From DEPARTMENT OF CLINICAL NEUROSCIENCE

Karolinska Institutet, Stockholm, Sweden

SOCIALLY DEPENDENT AVOIDANCE LEARNING

Mechanisms Of Adaptive Behavior

Björn Lindström

Stockholm 2014

Socially dependent avoidance learning
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Björn Lindström

Principal Supervisor:
Dr Andreas Olsson
Karolinska Institutet
Department of Clinical Neuroscience
Division of Psychology

Co-supervisor(s):
Professor Arne Öhman
Karolinska Institutet
Department of Clinical Neuroscience
Division of Psychology

Opponent:
Professor Philippe Tobler
University of Zurich
Department of Economics

Examination Board:
Professor Gustaf Gredebäck
Uppsala University
Department of Psychology

Professor Agneta Herlitz
Karolinska Institutet
Department of Clinical Neuroscience
Division of Psychology

Docent Pontus Strimling
Institute for Future Studies

ABSTRACT

Many aspects of human social behavior are likely to be shaped by punishment, threats of punishment, and avoidance of punishment. However, surprisingly little is known about the psychological mechanisms underpinning these influences. The goal of the research presented in this thesis was to investigate how basic reinforcement-based learning processes contribute to adaptive social behavior. To this end, we investigated how human behavior is shaped by punishment and threats of punishment in simple laboratory models aimed at capturing essential aspects of the social world. In order to probe the mechanisms underlying overt behavior, we tested the capacity of simple computational models to describe human behavior.

In **Study I**, we investigated how the capacity to control behavior was affected by threat of punishment for performance mistakes, motivated by the functional perspective that control is most needed when mistakes are costly, as common in social interaction. Furthermore, we investigated if activity in the facial corrugator supercilii muscle might function as a novel index of this control process. We found that performance was enhanced under threat of punishment relative to non-threat, and that activity in the corrugator supercilii responded to performance errors during threat of punishment, similar to the known electroencelographic correlate of performance errors, the error related negativity.

Studies II and **III** investigated how social stimuli, emotional and out-group faces, affected avoidance learning when serving as punishers (**Study II**) and cues for punishment (**Study III**). In both studies, we were interested in how pre-existing negative attitudes (racial bias) would affect avoidance learning, motivated by the fact that social groups and group animosity play an important role in modern societies. We found that individual differences in racial bias strongly modulated avoidance learning, by leading to an increased learning rate in response to out-group faces (**Study II**) and to more biased behavior when out-group faces served as cues for electric shocks (**Study III**). Furthermore, **Study III** revealed that several types of social and non-social stimuli, which previously have been shown to be preferentially associated with fear (snakes, threatening faces, out-group faces, and guns), all shared a common effect on avoidance learning relative to control stimuli, by promoting adaptive behavior when these stimuli were good predictors of aversive outcomes but corrupting behavior when serving as bad predictors of aversive outcome. This pattern could be explained by a computational model positing two competing valuation systems.

In **Study IV**, we asked how avoidance behaviors can be learned and transmitted socially by observing others´ behavior in a setting with threat of punishment. We used agent-based simulations to show that the omission of expected punishment together with social learning can explain the generation, maintenance and transmission of avoidance behaviors that form "avoidance traditions" on the group-level. The assumptions underlying this model were corroborated in four experiments, which together showed that humans are exceedingly prone to copy the behavior of others when threatened with punishment.

Together, these four studies indicate that important aspects of human social behavior can be understood as involving mechanisms of basic reinforcement learning, and that simple computational models of these learning processes can provide important insights into human behavior.

LIST OF PUBLICATIONS

This thesis is based on the following publications, which are referred to in the text by their roman numerals (**Study I-IV**)

I. **Lindström, B**., Berglund Mattsson-Marn, I., Golkar, A., Olsson, A. (2013). In your face: Risk of punishment enhances cognitive control and error-related activity in the corrugator supercilii muscle, *Plos One*, DOI:10.1371/journal.pone.0065692

II. **Lindström, B**., Selbing, I, Molapour, T., Olsson, A. (2014). Racial bias shapes social reinforcement learning, 25(2), *Psychological Science*.

III. **Lindström, B**., Golkar, A., Olsson, A. Intrinsic stimulus value causes adaptive and maladaptive voluntary behavior in aversive environments. *Submitted manuscript.*

IV. **Lindström, B**., Olsson, A. Threat of punishment can create adaptive and arbitrary behavioral traditions through social learning in humans. *Submitted manuscript.*

ADDITIONAL PUBLICATIONS

Publications by the author from the Department of Clinical Neuroscience which are not included in the thesis:

I. **Lindström, B. R**., & Bohlin, G. (2011). Emotion processing facilitates working memory perfomance. *Cognition and Emotion, 25(7),* 1196-1204.

II. **Lindström, B. R**., & Bohlin, G. (2012). Threat-relevance impairs executive functions: negative impact on working memory and response inhibition. *Emotion, 12(2),* 384-393

III. Öhman A., Soares S.C., Juth, P., **Lindström, B.R**., Esteves, F. (2012). Evolutionary derived modulations of attention to two common fear stimuli: Serpents and hostile humans. *Journal of Cognitive Psychology, 24(1),* 17-32.

IV. Selbing, I., **Lindström,** B., Olsson, A. (2014). Demonstrator skill modulates observational aversive learning, *Cognition, 133(1)*, 128-139.

V. Soares, S., **Lindström,** B., Esteves, F. Öhman, A. Automatic Capture of Attention by Snakes may reflect an Evolutionary Heritage.
Submitted manuscript.

VI. **Lindström,** B., Selbing, I., & Olsson, A. Evolutionary preparedness and social learning interact to promote survival in dangerous environments. *Submitted manuscript*.

CONTENTS

LIST OF ABBREVIATIONS

ACC	Anterior Cingulate Cortex
cEMG	Corrugator EMG
CS	Conditioned Stimulus
EEG	Electroencephalography
EMG	Electromyography
ERN	Error-related Negativity
fMRI	Functional Magnetic Resonance Imaging
RL	Reinforcement Learning
R-W	Rescorla-Wagner model
US	Unconditioned Stimulus

1 INTRODUCTION

1.1 LEARNING IN THE AVERSIVE AND SOCIAL DOMAINS

The overall theme of this thesis is how humans learn to avoid different types of punishment, such as uncomfortable electric shocks or frowning facial expressions (the *aversive domain*), in simple experimental laboratory models of social interaction (the *social domain*). The questions arising from this focus are likely: what is so important about learning in the aversive domain? And why is this relevant in the social domain? I will begin with sketching answers to these questions by providing an overview of the wider context this thesis can be placed in and thereafter summarize the structure of the thesis.

Dangers, such as predation, noxious foods or hostile con-specifies are ubiquitous in the natural world. Exposure to such dangers is by definition *costly*, in terms of Darwinian fitness (Johnson, Blumstein, Fowler, & Haselton, 2013). An erroneous decision in a dangerous environment might be the last decision made by that individual, providing a strong selective pressure for optimizing behavior to avoid potentially dangerous consequences. In simple organisms living in stable environmental conditions, avoiding danger can be handled by hard-wired genetic behavioral programs (Fawcett, Hamblin, & Giraldeau, 2012). However, in most species, the environment is complex and stochastic, which requires behavioral flexibility in the form of learning (Johnston, 1982). Animals learn to avoid objects or behaviors followed by naturally aversive consequences, such as pain or nausea. Thus, understanding how animals learn avoidance of punishments provides an answer to a fundamental aspect of survival (Bolles, 1970).

In humans, punishment and avoidance of punishment can take many forms. Aside from simple primary punishments such as pain, which motivates behavior across species, humans are influenced by many other types of punishments, many of which are distinctly social (Seymour, Singer, & Dolan, 2007). Loss of money, threats to social status, social disapproval, gossip, ostracism and so forth are powerful influences on human behavior (note that some of these punishment types can also be relevant for other animals, Cant, 2011). Most importantly, punishment and threat of punishment are believed to be fundamental mechanisms for maintaining social norms, which can be defined as group-shared expectations of how members of the group should behave in a given situation (Bicchieri, 2005; Fehr & Fischbacher, 2004). The perhaps most important social norms are those prescribing cooperation between unrelated individuals (Bowles & Gintis, 2011). Cooperation is often viewed as a mystery when viewed from an evolutionary perspective; how can cooperation emerge and be sustained between self-interested individuals? Punishment, in a variety of social and non-social currencies, appears to be an important explanatory mechanism for this evolutionary puzzle (Boyd, Gintis, & Bowles, 2010; Boyd, Gintis, Bowles, & Richerson, 2003; Boyd & Richerson, 1992; Fehr & Gächter, 2002): if people are aware of the existence of norm prescribing cooperation and believe that transgressions might be punished, cooperation can be maintained. A more mundane example might be the norms prescribing

which side of a street or an escalator one should use when walking (Bicchieri, 2005). Transgressions of such norms are likely to incur social punishment, in the form of facial expressions of disapproval or even verbal threats. Importantly, information about possible sources of punishment in the environment can often be acquired by observing the behavior of others. Such social learning is thought to be important for many species (Griffin, 2004; Olsson & Phelps, 2007), but takes especially complex forms in humans (Boyd, Richerson, & Henrich, 2011). In summary, much human social behavior is likely to be shaped by punishment, threats of punishment, and avoidance of punishment. However, surprisingly little is known about the psychological mechanisms underpinning these influences.

The overall aims of this thesis was therefore to identify how punishment affects adaptive behavior in the social domain, by asking how threat of punishment influences the ability to control behavior (**Study I**), how the value of social stimuli, such as emotional faces, influences avoidance learning when serving as punishments or as cues predicting punishment (**Study II-III**), and how humans acquire and transmit avoidance behavior through observing the actions of others when threatened with punishment (**Study IV**). Together, these studies seek to triangulate some of the basic psychological mechanisms underpinning aversive learning in the social domain. The first part of this thesis (Introduction) will start with clarifying the theoretical framework on which this thesis builds by discussing the relationship between evolutionary function and psychological mechanisms. Then I will review the functional basis for learning, as well as characterizing two primary forms of associative learning. Thereafter, I will briefly describe the basic computational mechanisms of learning and decision making. Finally, I will summarize some important aspects of the neural mechanisms underlying learning with a specific focus on how these might relate to, and possibly explain, some of the behavioral findings presented in this thesis.

1.2 FUNCTION AND MECHANISM IN THE BEHAVIORAL SCIENCES

The questions posed in the studies included in this thesis are derived from a particular way of viewing both the purpose and the explanations given by the behavioral and cognitive sciences (which will be used synonymously for simplicity, as even animal learning theory involve "cognitive" constructs such as attention). In short, I believe that understanding behavior both in the light of its putative evolutionary function and it´s mechanistic basis are important and mutually illuminating. This view is best defined by considering the relationship between two classical models, or frameworks, in biology and cognitive science: Tinbergen´s "four questions" and Marr´s "levels of explanation". Starting with the former, the Nobel laureate Nikolas Tinbergen (Tinbergen, 1963) outlined the "major problems of biology" which biologists working with behavior were trying to address. Full understanding of a behavioral trait requires answering all four questions, but they are in principle thought to be independent from each other.

These problems (sometimes phrased as questions) are; (*i*) what is the survival value of a behavioral trait, (*ii*) what is it´s ontogeny, (*iii*) how did it evolve, and (*iv*) what is it´s causation. Of these four questions, *i* (survival value) and *iv* (causation) are of direct relevance for this thesis. In a modern treatise on Tinbergen´s question, Bateson and Laland (Bateson & Laland, 2013) suggested that survival value should be rephrased as *current utility*, to clarify that the current survival value of a trait could be distinct from the survival value during earlier points of the species evolutionary trajectory (e.g., feathers originated for temperature regulation but were co-opted for flight). Regardless, Tinbergen´s first question refer to the function of a trait in terms of how it benefits the survival of the organism. Investigating the function of behavioral traits is typically the area of behavioral ecology (Winterhalder, Smith, & Bruce Winterhalder, 2000). The forth question, causation, refer to the mechanisms that generate a particular behavior. Most fields of psychology and cognitive science are typically interested in mechanistic explanations of behavior. Mechanistic explanations are complicated by the fact that causation can be understood on different levels of explanation simultaneously. For example, explanations of human behavior in crowds can be posed either as due to physical interactions between individual agents, or by the psychological motivations governing these interactions, or even by the neural events underlying such motivations at the level of the brain (and as the interaction of atoms, and so on). In biology, a successful example of how several levels of explanation can be bridged is the study of birdsong, which integrates mechanisms involved in gene expression, specific brain nuclei and the sensory input that underlie expression and learning of song (Bateson & Laland, 2013). The problem of linking levels in understanding psychological mechanisms will introduced in relation to Marr´s levels of explanation below. Tinbergen´s questions *i* and *iv* can be restated as a distinction between the ultimate and proximate level of explanation. The ultimate level of explanation asks why a behavior exists, whereas the proximate level of explanation asks how it works (Laland, Odling-Smee, Hoppitt, & Uller, 2012). Human behavioral science (psychology and cognitive science) have traditionally focused on the latter level by trying to explain how psychological mechanisms, such as attention or decision-making, works, mainly through direct experimental studies. This thesis can be placed in this tradition, but Tinbergen´s first question has been an underlying inspiration for many of the topics and questions addressed in the studies. Mostly, this inspiration is implicit, but in **Study III** and **IV** some more direct allusions to function or adaptive value are made.

Motivating this mixture of explanational levels is my conviction that the questions worth asking about human psychology are those where asking about the survival value of the psychological trait also is meaningful. The specific perspective referred to at the outset of the text is thus to ask questions about psychological mechanisms in the light of evolution. Here, it should be emphasized that this perspective is not based upon speculation about human evolutionary history (so called "just so" stories), which unfortunately is relatively common (Bolhuis, Brown, Richardson, & Laland, 2011; Laland & Brown, 2011; Panksepp & Panksepp, 2000). Rather, a fruitful approach might be to investigate the psychological mechanisms underlying behavioral predictions from formal (evolutionary) models, or to

device such models that might explain why a particular trait or psychological mechanism is observed (i.e., what is its current utility: McNamara & Houston, 2009; Trimmer, McNamara, Houston, & Marshall, 2012). More informally, meaningful questions can be posed by devising laboratory experiments where the dependent variable by some stretch of imagination is related to fitness (e.g., the ability to avoid punishment). In the fields of biology and behavioral economy, as well as intersections of these fields such as behavioral ecology and neuroeconomics, investigating how a behavior works is intimately tied to why (e.g., human cooperation) it exists. Economy and evolutionary biology share normative theories, which make clear predictions of what behaviors should be observed given some underlying assumptions forming an analytical core (such as rationality or maximization of expected value) (Hammerstein & Hagen, 2005; McNamara & Houston, 2009). While such predictions are not always confirmed, failures commonly result in theory development. In contrast, many areas of psychology seem to lack both normative theories and theory development, and thus little progress appears to be made (Gintis, 2007). For example, many psychologists appear to glee in the fact that humans are not perfectly rational, without asking how such deviations from rationality can be understood from a normative or evolutionary perspective (Houston, McNamara, & Steer, 2007). Unsurprisingly, I have tried to place this thesis closer to the former than the latter approach.

The topics touched upon in this thesis can be seen as related to adaptive behavior, which in theory function to maximize the survival of the individual. Adaptive behavior must however be realized by psychological mechanisms. Psychological mechanisms might have built in constraints with important implications for understanding how human behavior both comply to the predictions of normative theories, but also divergence from such predictions (Fawcett et al., 2014, 2012; McNamara & Houston, 2009). Closely related to this issue is the growing realization within evolutionary biology that adaptive function is constrained by the workings of evolved mechanisms, which often are co-opted between different behavioral domains (Fawcett et al., 2012; McNamara & Houston, 2009). One such fundamental mechanism, which is the major focus of this thesis, is the capacity for learning which contribute to adaptive behavior in a wide variety of settings (Bouton, 2007). Next, I expand upon the idea of psychological mechanisms by referring to the framework posed by Marr (Marr, 1982; McClamrock, 1991).

