftc		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-0.00	-0.46 – 0.45	0.986	-0.06	-0.39 – 0.26	0.690	-0.05	-0.24 – 0.14	0.580	0.03	-0.21 – 0.26	0.820
Condition	-0.00	-0.07 – 0.07	0.963	0.03	-0.02 – 0.08	0.269	0.02	-0.00 – 0.05	0.102	0.01	-0.03 – 0.04	0.635
Random Effects												
σ^2	0.27			0.14			0.04			0.07		
τ00	0.01 Animal			0.00 Animal			0.01 Animal			0.00 Animal		
ICC	0.02			0.02			0.17			0.00		
N	8 Animal			8 Animal			8 Animal			8 Animal		
Observations	24			24			24			24		
Marginal R^2 / Conditional R^2	0.000 / 0.023			0.053 / 0.074			0.099 / 0.252			0.010 / 0.010		

Supplementary material Appendix 1. Table 2B

Discussion.

5.1 Introduction.

The following discussion aims to prove that the two introduced tasks and their results have unlocked the value of studying social associative learning, its behavioral repertoires, and underlying brain substrate(s) in rats and therefore fulfils the criteria of a meaningful and useful thesis. The focus for Section 5.2 lies on the implications the social unblocking study has for the field of vicarious learning and prosocial choice and also includes a comparative approach that attempts to bridge the differences between animal and human social unblocking. A proposal is then presented of how rats can vicariously learn about mutual rewards and how exactly social information exchanges occur between two interacting rats. Afterward, minor problems and their solutions and other considerations concerning the social reinforcement task are discussed that could inform future task improvements. Section 5.3 presents a discussion of the posterior-lateral OFC's role in appetitive unblocking and why it does not play a prominent role in social associative learning. Other brain regions are then examined that do potentially play an important role in driving social associative learning, and experiments are proposed that could test a differentiation between their respective involvements. Finally, a short discussion follows regarding the role of credit assignment in social blocking. For the final section, 5.4, how rats learn to discriminate between various non-social rewards and how these non-social rewards affect social interaction behavior are outlined. Here, the role of ultrasonic vocalizations in social versus non-social reward processing is also discussed. Finally, possible ways to improve the XCST task's analyses are presented.

5.2 Vicarious reward unblocks associative learning about novel cues.

5.2.1 General Implications.

Unblocking social associative learning. It was previously unknown whether rats could learn the value of more complex social stimulus-outcome associations in the positive domain, alongside simpler social Pavlovian reward associations found in social place preference. We find here that rats can indeed learn through vicarious reinforcement, showing a conditioned response to a cue predicting mutual reward delivery. This conditioned response signifies that there is indeed a value difference between the expected outcomes of a standard food reward and the actual outcome of a mutually shared food reward. The most powerful aspect of this difference is that the appetitive self-reward aspect of the surprise is blocked because of previous learning; therefore, what is surprising is purely of an appetitive other-regarding nature. The main conclusion from this study is that observing the partner rat receiving food pellets is more rewarding for the actor rat than the

absence of the partner receiving food because we observed unblocking for the cue that was reinforced with other rewards, but not for the cue that predicts no reward for the other (pCS+ > pCS-). Furthermore, control experiments showed that social learning (i.e., social-information exchange or observational learning) was necessary for the social unblocking effect to occur. While these novel cues had no initial value for the actor rat, during compound conditioning, their value estimates were updated through a social learning process that took into account vicarious other-regarding outcomes, possibly by making use of vicarious prediction errors (Joiner, Piva, Turrin, & Chang, 2017). The actor rat thus optimized its behavior by learning vicariously which environmental cue predicted a valuable social or food outcome and which cue did not. In this way, the actor rat gained the adaptive benefit of finding and remembering the locations and stimuli in the environment better when it observed the partner rat find food (Leblanc & Ramirez, 2020). Our results provide further evidence for Trezza et al.'s (2011) proposal that the behavioral repertoire associated with social valuation can be incorporated with the incentive salience theory (Berridge et al., 2009). To be rewarded together is desired or "wanted" and associated with the incentive salience component of "liking." Meanwhile, observing the other rat not being rewarded is potentially disliked and undesirable. We extended this framework by showing that the social valuation of others regarding outcomes and the formation of social stimulus-outcome associations can be interpreted within the framework of incentive salience theory (Figure 1).

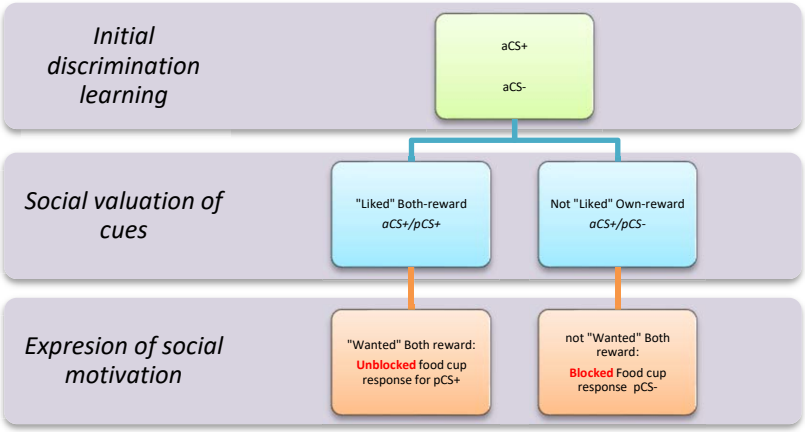


Figure 1. Social unblocking and incentive salience theory.

Socially unblocked cues could strengthen prosocial choice. The finding that a mutual reward unblocks novel cues is an extension of the work by Hernandez-Lallement et al., 2015 and Márquez et al., 2015 that showed that both-rewards are preferred over own-rewards when rats are given a direct choice. This evidence shows that whenever the rats make their choices in these task settings, it is very likely they form their decisions partly based on learning to correctly associate the sensory features of the environment (cues, locations, actions) with the consequence of their actions (self- and other-rewards). How do these sensory cues help strengthen the expression of social preferences, and what are the benefits? Two potential benefits are likely to be present if, during the act of prosocial cues, either reward outcome is indicated by cues. First, the prosocial choice

expression could be strengthened if the presentation of both-rewards would be contingent on the presentation of a flashing light, while the own-reward would be presented together with a green light, or vice versa. The reason that this could work comes from a study by Burke et al. (2008) that demonstrated that if a rat first learns that a cue predicts a grape-flavored pellet when the identity of the taste is then switched to banana-flavored, and a new cue is added, this cue becomes unblocked when presented in isolation. Importantly, they also showed that the rat pressed the lever more frequently for the cue that predicted a change in flavor identity over a cue that predicted the same flavor. It is likely that the social reward-identity characteristics of a mutual reward (partner approach, food consumption, and USV expression) unblock the cue associated with it but not the cue associated with the own-reward, similar to what happens during social unblocking, just as we hypothesized (see Chapter 2's introduction). Notably, the mutual reward choice-action could be enhanced like the enhanced lever press for the unblocked cue during identity unblocking, but via social reward-identity unblocking. A critical test, similar to the trial test during Pavlovian to instrumental transfer (PIT), would present the both-reward together with a flashing light cue during the entry of a prosocial choice. The time to reach the food cup should be shorter in the mutual reward coupled with a flashing light than in another group where the light is not presented. If so, then the social mutual reward's associated value of the conditioned stimuli would enhance the prosocial motor action. Besides this test, crucially, a generally higher preference for the prosocial choice option of both-reward over the own-reward should be observed in the task with the added cue lights, compared to when the task does not have added lights. This experiment could thus help to strengthen the expression of social preference, in addition to performing the proper control experiments such as with a toy rat (Hernandez-Lallement et al., 2015) or the partner rat indicating a preference (Márquez et al., 2015).

Social unblocking: Comparative approach. Next to that, as we discussed in the introduction, this thesis' primary aim is to find the social behavioral patterns similar to humans to cast light on seemingly related behavior in humans. A study by Seid-Fatemi and Tobler (2015) found that vicarious reinforcement was also present when rewards were delivered to other people (see Figure 2). When an added Y cue during compound training predicted a change (going from no social reward to a social reward delivery), participants increased the percentage of key presses for those trials that led to the social reward delivery. When the Y cue that predicted the added social reward was presented alone, participants showed a higher key press for this cue than for an X cue that predicted no social reward delivery. The enhanced key press in this paradigm is similar to the enhanced food cup behavior observed during vicarious unblocking; they both indicate an enhanced motivation to gain the social reward. The social behavior we observed in the rat is,

therefore, meaningful and relevant. There are, however, methodological differences between this vicarious reinforcement in people and the social unblocking effect I observed in rats. Vicarious reinforcement in Seid-Fatemi and Tobler's (2015) paradigm was achieved by first learning a cue that predicted a no-reward delivery for the other and then learning that an added cue predicted a monetary reward delivered to the other. This result, however, is a blocking control and thus different from traditional unblocking, where the cue first predicts a reward delivery and the added cue signals an increase in reward delivery. Blocking was defined as a cue that already predicted the reward delivered to the other. Subsequently, an added cue that predicted a similar reward to the others was blocked. Our task, however, dealt with food outcomes, and the social unblocking effect in rats was achieved differently: first by learning that a cue predicts a self-reward and then learning that an added cue predicts a both-reward (i.e., both actor and partner are rewarded).

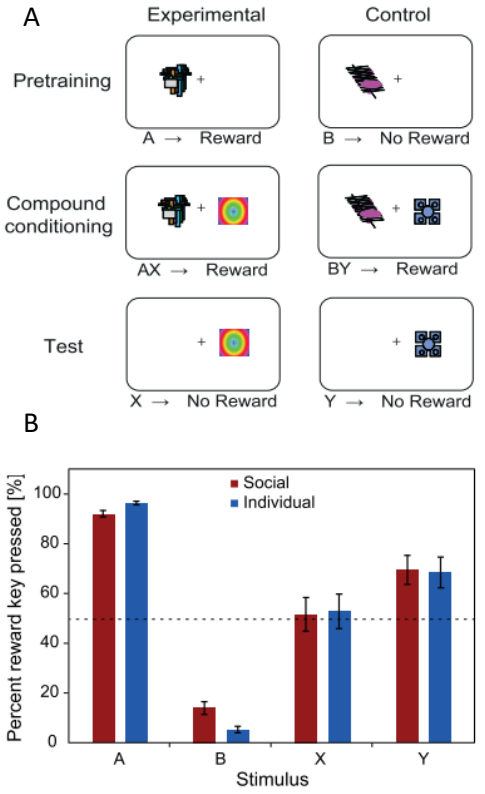


Figure 2. Social reinforcement in humans using the blocking task. (A) Task conditions. Left: blocking, right: control condition that measures vicarious reinforcement. (B) Enhanced key press responding to Y but not X. Reprinted from Seid-Fatemi and Tobler (2015)

Two questions must be answered to advance the comparative approach between animals and humans. *Firstly*, do people also show social appetitive unblocking when a cue added during unblocking predicts a food reward instead of a monetary reward delivered to the other person, in addition to a food self-reward (mutual reward), like what was observed in Chapter 2? *Secondly*, do the rats show unblocking if the paradigm by Seid-Fatemi and Tobler (2015) is translated to rats if the rat first learns by observation that an initial cue does not provide rewards to another rat (with no initial learning for the actor itself; naïve observer), and consequently, an added cue provides a reward to the partner rat (see Figure 2a; control)? Does the rat then press a lever (after learning that it can press this lever to obtain food) that presents the cue associated with a reward delivered to the partner more frequently than a lever that presents the cue predicting a no-reward delivered to the partner? This question is similar to the question raised regarding the unblocked pCS+ cue: Would the actor rat press a lever to receive the pCS+ that predicted a both-reward more often than a lever that presents the pCS- that predicted an own-reward?

Answering these questions will provide a clearer picture of whether social unblocking in the appetitive domain exists for humans. Furthermore, it will provide a clearer answer to whether cues that predict a reward for another rat show vicarious reinforcement value in rats, as has been found in monkeys (Chang et al., 2011).

5.2.2 What is learned, and how does the learning take place? We discuss two possible models that describe how the valuation of mutual rewards works and propose an integrative account for which we consequently describe the important behavioral repertoires involved with rats. Firstly, the reward to the other may be processed entirely separately or processed separately and integrated with the self-reward value. According to Apps et al.'s (2016) motivation and vicarious error model, vicarious processing of another's motivation to obtain a rewarding outcome, consists of three representations that the actor should develop. 1) The actor rat should form an expectation concerning the value of a reward for the partner itself, 2) it should update this reward value of the partner by monitoring their state, and 3) it should update this value through a vicarious error. Ruff and Fehr (2014), conversely, propose three different factors necessary for the valuation process of others: 1) Assessment of how the behavior of the other affects one's own behavior, 2) the vicarious processing of choice and outcomes for the other, and 3) how to guide your behavior to social normative principles. I propose using an additive combination of these two models to explain mutual reward or no-reward valuation during social unblocking. The actor rat should be able to 1) learn the value of outcomes delivered to himself by integrating own-reward prediction errors; 2) represent the value that certain outcomes represent to the partner rat, learn what the partner finds valuable or not valuable, and update this value through observation of reward-related behavior in the partner; and 3) learn an integrative state that represents the value of the reward delivered to the other for the self. Consequently, there are three distinct states that the actor rat either experiences at the same time or sequentially (Figure 3, example for both-reward). The integration of the value of rewards delivered to the partner, the value of rewards delivered to the actor, and consequently the meaning for the actor of these outcomes experienced together, define the mutual reward valuation (i.e., what does it mean and feel like to both be rewarded).

Notably, in social unblocking, the mutual reward valuation process becomes associated with the cues that predict them through, for example, positive 4) self- and 5) partner-related Pavlovian CS-US associations, as was described in the introduction. After learning, the presentation of the social CS in the probe trial elicits the integrated social US outcome. It would drive an approach if the social valuation process led to a positive value of the social outcome (Figure 3, both-reward example) and withhold the approach if the social outcome was not valuable for the self. The valuation of rewards delivered to others and the integration of that value with a self-reward value

could be supported by the social reward-identity signals present in both the actor and partner rat behavior. These behaviors include learning by observing others, transferring affective state or value through the expression of vocalizations, and finally, the exchange of olfactory and tactile information between actor and partner. These are discussed in the following sections.

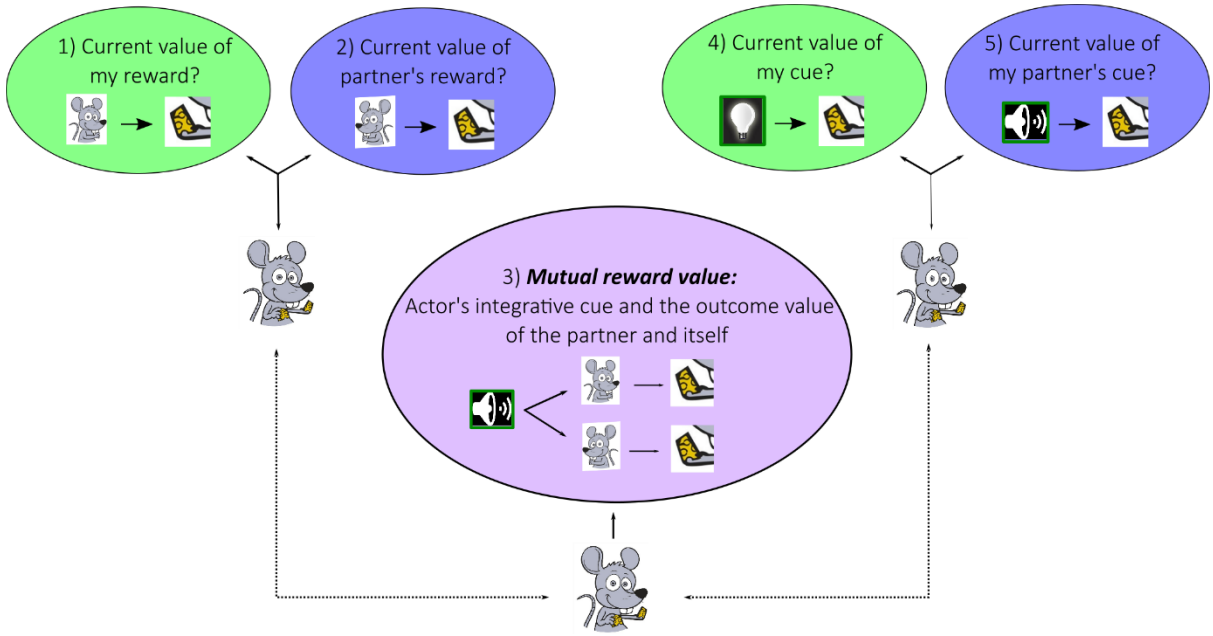


Figure 3. Mutual reward valuation and social unblocking.

Observational learning. Vicarious processing of rewards delivered to the partner rat could be supported by visual observational learning. Previous experiments in rats have shown that, during learning, animals can develop sign tracking behavior in the form of an orienting and enhanced response to the location of the CS or goal tracking toward the location of the US (Cleland & Davey, 1983). This enhanced response is indicative of the cue's enhanced incentive salience. We observed that some animals (both actor and partners) in our task indeed showed cue-directed behavior during the compound phase after the cues turn on; for example, the actor or partner moved their nose toward the cues and then toward the interaction window and subsequently observed the other's food cup approach and entry during the both-reward. This result conforms to the definition of stimulus enhancement in social learning: when the observation of an action leads the observer to increase the proportion of its behavior directed toward the location or object (in our case, the location of the lights and sounds) of the demonstrator's action (Spence, 1937). Therefore, such putative sign-tracking could facilitate learning in the actor rats about the cues and presumably increase sign tracking behavior for the actor rat toward the unblocked cue. As an

example of such stimulus enhancement, Heyes et al. (2000) showed that rats observing other rats performing a lever-pressing discrimination task had a higher lever-pressing rate for a previously observed CS+ lever. In our task, actor rats' response at the food cup might be enhanced through observational conditioning (i.e., by observing the partner's conditioned approach on BR trials compared to OR/NR trials). Trial by trial computational modeling of the influence the partner rat's sign tracking behavior and approach behavior has on the actor rat's actions could lead to further insights regarding the putative role of stimulus enhancement and observational learning in this task.

