From CLINICAL NEUROSCIENCE
Karolinska Institutet, Stockholm, Sweden

LEARNING FROM THE BEHAVIORS AND EXPERIENCES OF OTHERS

Ida Selbing

Stockholm 2017

Front page: One of the stimuli used in **Study IV**.

Learning From the Behaviors and Experiences of Others
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Ida Selbing

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

Co-supervisor(s):
Professor Christian Balkenius
Lund University
Department of Philosophy
Division of Cognitive Science

Opponent:
Professor Christian Ruff
University of Zürich
Department of Economics

Examination Board:
Professor Peter Juslin
Uppsala University
Department of Psychology

Professor Anna Dreber Almenberg
Stockholm School of Economics
Department of Economics

Dr Marc Guitart-Masip
Karolinska Institutet
Department of Neurobiology,
Care Sciences and Society
ARC (Aging Research Center)

ABSTRACT

Learning to fear and avoid what is dangerous is crucial for survival. Perhaps equally important is the ability to learn that something which was previously dangerous is now safe. Although we can learn about dangers individually, through our own experiences, it is likely more safe to learn about them from others, by observing their behaviors and reactions. In a sense, this allows us to learn through the experiences of others. The overarching goal of this thesis is to deepen our understanding of how we learn about fear and safety through observation of others.

In **Study I** we let participants undergo an observational extinction paradigm to investigate if safety learning was facilitated through observation of a calm learning model. In a direct conditioning stage participants first learned to associate a stimulus with fear. Next, they learned through that the previously feared stimulus was now safe. This extinction of fear was either direct or vicarious (observational). We demonstrated that attenuation of fear was greater following vicarious rather than direct extinction. We further showed that this was driven by the learning model's experience of safety.

Although learning through others is likely an efficient way of learning, observational learning also has to be applied critically, for instance by not copying the choices of someone that performs poorly. In **Study II** and **Study III** we investigated how people learned to make choices through observation of others, demonstrators, which had either a high or low ability. In both studies, participants learned a simple probabilistic two-choice task to avoid shock. Results from **Study II** demonstrated that people were able to use the observational information to improve performance regardless of the ability (skill) of the demonstrator. They only copied the choices of the demonstrator with high ability and they were able to learn from observing the consequences of a demonstrator's choice regardless of the demonstrator's ability. In **Study III** we also provided participants with descriptions of the abilities of the demonstrators. Our results showed that describing the demonstrator as low in ability impaired observational learning, regardless of the actual ability of the demonstrator and that this is likely driven by a difference in attention directed towards the observational information.

An inability to discriminate threatening from safe stimuli is typical for individuals suffering from anxiety. In **Study IV** we investigated how observational fear conditioning is affected by the learning model's expressed anticipatory anxiety. Results showed that participants were able to discriminate the threatening from the safe stimuli equally well from a learning model that behaved anxiously (i.e. did not discriminate) as from one that did not behave anxiously (i.e. did discriminate).

The results presented in this thesis increase our understanding of how healthy individuals learn about aversive events and stimuli through observation of the behaviors and reactions of others and how these reflect the observed individuals' experiences.

LIST OF SCIENTIFIC PAPERS

The scientific work underlying this thesis is included in the following papers, referred to by their roman numerals (**Study I-IV**):

I.   Golkar, A., **Selbing, I.**, Flygare, O., Öhman, A. & Olsson, A. (2013). Other people as means to a safe end: Vicarious extinction blocks the return of learned fear. *Psychological Science*, 24(11), 2182-2190. doi:10.1177/0956797613489890 *

II. **Selbing, I.**, Lindström, B. & Olsson, A. (2014). Demonstrator skill modulates observational aversive learning. *Cognition*, 133, 128-139. https://doi.org/10.1016/j.cognition.2014.06.010 **

III. **Selbing, I.** & Olsson, A. (2017). Beliefs about Other's Abilities Impair Learning from Observation. *Submitted manuscript*.

IV. **Selbing, I.** & Olsson, A. (2017). The Effect of a Learning Model's Anxious Behavior on Observational Fear Conditioning, *Manuscript in preparation*.

ADDITIONAL SCIENTIFIC PAPERS

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

i. Lindström, B., **Selbing, I.**, Molapour, T. & Olsson, A. (2013). Racial bias shapes social reinforcement learning. *Psychological Science*, 25(3), 711-719.

ii. Kleberg, J., **Selbing, I.**, Lundqvist, D., Hofvander, B. & Olsson, A. (2015). Spontaneous eye movements and trait empathy predict vicarious learning of fear. *International Journal of Psychophysiology*, 98(3), 577-583.

iii. Golkar, A., Haaker, J., **Selbing, I.** & Olsson, A. (2016). Neural signals of vicarious extinction learning. *Social cognitive and affective neuroscience*, 11(10), 1541-1549.

iv. Lindström, B., **Selbing, I.** & Olsson, A. (2016). Co-evolution of social learning and preparedness in dangerous environments. *PLoS ONE*, 11(8).

v. Haaker, J., Golkar, A., **Selbing, I.** & Olsson, A. (2017). Assessment of social transmission of threats in humans using observational fear conditioning. *Nature Protocols*, 12(7), 1378-1386. (All authors contributed equally.)

vi. Lindström, B., Jangard, S., **Selbing, I**. & Olsson, A. (in press). The role of a "common is moral" heuristic in the stability and change of moral norms. *Journal of Experimental Psychology: General.*

CONTENTS

LIST OF ABBREVIATIONS

ACC	Anterior cingulate cortex
AI	Anterior insula
BOLD	Blood-oxygen-level dependent
CR	Conditioned response
CS	Conditioned stimulus
fMRI	functional magnetic resonance imaging
FPS	Fear-potentiated startle
IADS	International Affective Digital Sounds
KDEF	Karolinska Directed Emotional Faces
RL	Reinforcement learning
SCL	Skin conductance level
SCR	Skin conductance response
UR	Unconditioned response
US	Unconditioned stimulus

INTRODUCTION

The thesis presented here deals with aspects of learning from others. The benefits of being able to learn from others are undoubtedly vast; it enables us to have culture and language and to communicate knowledge over time and space (Richerson & Boyd, 2008). It also allows us to safely learn about dangers that we might encounter (Rachman, 1977) and quickly learn which of a set of decisions that is optimal to select (McElreath et al., 2008).

It has been suggested that learning from others is especially beneficial in dangerous, or aversive, contexts, when individual learning is costly (H. C. Barrett & Broesch, 2012; Laland, 2004) and the work described in this thesis focuses on observational aversive learning, both associative learning of fear (**Study I)** and safety (**Study IV**), as well as instrumental avoidance learning (**Study II** and **Study III**). Many aspects in an observational learning situation are believed to influence learning. The focus of this thesis is on the influence from the behavior and expressed experiences of the person learnt from on observational learning.

In the following pages I will first briefly describe the background of the work presented in this thesis by giving a short description of how fear can be acquired and how we learn which actions to take. I will also discuss some core concepts related to social and observational learning. The aims of the thesis will then be stated, followed by a presentation of some of the methods used to reach these aims. Next follows a short summary of the studies included. Finally, I will discuss the findings from these studies from a more general perspective.

1.1   LEARNING FROM OTHERS

Historically, the scientific interest in social forms of learning has focused on several, somewhat separate, paths of research. These can roughly be categorized as research focusing on: the ability of (non-human) animals to learn a behavior through imitation (Galef, 1988), observational fear conditioning (Susan Mineka & Cook, 1988; Rachman, 1977) and, more recently, the cultural transmission of information such as beliefs, skills and language (Boyd & Richerson, 1985; Tomasello, Kruger, & Ratner, 1993).

The term social learning is usually defined broadly as learning that is influenced by observation of, or interaction with, another individual or its products (C. M. Heyes, 1994). In practice, the scientific field usually referred to as social learning has mainly focused on the adaptiveness and evolution of behavior (J. Kendal, Giraldeau, & Laland, 2009; Laland, 2004). In this sense, social learning is closely tied to the field of cultural evolution (Boyd & Richerson, 1985; Henrich et al., 2016), a field with an often strong theoretical approach (Kameda & Nakanishi, 2002; Luke Rendell et al., 2011) but with less interest in cognitive mechanisms (C. Heyes, 2012). Within the field of social learning it is common to talk about learning as a form of copying, and also to refer to adaptive 'strategies of copying' such as 'copy when uncertain' (Laland, 2004). It is important to understand that these references to strategies do not imply that the individuals who comply to them necessarily do this through

higher-order deliberation. Rather, it is often assumed that compliance to these (adaptive) strategies is the result of natural selection, culture or learning.

The term observational conditioning was first used by Mineka and colleagues (Michael Cook, Mineka, Wolkenstein, & Laitsch, 1985; Susan Mineka, Davidson, Cook, & Keir, 1984) and is often used interchangeably with the term vicarious learning (Askew & Field, 2008) or vicarious conditioning (Bandura, 1965). The scientific field of observational conditioning (including observational extinction) is rooted in the work on classical (Pavlovian) conditioning (Rescorla & Wagner, 1972) and refers to the learning of a stimulus-stimulus association through observation.

In the studies presented in this thesis we mainly refer to observational, rather than social, forms of learning (**Study II - IV**). In **Study I** we have used the term vicarious instead of observational. However, there is no established consensus on how to refer to the observed other individual during observational learning. We use the term 'learning model' in **Study I** and **Study IV,** where a more classical observational fear conditioning paradigm is used, and 'demonstrator' in **Study II** and **Study III** where a more typical decision making framework is used to study observational avoidance learning. The term 'demonstrator' is meant to describe an individual who is observed, not someone that actively or intentionally demonstrates something to the observer.

1.2   FEAR AND SAFETY

1.2.1   Conditioning

Conditioning is described as the learning of an association between two stimuli. During fear conditioning, an association is formed between a previously neutral stimulus, the so called conditioned stimulus (CS), and an aversive stimulus, the unconditioned stimulus (US), which by itself elicits an aversive response, the unconditioned response (UR). Following learning, exposure to the CS leads to a fear response, the conditioned response (CR), in the learning agent. Although it is often described that the learnt association depends on the pairing of the two stimuli, a more accurate view is that the association is dependent on an exposure to a relation between them (Rescorla, 1988). Thus, it is important that the CS is predictive of the US for an association to form, for instance by having a low probability of experiencing the US in absence of the CS. However, learning does not rely only on the relation between stimuli. Instead, some stimuli, such as snakes or threatening faces, believed to have been fear relevant from an evolutionary perspective, are more easily associated with fear (Öhman & Mineka, 2001). In addition, some associations are more easily formed than others, such as the association between a gustatory (taste) stimulus and a toxin rather than a gustatory stimulus and a shock (Garcia & Koelling, 1966). This is important because it means that the learning agent can be seen as an "information seeker using logical and perceptual relations among events, along with its own preconception, to form a sophisticated representation of its world" (Rescorla, 1988, p. 154).

Fear conditioning is often studied using discriminatory paradigms where participants are presented with two neutral stimuli, CSs, out of which only one is paired with an aversive stimulus, the US. The CS paired with the US is denoted the CS+ while the unpaired CS is denoted the CS-. The CS- here functions as a control stimulus. Following successful conditioning, presentation of the CS+ will elicit a stronger conditioned response, CR, than the CS- (e.g. Shmuel Lissek et al., 2009). There are several methods to measure fear in humans (Lonsdorf et al., 2017). Physiological measures include for instance the skin conductance response (SCR), fear-potentiated startle (FPS), and heart rate. It is also common to use verbal report such as US expectancy ratings and contingency ratings.

1.2.2 Extinction

Equally important to the ability to learn about threats is the ability to adapt to new information and learn that what was previously threatening is now safe. During extinction of a fear memory, it is not believed that the original CS-US association is weakened or erased. Rather, current theories suggest that extinction is caused by the formation of a safety memory, an association between the CS and the non-occurrence of the US (Myers & Davis, 2007). This safety memory appears to be more context dependent than the threat memory (Bouton, Westbrook, Corcoran, & Maren, 2006) as if an exception to the general rule, specific for the context, is learnt. The safety memory then coexists with the fear memory and activation of the safety memory inhibits expression of the fear memory (Bouton, 2004).

Research on the mechanisms of extinction learning is especially important from a clinical perspective since excessive fear is involved in anxiety disorders and phobias (Lang, Davis, & Öhman, 2000). Unfortunately, extinction learning is considered rather 'fragile', possibly due to its dependency on context, therefore resulting in inefficient behavioral interventions with frequent relapse of fear (Craske et al., 2008).

In extinction paradigms, participants first acquire fear through a conditioning procedure, as described above. After acquisition, the participants are presented with the CSs again, this time in absence of the US. The participant now learns that the previous relation between the CS+ and US no longer exists, leading to an attenuation of the heightened fear response to the CS+. The efficiency of extinction is then tested in an ensuing test phase where the participants are presented with the CSs again. The test phase is designed according to one of several principals known to induce the return of fear, for instance the administration of the test phase following the passing of time, spontaneous recovery (Rescorla, 2004), or a change in context, renewal (Bouton, 2004). In **Study I**, where we investigated the properties of observational extinction, the test phase was preceded by exposure to unsignaled USs, according to a reinstatement procedure. Reinstatement is believed to be caused by context conditioning, where the context is associated with the US, in turn leading to an activation of the CS-US rather than the CS-no US association (Bouton, 2004).

In addition to the ability to learn to associate stimuli with fear (or safety), humans can also learn to associate stimuli with valence, e.g. 'likes' and 'dislikes' through a procedure referred

to as evaluative conditioning (EC). In EC, the preference for a stimulus, the CS, is changed due to pairing with another liked or disliked stimulus, the US (Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010). It has also been shown that such changes in preferences can occur following observation of another individual exposed to the CS-US contingency and expressing liking or disliking the US (Baeyens, Kaes, Eelen, & Silverans, 1996; Baeyens, Vansteenwegen, De Houwer, & Crombez, 1996). However, EC lies beyond the scope of this thesis and will therefore not be discussed further.

