INDIVIDUAL DIFFERENCES IN DEVELOPMENT AND REPRESENTATION OF NOVEL AFFECTIVE ASSOCIATIONS

by

Mana R. Ehlers

B.Sc., Carl-von-Ossietzky University Oldenburg, Germany, 2012

M.Sc., University of Bremen, Germany, 2014

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Neuroscience)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Individual differences in development and representation of novel affective associations

submitted by Mana R. Ehlers in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

In Neuroscience

Examining Committee:

Prof Rebecca Todd
Supervisor

Prof Luke Clark
Supervisory Committee Member

Prof Lara Boyd
Supervisory Committee Member

Prof Karon MacLean
University Examiner

Prof Catharine A. Winstanley
University Examiner

Additional Supervisory Committee Members:

Supervisory Committee Member

Supervisory Committee Member

Abstract

Emotionally arousing events are typically better remembered than mundane ones, in part because emotionally relevant aspects of our environment are prioritized in attention. Such biased attentional tuning is itself the result of associative processes through which we learn affective and motivational relevance of cues. While such affective biases in cognition can be highly adaptive, extreme biases to specific categories of aversive or rewarding stimuli can be symptomatic of psychopathology. That raises the question which factors contribute to individual differences in development of affective biases via emotional learning processes and how emotional associations come to be represented in the brain.

More specifically, the present thesis aimed to investigate the role of individual differences in the norepinephrine and stress system in emotional learning processes. In Study I, I demonstrated that a common genetic variation putatively influencing norepinephrine availability is associated with subjective perception of ambiguous stimuli as more rewarding. Moreover, change in affective bias was mediated by acute stress. Thus, in the first study I established that individual differences in the locus coeruleus-norepinephrine (LC-NE) and stress system play a role in affective perception and the flexibility of the underlying subjective biases. In Study II and III, I found that acute stress affects both classical and operant conditioning and that the direction of those effects depends on the timing of the stressor relative to the learning experience. Study IV aimed to investigate the neural representation of the development of novel affective associations using functional magnetic resonance imaging (fMRI). By means of representational similarity analysis (RSA) - a multivariate approach to analyzing neuroimaging data - the study revealed that conditioned stimuli reactivate the representational pattern elicited by the

unconditioned stimuli. I further observed that it is specifically the hedonic response to the unconditioned stimulus that is being reproduced by the conditioned stimulus.

Together these studies demonstrate a role for the norepinephrine and stress system in reward-based learning as well as providing new information of neural mechanisms underlying emotional learning. This research provides insight into individual differences in emotional learning processes that can underlie formation of affective biases.

Lay Summary

We typically pay more attention to emotionally relevant aspects of our environment. Those affective biases are themselves the result of emotional learning processes, raising the question of which factors promote associative learning processes and what the underlying neural mechanisms are. My research revealed that flexibility of affective biases is influenced by genetic differences in our arousal system. Moreover, I demonstrated that both perception of affective information as well as learning which stimuli or actions are associated with a positive outcome is affected by exposure to acute stress. In addition, I was able to show that developing those novel associations, especially our emotional response to them, is represented in brain regions important for emotional processing. Overall my research suggests that individual differences in our stress and arousal system influence learning of positive associations. In the brain, cues that predict positive or negative outcomes reproduce the feeling that we experienced in response to those emotional stimuli.

Preface

I wrote this dissertation with minor edits from RM Todd. I was responsible for the research described in Chapter 2 to 5 and the theoretical framing and summary of findings presented in Chapter 1 and 6. Studies presented in Chapter 2 and 3 have been published as described below. The theoretical background presented in Chapter 1 and parts of Chapter 6 have also been published as a review paper.

A version of Chapter 1 and parts of Chapter 6 have been published as Ehlers MR, Todd RM (2017) Genesis and maintenance of attentional biases: The role of the locus coeruleus-noradrenaline system, Neural Plasticity, 2017, 6817349. I was responsible for manuscript composition. RM Todd was responsible for manuscript composition and critical review of the manuscript.

A version of Chapter 2 has been published as Ehlers MR, Ross CJD**,** Todd RM (2018) The influence of the noradrenergic/stress system on perceptual biases for reward, Cognitive, Affective & Behavioural Neuroscience, https://doi.org/10.3758/s13415-018-00657-0. I was responsible for study design, data collection, most analysis and interpretation as well as for writing the manuscript. CJD Ross was responsible for the analysis of genotyping data and critical review of the manuscript. RM Todd was involved in study design, interpretation and manuscript review. The study was approved by the University of British Columbia's Research Ethics Board, H13-01385: DAMP.

A version of Chapter 3 has been published as Ehlers MR, Todd RM (2017) Acute psychophysiological stress impairs human associative learning, Neurobiol of Learn and Mem, 145, 84-93. I was responsible for study design, data collection, analysis and interpretation as well as write up. RM Todd was involved in study design, data analysis, and interpretation and critical review of the manuscript. The study was approved by the University of British Columbia's Research Ethics Board, H15-00711: Emotional Learning.

A version of Chapter 4 is currently under review as Ehlers MR**,** Todd RM (under review) Appetitive learning is facilitated by immediate but not delayed stress, Emotion, manuscript number: EMO-2019-1533. I was responsible for study design, data collection, analysis and interpretation as well as manuscript composition. RM Todd was responsible for study design, interpretation and manuscript review. The study was approved by the University of British Columbia's Research Ethics Board, H15-00711: Emotional Learning.