Auditory transfer of the affective state or value. Our data showed that the possibility for social information exchange was necessary to produce the social unblocking effect. One possibility is that social information was transferred through an auditory transfer of affective state or value between rats. It has been found that rats observing conspecifics being rewarded produce 50 kHz and 22 kHz vocalizations and show an increase in dopamine release (Kashtelyan et al., 2014). Furthermore, playback of 50 kHz vocalizations motivates rats to approach the location of playback and increases their USV production (Wöhr & Schwarting, 2007). During the playback of 50 kHz calls, the dopamine released in the nucleus accumbens is transiently increased (Ingo Willuhn et al., 2014). In our task and instrumental choice tasks, the mutual reward value could be processed through the expression, perception, and valuation of 50 kHz and 22 kHz, which do not need to be visually attended. Prosocial choices in rats, as found by Hernandez-Lallement (2015) in the PCT (Hernandez-Lallement et al., 2015), are therefore likely to be driven by communicative social reinforcement signals (Hernandez-Lallement et al., 2016a), such as 50 kHz or 22 kHz vocalizations expressed by the partner rat or actor rat. Recently, it has been found that suppressing VTA dopamine transients through optogenetics (definition, see Table 1) during the compound phase of a non-social blocking task impairs the unblocking of novel cue values and features (Chang, Gardner, Di Tillio, & Schoenbaum, 2017).

Furthermore, increasing VTA dopamine activity can drive learning about novel artificial cue-reward associations in an unblocking paradigm (Keiflin et al., 2019). These findings, when taken together, suggest that unblocking in our task could occur through social reinforcement via 50 kHz calls and associated dopamine release. Some evidence that 50 kHz can enhance the value of novel cues comes from a study by Saito et al. (2016), which found that the playback of 50 kHz ultrasonic vocalizations 20 minutes before hearing a neutral tone increased rats' lever presses during the subsequent presentation of the neutral tone more so than in a control condition where there was no playback. The number of lever presses became more similar to the number of lever presses observed in response to a previously conditioned rewarding stimulus (positive bias), indicating that

the playback of 50 kHz USVs induced a positive affective state that transferred onto the neutral stimuli, thereby enhancing its value. In our task, positive reinforcement of the partner can be perceived as both surprising and rewarding to the actor rat, which could increase the partner's or target's 50 kHz vocalizations. Future work would indicate whether the production of 50 kHz USVs (unpublished observations) during the BR trials in the compound phase can contribute to dopamine release in the actor rat and thus drive learning (i.e., unblocking) of the additional compound cue associated with partner rewards. Preliminary analyses of the 50 kHz USV's expressed during a both-reward, however, suggest that these calls are not produced often (one or two per trial and mainly on the first trials of each day of compound training), and many later trials show a complete absence of this call type. These findings suggest that the reinforcement value of 50 kHz vocalizations could be low. However, these observations were done in only four rats and due to the high inter-individual difference in call production between rats, further analyses of call production and more specifically how the call production is related to an approach behavior, social interaction, and food consumption, should provide a more definite answer on its role in social unblocking.

Olfactory and tactile information exchange. Finally, social information exchange about partner rewards could have occurred through olfactory and tactile information exchanges. It has been found that in rats, food preferences can be transferred from demonstrators to observers. For this effect to occur, the scent of the food together with the presence of carbon disulfide (CS2) on the demonstrator's breath is necessary (Galef, 2001). Preliminary observations of the interaction between our rats in this task showed that facial touch occurred throughout all compound conditions in the interaction window (Chapter 2, Figure. 1A). The presence of pellet consumption-related olfactory cues and the associated CS2 on the partner's breath could thus have aided the target rats in identifying BR trials, while an absence or decrease in these indicator smells may have indicated OR NR trials. In addition, it has been found that facial touch is associated with an increase in the production of 50 kHz calls (Rao, Mielke, Bobrov, & Brecht, 2014b). Within a BR condition, where sniffing of dyadic olfactory cues might be enhanced, social face touch could potentially enhance the production of 50 kHz calls further and thus contribute to unblocking.

Mentalizing and empathizing. The identification of rewards delivered to others by the above-mentioned factors could give rise to a valuation process indicative of something that transcends the valuation of rewards delivered to others. This higher-order valuation process could rely on a mentalizing or empathizing ability by which identification of the mutual, cooperative reward has its own value (Rilling et al., 2002). The positive experience of mutual reward consumption could, in this case, be explained by using the perception–action model in which the rat senses the

emotional state of the other's experience (de Waal & Preston, 2017) by way of emotional contagion that arises during social observational learning, auditory transfer, and olfactory and tactile information exchange. This state would be mapped onto the actor rat's own state spontaneously, driving the formation of the positive vicarious emotions experienced during mutual reward. It is, however, hard to disentangle if such a positive vicarious emotion is indeed elicited during a both-reward and possibly indicative of a willingness to share food with others. One possible way to further understand the emotional state that the rat experiences during a both-reward and an own-reward or no-reward is by looking at facial expressions and their associated whisker positions. That this is indeed possible and important comes from a recent study in mice, where the facial expression of the mouse was analyzed using computer vision. They found that the facial expressions when drinking sweet sucrose and bitter quinine, for example, showed different facial expressions that were indicative of the mouse's internal state (Dolensek et al., 2020; see Figure 4). Importantly, they found that these facial expressions as an indicator of emotions depended on the mouse's internal state, valence, and the learned value of stimuli. This finding suggest that a video recording could be made of the actor rat's facial expressions while observing the partner rat receiving a reward. The rat's facial expressions could potentially indicate which internal state the rat was feeling during observations of mutual reward. This observation would be possible in a dual head-fixed setup, similar to the setup that measures vicarious reinforcement in monkeys (Chang et al., 2011). If the observed expressions during observations of mutual reward were positive, it might show similarities with the sucrose consumption. The actor rat might experience disadvantageous inequity aversion during the partner-only reward, which possibly bears similarities to the aversive Quinine facial expression. Importantly though, as the above example is hypothetical, a clearer readout of emotional contagion would be possible if, for example, the actor rat learned through observation that one cue light (c1) would lead to the satiation of the partner rat with one type of reward. In contrast, another cue-light (c2) would lead to a reward presentation of another equally preferred type of reward (but not satiated). If the actor rat learned the partner's internal state, the cue's presentation would show a happier rat (as indicated by its facial expression) for the c1 cue over the c2 cue. Finally, if the actor

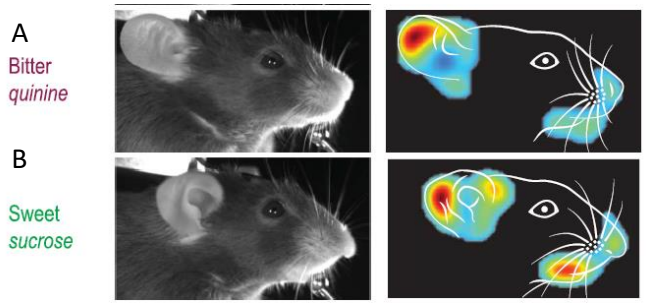


Figure 4. Analyzing the mouse's internal state by looking at facial expressions. Examples showing the facial expressions when drinking something (A) bitter or (B) sweet. Reprinted from Dolensek et al (2020).

could press a lever to obtain the cue light, it would have to show an enhanced lever press for c1 over c2.

Overall, a mentalizing or empathizing ability could, with the help of new experiments and advances in the detailed analyses of specific behavior such as facial expression, lead to further insights into the valuation of rewards delivered to the partner through the actor's expression patterns. Improvements and considerations concerning the social reinforcement learning task are presented in the next section.

5.2.3 Necessary improvements and considerations for the social reinforcement-learning task. I will address here the issues that could influence the reliability of the social unblocking effect. The first issue I discovered was by looking carefully at the literature. It becomes clear that Peter Holland's lab (Chang et al., 2012; Holland, 1984) mainly used a between-subject task configuration, where the control group was a different group of rats. Conversely, McDannald et al. (2011) and Tobler (2005) used a within-subject model where the experimental and control conditions were randomized within a group (Figure 5 A, B).

Task design. In the social unblocking task used in Chapters 2 and 3 of this thesis, based on the approach by McDannald et al. (2011), I have also used the within-subject approach. The within-subject approach during social unblocking created one potential problem due to the addition of the no-reward condition. We initially decided on adding this condition because, in the McDannald et al. (2011) paper, the cues were presented separately during the compound phase as reminders. Instead, we decided that the no-reward condition would be a good reminder of the learned CS-response (Figure 5 C). ***Observation.*** However, Rescorla and Colwill (1983) found that if one first learns that the A light cue predicts two shocks and then a tone Y is added that predicts only one shock, there would be a decreased unblocking response (diminished conditioned suppression) if non-reinforced trials of the light presentation are added in between.

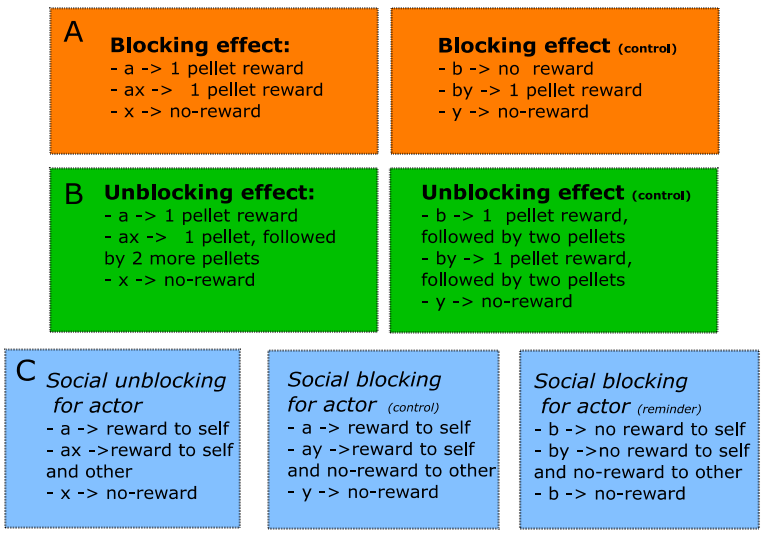


Figure 5. Design comparison. (A) Blocking effect (B) Unblocking effect (C) Social unblocking effect Adapted from (A) Tobler (2005) (B) McDannald et al. (2011) (C) van Gurp et al. (2020)

These results showed that a within-compound association of the A cue's associated value changed the Y cue associability and influenced the strength of unblocking in addition to changing the US (moving from a double shock to only one shock).

Problem. In our task, it is possible that the coupling of the no-reward aCS-learned value could have transferred to the additional pCS-. In this case, it could be that blocking was caused by a no change in value during own-reward. Simultaneously, during no-reward trials, the aCS- value might have caused an inhibitory influence on the pCS- cue's value. However, this is unlikely, as pCS- responded at least descriptively higher than aCS- responded.

Solution 1. What would have been a better option was to show the cues not simultaneously but separately, thereby not allowing an association to form due to the absence of contingent presentation. A direct experiment that could test this hypothesis is task configuration where the own-reward is still a compound with the partner no-reward cue (thus leading to the blocking effect), but the additional no-reward presentation would be left out or only shown with a non-contingent presentation of aCS- and pCS- separately.

Solution 2. Another option would have been to have an experiment where blocking and unblocking are uncoupled (i.e., the both-reward and own-reward learners are actually separate groups or separate learning episodes). In this case, there would have been no spill-over effect from social learning on the aCS+ and pCS- cues associated with two different social outcomes. Please note that this would not affect the crucial pCS+ that is only present in BR but could address a potential "value surplus" in the BR through a concurrent reduction in aCS+ value by virtue of it also being associated with a partner no-reward in the OR compound.

Secondary reinforcement. The second improvement that could be made refers to the addition of pellet dispensers. ***Observation.*** The additional pellet dispenser was put outside the Phenotyper, on the other side, where the partner rat would be rewarded during the compound phase. Already during the discrimination phase, pellets were dispensed during an aCS+ trial into a soft cup to diminish sound so as not to introduce a difference in pellet-delivery-related sounds that could act as secondary reinforcers, irrespective of whether a partner rat was collecting the pellets.

Problem. It is still unclear how much the dropping of pellets on the partner's side influenced discrimination learning and consequent unblocking. There was a descriptively lower unblocking for the group where these pellet dispensers were placed versus the one with no dispensers. This difference was tested in two different experiments. It would, however, be better to directly compare the two different groups to see if the inclusion of an added pellet dispenser influences

the unblocking effect. Besides that, we used soft cups to catch the pellets, so the actual sound of a pellet falling differed from when a pellet typically falls in the food cup. Significantly, for the control experiment with no rat present, we changed the pellet drops in such a way that the pellet would fall through a customized food cup (preventing the stacking of pellets in view,

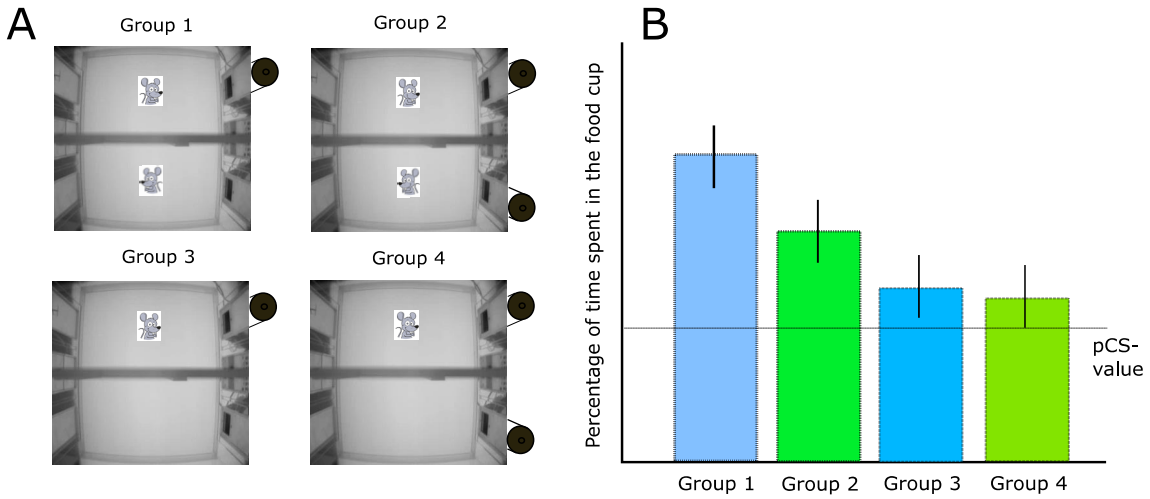


Figure 6. Secondary reinforcement experiment. (A) Setup configuration with the presence and absence of a pellet dispenser. (B) Hypothesized levels of social unblocking for the pCS+. The horizontal line is the hypothesized pCS- level.

but retaining the distinct sound related to the pellet falling on the metal food cup).

Solution. Therefore, it would be best to do the social unblocking experiment with four groups and use the configuration with the customized food cup (Figure 6) to investigate pellet dispenser sounds' influence during social unblocking precisely. (Group 1) Social unblocking without an added pellet dispenser during discrimination, but with a rat present to receive rewards during the compound phase. (Group 2) Social unblocking with an added pellet dispenser during discrimination and a rat present to receive rewards during the compound phase. (Group 3) Social unblocking without an added pellet dispenser during discrimination and with no rat present to receive rewards during the compound phase. (Group 4) Social unblocking with an added pellet dispenser during discrimination and with no rat present to receive rewards during the compound phase. The unblocking strength is hypothesized as being in the following order: Group 1 > Group 2 > Group 3 > Group 4.

Task benefits. That the actor rat has a preference or forms a valuable stimulus outcome association for a both-reward over an own-reward is most likely only adaptive if benefits outweigh the cost (Galef & Laland, 2005) to attend to the places (and the associated features) where and when the partner rat finds food. The benefits depend on whether the paradigm can be transformed to acquire ecological validity (i.e., how does the test performance predict behavior in a real-world setting; Gouvier et al., 2014). Arbilly and Laland (2013) used a modeling approach to demonstrate that local and stimulus enhancements (i.e., the probability that an actor observes a partner receiving a reward at a specific location or is contingent on stimuli present) can quickly become extinct when previously rewarding places in the environment suddenly no longer provide food. They also found that social learning is ineffective when it is difficult to remember the place with the highest probability of giving a food reward because there are too many locations to explore with too little time. In other words, if there is one place where the partner can find food with a high probability, the partner rat will most likely still explore other places for food and then has little time to return to the high probability food location. Therefore, social reinforcement of the high probability location would be limited. Arbilly and Laland proposed that for social learning to be effective when locations always have a different sensory context, the locations must be limited and the location with the highest reward probability must be significantly better than the other locations. If the actor observes the partner reliably acquiring food, then local and stimulus enhancements could take place. Ecological validity could be achieved if rats were observed in a natural habitat (possibly emulated in the lab) and the probability of an actor rat observing a partner rat receiving food in the presence of discriminable stimuli should then be calculated. These probabilities (instead of the fixed pseudo-randomized probabilities used in the current social unblocking task) must then be used in the lab and replicated in many more rats to show the strength of the social unblocking effect in a real-world setting. This result would indicate precisely when the benefits outweigh the costs for the actor rat to follow the partner rat and observe it finding a reward.