1.2.3 Observational learning of fear and safety

Observational learning of fear and safety is an important way for an individual to safely learn what is dangerous and what is safe. Observational learning of fear is also, together with direct learning, regarded as a pathway to the acquiring of excessive fear, phobias and anxiety, including post-traumatic stress disorder, PTSD (DSM-5, American Psychiatric Association, 2013; Rachman, 1977).

In most human observational fear conditioning paradigms, a learning agent, the observer, is exposed to the relation between a CS and US through observation of another individual, a learning model, who directly experiences them both (e.g. A. Golkar, Castro, & Olsson, 2015; Olsson, Nearing, & Phelps, 2007). The learning model reacts aversively to the US, expressing an unconditioned response (UR). The observer can then infer a US from the learning model's UR and associate it with the CS. In these paradigms, learning critically depends on the learning model's UR and the ability of the observer to correctly interpret or understand it. As an example, Hygge (1976) demonstrated that observational conditioning was dependent on the emotional strength of the UR. Recent work has shown that empathic appraisal, as well as trait empathy, influence observational conditioning during learning from a learning model's UR (Olsson et al., 2016). It has also been demonstrated that avoidance learning of a CS associated with fear can be learned observationally (Cameron, Schlund, & Dymond, 2015).

In addition to learning from observation of the learning model's reaction to the US, the UR, the observer can also learn the relation between a CS and a US by observing the reaction to the CS (rather than the US), the learning model's conditioned response (CR). For instance, in the seminal work by Mineka and Cook (Susan Mineka & Cook, 1988) on observational fear conditioning in rhesus monkeys, lab-reared observer monkeys acquired persistent fear towards snakes following the observation of wild-reared model monkeys that expressed an intense fear towards a lifelike toy snake. Here, learning did not depend on the observation of the UR or on any other information of a potential US.

Similarly, research on vicarious learning in children has demonstrated fear conditioning following exposure to neutral or fear-relevant stimuli paired with pictures of faces expressing negative, rather than positive or neutral expression, without any clear link to some sort of US (Askew & Field, 2007; Gerull & Rapee, 2002).

The work on observational conditioning in children is often motivated by the need to understand the transmission of ('irrational') fear between individuals and the acquiring of

anxious behavior following observation of anxious others, such as parents (Gerull & Rapee, 2002). It is therefore reasonable to use an experimental model where the observed response is paired with the feared or anxiety-evoking event without involving further aversive stimuli such as shocks, i.e. learning by pairing the CS with a CR. Otherwise, the learnt fear would reflect a real relationship between the CS and the US, making it 'rational'. Further, since fear leads to avoidance (Krypotos, 2015) it is possible that observations (especially repeated observations) of others directly experiencing the aversive consequence of a stimulus, are less common than observations of others expressing fear to the stimulus predicting aversive consequences. This highlights the need to systematically investigate the contributions to learning from observation of others' UR and CR, respectively. A first attempt to do this is presented in **Study IV** of this thesis.

Most studies on observational learning focus on the learning of fear while less is known about observational extinction. As described earlier, research on extinction of previously acquired fear is especially important from a clinical point of view. In exposure treatment of phobias, it is sometimes common for the phobic to observe a therapist acting as a calm learning model, interacting with the feared stimulus, before the phobic is allowed to interact with it directly (Seligman & Wuyek, 2005). This clinical relevance emphasizes the need to understand the conditions under which observational extinction is efficient.

In **Study I** of this thesis we developed an experimental paradigm to study observational extinction. In this paradigm, the participant first undergoes a direct conditioning procedure followed by observational extinction. The efficiency of the extinction procedure is then tested when the CSs are presented to the participant directly. It is important that testing is done by presenting the CSs directly to the participant, both since this is more informative of how effective the extinction learning is in situations where the participant is confronted with the stimuli directly, but also because this testing of extinction should be carried out in a context different from the extinction context since change in context can lead to renewal of fear.

Extinction learning is a form of safety learning that occurs after initial fear conditioning to reduce fear. Fear conditioning can also be effectively reduced when it is preceded by safety learning, referred to as latent inhibition (Meulders et al., 2012). It has been shown that observational latent inhibition, or observational immunization, is an efficient way to prevent fear (Armita Golkar & Olsson, 2016; Susan Mineka & Cook, 1986), see also section 5.2.

1.2.4 Comparing direct and observational conditioning

If fear and safety can be learned both directly and through observation, are the underlying learning mechanisms also the same? The work by Mineka and Cook (M Cook & Mineka, 1989; Michael Cook & Mineka, 1987; S Mineka & Cook, 1993) have shown that observational conditioning share some features with classical conditioning, for instance second order conditioning (Rizley & Rescorla, 1972). However, other studies have failed to demonstrate typically Pavlovian phenomena in observational conditioning (Galef, JR. & Durlach, 1993). Nevertheless, most researchers agree that there is at least a considerable

overlap in the mechanisms involved in observational and classical conditioning (Askew & Field, 2008; C. Heyes, 2012; Susan Mineka & Cook, 1988; Olsson & Ochsner, 2008)

When discussing observational conditioning using the same framework as the one used for classical conditioning, the question arises: what is it that serves as the observer's US? Does observational conditioning require overt information about the learning model's US, for instance information that the learning model is reacting to a shock, or is the learning model's UR enough? The work by Mineka and Cook described above clearly show that in monkeys, information about the US is not necessary for observational conditioning to occur. Work in humans (Hygge, 1976) demonstrating observational conditioning in a paradigm where the learning model's US was safe for the observer, appears to point in the same direction. One possibility that is sometimes suggested is that observation of the learning model's UR evokes an aversive response and serves as the observer's US (Askew & Field, 2008). This could for instance explain the results by Mineka and Cook. However, work in children have demonstrated that images of fearful faces, used as the observed UR, resulted in successful fear learning although they did not in themselves appear to evoke fear (Askew & Field, 2007).

From a neural perspective, observational conditioning activates many of the structures involved in direct conditioning, such as the amygdala, the key structure in the brain's fear learning network, known to be involved in the processing and storing of the CS-US contingency (Olsson & Phelps, 2007). In addition, there appear to be additional involvement of the anterior insula (AI) and the anterior cingulate cortex (ACC), both being a part of processing self-experienced, nociceptive, pain and the pain of others (empathic pain) (Betti & Aglioti, 2016; Zaki, Wager, Singer, Keysers, & Gazzola, 2016).

1.3 CHOICES AND VALUES

1.3.1 Learning from reinforcement

In instrumental, or operant, conditioning, the frequency with which an individual performs a certain behavior changes as an effect of the consequences or reinforcement of that behavior. If the reinforcer which follows a certain behavior is appetitive, the frequency of the behavior increases, and if it is aversive it decreases. This simple principal, first formulated by Thorndike (Thorndike, 1911) and later developed into the theory of behaviorism by Watson (Watson, 1958) and Skinner (Skinner, 1938), is demonstrated to guide behavior in a range of animals, from the sea slug to humans. However, whereas Skinner had much of a black-box-approach to behavior, present day theories are highly concerned with the mechanisms of behavior.

The process of instrumental learning through reinforcement can be formalized computationally. This area is often referred to simply as reinforcement learning or RL. Within the general reinforcement learning framework, actions are assigned with values, representing the mean expected outcome of that action. After taking an action, these values are updated proportional to a learning rate and the difference between the actual outcome and

the expected value, the so called prediction error. The learning agent then makes a choice based on the values of the set of actions in order to maximize some utility function (Sutton & Barto, 1998). Reinforcement learning, as a way to formalize and model both learning and behavior and solve problems, is widespread and used in areas such as psychology, economics, ethology and machine learning.

One of the strengths of reinforcement learning as a framework to understand behavior is the link between reinforcement learning algorithms and neural mechanisms (Niv, 2011). This was first demonstrated in the influential paper by Schultz and colleagues (Schultz, Dayan, & Montague, 1997) where it was shown that dopaminergic neurons in the primate midbrain signaled errors in the prediction of future rewards.

Reinforcement learning can be understood as a general method for an agent to optimize behavior or to solve problems rationally (Chater, 2009). As an example, through reinforcement learning you can learn by repeated trial and error that it is faster to take the elevator rather than the stairs to get from your office to the street outside. Since you want to get out as quick as possible you take the elevator. The learning rate with which an agent learns reflects the impact a prediction error has on the updated value and thus a higher learning rate is not necessarily a better learning rate. Instead, the learning rate reflects how much the updated value relies on more recent information. This also means that for learning to be optimal, the learning rate should be sensitive to the agent's uncertainty in estimating the expected value. It has been demonstrated that humans are able to track uncertainty in an optimal manner and that they can adapt learning accordingly (Behrens, Woolrich, Walton, & Rushworth, 2007).

In the elevator versus stairs example above it seems easy to quantify the outcome or value of each choice, simply by estimating the time it took to reach the street for both choices. In laboratory or computerized experiments of human behavior it is common to use easily quantifiable reinforcers such as money or points (e.g. Niv, Edlund, Dayan, & O'Doherty, 2012). From the experimenter's point of view, quantifiable reinforcers facilitate the formalization of the learning process and it is therefore a useful tool in the investigation of both the behavioral and neural properties of learning. Of course, in many real world situations, values of actions are much more oblique, such as the value of eating a good dinner or the (negative?) value of getting fired. It is to some extent unclear how findings from research using secondary reinforcers, such as money, generalize to situations which include primary reinforcers. This is relevant for the work presented in **Study III**.

Analogous to the separately learned associations to fear and safety, learning from positive feedback and learning from negative feedback are not considered to be two sides of the same coin. Instead, positive and negative reinforcement, through the neuromodulator dopamine, separately activates the "Go" and "NoGo" pathways of the basal ganglia such that learning to take an action which leads to a reward is different than learning to avoid an action which leads to a punishment or bad outcome (Frank, Seeberger, & O'Reilly, 2004).

1.3.2 Decision making

Decisions are made based on the (learnt) expected values of a set of choices according to some decision rule (Sutton & Barto, 1998). The most straightforward rule is to behave "greedy" and exploit the current information by selecting the choice with the highest expected value. Exploitation of acquired information does however come at the cost of exploration of new information and most decision rules try to balance this trade-off. For instance, according to the "ε-greedy" method, choices are near greedy but there is a small probability ε that the agent makes a purely random choice, independent of the estimated choice values. A third method is to use the softmax selection rule, according to which the agent is most likely to select the choice with the highest expected value but where exploration is driven by the distribution of expected values of the whole choice set. In a study set out to investigate the neural substrates of exploitative and exploratory choices in humans, participants' choices where best modeled by the softmax rule (compared to ε-greedy and the softmax rule with an uncertainty bonus) and analyses of fMRI data revealed differential neural activation for choices characterized as exploitative versus exploratory (Daw, O'Doherty, Dayan, Dolan, & Seymour, 2006). Interestingly, even though choices likely reflect an agent's underlying distribution of expected values, it is difficult for an observer to extract this information based on observation of choices only, see Fig.1.

The reinforcement learning framework provides a rational account for how the values of choices are learned and expressed. However, both humans and non-human animals appear at times to violate their goals. For instance, individuals often exhibit the seemingly irrational behavior to prefer a smaller reward today to a larger reward tomorrow. This behavior can be
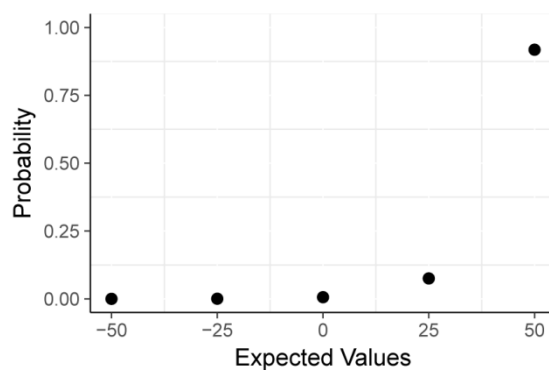
Fig.1. According to the softmax function, the probability of selecting a choice depends on the expected values of the choices under consideration, here exemplified with a set of five choices, ranging in expected value from -50 to 50 (β set to 10, see section 5.3, Eq.1). From an observer's point of view, even though it might be easy to distinguish the best choice from the rest by observing the choice frequencies, it appears hard to distinguish the worst, most dangerous choice, even from the second best.

described as a form of temporal discounting (Frederick, Loewenstein, & O'Donoghue, 2008). Individuals also tend to be risk-aversive, showing that in order to understand how humans and other animals make decisions it is important to take into consideration the probabilities of an outcome not just the expected values (i.e. magnitude × probability) (Bernoulli, 1954; Kacelnik & Bateson, 2016).

Perhaps more interestingly from a fear conditioning perspective is the interaction between instrumental and Pavlovian control systems in producing behavior, the so called Pavlovian to instrumental transfer. This interaction explains why it is harder to learn to make an action, to "go", to avoid a punishment compared with learning to "go" to earn a reward, although the two options should be equivalent using a strictly instrumental approach (Guitart-Masip et al., 2012). The Pavlovian system has also been shown to influence behavior involving sequences of decisions by removing sequences that contain large losses from consideration, even when this leads to suboptimal performance (Huys et al., 2012).

1.3.3 Observational reinforcement learning

Prediction errors, the difference between actual and expected outcome, are considered crucial for the ability to learn from reinforcement. Research on the neural and behavioral mechanisms of observational learning has shown that such prediction errors also drive observational learning (Burke, Tobler, Baddeley, & Schultz, 2010). Neural activity in the dorsolateral prefrontal cortex was shown to correlate with an observational action prediction error, signaling how much the observed choice deviated from the expected choice. Activity in the ventral striatum correlated with outcome prediction errors both during the participants' individual outcome and the observed demonstrator's outcome, although the correlation was reversed during observation. The study demonstrates that observational learning can be described within the general reinforcement learning framework, although with distinct computations related to social information. The paradigm upon which it is based has been used (with modifications) in **Study II** and **Study III** of this thesis.