Chapter 5 is based on work conducted in UBC's Motivated Cognition Lab and Cornell University's Affect and Cognition Lab. I was responsible for study design, data collection, analysis and interpretation and write up of the chapter. SM Moore was responsible for study design and data collection. AO Beukers and JH Kryklywy were involved in data analysis. JH Krykylwy, RM Todd and AK Anderson were further responsible for study design and result interpretation. Data collection has been completed at Cornell University and the study was approved by the Cornell University's Research Ethics Board.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACC	anterior cingulate cortex
ACTH	adrenocorticotrophic hormone
AFNI	Analysis of Functional NeuroImages
aINS	anterior insula
Amy	amygdala
ANOVA	analysis of variance
ANS	autonomic nervous system
BDI	Beck Depression Inventory
BIC	Bayesian Information Criterion
BOLD	blood-oxygen-level-dependent
CPP	conditioned place preference
CR	conditioned response
CRF	corticotophin releasing factor
CS	conditioned stimulus
CTQ	Childhood Trauma Questionnaire
DA	dopamine
DNA	deoxyribonucleic acid
fMRI	functional magnetic resonance imaging
FOV	field of view
GFBS	greedy best-first algorithm
HPA	hypothalamic-pituitary-adrenal
ICA	independent component analysis
IRM	ideal representation model
LC	locus coeruleus
LSAS	Liebowitz Social Anxiety Scale
ME	multi-echo
mm	millimeter
MNI	Montreal Neurological Institute
MRI	magnetic resonance imaging
ms	milliseconds

NE	norepinephrine
OC	optimally combined
OFC	orbitofrontal cortex
PANAS	Positive and Negative Affect Schedule
PFC	prefrontal cortex
pIns	posterior insula
PIT	Pavlovian-Instrumental Transfer
PTSD	post-traumatic stress disorder
ROI	region of interest
RSA	representational similarity analysis
rUS	reconstructed unconditioned stimulus
s	seconds
S1	primary somatosensory cortex
SAM	sympathetic-adrenal-medullary
SECPT	socially evaluated cold pressor test
S-R	stimulus-response
S-S	stimulus-stimulus
STAI	State-Trait Anxiety Inventory
TE	echo time
TR	repetition time
UR	unconditioned response
US	unconditioned stimulus
V1	primary visual cortex
vmPFC	ventromedial prefrontal cortex
VVS	ventral visual stream

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Rebecca Todd for her incredible mentorship and guidance over the last 5+ years. When I first came to Vancouver to work with her, I would not have dreamed that she would offer to take me on as a PhD student. I have never once regretted that decision and I am well aware that I would not be where I am today without her. The balance of personal and professional that she demonstrates every day, created a lab environment of friends not just colleagues. Spending each year's lab retreat at her farm, will forever be one of my fondest memories of grad school. But what really sets her apart is her positivity, optimism and belief in me and all her other students. I remember many meetings with her that I was anticipating with low motivation and self-doubt and I do not know how she manages it, but the majority of time I left those meetings feeling light, optimistic and full of new hope and motivation. And let's be honest, those times are common in grad school and are when you need a mentor more than ever (even if you disagree on the subject of carob).

Thanks to the members of my supervisory committee Dr. Luke Clark, Dr. Lara Boyd and Dr. Colin Ross for their ongoing guidance, feedback and mentorship as well as to my collaborators Dr. Adam Anderson and Dr. Sarah Moore, without whom this thesis would not be the same.

In addition to my supervisor, I want to thank the entire MC Lab for putting up with me and each in your own way helping me smooth out some of my edges (and develop others) and making me a better person overall. Special thanks to Joey, Jowi, Geoy, Glowi or whichever name you prefer today. You are one of the most loyal, selfless and kind people that I was ever allowed to call my friend. I am glad that I have you on my side (we all know what you would be able to do with your machete if someone hurt the people closest to you). Thanks Alex for being

my therapist and allowing me to feel like yours. I admire your dedication to other people's well-being and your own self-improvement. Thank you for being a light in my life when you were the only person I could stand to be around. Thanks Jamie for being my best friend, my most successful collaboration and simply for all that. Thank you for always being there, no matter what, for believing in me when I couldn't and challenging me in ways that allowed me to grow and improve. Thank all of you as well as Kevin, Lia and Max for teaching me how to science, for trying all my attempts to make cake and for being there through my highs and lows. Thank all of you for being the lab members and friends that you were over the years. I would not have been able to go through grad school without you.

To my parents, for never putting any pressure on me about what I should do or become when or how; for teaching me what is really important in life. To my friends at home who patiently waited for me all these years, who continued to make time for me the few times that I made it home and who will hopefully welcome me back with open arms.

Last but by no means least, I would like to thank Vancouver for making me feel home from the beginning. To this date, I am in awe of the beautiful nature surrounding this city. While I had my moments of doubt in life in general and in grad school in particular, I have never doubted I belong here.

Chapter 1: Introduction

1.1 Introduction

It is well established that emotional salience - that is the degree to which information stands out due to its positive or negative or arousing qualities - modulates both learning and memory. For example, we typically remember emotionally arousing events better than mundane ones, reliving the birth of a child or a teenage humiliation with a high degree of vividness decades later (Cahill & McGaugh, 1998; Kensinger & Corkin, 2003; LaBar & Cabeza, 2006). We remember these events better in part because we pay heightened attention to emotionally relevant aspects of our environment that signal potential punishment and reward (Markovic, Anderson, & Todd, 2014; Pourtois, Schettino, & Vuilleumier, 2013). In turn, such patterns of heightened attention are themselves the result of emotional learning processes that tune our perceptual systems to prioritize such affectively and motivationally relevant cues [e.g. (Chelazzi et al., 2014; Lim, Padmala, & Pessoa, 2008, 2009)]. Visual selective attention, or attentional prioritization, is the process by which we tune ourselves to the world so that, of the millions of bits per second transmitted by the retina (Koch et al., 2006), the information that is most important, or salient to us, reaches awareness and guides action. Affect-biased attentional prioritization (Todd, Cunningham, Anderson, & Thompson, 2012), or selective prioritization of what is emotionally or motivationally relevant, can be highly adaptive, as emotional arousal signals events that are important to attend and remember in the interest of survival. Yet at the extreme ends of the spectrum, affect-biased attentional prioritization of specific categories of stimulus, which are often unconscious and automatic, are symptomatic of psychopathology. For example, implicit biases toward stimuli associated with threat characterize anxiety disorders (Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van, 2007), and biases to attend