Influence of food deprivation. Furthermore, the rats' food deprivation during the task must also be taken into account. The cues that they both learn to predict the food for each other could potentially induce a motivational state or craving for the food that the other has because of their food deprivation state. This craving is present in a cue-induced feeding paradigm where rats learn that specific cues predict food under food deprivation. Consequently, when these cues are presented when the rats are sated, they consume more food (Petrovich, 2011). In our case, the cue represents another rat eating that could very likely trigger a similar sort of craving born out of a hunger for the other's food while being adaptive to learn where the other finds food, as it costs more energy to learn this all by itself. We find, however, that this is not the case because when an

added cue during unblocking predicted a partner but not an actor reward, this cue remained blocked (see Chapter 2, control group 2). The actor rat likely did not learn to approach this cue because rats show an aversion toward outcomes where a partner gets a greater reward than the actor (Oberliessen et al., 2016). This disadvantageous inequity aversion, therefore, counteracts the effect of food deprivation on learning. However, this does not preclude that there might be a general enhancement effect of food deprivation on both-reward learning. A simple test could be to have two groups, one food-deprived and one not. In both groups, actors and partner rats first learn that one cue predicts a both-reward and one cue does not. After learning, one would show both cues without the presentation of the both-reward to the actors and partners. If food deprivation has a general effect, then the cue predicting a both-reward would have a higher food cup response on the first trial or during the extinction than the non-food deprived cue.

5.3 Posterior-lateral orbitofrontal cortex lesion leaves social and appetitive unblocking intact but impairs the blocking effect related to partner no-reward.

5.3.1 General Implications. I hypothesized that rats learn the value of more complex social stimulus-outcome associations by using social reward-identity features that signal the distinction between both-reward and own-reward. These signals include the actor's observation that the partner is rewarded (observing the partner's approach behavior directed at the food cup), consumes the food (hearing the consummatory sound), and potentially the exchange of olfactory-related information (scent of food on the partner's breath) in both-reward but not own-reward. This hypothesis was formed because it was found that an inserted wall or partner absence during mutual reward presentation impaired the social unblocking effect. This finding would suggest that the rat forms a model-based cognitive representation of the actions and outcomes delivered to the partner rat to learn the value of mutual reward outcomes and demonstrate unblocking to the cues predicting this outcome. We hypothesized that these signals could be processed in the lateral OFC, which has been found to form a model-based cognitive representation of identity features of outcomes in both the social (Azzi et al., 2012) and appetitive domains (McDannald et al., 2011). Thus, (social) reward-identity signals in this region could support social unblocking similar to how they support identity unblocking (McDannald et al., 2011). We found, however, that the posterior-lateral OFC is not involved, as lesioning this region before the compound phase did not impair the social unblocking effect. These findings suggest that social reward-identity features of the partner rat's behaviors related to rewards delivered to others are not encoded in this region and do not drive the social unblocking effect. Moreover, it is possible that the social reward-identity signals are completely unnecessary, and simpler model-free learning would suffice to show social unblocking.

However, below I propose that social reward-identity signals are necessary but that the PLO is not involved because it only supports appetitive identity unblocking based on a specific appetitive circuit related to coding changes in gustatory flavor-taste or odor-taste identity signals. In contrast, social reward-identity signals related to the valuation of rewards delivered to others are more likely to be processed in a circuit that includes the prelimbic and anterior cingulate regions of the medial prefrontal cortex and the VTA to NAcc circuit and do not include the PLO. However, we propose ways in which the lateral OFC could be implicated when both appetitive identity-specific action-outcome associations are coupled with social stimulus-outcome associations.

5.3.2 Appetitive versus social unblocking: Different neural circuitry? Identity-based unblocking is thus dependent on the lateral OFC when the change in outcomes is related to the taste identity (banana or grape; McDannald et al., 2011) or the (valueless) odor's identity (Lopatina et al., 2015). The origins for this modality-specific coding come from a recent study that investigated the input connections to subparts of the lateral OFC. Barreiros et al. (2021) found that posterior-lateral OFC receives the most input from the BLA, mediodorsal thalamus, anterior and agranular insula, and piriform cortex but not from the anterior cingulate cortex area 24a and all sensory areas. The PLO receives mostly gustatory information from the anterior and granular insula and olfactory information from the piriform cortex. Importantly, the agranular insular cortex signals the degree of familiarity to specific tastes (Bahar, Dudai, & Ahissar, 2004), while the anterior insula contains neurons that encode the intensity of the sucrose reward (Fonseca et al., 2018). This information can be sent to the PLO and contributes to forming identity-specific stimulus-outcome predictions. Flavor-taste associations are furthermore formed in the BLA. For example, when rats learned to associate a tangerine or kiwi Kool-Aid flavor with a fructose outcome (CS+) or no fructose (CS-), lesions of the BLA impaired this flavor-taste learning (Dwyer, 2011). This result adds to the fact that the BLA is only involved in unblocking when the reward's identity is changed to a different flavor (moving from food pellets to an 8% sucrose water reward) but not when changing just the pellet amount (Chang et al., 2012).

PLO neurons further contribute to this kind of model-based flavor-taste learning, as firing rates changed when the identity of a vanilla milk reward changed to a new chocolate milk reward that had not been experienced yet. A new identity was signaled (Stalnaker et al., 2014), possibly due to the processed taste- and flavor-related information coming from the BLA and granular and anterior insula. However, social reward-identity feature changes that signal rewards delivered to others require (processed) visual, auditory, or tactile information, and the PLO does not receive this input. A more likely cortical candidate involved in coding the social reward-identity signals necessary for social unblocking is the subparts of the medial prefrontal cortex (mPFC), such as the

anterior cingulate cortex (ACC; area 24) or the prelimbic region (PL), and the BLA. Lesions to the anterior cingulate cortex (including areas 24a and b) have been found to decrease the social interaction time between two adult rats from different cages (Rudebeck et al., 2007).

In a recent article by Guo et al. (2019), the authors studied the role of the SHANK3 gene in the anterior cingulate cortex. Single mutations to this gene are a causal factor that leads to autism. They found that deleting the SHANK3 gene, specifically in the anterior cingulate cortex (area 24 a/b), caused the mice to spend less time with a stranger mouse in the three-chamber social preference test. They then demonstrated that mice whose ACC pyramidal cells were optogenetically stimulated (after inserting channelrhodopsin-2) spent more time with a stranger mouse and initiated more interaction, than the mice exposed to a control stimulation. Notably, besides coding only social interactions, it was also found that these neurons signal the competitive effort to obtain food in a social situation. In this competitive effort task by Hillman and Bilkey (2012), rats could either 1) eat a small two-pellet reward without competition for another rat or 2) 12 pellets were given on the other side of a maze, of which both rats had access to and competed with each other to gain access. In this situation, the rats developed no behavioral preference for the higher reward number, but cells in the anterior cingulate cortex nonetheless increased their firing rate when the rat approached this highly competitive outcome.

Another region ventral to the anterior cingulate cortex that could likely be involved in social unblocking is the prelimbic region of the medial prefrontal cortex (mPFC). Lee et al. (2016) found that, in a three-chamber social preference test where mice preferred to explore a stranger mouse in one chamber over an inanimate object in another chamber, a significant proportion of cells in the mPFC prelimbic area enhanced their firing rate during social interaction, but not when investigating the inanimate object. Besides a direct approach, a study by Teresa Jurado-Parras, Gruart, & Delgado-García, 2012 found that observational learning of an instrumental lever press by a trained demonstrator rat for food improved the criterion to reach twenty lever presses for twenty food pellets in the actor itself. The researchers found that the actor rat's observational learning was impaired when they electrically stimulated the PL in the observer rat at the critical moment when the demonstrator pressed the lever. Finally, another crucial region is the BLA as it was found that BLA lesions abolish prosocial choice behavior (Hernandez-Lallement et al., 2016b), socially transmitted food preference in rats (Wang et al., 2006), and reduce the response to 50 kHz playback (Schönfeld et al., 2020) in rats.

Conversely, the BLA is important for multimodal learning as its connections to the PLO are also crucial for taste-flavor identity learning, while at the same time, activation of prelimbic cortex

projections to the BLA impairs social place preference learning (Huang, Zucca, Levy, & Page, 2020). These results all point to the anterior cingulate cortex areas 24a and b as having an essential role in calculating the value of direct social interaction and social effort. In contrast, the prelimbic region of the medial prefrontal cortex plays a vital role in social approach and social observational learning of action-outcome association. The BLA is important for both social and appetitive outcomes and identity-related information, possibly integrating social-appetitive outcomes.

5.3.3 Interaction between social and appetitive learning. We hypothesized in Chapter 2 that social unblocking depends on social reward-identity signals related to observing social approaches and consummatory responses during a both-reward and the transfer of olfactory information of food-related information after a BR that needs direct social interaction. It is, as discussed above, highly likely that the PL, ACC, and BLA could encode social reward-identity value signals and aid in the social unblocking of novel stimulus outcome associations. Notably, the posterior ventral orbito-frontal cortex (VO) but not the posterior-lateral orbito-frontal cortex (PLO) receives projections from both the anterior cingulate cortex area 24b and the prelimbic cortex regions (Barreiros et al., 2021), and vice versa, the PL and ACC receive projections from the VO and MO but not the PLO (Hoover & Vertes, 2007). It is, therefore, possible that instead of PLO, the VO/MO can receive social reward-identity information necessary for social unblocking. A social version of the outcome-specific versus general Pavlovian-to-instrumental paradigm (Social PIT; general Figure 7A, B and C) could be used to dissociate the role of the PLO and the ACC. The original paradigm could be adapted in a specific PIT so that the actor rat first learns that a tone predicts a both-reward and a clicker predicts an own-reward (Figure 7A). In this specific social PIT, the actor rat needs to form a cognitive representation of social outcome, mapped onto a direct appetitive action. Afterward, the actor rat learns that a left lever press gives it a food pellet for both the actor and partner and a right lever press results in a pellet reward the actor only (Figure 7B). The crucial test comes when presenting the BR cue simultaneously as the rat pressing the left lever and presenting the OR cue when the rat presses the right lever (Figure 7C). Figures 7C, D, and E show the hypothetical predictions that I think will occur when lesioning the PLO compared to lesioning the ACC/PL. I expect to find that the PLO lesion only impairs the appetitive aspect of the enhancement when presenting the OR cue during the right lever press, similarly as has to what has been found in specific non-social PIT (Ostlund & Balleine, 2007). It would not impair the social aspect of the BR associated left lever press enhancement (Figure 7B). The ACC/PL lesion would impair only the social aspect of the BR cue-induced enhancement of the left lever press but not the appetitive OR cue-induced enhancement of the right lever press (Figure 7B). These hypothesized results are derived from the fact that the ACC is not involved in specific non-

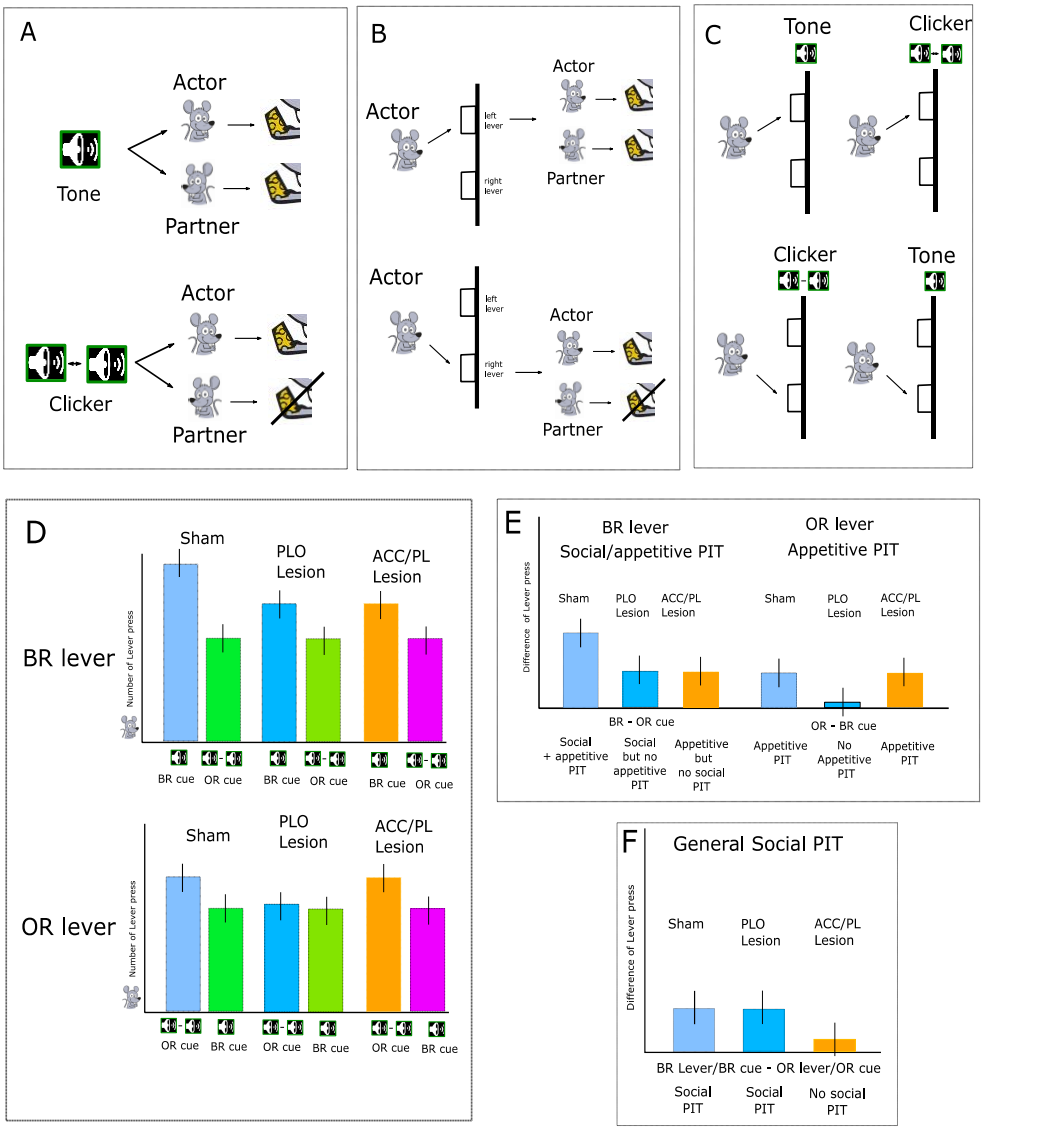


Figure 7. Outcome-specific social Pavlovian-to-instrumental paradigm (PIT) and hypothetical lesion effects on social PIT. (A) Actor learns that a tone predicts a both-reward and a clicker an own-reward (B) Actor learns that a left lever press results in a both-reward and a right lever press in the own-reward. Presenting the tone during the lever press (C) results in the actor's enhanced response, compared to when the clicker is presented for the left lever in a sham group and that social/appetitive PIT is higher than a PLO and ACC/PL lesion group (D; upper bar plot). Presenting the clicker during the right press (C) results in the actor's enhanced response, compared to when the tone is presented for the right lever in a sham group and that this appetitive PIT is impaired in the PLO lesion group, but not the ACC/PL lesion group (D; lower bar plot). (E; BR lever) The bar plot shows the difference between the BR cue's and OR cue's influence on pressing the left BR lever. The sham group shows specific social/appetitive PIT, the PLO lesion group shows social only PIT, the ACC/PL lesion group shows appetitive only PIT. (E; OR Lever) The bar plot shows the difference between the BR cue's and OR cue's influence on pressing the right BR lever. The sham group shows specific appetitive PIT, the PLO lesion group does not show appetitive PIT, and the ACC/PL lesion shows appetitive PIT. (F) Shown here is the general Social PIT expressed as the difference between the BR cue enhancement of the BR lever—OR cue enhancement of the OR lever.

social PIT (Cardinal et al., 2003) but has been found to be involved in vicarious reinforcement as discussed above. Here, the ACC lesion impairs general social enhancement, while the PLO lesion only impairs general appetitive enhancement (Figure 7E). Overall, a social PIT paradigm could be used to further establish the different roles of these brain regions in vicarious reinforcement learning.

5.3.4 Involvement of the Dopamine system in social learning. Alternatively, it could well be that the actual association between the stimulus and the reward delivered to the other is, next to being driven by the observational learning of social approach, learned by enhanced attention to environmental cues because of the other's reward delivery. In a task by Teresa Jurado-Parras et al. (2012), it was found that electrical stimulation of the NAcc in the observer rat, done while it is observing the demonstrator make a lever press to obtain food from a feeder, caused the observer rat to spend more time investigating the feeder and the lever in its first training session after the observation learning. This result and the observation of synchronized orienting to the stimuli associated with a both-reward (unpublished observation) during social unblocking suggest the involvement of the dopamine circuit. Observation of rewards delivered to a partner rat can increase the dopamine release from the dopamine-producing neurons in the VTA (Kashtelyan et al., 2014). This social observation-related dopamine release could potentially transfer predictive and social motivational properties to the cues present.

In a study by Saunders et al. (2018), dopamine neurons in the VTA were optogenetically activated in temporal association with a cue. They found that these stimuli then become conditioned as the rats would show approach behavior toward the cue when presented, and the cue would evoke dopamine neuron activity on its own. These results, coupled with the fact that dopamine activation in the VTA can cause unblocking (Keiflin et al., 2019), the activation of VTA to NAcc projections causes an increase in social interaction (Gunaydin et al., 2014), and NAcc stimulations can enhance behavior directed at reward locations during observational learning (Teresa Jurado-Parras et al., 2012), make it likely that the dopamine system plays an additional role in the social unblocking effect. Thus, the overall implication is that social unblocking likely relies on a circuit involving the ACC, PL, and the VTA to NAcc connection and possibly the VO/MO but not the PLO unless, perhaps when an outcome-specific identity is modulated.