Results from several studies seem to indicate that neural value computations related to decision making and learning appear to be shared between individual and observational (or social) contexts (Behrens, Hunt, Woolrich, & Rushworth, 2008; Cooper, Dunne, Furey, & O'Doherty, 2012). However, other studies appear to demonstrate the presence of distinctly social prediction errors, for instance during prediction of observed behavior (Burke et al., 2010; Suzuki et al., 2012). For a discussion and review of the topic of the differences and similarities in the neurobiology of values between social and non-social decision making, see Ruff and Fehr (2014)

1.4 SOCIAL AND CULTURAL LEARNING

1.4.1 Adaptive strategies

Historically, although the ability to learn about threats through others has a clear evolutionary advantage, the issue of the adaptiveness of fear conditioning has received little attention. One

exception is the research focusing on preparedness (Davey, 1995; Öhman & Mineka, 2001), i.e. the predisposition of certain fear relevant stimuli to be associated with fear.

In contrast, much is known about the adaptiveness of social transmission of behavior, beliefs and culture (Cavalli-Sforza & Feldman, 1981; Mesoudi, 2017; Richerson & Boyd, 2008; Tomasello et al., 1993), what we here refer to as social learning. The adaptiveness of social learning is often studied by investigating the tradeoff between individual (asocial) learning and social learning. Here, individual learners are considered to produce information or knowledge by exploration while social learners instead acquire information by exploiting, or scrounging, the information of others. Hence, individual learners are often referred to as producers and social learners as scroungers (Barnard & Sibley, 1981). Individual learning is considered accurate but slow and social learning fast but potentially less accurate or even outdated.

Theoretical work using game theory and mathematical modeling has demonstrated that the payoffs for individual versus social learning vary between different environments and situations. For instance, it is suggested that the benefits of social learning are greater in more stable environments (Feldman, Aoki, & Kumm, 1996). The benefits of social learning also depend on the proportion of individual versus social learners in a given population (Rogers, 1988). If individual learning is common, the information available in the population will be accurate and updated, increasing the benefits of social learning. If individual learning is rare, information in the population will be less accurate, decreasing the benefits of social learning. It is proposed that one solution to this problem, that social learning does not necessarily improve the fitness of a population, is to apply social learning selectively or critically (Enquist, Eriksson, & Ghirlanda, 2007; Laland, 2004). Instead of indiscriminately copying others, social learners can adopt several 'strategies' to know when and who to copy. Examples of such strategies can be to 'copy when asocial learning is costly', 'copy the majority' or 'copy successful individuals' (Laland, 2004).

Of specific interest to **Study II** and **Study III** of this thesis are the strategies or social learning heuristics referred to as payoff- or prestige-biased. These include strategies to copy others when they receive a higher payoff (Boyd & Richerson, 1985) and to copy others whose behavior is productive. The latter is usually referred to as a strategy to 'copy-successful-individuals' and can be related to both the 'indirect bias', i.e. the tendency to copy others whose behavior appears to be productive according to cues such as health or reproductive success (Boyd & Richerson, 1985), and the strategy to copy prestigious or high-ranked individuals (Henrich & Gil-White, 2001). Although it is suggested that these strategies or biases can lead to maladaptive cultural traits, the benefits appear to be relatively robust (Laland, 2004). There is empirical evidence of payoff-biased learning in both humans (McElreath et al., 2008) and wild primates (B. J. Barrett, McElreath, & Perry, 2017). However, although experimental findings from humans show that payoff-biased learning is adaptive it also appear to be underused (Acerbi, Tennie, & Mesoudi, 2016; Mesoudi, 2011). There is empirical evidence of humans adopting prestige-biased social learning strategies,

such as increased copying from others perceived as knowledgeable or successful (Henrich & Broesch, 2011) or increased copying of a demonstrator that others had previously attended positively to (Chudek, Heller, Birch, & Henrich, 2012). In primates, it has been shown that chimpanzees prefer to copy older, more high-ranked individuals compared to younger individuals with low rank (Horner, Proctor, Bonnie, Whiten, & de Waal, 2010; R. Kendal et al., 2015)

1.4.2   A few words of caution

Much of the research on social learning is grounded in quantitative theoretical work, applying mathematical models where social and behavioral phenomena are precisely and explicitly defined. Although this theoretical approach has proved to be fruitful it has been criticized for its focus on the products of learning while lacking an interest in the underlying cognitive mechanisms (C. Heyes, 2012, 2015). For instance, in discussing the mechanisms underlying the adoption of specific social learning strategies, Laland (2004, p. 5) claims that "it does not matter whether animals adopt such strategies as a consequence of evolved psychological mechanisms, learning, culture, or some combination of processes". Heyes, in her critique, argues that social and asocial learning depend on the same associative learning mechanisms and that social learning should not be viewed as something special or adaptively specialized (C. Heyes, 2012). Instead, she argues that if any mechanism is adaptively specialized in social learning it would be the input mechanisms, e.g. attentional or motivational processes, rather than specific learning mechanisms.

Some of the assumptions or conventions often used in the field of social learning fit poorly with what is known about the cognitive, psychological or neural mechanisms of social learning. For instance, according to the definition, social learning involves the transmission of information between individuals. Such transmission can occur through copying of information. However, in the literature, 'copying' is usually meant to mean copying of behavior, with the underlying assumption that observed behavior reflects an individual's knowledge. The distinction between copying of knowledge or information and copying of behavior might be of specific interest in the study of fear and avoidance learning. Here, it is apparent that copying an observed choice (even if that choice can be regarded as optimal) does not transmit information about potential dangers, even if the observed individual has that information (see also Fig. 1.).

The approach to formalize behavior and study it on the population level has yielded many insights into, and testable predictions about, the nature of social learning. However, if the underlying assumptions that the theoretical models are based on were better anchored in the knowledge we have of the mechanisms of learning, both individual and observational (social), I believe that the models would be more accurate. As an example, consider the influential study by Rendell and colleagues (L Rendell, Boyd, Cownden, Enquist, Eriksson, Feldman, Ghirlanda, et al., 2010), often used as an example of the surprising efficiency of the copying of others' behavior. The study is based on a computer tournament where entrants could specify strategies for agents to learn and behave in a complex environment. Even

though this is an innovative approach to gain new ideas, I believe the assumptions built into the structure of the tournament make the results uninformative as to why, and when, copying is successful. For instance, agents were only allowed to copy the behavior of another agent if that agent was currently exploiting the information it had previously gained. As we have discussed earlier (see section 1.3.2), it is very difficult for an observer to differentiate between an observed exploratory or exploitative decision. And, if we assume that agents cannot learn from exploratory behavior this might give copying-heavy strategies an advantage in this type of game.

The work presented in this thesis is grounded in classical conditioning and reinforcement learning but has to some extent also been inspired by theories from the field of social learning. Although these research areas differ they are all concerned with understanding how individuals are able to extract valuable information about the statistics of their environments. This has also been my concern while writing this thesis.

2 AIMS

The overarching goal of this thesis is to deepen our understanding of learning about fear and safety through observation of others (observational learning), and how this learning is shaped by observed behaviors and experiences of the other as well as the inferred competence of that individual. For each of the four studies of this thesis we formulated a more detailed aim:

- To investigate if safety learning is facilitated through the observation of a calm learning model (**Study I**).

- To extend the literature on observational instrumental learning into the aversive domain using primary reinforcers and to investigate if and how observational learning is modulated by the skill (or ability) of a demonstrator (**Study II**).

- To investigate how the description of a demonstrator's ability affects observational learning and how it interacts with the actual ability of the demonstrator (**Study III**).

- To investigate how observational learning differs as a function of a learning model's expressed anticipatory anxiety (**Study IV**)

3 METHODS

3.1 PARTICIPANTS

For the studies in this thesis we recruited a total of 239 self-reportedly healthy participants. All participants signed an informed consent prior to participation and were payed for their participation at the end of their experiment.

3.2 SKIN CONDUCTANCE

Electrodermal responses have been used as psychophysiological measures since the late 19th century (Neumann & Blanton, 1970). The skin conductance response, SCR, captures the phasic component of skin conductance, and is often treated as independent from the tonic component, the skin conductance level (SCL) (Lim et al., 1997). The SCR reflects event-related sympathetic arousal (through activation of the sweat glands) and is used as a non-invasive measure of emotional arousal and attention (Boucsein, 2012).

In the conditioning literature, SCR is a used as a measure of an acquired fear, or threat, response (Öhman & Mineka, 2001). Learning is measured either as the change in SCR over time, for instance as an increase in the SCR to a CS following pairing with a shock, or as a differential change in SCR in discriminatory paradigms, i.e. larger responses over time to the CS+ compared with the CS-. The SCR has also been used in studies of decision-making, for instance to capture arousal responses related to anticipation (Dawson, Schell, & Courtney, 2011). SCR was used as a measure of fear in **Study I** and to capture effects of surprise and attention in **Study III**.

3.3 FEAR-POTENTIATED STARTLE

The startle reflex, a defensive response to threatening sudden stimuli, is a reflex in both humans and non-human animals which was well documented already in the 1930s (Landis & Hunt, 1939) and it has been used as a psychophysiological measure to study emotional processes in humans for almost three decades (Vrana, Spence, & Lang, 1988).

Experienced fear increases the magnitude of the startle response and this fear-potentiated startle response, FPS, is besides SCR the most common used measure of conditioned fear. In comparison to SCR, which captures effects of both emotion and attention (Filion, Dawson, Schell, & Hazlett, 1991), FPS is considered less of a measure of attention and instead more of a measure of emotion or emotional valence (Lang, Bradley, & Cuthbert, 1998). Further, the startle reflex can be linked to the cross-species threat response system (Grillon & Baas, 2003) which allows it to be used as a translational tool. FPS is a non-event-related measure that can be elicited both during and in between presentations of discrete stimuli, allowing investigation of baseline reactivity and context conditioning. This is especially important during studying of anxiety disorders, since anxiousness is often described as a form of contextual, rather than cue-specific, fear (Grillon, 2002b).

In humans, FPS is often measured as the magnitude of the eyeblink reflex, elicited by a brief, rapid, intense stimulus, e.g. a short burst of noise. FPS has been used in **Study IV** of this thesis.

3.4 EYE TRACKING

The method to track the gaze of the eye was first used to study reading. In the 1950s, Yarbus showed in his influential work that gaze trajectories depend on the nature of the task during viewing (Yarbus, 1967). His work made eye tracking an important tool for psychological and cognitive research. Shortly after, Hess and Polt (Hess & Polt, 1960) showed that in humans, pupil size varies as a function of the 'interest value' of a visual stimulus, adding pupil size as a measure of interest to psychological research.

Most modern-day eye trackers used in psychological laboratory experiments employ a method using light, often infrared light, which is reflected from the eye and then recorded. This non-invasive method allows recording of both gaze and pupil size.

During analyses of gaze patterns, location and length of fixations as well saccades and blinks are of interest. The direction of gaze is often used as a measure of visual attention (e.g. Wadlinger & Isaacowitz, 2006) although it is clear that visual attention does not necessarily overlap with gaze direction (Brefczynski & DeYoe, 1999). The size of the pupil changes as a response to both lightning conditions and cognitively relevant stimuli, and can either dilate or contract. Increases in pupil size have been demonstrated following for instance presentation of emotionally arousing stimuli (Bradley, Miccoli, Escrig, & Lang, 2008) and during increased executive or working memory load (Chatham, Frank, & Munakata, 2009; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004). It has been proposed that the size of the pupil is under control of the locus coerulus (Joshi, Li, Kalwani, & Gold, 2016), suggested to mediate the brain's attentional network and that the pupil size therefore can be used as an index of attention (Laeng, Sirois, & Gredeback, 2012). Eye tracking has been used in **Study III**, to assess gaze location and pupil size.

3.5 REINFORCEMENT LEARNING

Reinforcement learning, RL, studies how agents (natural or artificial) learn to make actions to maximize reward (Sutton & Barto, 1998). Actions can lead to a change of the state of the agent or to a reward or punishment. RL is used in many fields, ranging from game theory (Erev & Roth, 1998) to robotics (Kober, Bagnell, & Peters, 2013).

RL modeling can be used to predict and describe behavior with great accuracy and can therefore be used as a tool to analyze data. Especially, RL allows modeling of reinforced behavior on a fine-grained trial-by-trial level rather than more standard aggregated approaches (Daw, 2011). An additional strength with the RL framework is its link to the neurobiology of decision making (Daw & Doya, 2006; Lee, Seo, & Jung, 2012), which took its starting point in the finding of a link between RL error signals and the firing of dopamine neurons in the primate midbrain (Schultz et al., 1997). Linking computational models of

learning to both neural and behavioral data aides our understanding of the mechanisms underlying cognition.

In this thesis, RL modelling has been used to analyze data in **Study II** and **Study III**. The RL framework that we have used is based on the Q-learning technique (Watkins & Dayan, 1992), applied to a simple two-armed bandit task. Q-learning is a model-free form of reinforcement learning where an agent uses the consequences of its actions to update the expected values of a (finite) set of actions.

4 SUMMARY OF THE STUDIES

4.1 STUDY I

4.1.1 Background & Aims

Acquiring fear through social forms of learning is believed to be an important form of learning about dangers (Askew & Field, 2008; Olsson & Phelps, 2007). Extinction, the attenuation of fear to something that was previously dangerous but that is now safe, through social means, can be considered equally important. For instance, treatment of phobic individuals often includes exposure treatment where the therapist calmly interacts with the feared stimulus, a form of vicarious attenuation or extinction of fear (Seligman & Wuyek, 2005). Although there is a rich literature focusing on extinction through direct exposure (Milad & Quirk, 2012; Myers & Davis, 2007; Phelps, Delgado, Nearing, & Ledoux, 2004) research on social forms of safety learning is sparse and methodologically limited (Bandura, Grusec, & Menlove, 1967; Hill, Liebert, & Mott, 1968; Ritter, 1968; Seligman & Wuyek, 2005). In **Study I,** we aimed to bring together research on direct extinction learning with the work focusing on socially acquired fear to investigate if extinction of fear can be facilitated through observation of a calm learning model undergoing extinction learning. We predicted that such vicarious extinction, compared to direct extinction, would facilitate the attenuation of expressed fear as measured in a subsequent test.