Marr proposed a general framework for relating different levels of explanation in the cognitive sciences (where I include psychology for simplicity). Conceptually, this framework resembles the distinctions made by Tinbergen, but the domain of Marr´s framework is limited to Tinbergen´s causation question. The three levels in Marr´s framework are: (i) the computational level), (ii) the algorithmic level, and (iii) the implementation level. These should, as in Tinbergen´s framework, be seen as complementary and near independent perspectives on how a cognitive process works. The computational level concern the nature of the problem being solved by the cognitive system, the information involved in solving it, and the logic by which it can be solved. The algorithmic level specifies the representations and processes by which solutions to the problem are computed. Finally, the implementational

level specifies how these representations and processes are realized in neural terms. In this thesis, the primary focus is on the two first levels of explanation, although the knowledge of the implementational level is occasionally used to illuminate or constrain explanations on the other levels (especially in **Study I**). Ultimately, all three levels have to be understood, as well as the mapping between them, to explain how adaptive behavior is instigated. In **Studies II-IV**, the computational and algorithmic – level analysis is conducted by constructing formal computational models in seeking to understand the mechanisms underlying behavior.

1.3 THE FUNCTION OF LEARNING

The basic capacity for learning, evident in humans, all vertebrates and many other species is an evolutionary adaption that allows flexibility in the mapping between input stimuli and behavioral output (Fawcett et al., 2012; McNamara & Houston, 2009; Mery & Burns, 2009; Snell-Rood, 2013). At its core, the function of learning is to allow the organism to procure rewards, such as food and mates, and avoid danger, such as predators, hostile con-specifics and toxic food. Such survival promoting functions can be implemented by wholly genetic "hardcoded" programs if the environment is highly stable (Johnston, 1982; Snell-Rood, 2013). However, if conditions change, for example if new predators emerge or the distribution of food sources change, such hardcoded programs are likely insufficient for adaptive behavior. Learning allows the organism to adapt its behavior to such changes through experience, by associating stimuli or actions with aversive or beneficial outcomes.

A basic distinction is made between two types of associative learning: *instrumental* (or operant) *conditioning*, where animals learn to perform behaviors to enhance their fitness, and *classical* (or Pavlovian) *conditioning*, where animals learn that a stimulus (referred to as conditioned stimulus, CS) predicts a biologically relevant outcome, such as food or pain, (referred to as unconditioned stimulus, US, or reinforcer synonymously throughout the thesis) (Bouton, 2007). In instrumental conditioning, the relationship between behavior and reinforcement can be described by the *law of effect*, first described by Thorndike (Thorndike, 1898). The law of effect states that the probability or frequency of a behavior is affected by if performing the behavior lead to the increased probability of a positive reinforcement or decreased probability of an aversive reinforcement. The law of effect is summarized in Table 1. Here, it is assumed that a "Positive reinforcement" refers to an outcome with positive survival value (such as food or mating) and "Aversive reinforcement" refers to an outcome with negative survival value (such as pain or food toxicity). The consequence of this law is that animals generally will behave to maximize beneficial outcomes and minimize costly outcomes, in conceptual resemblance to how genetic evolution adapts organisms to their environment by enhancing beneficial traits and weeding out costly traits (as famously noted by Skinner, 1981). The studies included in this thesis exclusively focus on aversive reinforcement. Note that many stimuli can acquire reinforcing properties by pairing with primary reinforcers. Such secondary reinforcers are exceedingly important for human

behavior, as exemplified with money which functions as a powerful reinforcer in its own right.

Effect of behavior	Positive reinforcement	Aversive reinforcement
Produces reinforcement	Reward ↑	Punishment ↓
Prevents reinforcement	Omission ↓	Avoidance/escape ↑

Table 1. The law of effect. Arrows indicate whether behavior increases or decreases in frequency. After Bouton (2007)

Classical conditioning similarly functions to adapt the animal to its environment (Domjan, 2005). By learning that a stimulus predicts the occurrence of reinforcement, the animal can respond beforehand. Generally described, a stimulus predicting reinforcement (a CS), elicits responses related to the valence of the reinforcer, leading to approach of a CS predicting a positive reinforcement and avoidance of a CS predicting an aversive reinforcement. Such responses are called *prepatory* because they prepare the animal for interaction with the reinforcer (Bouton, 2007). Importantly, prepatory behaviors are not identical to those elicited by the reinforcer, and also are specific to different classes of reinforcers (Bouton, 2007; Dayan & Seymour, 2007). As shown in Table 2, classical conditioning shares the same mirror-pattern as instrumental conditioning; animals tend to approach stimuli predicting a reduction of aversive reinforcement and avoid stimuli predicting a reduction of positive reinforcement. Apart from the overt behavioral responses, the CS comes to elicit a host of bodily responses, such as physiological arousal or salivation. It is thought that the behavioral and physiological responses to the CS reflect a common *central* (i.e., brain) *state*, which often is identified with specific emotions. For example, classical conditioning with aversive reinforcement (typically electric shock) is termed *fear conditioning*. Fear conditioning evokes a constellation of responses including changes in blood pressure, skin conductance, heart rate and respiration that prepare the animal for defense, which can be summarized as *fear* (Ohman & Mineka, 2001). In humans, the state of fear also involves subjective negative feelings, but these are not necessarily considered to have a causal role or function for the physiological and behavioral indices (LeDoux, 2012). Several influential theories of emotion and motivation are based in the responses to, and in prediction of, various reinforcers (Rolls, 2000; Solomon, 1980). For example, while predictions of aversive reinforcers can result in fear, omission of a predicted aversive reinforcer results in relief (Fujiwara, Tobler, Taira, Iijima, & Tsutsui, 2009; Kim, Shimojo, & O'Doherty, 2006; Solomon, 1980).

14

	Positive reinforcement	Aversive reinforcement
S predicts reinforcement	Approach S	Avoid S
S predicts no reinforcement	Avoid S	Approach S

Table 2. Response tendencies in classical conditioning. S = stimulus. After Bouton (2007)

There are many important similarities between instrumental and classical conditioning. When a response (in instrumental conditioning) or a stimulus (in classical conditioning) stops being predictive of reinforcement, both learning processes exhibit *extinction*, in which the probability of a responses decreases, often gradually over time. Similarly, both processes are influenced by the magnitude of the reinforcer (higher magnitude typically generates stronger and faster acquisition of responding), as well as the timing of the reinforcer (subsequent to the response or the CS, and most efficient when presented in close temporal proximity). However, a range of both behavioral and neurological evidence indicates some degree of independence between instrumental and classical conditioning (Rescorla & Solomon, 1967). A CS, which has acquired predictive value through classical conditioning, can either enhance or suppress instrumental behavior if presented concurrently with the instrumental task. As predicted by combining Tables 1 and 2, the direction of the influence depend on the congruency with the instrumental action; an aversive CS suppresses instrumental behavior for positive reinforcers but enhances instrumental avoidance behaviors, while an positive CS enhances instrumental behavior motivated by positive reinforcers, and suppresses or corrupts instrumental avoidance behaviors (although not much research has been conducted regarding this last combination: Dayan & Seymour, 2007). This phenomenon is termed *Pavlovian-instrumental transfer* (Bouton, 2007; Dayan & Seymour, 2007; Talmi, Seymour, Dayan, & Dolan, 2008), and serves as the conceptual basis for **Study II** and **III**.

Importantly, both instrumental (Burke, Tobler, Baddeley, & Schultz, 2010; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012) and classical associations (Olsson & Phelps, 2007) can be socially learned by observing the actions of others or by observing the reinforcements delivered to others. As mentioned above, social learning is highly functional as it mitigates the need for direct experiences to support learning (Boyd et al., 2011; Griffin, 2004; Olsson & Phelps, 2007; Rendell et al., 2010; Rendell, Fogarty, & Hoppitt, 2011). In humans, the most well-studied type of social learning, particularly in the aversive domain, is observational fear learning (Olsson & Phelps, 2007). In observational fear learning, a CS acquires fear-invoking properties through its association with overt indices of a con-specific´s pain or discomfort. This suggests that the emotional behavior of others can function as reinforcers

(Olsson & Phelps, 2007). It is in principle possible that the reinforcing properties of others pain expressions reflects second-order conditioning (e.g., if others´ pain expressions reliably are paired with direct aversive experiences) (Dawson, Avarguès-Weber, Chittka, & Leadbeater, 2013), but studies involving non-human primates with controlled rearing conditions suggest that this probably is not the case (Cook & Mineka, 1990). Because most previous research has involved classical conditioning, it is unclear of humans learn avoidance behaviors from others. In **Study IV**, we sought to extend observational fear learning to instrumental behaviors, motivated by the central role threat of punishment is thought to play in human social behavior (Bicchieri, 2005; Bowles & Gintis, 2011; Fehr & Fischbacher, 2004).

1.4 COMPUTATIONAL AND ALGORITHMIC MECHANISMS OF LEARNING

The most classic and widely cited theory of classical conditioning is the Rescorla-Wagner (R-W) model (Rescorla & Wagner, 1972) (which builds upon the earlier model by Bush & Monsteller: Glimcher, 2011). The R-W model was designed to explain the fundamental computational mechanisms involved in classical conditioning, and was able to account for a range of, at that time, puzzling phenomena (R. R. Miller, Barnet, & Grahame, 1995). The predictive (or associative) value (V) of the CS on trial $t + 1$ is given by the following equations:

$$V(t + 1) = V(t) + \alpha\beta\delta_i(t) \tag{1}$$

$$\delta_i(t) = R(t) - \sum V(t) \tag{2}$$

where α ($0 < \alpha < 1$) is a parameter (commonly referred to as the learning rate) which is determined by the salience of the CS, β ($0 < \beta < 1$) determines the salience of the reinforcer, and R is the reinforcement on trial t. The most important part is equation two (which commonly is referred to as the *prediction error*), which indicates how surprising the reinforcement was. When the CS perfectly predicts R, $\delta = 0$. The term $\sum V$ is the summed predictive value all stimuli presented on trial t. This summation is critical for the success of the R-W model as it allows explaining how compound stimuli affect learning, such as the blocking phenomenon (Rescorla, 1988). However, in many situations, the model can be simplified by just considering the specific CS presented on trial t. Furthermore, the β parameter is often removed to limit the number of free parameters in the model. This simple model can account for the effect of reinforcement magnitude (by the value of R), extinction after non-reinforcement, and that the CS must be *predictive* (i.e., the correlation between the CS and the reinforcer must be stronger than the correlation of no-CS with the reinforcer, Rescorla, 1988) of the reinforcement to produce learning. It is worth noting that the R-W model assumes the measured performance variable (e.g., skin conductance) to be linearly related to the predictive value.

In contrast to the algorithmic account of classical conditioning represented by the R-W model, the most famous models of instrumental conditioning aim to describe equilibrium behavior across an experimental session. The impetus for these models was discovery of the probability matching phenomenon (Herrnstein, 1970), which entail that animals on average do not maximize reception of reward when choosing between two options (i.e., concurrent schedules of reinforcement), but rather perform each action in proportion to the amount of reward received by the action. This relationship, termed the matching law, can in its simplest form be summarized as

$$\frac{B_1}{B_2} = \frac{R_1}{R_2} \qquad\qquad [3]$$

where B_1 is the frequency of the first action, B_2 the frequency of the second action, and R_1 and R_2 the corresponding frequencies of reinforcement. This empirical relationship has motivated a very large literature because probability matching is sub-optimal; the optimal strategy would be to maximize performance of the action with the highest average reward. It has for example been suggested that probability matching emerge because animals seek to maximize the momentary rate of reinforcement (Herrnstein, Loewenstein, Prelec, & Vaughan, 1993). It is however clear that many behavioral strategies can result in matching (Houston & McNamara, 1981). The matching phenomenon is outside the scope of this thesis, but it can be noted for completeness that average behavior in **Study II** and **III**, which most closely resembles concurrent schedules of reinforcement, was consistent with probability matching.

1.5 FROM LEARNING TO VALUE-BASED DECISION MAKING

On a very general level of description, the function of learning is to sub-serve value-based *decision making* (Gintis, 2007). Value-based decision making takes places whenever animals choose between different actions based on estimates of their value, from foraging to financial decisions (Rangel, Camerer, & Montague, 2008). Such estimates can be influenced by experience through individual learning, by observing the actions of others through social learning as well as by evolutionary predispositions tied to specific stimuli or actions. All studies included in this thesis can be viewed as involving value-based decision making in the aversive domain. In recent years, the understanding of the mechanisms underlying value-based decision making has made great progress, and an integrative framework has been proposed: value-based decision making require the joint operation of a number of computational mechanisms; *representation* of the feasible actions, internal states, and external states, *valuation* of each action given the internal and external states, *action selection* based on these valuations, *outcome evaluation* of the outcomes and states that followed the action, and *learning*, by which the representations and valuations are updated based on the outcomes (Rangel et al., 2008). Naturally, the current knowledge about these mechanisms differs. Ideally, each mechanisms, and how they interact, should be described on all three of Marr´s levels, but for most of these mechanisms, the current state of knowledge is far from ideal. For the representation process, almost nothing is currently known. In the studies included in this thesis, representation is assumed to accurately correspond to the stimuli

presented and actions possible in the experimental task. Much more is known about the valuation process, which is of particular importance for **Study II** and **III** wherefore a brief summary of the computational and implementation mechanisms of valuation is presented below.

On the basis of a sizable body of human and animal behavioral evidence, the existence of three different types of valuation systems has been postulated (Dayan, 2009; Dayan & Niv, 2008; Rangel et al., 2008). However, the exact number of systems, and how independent they are from each other, is currently unknown. The three types of systems should be viewed as an operational division of the valuation processes according to the different types of computations that are known to be involved (Rangel et al., 2008). These systems partially mirror the different types of learning, instrumental and classical (Pavlovian) conditioning, which were described above. The *Pavlovian systems* assign value to a small set of behaviors in response to stimuli with intrinsic or acquired evolutionary significance, as when paired with an evolutionary valuable reinforcer (Dayan, Niv, Seymour, & Daw, 2006). As described in Table 2, these Pavlovian responses are organized along an approach-avoidance dimension and serve to optimize the interactions of the animal with its environment. The limited behavioral repertoire of the Pavlovian system is it´s hallmark relative to the instrumental systems. Importantly, the expression of Pavlovian behaviors are not automatic; the fact that Pavlovian responses can be interrupted by other brain systems (as when humans override the Pavlovian impulse to overeat in the presence of food) indicate that they must be assigned a value which competes with other possible actions (Dayan et al., 2006; Rangel et al., 2008). As noted above, such competitive interactions between Pavlovian responses and instrumental responding are expressed in the Pavlovian-instrumental transfer phenomena. The Pavlovian responses can thus be viewed as a set of action priors over which the instrumental systems operate (Seymour & Dolan, 2008). On the level of functional neurobiology, the Pavlovian system involves sub-cortical brain structures such as the amygdala and periaquductal grey (LeDoux, 2012; Seymour & Dolan, 2008). Furthermore, Pavlovian responses to certain, so called *evolutionary prepared*, stimuli appear to underlie important learning biases in humans and other primates. Stimuli such as snakes or threatening con-specific faces exhibit enhanced fear conditioning when paired with an aversive reinforcer relative to control stimuli, which have been explained as due to the evolutionary history of these stimuli (Ohman, 2009; Ohman & Mineka, 2001). The preparedness phenomena is thought to underlie the uneven distribution of human fears and phobias, where for example snakes are common objects of phobias relative to cars or electric outlets which likely are more common dangers in the modern world (Ohman & Mineka, 2001). The influence of the intrinsic value of prepared stimuli on value-based decision making is explored in **Study III**. Aside from the Pavlovian valuation system, current evidence indicates the existence of at least two *instrumental* valuation systems; the *habitual system* and the *goal-directed system*. As implied by the name, these systems underlie instrumental conditioning, and can flexibly map arbitrary responses to desired outcomes. The distinction between the instrumental systems are not of primary importance for this thesis, but is briefly summarized for completeness. It is generally thought

that the goal-directed system, which involves a "cognitive map" of the relationship between actions and outcomes, is recruited when confronted with novel situations, and the habitual system, which involves simpler stimulus-response associations, successively assumes control during the course of training (Dolan & Dayan, 2013). Behaviorally, these systems can be dissociated by manipulations of the value of an outcome (e.g., water), for example by allowing the animal to satisfy its need for it (by drinking to satiety; reinforcer devaluation). Because the goal-directed system involves a representation of the identity of the outcome, such devaluations will directly lead to reduced responding. In contrast, the habitual system, which is believed to only keep a "cached" one-dimensional (scalar) representation of the value of the action, is insensitive to the devaluations; the animal will perform as if still thirsty, until the reinforce is consumed in the new devalued state. Neurally, the goal-directed system appears to mainly involve parts of the pre-frontal cortex (Dolan & Dayan, 2013). In contrast, the habitual system mainly involves the striatum (Dolan & Dayan, 2013). Both systems are important for adaptive behavior, but the precise computational mechanisms governing their interaction are an area of active research (Daw, Niv, & Dayan, 2005; Dolan & Dayan, 2013).