5.3.5 Role of credit assignment in social blocking. Finally, I observed enhanced responses to an added cue in lesioned animals that predicted no additional reward to a partner but still gained associative value. The observed conditioned response that was a consequence of an impaired blocking was discussed in the framework of improper credit assignment because of the inherent

uncertainty in the partner CS- associated with both an own-reward (aCSp, pCS-) and a no-reward (aCS-, pCS-). Credit assignment, as discussed in Chapter 3, means assigning the proper outcomes to the right choice. This process becomes even more challenging if one must integrate a history of multiple stimulus-outcome associations presented in a randomized manner. As shown by Walton et al. (2010), lateral OFC lesions lead to a more pronounced reliance on the recent integrated history of reinforcement but not on the *specific* contingency of which stimulus caused which outcome, leading to the spread of effect. In the case of the actor rat that reacts to cues that predict no additional rewards delivered to others, this response has an additional cost for the actor rat but no benefits. Notably, there is the idea that there are eligibility traces that consist of memories from previous states and actions, which Seo and Lee (2010) discussed are incorrectly assigned to action and cues that do not predict rewards. In the case of reversal learning, when the CS+ becomes the CS-, the animals with a lateral OFC lesion will respond longer to the CS- because of an eligibility trace of the CS+ that is persistent. The actor rat with a lesioned posterior-lateral OFC could certainly have an eligibility trace of the aCS+ in the own-reward condition (aCS+, pCS-) that is still active in the NR (aCS-, pCS-) and thereby transfers aCS+ value and thus incorrectly keeps the pCS- value at a higher level.

5.4. Distinct profiles of 50 kHz vocalizations differentiate between social and non-social reward approaches and consumption.

5.4.1 Sucrose discrimination. We first established that adolescent rats could learn to distinguish between different sucrose levels (2% versus 5%, 2% versus 10%, and 5% versus 10%) as they spent more time in higher sucrose concentrations. The task by Fonseca et al. (2018), as explained in the introduction, found that rats form such preferences during discrimination learning because they can taste the sucrose intensity. In their study, they also found that the latency to stop licking increased with the higher sucrose concentrations, while at the same time, the rats' licking rhythm did not change. From this study, it also becomes clear that there is a linear increase in the time it took the rats to stop licking, with 10% > 5% > 2%. These results mirror our results of more time spent in the arm with a higher concentration. We, however, extend this result by showing that over 3 days, this sucrose place preference is enhanced, indicating that either the 10% sucrose became more valuable after multiple visits or the lower sucrose levels become less valuable, or that the rats had a more accurate mental map and favored exploitation over exploration of the maze.

5.4.2 How does sucrose-discrimination learning take place in the maze? Rats used fast spatial learning in our task to associate the different outcomes with the different arms. This spatial learning process, similar to within the T-maze, likely involved assessing reference memory (i.e., the

direction and motor sequence associated with the direction that the rat takes when entering a specific arm) and possibly using the surroundings' remote sensory cues (Wenk, 1998), but not local cues since the maze arms remain the same across multiple days. While spatial reward learning would be expected to be strong, tactile texture learning would be considered marginal as it takes roughly 300 to 500 trials to create a texture-reward association (Guić-Robles, Valdivieso, & Guajardo, 1989), which is not possible in our task. However, it is possible that, during a 10-minute exploration for six trials (one session per day), the texture within one arm (fixed position) was crossed multiple times. These multiple crossings could at least contribute to learning, while rats that established discrimination between two textures need only three touches to decide whether the texture will provide a reward or not (von Heimendahl, Itskov, Arabzadeh, & Diamond, 2007). Furthermore, the rats learn to associate these different spatial texture stimuli with the various sucrose levels. During downshifts in unblocking, rats spend less time in the food cup, while upshifts increased the time spent in the food cup during the probe trials (Lopatina et al., 2015) when an added olfactory cue during a compound phase signaled a decrease or increase in the outcome. This result aligns with reinforcement learning theory, which states that a negative prediction error can cause inhibitory learning while a positive prediction error can cause excitatory learning (Rescorla & Wagner, 1972). On the maze the rat encountered negative and positive prediction errors when traversing the maze going from a lower to a higher sucrose and vice versa and that these prediction errors indeed can drive excitatory learning for drinking the 10% sucrose learning after drinking 5% and 2% due to the higher prediction error that occurs. The higher sucrose level likely caused a "liking" response consisting of positive hedonic facial expressions such as rhythmic tongue protrusions (Berridge & Kringelbach, 2015). These "liking reactions have been found to increase in a mouse when going from drinking a 1% to 4% to 20% sucrose drink (Dolensek et al., 2020) and are therefore a good candidate for the received US value component necessary to calculate the excitatory prediction error (received—predicted) that drives learning. The stronger spatial-texture sucrose association of the 10% sucrose thus resulted in a stronger sucrose place preference than with a 5% or 2% sucrose concentration.

5.4.3 Sucrose intensity and 50 kHz vocalizations. As the rat spent more time in the 10% sucrose and it is clear that the rat can sense the sucrose intensity and the hedonic value is high, we expected it to vocalize more during the 10% than the 2% or 5% sucrose. We, however, found that rats vocalize more than 50 kHz in the lower 2% and specifically trill. Some insight into this comes from a study by Brenes and Schwarting (2014), where they also found that during food deprivation, 50 kHz vocalizations did not differ when drinking sweetened condensed milk or tap water. Strikingly, the rats started to vocalize more when drinking the sweet milk after the rats were put

back on ad-libitum (full) food access. Furthermore, the researchers showed that food deprivation directly diminishes the rats' vocalization rate when going from ad libitum to food deprivation. During the SDT test, we observed a smaller amount of vocalizations than the SSPT (in which rats were fed ad-libitum). The food deprivation may have diminished the vocalizations expressed during the SDT task. Brenes and Schwarting (2014) have shown more vocalizations during food deprivation for the tone that predicts the sweet milk over the tone that predicts water. In the SDT vocalization score, we only included the period of drinking, but not the vocalizations during the approach behavior, before the start of drinking, and social interaction. We potentially missed the evoked USV's associated with the conditioned spatial and texture stimulus that predicts the 10% by not including this period. These results by Brenes and Schwarting (2014) suggest that it is possible that attributing incentive salience to cues that predict a hedonic 10% might still be wanted more and associated with a higher expression of 50 kHz calls during cue presentation. At the same time, the food deprivation state could inhibit calls during consumption.

The following example is somewhat hypothetical; nonetheless, I would still like to discuss it. Steiner and Redish (2014) have found that rats who left early on a low auditory cost (short delay) highly valued reward trial and subsequently encountered a high auditory cost (long delay) low valued reward experienced regret, which was visible by the fact that the rats looked back at the previously missed choice. When spatial and potentially texture cue-outcomes and spatial action-outcome reward associations were formed in our task, it is also possible that the rat experienced regret after leaving the highly valued 10% sucrose and subsequently encountered the 2% or 5% sucrose. Vocalization results show that in the SDT tests, there was an increase in 50 kHz vocalizations in lower sucrose concentrations. It is possible that the surprise of lower-level sucrose after a visit to the higher-level sucrose caused regret, and consequently, the rat would look back and think of the 10% sucrose. A wanting for the 10% sucrose could, at that point, induce the enhanced expression of 50 kHz calls (primarily represented in the expression of a trill). Importantly though, we do not know yet if regret is induced here, but it is clear from Steiner and Redish (2014) that it is possible. Another result from Coffey et al. (2013) showed that rats, when going from a stimulus that predicts a reward with a probability of 100% to a stimulus that predicts a reward with a probability of 25%, decrease their head entries in a food cup and express an increase in 22 kHz vocalizations. They discuss this in terms of a negative mood state produced by prediction errors. It is possible that in the SSPT task, rats, next to the positive wanting USV's for the 10% percent, also expressed 22 kHz calls (as we do observe those calls but did not include them in the analyses) when going from 10% to 2%.

5.4.4 Sucrose versus social interaction. In a social place preference experiment by Calcagnetti and Schechter (1992), social interaction of one adolescent rat with an active partner was first paired with tactile and visual cues in one room on one day. The rat was then placed in another room with different cues and paired with an inactive partner. Afterward, the rats showed a social place preference, which is indicated by the fact that they preferred to spend time in the room where the socially active partner was, even though the active rat was not there at the time of the test. During the direct choice in the XCST task, between a juvenile and sucrose water, we found that interaction with a juvenile was preferred when the sucrose concentration was low, while the sucrose was preferred over the juvenile when the sucrose concentration was high (10%). In our task, the social interaction of the adolescent rat with the juvenile rat thus caused a similar social place preference to be established. This preference for the social interaction over sucrose consumption indicates that the US value of social interaction in this condition was higher than the sucrose US value. However, with a higher sucrose level, we find that at 5%, a preference for social interaction is no longer observed, but neither is a preference for drinking the 5% sucrose water. This point is called the indifference point and occurs when the reward value of the two different outcomes is equal. The sucrose consumption in the 10% percent, however, caused a sucrose place preference. This preference for sucrose consumption over social interaction indicates that the US value of sucrose in this condition was higher than the social US, possibly due to the strong hedonic value associated with drinking the sweet sucrose water, as discussed above.

5.4.5 50 kHz vocalization expression patterns when choosing between a sucrose reward and social interaction. As discussed in the introduction, the sucrose place preference observed here is part of the feeding motivation cycle caused by increased feeding motivation due to the highly hedonic sucrose reward. This sucrose reward causes rats to develop the motivation to drink longer and thus spend more time drinking. In contrast, the social place preference is part of the social motivation cycle, whereby the increased social motivation to spend time with the socially rewarding event of interacting with a juvenile rat is stronger than the feeding motivation (i.e., to drink a less preferred 2% sucrose reward). Interestingly, when the sucrose level was higher, and the rat decided to spend less time with the juvenile rat and more time drinking the 10% sucrose, there was a higher rate of 50 kHz vocalization (specifically the subtype trill) during the social interaction.

One would expect to find an association between the vocalization rate and social interaction. Brudzynski and Pniak (2002) found that rats vocalize dramatically more during periods of social interaction than periods of being alone in an observation cage. We extend this finding by showing that rats generally also vocalize more during social interaction versus drinking sucrose water

solutions with different concentrations. It was, however, not expected that even though the rats spent less time interacting, they would still vocalize more when the sucrose concentration was higher. These vocalizations may indicate that a 10% visit on the maze caused a high-arousal state, which includes the hedonic "liking" properties such as rhythmic tongue protrusions that rats express when, for example, they are given a 10% sucrose reward orally (Wilmouth & Spear, 2009). This positive motivational state was likely associated with nibbles, sniffs, and bites directed at the drinking cup (even though we did not measure this) that are performed by the rats when a cue comes to predict a highly hedonic outcome and can be amplified by central amygdala stimulation (DiFeliceantonio & Berridge, 2012). It is currently unclear if this liking state persists after drinking the 10% solution when the rat traverses the maze to meet the juvenile. I propose that it is indeed the increased arousal state that causes the increase in trill calls during social interaction, similar to the hedonic state that results from amphetamine injections and which can cause an enhancement in conditioned 50 kHz vocalizations and a stronger social place preference in rats that generally call more (Ahrens et al., 2013). Moreover, it is also possible that USV production is inactive during the licking and drinking response because USV production is bound to other behaviors such as active sniffing, which is again phase-locked with whisking (Kleinfeld, 2014; Sirotin, Costa, & Laplagne, 2014). Importantly, 50 USV's subtype trills and complex calls increased during social facial touching between two rats, but not when the rat was alone (Rao, Mielke, Bobrov, & Brecht, 2014a). Therefore, this finding would argue in favor of the observed enhancement of expressed USV and the subtype trill during nose-to-nose contact between the adolescent and juvenile rats, which is probably associated with an active sniffing pattern, but not during active drinking. Another option would be that the juvenile rat sensed the emotional state of the rat who had just consumed the reward and emotional contagion (de Waal & Preston, 2017) took place, whereby the juvenile matched its emotional state with that of the rat that had just drank the hedonic 10% solution, and this caused the juvenile to vocalize more.

Overall, different improvements to our analyses could help provide a clearer answer for the question as to why we observe more than 50 kHz and especially the subtype trill here during social interaction, while at the same time the rats spend less time in the juvenile zone (when the other option is the 10% sucrose). A more precise analysis of the USV production for the specific behavioral components of the social approach versus social interaction could be undertaken. A more detailed time-resolved analysis of the sucrose consumption response (is there a hedonic response?) and social interaction could be used, which would include a representation of when, where, and which subtype was vocalized. Finally, to see if the 10% sucrose genuinely induced a hedonic state, simple task manipulations could be done that include one group where the rats are

satiated on the 10% sucrose and another group where the rats are not satiated. The hypothesis would be that the satiated rats would decrease their 50 kHz vocalizations during social interaction, while the other group would show an enhancement similar to what we observed.

It is clear, though, that we have found that the "hedonic" value of sucrose can influence social interaction and its associated 50 kHz vocalizations and subtypes in rats. Therefore, the task and its observed behavioral repertoire are helpful for understanding how animals evaluate social interaction in the presence of "hedonic" appetitive alternatives.

6. Conclusions.

Vicarious reward unblocks associative learning about novel cues. Rats vicariously learn the value of stimuli that predict appetitive outcomes delivered to others. Food cup behavior indicative of a conditioned response is unblocked for cues that predict a mutual Both-Reward, while it is blocked for cues that predict an OR outcome (pCS+ > pCS-). The Social unblocking effect is dependent on social information observation and/or exchange. Partner absence and a wall that impedes visual observation during the compound phase both impair the effect. Finally, the reward configuration that includes a Both-reward versus Own-reward distinction enables the social unblocking effect, while if the added cue predicts Partner-reward only, in the absence of actor reward, no unblocking is observed.

Posterior-lateral Orbitofrontal cortex lesion leaves social and appetitive unblocking intact but impairs the blocking effect related to partner no-reward. We hypothesized that the unblocking effect was dependent on social-identity signals, such as cue triggered partner approach and pellet consumption, indicative of a Both-reward versus Own-reward distinction and that these identity signals drove social unblocking by ways of neuronal activity in the Posterior-lateral orbitofrontal cortex, as it does in appetitive unblocking. We find, however, that Posterior-lateral orbitofrontal cortex lesion does *not* impair the social unblocking effect and are therefore interpret that PLO is not required for driving the upward valuation of cues associated with a Both-reward. We do on the other hand find that the blocking of a cue that predicts no additional reward to the other rat is impaired, possibly due its association with self-reward and self no-reward or to an enhanced reliance on cue-reward value of nearby trials, over assigning the right credit to each stimulus outcome.

Distinct profiles of 50 kHz vocalizations differentiate between social and non-social reward approach and consumption. Rats express themselves emotionally by vocalizing 50 kHz calls, when encountering either social or non-social rewards and depending on the situation 50 KHz call subtypes differ. It was however not yet known, how these vocalizations in non-social situation affect 50 KHz vocalizations in social situations and vice versa. In this study, we first established that rats prefer to spend time in the arms of an X shaped maze that signal a high sucrose reward (10%) over a lower sucrose reward (2% or 5%). Afterwards, we directly compared how the rats behaved when they had the opportunity to explore a maze and choose to either interact in one arm with a juvenile rat or drink one of three different sucrose solutions with different sucrose concentrations (2%, 5% and 10%) in the opposite arm. We found that rats prefer to interact with the juvenile in the 2% versus Juvenile condition, had no preference in the 5%

versus juvenile condition and switched their preference to spend more time drinking the sucrose solution in the 10% versus juvenile condition. We finally found that, even though they spent less time with a higher sucrose reward, rats generally increased their 50 kHz vocalisations when interacting with a juvenile rat, when sucrose water concentrations increased from 2% to 5% to 10%. Specifically, the subtypes of trill and composite showed a linear association between sucrose level and their difference in vocalisation rate between reward sites. In conclusion, the rat's behavior and its expressed 50 KHz vocalisation, including different subtypes, are directly influenced by sucrose concentration.

7. References

Achterberg, E. M., Wm Van Kerkhof, L., Servadio, M., Mh Van Swieten, M., Houwing, D. J., Aalderink, M., … Vanderschuren, L. J. (2016). Contrasting Roles of Dopamine and Noradrenaline in the Motivational Properties of Social Play Behavior in Rats. *Neuropsychopharmacology*, *41212*(10), 858–868. https://doi.org/10.1038/npp.2015.212

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, *32*(3), 537–551. https://doi.org/10.1016/S0896-6273(01)00491-3

Ahrens, A. M., Nobile, C. W., Page, L. E., Maier, E. Y., Duvauchelle, C. L., & Schallert, T. (2013). Individual differences in the conditioned and unconditioned rat 50-kHz ultrasonic vocalizations elicited by repeated amphetamine exposure. *Psychopharmacology*, *229*(4), 687–700. https://doi.org/10.1007/s00213-013-3130-9

Al-imari, L., & Gerlai, R. (2008). Sight of conspecifics as reward in associative learning in zebrafish (Danio rerio), *189*, 216–219. https://doi.org/10.1016/j.bbr.2007.12.007

Albrecht, K., von Essen, E., Fliessbach, K., & Falk, A. (2013). The influence of status on satisfaction with relative rewards. *Frontiers in Psychology*, *4*(October), 1–9. https://doi.org/10.3389/fpsyg.2013.00804

American Psychological Association. (n.d.). Conditioned place preference. In *APA dictionary of psychology*. Retrieved April 12, 2021, from https://dictionary.apa.org/Conditioned place preference

American Psychological Association. (n.d.). Observational learning. In *APA dictionary of psychology*. Retrieved April 12, 2021, from https://dictionary.apa.org/Conditioned place preference

American Psychological Association. (n.d.). Social reinforcement. In *APA dictionary of psychology*. Retrieved April 12, 2021, from https://dictionary.apa.org/Social reinforcement

American Psychological Association. (n.d.). Social stimuli. In *APA dictionary of psychology*. Retrieved April 12, 2021, from https://dictionary.apa.org/Social stimuli.