trauma-related cues characterize post-traumatic stress disorder (PTSD) (Todd, MacDonald, et al., 2015). According to popular models of PTSD, such trauma-related biases are themselves the result of Pavlovian associative learning processes (Mahan & Ressler, 2012). Moreover, altered biases in attention to reward-related cues are linked to both depression (B. A. Anderson, Leal, Hall, Yassa, & Yantis, 2014; Peckham, McHugh, & Otto, 2010) and addictive behaviours (B. A. Anderson, 2016b; B. A. Anderson, Faulkner, Rilee, Yantis, & Marvel, 2013; Waters, Heishman, Lerman, & Pickworth, 2007). In addiction as well, biases to addictive cues are thought to result from learning associations between the cue predicting reward and the actual reward (B. A. Anderson, 2016b).

Observing those maladaptive attentional pattern in some individuals but not others raises the question which factors create individual differences in attentional or perceptual biases, as well the associative learning processes that give rise to them. The locus coeruleus-norepinephrine (LC-NE) and stress system have long been known to play a role in the inception and development of addictive behaviours, by elevating drug use and promoting relapse (Piazza & Le Moal, 1998; Sinha, 2008) and in the etiology of depression (Yang et al., 2015). However, much less is known about how the LC/stress system affects appetitive learning processes that might underlie the development of psychopathological conditions such as addiction or depression (Ehlers & Todd, 2017b). In addition, while the basic neural circuits underlying associative learning processes are relatively well mapped out, there is a lack of understanding of how emotional associations develop in the brain. More specifically, it is unclear what meaning becomes attached to a stimulus that now represents something good or bad. Thus, this dissertation will focus on the role of the norepinephrine and stress system in associative learning processes and the neural mechanisms underlying it.

1.2 Genotype-dependent differences in affective biases

One approach to investigating the role of individual differences in the norepinephrine system is to employ genotyping. *ADRA2b* is a common [~50% of population (Li, Weerda, Guenzel, Wolf, & Thiel, 2013; Todd et al., 2014; Todd, Schmitz, Susskind, & Anderson, 2013)] deletion variant of the *ADRA2b* gene, which codes for the alpha2b adrenoceptor (Small, Brown, Forbes, & Liggett, 2001). This genetic variation has been related to differences in blood pressure (Zhang et al., 2005) and vasoconstriction (Muszkat et al., 2005) acting on noradrenergic receptors in the periphery. Genetically modified mice carrying the deletion variant have been shown to have increased plasma NE levels (Makaritsis, Johns, Gavras, & Gavras, 2000). In addition, genetically-based differences in cortisol secretion have been observed between trauma victims and matched controls (Fridman, van IJzendoorn, Sagi-Schwartz, & Bakermans-Kranenburg, 2012). Results of behavioural studies in humans suggested that the deletion variant is associated with effects that are similar to those of an alpha2b receptor antagonist (de Quervain et al., 2007), suggesting increased NE availability. While the mechanisms underlying the effects of *ADRA2b* on emotional processing are not fully understood, a recent meta-analysis supports the robustness of genetically-based differences in emotion and cognition (Xie, Cappiello, Meng, Rosenthal, & Zhang, 2018), which are further discussed in the following paragraph.

Previous research has demonstrated more pronounced affective biases in deletion carriers relative to non-carriers in both memory and attention: In a seminal paper, de Quervain et al. (2007) showed greater emotional memory enhancement for deletion carriers. We subsequently found the deletion variant to be associated with enhanced attentional biases towards (Todd, Muller, et al., 2013b) and more vivid perception of (Todd, Ehlers, et al., 2015) emotionally salient stimuli, indicating it plays a role in prioritized encoding of emotional information. An

outstanding question concerns whether more pronounced affective biases in deletion carriers result from the influence of NE on learning. A hypothesized role for alpha2b noradrenergic receptors in emotional learning is supported by rodent studies showing reduced emotional learning upon full development of inhibitory alpha2b receptors (Moriceau & Sullivan 2004). Moreover, human studies investigating *ADRA2b* have suggested that deletion carriers show greater cognitive-affective flexibility relative to non-carriers (Mammarella et al., 2016). Combining an emotional working memory task with a task requiring action to switch from negative to positive affective intonation, the authors found that deletion carriers remembered more positive words. In addition, despite the fact that deletion carriers were less willing to switch the intonation from negative to positive, they remembered more positive information suggesting higher flexibility than non-carriers. In summary, a number of studies have shown genotype-dependent differences in affective biases, and some investigations have suggested differences in flexibility. What remained to be investigated was how *ADRA2b* genotype influences flexibility in shifting pre-existing biases. Thus, in Chapter 2, I address the question of whether putative differences in NE availability influence flexibility in shifting biases based on experience and, more broadly speaking, whether individual differences in the NE system influence affective bias manifestation.