4.1.2 Method

We recruited 49 male participants, randomized into two groups (the direct-extinction group and the vicarious-extinction group), to take part in the study. Participants underwent an extinction learning paradigm where they were conditioned to two CSs consisting of pictures of angry male faces from the Karolinska Directed Emotional Faces, KDEF, database (Lundqvist, Flykt, & Öhman, 1998) displayed for 6 s. Following initial habituation (each CS presented twice), participants were exposed to a direct acquisition stage where each CS was presented 9 times. The CS+ co-terminated with the administration of a shock (100 ms DC-pulse electric stimulation, administered using the STM200, BIOPAC Systems, CA) on 6 of these presentations, while the CS- was never reinforced. Next, participants underwent extinction by observing two videos, one for each group, showing a screen on which the CSs were presented 6 times each. The direct-extinction group observed a video displaying the screen only, preceded by information that they were to observe a movie described as depicting an experiment similar to the one they had been doing. The vicarious-extinction group observed an identical video, except for the presence of a calm learning model placed in front of the screen watching the display, preceded by information that they were to observe a movie of a person participating in an experiment similar to the one they had participated in. During extinction, no CS was paired with a shock. To test the subsequent expression of fear, participants underwent a reinstatement test stage (Hermans et al., 2005) which began by administering 3 shocks to the participants while they observed a black screen for 30 s. Next, the CSs were presented 6 times each without shock directly on the screen.

To assess fear responses, we measured the skin conductance response (SCR) to each CS using a pair of Ag-AgCl electrodes attached to the middle and index finger of the left hand using BIOPAC Systems (Santa Barbara, CA). Data were analyzed using AcqKnowledge software (BIOPAC Systems).

4.1.3 Results & Conclusion

Our analyses of SCR responses showed that during acquisition, participants expressed successful differentiation, with higher responses to the CS+ compared to the CS-, but the groups did not differ. However, during the extinction stage the direct-extinction group still differentiated between the CSs while the vicarious-extinction did not. In the subsequent reinstatement test stage the direct-extinction group showed recovery of fear while the vicarious-extinction group extinguished more efficiently and did not differentiate between the CS+ and CS-.

To confirm that the efficiency of observing the calm learning model in the vicarious-extinction group did not depend on simply the presence of the learning model but rather the expressed experience of the learning model we compared the vicarious-extinction group with a new group, vicarious-reinforcement, of 19 males that underwent the same extinction paradigm as the vicarious-extinction group but where the learning model during the extinction stage received four shocks to CS+ and reacted to them by twitching the arm and blinking. As before, the participants did not receive any shocks during extinction. The vicarious-reinforcement group did not differ from the vicarious-extinction group during acquisition but marginally so during extinction. In the reinstatement test stage, the vicarious-reinstatement group differentiated significantly more between the CSs than the vicarious-extinction group showing that the efficiency of extinction in the vicarious-extinction stage was dependent on the experience rather than mere presence of the learning model.

4.2 STUDY II

4.2.1 Background & Aims

Learning to make an action or choice through observation of someone else, here referred to as observational learning, is often more efficient than individual learning (Feldman et al., 1996; Merlo & Schotter, 2003; L Rendell, Boyd, Cownden, Enquist, Eriksson, Feldman, Fogarty, et al., 2010). It has further been theorized that social forms of learning are especially valuable when consequences are costly (R. L. Kendal, 2004; Webster & Laland, 2008). Despite this, research on human observational learning has focused almost exclusively on the appetitive domain (Apesteguia, Huck, & Oechssler, 2007; McElreath et al., 2008), often using secondary reinforcers such as money (Burke, Tobler, Baddeley, & Schultz, 2010; Nicolle, Symmonds, & Dolan, 2011; Suzuki et al., 2012), in contrast to research on observational associative learning which often uses primary aversive reinforcers, such as shocks (Olsson, Nearing, & Phelps, 2007). The primary aim of **Study II** was to extend the literature on observational learning into the aversive domain using primary reinforcers.

Even though observational learning is efficient, observational information can be outdated or unreliable. For instance, learning from observing only the choices (not the consequences of the choices) of someone, a demonstrator, who is unskilled and makes poor choices is likely inefficient. However, the efficiency of learning by pairing a demonstrator's choice with its consequence is less influenced by the demonstrator's skill level. It has been proposed that observational learning should be applied critically (Enquist et al., 2007; Rachel L. Kendal, Coolen, van Bergen, & Laland, 2005) for instance by adhering to strategies such as "copy successful individuals" (Laland, 2004). There are studies showing that copying is more common in humans when the demonstrator is successful (Apesteguia et al., 2007; Mesoudi, 2008; Morgan, Rendell, Ehn, Hoppitt, & Laland, 2012), but these studies all provide the observer with very explicit information of the demonstrator's performance level. The second aim of **Study II** was thus to investigate if and how observational learning is modulated by the skill level of the demonstrator, without giving the observer explicit information about the demonstrator's performance level. To better understand the effect of the demonstrator's skill level we also varied the amount of observable information available to the observer.

4.2.2 Method

We recruited 42 participants who learned a probabilistic two-choice task to avoid shock. They learned the task both directly and through the observation of a demonstrator. The experimental design was based on the paradigm by Burke and colleagues (2010). To each participant, the demonstrator was presented as another participant but was in fact a sex-matched confederate. The behaviors of the demonstrators were controlled by a computerized learning algorithm. Participants were randomly assigned to either a group that observed a skilled demonstrator (SD group) that learned the task quickly and performed well or a randomly behaving, unskilled, demonstrator (UD group). The amount of available observable information varied between pairs of choice-stimuli according to three observational learning

conditions: 1) individual learning (No Observation), 2) observable information of the demonstrator's choices (Choice Observation), 3) observable information of both the demonstrator's choices and the consequences of those choices (Choice-Consequence Observation). At each trial during the Choice Observation condition and the Choice-Consequence condition participants saw which choice the demonstrator made. During the Choice-Consequence condition, participants could also observe a symbol indicating if the demonstrator received a shock or not. During the No Observation condition, participants received no information regarding the demonstrator's behavior. Importantly, no information regarding the skill or performance of the demonstrator was given.

In the experiment, participants learned to make choices between pairs of picture stimuli (randomly generated fractals). Within each pair, one stimulus was associated with a low probability of being followed by a shock (20%, optimal choice) and the other with a high probability (80%, suboptimal choice). The shock consisted of a 100 ms DC-pulse (STM200, Biopac Systems) individually calibrated to be uncomfortable but not painful. At each trial, participants first saw the demonstrator make a choice (depending on the observational learning condition) before they made a choice themselves. Each pair of choice stimuli was presented for 15 trials. Trials were divided into three blocks, each consisting of three pairs of choice stimuli, one for each observational learning condition.

Data were analyzed using both statistical methods and reinforcement learning modeling.

4.2.3 Results & Conclusions

Our results showed that observing a skilled, compared to an unskilled demonstrator increased performance in the Choice Observation condition but had no effect on performance in the No Observation condition or the Choice-Consequence condition. In the Choice Observation condition, participants observing a skilled demonstrator imitated or copied the demonstrator's choices more than those observing an unskilled demonstrator and did so increasingly over time. In the Choice-Consequence condition RL analyses showed that participants did not appear to imitate the choices of neither the skill nor unskilled demonstrator (at least not to any significant extent) but that they instead learned by associating the observed choice with its consequence. We concluded that participants in both groups used observational information to improve performance compared to individual learning when observational information was valuable (i.e. not during observation of an unskilled, randomly behaving, demonstrator in the Choice Observation condition). Further, the increase in imitation of the skilled demonstrator over the course of the experiment suggested that participants over time learned the value of observing the skilled demonstrators' choices only.

4.3 STUDY III

4.3.1 Background & Aims

Learning to avoid dangers through observation of others is relatively safe and efficient. When learning by observing and copying the choices of others, performance is influenced by the ability of the observed individual, the demonstrator. The higher ability the demonstrator has, the more efficient it is to copy the behavior of the demonstrator. This is reflected in proposed efficiency of strategies such as payoff and prestige biased copying (Henrich & Gil-White, 2001; Laland, 2004; Luke Rendell et al., 2011). However, when learning by observing both the demonstrator's choice and the outcome of that choice, the ability of the demonstrator is of less importance since we can learn from both positive and negative consequences (for a more detailed discussion, see section 5.3).

Previous research has shown that people show a preference for attending to individuals perceived as high in ability (Bruin & Lange, 2015). This has also been shown in non-human animals (R. Kendal et al., 2015). Research has suggested that this attentional bias could account for persistence of group stereotypes (Denrell, 2005) and poor organizational decisions (Denrell, 2003) simply by biasing sampling of information.

In **Study II** we showed that the ability, or skill, of a demonstrator did not affect the efficiency of observational learning in a simple two-choice task to avoid shocks when both the demonstrator's choices and the consequences of those choices where available and when no explicit information about the demonstrator's ability was given. The aim of **Study III** was to extend the findings from **Study II** and investigate how the description of a demonstrator's ability affects observational learning and how it interacts with the actual ability of the demonstrator. We hypothesized that describing a demonstrator's ability as low rather than high would impair observational learning by reducing attention towards that demonstrator. We further suggested that described ability might interact with actual ability such that the effect of described ability would be greater during observation of a demonstrator that had an actual low ability, since learning from a poor performing demonstrator is a more cognitively demanding task.

4.3.2 Method

The experimental paradigm was based on the design used in **Study II** during the observational learning condition where both the observers' choices and the consequences of those choices were available.

We recruited 46 participants who each observed two demonstrators, described as previous participants in a similar study but unbeknownst to the participants controlled by a simple computerized learning algorithm. The demonstrators made repeated choices in a simple two-choice task to avoid shock. After each observed choice and outcome, it was the participant´s turn to make a choice. Simple geometric figures with equal luminance were used as choice stimuli. For each pair of choice stimuli, one stimulus became the optimal choice and the other

became the suboptimal, associated with a probability of 20% and 80%, respectively, of being followed by a shock. The shock to the participant consisted of a 100 ms DC-pulse (STM200; Biopac Systems Inc), individually calibrated to be uncomfortable but not painful. When the demonstrator received a shock, this was signaled to the observer/participant by a short neutral sound presented in the participant's headphones.

To vary the described ability of the demonstrator between participants, they were randomly assigned to one of two groups; the Described-High group, where participants were told the two demonstrators had performed well, or the Described-Low group, where participants were told the two demonstrators had performed poorly. To vary the actual ability within participant we controlled the choices of the demonstrators so that one performed well (Actual-High) and the other performed at chance (Actual-Low).

To measure effects on attention we collected eye tracking data with a resolution of 50 Hz through iViewX 1.6 using an SMI remote Red II eye tracker placed in front of the participants. In the eye tracking data, we were specifically interested in the pupil dilation response since this is commonly used as an index of surprise, attention and cognitive load (Beatty & Lucero-Wagoner, 2000; Laeng et al., 2012; O'Reilly et al., 2013). To assess psychophysiological effects of learning we also collected skin conductance data, SCR with BIOPAC Systems (Santa Barbara, CA) using a pair of Ag-AgCl electrodes attached to the index and middle finger of the left hand. The SCR is commonly used in conditioning paradigms (Öhman & Mineka, 2001) and also as an index of attentional processes (Dawson et al., 2011).

In addition to behavioral analyses, we used reinforcement learning (RL) modelling to try to capture and describe the learning process.

4.3.3 Results & Conclusions

Behavioral results showed that participants in the Described-Low group performed worse than participants in the Described-High group, in line with our hypothesis. There was no main effect of actual ability or any interaction between actual ability and described ability.

Analyses of preparatory pupil dilation responses just before the presentation of the demonstrator's choice revealed smaller pupil dilation responses before the choice in the Described-Low group compared to the Described-High group. These results indicate that participants in the Described-Low group paid less attention to the demonstrator's choices than the Described-High group. This was further corroborated by data from ratings provided by the participants of the number of shock administered to the demonstrator during each block. These data showed that the Described-Low group made more mistakes in their ratings than the Described-High group.

RL modelling suggests that the effect of the described ability could be attributed to an impaired ability to integrate observational information over time.

4.4   STUDY IV

4.4.1   Background & Aims

Fear can be acquired by observing other individual's reactions and responses. Such learning can be efficient (S Mineka & Cook, 1993) but it is also believed that observational learning can be a cause behind the development of anxiety disorders (Rachman, 1977). Reactions of others that we observe and learn from can be very direct, such as the response to feeling pain, or more indirect, such as the anticipation of pain. Anticipatory fear reactions are likely less reliable cues about what constitutes a real threat and not, especially since individuals vary in their reactions to safe and unsafe stimuli. For instance, both clinically (Grillon, 2002a; Jovanovic, Kazama, Bachevalier, & Davis, 2012; Schmuel Lissek et al., 2010; Shmuel Lissek et al., 2005) and non-clinically (Gazendam, Kamphuis, & Kindt, 2013; Lenaert et al., 2014) anxious individuals show reduced discriminatory learning, often driven by heightened fear responses to safe stimuli. The main aim of **Study IV** was to investigate how observational learning differs as a function of expressed anxiety in an observed other, the learning model. We hypothesized that learning through observation of a typically anxious individual, expressing anticipatory fear to both safe and unsafe stimuli, would reduce discriminatory fear learning. The second aim of **Study IV** was to develop an observational conditioning paradigm which used sound as the aversive stimuli presented to the learning model.