Anderson, B. A. (2016). Social reward shapes attentional biases. *Cognitive Neuroscience*, *7*(1–4), 30–36. https://doi.org/10.1080/17588928.2015.1047823

Anderson, J. R. (1998). Social stimuli and social rewards in primate learning and cognition.

Behavioural Processes, *42*(2–3), 159–175. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24897460

Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron*, *90*(4), 692–707. https://doi.org/10.1016/j.neuron.2016.04.018

Arbilly, M., & Laland, K. N. (2013). The local enhancement conundrum: In search of the adaptive value of a social learning mechanism. *Theoretical Population Biology*. https://doi.org/10.1016/j.tpb.2013.09.006

Avital, A., Aga-Mizrachi, S., & Zubedat, S. (2016). Evidence for social cooperation in rodents by automated maze. *Scientific Reports*, *6*(1), 29517. https://doi.org/10.1038/srep29517

Azzi, J. C. B., Sirigu, A., & □ Duhamel, J. R. (2012). Correction for Azzi et al., Modulation of value representation by social context in the primate orbitofrontal cortex. *Proceedings of the National Academy of Sciences*, *109*(10), 4020–4020. https://doi.org/10.1073/pnas.1201821109

Bahar, A., Dudai, Y., & Ahissar, E. (2004). Neural Signature of Taste Familiarity in the Gustatory Cortex of the Freely Behaving Rat, (Lubow 1989), 3298–3308. https://doi.org/10.1152/jn.00198.2004

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: an integrated view of amygdala function. *Trends in Neurosciences*, *29*(5), 272–279. https://doi.org/10.1016/j.tins.2006.03.002

Bandura, A. (1971). Vicarious and self-reinforcement processes. *The Nature of Reinforcement*.

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F. F., & Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. *Nature*, *585*(7824), 245–250. https://doi.org/10.1038/s41586-020-2704-z

Barreiros, I. V, Panayi, M. C., & Walton, M. E. (2021). Organization of Afferents along the Anterior–posterior and Medial–lateral Axes of the Rat Orbitofrontal Cortex. *Neuroscience*, *460*, 53–68. https://doi.org/10.1016/j.neuroscience.2021.02.017

Ben-Ami Bartal, I., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science (New York, N.Y.)*, *334*(6061), 1427–1430. https://doi.org/10.1126/science.1210789

Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *The European Journal of Neuroscience*, *35*(7), 1124–1143.

https://doi.org/10.1111/j.1460-9568.2012.07990.x

Berridge, K. C., & Kringelbach, M. L. (2015). Pleasure Systems in the Brain. https://doi.org/10.1016/j.neuron.2015.02.018

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28*(3), 309–369. https://doi.org/10.1016/S0165-0173(98)00019-8

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences*, *26*(9), 507–513. https://doi.org/10.1016/S0166-2236(03)00233-9

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: "liking", "wanting", and learning. *Current Opinion in Pharmacology*, *9*(1), 65–73. https://doi.org/10.1016/j.coph.2008.12.014

Binkley, K. A., Webber, E. S., Powers, D. D., & Cromwell, H. C. (2014). Emotion and relative reward processing: An investigation on instrumental successive negative contrast and ultrasonic vocalizations in the rat. *Behavioural Processes*. https://doi.org/10.1016/j.beproc.2014.07.011

Birch, L. L., Zimmerman, S. I., & Hind, H. (1980). The Influence of Social-Affective Context on the Formation of Children's Food Preferences. *Child Development*, *51*(3), 856. https://doi.org/10.2307/1129474

Bolles, R. C., & Fanselow, M. S. (1980). A perceptual-defensive-recuperative model of fear and pain. *Behavioral and Brain Sciences*, *3*(2), 291–301. https://doi.org/10.1017/S0140525X0000491X

Börner, A., Hjemdahl, R., Götz, T., & Brown, G. R. (2016). Ultrasonic vocalizations of female norway rats (Rattus norvegicus) in response to social partners. *Journal of Comparative Psychology*. https://doi.org/10.1037/com0000017

Brenes, J. C., & Schwarting, R. K. W. (2014). Attribution and expression of incentive salience are differentially signaled by ultrasonic vocalizations in rats. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0102414

Brosnan, S. F., & de Waal, F. B. M. (2003). Monkeys reject unequal pay. *Nature*, *425*(6955), 297–299. https://doi.org/10.1038/nature01963

Browning, J. R., Browning, D. A., Maxwell, A. O., Dong, Y., Jansen, H. T., Panksepp, J., & Sorg,

B. A. (2011). Positive affective vocalizations during cocaine and sucrose self-administration: A model for spontaneous drug desire in rats. *Neuropharmacology*. https://doi.org/10.1016/j.neuropharm.2011.04.012

Brudzynski, S. (2015). Pharmacology of Ultrasonic Vocalizations in adult Rats: Significance, Call Classification and Neural Substrate. *Current Neuropharmacology*. https://doi.org/10.2174/1570159x13999150210141444

Brudzynski, S. M. (2013). Ethotransmission : communication of emotional states through ultrasonic vocalization in rats. *Current Opinion in Neurobiology*, *23*(3), 310–317. https://doi.org/10.1016/j.conb.2013.01.014

Brudzynski, S. M. (2014). Social origin of vocal communication in rodents. In *Biocommunication of Animals*. https://doi.org/10.1007/978-94-007-7414-8_5

Brudzynski, S. M., & Holland, G. (2005). Acoustic characteristics of air puff-induced 22-kHz alarm calls in direct recordings. In *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2005.04.007

Brudzynski, S. M., & Pniak, A. (2002). Social contacts and production of 50-kHz short ultrasonic calls in adult rats. *Journal of Comparative Psychology*. https://doi.org/10.1037/0735-7036.116.1.73

Brudzynski, S. M., & Zeskind, P. S. (2018). Introduction and Overview of the Handbook of Ultrasonic Vocalization. In *Handbook of Behavioral Neuroscience*. https://doi.org/10.1016/B978-0-12-809600-0.00001-9

Burgdorf, J., & Panksepp, J. (2006). The neurobiology of positive emotions. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2005.06.001

Burgdorf, J., Panksepp, J., Brudzynski, S. M., Beinfeld, M. C., Cromwell, H. C., Kroes, R. A., & Moskal, J. R. (2009). The effects of selective breeding for differential rates of 50-kHz ultrasonic vocalizations on emotional behavior in rats. *Developmental Psychobiology*. https://doi.org/10.1002/dev.20343

Burgdorf, J., Panksepp, J., & Moskal, J. R. (2011a). Frequency-modulated 50kHz ultrasonic vocalizations: A tool for uncovering the molecular substrates of positive affect. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2010.11.011

Burgdorf, J., Panksepp, J., & Moskal, J. R. (2011b). Neuroscience and Biobehavioral Reviews

Frequency-modulated 50 kHz ultrasonic vocalizations : a tool for uncovering the molecular substrates of positive affect. *Neuroscience and Biobehavioral Reviews*, *35*(9), 1831–1836. https://doi.org/10.1016/j.neubiorev.2010.11.011

Burke, K. A., Franz, T. M., Miller, D. N., & Schoenbaum, G. (2008). The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. *Nature*, *454*(7202), 340–344. https://doi.org/10.1038/nature06993

Calcagnetti, D. J., & Schechter, M. D. (1992). Place conditioning reveals the rewarding aspect of social interaction in juvenile rats. *Physiology and Behavior*, *51*(4), 667–672. https://doi.org/10.1016/0031-9384(92)90101-7

Capozzi, F., & Ristic, J. (2018). How attention gates social interactions. *Annals of the New York Academy of Sciences*, *1426*(1), 179–198. https://doi.org/10.1111/nyas.13854

Cardinal, R. N., Parkinson, J. A., Marbini, H. D., Toner, A. J., Bussey, T. J., Robbins, T. W., & Everitt, B. J. (2003). Role of the anterior cingulate cortex in the control over behavior by Pavlovian conditioned stimuli in rats. *Behavioral Neuroscience*, *117*(3), 566–587. https://doi.org/10.1037/0735-7044.117.3.566

Carrillo, M., Han, Y., Migliorati, F., Liu, M., Gazzola, V., & Keysers, C. (2019). Emotional Mirror Neurons in the Rat's Anterior Cingulate Cortex. *Current Biology*, *29*(8), 1301-1312.e6. https://doi.org/10.1016/j.cub.2019.03.024

Chakrabarti, B., Haffey, A., Canzano, L., Taylor, P., & Mcsorley, E. (2017). Individual differences in responsivity to social rewards : Insights from two eye-tracking tasks, *1*, 1–13.

Chang, C. Y., Gardner, M., Di Tillio, M. G., & Schoenbaum, G. (2017). Optogenetic Blockade of Dopamine Transients Prevents Learning Induced by Changes in Reward Features. *Current Biology : CB*, *27*(22), 3480-3486.e3. https://doi.org/10.1016/j.cub.2017.09.049

Chang, S. E., McDannald, M. A., Wheeler, D. S., & Holland, P. C. (2012). The effects of basolateral amygdala lesions on unblocking. *Behavioral Neuroscience*, *126*(2), 279–289. https://doi.org/10.1037/a0027576

Chang, S. W. C., Gariépy, J. F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience*, *16*(2), 243–250. https://doi.org/10.1038/nn.3287

Chang, S. W. C., Winecoff, A. A., & Platt, M. L. (2011). Vicarious reinforcement in rhesus

macaques (macaca mulatta). *Frontiers in Neuroscience*, *5*(March), 27. https://doi.org/10.3389/fnins.2011.00027

Cleland, G. G., & Davey, G. C. L. (1983). Autoshaping in the rat: The effects of localizable visual and auditory signals for food. *Journal of the Experimental Analysis of Behavior*, *40*(1), 47–56. https://doi.org/10.1901/jeab.1983.40-47

Coffey, K. R., Barker, D. J., Ma, S., Root, D. H., Martinez, L., Horvitz, J. C., & West, M. O. (2013). Effects of varying reinforcement probability on pavlovian approach behavior and ultrasonic vocalizations in rats. *Behavioural Brain Research*, *237*(1), 256–262. https://doi.org/10.1016/j.bbr.2012.09.041

Coffey, K. R., Marx, R. G., & Neumaier, J. F. (2019). DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology*. https://doi.org/10.1038/s41386-018-0303-6

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early Social Attention Impairments in Autism: Social Orienting, Joint Attention, and Attention to Distress. *Developmental Psychology*, *40*(2), 271–283. https://doi.org/10.1037/0012-1649.40.2.271

Day, J. E. L., Kyriazakis, I., & Rogers, P. J. (1998). Food choice and intake: towards a unifying framework of learning and feeding motivation. *Nutrition Research Reviews*, *11*(1), 25–43. https://doi.org/10.1079/nrr19980004

de Waal, F. B. M., & Preston, S. D. (2017). Mammalian empathy: behavioural manifestations and neural basis. *Nature Reviews Neuroscience*, *18*(8), 498–509. https://doi.org/10.1038/nrn.2017.72

de Waal, F. B., & Suchak, M. (2010). Prosocial primates: selfish and unselfish motivations. *Philos Trans R Soc Lond B Biol Sci*, *365*(1553), 2711–2722. https://doi.org/365/1553/2711 [pii]10.1098/rstb.2010.0119

Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys Pay Per View: Adaptive Valuation of Social Images by Rhesus Macaques. *Current Biology*, *15*(6), 543–548. https://doi.org/10.1016/j.cub.2005.01.044

Deisseroth, K., Feng, G., Majewska, A. K., Miesenbock, G., Ting, A., & Schnitzer, M. J. (2006). Next-Generation Optical Technologies for Illuminating Genetically Targeted Brain Circuits. *Journal of Neuroscience*, *26*(41), 10380–10386.

https://doi.org/10.1523/JNEUROSCI.3863-06.2006

Desor, J. A., & Beauchamp, G. K. (1987). Longitudinal changes in sweet preferences in humans. *Physiology & Behavior*, *39*(5), 639–641. https://doi.org/10.1016/0031-9384(87)90166-1

Dickinson, A., Hall, G., & Mackintosh, N. J. (1976). Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(4), 313–322. https://doi.org/10.1037/0097-7403.2.4.313

DiFeliceantonio, A. G., & Berridge, K. C. (2012). Which cue to "want"? Opioid stimulation of central amygdala makes goal-trackers show stronger goal-tracking, just as sign-trackers show stronger sign-tracking. *Behavioural Brain Research*, *230*(2), 399–408. https://doi.org/10.1016/j.bbr.2012.02.032

Dolensek, N., Gehrlach, D. A., Klein, A. S., & Gogolla, N. (2020). Facial expressions of emotion states and their neuronal correlates in mice. *Science*, *368*(6486), 89–94. https://doi.org/10.1126/science.aaz9468

Dwyer, D. M. (2011). Lesions of the basolateral , but not central , amygdala impair flavour-taste learning based on fructose or quinine reinforcers. *Behavioural Brain Research*, *220*(2), 349–353. https://doi.org/10.1016/j.bbr.2011.02.007

Ellenbroek, B., Abizaid, A., Amir, S., de Zwaan, M., Parylak, S., Cottone, P., … Hutton, S. B. (2010). Extracellular Recording. In *Encyclopedia of Psychopharmacology* (pp. 522–525). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68706-1_290

Fanselow, M. S., & Wassum, K. M. (2016). The origins and organization of vertebrate pavlovian conditioning. *Cold Spring Harbor Perspectives in Biology*, *8*(1), 1–28. https://doi.org/10.1101/cshperspect.a021717

Fehr, E., & Rockenbach, B. (2004). Human altruism: economic, neural, and evolutionary perspectives. *Current Opinion in Neurobiology*, *14*(6), 784–790. https://doi.org/10.1016/j.conb.2004.10.007

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum. *Science*, *318*(5854), 1305–1308. https://doi.org/10.1126/science.1145876

Fonseca, E., de Lafuente, V., Simon, S. A., & Gutierrez, R. (2018). Sucrose intensity coding and decision-making in rat gustatory cortices. *ELife*, *7*, 1–33.

https://doi.org/10.7554/eLife.41152

Fulford, D., Campellone, T., & Gard, D. E. (2018). Social motivation in schizophrenia: How research on basic reward processes informs and limits our understanding. *Clinical Psychology Review*, *63*(September 2017), 12–24. https://doi.org/10.1016/j.cpr.2018.05.007

Galef, B. G. (2001). Social Influences on Food Choices of Norway Rats and Mate Choices of Japanese Quail. *International Journal of Comparative Psychology*, *14*(1), 1–24. https://doi.org/https://doi.org/10.5070/P4141009841

Galef, B. G., Iliffe, C. P., & Whiskin, E. E. (1994). Social influences on rats' (Rattus norvegicus) preferences for flavored foods, scented nest materials, and odors associated with harborage sites: Are flavored foods special? *Journal of Comparative Psychology*, *108*(3), 266–273. https://doi.org/10.1037/0735-7036.108.3.266

Galef, B. G., & Laland, K. N. (2005). Social Learning in Animals: Empirical Studies and Theoretical Models. *BioScience*, *55*(6), 489. https://doi.org/10.1641/0006-3568(2005)055[0489:SLIAES]2.0.CO;2

Gao, S. C., Wei, Y. C., Wang, S. R., & Xu, X. H. (2019). Medial Preoptic Area Modulates Courtship Ultrasonic Vocalization in Adult Male Mice. *Neuroscience Bulletin*. https://doi.org/10.1007/s12264-019-00365-w

Garcia, E. J., McCowan, T. J., & Cain, M. E. (2015). Harmonic and frequency modulated ultrasonic vocalizations reveal differences in conditioned and unconditioned reward processing. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2015.03.049

Guić-Robles, E., Valdivieso, C., & Guajardo, G. (1989). Rats can learn a roughness discrimination using only their vibrissal system. *Behavioural Brain Research*, *31*(3), 285–289. https://doi.org/10.1016/0166-4328(89)90011-9

Gunaydin, L. A., Grosenick, L., Finkelstein, J. C., Kauvar, I. V., Fenno, L. E., Adhikari, A., … Deisseroth, K. (2014). Natural Neural Projection Dynamics Underlying Social Behavior. *Cell*, *157*(7), 1535–1551. https://doi.org/10.1016/j.cell.2014.05.017

Guo, B., Chen, J., Chen, Q., Ren, K., Feng, D., Mao, H., … Wu, S. (2019). Anterior cingulate cortex dysfunction underlies social deficits in Shank3 mutant mice. *Nature Neuroscience*, *22*(8), 1223–1234. https://doi.org/10.1038/s41593-019-0445-9

Hamilton, W. D. (1963). The Evolution of Altruistic Behavior. *American Naturalist, The*, *97*(896),

354.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural Responses to Taxation and Voluntary Giving Reveal Motives for Charitable Donations. *Science*, *316*(5831), 1622–1625. https://doi.org/10.1126/science.1140738

Hari, R., Henriksson, L., Malinen, S., & Parkkonen, L. (2015). Centrality of Social Interaction in Human Brain Function. *Neuron*, *88*(1), 181–193. https://doi.org/10.1016/j.neuron.2015.09.022

Haselgrove, M., Tam, S. K. E., & Jones, P. M. (2013). Enhanced unblocking from sustained post-trial surprise. *Journal of Experimental Psychology: Animal Behavior Processes*, *39*(4), 311–322. https://doi.org/10.1037/a0032569

Hayward, D. A., Pereira, E. J., Otto, A. R., & Ristic, J. (2018). Smile! Social reward drives attention. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(2), 206–214. https://doi.org/10.1037/xhp0000459