1.3 Attentional bias as product of emotional learning

Implicit biases in attentional prioritization not only influence what we encode and remember, they are themselves the product of learning and memory. Research in my lab has found that in "real life" the categories of stimulus for which attentional selection is biased are strongly shaped by traumatic experiences. Through these experiences, neutral stimuli are linked

to high emotional arousal through associative learning processes (Lee, Todd, Gardhouse, Levine, & Anderson, 2013; Todd, MacDonald, et al., 2015). Moreover, the degree of this bias predicts PTSD diagnosis, and is highly correlated with anxiety symptoms. Such examples of high-arousal associative learning experiences mirror effects found in controlled laboratory experiments using fear conditioning, and complement a wide literature linking fear conditioning to anxiety disorders (Lissek et al., 2008; Lissek et al., 2005; Lissek et al., 2009; Wilker, Elbert, & Kolassa, 2014). On the other end of the valence spectrum, attentional biases for substance-related stimuli, or cues, which predict craving in addiction can also be created through classical conditioning processes (B. A. Anderson, 2016a; Field & Cox, 2008). Thus, considerable evidence suggests that attentional biases towards specific categories of salient stimuli develop through associative learning processes, and they do so at time scales that can range from minutes to decades.

Moreover, evidence in humans and non-human animals suggests that the NE/stress system also plays a role in such associative learning processes raising the question whether individual differences in exposure to stress can facilitate or impair the development of novel affective associations which in turn lead to observable differences in affective biases, potentially contributing to pathological disorders. In the following section I will review the basic mechanisms and current state of the literature on different associative learning processes.

1.4 Associative learning in humans and non-human animals

Associative learning is used as an umbrella term to refer to forms of learning that are characterized by the development of conscious or unconscious associations between a certain cue or action and the occurrence of a specific stimulus (Skinner, 1938). For example, in an aversive *classical conditioning* paradigm, an animal learns to associate an initially neutral conditioned

stimulus (CS+) with an aversive unconditioned stimulus (US) or event that elicits an innate response (Rescorla, 1968). After learning, the presentation of the CS+ alone leads to the aversive response. In *operant conditioning* or *reinforcement learning*, an animal learns that performing a certain action (e.g. pressing a lever) is followed by a specific outcome (e.g. delivery of food reward). Similar paradigms have been developed to study associative learning in humans.

1.4.1 Aversive conditioning

Aversive classical conditioning, or Pavlovian conditioning, is any type of conditioning wherein an animal or human rapidly (often unconsciously) learns to associate an initially neutral stimulus (CS+) with an aversive stimulus or event (US) (Pavlov, 2010). Fear conditioning paradigms in which after learning, the presentation of the CS+ alone leads to a fear response, are the most prevalent form of aversive conditioning in the literature [for reviews see (Delgado, Olsson, & Phelps, 2006; J. E. LeDoux, 2000; Maren, 2001)]. Because of its reliability and simplicity, as well as the high degree of control the experimenter has over all aspects of the learning process, fear conditioning has been used as a tool to understand general mechanisms of associative learning (J. E. LeDoux, 2000).

A wide range of possible CS and US can be chosen from when studying aversive conditioning in humans (Sehlmeyer et al., 2009). While auditory or even olfactory CS can be found in the literature, visual conditioned stimuli are by far the most common type. Many researchers employ photographs of human faces or geometrical figures. Similarly, US differ in stimulus modality as well as in salience and unpleasantness. Among the most common US are electric shocks, air blasts, auditory cues or painful stimuli. Behaviourally, successful

conditioning can be assessed by physiological measures such as skin conductance response (SCR) or changes in heart rate as well as subjective stimulus ratings (LaBar & Cabeza, 2006).

1.4.2 Appetitive conditioning

Appetitive conditioning is an associative learning process by which initially neutral stimuli or events become associated with a reward and hence gain motivational salience. In appetitive classical conditioning the presentation of a cue (CS+) becomes passively associated with a reward (US). Reward learning is more often studied in the form of appetitive operant conditioning or reinforcement learning. In the following section, I will review what is known about the role of the norepinephrine and stress system and which open question my research program addressed.

1.5 The stress and norepinephrine system

1.5.1 Acute stress: the two systems

Stress, defined as a state of real or perceived threat to an organism's homeostasis (S. M. Smith & Vale, 2006), has long been known to effect cognition [for reviews see (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007; Sandi, 2013)]. In the current dissertation, acute stress is identified as one source of individual differences in appetitive learning processes underlying the development of affective biases.

When an organism is challenged by an acute stressor, it responds with a range of physiological and behavioural changes in order to maintain or restore homeostasis. The biological stress response is mediated by the activation of two different systems: The immediate reaction of the autonomic nervous system (ANS) and the slower response of the hypothalamic-

pituitary-adrenal (HPA) axis. Activation of the fast-acting stress system leads to a release of mostly catecholamines such as norepinephrine (NE) and dopamine (DA) in the brain (Arnsten, 2009; Joels, Pu, Wiegert, Oitzl, & Krugers, 2006; Shansky & Lipps, 2013). Through activation of the sympathetic-adrenal-medullary (SAM) axis central (e.g. amygdala, hippocampus and prefrontal cortex) and peripheral release of adrenaline and noradrenaline is initiated enabling the organism to fight, flight or freeze [for review see (Roelofs, 2017)]. The fast autonomic response is followed by the slower-acting stress response, the activation of the HPA axis. Secretion of corticotrophin releasing factor (CRF) from the hypothalamus leads to a release of adrenocorticotrophic hormone (ACTH) from the pituitary, which in turn stimulates glucocorticoid (cortisol in humans) secretion from the adrenal glands [for review see (Schwabe, Wolf, & Oitzl, 2010)]. The activation of these two different stress systems has been known for decades to influence cognitive functions such as learning and memory (de Kloet, Oitzl, & Joels, 1999; McGaugh, 1966). However, the influence of the two stress systems on appetitive learning processes has been mostly neglected (Ehlers & Todd, 2017b; Weinshenker & Schroeder, 2007). A detailed review of the current literature on this topic will follow in the next section.