4.4.2   Method

For **Study IV**, we recruited eighty-three participants to take part in an observational fear conditioning paradigm based on previous work in our lab (Haaker, Golkar, Selbing, & Olsson, 2017). Participants were randomized into two groups, the Anxious Model group and the Non-Anxious Model group.

Participants underwent an experiment where they were observationally conditioned to two colored squares, the conditioned stimuli (CSs). Conditioning took place following an initial pre-acquisition phase where participants were presented with the CSs directly (4 presentations each). During observational conditioning, participants observed a video depicting a learning model who was placed in front of a screen on which the CSs were presented. On 9 out of 12 presentations, the CS+ was followed by a 6s long aversive sound, the US, taken from the International Affective Digital Sounds (IADS) collection (Bradley & Lang, 2007). The US was presented in the learning model's headphones but was not available to the observer (the participant). The CS- was never paired with the US. The videos for the two groups differed in when the learning model expressed anticipatory fear to the CSs (see Fig.2.). To mimic the reduced discriminatory learning typical in anxious individuals, we let the learning model in the Anxious Model group express anticipatory fear during presentation of both the CS+ and CS-, following the first US presentation. To mimic successful discriminatory learning, the learning model in the Non-Anxious Model group expressed anticipatory fear to the CS+ only, following the first US presentation. In both groups, the learning model reacted aversively to the US. After observational conditioning, learning was
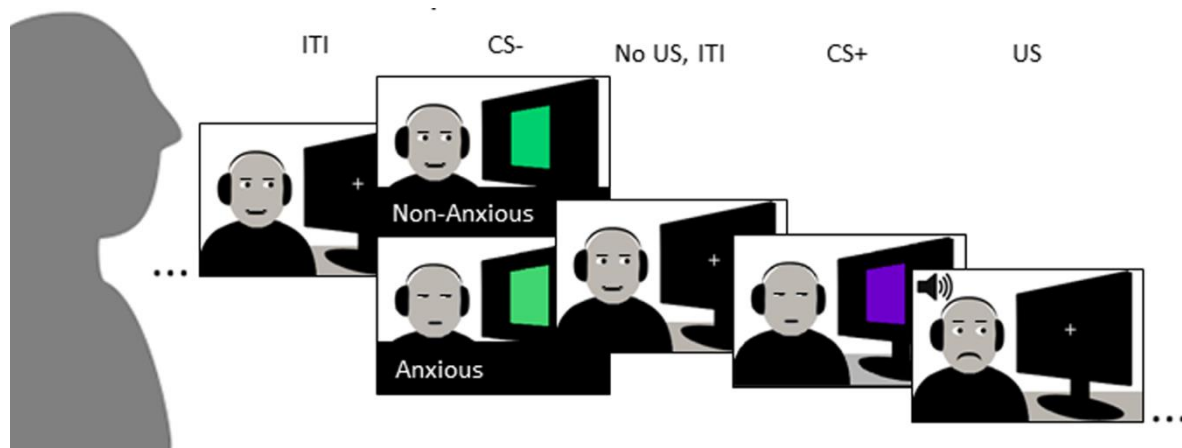
Fig.2. In the video shown during the observational acquisition, the learning model expressed anticipatory fear to the safe stimulus, CS-, only in the video shown to the Anxious Model group. The learning model expressed anticipatory fear to the CS+ and reacted aversively to the US in the videos shown to both groups.

tested in a direct test phase where the CS+ and CS- were presented directly to the participant. A novel CS, CSn, was introduced to assess possible effects on generalization learning. Each CS was presented 12 times.

To assess the conditioned fear responses, we measured FPS on half of the presentations of each CS and on corresponding inter-trial-intervals, ITIs. Participants were also asked to use a slider to constantly rate their expectation that the aversive sound, the US, would soon be presented (either to themselves during pre-acquisition and the direct test phase or to the learning model during observational acquisition). Ratings were continuous and ranged from "not at all likely" to "very likely". Expectancy ratings for the CSs where then measured by calculating the mean rating for the CS presentations for which FPS was not measured. CS presentations, startle administration and expectancy ratings were ordered in blocks such that in a block, each type of CS is presented twice to collect both FPS and expectancy ratings. Within each block, FPS for one ITI is collected.

4.4.3 Results & Conclusions

Data from both the expectancy ratings and FPS showed that participants in both groups learned to discriminate between the CS+ and the CS-. However, there was no support for our hypothesis that learning through observation of a typically anxious individual would reduce discriminatory fear learning. The Anxious Model group discriminated between the CS+ and the CS- at least as well as the Non-Anxious group. In fact, contrary to our hypothesis, results from analyses of the expectancy ratings showed a reduced extinction rate during testing in the Anxious Model group compared to the Non-Anxious Model group. There were no group differences with regards to generalization. As for the second aim of **Study IV,** to develop an observational conditioning paradigm which used sound as the aversive stimuli, our results showed that the paradigm served this purpose.

5 DISCUSSION

5.1 SUMMARY OF THE RESULTS

The studies presented in this thesis all deal with the impact that the behavior of the person learnt from has on observational learning. In addition, in **Study I** and **Study II** we also contrasted observational learning with individual learning. Importantly, all studies used healthy participants and the variations in behavior of the persons learnt from are variations that we can expect to encounter in real life. We have demonstrated that observational fear extinction is more efficient than direct extinction (**Study I**) and that individuals can adapt their use of observational information during avoidance learning, for instance by not copying other individual's poor choices (**Study II** and **Study III**). We have further shown that a description of those we learn from can bias observational avoidance learning, in such a way that it becomes less efficient during observation of someone described as low in ability (**Study III**). Finally, we have shown that expressed anxiety of the person learnt from does not have any substantial effect on observational learning of fear (**Study IV**).

5.2 UNDERSTANDING OBSERVATIONAL SAFETY LEARNING

One of the more interesting findings presented in this thesis in the result from **Study I.** We found that the observational (vicarious) extinction procedure, compared to the direct extinction, resulted in more extinction and blocked the return of fear. These results have implications for clinical practices by suggesting that model-based exposure could be an efficient form of treatment of fears and phobias. Yet, our study, as well as other previous and later ones, did not provide an explanation as to why observational extinction was *more* efficient than direct extinction. For instance, in a later study conducted in our lab using the same type of video stimuli as in **Study I**, it was demonstrated that observational, compared to direct, safety learning enhanced the prevention of subsequent observational fear learning (Armita Golkar & Olsson, 2016). This form of 'immunization' through observational pre-exposure to safe models has also previously been shown to reduce subsequent fear learning more efficiently than direct pre-exposure (Susan Mineka & Cook, 1986; Poser & King, 1975).

One explanation to the increased efficiency of observational, compared to direct, safety learning is provided by Mineka and Cook (Susan Mineka & Cook, 1986) who suggest that observation of a safe model leads to more active safety learning compared to direct exposure. Of course, pre-exposure to non-reinforced stimuli is not the same as extinction learning. Still, the suggestion that observational extinction might be a more active form of safety learning could explain our results. A second possibility could be that observational safety learning promotes more efficient learning by providing additional safety information compared to direct safety learning. This would fit well with the findings in **Study I** where the observational and the direct condition differed by the added presence of a learning model in the observational condition. It might also explain the findings from Mineka and Cook. A third possibility could be that our implementation of the direct extinction was interpreted

differently than a more standard direct extinction procedure. To keep everything but the presence of the calm learning model constant between groups, participants in the direct extinction condition observed a video showing a screen slightly turned to the side. Even though this meant that participants did observe the CSs in the absence of receiving shocks it is possible that they interpreted the video as if no one was presented with the CSs. This would make the relationship between the CSs and the lack of shocks less pronounced and extinction learning possibly less efficient. To summarize, even though there exist some evidence that observational safety learning might be more efficient than direct safety learning, it is unclear both why this is the case and how well these results translate to real life situations. Further research is needed to bring clarity to these issues.

In **Study I**, we also demonstrated that the efficiency of observational extinction depended on the experience of the learning model rather than just the presence of the learning model per se. Extinction learning through observation of a calm learning model was more efficient than learning through observation of a learning model that received, and reacted to, shocks. We have since replicated these findings in a study designed to explore the neural circuitry underlying observational extinction using functional magnetic resonance imaging (fMRI) (Armita Golkar, Haaker, Selbing, & Olsson, 2016). Here, we used a within-subject design where the participants were exposed to three CSs: $CS+_{vic\ reinf}$, $CS+_{vic\ safety}$ and CS-. CS- was never paired with the US while both $CS+_{vic\ reinf}$ and $CS+_{vic\ safety}$ were paired with the US during an initial acquisition phase. During extinction, participants observed a learning model that was presented with the CSs. Here, the learning model experienced and reacted to the US following $CS+_{vic\ reinf}$ but not $CS+_{vic\ safety}$ or CS-. Return of fear was efficiently attenuated for $CS+_{vic\ safety}$ but not $CS+_{vic\ reinf}$, again showing the efficiency of observational extinction. Analysis of the BOLD (blood-oxygen-level dependent) response during extinction revealed a higher activity in the ventromedial prefrontal cortex (vmPFC) to the $CS+_{vic\ reinf}$ and a decrease in activity to the $CS+_{vic\ safety}$. This pattern was later reversed in the reinstatement test where activity in the vmPFC was higher for the $CS+_{vic\ safety}$ and lower for the $CS+_{vic\ reinf}$. Although an increase in vmPFC activity during extinction procedures is common (Milad et al., 2007; Phelps et al., 2004), explained as an involvement by the vmPFC in suppression of fear, it is less clear why we observed a decrease in activity to the $CS+_{vic\ safety}$ during extinction learning. We suggested that observational extinction learning might have somehow reduced the need for vmPFC involvement during extinction. Further research is, however, needed to more clearly understand the neural mechanisms underlying observational extinction.

In an attempt to understand the influence of who we learn from, later research has demonstrated that the efficiency of observational safety (as well as fear) learning depends on the group belonging of the learning model (A. Golkar et al., 2015). Both observational fear and safety learning were facilitated when learning from a learning model who belonged to the observer's in-group rather than the out-group.

5.3 PRESTIGE BIAS REVISITED

In **Study III**, where participants performed a simple two-choice task which included observation of a demonstrators that were described as either high or low in ability, pupil dilation and behavioral measures indicated that more attention was payed to the demonstrator described as high in ability. We interpret this as an effect of participants estimating that the value of observational information is higher when a demonstrator has a high ability. Participants would subsequently pay more attention or exert more effort while observing the demonstrator that is described as high in ability, in line with theories of attention as a utility maximizing system that mediates search for information (Gottlieb, 2012). However, results from both **Study II** (when the demonstrator's choice and outcome was observed) and **Study III** showed no effect of the demonstrator's ability on performance. Thus, from a performance perspective, the value of information in this type of task does not depend on the demonstrator's ability, indicating that the difference in attention between groups does not reflect any real difference in informational value. Participants in **Study III** appeared to wrongfully estimate that the value of informational is higher when a demonstrator has a high, rather than low ability. Therefore, we suggested that participants attended more to the demonstrators described as high in ability since high ability is argued to be more informative in a general sense (Martijn, Spears, Van Der Pligt, & Jakobs, 1992). We also suggested that the participants' attentional bias and subsequent performance could be explained by a prestige or payoff biased heuristic (Henrich & Gil-White, 2001; Laland, 2004) which can be an efficient strategy for copying of behavior. In addition, we also briefly discussed the possible explanation that the learning bias demonstrated in **Study III** could be more beneficial in other, more complex, tasks. To investigate the efficiency of this learning bias in other tasks and situations, we ran a series of simulations, separate from the study. In these simulations we varied not just the knowledge of the demonstrator but also the task size and the observer's initial expectations of the values of the choices since these factors often vary in real world tasks. We were thus able to capture at least some of the normally occurring variability that learning agents face. Results from these simulations are briefly presented here (for details, see Selbing & Olsson, 2017).

To explore how the ability of the demonstrator influences the efficiency of observational learning in different situations we performed a set of simulations. In these simulations, agents, referred to as observers, learned from demonstrators with varying knowledge ("Full", "No" and "Some" knowledge) of each task's reward structure in probabilistic n-choice tasks of varying size (n = 2,3 and 10). In addition, to explore the effect of the observer's initial expectations of the task's reward structure, we also varied the initial expected values of the choices in the task ("Mean", "Low" and "High" initial expectations). Simulations were based on the Q-learning algorithm (Watkins & Dayan, 1992). The demonstrators made choices according to the softmax function, which calculates the probability p of selecting each choice i at every time step t, given the expected values Q of all n choices and the inverse temperature parameter β (set to 0.5), controlling the tendency to either explore or exploit:

$$p_i(t) = e^{Q_i(t)/\beta} / \sum_{k=1}^{n} e^{Q_k(t)/\beta} \qquad \text{Eq.1}$$

Observers learned from observing both the choices and subsequent outcomes of the demonstrators. They learned according to a simple learning rule where the observer at each time step t updates its expected value Q_n^{obs} for the demonstrator's selected choice s by the observational prediction error (the difference between observed and expected outcome) multiplied by the learning rate α (set to 0.3):

$$Q_s^{obs}(t+1) = Q_s^{obs}(t) + \alpha * (outcome^{dem} - Q_s^{obs}(t)) \qquad \text{Eq.2}$$

Observers made choices based on their expected values Q_n^{obs} according to the softmax function in Eq. 1, but did not learn from their own choices (in order to study the efficiency of observational learning in isolation). The knowledge levels of the demonstrators were fixed at the start of the simulation and from the on demonstrators did not learn.