Heckman, J. J., Proville, R., Heckman, G. J., Azarfar, A., Celikel, T., & Englitz, B. (2017). High-precision spatial localization of mouse vocalizations during social interaction. *Scientific Reports*. https://doi.org/10.1038/s41598-017-02954-z

Henriksen, R. E., Torsheim, T., & Thuen, F. (2014). Loneliness, social integration and consumption of sugar-containing beverages: Testing the social baseline theory. *PLoS ONE*, *9*(8). https://doi.org/10.1371/journal.pone.0104421

Hernandez-Lallement, J., van Wingerden, M., Marx, C., Srejic, M., & Kalenscher, T. (2015). Rats prefer mutual rewards in a prosocial choice task. *Frontiers in Neuroscience*, *8*(JAN), 1–9. https://doi.org/10.3389/fnins.2014.00443

Hernandez-Lallement, J., van Wingerden, M., Schäble, S., & Kalenscher, T. (2016a). A Social Reinforcement Learning Hypothesis of Mutual Reward Preferences in Rats. In *Current topics in behavioral neurosciences* (Vol. 7, pp. 159–176). https://doi.org/10.1007/7854_2016_436

Hernandez-Lallement, J., van Wingerden, M., Schäble, S., & Kalenscher, T. (2016b). Basolateral amygdala lesions abolish mutual reward preferences in rats. *Neurobiology of Learning and Memory*, *127*, 1–9. https://doi.org/10.1016/j.nlm.2015.11.004

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews*, *69*, 207–231. https://doi.org/10.1111/j.1469-185X.1994.tb01506.x

Heyes, C. M., & Dawson, G. R. (1990). A Demonstration of Observational Learning in Rats Using a Bidirectional Control. *The Quarterly Journal of Experimental Psychology Section B*, *42*(1), 59–71. https://doi.org/10.1080/14640749008401871

Heyes, C. M., Jaldow, E., Nokes, T., & Dawson, G. R. (1994). Imitation in rats (Rattus norvegicus): The role of demonstrator action. *Behavioural Processes*, *32*(2), 173–182. https://doi.org/10.1016/0376-6357(94)90074-4

Heyes, C. M., Ray, E. D., Mitchell, C. J., & Nokes, T. (2000). Stimulus Enhancement: Controls for Social Facilitation and Local Enhancement. *Learning and Motivation*, *31*(2), 83–98. https://doi.org/10.1006/lmot.1999.1041

Hill, C. A. (1987). Affiliation motivation: People who need people… but in different ways. *Journal of Personality and Social Psychology*, *52*(5), 1008–1018. https://doi.org/10.1037/0022-3514.52.5.1008

Hillman, E. M. C. (2014). Coupling Mechanism and Significance of the BOLD Signal: A Status Report. *Annual Review of Neuroscience*, *37*(1), 161–181. https://doi.org/10.1146/annurev-neuro-071013-014111

Hillman, K. L., & Bilkey, D. K. (2012). Neural encoding of competitive effort in the anterior cingulate cortex. *Nature Neuroscience*, *15*(9), 1290–1297. https://doi.org/10.1038/nn.3187

Himmler, B. T., Kisko, T. M., Euston, D. R., Kolb, B., & Pellis, S. M. (2014). Are 50-kHz calls used as play signals in the playful interactions of rats? I. Evidence from the timing and context of their use. *Behavioural Processes*, *106*, 60–66. https://doi.org/10.1016/j.beproc.2014.04.014

Hofer, M. A. (1996). Multiple regulators of ultrasonic vocalization in the infant rat. *Psychoneuroendocrinology*, *21*(2), 203–217. https://doi.org/10.1016/0306-4530(95)00042-9

Holland, P. C. (1977). Conditioned Stimulus as a Determinant of the Form of the Pavlovian Conditioned Response, *3*(1), 77–104.

Holland, P. C. (1984). Unblocking in Pavlovian appetitive conditioning. *Journal of Experimental Psychology. Animal Behavior Processes*, *10*(4), 476–497. https://doi.org/10.1037/0097-7403.10.4.476

Holland, P. C., & Gallagher, M. (1993a). Amygdala Central Nucleus Lesions Disrupt Increments, But Not Decrements, in Conditioned Stimulus Processing. *Behavioral Neuroscience*.

https://doi.org/10.1037/0735-7044.107.2.246

Holland, P. C., & Gallagher, M. (1993b). Effects of amygdala central nucleus lesions on blocking and unblocking. *Behavioral Neuroscience*, *107*(2), 235–245. https://doi.org/10.1037/0735-7044.107.2.235

Honey, R. C., & Hall, G. (1989). Enhanced discriminability and reduced associability following flavor preexposure. *Learning and Motivation*, *20*(3), 262–277. https://doi.org/10.1016/0023-9690(89)90008-8

Hoover, W. B., & Vertes, Æ. R. P. (2007). Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat, 149–179. https://doi.org/10.1007/s00429-007-0150-4

Horner, V., Carter, J. D., Suchak, M., & de Waal, F. B. M. (2011). Spontaneous prosocial choice by chimpanzees. *Proceedings of the National Academy of Sciences*, *108*(33), 13847–13851. https://doi.org/10.1073/pnas.1111088108

Huang, W. C., Zucca, A., Levy, J., & Page, D. T. (2020). Social Behavior Is Modulated by Valence-Encoding mPFC-Amygdala Sub-circuitry. *Cell Reports*, *32*(2), 107899. https://doi.org/10.1016/j.celrep.2020.107899

Hughes, R. N. (1997). Intrinsic exploration in animals: Motives and measurement. *Behavioural Processes*, *41*(3), 213–226. https://doi.org/10.1016/S0376-6357(97)00055-7

Humphreys, A. P., & Einon, D. F. (1981). Play as a reinforcer for maze-learning in juvenile rats. *Animal Behaviour*, *29*(1), 259–270. https://doi.org/10.1016/S0003-3472(81)80173-X

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*. https://doi.org/10.1109/MCSE.2007.55

Izquierdo, A. (2018). Functional Heterogeneity within Rat Orbitofrontal Cortex in Reward Learning and Decision Making. *The Journal of Neuroscience*, *37*(44), 10529–10540. https://doi.org/10.1523/jneurosci.1678-17.2017

Jennings, J. H., Kim, C. K., Marshel, J. H., Raffiee, M., Ye, L., Quirin, S., … Deisseroth, K. (2019). Interacting neural ensembles in orbitofrontal cortex for social and feeding behaviour. *Nature*, *565*(7741), 645–649. https://doi.org/10.1038/s41586-018-0866-8

Johnson, P. E. (2019). Regression Estimation and Presentation.

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. C. (2017). Social learning through prediction error in the brain. *Npj Science of Learning*, *2*(1), 8. https://doi.org/10.1038/s41539-017-0009-2

Jones, R. M., Somerville, L. H., Li, J., Ruberry, E. J., Libby, V., Glover, G., … Casey, B. J. (2011). Behavioral and neural properties of social reinforcement learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *31*(37), 13039–13045. https://doi.org/10.1523/JNEUROSCI.2972-11.2011

Kalenscher, T. (2020). Social Neuroscience: How the Brain Attends to the Joys and Pains of Others. *Current Biology*. https://doi.org/10.1016/j.cub.2020.07.051

Kalenscher, T., Schönfeld, L.-M., Löbner, S., Wöhr, M., van Berkel, M., Zech, M.-P., & van Wingerden, M. (2020). Rat Ultrasonic Vocalizations as Social Reinforcers—Implications for a Multilevel Model of the Cognitive Representation of Action and Rats' Social World. In *Cognitive Structures*.

Kamin, L. J. (1969). Predictability, surprise, attention and conditioning. *In B. A. Campbell, & R. M. Church (Eds.) Punishment Aversive Behavior. New York: Appleton- Century-Crofts.*, 279–296.

Kashtelyan, V., Lichtenberg, N. T., Chen, M. L., Cheer, J. F., & Roesch, M. R. (2014). Observation of Reward Delivery to a Conspecific Modulates Dopamine Release in Ventral Striatum. *Current Biology*, *24*(21), 2564–2568. https://doi.org/10.1016/j.cub.2014.09.016

Kawamichi, H., Sugawara, S. K., Hamano, Y. H., Makita, K., Kochiyama, T., & Sadato, N. (2016). Increased frequency of social interaction is associated with enjoyment enhancement and reward system activation. *Scientific Reports*, *6*(1), 24561. https://doi.org/10.1038/srep24561

Keiflin, R., Pribut, H. J., Shah, N. B., & Janak, P. H. (2019). Ventral Tegmental Dopamine Neurons Participate in Reward Identity Predictions. *Current Biology*, *29*(1), 93-103.e3. https://doi.org/10.1016/j.cub.2018.11.050

Kelm-Nelson, C. A., Lenell, C., Johnson, A. M., & Ciucci, M. R. (2018). Laryngeal Activity for Production of Ultrasonic Vocalizations in Rats. In *Handbook of Behavioral Neuroscience*. https://doi.org/10.1016/B978-0-12-809600-0.00004-4

Kisko, T. M., Braun, M. D., Michels, S., Witt, S. H., Rietschel, M., Culmsee, C., … Wöhr, M. (2018). Cacna1c haploinsufficiency leads to pro-social 50-kHz ultrasonic communication deficits in rats. *DMM Disease Models and Mechanisms*. https://doi.org/10.1242/dmm.034116

Kleinfeld, D. (2014). More than a rhythm of life: Breathing as a binder of orofacial sensation, *17*(5), 647–651. https://doi.org/10.1038/nn.3693.More

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., … Willing, C. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas - Proceedings of the 20th International Conference on Electronic Publishing, ELPUB 2016*. https://doi.org/10.3233/978-1-61499-649-1-87

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *21*(16), 1–5. https://doi.org/10.1523/jneurosci.21-16-j0002.2001

Knutson, Brian, Burgdorf, J., & Panksepp, J. (1999). Anticipation of play elicits high-frequency ultrasonic vocalizations in young rats. *Journal of Comparative Psychology*. https://doi.org/10.1037//0735-7036.112.1.65

Knutson, Brian, Burgdorf, J., & Panksepp, J. (2002). Ultrasonic vocalizations as indices of affective states in rats. *Psychological Bulletin*, *128*(6), 961–977. https://doi.org/10.1037//0033-2909.128.6.961

Kohls, G., Chevallier, C., Troiani, V., & Schultz, R. T. (2012). Social "wanting" dysfunction in autism: neurobiological underpinnings and treatment implications. *Journal of Neurodevelopmental Disorders*, *4*(1), 10. https://doi.org/10.1186/1866-1955-4-10

Kruse, J. M., Overmier, J. B., Konz, W. A., & Rokke, E. (1983). Pavlovian conditioned stimulus effects upon instrumental choice behavior are reinforcer specific. *Learning and Motivation*, *14*(2), 165–181. https://doi.org/10.1016/0023-9690(83)90004-8

Latane, B., Joy, V., Meltzer, J., Lubell, B., & Cappell, H. (1972). Stimulus Determinants of Social Attraction in Rats. *Journal of Comparative and Physiological Psychology*, *79*(1), 13–21. https://doi.org/10.1037/h0032563

Latané, B., & Steele, C. (1975). The persistence of social attraction in socially deprived and satiated rats. *Animal Learning & Behavior*, *3*(2), 131–134. https://doi.org/10.3758/BF03209115

Leadbeater, E., & Chittka, L. (2009). Bumble-bees learn the value of social cues through experience. *Biology Letters*, *5*(3), 310–312. https://doi.org/10.1098/rsbl.2008.0692

Leblanc, H., & Ramirez, S. (2020). Linking social cognition to learning and memory. *Journal of Neuroscience*, *40*(46), 8782–8798. https://doi.org/10.1523/JNEUROSCI.1280-20.2020

Lee, E., Rhim, X. I., Lee, J. W., Ghim, J., Lee, S., Kim, E., & Jung, M. W. (2016). Enhanced Neuronal Activity in the Medial Prefrontal Cortex during Social Approach Behavior, *36*(26), 6926–6936. https://doi.org/10.1523/JNEUROSCI.0307-16.2016

Lehner, R., Balsters, J. H., Herger, A., Hare, T. A., Wenderoth, N., Garofalo, S., & di Pellegrino, G. (2017). Commentary: Monetary, Food, and Social Rewards Induce Similar Pavlovian-to-Instrumental Transfer Effects. *Frontiers in Behavioral Neuroscience*, *10*(January), 136. https://doi.org/10.3389/fnbeh.2017.00136

Lichtenberg, N. T., Pennington, Z. T., Holley, S. M., Greenfield, V. Y., Cepeda, C., Levine, M. S., & Wassum, K. M. (2017). Basolateral Amygdala to Orbitofrontal Cortex Projections Enable Cue-Triggered Reward Expectations. *The Journal of Neuroscience*, *37*(35), 8374–8384. https://doi.org/10.1523/JNEUROSCI.0486-17.2017

Lickliter, R. (2004). The Aims and Accomplishments of Comparative Psychology. *Developmental Psychobiology*, *44*(1), 26–30. https://doi.org/10.1002/dev.10157

Litvin, Y., Blanchard, D. C., & Blanchard, R. J. (2007). Rat 22kHz ultrasonic vocalizations as alarm cries. *Behavioural Brain Research*, *182*(2), 166–172. https://doi.org/10.1016/j.bbr.2006.11.038

Lopatina, N., McDannald, M. A., Styer, C. V, Peterson, J. F., Sadacca, B. F., Cheer, J. F., & Schoenbaum, G. (2016). Medial Orbitofrontal Neurons Preferentially Signal Cues Predicting Changes in Reward during Unblocking. *Journal of Neuroscience*, *36*(32), 8416–8424. https://doi.org/10.1523/JNEUROSCI.1101-16.2016

Lopatina, Nina, McDannald, M. A., Styer, C. V, Sadacca, B. F., Cheer, J. F., & Schoenbaum, G. (2015). Lateral orbitofrontal neurons acquire responses to upshifted, downshifted, or blocked cues during unblocking. *ELife*, *4*, e11299. https://doi.org/10.7554/eLife.11299

Łopuch, S., & Popik, P. (2011). Cooperative Behavior of Laboratory Rats (Rattus norvegicus) in an Instrumental Task. *Journal of Comparative Psychology*. https://doi.org/10.1037/a0021532

Lüdecke, D. (2017). sjPlot: Data Visualization for Statistics in Social Science, R package version 2.4.0. *R Package Version 2.4.0*.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*(4), 276–298. https://doi.org/10.1037/h0076778

Madsen, H. B., & Ahmed, S. H. (2015). Drug versus sweet reward: Greater attraction to and

preference for sweet versus drug cues. *Addiction Biology*, *20*(3), 433–444. https://doi.org/10.1111/adb.12134

Maia, T. V. (2009). Reinforcement learning, conditioning, and the brain: Successes and challenges. *Cognitive, Affective and Behavioral Neuroscience*, *9*(4), 343–364. https://doi.org/10.3758/CABN.9.4.343

Manduca, A., Campolongo, P., Palmery, M., Vanderschuren, L. J. M. J., Cuomo, V., & Trezza, V. (2014). Social play behavior, ultrasonic vocalizations and their modulation by morphine and amphetamine in Wistar and Sprague-Dawley rats. *Psychopharmacology*. https://doi.org/10.1007/s00213-013-3337-9

Márquez, C., Rennie, S. M., Costa, D. F., & Moita, M. A. (2015). Prosocial Choice in Rats Depends on Food-Seeking Behavior Displayed by Recipients. *Current Biology*, *25*(13), 1736–1745. https://doi.org/10.1016/j.cub.2015.05.018

Martianova, E., Aronson, S., & Proulx, C. D. (2019). Multi-fiber photometry to record neural activity in freely-moving animals. *Journal of Visualized Experiments*, *2019*(152), 1–9. https://doi.org/10.3791/60278

Mateus-Pinheiro, A., Patrício, P., Alves, N. D., Machado-Santos, A. R., Morais, M., Bessa, J. M., … Pinto, L. (2014). The Sweet Drive Test: Refining phenotypic characterization of anhedonic behavior in rodents. *Frontiers in Behavioral Neuroscience*. https://doi.org/10.3389/fnbeh.2014.00074

McDannald, M. A., Lucantonio, F., Burke, K. A., Niv, Y., & Schoenbaum, G. (2011). Ventral Striatum and Orbitofrontal Cortex Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *Journal of Neuroscience*, *31*(7), 2700–2705. https://doi.org/10.1523/JNEUROSCI.5499-10.2011

McDannald, Michael A., Takahashi, Y. K., Lopatina, N., Pietras, B. W., Jones, J. L., & Schoenbaum, G. (2012). Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *European Journal of Neuroscience*, *35*(7), 991–996. https://doi.org/10.1111/j.1460-9568.2011.07982.x

McDannald, Michael A, Esber, G. R., Wegener, M. A., Wied, H. M., Liu, T., Stalnaker, T. A., … Schoenbaum, G. (2014). Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *ELife*, *3*, 1–14. https://doi.org/10.7554/eLife.02653

McKinney, W. (2010). Data Structures for Statistical Computing in Python. In *Proceedings of the 9th*

Python in Science Conference. https://doi.org/10.25080/majora-92bf1922-00a

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., … Dalgleish, T. (2009). A Key Role for Similarity in Vicarious Reward. *Science*, *324*(5929), 900–900. https://doi.org/10.1126/science.1170539

Mulvihill, K. G., & Brudzynski, S. M. (2018a). Individual behavioural predictors of amphetamine-induced emission of 50 kHz vocalization in rats. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2018.05.009

Mulvihill, K. G., & Brudzynski, S. M. (2018b). Non-pharmacological induction of rat 50 kHz ultrasonic vocalization: Social and non-social contexts differentially induce 50 kHz call subtypes. *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2018.09.005