1.5.2 The role of glucocorticoids and stress in appetitive learning

Most research on the role of glucocorticoids in appetitive learning processes conducted to date has been centered on habit formation. Habit formation involves a shift from goal-directed to habitual behaviour in operant learning: Early in the learning process, animal behaviour is predominantly goal-directed; the animal performs the action leading to a reward (e.g. drug taking), and the action-outcome association is developed (Balleine & Dickinson, 1998). Later behaviour becomes much more habitual or even compulsive: It is no longer the reinforcing

property of the reward (e.g. a drug) that leads to action completion but the action is performed irrespective of the actual outcome and even despite negative consequences (Everitt & Robbins, 2016).

This shift has been shown to be facilitated by glucorticoid action (Schwabe, Tegenthoff, Hoffken, & Wolf, 2012). That is under the influence of glucocorticoids behaviour shifts from flexible, goal-directed behaviour to more rigid, habitual control of behaviour. It is no longer the rewarding outcome driving ones behaviour but simple stimulus-response mechanisms that have been established (Schwabe & Wolf, 2011). This has been further demonstrated by additional studies (Schwabe & Wolf, 2009, 2010b) in which participants were exposed to acute psychophysiological stress or a control condition either before or after operant training tasks. Participants in the stress group showed a more persistent habitual performance - that is, the performed operant action even in the absence of reward both when stress was induced before and after contingencies were learned. Thus, those studies suggest that glucocorticoids can act in at least two different ways to promote habitual behaviours: the initial learning stage and the extinction period in which habitual behaviours become apparent. What remains to be investigated however is how early stages of habit formation, that is simple associative learning processes, are affected by glucorticoid action.

Another recent human study (Pool, Brosch, Delplanque, & Sander, 2015) investigating Pavlovian-Instrumental Transfer (PIT) - a learning process under motivational not habitual control - showed that stress induced cortisol release increased the craving for a rewarding outcome without affecting the pleasure of consuming it. The 3-stage PIT task employed (Talmi, Seymour, Dayan, & Dolan, 2008) involves three distinct processes. In the operant conditioning phase, the association between an action and reward is established via operant conditioning

(Balleine, 2011; Skinner, 1938). In the second, Pavlovian learning phase, a passive association is made between a stimulus and reward. Finally, during the subsequent extinction phase, transfer behaviour is measured by strength and persistence of instrumental action in response to the Pavlovian stimulus in the absence of reward. Here the group exposed to higher cortisol levels mobilized more effort in response to the now-unrewarded Pavlovian stimulus than the control group, which was interpreted as increased cue-triggered 'wanting' (Pool et al., 2015). Critically, in the described study participants were exposed to an acute stress or a no-stress control condition *after* the learning phase in order to investigate effects on habitual transfer. Again the question arises how initial learning in operant and Pavlovian conditioning is affected if stress is induced *before* learning.

Increased wanting of rewards has also been demonstrated in rodents, who have been injected with corticotropin-releasing factor, which mediates the release of glucocorticoids (see section 1.4.1) (Pecina, Schulkin, & Berridge, 2006). In an effort to compare the effects of glucocorticoids on aversive and appetitive conditioning, in an animal study, rats were trained in two Pavlovian conditioning paradigms followed by an immediate injection with a glucocorticoid agonist (Zorawski & Killcross, 2002). The highest dose of agonist significantly enhanced learning suggesting that glucocorticoids may facilitate appetitive and aversive associative learning over several sessions. In a follow up study the authors demonstrated again that post learning glucocorticoid action can facilitate appetitive conditioning but furthermore that the development of associations with specific outcomes was disrupted (Zorawski & Killcross, 2003). Another study reported that repeated low- and high-dose injections of a glucocorticoid receptor agonist after training sessions did not affect Pavlovian conditioning, while instrumental learning

was impaired by high doses (Pielock, Braun, & Hauber, 2013). Most critically, Pavlovian to instrumental transfer was reduced in both experimental groups.

In summary, previous investigations on glucocorticoids effects on appetitive learning show a range of results. Human studies seems to suggest that glucocorticoids promote habitual behaviours while results from animal studies are more equivocal. What becomes apparent, however, is that knowledge gaps remain in the area of simple associative learning processes that are not overlearned or habitual. In Chapter 3 I will address this question by demonstrating a series of studies investigating the effects of the delayed stress response or glucocorticoid action on both classical and operant conditioning tasks in humans.

1.5.3 The role of norepinephrine and stress in appetitive learning

The shift from goal-directed to habitual behaviour does not only rely on glucocorticoid action. In fact, only the simultaneous release of norepinephrine and glucocorticoids, i.e. the full spectrum of stress hormones, leads to the described shift (Schwabe et al., 2012). In addition, increasing evidence suggests that the locus coeruleus-norepinephrine (LC-NE) system – a major mediator of the immediate stress response - is not only important for aversive conditioning [for example (LaLumiere, Buen, & McGaugh, 2003; McGaugh, 2004; Tully & Bolshakov, 2010)], but also plays a role in reward learning. Decades of research have established that dopamine (DA) is essential for the reinforcing effects of various rewards such as drugs (K. C. Berridge, 2007; Flagel et al., 2011; Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). A selective role of DA in reward learning has been shown to be that of a mediator of incentive salience that is the motivational properties that a stimulus develops through conditioning (K. C. Berridge & Robinson, 1998; Flagel et al., 2011). In other words, DA has been shown to be essential for the

'wanting' of a reward, but not for the associated pleasure, or 'liking,' or for the associative learning process. Furthermore, DA has been shown to be key for the coding of reward prediction errors, operationalized as the difference between anticipated and actual reward (Schultz, 2002). However, it has been argued that DA cannot account for effects of all addictive substances (Nutt, Lingford-Hughes, Erritzoe, & Stokes, 2015). Critically, the contribution of NE has been relatively neglected (Weinshenker & Schroeder, 2007) despite its abundance throughout the brain and its central role in arousal, attention, cognitive flexibility, and adaptation (Aston-Jones & Cohen, 2005; Sara, 2009). However, recent investigations have linked activation of the noradrenergic system to motivation as well. NE has been shown to be important for morphine-associated conditioned place preference (CPP) (Zarrindast, Bahreini, & Adl, 2002) as well as the rewarding effects of the drug (Drouin et al., 2002): Decreasing noradrenergic activity (by stimulating alpha2-adrenergic autoreceptors) inhibits the development of CPP, while enhancing NE availability (by receptor inhibition) facilitates conditioning for actual reward learning processes.