The learning tasks were probabilistic n-choice task of sizes 2, 3 or 10, where the outcome of each choice was drawn from a normal distribution with mean m_n, initialized for each simulation round. Demonstrators' knowledge of the task was varied by varying the demonstrators' expected values, Q_n^{obs}, of the choices. In simulations where the demonstrators had "Full" knowledge, the expected values, Q_n^{obs}, corresponded exactly to the means of each choice outcome distribution m_n. In simulations where the demonstrators had "No" knowledge, the expected values were all set to the average outcome of all choices, $\overline{m_n}$. In simulations where the demonstrators had "Some" knowledge, the expected values were less accurate compared to "Full" knowledge but more accurate compared to "No" knowledge. This resulted in high levels of performance for the demonstrators with "Full" knowledge, relatively good performance for the demonstrators with "Some" knowledge and very poor performance (random choices) for the demonstrators with "No" knowledge (see also Fig. 3.). The observers' initial expectations of choice values were varied by varying $Q_n^{obs}(t)$ at $t = 0$, before learning had taken place. When observers' expectations were at "Mean", the initial expected values were set to the mean overall outcome, $Q_n^{obs}(t = 0) = \overline{m_n}$. When observers' expectations were "High", the initial expected values were set to the mean outcome of the optimal choice, $Q_n^{obs}(t = 0) = \max(m_n)$. When observers' expectations were at "Low", the initial expected values were set to the mean outcome of the most suboptimal choice, $Q_n^{obs}(t = 0) = \min(m_n)$.

Results from the simulations show that the efficiency of learning a probabilistic n-choice task through observation of a demonstrator depends on an interaction between the demonstrator's knowledge of the task's reward structure, the task size (n) and the observer's initial expected choice values (see Fig.3.). Learning through observation of a high performing demonstrator is beneficial for tasks sizes larger than two ($n > 2$) when the observer's initial expected choice values are "Low" or "Mean".
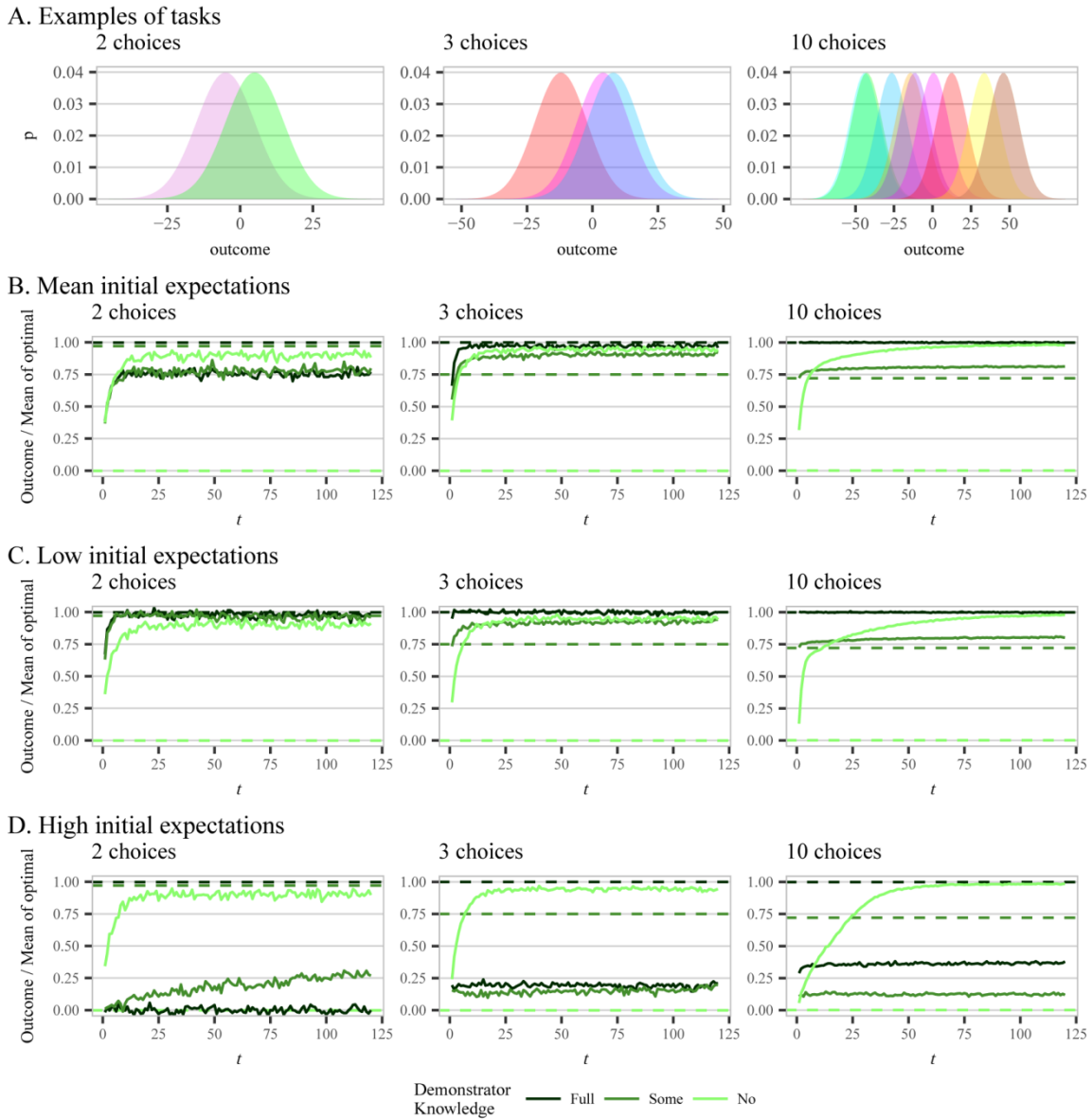
Fig.3. A. Examples of the probability distributions of outcomes for n-choice tasks of varying size n (2, 3 and 10). **B-D.** Performance plotted as the mean outcome per t over 10 000 simulation runs of tasks of varying size, n, following observation of a demonstrator with varying levels of knowledge. Dashed lines show mean demonstrator performance over all timepoints (demonstrators do not learn and their performance does not change over time). **B.** Mean initial expectations of choice values. **C.** Low initial expectations of choice values. **D.** High initial expectations of choice values.

Learning through observation of a poor performing demonstrator is especially beneficial in cases where the observers' initial expectations of choice values are high. To understand these results, we need to consider how the demonstrators function as explorers and what information the observer needs to make a good decision. A poor performing demonstrator will always explore the task environment more than a high performer since the poor performer behaves randomly and the high performer's choices are biased. Thus, learning

from observation of a poor performer will faster lead to the observer having a more correct view of the contingencies of the whole task. However, in order to make a good decision, it is more important to be able to differentiate the most optimal choice(s) from less optimal rather than to have correct expectations of all choice values, including the poor ones. In the case of learning a two-choice task, there is by definition only one poor choice (given that the choices are not equally good) which makes the distinction between valuable limited knowledge and full correct knowledge of the choice contingencies smaller.

These results suggests that the general strategy of paying more attention to choices and outcomes of demonstrators believed to be high in ability is useful, or adaptive, given that task sizes are sufficiently large and initial expectations of choice values are not too high. The results also demonstrate that in the special case of two-choice tasks, the task size used in both **Study II** and **Study III**, observation of a poor performer is slightly more beneficial than observation of a high performer.

Results from the simulations demonstrate that learning from demonstrators that are high in ability is especially inefficient when initial expectations are high. This is important if we consider the possibility that the initial expectations would also be observationally learned and biased in the same manner. For instance, it seems plausible that a success-oriented media coverage of the stock market could lead stock investors to have overly optimistic views of average expected return. This type of observationally learned initial expectations has been demonstrated in a study investigating social avoidance learning (Lindström & Olsson, 2015). In addition, the results from the simulations show that findings from experiments using a two-choice task cannot necessarily be understood correctly without accounting for effects using larger choice sets, since strategies that appear suboptimal in small choice sets might be optimal given larger sets.

The results from the simulations suggests that it is possible that participants in **Study III** employed a heuristic that would have been profitable in other tasks with larger choice sets although it leads to poor performance in smaller two-choice tasks. Thus, the simulation shows that findings from experiments using a two-choice task might not necessarily be understood correctly without accounting for effects using larger choice sets.

The simulations presented here are based on a very simple model of observational learning, where for instance the learning rate is fixed and only the expected values of observed choices are updated. This naturally limits the generalizability of our results.

It is important to note that the simulations are not directly based on the task used in **Study III** since individual learning is not included in the simulations. Also, it is important to note that even though learning through observation of a high performing demonstrator is more efficient than observation of a poor performing demonstrator, observation of any demonstrator is better than no demonstrator. This is relevant for the interpretation of the findings in **Study III**. Here, participants could not choose between observation of a demonstrator described as high in ability and one described as low in ability. Instead, paying less attention to a particular

demonstrator would always mean paying less attention to a source of valuable information. Therefore, participants' behavior cannot be described as optimal with regards to any task size. However, attention should be thought of as a cost and in this perspective it can be considered rational to pay a smaller cost (less attention) for something believed to be of less value.

5.4 OBSERVED ANTICIPATORY FEAR

In **Study IV** we explored if observational conditioning was influenced by the expressed anxiousness of the learning model, as evident by the ability of the model to discriminate between the threatening stimulus, CS+, and the safe stimulus, CS-. The study was in part motivated by the relatively high prevalence of anxiety in humans. Thus, observation of typically anxious behavior, i.e. fear towards safe stimuli and situations, is likely to be common in every-day life. However, heightened fear responses to safe stimuli are not seen only in anxious individuals but can also result from the common phenomena of fear generalization (Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015), even though exaggerated generalization, overgeneralization, may be linked to heightened levels of anxiety. Interestingly, generalization is demonstrated to be wider when stimuli are reinforced with outcome of negative, compared to positive, valence (Schechtman, Laufer, & Paz, 2010). It is argued that this makes sense since it is potentially more costly to approach dangers than it is to avoid safe situations. For instance, it would be more costly to approach a lion than to avoid a dog. Along the same reasoning, it is possible that fear generalization might be sensitive to the strength of the aversive outcome. If this would be the case, observation of a learning model expressing increased fear responses to safe stimuli could lead the observer to expect the outcome to be more aversive. If anxious behavior, i.e. anticipatory fear towards safe stimuli, would signal heightened aversiveness, this could explain our finding that fear conditioning is more robust following observation of an anxiously behaving learning model compared to a non-anxious one. However, to our knowledge, there exists no study to support the claim that generalization of fear is sensitive to the strength of the aversive outcome, and this proposal should thus be considered as hypothetical.

We see **Study IV** as a first attempt in trying to investigate effects of the learning model's anticipatory fear reaction to the CS. In the manuscript, we discuss some of the limitations of the design, for instance the combination of an auditory US with the use of an auditory startle probe to measure fear-potentiated startle. Also, since the aim of the study was to investigate the effect of observing expressed fear to both dangerous and safe stimuli, it was crucial that information about the CS-US was evident. Thus, the learning model also reacted strongly to the presentation of the US. In order to investigate observational conditioning through observation of the learning models anticipatory fear response, future research should not include observation of the learning models UR.

5.5 LEARNING, DANGER AND INFORMATION

Learning, whether individual or observational, should be sensitive to the predictive relationships in the environment. It has been argued that this underlies associative learning

mechanisms in both humans and non-human animals (C. Heyes, 2012; Rescorla, 1988). To be able to understand how, and what, we can learn through observation of others, it is important to first understand how the underlying statistics of the environment are expressed in the behavior and reactions of others.

From what we know about human decision making, learning about dangers simply through observation of the choices other people make appears difficult (see for instance Fig.1.). On the other hand, observing other's reactions when directly experiencing negative consequences can be very informative of dangers, for instance observing someone eating a poisonous fruit. However, we can hardly expect someone to eat a particular poisonous fruit more than once and if we happen to miss that occasion we can still learn by observing the persons reactions towards the fruit. In this sense, observational learning about dangers differs from observational learning about rewards. In the latter case, it is possible for the observer to extract information about the observed other's knowledge about what is rewarding from observing anticipatory reactions, choice behavior and finally by observing the direct consequences, for instance the outcome of making a choice. In the former case, observational learning should rely heavily on the frequent observations of anticipatory reactions or expressed fear but also of course on observed aversive outcomes of behavior. I hope that future research on observational and social learning will be able to combine learning about rewards with learning about dangers in more realistic settings which allows individuals to learn from both behaviors and emotional reactions.

6 FUTURE DIRECTIONS

For the future, there are several paths to take to further develop the ideas put forward in this thesis and to investigate some of the questions raised along the way.

Results from **Study III** suggested that the effects on observational learning of describing the observed demonstrators as either high or low in ability was driven by biased attention. This was evident by a smaller anticipatory pupil dilation response in those participants that observed demonstrators described as low in ability. This finding needs to be replicated, preferably in a study which varies the described ability within instead of between participants. To further investigate the underlying learning mechanisms and the attentional bias, fMRI could be applied.

During the work to investigate the effect of the learning model's conditioned response, CR, on subsequent observational conditioning, presented in **Study IV**, it became clear that although research on observational conditioning in non-human animals have included observation of a learning model's learned response (Susan Mineka & Cook, 1988) this is lacking from the research carried out in humans. If we would like to understand transmission of fear between humans, future research needs to take the learning model's learned fear into account, not just the learning model's direct reaction to an aversive US. For instance, this could be done by extending the work presented in **Study I** and investigate if the efficiency of observational extinction varies depending on the learning model's expressed fear to the CS+.

In section 5.3., results from simulations of observational reinforcement learning are presented. These results provide several testable predictions. For instance, our results showed that observational learning from a demonstrator high in ability is more beneficial than observational learning from a demonstrator low in ability when the task size is sufficiently large. This could be tested experimentally by letting participants observationally learn probabilistic tasks with varying sizes. In addition, further developing the learning framework upon which the simulations are based would allow us to make more precise predictions. For instance, learning should be better based on what is known about learning in uncertain environments.

7 ACKNOWLEDGEMENTS

As we all know, practically nothing is possible to achieve without the help of others. The work presented in this thesis is no exception. Here is my chance to show some appreciation.