In comparison to the relatively well-characterized valuation process, little is known about the mechanisms of action selection. It is clear that the estimates by the different valuation systems must be compared and possibly integrated in some way to produce adaptive behavior. The most sophisticated proposal of this comparison process pertains to goal-directed versus habitual instrumental control. Daw and colleagues (Daw et al., 2005) suggested that these systems compete for behavioral control, and that the contribution of each system to action selection is weighted by the uncertainty of their respective value estimates, thus producing optimal control. The notion of competition between valuation systems for action selection have also been extended to the Pavlovian system, which is viewed as a hardwired influence on action selection which can either enhance or corrupt instrumental behavior depending on the correspondence between the Pavlovian responses and optimal action (Dayan et al., 2006). In **Study III,** we suggest that stimuli with intrinsic (e.g., prepared) negative value influences value-based decision making as predicted by a computational model where the Pavlovian and the instrumental systems compete for action selection at the time of choice. In many models of value-based decision making, the action selection stage is implemented with a logistic function known as *Softmax* (Daw, 2011):

$$P_A(t) = \frac{e^{Q_A(t)/\beta}}{\sum_{i=1}^{n} e^{Q_i(t)/\beta}} \qquad\qquad [4]$$

Here, Q_A refer to the expected value of choosing action A on trial $t + 1$. As seen, the probability of choosing action A is determined by the ratio of the expected value of A and the sum of all possible actions, scaled by the parameter β (not to be confused with the β parameter in the R-W model). The parameter β is commonly referred to as the *exploration vs. exploitation* parameter. Higher values of β increase the probability of choosing another action than the one with the highest expected value, which can be beneficial for ensuring continued exploration of the action space (Sutton & Barto, 1998). The Softmax function describes human decision making well (Daw, 2011; Daw, O'Doherty, Dayan, Seymour, & Dolan,

2006), although alternative decision functions are possible (e.g., ε-greedy). If Q_i is interpreted as the valuation of the instrumental system, it is clear that other terms can added within the Softmax function to represent competition between valuation systems (e.g., Huys et al., 2011). This approached is used in **Study III** to model competition between the Pavlovian and instrumental systems.

As for action selection, little is known about outcome evaluation. Human fMRI studies have identified the orbitofrontal cortex (OFC) as a nexus for value related signals (Hare & O'Doherty, 2008; Padoa-Schioppa & Assad, 2006; Plassmann, 2010; Rangel et al., 2008) . Such signals generally correlate with subjective reports about the quality of the experience, across modalities. Furthermore, such responses can also be identified for secondary reinforcers, such as money. Thus, the OFC appears to represent valuation of outcomes as a common currency.

Finally, the learning process has been focus of intense attention within psychology and cognitive neuroscience under the guise of *reinforcement learning* (RL). Reinforcement learning was originally formulated to solve adaptive control problems within the field of artificial intelligence, but the inspiration mainly came from animal learning theory (Glimcher, 2011). A simple formulation of a RL model (Q-learning) is described below:

$$Q_i(t + 1) = Q_i(t) + \alpha \delta_i(t) \qquad\qquad [5]$$

$$\delta_i(t) = R_i(t) - Q_i(t) \qquad\qquad [6]$$

As above, Q_i refer to the expected value of choosing action i on trial $t + 1$. As seen, this formulation is identical to the R-W model, but the Q-values represent estimates of expected value instead of predictive value/associative strength as in the original R-W formulation (Glimcher, 2011). Combined with the *Softmax* function (equation 4), equations 5-6 constitutes a simple model of value-based decision making. Variations of this model (equation 4-6) are used in **Study II-IV** to model human behavior. However, much of the enthusiasm for RL within psychology and cognitive neuroscience can be traced to the finding that the mammalian brain appears to actually compute prediction errors during the course of learning. The implementational mechanisms sub-serving RL are briefly reviewed below.

1.6 IMPLEMENTATIONAL MECHANISMS OF LEARNING AND DECISION MAKING

The best understood aspect of learning concern how the habitual system learns by trial and error in a manner resembling the R-W model. The basic idea is that a prediction error signal is computed by comparing the actual outcome of an action or state with its predicted outcome (equation 2). Such prediction error signals have been identified in the mammal brain, originally using single cell recordings in the dopaminergic midbrain of macaques during reward learning (Waelti, Dickinson, & Schultz, 2001) and later using functional magnetic resonance imaging (fMRI) in humans (Seymour et al., 2004). Recently, causal manipulations have also been conducted in behaving rodents using optogenetic methods, thus showing that

dopaminergic prediction errors have a casual role in learning (Steinberg et al., 2013). Activity in dopaminergic neurons during learning shows a pattern resembling how reinforcements affect behavior (Table 1). An unexpected positive reinforcement leads to increased phasic firing in these neurons, while the omission of an expected positive reinforcement leads to reduced firing (Tobler, Dickinson, & Schultz, 2003). As expected from a prediction error account, fully expected positive outcomes do not induce increased firing, showing that these neurons code for prediction errors rather than for the sensory aspects of reward. For aversive reinforcements, the picture is less clear. Aversive outcomes generally result in decreased firing rates, but due to the low baseline firing rates of midbrain dopaminergic neurons, it has been argued that the possible range of this reduction is too limited for coding a range of aversive outcomes with any resolution (Dayan & Niv, 2008). Rather, it has been proposed that some other neural system, possibly involving serotonin, contributes to positive aversive prediction errors (i.e., prediction errors where an unexpected aversive outcome results in increased firing rates), although this is presently unclear (Boureau & Dayan, 2011; Dayan & Huys, 2009). Thus, the extensive evidence in support of reward prediction errors in the dopaminergic midbrain do not speak directly to the studies included this thesis, which exclusively focuses on the aversive domain. However, the evidence do show that the RL framework is realistic, which makes a generalization to the aversive domain plausible (Delgado, Li, Schiller, & Phelps, 2008; McNally, Johansen, & Blair, 2011). For example, aversive prediction errors have been shown in the mice amygdala (McHugh et al., 2014) and in the human striatum (Delgado et al., 2008), although the generating sources of these signals are still unknown.

It should be noted that the dopaminergic reward prediction errors do not correspond to those computed by the simple R-W model, but rather the slightly more complicated *temporal difference model* (Glimcher, 2011; Waelti et al., 2001). The temporal difference model of RL was developed to solve problems of how autonomous agents should adapt their behavior to the environment (Sutton & Barto, 1998). The temporal difference model is basically a continuous time- extension of the basic R-W model, where prediction errors and updated expected value estimates are computed on every time-step (rather than once per experimental episode). Importantly, the expected value of performing an action in response to a stimulus or in a particular state depends not only on the outcomes (if any) on that precise time step, but also on the discounted value of all expected future outcomes (which can depend on the chosen action in that state). Temporal difference RL allows a solution to the temporal credit assignment problem, which refer to the fact that if rewards are delayed it is difficult to correctly determine the action or stimulus that predicted the reward. The solution is accomplished by attributing each obtained reward not only to the current time-step but also a few of the preceding moments (Glimcher, 2011). Dopaminergic neurons apparently reflect this process in the propagation of phasic activity from the delivery of an unexpected reward to the stimulus that predicted it over the course of experimental trials (Waelti et al., 2001).

Reinforcement learning prediction errors in the dopaminergic midbrain are also widely believed to underlie neural activity in the human anterior cingulate cortex (ACC) in response

to outcomes, particularly response errors, which are worse than expected (Holroyd & Coles, 2002, but see Ullsperger, Fischer, Nigbur, & Endrass, 2014). This activity, as measured using electroencephalography (EEG) in humans, is known as the *error related negativity* (ERN; Hajcak, 2012). Importantly, the ERN is amplified as a function of the severity of the consequences of errors, such as monetary punishment (Hajcak, Moser, Yeung, & Simons, 2005), social evaluation (Hajcak et al., 2005), and punishment with an aversive noise (Riesel, Weinberg, Endrass, Kathmann, & Hajcak, 2012), in concordance with enlarged aversive outcomes, and thus larger negative prediction errors. Based on such findings, researchers have proposed that the ERN might index "affective" qualities of error monitoring (Hajcak, 2012) in addition to the traditional emphasis on cognitive control (Gehring, Goss, Coles, Meyer, & Donchin, 1993). A recent meta-analysis of fMRI studies showed extensive overlapping activation to cognitive control, negative emotion and pain in the anterior mid cingulate cortex (aMCC) (Shackman et al., 2011). It was postulated that this regions integrates information about negative reinforcers (e.g., pain) arriving from cortical and subcortical afferents (e.g., insula, striatum, amygdala), to bias behavioral selection away from punishment. The bias of behavioral selection would foremost needed in demanding or potentially costly situations, for example, when the consequences of action are uncertain (e.g., probabilistic learning), multiple conflicting response alternatives are active, or when failure of an intended action is associated with potential punishment (Shackman et al., 2011). This hypothesis is of particular relevance for **Study I**.

2 AIMS

The overall aim of this thesis was to investigate how aversive learning shapes adaptive behavior in the social domain. To this end, the following objectives were specified:

- To investigate how threat of punishment affects action selection, and to evaluate corrugator EMG as a novel index of the control processes active under threat of punishment (**Study I**)

- To assess how facial expression function as reinforcers in social avoidance learning, and to understand how individual differences in negative racial bias modulates the effect of social reinforcers (**Study II**)

- To understand how the intrinsic negative value of "preparded" social and non-social stimuli affects avoidance learning (**Study III**)

- To investigate how threat of punishment can generate and propagate avoidance behaviors between individuals through social learning (**Study IV**)

3 OVERVIEW OF THE STUDIES

3.1 STUDY 1

Background and objectives

Cognitive control over action selection has received extensive attention over the last three to four decades, and much is known both about the psychological and neural mechanisms involved (Botvinick, Braver, Barch, Carter, & Cohen, 2001; E. K. Miller & Cohen, 2001). However, this vast body of research has generally disregarded how the *consequences* of failed control affect behavior, which surprising when viewed from a functional perspective: control is most needed when control failure is costly. Such costly consequences might take place both in the non-social (e.g., failing to override a habitual tendency to look to the right when crossing the street in Britain) and the social (e.g., failing to control verbal behavior in a hostile interaction) domains. In this study, we investigated how cognitive control over action selection was influenced by threat of punishment and response conflict in a simple experimental model designed to capture situations where failure to respond correctly is costly. As described in the introduction, the ACC integrates signals related to danger and control demands, with the hypothetical function of promoting adaptive behavior (the Adaptive Control Theory, Shackman et al., 2011). We investigated if activity in the corrugator supercilii muscle of the upper face would potentially index the computations of the ACC. The rational for this hypothesis was the known electro-physiological correlate of ACC activity following response errors (i.e., control failure) and response conflict, the ERN, which also is responsive to aversive stimuli. The ACC enervates the corrugator supercilii muscle (Morecraft, Louie, Herrick, & Stilwell-Morecraft, 2001; Shackman et al., 2011), which motivated us to look for functional similarities between activity in this muscle and the ERN. We hoped that activity in this muscle could function as an index of the implementational level mechanisms involved in action selection during threat of punishment, and thus shed light on the underlying process. We predicted that (i) behavioral performance should be enhanced by punishment threat, and (ii) cEMG activity should be sensitive to the combination of response conflict, punishment threat, and error commissions.

Methods

Forty participants took part in the experiment. Participants were randomly assigned to the Punishment Risk group (n = 22, 11 male), and the Control group (n = 18, 10 male). Four participants were excluded from the Control group, as they faultily reported a relationship between their performance and the number of received electric shocks (see below for details). The participants performed a two-choice speeded gender decision task (is the individual displayed male or female), and responded behaviorally with a keyboard. The probability of each target gender was asymmetric (75% male faces/25% female faces, or the reverse), giving a correspondingly asymmetric response ratio. This manipulation was based on the Go/No-Go task (Simmonds, Pekar, & Mostofsky, 2008), where the high probability target induces a pre-potent tendency to respond, which has to be inhibited for low probability

targets. However, the standard Go/No-Go task has apparent limitations due to the lack of a recorded response for the critical low probability (no-go) condition. For this reason, responses to both targets were collected. It should be noted that evidence from fMRI (Kenner et al., 2010; Mostofsky & Simmonds, 2008) and computational modeling (Gomez, Ratcliff, & Perea, 2007) indicate that response inhibition and response selection is highly related, or even overlapping, processes, suggesting that the task used in the present study and the standard Go/No-Go task are comparable. The 75% target condition is referred to as Low Conflict and the 25% target condition is referred to as High Conflict. The Low Conflict and High Conflict gender was counterbalanced across participants.

Punishment threat (referred to as Punishment risk in the article, and below) was manipulated block-wise, with three levels; No Risk, Low Risk, and High Risk. For Low- and High Risk blocks, participants were instructed that any errors (both commissions and omissions) during the block might be punished with a mild electric shock *after* the block. Critically, the actual number of delivered shocks was identical for all participants. The intention was to manipulate punishment risk without introducing performance-contingent variability in the amount of experienced shocks across participants. Participants received 0-2 electric shocks after Low Risk blocks, and 2-4 electric shocks after High Risk blocks. Thus, no actual relationship between individual performance and number of shocks existed. Importantly, funneled interviews after the experiment showed that all 22 participants in the Punishment Risk group believed that there was a direct relation between their own performance and the number of shocks they received.

 As noted above, 14 participants were randomly assigned to the Control group. The Control group performed the same experimental task and received the same number of electric shocks as the Punishment Risk group, but was explicitly informed that there was *no* contingency between performance and punishment. The purpose of this manipulation was to delimitate the effect of punishment threat on corrugator activity from general shock anxiety. Four participants in the control group faultily reported a relationship between their performance and the number of shocks they received, and were therefore excluded from the analyses.

Both groups completed 12 blocks (4 blocks per Punishment Risk level) of 80 trials (total of 960 trials). Both block order (Punishment Risk) and trial order (Response Conflict) were fully randomized for each participant. A colored frame surrounding the target stimulus indicated Punishment Risk level (control group: ordinal amount of expected shocks) during the blocks (No Risk = green, Low Risk = yellow, High Risk = red).