Nadler, J. J., Moy, S. S., Dold, G., Trang, D., Simmons, N., Perez, A., … Crawley, J. N. (2004). Automated apparatus for quantitation of social approach behaviors in mice. *Genes, Brain and Behavior*, *3*(5), 303–314. https://doi.org/10.1111/j.1601-183X.2004.00071.x

Nasser, H. M., Calu, D. J., Schoenbaum, G., & Sharpe, M. J. (2017). The dopamine prediction error: Contributions to associative models of reward learning. *Frontiers in Psychology*, *8*(FEB), 1–17. https://doi.org/10.3389/fpsyg.2017.00244

Nicol, C. J. (1995). The social transmission of information and behaviour. *Applied Animal Behaviour Science*, *44*(2–4), 79–98. https://doi.org/10.1016/0168-1591(95)00607-T

Noritake, A., Ninomiya, T., & Isoda, M. (2018). Social reward monitoring and valuation in the macaque brain. *Nature Neuroscience*, *21*(10), 1452–1462. https://doi.org/10.1038/s41593-018-0229-7

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560–1563. https://doi.org/10.1126/science.1133755

Oberliessen, L., Hernandez-Lallement, J., Schäble, S., van Wingerden, M., Seinstra, M., & Kalenscher, T. (2016). Inequity aversion in rats, Rattus norvegicus. *Animal Behaviour*, *115*, 157–166. https://doi.org/10.1016/j.anbehav.2016.03.007

Ostlund, S. B., & Balleine, B. W. (2007). Orbitofrontal Cortex Mediates Outcome Encoding in Pavlovian But Not Instrumental Conditioning. *Journal of Neuroscience*, *27*(18), 4819–4825. https://doi.org/10.1523/JNEUROSCI.5443-06.2007

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic

value. *Nature*, *441*(7090), 223–226. https://doi.org/10.1038/nature04676

Panayi, M. C., & Killcross, S. (2018). Functional heterogeneity within the rodent lateral orbitofrontal cortex dissociates outcome devaluation and reversal learning deficits. *ELife*, *7*, 1–27. https://doi.org/10.7554/eLife.37357.001

Panksepp, J. (1981). The ontogeny of play in rats. *Developmental Psychobiology*, *14*(4), 327–332. https://doi.org/10.1002/dev.420140405

Panksepp, J. (2007). Neuroevolutionary sources of laughter and social joy: modeling primal human laughter in laboratory rats. *Behavioural Brain Research*, *182*(2), 231–244. https://doi.org/10.1016/j.bbr.2007.02.015

Panksepp, J., & Burgdorf, J. (2000). 50-kHz chirping (laughter?) in response to conditioned and unconditioned tickle-induced reward in rats: Effects of social housing and genetic variables. *Behavioural Brain Research*. https://doi.org/10.1016/S0166-4328(00)00238-2

Panksepp, J., Gordon, N., & Burgdorf, J. (2002). Empathy and the action-perception resonances of basic socio-emotional systems of the brain. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X0247001X

Park, S. Q., Kahnt, T., Dogan, A., Strang, S., Fehr, E., & Tobler, P. N. (2017). A neural link between generosity and happiness. *Nature Communications*, *8*(May), 15964. https://doi.org/10.1038/ncomms15964

Parsana, A. J., Li, N., & Brown, T. H. (2012). Positive and negative ultrasonic social signals elicit opposing firing patterns in rat amygdala, *226*, 77–86. https://doi.org/10.1016/j.bbr.2011.08.040

Patil, I. (2018). ggstatsplot: 'ggplot2' based plots with statistical details. *CRAN*.

Pavlov, I. P. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences*, *17*(3), 153. https://doi.org/10.5214/ans.0972-7531.1017309

Petrovich, G. D. (2011). Learning and the motivation to eat: Forebrain circuitry. *Physiology & Behavior*, *104*(4), 582–589. https://doi.org/10.1016/j.physbeh.2011.04.059

Phillipson, O. T., & Griffiths, A. C. (1985). The topographic order of inputs to nucleus accumbens in the rat. *Neuroscience*, *16*(2), 275–296. https://doi.org/10.1016/0306-4522(85)90002-8

Prochazkova, E., & Kret, M. E. (2017). Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neuroscience and Biobehavioral Reviews*, *80*(May), 99–114. https://doi.org/10.1016/j.neubiorev.2017.05.013

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413–425. https://doi.org/10.1016/j.tics.2013.06.003

Rao, R. P., Mielke, F., Bobrov, E., & Brecht, M. (2014a). Vocalization-whisking coordination and multisensory integration of social signals in rat auditory cortex. *ELife*, *3*, 1–20. https://doi.org/10.7554/eLife.03185

Rao, R. P., Mielke, F., Bobrov, E., & Brecht, M. (2014b). Vocalization–whisking coordination and multisensory integration of social signals in rat auditory cortex. *ELife*, *3*. https://doi.org/10.7554/eLife.03185

Rescorla, R. A. (1994). Control of Instrumental Performance by Pavlovian and Instrumental Stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(1), 44–50. https://doi.org/10.1037/0097-7403.20.1.44

Rescorla, R. A. (1999). Learning about qualitatively different outcomes during a blocking procedure. *Animal Learning and Behavior*, *27*(2), 140–151. https://doi.org/10.3758/BF03199671

Rescorla, R. a., & Colwill, R. M. (1983). Within-compound associations in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(4), 390–400. https://doi.org/10.1037/0097-7403.9.4.390

Rescorla, & Wagner. (1972). Theory of Classical Conditioning.

Richards, J. B., Mitchell, S. H., de Wit, H., & Seiden, L. S. (1997). Determination of Discount Functions in Rats With an Adjusting-Amount Procedure. *Journal of the Experimental Analysis of Behavior*, *67*(3), 353–366. https://doi.org/10.1901/jeab.1997.67-353

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, *35*(2), 395–405. https://doi.org/10.1016/S0896-6273(02)00755-9

Rudebeck, P. H., Walton, M. E., Millette, B. H. P., Shirley, E., Rushworth, M. F. S., & Bannerman, D. M. (2007). Distinct contributions of frontal areas to emotion and social behaviour in the rat, *26*(August), 2315–2326. https://doi.org/10.1111/j.1460-

9568.2007.05844.x

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature Reviews Neuroscience*, *15*(8), 549–562. https://doi.org/10.1038/nrn3776

Rutte, C., & Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (Rattus norvegicus): Direct vs generalised reciprocity The role of signals for social decisions in Norway rats View project Cooperative behaviour and ecology in animals View project. *Article in Behavioral Ecology and Sociobiology*. https://doi.org/10.1007/s00265-007-0474-3

Saito, Y., Yuki, S., Seki, Y., Kagawa, H., & Okanoya, K. (2016). Cognitive bias in rats evoked by ultrasonic vocalizations suggests emotional contagion. *Behavioural Processes*, *132*, 5–11. https://doi.org/10.1016/j.beproc.2016.08.005

Saunders, B. T., Richard, J. M., Margolis, E. B., & Janak, P. H. (2018). Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. *Nature Neuroscience*, *21*(8), 1072–1083. https://doi.org/10.1038/s41593-018-0191-4

Schneider, K. N., Sciarillo, X. A., Nudelman, J. L., Cheer, J. F., & Roesch, M. R. (2020). Anterior Cingulate Cortex Signals Attention in a Social Paradigm that Manipulates Reward and Shock. *Current Biology*, *30*(19), 3724-3735.e2. https://doi.org/10.1016/j.cub.2020.07.039

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2011). *Orbitofrontal Cortex and Outcome Expectancies: Optimizing Behavior and Sensory Perception*. (J. A. Gottfried, Ed.), *Neurobiology of Sensation and Reward*. Boca Raton (FL). Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22593899

Schönfeld, L.-M., Zech, M.-P., Schäble, S., Wöhr, M., & Kalenscher, T. (2020). Lesions of the rat basolateral amygdala reduce the behavioral response to ultrasonic vocalizations. *Behavioural Brain Research*, *378*(May), 112274. https://doi.org/10.1016/j.bbr.2019.112274

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schultz, Wolfram. (2016). Dopamine reward prediction-error signalling: a two-component response. *Nature Reviews. Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schwarting, R. K. W., Jegan, N., & Wöhr, M. (2007). Situational factors, conditions and individual variables which can determine ultrasonic vocalizations in male adult Wistar rats.

Behavioural Brain Research. https://doi.org/10.1016/j.bbr.2007.01.029

Seffer, D., Schwarting, R. K. W., & Wöhr, M. (2014). Pro-social ultrasonic communication in rats: insights from playback studies. *Journal of Neuroscience Methods*, *234*, 73–81. https://doi.org/10.1016/j.jneumeth.2014.01.023

Seid-Fatemi, A., & Tobler, P. N. (2015). Efficient learning mechanisms hold in the social domain and are implemented in the medial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, *10*(5), 735–743. https://doi.org/10.1093/scan/nsu130

Seo, H., & Lee, D. (2010). Orbitofrontal Cortex Assigns Credit Wisely. *Neuron*, *65*(6), 736–738. https://doi.org/10.1016/j.neuron.2010.03.016

Silk, J. B., Brosnan, S. F., Vonk, J., Henrich, J., Povinelli, D. J., Richardson, A. S., … Schapiro, S. J. (2005). Chimpanzees are indifferent to the welfare of unrelated group members. *Nature*, *437*(7063), 1357–1359. https://doi.org/10.1038/nature04243

Sirotin, Y. B., Costa, M. E., & Laplagne, D. A. (2014). Rodent ultrasonic vocalizations are bound to active sniffing behavior. *Frontiers in Behavioral Neuroscience*, *8*(NOV), 1–12. https://doi.org/10.3389/fnbeh.2014.00399

Smith, C. J.W., Wilkins, K. B., Mogavero, J. N., & Veenema, A. H. (2015). Social Novelty Investigation in the Juvenile Rat: Modulation by the μ-Opioid System. *Journal of Neuroendocrinology*, *27*(10), 752–764. https://doi.org/10.1111/jne.12301

Smith, Caroline J.W., Mogavero, J. N., Tulimieri, M. T., & Veenema, A. H. (2017). Involvement of the oxytocin system in the nucleus accumbens in the regulation of juvenile social novelty-seeking behavior. *Hormones and Behavior*. https://doi.org/10.1016/j.yhbeh.2017.05.005

Spector, A. C., Klumpp, P. A., & Kaplan, J. M. (1998). Analytical issues in the evaluation of food deprivation and sucrose concentration effects on the microstructure of licking behavior in the rat. *Behavioral Neuroscience*, *112*(3), 678–694. https://doi.org/10.1037/0735-7044.112.3.678

Spence, K. W. (1937). Experimental studies of learning and the higher mental processes in infra-human primates. *Psychological Bulletin*, *34*(10), 806–850. https://doi.org/10.1037/h0061498

Spreckelmeyer, K. N., Krach, S., Kohls, G., Rademacher, L., Irmak, A., Konrad, K., … Gründer, G. (2009). Anticipation of monetary and social reward differently activates mesolimbic

brain structures in men and women. *Social Cognitive and Affective Neuroscience*, *4*(2), 158–165. https://doi.org/10.1093/scan/nsn051

Stalnaker, T. A., Cooch, N. K., Mcdannald, M. A., Liu, T., Wied, H., & Schoenbaum, G. (2014). Orbitofrontal neurons infer the value and identity of predicted outcomes. *Nature Communications*, 1–13. https://doi.org/10.1038/ncomms4926

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*(7), 966–973. https://doi.org/10.1038/nn.3413

Steiner, A. P., & Redish, A. D. (2014). Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nature Neuroscience*, *17*(7), 995–1002. https://doi.org/10.1038/nn.3740

Stevens, J. R., Cushman, F. a., & Hauser, M. D. (2005). Evolving the Psychological Mechanisms for Cooperation. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 499–518. https://doi.org/10.1146/annurev.ecolsys.36.113004.083814

Suchak, M., Eppley, T. M., Campbell, M. W., & de Waal, F. B. M. (2014). Ape duos and trios: Spontaneous cooperation with free partner choice in chimpanzees. *PeerJ*, *2014*(1), 1–19. https://doi.org/10.7717/peerj.417

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, *88*(2), 135–170. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7291377

Taborsky, M., Frommen, J. G., & Riehl, C. (2016). Correlated pay-offs are key to cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1687), 20150084. https://doi.org/10.1098/rstb.2015.0084

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. a., Haney, R. Z., Calu, D. J., Taylor, A. R., … Schoenbaum, G. (2009). The Orbitofrontal Cortex and Ventral Tegmental Area Are Necessary for Learning from Unexpected Outcomes. *Neuron*, *62*(2), 269–280. https://doi.org/10.1016/j.neuron.2009.03.005

Tamir, D. I., & Hughes, B. L. (2018). Social Rewards: From Basic Social Building Blocks to Complex Social Behavior. *Perspectives on Psychological Science*, *13*(6), 700–717. https://doi.org/10.1177/1745691618776263

Tan, L., & Hackenberg, T. D. (2016). Functional analysis of mutual behavior in laboratory rats (Rattus norvegicus). *Journal of Comparative Psychology*, *130*(1), 13–23. https://doi.org/https://doi.org/10.1037/com0000015

Teresa Jurado-Parras, M., Gruart, A., & Delgado-García, J. M. (2012). Observational learning in mice can be prevented by medial prefrontal cortex stimulation and enhanced by nucleus accumbens stimulation. *Learning and Memory*, *19*(3), 99–106. https://doi.org/10.1101/lm.024760.111

Thompson, B., Leonard, K. C., & Brudzynski, S. M. (2006). Amphetamine-induced 50 kHz calls from rat nucleus accumbens: A quantitative mapping study and acoustic analysis. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2005.10.012

Thorndike, E. L. (1933). A PROOF OF THE LAW OF EFFECT. *Science*, *77*(1989), 173–175. https://doi.org/10.1126/science.77.1989.173-a

Timberlake, W., & Grant, D. L. (1975). Auto shaping in rats to the presentation of another rat predicting food. *Science*, *190*(425), 690–692. https://doi.org/10.1126/science.190.4215.690

Tobler, P. N. (2005). Human Neural Learning Depends on Reward Prediction Errors in the Blocking Paradigm. *Journal of Neurophysiology*, *95*(1), 301–310. https://doi.org/10.1152/jn.00762.2005

Tobler, Philippe N, O'doherty, J. P., Dolan, R. J., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, *95*(1), 301–310. https://doi.org/10.1152/jn.00762.2005

Trezza, V., Campolongo, P., & Vanderschuren, L. J. M. J. (2011). Evaluating the rewarding nature of social interactions in laboratory animals. *Developmental Cognitive Neuroscience*, *1*(4), 444–458. https://doi.org/10.1016/j.dcn.2011.05.007

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*(1), 35–57. https://doi.org/doi:10.1086/406755

Tschida, K., Michael, V., Takatoh, J., Han, B. X., Zhao, S., Sakurai, K., … Wang, F. (2019). A Specialized Neural Circuit Gates Social Vocalizations in the Mouse. *Neuron*. https://doi.org/10.1016/j.neuron.2019.05.025

Van Den Berg, C. L., Pijlman, F. T. A., Koning, H. A. M., Diergaarde, L., Van Ree, J. M., & Spruijt, B. M. (1999). Isolation changes the incentive value of sucrose and social behaviour

in juvenile and adult rats. *Behavioural Brain Research*, *106*(1–2), 133–142. https://doi.org/10.1016/S0166-4328(99)00099-6

van der Staay, F. J. (2006). Animal models of behavioral dysfunctions: Basic concepts and classifications, and an evaluation strategy. *Brain Research Reviews*, *52*(1), 131–159. https://doi.org/10.1016/j.brainresrev.2006.01.006

van Gurp, S., Hoog, J., Kalenscher, T., & van Wingerden, M. (2020). Vicarious reward unblocks associative learning about novel cues in male rats. *ELife*, *9*, 1–25. https://doi.org/10.7554/eLife.60755

von Heimendahl, M., Itskov, P. M., Arabzadeh, E., & Diamond, M. E. (2007). Neuronal activity in rat barrel cortex underlying texture discrimination. *PLoS Biology*, *5*(11), e305. https://doi.org/10.1371/journal.pbio.0050305

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., & Rushworth, M. F. S. (2010). Separable Learning Systems in the Macaque Brain and the Role of Orbitofrontal Cortex in Contingent Learning. *Neuron*, *65*(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027

Wang, Y., Fontanini, A., & Katz, D. B. (2006). Temporary basolateral amygdala lesions disrupt acquisition of socially transmitted food preferences in rats. *Learning & Memory*, *13*(6), 794–800. https://doi.org/10.1101/lm.397006

Waskom, M. (2018). Seaborn: Statistical Data Visualization — Seaborn 0.9.0 Documentation. *Sphinx 1.7.4*.

Watson, K. K., & Platt, M. L. (2012). Report Social Signals in Primate Orbitofrontal Cortex. *Current Biology*, *22*(23), 2268–2273. https://doi.org/10.1016/j.cub.2012.10.016

Weingarten, H. P. (1984). Meal initiation controlled by learned cues: Basic behavioral properties. *Appetite*, *5*(2), 147–158. https://doi.org/10.1016/S0195-6663(84)80035-5

Weiss, O., Dorfman, A., Ram, T., Zadicario, P., & Eilam, D. (2017). Rats do not eat alone in public: Food-deprived rats socialize rather than competing for baits. *PLoS ONE*, *12*(3), 1–14. https://doi.org/10.1371/journal.pone.0173302

Wenk, G. L. (1998). Assessment of Spatial Memory Using the T Maze. *Current Protocols in Neuroscience*, *4*(1), 1–7. https://doi.org/10.1002/0471142301.ns0805bs04

Westerink, B., Kwint, H., & DeVries, J. (1996). The pharmacology of mesolimbic dopamine

neurons: a dual-probe microdialysis study in the ventral tegmental area and nucleus accumbens of the rat brain. *The Journal of Neuroscience*, *16*(8), 2605–2611. https://doi.org/10.1523/JNEUROSCI.16-08-02605.1996

Whishaw, I. Q., & Kolb, B. (2009). *The Behavior of the Laboratory Rat: A Handbook with Tests*. *The Behavior of the Laboratory Rat: A Handbook with Tests*. https://doi.org/10.1093/acprof:oso/9780195162851.001.0001

White, N. R., Cagiano, R., Moises, A. U., & Barfield, R. J. (1990). Changes in mating vocalizations over the ejaculatory series in rats (Rattus norvegicus). *Journal of Comparative Psychology (Washington, D.C. : 1983)*. https://doi.org/10.1037/0735-7036.104.3.255

Wickham, H. (2017). The Tidyverse style guide. *R Style Guide*.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*. https://doi.org/10.21105/joss.01686

Wickham, H., Bryan, J., RStudio, Kalicinski, M., Valery, K., Leitienne, C., … Miller, E. (2019). readxl: Read Excel files. R package version 1.3.1.