First and foremost, I would like to thank my supervisor Andreas Olsson. You have always been happy, enthusiastic and supportive, and I am very grateful for the trust you have had in me. I would also like to thank Professor Arne Öhman who co-supervised me for the first half of my PhD project and Professor Christian Balkenius who has been my co-supervisor for the final half of my PhD project. Your knowledge has been greatly appreciated. Further, I would like to thank John Axelsson who has been my mentor through the years, sharing his thoughts on academia, research and life.

I have had the great fortune to work with talented colleagues and friends in the Emotionlab. Armita Golkar, Björn Lindström and Tanaz Molapour – I simply don't know where I would have been without you and I don't know how to thank you enough. Philip Pärnamets – thank you for your wisdom and appreciation of the small things in life.

And to all other former and current members of the Emotionlab and co-authors (in alphabetic order): Martin Bellander, Isak Berglund Mattsson-Mårn, Troy Dildine, Kristian Ekeroth, Lisa Espinosa, Jan Haaker, Tove Hensler, Björn Hofvander, Simon Jangard, Johan Lundin Kleberg, Daniel Lundqvist, Åsa Michelgård Palmquist, Olof Roshanbin Hjort, Elin Schröder, Charlotte Sjöberg, Victoria Spring, Irem Undeger, Jonathan Yi - It has been my very great pleasure to work with you all. I would also like to take the opportunity to thank some of the students that have contributed to my work: Oskar Flygare, Liisa Veerus, Laura Mandt, Gjermund Glimsdal & Michaela Davilén.

I would also like to thank all my colleagues and friends at the Division for Psychology at KI as well as people I have met and worked with at KI during my PhD project, especially Rita Almeida, Jonathan Berrebi and everyone in the Neurocomputational journal club.

I would like to thank my friends simply for making my life better, especially Sara Lindström, Kristina Stark and Martina Phalén. I hope I will get to see you more often now.

I am especially grateful to my parents. You have always, throughout my life, helped me out when I have been overwhelmed by deadlines and to-do-lists. The last year has been especially tough and I couldn't have managed without you.

Finally, Anders – you have supported me more than I could have asked for, giving me space to work when I needed it and kindly reminding me to take care of myself from time to time. Also, to Edward & Emanuel – thank you for never being considerate of my work.

8 REFERENCES

Acerbi, A., Tennie, C., & Mesoudi, A. (2016). Social learning solves the problem of narrow-peaked search landscapes: experimental evidence in humans.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5).* Arlington, VA : American Psychiatric Publishing.

Apesteguia, J., Huck, S., & Oechssler, J. (2007). Imitation — theory and experimental evidence. *Journal of Economic Theory*, *136*, 217–235. http://doi.org/10.1016/j.jet.2006.07.006

Askew, C., & Field, A. P. (2007). Vicarious learning and the development of fears in childhood. *Behaviour Research and Therapy*, *45*(11), 2616–27. http://doi.org/10.1016/j.brat.2007.06.008

Askew, C., & Field, A. P. (2008). The vicarious learning pathway to fear 40 years on. *Clinical Psychology Review*, *28*, 1249–1265. http://doi.org/10.1016/j.cpr.2008.05.003

Baeyens, F., Kaes, B., Eelen, P., & Silverans, P. (1996). Observational evaluative conditioning of an embedded stimulus element. *European Journal of Social Psychology*, *26*(1), 15–28. http://doi.org/10.1002/(SICI)1099-0992(199601)26:1<15::AID-EJSP729>3.0.CO;2-5

Baeyens, F., Vansteenwegen, D., De Houwer, J., & Crombez, G. (1996). Observational conditioning of food valence in humans. *Appetite*, *27*(3), 235–250.

Bandura, A. (1965). Vicarious processes: A case of no-trial learning. *Advances in Experimental Social Psychology*, *2*.

Bandura, A., Grusec, J. E., & Menlove, F. L. (1967). Vicarious extinction of avoidance behavior. *Journal of Personality and Social Psychology*, *5*(1), 16–23.

Barnard, R. M., & Sibley, C. J. (1981). Producers and scroungers: a general model and its application to a captive flock of house sparrows. *Animal Behaviour*, *29*(2), 543–550.

Barrett, B. J., McElreath, R., & Perry, S. (2017). Payoff-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate. *Proceedings of the Royal Society B*. http://doi.org/10.1101/110221

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

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of Psychophysiology* (Vol. 2nd, pp. 142–162).

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature*, *456*(7219), 245–9. http://doi.org/10.1038/nature07538

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–21. http://doi.org/10.1038/nn1954

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, *22*(1), 23–36. http://doi.org/10.2307/1909829

Betti, V., & Aglioti, S. M. (2016). Dynamic construction of the neural networks underpinning empathy for pain. *Neuroscience and Biobehavioral Reviews*, *63*(February), 191–206. http://doi.org/10.1016/j.neubiorev.2016.02.009

Boucsein. (2012). *Electrodermal activity*. Springer Science & Business Media.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*

(Cold Spring Harbor, N.Y.), *11*(5), 485–494. http://doi.org/10.1101/lm.78804

Bouton, M. E., Westbrook, R. F., Corcoran, K. a, & Maren, S. (2006). Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biological Psychiatry*, *60*(4), 352–60. http://doi.org/10.1016/j.biopsych.2005.12.015

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press.

Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. University of Florida, Gainesville, Fl.

Bradley, M. M., Miccoli, L., Escrig, M. a, & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–7. http://doi.org/10.1111/j.1469-8986.2008.00654.x

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the "spotlight" of visual attention. *Nature Neuroscience*, *2*(4), 370–374. http://doi.org/10.1038/7280

Bruin, E. N. M. De, & Lange, P. A. M. Van. (2015). What People Look for in Others : Influences of the Perceiver and the Perceived on Information Selection. *Personality and Social Psychology Bulletin*, *26*(2), 206–219.

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. http://doi.org/10.1073/pnas.1003111107

Cameron, G., Schlund, M. W., & Dymond, S. (2015). Generalization of socially transmitted and instructed avoidance. *Frontiers in Behavioral Neuroscience*, *9*(June), 159. http://doi.org/10.3389/fnbeh.2015.00159

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Princeton University Press.

Chater, N. (2009). Rational and mechanistic perspectives on reinforcement learning. *Cognition*, *113*(3), 350–64. http://doi.org/10.1016/j.cognition.2008.06.014

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control.

Chudek, M., Heller, S., Birch, S., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, *33*(1), 46–56. http://doi.org/10.1016/j.evolhumbehav.2011.05.005

Cook, M., & Mineka, S. (1987). Second-order conditioning and overshadowing in the observational conditioning of fear in monkeys. *Behaviour Research and Therapy*, *25*(5), 349–364. http://doi.org/10.1016/0005-7967(87)90013-1

Cook, M., & Mineka, S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *Journal of Abnormal Psychology*, *98*(4), 448–459. http://doi.org/10.1037/0021-843X.98.4.448

Cook, M., Mineka, S., Wolkenstein, B., & Laitsch, K. (1985). Observational Conditioning of Snake Fear in Unrelated Rhesus Monkeys. *Journal of Abnormal Psychology*, *94*(4), 591–610.

Cooper, J. C., Dunne, S., Furey, T., & O'Doherty, J. P. (2012). Human Dorsal Striatum Encodes Prediction Errors during Observational Learning of Instrumental Actions. *Journal of Cognitive Neuroscience*, *24*(1), 106–18. http://doi.org/10.1162/jocn_a_00114

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A.

(2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*(1), 5–27. http://doi.org/10.1016/j.brat.2007.10.003

Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias? *Behavioral and Brain Sciences*, *18*(May), 289. http://doi.org/10.1017/S0140525X00038498

Daw, N. D. (2011). Trial-by-trial data analysis using computational models. In *Decision Making, Affect, and Learning: Attention and Performance XXIII* (pp. 3–38). Oxford University Press.

Daw, N. D., & Doya, K. (2006). The computational neurobiology of learning and reward. *Current Opinion in Neurobiology*, *16*, 199–204. http://doi.org/10.1016/j.conb.2006.03.006

Daw, N. D., O'Doherty, J. P., Dayan, P., Dolan, R. J., & Seymour, B. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–9. http://doi.org/10.1038/nature04766

Dawson, M. E., Schell, A. M., & Courtney, C. G. (2011). The skin conductance response, anticipation, and decision-making. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 111–116. http://doi.org/10.1037/a0022619

Denrell, J. (2003). Vicarious Learning, Undersampling of Failure, and the Myths of Management. *Organization Science*, *14*(3), 227–243. http://doi.org/10.1287/orsc.14.2.227.15164

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

Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear Generalization in Humans: Systematic Review and Implications for Anxiety Disorder Research. *Behavior Therapy*, *46*(5), 561–582. http://doi.org/10.1016/j.beth.2014.10.001

Enquist, M., Eriksson, K., & Ghirlanda, S. (2007). Critical Social Learning: A Solution to Rogers's Paradox of Nonadaptive Culture. *American Anthropologist*, *109*(4), 727–734. http://doi.org/10.1525/AA.2007.109.4.727.728

Erev, I., & Roth, A. E. (1998). Predicting How People Play Games: Reinforcement Learning Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review*, *88*(4), 848–881.

Feldman, M. W., Aoki, K., & Kumm, J. (1996). Individual Versus Social Learning: Evolutionary Analysis in a Fluctuating Environment. *Anthropological Science*, *104*(3), 209–231.

Filion, D. L., Dawson, M. E., Schell, A. M., & Hazlett, E. A. (1991). The Relationship Between Skin Conductance Orienting and the Allocation of Processing Resources. *Psychophysiology*, *28*(4), 410–424. http://doi.org/10.1111/j.1469-8986.1991.tb00725.x

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science*, *306*(5703), 1940–1943. http://doi.org/10.1126/science.1102941

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2008). Time Discounting and Time Preference : A Critical Review. *Journal of Economic Literature*, *40*(2), 351–401.

Galef, JR., B. G., & Durlach, P. J. (1993). Absence of blocking, overshadowing, and latent inhibition in social enhancement of food preferences. *Animal Learning & Behavior*, *21*(3), 214–220.

Galef, B. G. (1988). Imitation in animals: history, definition, and interpretation of data from

the psychological laboratory. In T. R. Zentall & B. G. Galef (Eds.), *Social learning: Psychological and biological perspectives*. L. Erlbaum Hillsdale, New Jersey.

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence. *Psychonomic Science*, *4*(1), 123–124.

Gazendam, F. J., Kamphuis, J. H., & Kindt, M. (2013). Deficient safety learning characterizes high trait anxious individuals. *Biological Psychology*, *92*(2), 342–52. http://doi.org/10.1016/j.biopsycho.2012.11.006

Gerull, F. C., & Rapee, R. M. (2002). Mother knows best: Effects of maternal modelling on the acquisition of fear and avoidance behaviour in toddlers. *Behaviour Research and Therapy*, *40*(3), 279–287. http://doi.org/10.1016/S0005-7967(01)00013-4

Golkar, A., Castro, V., & Olsson, A. (2015). Social learning of fear and safety is determined by the demonstrator's racial group. *Biology Letters*, *11*(1), 2015–2018. http://doi.org/10.1098/rsbl.2014.0817

Golkar, A., Haaker, J., Selbing, I., & Olsson, A. (2016). Neural signals of vicarious extinction learning. *Social Cognitive and Affective Neuroscience*, *11*(10), 1541–1549. http://doi.org/10.1093/scan/nsw068

Golkar, A., & Olsson, A. (2016). Immunization Against Social Fear Learning. *Journal of Experimental Psychology. General*, *145*(April), 665–671. http://doi.org/10.1037/xge0000173

Gottlieb, J. (2012). Attention, learning, and the value of information. *Neuron*, *76*(2), 281–95. http://doi.org/10.1016/j.neuron.2012.09.034

Grillon, C. (2002a). Associative Learning Deficits Increase Symptoms of Anxiety in Humans. *Biological Psychiatry*, *51*, 851–858.

Grillon, C. (2002b). Startle Reactivity and Anxiety Disorders : Aversive Conditioning, Context, and Neurobiology. *Biological Psychiatry*, *52*, 958–975.

Grillon, C., & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology*, *114*(9), 1557–1579. http://doi.org/10.1016/S1388-2457(03)00202-5

Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: interactions between affect and effect. *NeuroImage*, *62*(1), 154–66. http://doi.org/10.1016/j.neuroimage.2012.04.024

Haaker, J., Golkar, A., Selbing, I., & Olsson, A. (2017). Assessment of social transmission of threats in humans using observational fear conditioning. *Nature Protocols*, *12*(7), 1378–1386. http://doi.org/10.1038/nprot.2017-027

Henrich, J., Boyd, R., Derex, M., Kline, M. a, Mesoudi, A., Muthukrishna, M., … Thomas, M. G. (2016). Understanding cumulative cultural evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(44), E6724–E6725. http://doi.org/10.1073/pnas.1610005113

Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1567), 1139–1148. http://doi.org/10.1098/rstb.2010.0323

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*, 165–196.

Hermans, D., Dirikx, T., Vansteenwegenin, D., Baeyens, F., Van Den Bergh, O., & Eelen, P.

(2005). Reinstatement of fear responses in human aversive conditioning. *Behaviour Research and Therapy*, *43*, 533–551. http://doi.org/10.1016/j.brat.2004.03.013

Hess, E. H., & Polt, J. M. (1960). Pupil Size as Related to Interest Value of Visual Stimuli. *Science*, *132*(3423), 349–350.

Heyes, C. (2012). What's social about social learning? *Journal of Comparative Psychology*, *126*(2), 193–202. http://doi.org/10.1037/a0025180

Heyes, C. (2015). When does social learning become cultural learning? *Developmental Science*, (June), 1–14. http://doi.org/10.1111/desc.12350

Heyes, C. M. (1994). Social learning in animals: categories and mechanisms. *Biological Reviews of the Cambridge Philosophical Society*, *69*(2), 207–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8054445

Hill, J. H., Liebert, R. M., & Mott, D. E. W. (1968). Vicarious extinction of avoidance behavior through films: an initial test. *Psychological Reports*, *22*(1).