Electromyographic (cEMG) activity of the left corrugator supercilii muscle was recorded using a BioPac (MP100; Biopac Systems Inc, www.biopac.com) device equipped with two miniature Ag/AgCl electrodes filled with electrolyte gel (Fridlund & Cacioppo, 1986). A third ground electrode was placed on the mid forehead, proximal to the hairline. The raw EMG signal (sample rate 1000 Hz) was amplified and filtered through a 28–500 Hz IIR band pass, followed by a 50 Hz IIR band stop. The signal was rectified and integrated with a time constant of 20 ms.

The cEMG data was extracted in 100 ms time bins locked to the behavioral response on each trial using in-house software. Time-bins were extracted both prior (Pre) and following (Post) the behavioral response. The number of Pre-response relative to Post-response time-bins extracted on each trial was dependent on RT (e.g., for a trial with relatively long RT, more Pre-response time-bins and fewer Post-response time-bins were extracted, compared to a short RT trial). A 100 ms pre-stimulus baseline (mean cEMG amplitude) was subtracted from all time-bins to reduce slow signal drift and tonic level differences. Baseline measures below or exceeding 3 standard deviations were replaced and interpolated as an un-weighted average from the six adjacent baseline means. The peak cEMG amplitude was computed for each time-bin.

Results & Conclusion

First, as predicted from a functional perspective, the results showed that the perceived risk of punishment for error commissions attenuated the detrimental effect of high response conflict on performance. This effect was however non-linear: Low Risk enhanced performance, while High Risk had little effect (Figure 1). Second, the cEMG was highly sensitive to both punishment risk and response conflict, showing the predicted properties of a signal that integrates cognitive control demands and avoidance motivation.
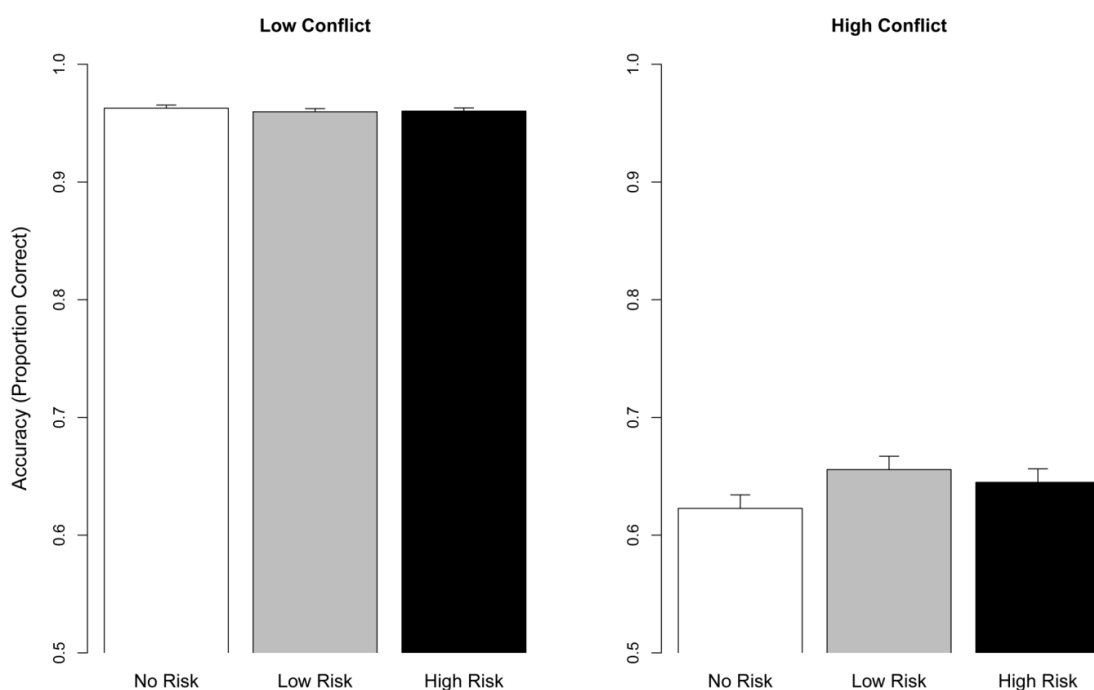

Figure 1. Study 1. Performance accuracy as a function of punishment risk and conflict.

In particular, the study provided the first report of cEMG activity as a correlate of error monitoring (Figure 2), and showed that this correlate (i) operates on a similar time-scale as the ACiC generated index of error processing, the ERN, (ii) is modulated by the expected

aversive consequences of errors, and (iii) related to behavioral adjustments following errors (post-error slowing).
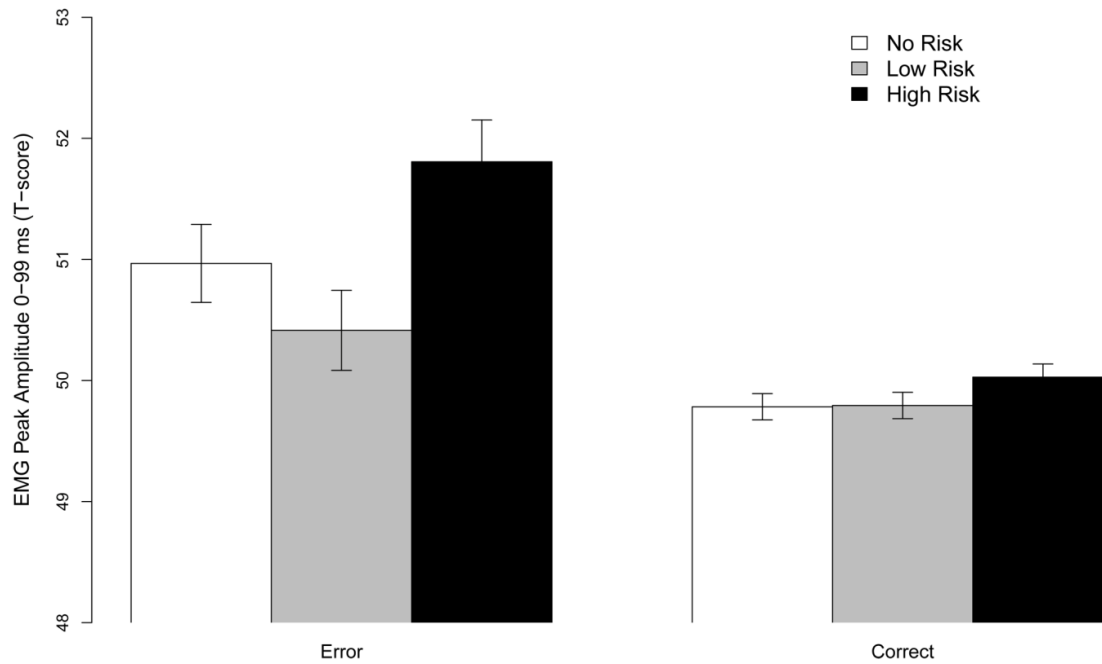
Figure 2. Study I. Corrugator EMG activity within 100 milliseconds following Correct and Error responses for the different levels of punishment risk.

These finding stresses the importance of considering not only the goals of behavior, but also the potentially aversive or costly consequences of failure, to fully characterize cognitive control over action selection. We also for the first time demonstrated that cEMG is sensitive to the interaction of cognitive control and punishment avoidance. This finding is concordant with the proposed integrative role of the ACC in potentially dangerous and control demanding situations (Shackman et al., 2011), and suggest that further investigation of the response properties of the cEMG might be highly fruitful for understanding the implementational processes underlying action selection, reinforcement learning and cognitive control in the aversive domain.

3.2 STUDY 2

Background and objectives

Reinforcement learning is likely to play an important role in adaptive social behavior, by allowing one´s behaviors toward others to be based on the emotional significance of their previous actions toward oneself (Ruff & Fehr, 2014). Facial expressions is a class of reinforcers that might play an important role in such social RL by signaling benign and malign behavioral intentions (Blair, 2003; Heerey, 2013). The face also conveys information about an individual's racial and ethnic group belonging, and such markers of social group membership might affect learning from social reinforcement. In support of this conjecture,

previous research has shown superior memory for angry, relative to neutral or happy, racial out-group faces (Ackerman et al., 2006), and more persistent fear conditioning to racial out-group versus in-group faces (Navarrete et al., 2009; Olsson, Ebert, Banaji, & Phelps, 2005). It is however unknown whether the group-membership (race) of a face can modulate the value of its emotional expression, and thus shape behavior during social interaction. Related to this, little is known about the possible mechanisms underlying the impact if social reinforcement on behavior. To address these issues, we examined how racial bias shaped avoidance behavior reinforced by threatening (angry) and friendly (happy) facial expressions posed by racial in- or out-group individuals, which is analogous to how our behavior is shaped as a function of the facial feedback we receive from others in response to our actions (Blair, 2003). We used a novel experimental model of a dynamic social situation in which facial race and emotion reinforce behavior to elucidate (1) how individual differences in racial bias modulate social RL from facial expressions posed by racial in- or out-group members, and (2) the underlying computational mechanisms, and how these are shaped by racial bias, by computational modeling. First, we reasoned that because angry out-group faces (Black for the White experimental group) were potentially perceived as more threatening than angry in-group and happy out-group faces, learning to avoid them should be easier. Importantly, we predicted that this learning bias would be directly related to individual differences in racial bias. Second, we aimed to understand not only *if*, but also *how*, social reinforcers (here emotional expressions) and racial belonging together would affect RL by elucidating the underlying computations. As described in the introduction, value-based decision making can be parsed into a number of separate sub-processes, where two of the most important are *outcome evaluation* (OE; what was the positive or negative value of the outcome?) and *outcome learning* (OL; how fast, i.e., at what rate, does an unexpected outcome affect subsequent behavior?) (Rangel et al., 2008). We sought to determine which of these processes were most sensitive to the emotional expression and racial belonging of the reinforcing faces.


Methods

Thirty participants of European decent (20 women) performed a probabilistic two-choice learning task. They were instructed (by a 1 s text message prior to each of the 16 blocks of the experiment) to choose the one out of two abstract fractal stimuli that lead to least exposure ("avoid") of the emotional face (angry or happy depending on block). The rational for contrasting angry and happy faces was to assess the specificity of the predicted interaction of racial bias, emotion and race. Each fractal differed in terms of the probability that it was followed by an emotional face ($P = .3$, the instrumentally optimal choice, or $P = 0.7$, the instrumentally sub-optimal choice). The presentation of the neutral or emotional face, subsequent to the participant's choice, served as the only feedback on the choice made. By repeatedly choosing among these two fractals, the participants learned across trials which choice was optimal in each block. No instructions regarding the racial group belonging of the

stimuli were provided. The stimuli gender was randomized for each block of the 12 experimental blocks (four block for each experimental condition), each consisting of 30 trials. Because the racial group of the facial stimuli was irrelevant to the instructed goal of the task, this design allowed us to probe the indirect effects of emotion and race in interaction on instrumental learning. Following the experiment, the participants completed a validated Swedish version of the Modern Racism Scale to assess explicit negative attitudes toward non-European immigrants in Sweden (Akrami, Ekehammar, & Araya, 2000), and the Race Implicit Association Test (Greenwald & Banaji, 1995) to measure implicit racial bias.

Results & Conclusion

Individual differences in explicit racial bias strongly, and uniquely, modulated the impact of out-group faces on behavior, so that the higher the explicit racial bias, the better learned avoidance of threatening out-group individuals. Previous findings have shown more persistent fear conditioning to racial out-group, relative to racial in-group faces (Navarrete et al., 2009; Olsson et al., 2005), and to angry, compared to neutral, in-group faces (Öhman & Mineka, 2001). The present results provide an important extension of previous research by showing that reinforcing facial expressions and markers of racial group-belonging interact in their impact on instrumental learning as a function of individual differences in racial bias (see Figure 3).
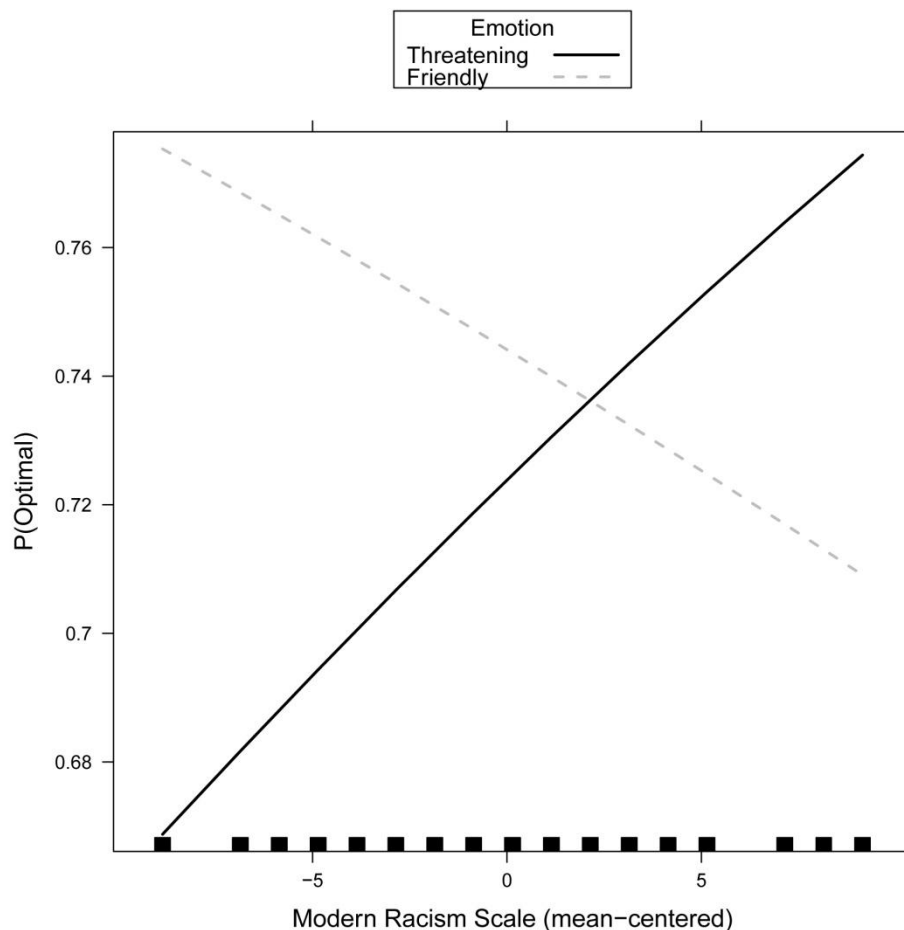
Figure 3. Study II. The probability of optimal avoidance when out-group faces served as reinforcers as a function of their emotional expressions and individual differences in racial bias.

By applying a RL model-based analysis, we distinguished between two competing hypotheses about the computational processes underlying social avoidance learning; OE and OL (Lin, Adolphs, & Rangel, 2011; Rangel et al., 2008). The model comparison provided strong support for the OL-hypothesis. We found that social reinforcement and racial group-belonging together affected how rapidly unexpected outcomes were transformed into future behaviors. Importantly, individual differences in both explicit and implicit racial bias were positively and selectively correlated with the learning rate associated with threatening racial out-group faces. These correlations indicate that individuals with higher, as compared to lower, racial bias more rapidly update the value of each action to adjust their subsequent behavior when exposed to threatening racial out-group faces. This rapid value updating was functional; individuals with high explicit racial bias exhibited enhanced performance when their actions were reinforced by threatening out-group faces. Taken together, these results suggest a computational mechanism for how racial bias affects social reinforcement, dependent on the racial belonging of the reinforcing individual. These findings have implications for understanding the role of learning in social interaction, and how this learning is modulated by the racial group belonging of the people that interact.