A series of single-cell recording studies conducted in monkeys by Bouret and colleagues further supports the involvement of the LC-NE system in reward learning. Recordings from LC neurons during a task with both Pavlovian and operant components revealed that LC neurons are activated during conditioned responses and their response is modulated by goal-directed processes (Bouret & Richmond, 2009). Directly comparing activity of noradrenergic LC and dopaminergic substantia nigra pars compacta neurons suggests that these neurotransmitters play slightly different roles, with DA responding to rewarded actions - possibly related to value - while NE neurons fire in response to unrewarded action, potentially suggesting it signals the cost associated with an action (Bouret, Ravel, & Richmond, 2012). More recent research further

suggests that the LC plays a role in reward processing by integrating motivationally relevant information such as cue information and reward size (Bouret & Richmond, 2015). The authors extend their interpretation of the results to conclude that the LC is necessary to trigger actions requiring a high amount of energy because the incentive salience is low. This idea is supported by their findings showing that noradrenergic neurons increase their firing rate with increased effort in an effort-based decision making task (Varazzani, San-Galli, Gilardeau, & Bouret, 2015). That is, LC activation is necessary to produce behavioural energy in such a task after a cost-benefit analysis, while dopaminergic activity codes information about the costs and benefits involved.

As summarized in a recent theoretical paper, empirical evidence also supports the idea that the LC-NE system works as an uncertainty signal driving behaviour to adapt to environmental changes (Sadacca, Wikenheiser, & Schoenbaum, 2016). In turn, the activation of the LC-NE system in situations of uncertainty with respect to reward expectations might facilitate attentional biases for reward-related cues. Such biases might then allow for more efficient and eventually habitual tracking (B. A. Anderson, 2016a) of cue-outcome relations. Failures of reward evaluations may give rise to the excessive attentional biases for reward-related cues that have been found to characterize addiction (B. A. Anderson, 2016a).

Putative neuronal mechanisms underlying the role of the LC-NE system in attentional mechanisms related to reward have been further elucidated in a recent study suggesting a major role of the LC-NE system in modulating neural gain (Eldar, Cohen, & Niv, 2013). Under some circumstances, increased gain, which is associated with greater NE availability, narrows attention to those categories of stimulus that individuals are already predisposed to attend to and strengthens only the strongest neural connections. As a result, behaviour can become more rigid,

flexibility can be impaired and habitual behaviours are favored (Eldar et al., 2013). This model is in line with an existing theory relating the LC-NE system to neural gain (Aston-Jones & Cohen, 2005) as well as with empirical evidence showing that pupil diameter as an index of LC activity predicts exploration vs exploitation between individuals as well as across trials (Jepma & Nieuwenhuis, 2011). The model has important implications for reward learning as it can explain the observed shift from goal-directed to habitual behaviour in operant learning described earlier. Research has shown the shift to not only be facilitated by glucocorticoid but also by noradrenergic action (Schwabe et al., 2012; Schwabe & Wolf, 2011) and it is prevented when noradrenergic activity is blocked (Schwabe, Hoffken, Tegenthoff, & Wolf, 2011). That is under conditions of high gain or high NE levels, behaviour shifts from flexible, goal-directed behaviour to more rigid, habitual control of behaviour.

Taken together, previous research suggests that both glucocorticoid and noradrenergic action are critical for the development of habitual behaviours. Moreover, as discussed above the LC-NE system has further been suggested to be critical for cost-benefit analysis and action initiation. It has also been associated with uncertainty in decision-making and neural gain. Thus, while the stress response as a whole seems to support some processes related to reward learning, findings specific to the noradrenergic system and the nature of the two-part stress response suggest that the immediate and delayed stress response must have distinct effects on learning processes. A compelling theory (Joels et al., 2006) proposed that the immediate NE-driven stress response facilitates learning and memory processes while glucocorticoids raise the processing threshold for incoming information leading to learning and memory impairments. In Chapter 4, I aim to directly test this theory in the context of appetitive learning in order to better understand the role of the immediate and delayed stress response in human appetitive conditioning.

1.6 Neural circuits underlying conditioning

After establishing the unique roles of the immediate and delayed stress response on behavioural indices of associative learning, I will next review neural circuitry underlying both aversive and appetitive conditioning. As discussed in the following sections, previous research has thoroughly mapped the neural systems recruited for the development of novel associations. However, little is known about how conditioned associations are represented in different brain regions, i.e. what quality of a positive or negative stimulus becomes attached to the initially neutral stimulus – the hedonic response or sensory sensation.