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: a meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. http://doi.org/10.1037/a0018916

Horner, V., Proctor, D., Bonnie, K. E., Whiten, A., & de Waal, F. B. M. (2010). Prestige affects cultural learning in chimpanzees. *PloS One*, *5*(5), e10625. http://doi.org/10.1371/journal.pone.0010625

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3), e1002410. http://doi.org/10.1371/journal.pcbi.1002410

Hygge, S. (1976). Information about the model's unconditioned stimulus and response in vicarious classical conditioning. *Journal of Personality and Social Psychology*, *33*(6), 764–771. http://doi.org/10.1037/0022-3514.33.6.764

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, *89*(1), 221–234. http://doi.org/10.1016/j.neuron.2015.11.028

Jovanovic, T., Kazama, A., Bachevalier, J., & Davis, M. (2012). Impaired safety signal learning may be a biomarker of PTSD. *Neuropharmacology*, *62*(2), 695–704. http://doi.org/10.1016/j.neuropharm.2011.02.023

Kacelnik, A., & Bateson, M. (2016). Risky Theories: The Effects of Variance on Foraging Decisions. *American Zoologist*, *36*(4), 402–434. http://doi.org/10.1093/icb/36.4.402

Kameda, T., & Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a nonstationary uncertain environment An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, *23*, 373–393.

Kendal, J., Giraldeau, L., & Laland, K. (2009). The evolution of social learning rules: Payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*, *2*. http://doi.org/10.1016/j.jtbi.2009.05.029

Kendal, R., Hopper, L. M., Whiten, A., Brosnan, S. F., Lambeth, S. P., Schapiro, S. J., & Hoppitt, W. (2015). Chimpanzees copy dominant and knowledgeable individuals: Implications for cultural diversity. *Evolution and Human Behavior*, *36*(1), 65–72. http://doi.org/10.1016/j.evolhumbehav.2014.09.002

Kendal, R. L. (2004). The role of conformity in foraging when personal and social information conflict. *Behavioral Ecology*, *15*(2), 269–277. http://doi.org/10.1093/beheco/arh008

Kendal, R. L., Coolen, I., van Bergen, Y., & Laland, K. N. (2005). Trade-Offs in the Adaptive Use of Social and Asocial Learning. *Advances in the Study of Behavior*, *35*, 333–379.

Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement Learning in Robotics: A Survey. *International Journal of Robotics Research*, *32*(11), 1238–1278.

Krypotos, A.-M. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, *9*(July). http://doi.org/10.3389/fnbeh.2015.00189

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27. http://doi.org/10.1177/1745691611427305

Laland, K. N. (2004). Social learning strategies. *Learning & Behavior*, *32*(1), 4–14. http://doi.org/10.1098/rspb.2008.0817

Landis, C., & Hunt, W. (1939). *The Startle Pattern*. New York: Farrar & Rinehart.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, Motivation, and Anxiety: Brain Mechanisms and Psychophysiology. *Biological Psychiatry*, *44*, 1248–1263.

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*, 137–159. http://doi.org/10.1016/S0165-0327(00)00343-8

Lee, D., Seo, H., & Jung, M. W. (2012). Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of Neuroscience*, *35*, 287–308. http://doi.org/10.1146/annurev-neuro-062111-150512

Lenaert, B., Boddez, Y., Griffith, J. W., Vervliet, B., Schruers, K., & Hermans, D. (2014). Aversive learning and generalization predict subclinical levels of anxiety: a six-month longitudinal study. *Journal of Anxiety Disorders*, *28*(8), 747–53. http://doi.org/10.1016/j.janxdis.2014.09.006

Lim, C. L., Rennie, C., Barry, R. J., Bahramali, H., Lazzaro, I., Manor, B., & Gordon, E. (1997). Decomposing skin conductance into tonic and phasic components. *International Journal of Psychophysiology*, *25*, 97–109.

Lindström, B., & Olsson, A. (2015). Mechanisms of Social Avoidance Learning Can Explain the Emergence of Adaptive and Arbitrary Behavioral Traditions in Humans. *Journal of Experimental Psychology: General*, *144*(3), 688–703. http://doi.org/10.1037/xge0000071

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy*, *43*(11), 1391–424. http://doi.org/10.1016/j.brat.2004.10.007

Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of Conditioned Fear as a Pathogenic Marker of Panic Disorder. *American Journal of Psychiatry*, *167*(1), 47–55.

Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., … Grillon, C. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behaviour Research and Therapy*, *47*(2), 111–8. http://doi.org/10.1016/j.brat.2008.10.017

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., … Merz, C. J. (2017). Don't fear "fear conditioning": Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of

fear. *Neuroscience & Biobehavioral Reviews*, *77*, 247–285. http://doi.org/10.1016/j.neubiorev.2017.02.026

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF.

Martijn, C., Spears, R., Van Der Pligt, J., & Jakobs, E. (1992). Negativity and positivity effects in person perception and inference: Ability versus morality. *European Journal of Social Psychology*, *22*(5), 453–463. http://doi.org/10.1002/ejsp.2420220504

McElreath, R., Bell, A. V, Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1509), 3515–28. http://doi.org/10.1098/rstb.2008.0131

Merlo, A., & Schotter, A. (2003). Learning by not doing: an experimental investigation of observational learning. *Games and Economic Behavior*, *42*, 116–136. http://doi.org/10.1016/S0899-8256(02)00537-7

Mesoudi, A. (2008). An experimental simulation of the "copy-successful-individuals" cultural learning strategy: adaptive landscapes, producer–scrounger dynamics, and informational access costs. *Evolution and Human Behavior*, *29*(5), 350–363. http://doi.org/10.1016/j.evolhumbehav.2008.04.005

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: Payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, *32*(5), 334–342. http://doi.org/10.1016/j.evolhumbehav.2010.12.001

Mesoudi, A. (2017). Pursuing Darwin's curious parallel: Prospects for a science of cultural evolution. *Proceedings of the National Academy of Sciences*, *114*(30), 7853–7860. http://doi.org/10.1073/pnas.1620741114

Meulders, A., Vervliet, B., Fonteyne, R., Baeyens, F., Hermans, D., & Vansteenwegen, D. (2012). Preexposure to (un)predictable shock modulates discriminative fear learning between cue and context: An investigation of the interaction between fear and anxiety. *International Journal of Psychophysiology*, *84*(2), 180–187. http://doi.org/10.1016/j.ijpsycho.2012.02.004

Milad, M. R., & Quirk, G. J. (2012). Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. *Annual Review of Psychology*, *63*, 129–151. http://doi.org/10.1146/annurev.psych.121208.131631.Fear

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of Fear Extinction in Humans Activates the Ventromedial Prefrontal Cortex and Hippocampus in Concert. *Biological Psychiatry*, *62*(5), 446–454. http://doi.org/10.1016/j.biopsych.2006.10.011

Mineka, S., & Cook, M. (1986). Immunization against the Observational Conditioning of Snake Fear in Rhesus-Monkeys. *Journal of Abnormal Psychology*, *95*(4), 307–318. http://doi.org///dx.doi.org.proxy.bc.edu/10.1037/0021-843X.95.4.307

Mineka, S., & Cook, M. (1988). Social Learning and the Acquisition of Snake Fear in Monkeys. In T. R. Zentall & B. G. Galef, JR. (Eds.), *Social Learning: Psychological and Biological Perspectives*. L. Erlbaum Hillsdale, New Jersey.

Mineka, S., & Cook, M. (1993). Mechanisms involved in the observational conditioning of fear. *Journal of Experimental Psychology. General*, *122*(1), 23–38.

Mineka, S., Davidson, M., Cook, M., & Keir, R. (1984). Observational conditioning of snake fear in rhesus monkeys. *Journal of Abnormal Psychology*, *93*(4), 355–372.

http://doi.org/10.1037/0021-843X.95.4.307

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., & Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B*, *279*, 653–662. http://doi.org/10.1098/rspb.2011.1172

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. *Molecular Psychiatry*, *12*, 120–150. http://doi.org/10.1038/sj.mp.4001939

Neumann, E., & Blanton, R. (1970). The early history of electrodermal research. *Psychophysiology*, *6*(4), 453–476.

Nicolle, A., Symmonds, M., & Dolan, R. J. (2011). Optimistic biases in observational learning of value. *Cognition*, *119*(3), 394–402. http://doi.org/10.1016/j.cognition.2011.02.004

Niv, Y. (2011). Reinforcement learning in the brain. *Learning*, 1–38. http://doi.org/10.1016/j.jmp.2008.12.005

Niv, Y., Edlund, J. A., Dayan, P., & O'Doherty, J. P. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience*, *32*(2), 551–562. http://doi.org/10.1523/JNEUROSCI.5498-10.2012

O'Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 1–10. http://doi.org/10.1073/pnas.1305373110

Olsson, A., McMahon, K., Papenberg, G., Zaki, J., Bolger, N., & Oschner, K. (2016). Vicarious fear learning depends on empathic appraisals and trait empathy. *Psychological Science*, *27*(1), 25–33. http://doi.org/10.1111/1467-9280.00327

Olsson, A., Nearing, K. I., & Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. *Social Cognitive and Affective Neuroscience*, *2*(1), 3–11. http://doi.org/10.1093/scan/nsm005

Olsson, A., & Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends in Cognitive Sciences*, *12*(2), 65–71. http://doi.org/10.1016/j.tics.2007.11.010

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

Phelps, E. A., Delgado, M. R., Nearing, K. I., & Ledoux, J. E. (2004). Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron*, *43*(6), 897–905. http://doi.org/10.1016/j.neuron.2004.08.042

Poser, E. G., & King, M. C. (1975). Strategies for the prevention of maladaptive fear responses. *Canadian Journal of Behavioural Science*, *7*(4), 279–294. http://doi.org/10.1037/h0081915

Rachman, S. (1977). The conditioning theory of fearacquisition: A critical examination. *Behaviour Research and Therapy*, *15*(5), 375–387. http://doi.org/10.1016/0005-7967(77)90041-9

Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., … Laland, K. N. (2010). Why Copy Others? Insight from the Social Learning Strategies Tournament. *Science*, *328*(5975), 208–213. http://doi.org/10.1126/science.1184719.Why

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in Cognitive Sciences*, *15*(2), 68–76. http://doi.org/10.1016/j.tics.2010.12.002

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

Rescorla, R. A. (2004). Spontaneous Recovery. *Learning & Memory*, *11*, 501–509. http://doi.org/10.1101/lm.77504.ing

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current Theroy and Research* (pp. 64–99). Appleton-Century-Crofts, New York Educational Division, Meredith Corporation.

Richerson, P. J., & Boyd, R. (2008). *Not By Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press.

Ritter, B. (1968). The group desensitization of children's snake phobias using vicarious and contact desensitization procedures. *Behaviour Research and Therapy*, *6*, 1–6.

Rizley, R. C., & Rescorla, R. a. (1972). Associations in second-order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology*, *81*(1), 1–11. http://doi.org/10.1037/h0033333

Rogers, A. R. (1988). Does Biology Constrain Culture. *American Anthropologist*, *90*(4), 819–831. http://doi.org/10.1525/aa.1988.90.4.02a00030

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

Schechtman, E., Laufer, O., & Paz, R. (2010). Negative valence widens generalization of learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(31), 10460–4. http://doi.org/10.1523/JNEUROSCI.2377-10.2010

Schultz, W., Dayan, P., & Montague, P. R. (1997). A Neural Substrate of Prediction and Reward. *Science*, *275*, 1593–1599. http://doi.org/10.1126/science.275.5306.1593

Selbing, I., & Olsson, A. (2017). Strategies of Observational Reinforcement Learning. Poster presentation, The Multi-disciplinary Conference on Reinforcement Learning and Decision Making; Ann Arbor, MI.

Seligman, L. D., & Wuyek, L. A. (2005). Vicarious extinction. In *Encyclopedia of behavior modification and cognitive behavior therapy: Vol 3* (pp. 1085–1086). Thousand Oaks, CS: Sage.

Skinner, B. F. (1938). *The Behavior of Organisms: An experimental analysis*. Oxford, England: Appleton-Century. http://doi.org/10.1037/h0052216

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, Massachusetts: The MIT Press.

Suzuki, S., Harasawa, N., Ueno, K., Gardner, J. L., Ichinohe, N., & Haruno, M. (2012). Learning to Simulate Others' Decisions. *Neuron*, *74*(6), 1125–1137. http://doi.org/10.1016/j.neuron.2012.04.030

Thorndike, E. L. (1911). *Animal Intelligence*. New York: Macmillan.

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*, 495–552.

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*, 167–174. http://doi.org/10.1111/j.1469-8986.2003.00148.x

Vrana, S. R., Spence, E. L., & Lang, P. J. (1988). The Startle Probe Response : A New Measure of Emotion? *Journal*, *97*(4), 487–491.

Wadlinger, H. A., & Isaacowitz, D. M. (2006). Positive mood broadens visual attention to positive stimuli. *Motivation and Emotion*, *30*(1), 87–99. http://doi.org/10.1007/s11031-006-9021-1

Watkins, C. J. C. H., & Dayan, P. (1992). Q-Learning. *Machine Learning*, *8*, 279–292.

Watson, J. B. (1958). *Behaviorism*. Transaction Publishers.

Webster, M. M., & Laland, K. N. (2008). Social learning strategies and predation risk: minnows copy only when using private information would be costly. *Proceedings of the Royal Society B*, *275*(1653), 2869–2876. http://doi.org/10.1098/rspb.2008.0817

Yarbus, A. L. (1967). *Eye Movements And Vision*. New York.

Zaki, J., Wager, T. D., Singer, T., Keysers, C., & Gazzola, V. (2016). The Anatomy of Suffering : Understanding the Relationship between Nociceptive and Empathic Pain. *Trends in Cognitive Sciences*, *20*(4), 249–259. http://doi.org/10.1016/j.tics.2016.02.003

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, *108*(3), 483–522. http://doi.org/10.1037//0033-295X.108.3.483