3.3 STUDY 3

Background and objectives

Humans and non-human primates preferentially acquire conditioned fear to certain archetypical, "prepared" threatening stimuli (referred to as *Pavlovian triggers* in **Study III**), several of which are highly important for the social domain (i.e., threatening faces and out-group faces). Yet, how these learning biases influence voluntary behavior, their basic mechanistic underpinnings, and how they interact with learning experiences during the life span of the individual remain unknown. In **Study III**, we investigated how four classes of such stimuli that previously have been shown to be preferentially associated with fear (snakes, threatening in-group faces, racial out-group faces, and guns) influenced voluntary behavior. We investigated the behavioral consequences of these stimuli on behavior using a novel experimental model and computational analyses. Furthermore, we sought to characterize the computational mechanisms through which stimulus driven biases interact with learning experiences using RL modeling. We also aimed to clarify how the negative value associated with the Pavlovian triggers was acquired by conducting two independent ratings studies. These rating studies probed the degree of negative individual and social exposure the participants had with the different Pavlovian triggers.

Methods

Experiment 1 (snake vs. flower stimuli) included 32 (21 female) participants, Experiment 2 (angry face vs. happy face stimuli) included 42 (22 female) participants, Experiment 3 (out-group face vs. in-group face stimuli) included 48 (34 female) participants, and Experiment 4 (gun vs. flower stimuli) included 34 (18 female) participants. No participant took part in more than one experiment. The first ratings sample involved 59 (39 female) participants, and the second ratings sample 35 (26 female) participants. All four experiments had identical experimental design, and only differed in the stimuli selection. The experimental task was a two-choice decision task with probabilistic aversive reinforcement (electric shocks). The participants were instructed to try to learn to avoid shocks, and that one choice stimulus might be better than the other for doing so. The task had a mixed two-by-two design, with one between-participants factor (Pavlovian Compatibility: Compatible To Incompatible/Incompatible to Compatible) and one within-participants factor (Phase: First/Second). After half the experiment (35 trials) the choice – reinforcement contingency reversed. For the Compatible To Incompatible - groups, choosing the Pavlovian trigger was the sub-optimal instrumental action (P(Shock) = .7) in the First half and optimal (P(Shock) = .3) in the Second half, while the reverse was true for the Incompatible to Compatible - groups. This design allowed us to control for the order effect of presentation, and disentangle transient and sustained effects of Pavlovian triggers on instrumental learning. The

participants were not informed about the contingency reversal. The participants performed 6 practice trials (no reinforcement), followed by the experimental session (70 trials). Participants used the left and right arrow keys on a computer keyboard to indicate their choices. The location of the choice stimuli varied randomly between the left and right position across trials to prevent spatial selection strategies.

We conducted two rating studies with independent samples of participants, who rated the stimuli used in Exp. 1-4. In the first rating study, the dimensions were "Threat" and "Danger", rated on a 7 – graded scale, and anchored with "Not at all"(1) and "Very"(7) . The participants also reported how many individual negative experiences they had experienced in relation to what the images depicted, using a three-graded scale consisting of "None", "Few", and "Many". In the second ratings study, we asked the participants to estimate how much negatively valenced exposure they had experienced to the different Pavlovian triggers through (i) individual experience, (ii) family and friends, (iii) news media, and (iv) popular culture.

Results & Conclusion

Collectively, our findings showed that the intrinsic value of stimuli predicting danger, Pavlovian triggers, has a strong influence on voluntary behavior. More specifically, the influence of Pavlovian triggers took the form of a stimulus-driven bias that either enhanced or corrupted voluntary behavior depending on the compatibility with the optimal instrumental action. This pattern was consistent across all four types of Pavlovian triggers; snakes, guns, threatening faces and out-group faces, showing that both archetypically phylogenetic (e.g., snakes) and ontogentic (e.g., guns) stimuli can affect voluntary behavior (Figure 4). For out-group faces (Exp. 3) the pattern was qualified by the fact that only participants with above average explicit racial bias exhibited this stimulus-driven bias.
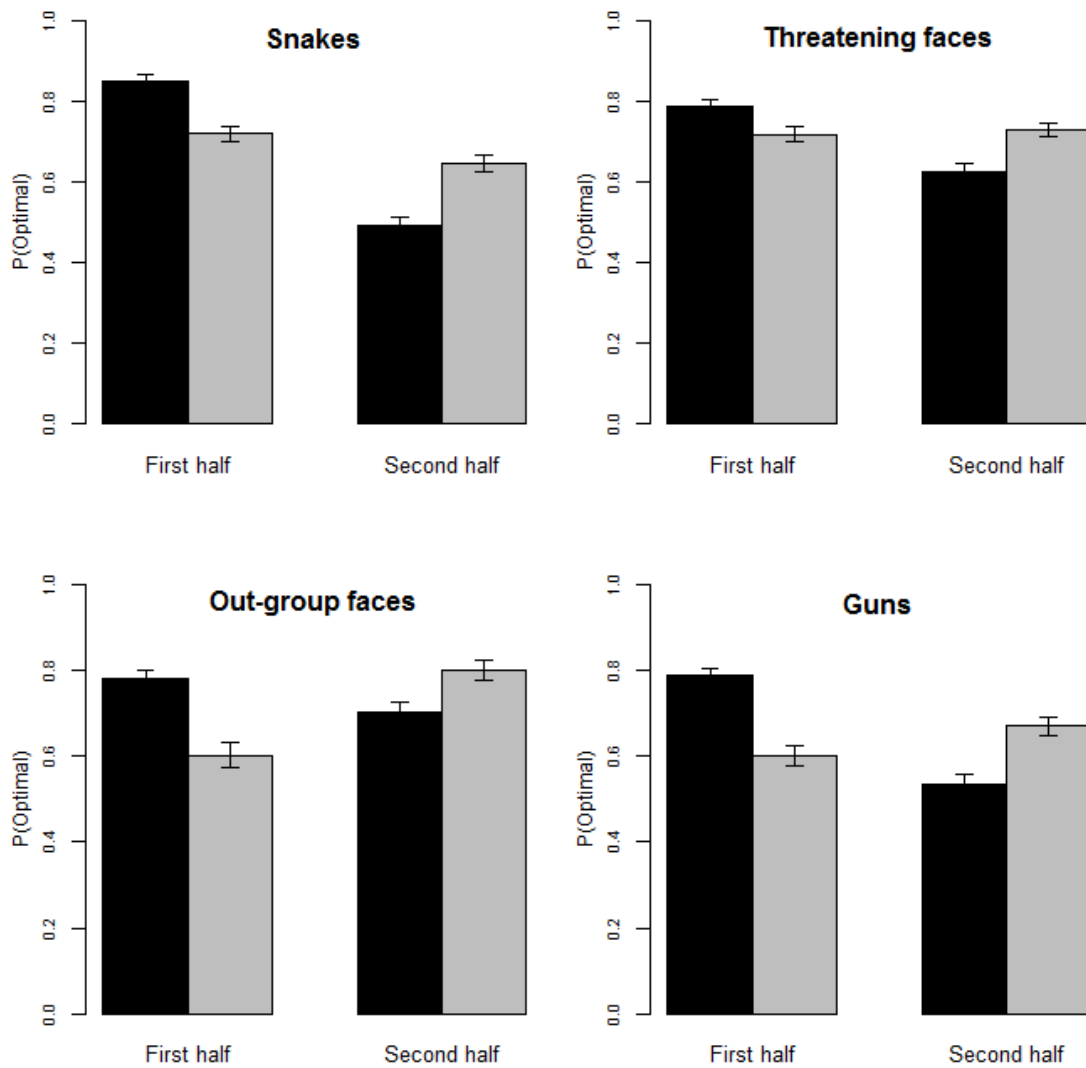
Figure 4. Study III. Probability of optimal behavior for the four classes of Pavlovian triggers: (**A**) Snakes (interaction: $p < .01$), (**B**) Threatening faces (interaction: $p < .01$), and (**C**) Out-group faces for participants *high* in racial bias (interaction: $p < .05$), (**D**) Guns (interaction: $p < .01$). Black bars = Compatible to Incompatible, grey bars = Incompatible to Compatible. Error bars denote SEM.

Our analyses of the computational basis of the Pavlovian influence on instrumental behavior provide support for a simple proposal: the intrinsic value of Pavlovian triggers primarily affect the action selection, rather than the updating or learning, process. In the independent systems - model, which provided the most parsimonious account of the data, the Pavlovian value down-weighted the instrumentally learned value of the action represented by the Pavlovian trigger at the time of choice, resulting in avoidance. The Pavlovian valuation process appears to be learning-independent (within the time-scale of the experiment) and automatically triggered, as shown by the shift in the direction of the Pavlovian bias after the contingency reversal in our experiments (see Figure 4). The fact that guns were associated

with a Pavlovian value, as estimated in the RL model, comparable to snakes and higher than both out-group and threatening faces clearly demonstrate that evolutionary predispositions are not necessary for the results of **Study III**. Instead, the pattern of estimated Pavlovian value of the stimuli types matched their perceived dangerousness, as rated by an independent sample. Perceived dangerousness or deadliness has been suggested as the factor underlying both phylogenetic and ontogenetic fear-relevant stimuli, thought to result in fear learning biases through an enhanced expectancy for associated negative events (Davey, 1992, 1995). A classical account of the etiology of human fears suggested that multiple pathways jointly contribute to the acquisition of fears (Rachman, 1977). These pathways are direct experience, observation (or vicarious) and verbal transmission of threat information. The effectiveness of the last two pathways, about which the preparedness literature generally has been silent, has been repentantly demonstrated in both children and adults (Askew & Field, 2008; Olsson & Phelps, 2007; Phelps et al., 2001). Our second ratings study, where we asked participants to estimate the amount of negatively valenced exposure from different sources to the four types of Pavlovian triggers, provide qualitative support for the idea that the Pavlovian value of all types except threatening faces is, at least partially, acquired socially, through observational and/or verbal pathways.

3.4 STUDY 4

Background and objectives

The results of **Study III** suggested that socially acquired negative valuations can powerfully impact individual avoidance learning and decision making. It´s long been recognized in behavioral ecology that animals are prone to influence the behavior of each other, especially if the environment is dangerous (Rendell et al., 2011; Webster & Laland, 2008). This is an adaptive strategy: social learning allows avoiding the costs associated with individual learning when the environment is dangerous, for example through threat of predation (Griffin, 2004). A limited research literature has shown that humans can acquire conditioned fear responses through observing how another individual (the Demonstrator) is exposed to aversive reinforcement (Olsson & Phelps, 2007). However, very little is known about how actual (instrumental) avoidance behaviors are transmitted socially in humans. In **Study IV**, we aimed to extend the understanding of human social learning about aversive outcomes by directly exploring how social information can influence avoidance behaviors. As outlined in the introduction, many uniquely human behavioral patterns, such as social norms and taboos, are maintained through the threat of punishment. Collectively, such large scale behavioral patterns involving threat of punishment can be described as "avoidance traditions". In humans, many avoidance traditions are likely to be adaptive, such as cooperative norms maintained by threat of punishment (Fehr & Fischbacher, 2004) and taboos prohibiting dangerous food (Henrich & Henrich, 2010). Other avoidance traditions are arbitrary or even destructive (Aunger, 1994; Mani, 1987), such as when behaviors that were adaptive in the past has become maladaptive in modern societies (Boyd et al., 2011). We addressed a little

studied, but crucial aspect for understanding how punishment shapes avoidance traditions: the motivational influence of punishment on the *punishee*, and how this influence can support the emergence, maintenance, and transmission of avoidance traditions. Despite the large theoretical and empirical literature investigating the influence of punishment on cooperation and other behavioral traditions, surprisingly little is known about the psychological and computational mechanisms of learning and decision-making underpinning this influence. We hypothesized that the rewarding effect of avoiding a possible punishment (see Table 1) together with a tendency to copy the behavior of others can provide building blocks to help explain the emergence, maintenance and transmission of avoidance traditions. Their joint impact on behavior would thereby present an important complement to established game theoretic analyses of how rare punishment affects large scale behavior patterns. To address our hypothesis, we first developed an agent-based simulation model that mimicked the core features of avoidance traditions that we discuss above: (i) one´s actions risk being punished, (ii) others' actions can be observed, and (iii) there is an uncertainty about the optimal course of action. Agent-based simulations is a computational modeling framework commonly used within biology and social sciences such as economy, but which yet has to receive much attention within psychology (Smith & Conrey, 2007). The premise of agent-based simulations is that complex systems can be understood and explained as emergent phenomena generated by the interaction of many simple individual agents. We used this modeling technique to investigate if the combination of rewarding punishment omission and behavior copying can in principle generate, maintain and transmit inter-generational avoidance traditions. We then confirmed the assumptions underlying this model in four separate experiments that each addressed specific aspects of the agent-based simulation model.

Methods

The agent-based simulation model was intended to capture the essential structure common to avoidance traditions in both social and non-social domains. Individual agents in a population (N = 100 in Figure 5) made repeated decisions between two possible actions [*A, B*] under threat of punishment. None of the actions were associated with external reward. Instead, we varied the probability that action *B* was punished to explore the role of rare punishment in avoidance traditions. In contrast to standard game theoretic models of punishment, there was no strategic interaction between the individuals. Instead, each agent first (Observation phase) observed *M* number of choices by another, randomly selected, agent (the "Demonstrator" agent), and subsequently chose individually *M* times (Choice phase). To generate a simple "birth-death" process, the agents were probabilistically removed (average lifespan = 2*M*) and replaced with completely naïve individuals. The model did not include any heritable traits. Each individual was based on a simple RL model (see Equations 4-6), but with an additional learning rate parameter that regulated the influence of observed actions on the individuals value estimates. The experiments mimicked the structure of the agent-based model, and had a computerized setup involving repeated choices between the two options [*A, B*]. A total of 120 human volunteers participated after providing written informed consent. Prior to the

experimental procedure, each participant was attached to the shock electrodes (all experiments except Exp. 2) and the shock amplitude was determined individually to be "uncomfortable but not painful". The participants in all experiments, except Exp. 2, received identical instructions. They were informed to avoid electric shocks by choosing between two images, that they first would observe the choices, but not their consequences, of a previous participant (Observation phase), and that they thereafter would do the same task (i.e., choose between the same images) themselves during which they might receive shocks based on their choices (Choice phase). The purpose of this instruction was to induce an experience of threatening punishment. In Exp. 2, the subjects received the same information as in the other experiments, but were told that their choices could be awarded by points. These points would increase their chance of winning a lottery for extra movie vouchers. They were instructed that delivery of points would be signaled with a "$" - symbol. The "$" – symbol was never shown. To provide unambiguous and clearly defined social information, we used a computer-controlled Demonstrator that consistently chose option *A* in Exp. 1-3. In Exp. 4, we relaxed this premise, and allowed the participants to observe the choices of real former experimental participants. In Exp. 1 (n = 25), none of the options could actually be punished. In Exp. 2 (n = 20), the participants were instructed that they could be rewarded, but in reality, none of the actions led to reward. In Exp.3 (n = 20), the participants were randomly punished with electric shocks, and finally, Exp. 4 (n = 50) had the same setup as Exp. 1, but the participants observed the choices of actual prior participants in a transmission chain design (Mesoudi & Whiten, 2008).