1.6.1 Neural circuits underlying aversive conditioning

An extensive body of literature based on research from several laboratories performed in the 1980s provides us with a relatively clear picture of the neuroanatomy underlying fear conditioning in rodents (Fanselow, 1994; J. E. LeDoux, Cicchetti, Xagoraris, & Romanski, 1990). The amygdala plays a central role in fear conditioning, as it integrates information about the CS and US (J. E. LeDoux, 2000). In line with the animal literature, human research has revealed that fear conditioning and fear responses depend on an intact amygdala and hippocampus (LaBar, LeDoux, Spencer, & Phelps, 1995; Peper, Karcher, Wohlfarth, Reinshagen, & LeDoux, 2001). While it is well established that the ventromedial prefrontal cortex (vmPFC) is involved in extinction of learned fear association by suppressing amygdala function through interneurons (Maren & Quirk, 2004; Sotres-Bayon, Bush, & LeDoux, 2004), recent findings suggest that the prefrontal cortex (PFC) also modulates fear response in the amygdala and is essential for fear expression (Sotres-Bayon & Quirk, 2010). A systematic review revealed activation of the amygdala, insula as well as the anterior cingulate cortex (ACC)

independent of design parameters such as the specific US used (Sehlmeyer et al., 2009). Thus, while the exact role of the insula in aversive conditioning remains to be determined, convergent evidence suggests an involvement of this region. The involvement of the anterior cingulate cortex (ACC) in the processing of negative emotions in general and acquisition of fear or aversive conditioning more specifically has been demonstrated by numerous studies [for review see (Etkin, Egner, & Kalisch, 2011; Greco & Liberzon, 2016)].

In summary, the neural circuits underlying fear or aversive conditioning are well mapped but it remains to be investigated what a CS is representing once it becomes associated with the US. Given the role of the amygdala, vmPFC, insula and ACC in facilitating conditioning and playing an important role in fear expression and internal evaluation of emotional stimuli, I expect to see aspects of the US being represented by the CS in these regions.

1.6.2 Neural circuits underlying appetitive conditioning

Converging evidence from human and non-human studies suggests that the amygdala plays a key role in appetitive conditioning. The amygdala has been shown to be critical for outcome evaluation and cost estimation (Everitt, Cardinal, Parkinson, & Robbins, 2003; Wassum & Izquierdo, 2015) as well as for the development of CS-US associations and attentional modulation in reward processing (Martin-Soelch, Linthicum, & Ernst, 2007; Peck & Salzman, 2014a, 2014b). Due to its rich connections with the OFC and striatum, the basolateral amygdala is also important for integration and relay of information allowing for flexible, goal-directed behaviour (Everitt et al., 2003; Everitt & Robbins, 2005; Martin-Soelch et al., 2007). The OFC in turn receives information from the amygdala and is central for reward evaluation and outcome expectancies (O'Doherty, 2004). Besides the OFC, the anterior cingulate cortex (ACC) has been shown to be an essential node of circuitry required for normal contingency learning (Jackson,

Horst, Pears, Robbins, & Roberts, 2016) as well as for the discrimination of multiple conditioned stimuli (Cardinal et al., 2003). The striatum has been suggested to play a general role in the processing of stimulus salience (Zink, Pagnoni, Martin-Skurski, Chappelow, & Berns, 2004) and is also of major importance for the formation of habits (Yin & Knowlton, 2006).

Taken together, much is known about the involvement of different brain regions in aversive and appetitive conditioning. Though only few studies have ever combined those two types of learning and if so the unconditioned stimuli have not been of similar nature (Andreatta & Pauli, 2015) (Nasser & McNally, 2013; Segal, Disterhoft, & Olds, 1972). A central question in conditioning that remained to be unsolved to date is how novel associations are represented in the brain. More specifically, Chapter 5 aims to address the question what aspect of the US becomes associated with the CS in appetitive and aversive conditioning or what is carried forward in conditioning. Using representational similarity analysis (RSA) (Kriegeskorte, Mur, & Bandettini, 2008) - a contemporary multivariate approach to analyzing functional magnetic resonance imaging (fMRI) data - will allow me to determine the extent to which a CS reactivates the representational pattern of brain activity initially elicited by the US and moreover to investigate whether the CS takes on the sensory properties of the US or whether the CS merely reproduces the hedonic response to the US.

1.7 Thesis objectives and overview

The overarching objective of this thesis is to investigate sources of individual differences affecting the manifestation and development of affective biases through emotional learning processes as well as the representation of conditioned associations in the underlying neural

circuits. Previous research has focused on how the norepinephrine and stress system influence aversive or fear conditioning. In this thesis I examine the influence of genetic and environmental differences in the stress and norepinephrine system on the manifestation of appetitive biases and learning processes, which have been shown to be a critical factor in the vulnerability for pathological disorders such as depression or addiction. Moreover, a central unresolved question in the field of conditioning is how the development of novel associations is represented in the brain. More specifically, to date, it remained unknown which aspect of the US a CS develops to represent and carries forward in conditioning – is it the sensory stimulus properties of or the hedonic response to the US.

Thus, in the first three experimental chapters, I will present evidence for the role of the stress/norepinephrine system as one major individual difference factor in affective bias flexibility and formation. In the last experimental chapter, I will present novel findings about what is carried forward in associative learning processes.

The objective of Chapter 2 is to investigate how genetically based differences in NE availability affect the flexibility of pre-existing affective biases and how that effect might be modulated by acute stress induction. The primary hypothesis is that carriers of the deletion variant of the *ADAR2b* polymorphism will display a greater change in pre-existing affective biases. Thus, the overarching goal of Chapter 2 is to identify factors that lead to individual differences in the manifestation of affective biases.

Chapter 3 examines the effects of acute stress on both classical and operant conditioning with the hypothesis that delayed acute stress or glucocorticoid action will impair the two different forms of emotional learning.

Chapter 4 aims to demonstrate how the immediate and delayed stress response differentially modulate operant conditioning. My hypothesis is that the immediate stress response would facilitate learning while the delayed stress response will impair or not affect associative learning.