Wikipedia contributors. (2021). Functional magnetic resonance imaging. In *Wikipedia, The Free Encyclopedia*. Retrieved April 12, 2021, from https://en.wikipedia.org/w/index.php?title=Functional_magnetic_resonance_imaging&oldid=1036719244

Willey, A. R., & Spear, L. P. (2013). The effects of pre-test social deprivation on a natural reward incentive test and concomitant 50kHz ultrasonic vocalization production in adolescent and adult male Sprague-Dawley rats. *Behavioural Brain Research*. https://doi.org/10.1016/j.bbr.2013.02.020

Willuhn, I., Tose, A., Wanat, M. J., Hart, a. S., Hollon, N. G., Phillips, P. E. M., … Wohr, M. (2014). Phasic Dopamine Release in the Nucleus Accumbens in Response to Pro-Social 50 kHz Ultrasonic Vocalizations in Rats. *Journal of Neuroscience*, *34*(32), 10616–10623. https://doi.org/10.1523/JNEUROSCI.1060-14.2014

Willuhn, Ingo, Tose, A., Wanat, M. J., Hart, A. S., Hollon, N. G., Phillips, P. E. M., … Wöhr, M. (2014). Phasic dopamine release in the nucleus accumbens in response to pro-social 50 kHz ultrasonic vocalizations in rats. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *34*(32), 10616–10623. https://doi.org/10.1523/JNEUROSCI.1060-14.2014

Wilmouth, C. E., & Spear, L. P. (2009). Hedonic sensitivity in adolescent and adult rats: Taste reactivity and voluntary sucrose consumption. *Pharmacology Biochemistry and Behavior*, *92*(4), 566–573. https://doi.org/10.1016/j.pbb.2009.02.009

Witten, I. B., Steinberg, E. E., Lee, S. Y., Davidson, T. J., Zalocusky, K. A., Brodsky, M., … Deisseroth, K. (2011). Recombinase-Driver Rat Lines: Tools, Techniques, and Optogenetic Application to Dopamine-Mediated Reinforcement. *Neuron*, *72*(5), 721–733. https://doi.org/10.1016/j.neuron.2011.10.028

Wöhr, M., Houx, B., Schwarting, R. K. W., & Spruijt, B. (2008). Effects of experience and context on 50-kHz vocalizations in rats. *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2007.11.031

Wöhr, M., & Schwarting, R. K. W. (2007). Ultrasonic communication in rats: can playback of 50-kHz calls induce approach behavior? *PloS One*, *2*(12), e1365. https://doi.org/10.1371/journal.pone.0001365

Wöhr, M., & Schwarting, R. K. W. (2008). Ultrasonic calling during fear conditioning in the rat: no evidence for an audience effect. *Animal Behaviour*. https://doi.org/10.1016/j.anbehav.2008.04.017

Wöhr, M., & Schwarting, R. K. W. (2013). Affective communication in rodents: Ultrasonic vocalizations as a tool for research on emotion and motivation. *Cell and Tissue Research*. https://doi.org/10.1007/s00441-013-1607-9

Wright, J. M., Gourdon, J. C., & Clarke, P. B. S. (2010). Identification of multiple call categories within the rich repertoire of adult rat 50-kHz ultrasonic vocalizations : effects of amphetamine and social context, 1–13. https://doi.org/10.1007/s00213-010-1859-y

Yun, I. A., Wakabayashi, K. T., Fields, H. L., & Nicola, S. M. (2004). The Ventral Tegmental Area Is Required for the Behavioral and Nucleus Accumbens Neuronal Firing Responses to Incentive Cues. *Journal of Neuroscience*, *24*(12), 2923–2933. https://doi.org/10.1523/JNEUROSCI.5282-03.2004

Zentall, T. R., & Levine, J. M. (1972). Observational Learning and Social Facilitation in the Rat. *Science*, *178*(4066), 1220–1221. https://doi.org/10.1126/science.178.4066.1220

Zentall, Thomas R. (2012). Perspectives on observational learning in animals. *Journal of Comparative Psychology*, *126*(2), 114–128. https://doi.org/10.1037/a0025381

8. Summary

In this thesis, I set out to identify the behavioral and neural substrates that support the valuation of social reward in male rats. To accomplish this goal, my colleagues and I transformed two non-social behavioral tasks, so that they could be used to measure the value of social rewards.

In the first of the two tasks, the social reinforcement-learning task, we hypothesised that reinforcement of novel sensory stimuli in the environment of the rat with specific social outcomes, could drive the formation of vicarious stimulus-outcome associations (Chapter 1). In the social reinforcement-learning task, we find that male rats can be trained to form these associations and results show that the presentation of cues that are reinforced with a mutual reward (actor and partner reward) show a stronger conditioned response in the actor rat, than cues that are reinforced with an Own-reward (actor self-reward, but no-reward to a partner). We named this the social unblocking effect and find evidence that these stimulus-outcome associations are the product of vicarious reinforcement, as social information exchange between the two rats is necessary for the associations to form.

I then hypothesised in chapter two, that the formation of these social stimulus-outcome associations could be driven by an increased activity in the posterior-lateral Orbitofrontal cortex (PLO). We however find that a lesion in the PLO does not impair the social unblocking effect but it, surprisingly, impairs blocking of responding to cues that predict no-reward to the partner rat.

In the second novel social task, an X-shaped maze was used that contained both a social and non-social reward. We hypothesised that the value of a social reward could be measured by directly manipulating the value of the non-social reward. To determine these relative values, we compared the time the rat spent on the maze with a social reward compared to a non-social reward (Chapter 3). We show that adolescent rats' preference for spending time with a juvenile is directly influenced by the sweetness of an alternatively available sucrose-water reward. The rat spends more time with the juvenile than drinking the sucrose water, when the sucrose-water is less sweet (2%), has no preference when sucrose-water sweetness is 5% and spends more time drinking than interacting with the juvenile when the sucrose water sweetness is 10%. We furthermore find that a higher sweetness of the sucrose-water reward induces a higher rate of 50 KHz vocalizations and in particular the subtype Trill, when the adolescent rat interacts with the juvenile rat.

9. Zusammenfassung

In dieser Dissertation untersuche ich das Verhalten und die neuronalen Substrate, welche die Bewertung von sozialer Belohnung in männlichen Ratten bedingen. Um dieses Ziel zu erreichen haben meine Kollegen und ich, zwei ursprünglich nicht-soziale Paradigmen umgewandelt, sodass diese genutzt werden konnten um soziale Wertschätzung untersuchen zu können.

In dem ersten der beiden Paradigmen, dem sozialen Verstärkungslernen-Paradigma, stellten wir die Hypothese auf, dass soziale Hinweisreize im Umfeld der Ratte geeignet sind, um neue, ungelernte Reize stellvertretend in Stimulus- Konsequenz -Assoziationen zu etablieren (Kapitel 1). In dem sozialen-Verstärkungslernen Paradigma kamen wir zu dem Ergebnis, dass männliche Ratten trainiert werden können diese Assoziationen zu lernen. Unsere Ergebnisse zeigen, dass die Präsentation von Stimuli, welche durch eine gemeinsame Belohnung verstärkt wurden (Akteur und Partner Belohnung) eine stärkere, konditionierte, Reaktion hervorrufen als Reize, welche durch eine einseitige Belohnung verstärkt wurden (Akteur Belohnung – keine Belohnung für den Partner). Wir nennen dies den sozialen Entblockungs-Effekt und finden Evidenz, dass diese Stimulus-Konsequenz Assoziationen das Produkt von stellvertretend gelernter Verstärkung sind, da der Austausch von sozialen Informationen notwendig ist um diese Assoziationen zu etablieren.

Ich stellte weiterhin die Hypothese auf (Kapitel 2), dass die Etablierung dieser Stimulus-Konsequenz Assoziationen auf erhöhter Aktivität im posterioren lateralen Orbitofrontalkortex (PLO) beruhen. Allerdings zeigen unsere Ergebnisse, dass Läsionen im PLO den sozialen Entblockungs Effekt nicht beeinträchtigen. Überraschenderweise beeinträchtigen Läsionen im PLO jedoch Reaktionen auf Stimuli welche die Auslassung einer Belohnung für die Partner-Ratte vorhersagen.

Im zweiten, neuen sozialen Paradigma nutzten wir ein X-förmiges Labyrinth welches zwei Belohnungen bereithielt, eine non-soziale – und eine soziale Belohnung. Wir stellten die Hypothese auf, dass der Wert der sozialen Belohnung gemessen werden kann, indem man die non-soziale Belohnung verändert. Um diese relativen Wertigkeiten bestimmen zu können, nutzten wir als Maß die Zeit, welche die Ratten in dem Teil des Labyrinths mit der sozialen Belohnung verbrachten im Vergleich zu dem Teil mit der non-sozialen Belohnung (Kapitel 3). Wir zeigen, dass die Präferenz von adoleszenten Ratten, Zeit mit Jungtieren zu verbringen, direkt vom Zuckergehalt einer alternativ verfügbaren Zuckerwasser Belohnung beeinflusst wird. Die Ratten verbrachten mehr Zeit mit dem Jungtier als damit das Zuckerwasser zu trinken, wenn das Zuckerwasser weniger süß war (2%). Bei einer mittleren Süße (5%) zeigten die Ratten keine Präferenz und bei einer hohen Süße (10%) kehrte sich die Präferenz um und die Ratten

verbrachten mehr Zeit damit das Zuckerwasser zu trinken als Zeit mit dem Jungtier zu verbringen. Weiterhin fanden wir, dass eine höhere Süße der Zuckerwasser-Belohnung eine höhere Rate von 50KHz Vokalisationen, besonders des Subtyps „Thrill“ induziert, während die adoleszente Ratte mit dem Jungtier interagiert.

10 Affidavit.

"I, Sander van Gurp, declare that I have produced my thesis entitled "Unblocking social associative learning" independently and without any undue assistance by third parties under consideration of the 'Principles for the Safeguarding of Good Scientific Practice at Heinrich Heine University Düsseldorf'. The thesis, in the present or similar form, has not been submitted previously to any other institution. I have not made any unsuccessful attempts to earn a doctoral degree so far.

Sittard, August 2021 _____

11 Epilogue.

The story of my PhD experience starts in the fall of 2015. I was just temporally back at my parents' home after a turbulent time finishing my master thesis in Rotterdam and breaking up with my girlfriend at the time. Luckily, biking in the hills near Sittard, Limburg gave me enough rest to recover and quickly I was already exploring possibilities to continue in science. I saw the advertisement to work in the Social Rodent Lab and there was no doubt: I thought, this is what I want, a new challenge. After studying motor behavior, it was time to go back to studying the thing that was missing in Neuroscience: Psychology. The fact that I was accepted at the Department of Comparative Psychology was beyond my own expectations, as I thought it normally could take months to get a PhD position.

Those first days meeting my supervisor Marijn and seeing for the first time the lab I thought: I can never do this.. All the behavioral setups were still unpacked and it felt like I would never be able to get something running. Luckily, the team was great and very welcoming, both the German side and our little Dutch enclave. Together with my colleague Mireille and my supervisor Marijn (both Dutch), we started our journey to do new and exciting studies in the field of social Neuroscience. That first year I found out what it means to be a neuroscientist; it means becoming an engineer, programmer, animal caretaker and much more. It was great times to work together with students to make a Skinner box run (it is really challenging to make sure that cables run and are connected properly). We even made sure that the rats were not able to hear each other by building a custom made wooden box around the skinner boxes and Plexiglas walls to be able to train two rats at the same time. I have to thank Jochen for the great times and philosophical and technical discussions into how social interaction works and how we together could investigate its connections to the environment around us. At the same time me and Mireille learned new techniques from Marijn to do Electrophysiology and we formed a good team together with Boateng. We ventured into discovering whether playing positive and negative vocalizations to the rats, there were actually neurons in the brain of the rat actively listening. These long days in the red light in the Social Rodent Lab formed the strong basis for making it through my PhD.

There was struggle, and experiments often did not work as expected or were very intense. If I say Prosocial Choice Task, everybody in the Lab will know what I mean. This task was particularly challenging because you had to open a door for a rat and wait for it to enter, but sometimes he would either just not enter, enter but then change his mind and turn around to go into the other room. It was physically intense work but I also enjoyed it and I tried to read the mind of the rat. I discovered three types of rats: One which is very slow and careful and whisks around the corners

of the entry point before it enters, another one goes very fast to the food to eat it and then continues to explore quickly, and finally there is my favorite: the rat that enters in a very chill manner and then really, really enjoys eating those sugar pellets. I was furthermore fortunate enough to go to Marburg and try to spread the knowledge to a phenomenal student Shona there. I had a wonderful time in the fairytale city with castle trips and I learned a lot about the rat, its vocalization repertoire and stayed at the house of soon to be good friends and colleagues. After this more challenges came, but also our first good results and a very nice symposium to show our results to the world at Schloss Mickeln. From that point and onwards, I think I found my way around Düsseldorf, in the lab and made new friends and lovely colleagues.

Marijn. I want to sincerely thank you for all the good times we had. That you gave us the ability to show our work in Copenhagen at FENS, Bilbao at the EBBS and that we visited the US together for the SFN was awesome. Besides that, because of you my Matlab skills excelled and I was able to transform complicated data and make wonderful graphs. More importantly, you were always full of good uplifting energy and you always had our backs and at the same time, you answered all my questions and explained patiently things that I did not understand. Without you, my thesis wouldn't have been half as good as what it is now. Of course, it has to be said you were also the CompssySquad leader and the cool guy who surely liked a Zombie cocktail and has awesome dance moves. Of course, I also had some critique every now and then, but you were luckily always open for a change. I am forgetting now that it was also great that we actually for a short time formed a band and played song covers. Overall, I am very glad you were my mentor, supervisor and that we became good friends.

Tobias. Thanks for welcoming me in the comparative psychology lab and for always having an open door for a chat. At first, I thought the team meetings were challenging, as I am not so good in formal group setting. Luckily, we had an open meeting where most issues could easily be discussed. I would still suggest though to have more data meetings, as I always like to go into the data. I am always amazed by your constant optimism and good spirits and you have been an inspiration to make something of myself. I was also very glad to be part of the skategroup, which helped me a lot to get to know my German colleagues. It was also awesome to skateboard all the way to the city of Zons, were I found my stardom after beer was spilled on my pants, after which I got an orange cape. A day which I will never forget. Thanks also for your comments and helpful thoughts on how to analyze my data.

Adam. Brother, we had such a miserable time together in the lab. It was suffering the whole way through and the PhD made us feel bad, day in day out. However, I am glad we went through that

time together. Biking with you home every day after work got me often through the day, by realizing that there was actually another world out there, which did not take place in the red light rat room. Thanks for all the good philosophical, ultimate cynical, sarcastic and uplifting talks!

Mireille. Mireille, ik heb een leuke en uitdagende tijd gehad samen in het begin en ik vind het nog steeds erg jammer dat je weg ging. Maar ik ben ook erg blij voor je dat je nu bezig bent met je toekomst in te richten zoals jij dat graag wilt.

Social rodent lab Team! Sharing is Caring! :)

Douman; I am happy that you joined the lab and we became such good friends and I'm proud to have worked together with you and that we were able to finish our projects together! *Sammy;* computer programming wizz-kid, your good humor and optimism and dungeons and dragons characters were all awesome and also the fact that you made it through parties till 4 without drinking a single drop!:) *Jochen;* I'm happy that we were such a good duo together; building the boxes and drinking beer was our motto and it was good times. Thanks for being a good friend. *Sonja;* You made me realize that I was not at all very good at making a proper experiment time schedule, thanks so much for being a boss around the lab. *Marlene and Simone;* It was great that you two were there, when times were a bit tough and I think we learned a lot together and you were for sure my favorite students. All other students thank you! You have made me a better teacher and I'm glad that I was able to work with you on our research!

Comparative Psychology Team!

Yue; Thanks for all the good talks and much appreciated data discussions! *Irina;* Thanks for our good chats in the company of good beers and finding out the best place to get and enjoy pizzas! *Maurice;* Thanks for the good times on the balcony and always making such awesome cocktails and cakes. *Sandra;* Thanks to your valuable help and support in the physiolab, I have become a better scientist. *All other people;* I want to say thank you for the great times during all our lab trips; it will forever be good memories.

Finally, I want to express all the love in the world for my parents who have been such great support to me during these years in Düsseldorf. *Mam,* Ik ben zo blij dat je er altijd was en mijn emotioneel mentale leven begreep en me daarin sturing gaf. Ik hou heel veel van jou en ik ben blij dat je er altijd voor me bent. *Pap,* Ik ben zo blij met je altijd positieve en sturende hulp bij het oplossen van problemen en het uitzoeken van oplossingen. Met jouw wetenschappelijke en rationele kijk op de dingen ben ik een betere wetenschapper geworden. Ik ben super trots dat het nu volbracht is mede dankzij jullie altijd aanwezige support!