Results & Conclusion

Collectively, the results from our agent-based simulation model and the four empirical experiments show that omission of threatening punishment (referred to as rewarding punishment omission: RPO) in combination with social learning (referred to as behavior copying: BC) can underlie the emergence, maintenance, and transmission of both adaptive (when actual punishment is possible) and arbitrary (when no actual punishment is possible) avoidance traditions (see Figure 5).
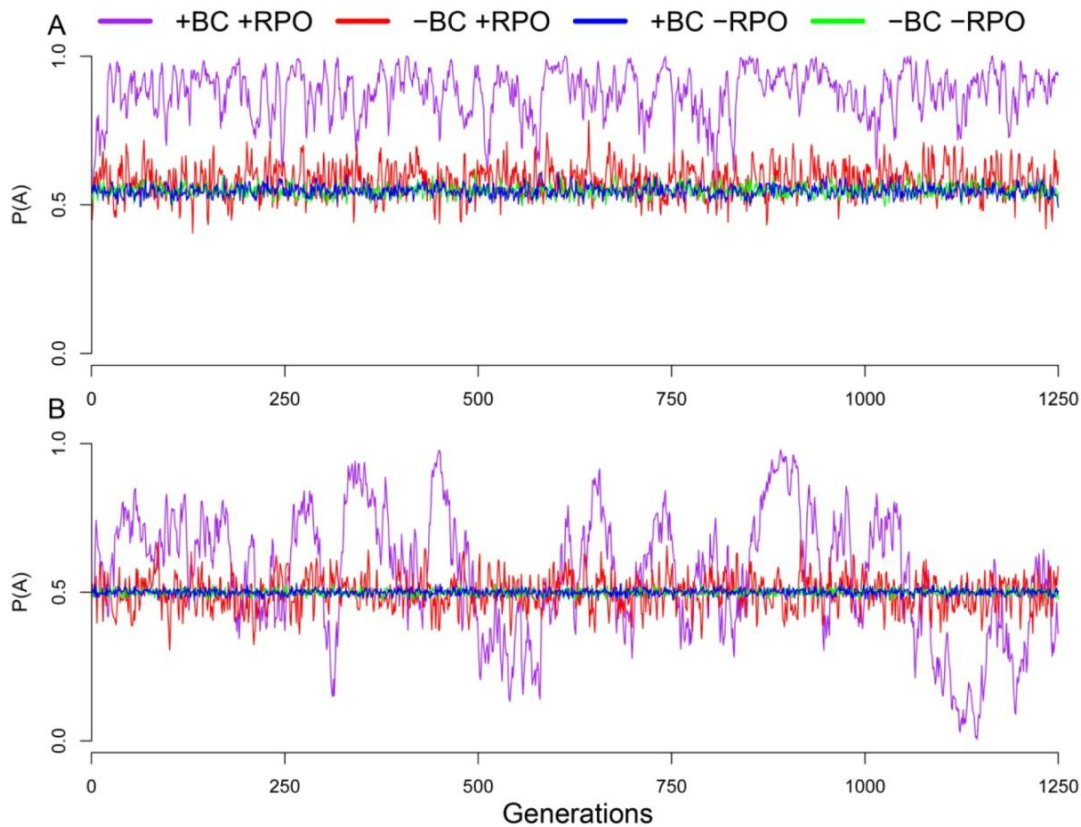
Figure 5. Study IV. Two typical simulation runs, with (A) and without (B) punishment. In (A), behavior B had a 0.5 % risk of being punished. BC = Behavior Copying, RPO = Rewarding Punishment Omission.+ indicates the presence of the learning mechanism and - absence of the mechanism. $M = 20$, $N = 100$.

The four empirical experiments showed that omission of threatening punishment (Exp. 1 & Exp. 4), but not reward (Exp. 2) leads to faithful behavioral imitation. Even if actions were randomly punished (Exp. 3), the action of the Demonstrator was imitated. Finally, in Exp. 4, we established that the rewarding omission of punishment together with social learning can transmit avoidance behaviors in an inter-generational fashion as assumed by the agent-based model.
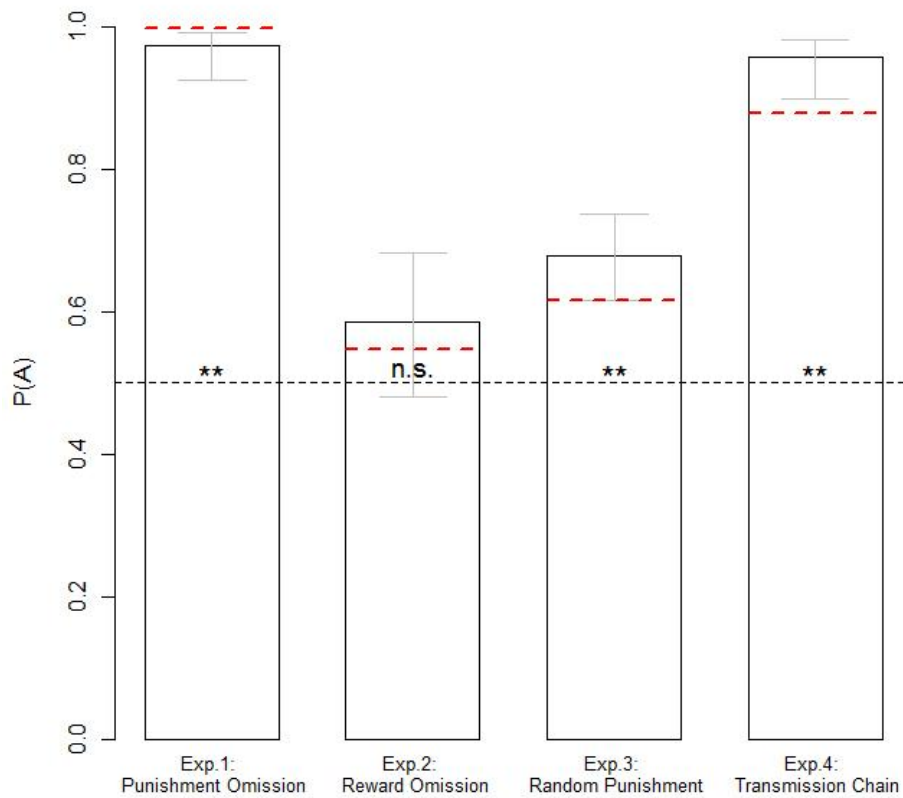
Figure 6. Study IV. The probability of copying the Demonstrator choice (A) in Experiments 1-4. ** = $p < .01$ compared to 0.5. The red lines show the simulated *P(A)* from the RL model with parameter values estimated from the empirical data. Error bars represent 1 SE.

These findings offer an important complement to traditional game theoretic analyses of punishment in strategic interaction by providing a mechanistic, motivational basis for how individuals are affected by the threat of (rare) punishment in systems where the actions of others can be observed. The proposed mechanisms are consistent with current knowledge about the computational and neurobiological basis of learning and value-based decision making (Rangel et al., 2008; Seymour et al., 2007), and thereby provides a plausible explanation for how exceedingly rare punishment can motivate behavior and how this behavior can spread in groups.

4 GENERAL DISCUSSION

The overall aim of the studies presented in this thesis was to illuminate the psychological and computational mechanisms underpinning adaptive behavior based on aversive leaning in (simple experimental models of) the social domain. The role of avoidance and punishment in the social domain is likely to be of paramount importance for understanding the behavior of

both human and non-human animals. Our general approach was to develop simple experimental model paradigms that captured the essential underlying structure of (aspects) of the social domain in order to characterize the psychological and computational mechanisms involved. Here below, I will begin with discussing some aspects of the results of each study, and thereafter touch upon some more general theoretical and methodological questions and conclusions suggested by these studies.

In **Study I**, we investigated how control over action selection in a simple social situation, involving decision about the gender of novel human faces, was affected by threat of punishment. This question is of importance in both social and nonsocial settings, because previous research on cognitive control has generally neglected the consequences of deficient control. In line with the functional hypothesis that humans are adapted to avoid costly errors, we found that behavioral performance was enhanced during threat of punishment. However, we found that this effect was curvilinear (See Figure 1). Performance was enhanced during Low threat of punishment, but not during High threat of punishment, relative to the No punishment baseline. We speculate that this pattern correspond to the classical inverted-U shape function between arousal and performance (Teigen, 1994). Although surprising, this finding is an intriguing demonstration that behavior is not necessarily maximally adaptive, even when it might be associated with potential fitness costs in natural settings. This pattern might be illuminated by the recent view in the field of behavioral ecology that behavior is not likely to be adaptive in *all* circumstances but rather constrained by evolved mechanisms shared across domains (Fawcett et al., 2014; McNamara & Houston, 2009) so that it adequate on average (i.e., satisficing rather than maximizing: Gigerenzer & Goldstein, 1996). However, strong conclusions about this pattern of results should await replication studies involving a fully parametric manipulation of the cost associated with punishment. Based on the proposal that the ACC integrates signals of danger and control demands in order to coordinate adaptive behavior (Shackman et al., 2011), and because this brain region enervates the muscles of the upper face (Morecraft et al., 2001), we predicted that activity in the corrugator supercilii (cEMG) muscle might index the underlying process. In support of this prediction, we found that cEMG activity was amplified after response errors, particularly during high threat of punishment, and that this activity predicted slower response times on the subsequent trial. Moreover, cEMG activity prior to the response predicted performance accuracy, in line with influential theories suggesting that the ACC resolves response conflict. As outlined in the introduction, one influential theory of ACC function in monitoring of action selection states that neural post-error activity in this region (i.e., the ERN) reflects a phasic decrease in the firing of mid-brain dopaminerigic following worse than expected outcomes or synonymously, a negative prediction error. Our cEMG data fits with this account, as the responses to more costly errors during the High threat of punishment condition were amplified. However, we did not directly test if cEMG activity correlates with prediction errors derived from a RL model, which make this reasoning speculative. Replication of our findings using a task suited for RL modeling or with the axiomatic approach to the necessary and sufficient criteria for a prediction error signal (Rutledge, Dean,

Caplin, & Glimcher, 2010) is thus needed. Such a replication would also tie the findings of **Study I** more tightly to the general value-based decision making framework outlined in the introduction. Regardless, **Study I** provides preliminary evidence for adaptive arbitration between habitual and controlled responding during threat of punishment.

In **Study II** and **III** we addressed a core process in value-based decision making by asking how the differential value of different social stimuli affects learnt avoidance behaviors. Furthermore, both studies shared a common methodological approach, by using two choice probabilistic decision making tasks and computational RL modeling as analysis strategy. However, the two studies approach the topic from different directions. In **Study II**, we sought to understand how emotional (angry and happy) facial expression could function as reinforcers of avoidance learning. Others facial expressions in response to one's own behavior have previously been suggested to play an important role in shaping social interaction (Blair, 2003), but research in support of this proposal have remained relatively scarce (Heerey, 2013; Lin et al., 2011; Shore & Heerey, 2011). In particular, we investigated how preexisting negative attitude to social out-groups (e.g., defined by racial belonging) shape such learning. In summary, **Study II** showed that negative racial bias specifically enhanced avoidance learning when one´s actions were reinforced by threatening out-group, but not in-group faces. Thus, individual differences in racial bias powerfully modified basic emotional learning from the facial expressions of others. Seeking to understand what value-based decision making process racial bias primarily affected, we formulated a series of RL models based on the R-W algorithm. These models implemented two competing hypothesis: (i) that differences between the experimental conditions primarily could be explained by differences in *learning* from different types of social reinforcers, or (ii) by differences in *outcome evaluation* of the different social reinforcers (akin to differences in reinforcer magnitude). The results supported the first hypothesis. Thus, the different types of social reinforcement were similarly valued, but differed in the rate with which they were transformed into future avoidance behaviors. Individual differences in racial bias were specifically related to the estimated learning rate associated with threatening out-group faces, suggesting a computational mechanism for how individual differences in racial bias affects learning from social reinforcement delivered by racial out-group individuals. Although not conceptualized as such in the published article (Lindström, Selbing, Molapour, & Olsson, 2014), the value of the learning rate parameters of the R-W model are typically interpreted as determined by salience of the CS, or similarly, the attention devoted to processing the CS (Pearce & Bouton, 2001). Thus, our results suggest that racial bias, which we understand as a type of negative valuation, affects attentional processing of the CS which appear to be in-line with research on out-group face recognition (Levin, 2000).

A related question concern how these effects of racial bias should be understood within the general learning theory framework. Regardless of if facial expressions best are conceptualized as primary or secondary reinforcers (Blair, 2003; Shore & Heerey, 2011), a simple explanation is that individual differences in racial bias reflect individual differences in negative experiences with out-group individuals, either by individual or socially-mediated

conditioning (see below for discussion). By this account, in- and out-group faces should differ in reinforcer magnitude (if secondary reinforcers) or the degree of change in associative value due to preconditioning (if primary reinforcers). This was however not what we found, as the models based on differences in learning rates (or salience) best explained behavior. Thus, the simplest R-W account appears insufficient for accounting for pre-existing individual differences. However, it is unclear what alternative account would fare better.

Instead of investigating social stimuli as *reinforcers* of avoidance learning, in **Study III** we asked how the intrinsic (i.e., pre-experimental) value of social (and non-social) stimuli affect avoidance learning when *predicting* an aversive reinforcer. The impetus for this question was the famous findings in classical conditioning showing enhanced fear acquisition and retarded fear extinction to "evolutionary prepared" stimuli such as snakes, angry faces (Ohman & Mineka, 2001), and out-group faces (Navarrete et al., 2009; Olsson et al., 2005). Despite the possible relevance for a range of real world behaviors, especially in the social domain, the influence of such stimuli on voluntary instrumental behavior has been unknown. We hypothesized that such stimuli (termed Pavlovian triggers in the manuscript) might trigger the Pavlovian valuation system during voluntary behavior and thus induce systematic deviations from the predicted instrumental behavior. Put differently, if the Pavlovian and the instrumental systems were put in conflict, the intrinsic value of the stimuli should be revealed. To asses if a putative evolutionary history would be necessary for these effects to emerge, we also included gun stimuli in one experiment. Briefly summarized, we found a strong and consistent effect of Pavlovian triggers on instrumental behavior. This influence was best accounted for by a RL model where the Pavlovian influence took the form of a fixed bias parameter within the Softmax function. Thus, we found that the Pavlovian influence biased the instrumentally learned value at the time of choice, resulting in either adaptive or maladaptive behavior depending on if the Pavlovian trigger best predicted the aversive reinforcer (Figure 4). This conclusion is in concordance with previous modeling-based analyses as well as the known neuro-anatomical correlates of the interaction between the Pavlovian and instrumental valuation systems (Guitart-Masip, Duzel, Dolan, & Dayan, 2014). The RL modeling approach allowed us to quantify the Pavlovian value of the different stimuli, which showed that both archetypically prepared stimulus, snakes, and the obviously non-prepared gun stimulus hade the same Pavlovian value. This finding, together with the rating studies we conducted, which showed a strong similarity in the Pavlovian influence and average threat ratings of each stimulus but not with the history of individual negative experience, suggest that social and cultural influences on the Pavlovian valuation system might be profound (Davey, 1995). An important feature of these data that testify to the importance of learning for Pavlovian triggers was the fact that negative racial bias was positively correlated with the estimated Pavlovian value of out-group faces (there was no comparable correlation for snakes).

The role of social learning in the acquisition of Pavlovian triggers thus provide an important avenue for future research, underscored by the fundamental impact of culture on human behavior (Boyd et al., 2011). Even an influential account of evolutionary preparedness theory

(Ohman & Mineka, 2001) acknowledges that cultural artifacts with strong cultural threat value, such as guns, can come to evoke the same aversive response by engaging the human "fear module" similarly to phylogentically fear-relevant stimuli. The details for how ontogenetic threats can piggyback on learning systems evolved for phylogenetic threats remain unclear. Our results do obviously not rule out an evolutionary contribution to the effect of Pavlovian triggers on behavior, although the magnitude of this contribution must remain unspecified. Ongoing work based on evolutionary simulations suggest that preparedness is likely to enhance social as well as individual learning (Lindström, Selbing, & Olsson, n.d.), but empirical studies remain scarce (Barrett & Broesch, 2012; Cook & Mineka, 1990; Curio, Ernst, & Vieth, 1978). More generally, it is obvious that the behavioral tendencies revealed in **Study III** can be adaptive in an evolutionary sense, *if* the Pavlovian triggers actually are more predictive of punishment than other stimuli. The Pavlovian influence might be viewed as a strong prior (in the Bayesian sense) that allows an approximation of Bayesian inference (Marshall, Trimmer, Houston, & McNamara, 2013). An interesting venture would therefore be to use evolutionary models to understand and predict the environmental conditions leading to such priors, akin to how recent work have modeled emotions and moods as ecologically rational priors for decision making (Trimmer, Paul, Mendl, McNamara, & Houston, 2013).

Directly comparing the results of **Study II** and **Study III** reveal several important features: first, and perhaps unsurprisingly, different social (and non-social) stimuli have clear differential influences on avoidance learning in the social domain, and these differences can be explained by RL modeling (Ruff & Fehr, 2014), and second, individual differences in racial bias are highly important for understanding the basic computational mechanisms underlying social interactions involving different social groups. The latter finding is surprising, and might have important implications for understanding biased behavior in real world settings. Especially, delineating how such biases can be reduced is obviously of pivotal importance. I will return to related issues below when discussing the consequences of adaptive sampling for social behavior.

A direct comparison also reveals a puzzling discrepancy: the results of the RL modeling gave contrasting accounts of the underlying mechanisms. As described above, these studies address two sides of the same topic; social stimuli as reinforcers (**Study II**), and as cues or targets for actions (**Study III**). It is likely that the difference between these processes is reflected in the discrepant results. In **Study II**, we found that behavior was best accounted for by an RL model where different social reinforcers were associated with different learning rates (and not by different outcome magnitudes) and that individual differences in racial bias correlated selectively with the estimated learning rate for threatening out-group faces. In contrast, in **Study III**, we found that behavior best was described by a model where the stimulus elicited a fixed (Pavlovian) bias at the time of choice, rather than affecting the learning rate. Similarly, individual differences in racial bias were correlated with this Pavlovian bias parameter. Currently, it is unclear how to best reconcile these two models into

a unified account, which obviously is needed for understanding the computational mechanisms involved in avoidance learning in the social domain.

The results of **Study II** and **Study III** share a common caveat: computational modeling is associated with important limitations. No computational model, especially not one at the high level of abstraction used in **Study II-III**, is likely to provide a complete account of the processes underlying complex behavior. Furthermore, conclusions from model comparisons are contingent on the selected set of candidate models. Although, in theory, an unknown number of other models are possible, only a limited number of models can be tested (McClelland, 2009). Thus, while the results support particular candidate hypothesis, there are likely (many) other hypotheses consistent with the behavioral data. This caveat points to the importance of using several kinds of data, such as neuroimaging or single cell recordings, and models placed at different levels of Marr´s explanational framework (Marr, 1982; McClamrock, 1991), to constrain theory construction. Ideally, the models formulated in **Study II-III** should also be reconstructed with a high degree of biological realism, for example as detailed neural networks, to show that the known principles of nervous systems operating within the neural architecture of the human value-based decision making systems could reproduce the behavioral results (O'Reilly & Munakata, 2000).

In **Study IV**, we expanded the basic value-based decision framework to include social learning from the actions of others (Seymour et al., 2007). Specifically, we sought to provide a computational basis for how rare or non-existent punishment can generate widespread behavioral avoidance traditions, and thereby promoting, for example, norms of cooperation or food taboos (Fehr & Fischbacher, 2004; Henrich & Henrich, 2010). The economic and evolutionary literatures on social norms and traditions have generally been silent on the psychological mechanisms underpinning observed behavior. We hypothesized that the rewarding quality of avoiding an aversive outcome (see Table 1) might provide the motivational building block needed for such avoidance traditions, together with a social learning mechanism allowing avoidance behavior to be transmitted between individuals. As predicted, we found in four experiments that these mechanisms together can underlie avoidance behaviors by showing that humans are extremely prone to copy the behavior of others when in situations where punishment (electric shock) is possible (see Figure 6) and to transmit this behavior to other individuals in an intergenerational fashion.

An important part of **Study IV** was to provide a proof of principle that punishment omission and social learning together can generate, maintain, and transmit *large scale* avoidance traditions (Figure 5). To do this, a novel methodological approach was introduced; we used agent-based simulation models to investigate the emergent consequences of punishment omission and social learning in tandem in populations of simple agents. This modeling approach is widely appreciated and used within biology (DeAngelis & Mooij, 2005) and parts of the social sciences (Bonabeau, 2002), but relatively unknown within psychology (Smith & Conrey, 2007). However, in my opinion, agent-based simulation has the potential to provide bridges between detailed individual-level psychological (and neural) mechanisms and the

resulting behavior in groups and populations; a mechanistic linkage between micro and macro - levels of explanation (Hedström & Ylikoski, 2010). Furthermore, such models allow investigating the evolution of psychological computational mechanisms in populations of agents, and thereby provide a link between the mechanistic and functional levels of analysis of Tinbergen´s framework (Fawcett et al., 2012; McNamara & Houston, 2009; Nolfi, Parisi, & Elman, 1994; Trimmer et al., 2012). Obviously, however, models are most powerful when clearly related to empirical data. In **Study IV**, we confirmed the assumptions underlying the agent-based simulation by testing the basic predictions across four experiments, where we also used a model-fitting approach to confirm that human behavior correspond to the parameter range needed to reliably produce traditions in the simulation model. Thus, the empirical experiments were used to constrain the simulation model, which in turn was used to extrapolate from the results of the empirical experiments (c.f., Rand, Tarnita, Ohtsuki, & Nowak, 2013).

However, it is worth asking how general the findings of **Study IV** are in terms of different time scales and types of punishments. The agent-based simulations model makes no assumptions about either: the general pattern of results is predicted to generalize to any situations with the same underlying structure and where the psychological impact of avoiding threatening punishment is potent enough. Such simplicity in the modeling formulation is commonly seen as a virtue because the model might reveal explanatory generalities across many situations (Evans et al., 2013). It can, however, also be criticized for not actually corresponding to any natural system (Evans et al., 2013). Our experimental tests of the model assumptions, however, are naturally limited to a specific timescale (brief) and one type of punishment (electric shocks). Furthermore, the experiments mainly pertain to arbitrary avoidance traditions since sufficiently rare punishment for adaptive avoidance traditions could not be experimentally administered. This situation is common to many experimental tests of theoretical models, such as the Prisoner´s Dilemma. The Prisoner´s Dilemma is commonly considered to have far reaching implications for human behavior at different time scales and in a variety of payoff currencies, but experimental tests are normally restricted to the behavioral laboratory with monetary payoffs (Guala, 2012). The discrepancy between theoretical scope and experimental tests is a topic of debate (Guala, 2012), but there is however a general agreement that lab experiments are important for model and theory development, and that simple models do contribute to understanding of complex phenomena (Bowles & Gintis, 2011; Fehr & Gächter, 2002; Janssen, Holahan, Lee, & Ostrom, 2010).

Finally, one might consider if there are any general mechanisms by which the specific mechanisms identified in the studies of this thesis might influence everyday behavior. One such mechanism likely to mediate the influence of socially mediated valuations on learning and decision making is *adaptive sampling*. Adaptive sampling refers to the simple fact that if the value of an option is unknown, the only way to discover its value is by *sampling* it. If the outcome is favorable (relative to other options for example) the probability of choosing it again will increase and similarly, if the outcome was unfavorable, the future probability of choosing the option will decrease. This is exactly what the simple RL models used in this

thesis assumes. Such adaptive sampling will, for example, result in risk aversion due to the simple fact that high variance options often will result in unfavorable outcomes and that the true expected value of the uncertain option thus never is correctly estimated (Denrell, 2007). Of specific relevance to the social domain is the fact that negative initial outcomes, due to high variance or other factors, will decrease the future willingness to sample the option again. If the options are interactions with specific individuals, it is easy to see that this process can result in avoidance of a particular individual if the initial experience was negative and there is no extrinsic motivation for interacting with the individual again (Denrell, 2005). False negative initial expressions will thus remain uncorrected, whereas false positive expression will be corrected due to continued interactions. This process might for example contribute to inter-group biases and ostracism (Denrell, 2005). It is easy to see that the findings of **Study II** and **III** imply that this sampling based avoidance process will be amplified for certain stimuli or due to prior negative attitudes. For example, adaptive sampling could cause even a mild Pavlovian bias to decrease the probability of learning the true value of an option, such as interacting with an out-group member, which could have dramatic consequences for the future behavioral trajectory (Denrell, 2005). This pattern could be seen in **Study III** when choosing the Pavlovian trigger represented the optimal action. Due to the Pavlovian influence, the participants instead tended to oversample the sub-optimal action, and thereby never accurately estimating the value of the other (optimal) option. Similarly, **Study II** showed that individuals with high racial bias learned faster than low bias individuals to avoid threatening out-group members, and would thus less be likely to revise their impression through future interactions. One factor that can either exaggerate or reduce such biases is social learning. Social learning can provide information about the outcomes of non-sampled options, which can reduce risk aversion (Arbilly, Motro, Feldman, & Lotem, 2011; Fazio, Eiser, & Shook, 2004). However, social learning can also propagate false negative impressions of the options, which thereafter are unlikely to be corrected by individual sampling (Fazio et al., 2004). In **Study IV**, we saw that social learning thus could underlie arbitrary avoidance traditions.

Together, the four studies included in this thesis show, in line with a growing literature (Ruff & Fehr, 2014), that basic value-based decision mechanisms play an important role in shaping social behavior in response to punishment and threats of punishment. Understanding the role of these mechanisms in social behavior might function as an important building block for bridging the questions of *how* and *why* (Bateson & Laland, 2013) through evolutionary modeling, and as a starting points for more detailed analysis at all levels of mechanistic explanation (Marr, 1982),

5 ACKNOWLEDGEMENTS

6 REFERENCES

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., Maner, J. K., et al. (2006). They all look the same to me (unless they're angry): from out-group homogeneity to out-group heterogeneity. *Psychological science*, *17*(10), 836-40. doi:10.1111/j.1467-9280.2006.01790.x

Akrami, N., Ekehammar, B., & Araya, T. (2000). Classical and modern racial prejudice: A study of attitudes toward immigrants in Sweden. *European Journal of Social …*, 410-421.

Arbilly, M., Motro, U., Feldman, M. W., & Lotem, A. (2011). Evolution of social learning when high expected payoffs are associated with high risk of failure. *Journal of the Royal Society, Interface / the Royal Society*, *8*(64), 1604-15. doi:10.1098/rsif.2011.0138

Askew, C., & Field, A. P. (2008). The vicarious learning pathway to fear 40 years on. *Clinical psychology review*, *28*(7), 1249-65. doi:10.1016/j.cpr.2008.05.003

Aunger, R. (1994). Are food avoidances maladaptive in the Ituri Forest of Zaire? *Journal of Anthropological Research*.

Barrett, H. C., & Broesch, J. (2012). Prepared social learning about dangerous animals in children. *Evolution and Human Behavior*, *33*(5), 499-508. doi:10.1016/j.evolhumbehav.2012.01.003

Bateson, P., & Laland, K. N. (2013). Tinbergen's four questions: an appreciation and an update. *Trends in ecology & evolution*, *28*(12), 712-8. doi:10.1016/j.tree.2013.09.013

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge, UK: Cambridge University Press.

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *358*(1431), 561-72. doi:10.1098/rstb.2002.1220

Bolhuis, J. J., Brown, G. R., Richardson, R. C., & Laland, K. N. (2011). Darwin in mind: new opportunities for evolutionary psychology. *PLoS biology*, *9*(7), e1001109. Public Library of Science. doi:10.1371/journal.pbio.1001109

Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review*, *77*(1), 32-48.

Bonabeau, E. (2002). Agent-based modeling: methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, *99 Suppl 3*, 7280-7. doi:10.1073/pnas.082080899

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict Monitoring and Cognitive Control, *108*(3), 624-652. doi:10.1037//0033-295X.I08.3.624

Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, *36*(1), 74-97. Nature Publishing Group. doi:10.1038/npp.2010.151

Bouton, M. (2007). *Learning and behavior: A contemporary synthesis.* Sunderland, USA: Sinauer Associates, Inc.

Bowles, S., & Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution* (p. 280). Princeton University Press.

Boyd, R., Gintis, H., & Bowles, S. (2010). Coordinated punishment of defectors sustains cooperation and can proliferate when rare. *Science (New York, N.Y.)*, *328*(5978), 617-20. doi:10.1126/science.1183665

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(6), 3531-5. doi:10.1073/pnas.0630443100

Boyd, R., & Richerson, P. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and sociobiology*.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl* , 10918-25. doi:10.1073/pnas.1100290108

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(32), 14431-6. doi:10.1073/pnas.1003111107

Cant, M. A. (2011). The role of threats in animal cooperation. *Proceedings. Biological sciences / The Royal Society*, *278*(1703), 170-8. doi:10.1098/rspb.2010.1241

Cook, M., & Mineka, S. (1990). Selective associations in the observational conditioning of fear in rhesus monkeys. *Journal of experimental psychology. Animal behavior processes*, *16*(4), 372-89.

Curio, E., Ernst, U., & Vieth, W. (1978). Cultural transmission of enemy recognition: one function of mobbing. *Science (New York, N.Y.)*, *202*(4370), 899-901. doi:10.1126/science.202.4370.899

Davey, G. C. L. (1992). An expectancy model of laboratory preparedness effects. *Journal of Experimental Psychology: General*.

Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias? *Behavioral and Brain Sciences*, *18*(02), 289-297. Cambridge University Press.

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In M. R. Delgado, E. Phelps, & T. W. Robbins (Eds.), *Decision Making, Affect, and Learning: Attention and Performance XXIII* (pp. 1-26). New York: Oxford University Press.

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*.

Daw, N. D., O'Doherty, J., Dayan, P., Seymour, B., & Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature*.

Dawson, E. H., Avarguès-Weber, A., Chittka, L., & Leadbeater, E. (2013). Learning by observation emerges from simple associations in an insect model. *Current biology : CB*, *23*(8), 727-30. doi:10.1016/j.cub.2013.03.035

Dayan, P. (2009). Goal-directed control and its antipodes. *Neural networks : the official journal of the International Neural Network Society*, *22*(3), 213-9. doi:10.1016/j.neunet.2009.03.004

Dayan, P., & Huys, Q. J. M. (2009). Serotonin in affective control. *Annual review of neuroscience*, *32*, 95-126. doi:10.1146/annurev.neuro.051508.135607

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology*, *18*(2), 185-96. doi:10.1016/j.conb.2008.08.003

Dayan, P., Niv, Y., Seymour, B., & Daw, N. D. (2006). The misbehavior of value and the discipline of the will. *Neural networks : the official journal of the International Neural Network Society*, *19*(8), 1153-60. doi:10.1016/j.neunet.2006.03.002

Dayan, P., & Seymour, B. (2007). Values and Actions in Aversion. (P. W. Glimcher, C. Camerer, E. Fehr, & R. A. Poldrack, Eds.)*Neuroeconomics: Decision making and the brain*, (November), 175-192. Academic Press.

DeAngelis, D., & Mooij, W. (2005). Individual-based modeling of ecological and evolutionary processes. *Annual Review of Ecology, Evolution, and ….*

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. a. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1511), 3787-800. doi:10.1098/rstb.2008.0161

Denrell, J. (2005). Why most people disapprove of me: experience sampling in impression formation. *Psychological review*, *112*(4), 951-78. doi:10.1037/0033-295X.112.4.951

Denrell, J. (2007). Adaptive learning and risk taking. *Psychological review*.

Dolan, R. J. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312-25. doi:10.1016/j.neuron.2013.09.007

Domjan, M. (2005). Pavlovian conditioning: a functional perspective. *Annual review of psychology*, *56*, 179-206. doi:10.1146/annurev.psych.55.090902.141409

Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., de Langhe, R., Lessells, C. M., Merz, M., et al. (2013). Do simple models lead to generality in ecology? *Trends in ecology & evolution*, *28*(10), 578-83. doi:10.1016/j.tree.2013.05.022

Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., & McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in cognitive sciences*, *18*(3), 153-161. Elsevier. doi:10.1016/j.tics.2013.12.012

Fawcett, T. W., Hamblin, S., & Giraldeau, L.-A. (2012). Exposing the behavioral gambit: the evolution of learning and decision rules. *Behavioral Ecology*, *24*(1), 2-11. doi:10.1093/beheco/ars085

Fazio, R. H. R., Eiser, J. R., & Shook, N. J. N. (2004). Attitude formation through exploration: valence asymmetries. *Journal of personality and social psychology*, *87*(3), 293-311. doi:10.1037/0022-3514.87.3.293

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human behavior*.

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*.

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*(5), 567-89.

Fujiwara, J., Tobler, P. N., Taira, M., Iijima, T., & Tsutsui, K.-I. (2009). A parametric relief signal in human ventrolateral prefrontal cortex. *NeuroImage*, *44*(3), 1163-70. doi:10.1016/j.neuroimage.2008.09.050

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385-390. SAGE Publications. doi:10.1111/j.1467-9280.1993.tb00586.x

Gigerenzer, G., & Goldstein, D. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*.

Gintis, H. (2007). A framework for the unification of the behavioral sciences. *The Behavioral and brain sciences*, *30*(1), 1-16; discussion 16-61. doi:10.1017/S0140525X07000581

Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 15647-54. doi:10.1073/pnas.1014269108

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of experimental psychology. General*, *136*(3), 389-413. doi:10.1037/0096-3445.136.3.389

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*(1), 4-27.

Griffin, A. (2004). Social learning about predators: a review and prospectus. *Animal Learning & Behavior*.

Guala, F. (2012). Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *The Behavioral and brain sciences*, *35*(1), 1-15. doi:10.1017/S0140525X11000069

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in cognitive sciences*, *18*(4), 194-202. doi:10.1016/j.tics.2014.01.003

Hajcak, G. (2012). What We've Learned From Mistakes: Insights From Error-Related Brain Activity. *Current Directions in Psychological Science*, *21*(2), 101-106. SAGE Publications. doi:10.1177/0963721412436809

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology*, *42*(2), 151-60. doi:10.1111/j.1469-8986.2005.00270.x

Hammerstein, P., & Hagen, E. H. (2005). The second wave of evolutionary economics in biology. *Trends in ecology & evolution*, *20*(11), 604-9. doi:10.1016/j.tree.2005.07.012

Hare, T., & O'Doherty, J. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *The Journal of ….*

Hedström, P., & Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology*.

Heerey, E. A. (2013). Learning From Social Rewards Predicts Individual Differences in Self-Reported Social Ability. *Journal of experimental psychology. General*. doi:10.1037/a0031511

Henrich, J., & Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings. Biological sciences / The Royal Society*, *277*(1701), 3715-24. doi:10.1098/rspb.2010.1191

Herrnstein, R. J. (1970). On the law of effect. *Journal of the experimental analysis of behavior*, *13*(2), 243-66.

Herrnstein, R. J., Loewenstein, G. F., Prelec, D., & Vaughan, W. (1993). Utility maximization and melioration: Internalities in individual choice. *Journal of Behavioral Decision Making*, *6*(3), 149-185. doi:10.1002/bdm.3960060302

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *109*(4), 679-709.

Houston, A. I., & McNamara, J. (1981). How to maximize reward rate on two variable-interval paradigms. *Journal of the experimental analysis of behavior*, *35*(3), 367-96.

Houston, A. I., McNamara, J. M., & Steer, M. D. (2007). Do we expect natural selection to produce rational behaviour? *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *362*(1485), 1531-43. doi:10.1098/rstb.2007.2051

Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and pavlovian responding. (A. Rangel, Ed.)*PLoS computational biology*, *7*(4), e1002028. Public Library of Science. doi:10.1371/journal.pcbi.1002028

Janssen, M. A., Holahan, R., Lee, A., & Ostrom, E. (2010). Lab experiments for the study of social-ecological systems. *Science (New York, N.Y.)*, *328*(5978), 613-7. doi:10.1126/science.1183532

Johnson, D. D. P., Blumstein, D. T., Fowler, J. H., & Haselton, M. G. (2013). The evolution of error: error management, cognitive constraints, and adaptive decision-making biases. *Trends in ecology & evolution*, *28*(8), 474-81. doi:10.1016/j.tree.2013.05.014

Johnston, T. D. (1982). Selective Costs and Benefits in the Evolution of Learning. *Advances in the Study of Behavior*, *12*, 65-106. doi:10.1016/S0065-3454(08)60046-7

Kenner, N. M., Mumford, J. A., Hommer, R. E., Skup, M., Leibenluft, E., & Poldrack, R. A. (2010). Inhibitory motor control in response stopping and response switching. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(25), 8512-8. doi:10.1523/JNEUROSCI.1096-10.2010

Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS biology*, *4*(8), e233. doi:10.1371/journal.pbio.0040233

Laland, K. N., & Brown, G. (2011). *Sense and Nonsense: Evolutionary Perspectives on Human Behaviour*.

Laland, K. N., Odling-Smee, J., Hoppitt, W., & Uller, T. (2012). More on how and why: cause and effect in biology revisited. *Biology & Philosophy*. doi:10.1007/s10539-012-9335-1

LeDoux, J. (2012). Rethinking the Emotional Brain. *Neuron*, *73*(4), 653-676. Elsevier Inc. doi:10.1016/j.neuron.2012.02.004

Levin, D. T. (2000). Race as a visual feature: using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of experimental psychology. General*, *129*(4), 559-74.

Lin, A., Adolphs, R., & Rangel, A. (2011). Social and monetary reward learning engage overlapping neural substrates. *Social cognitive and affective neuroscience*, *7*(3), 274-81. doi:10.1093/scan/nsr006

Lindström, B., Selbing, I., Molapour, T., & Olsson, A. (2014). Racial Bias Shapes Social Reinforcement Learning. *Psychological …*.

Lindström, B., Selbing, I., & Olsson, A. (n.d.). Evolutionary preparedness and social learning interact to promote survival in dangerous environments.

Mani, L. (1987). Contentious traditions: the debate on sati in colonial India. *Cultural Critique*.

Marr, D. (1982). Vision. A computational investigation into the human representation and processing of visual information. *WH San Francisco: Freeman and ….*

Marshall, J. A. R., Trimmer, P. C., Houston, A. I., & McNamara, J. M. (2013). On evolutionary explanations of cognitive biases. *Trends in ecology & evolution*, *28*(8), 469-73. doi:10.1016/j.tree.2013.05.013

McClamrock, R. (1991). Marr's three levels: A re-evaluation. *Minds and Machines*, *1*(2), 185-196. doi:10.1007/BF00361036

McClelland, J. L. J. (2009). The place of modeling in cognitive science. *Topics in Cognitive Science*, *1*(1), 11-38. doi:10.1111/j.1756-8765.2008.01003.x

McHugh, S. B., Barkus, C., Huber, A., Capitão, L., Lima, J., Lowry, J. P., & Bannerman, D. M. (2014). Aversive prediction error signals in the amygdala. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *34*(27), 9024-33. doi:10.1523/JNEUROSCI.4465-13.2014

McNally, G. P., Johansen, J. P., & Blair, H. T. (2011). Placing prediction into the fear circuit. *Trends in neurosciences*, *34*(6), 283-92. Elsevier Ltd. doi:10.1016/j.tins.2011.03.005

McNamara, J. M., & Houston, A. I. (2009). Integrating function and mechanism. *Trends in ecology & evolution*, *24*(12), 670-5. doi:10.1016/j.tree.2009.05.011

Mery, F., & Burns, J. G. (2009). Behavioural plasticity: an interaction between evolution and experience. *Evolutionary Ecology*, *24*(3), 571-583. doi:10.1007/s10682-009-9336-y

Mesoudi, A., & Whiten, A. (2008). Review. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *363*(1509), 3489-501. doi:10.1098/rstb.2008.0129

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual review of neuroscience*, *24*, 167-202. doi:10.1146/annurev.neuro.24.1.167

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychological bulletin*, *117*(3), 363-86.

Morecraft, R. J., Louie, J. L., Herrick, J. L., & Stilwell-Morecraft, K. S. (2001). Cortical innervation of the facial nucleus in the non-human primate: a new interpretation of the effects of stroke and related subtotal brain trauma on the muscles of facial expression. *Brain : a journal of neurology*, *124*(Pt 1), 176-208.

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings. Biological sciences / The Royal Society*, *279*(1729), 653-62. doi:10.1098/rspb.2011.1172

Mostofsky, S. H., & Simmonds, D. J. (2008). Response inhibition and response selection: two sides of the same coin. *Journal of cognitive neuroscience*, *20*(5), 751-61. doi:10.1162/jocn.2008.20500

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: the role of target gender. *Psychological science*, *20*(2), 155-8. doi:10.1111/j.1467-9280.2009.02273.x

Nolfi, S., Parisi, D., & Elman, J. L. (1994). Learning and Evolution in Neural Networks. *Adaptive Behavior*, *3*(1), 5-28. doi:10.1177/105971239400300102

Ohman, A. (2009). Of snakes and faces: an evolutionary perspective on the psychology of fear. *Scandinavian journal of psychology*, *50*(6), 543-52. doi:10.1111/j.1467-9450.2009.00784.x

Ohman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological review*, *108*(3), 483-522.

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The role of social groups in the persistence of learned fear. *Science (New York, N.Y.)*, *309*(5735), 785-7. doi:10.1126/science.1113551

Olsson, A., & Phelps, E. A. (2007). Social learning of fear. *Nature neuroscience*, *10*(9), 1095-102. doi:10.1038/nn1968

O'Reilly, R., & Munakata, Y. (2000). Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain.

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

Panksepp, J., & Panksepp, J. B. (2000). The Seven Sins of Evolutionary Psychology, *6*(2), 108-131.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual review of psychology*, *52*, 111-39. doi:10.1146/annurev.psych.52.1.111

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature neuroscience*, *4*(4), 437-41. doi:10.1038/86110

Plassmann, H. (2010). Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *The Journal of ….*

Rand, D. G., Tarnita, C. E., Ohtsuki, H., & Nowak, M. A. (2013). Evolution of fairness in the one-shot anonymous Ultimatum Game. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(7), 2581-6. doi:10.1073/pnas.1214167110

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience*, *9*(7), 545-56. doi:10.1038/nrn2357

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., Fogarty, L., et al. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*, *328*(5975), 208-13. doi:10.1126/science.1184719

Rendell, L., Fogarty, L., & Hoppitt, W. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in cognitive ….*

Rescorla, R. A. (1988). Pavlovian conditioning. It's not what you think it is. *The American psychologist*, *43*(3), 151-60.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological review*, *74*(3), 151-82.

Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research ….*

Riesel, A., Weinberg, A., Endrass, T., Kathmann, N., & Hajcak, G. (2012). Punishment has a lasting impact on error-related brain activity. *Psychophysiology*, *49*(2), 239-47. doi:10.1111/j.1469-8986.2011.01298.x

Rolls, E. T. (2000). On The brain and emotion. *Behavioral and Brain Sciences*, *23*(2), 219-228. Cambridge University Press. doi:10.1017/S0140525X00512424

Ruff, C. C., & Fehr, E. (2014). The neurobiology of rewards and values in social decision making. *Nature reviews. Neuroscience*, *15*(8), 549-62. doi:10.1038/nrn3776

Rutledge, R. B., Dean, M., Caplin, A., & Glimcher, P. W. (2010). Testing the reward prediction error hypothesis with an axiomatic model. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *30*(40), 13525-36. doi:10.1523/JNEUROSCI.1747-10.2010

Seymour, B., & Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron*, *58*(5), 662-71. doi:10.1016/j.neuron.2008.05.020

Seymour, B., O'Doherty, J. P. J., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., et al. (2004). Temporal difference models describe higher-order learning in humans. *Nature*, *429*(6992), 664-7. doi:10.1038/nature02581

Seymour, B., Singer, T., & Dolan, R. (2007). The neurobiology of punishment. *Nature reviews. Neuroscience*, *8*(4), 300-11. doi:10.1038/nrn2119

Shackman, A. J., Salomons, T. V., Slagter, H. a, Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature reviews. Neuroscience*, *12*(3), 154-67. Nature Publishing Group. doi:10.1038/nrn2994

Shore, D. M., & Heerey, E. a. (2011). The value of genuine and polite smiles. *Emotion (Washington, D.C.)*, *11*(1), 169-74. doi:10.1037/a0022601

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia*, *46*(1), 224-32. doi:10.1016/j.neuropsychologia.2007.07.015

Skinner, B. F. (1981). Selection by consequences. *Science (New York, N.Y.)*, *213*(4507), 501-4.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: a new approach for theory building in social psychology. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, *11*(1), 87-104. doi:10.1177/1088868306294789

Snell-Rood, E. C. (2013). An overview of the evolutionary causes and consequences of behavioural plasticity. *Animal Behaviour*, *85*(5), 1004-1011. doi:10.1016/j.anbehav.2012.12.031

Solomon, R. L. (1980). The opponent-process theory of acquired motivation: the costs of pleasure and the benefits of pain. *The American psychologist*, *35*(8), 691-712.

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), 966-73. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. doi:10.1038/nn.3413

Sutton, R., & Barto, A. (1998). Introduction to reinforcement learning.

Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human pavlovian-instrumental transfer. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *28*(2), 360-8. doi:10.1523/JNEUROSCI.4028-07.2008

Teigen, K. H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology*, *4*(4), 525-547. doi:10.1177/0959354394044004

Thorndike, E. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied*.

Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift für Tierpsychologie*, *20*(4), 410-433. doi:10.1111/j.1439-0310.1963.tb01161.x

Tobler, P. N., Dickinson, A., & Schultz, W. (2003). Coding of Predicted Reward Omission by Dopamine Neurons in a Conditioned Inhibition Paradigm. *J. Neurosci.*, *23*(32), 10402-10410.

Trimmer, P. C., McNamara, J. M., Houston, A. I., & Marshall, J. A. R. (2012). Does natural selection favour the Rescorla-Wagner rule? *Journal of theoretical biology*, *302*, 39-52.

Trimmer, P. C., Paul, E., Mendl, M., McNamara, J., & Houston, A. I. (2013). On the Evolution and Optimality of Mood States. *Behavioral Sciences*, *3*(3), 501-521. Multidisciplinary Digital Publishing Institute. doi:10.3390/bs3030501

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in cognitive sciences*, *18*(5), 259-67. doi:10.1016/j.tics.2014.02.009

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*(6842), 43-8. doi:10.1038/35083500

Webster, M. M., & Laland, K. N. (2008). Social learning strategies and predation risk: minnows copy only when using private information would be costly. *Proceedings. Biological sciences / The Royal Society*, *275*(1653), 2869-76. doi:10.1098/rspb.2008.0817

Winterhalder, B., Smith, E. A., & Bruce Winterhalder, E. A. S. (2000). Analyzing adaptive strategies: Human behavioral ecology at twenty-five. *Evolutionary Anthropology: Issues, News, and Reviews*, *9*(2), 51. doi:10.1002/(SICI)1520-6505(2000)9:2<51::AID-EVAN1>3.3.CO;2-Z