Chapter 5 examines the development of the neural representation of both aversive and appetitive conditioned associations over the course of the learning period. The multivariate analysis of fMRI data is guided by the question whether conditioned stimuli reproduce the pattern of activation elicited by the unconditioned stimulus and more specifically, what aspect of the US is carried forward in successful learning: Is it the sensory stimulus properties or the hedonic response?

Chapter 6 presents a general discussion of the findings and implications drawn from this research program.

Chapter 2: The influence of the noradrenergic/stress system on perceptual biases for reward

2.1 Introduction

Decades of research have supported the common observation that some people see the world through rose coloured glasses, others through lenses tinted with grey [e.g., (Eysenck & Eysenck, 1985)]. In general, we are all more likely to attend to and remember emotionally and motivationally salient environmental cues (Markovic et al., 2014; Pourtois et al., 2013). Yet when a cue is ambiguous in signaling reward or punishment, individuals differ in the habitual tendency to interpret information as negative or positive (Derryberry & Reed, 1994).

The norepinephrine (NE)/stress system is a key factor in sensitivity to emotionally or motivationally salient information as well as in the emotional/motivational learning processes that give rise to such biases [for review see (Ehlers & Todd, 2017b)]. Previous research has implicated *ADRA2b*, a common neurogenetic variation in the norepinephrine system, in affective biases in attention and subjective perception (Todd, Ehlers, et al., 2015; Todd, Muller, et al., 2013b). An outstanding question concerns whether more pronounced affective biases in carriers of a deletion variant of the *ADRA2b* gene result from the influence of NE on learning. Specifically, whether putative differences in NE availability influence flexibility in shifting biases based on experience. Thus, a goal of the present study was to examine the role of *ADRA2b* in bias flexibility.

Acute stress may also play a role in bias flexibility – either alone or in interaction with differences in *ADRA2b* variant. Acute stress leads to the activation of two sequentially-linked stress systems: Immediate activation of a fast-acting stress system leads to a release of mostly catecholamines such as norepinephrine and dopamine (Schwabe, Wolf, et al., 2010). This early

phase is followed by subsequent downstream activation of a glucocorticoid (cortisol in humans) pathway, typically leading to an elevated processing threshold for incoming information (Herman, McKlveen, Solomon, Carvalho-Netto, & Myers, 2012; Roozendaal, McEwen, & Chattarji, 2009). My own work has demonstrated that delayed acute stress can impair reward learning in healthy young adults (Ehlers & Todd, 2017a), which should reduce the initial formation of an attentional bias. Whether acute stress influences the flexibility of pre-existing affective biases remains to be investigated.

Thus, the overall goal of this pre-registered study (see Appendix A) was to examine the role of the norepinephrine (NE)/stress system in bias flexibility. Specifically, I wished to examine effects of naturally occurring differences in NE function and effects of acute stress on the flexibility of emotional judgments.

In order to examine the role of the NE/stress system in affective bias flexibility, healthy young adults were genotyped for the *ADRA2b* polymorphism and were exposed to a stress induction procedure with a task that assessed bias flexibility by probing affective bias before and after a training procedure (Penton-Voak et al., 2013). In the training procedure, I capitalized on facial emotion adaptation effects by repeatedly exposing participants to unambiguously angry faces. Bias flexibility was then assessed by the change in bias from the initial to the post-adaptation bias probe. Stress was induced before the initial bias probe using the commonly employed socially evaluated cold pressor test (SECPT) (Schwabe, Haddad, & Schachinger, 2008).

In this pre-registered study (see Appendix A), I predicted that the adaptation effect pushing face judgments in a positive direction would be more pronounced in *ADRA2b* deletion carriers compared to non-carriers following training, indexing greater flexibility linked to

putatively greater NE availability. I further predicted that greater initial NE activity in deletion carriers would be potentiated by stress induction leading to an enhanced adaptation effect in deletion carriers.

2.2 Materials and Methods

2.2.1 Participants

266 participants (192 females, mean age: 21.0 ± 3.9 years) took part in the experiment. All participants indicated that they were either of European or East-Asian descent. All participants were compensated for their participation with course credit. Sample size was based on power analyses included in the pre-registration protocol (see Appendix A). Power analysis for effects of *ADRA2b* was based on effect sizes found in previous studies comparing emotional processing in *ADRA2b* carriers and non-carriers (ηp of .05) (Rasch et al., 2009; Todd, Muller, et al., 2013b). For sufficient power for a repeated-measures analysis of variance (ANOVA) with *ADRA2b* and stress as between-subject factors we required a minimum sample size of 252. Data collection was continued until the end of the academic term in which the minimum was reached.

Participants were asked not to eat, consume alcohol or caffeine and exercise two hours before the experiment due to its known effects on the stress response (Kudielka, Hellhammer, & Kirschbaum, 2007). Participants were randomly assigned to stress and control conditions (129 and 137 participants respectively). The study was approved by the Human Research Ethics Board of the University of British Columbia.

2.2.2 Materials

2.2.2.1 Stimulus presentation. The MATLAB (The MathWorks, Natick, Massachusetts, USA) toolbox Cogent 2000 was used for all stimulus presentation.

2.2.2.2 Facial stimuli. Stimuli subtended a visual angle of approximately 15° x 19°. All stimuli were emotional faces taken from the NimStim Face Stimulus Set (Tottenham et al., 2009). Using morphing software (Abrosoft Fantamorph, Version 5.4.5), faces with happy and angry expressions of two females (Caucasian and East Asian) were morphed into two 15-image continua for the bias probe. Ten individual faces (5 females) from all ethnic face categories (Asian, Caucasian, African) in the NimStim set, all displaying angry expressions, were used in the adaptation phase.

2.2.2.3 Questionnaires. Participants were asked to complete a battery of questionnaires in order to control for possible interactions between psychopathology, life experience, and personality with task performance and stress response. In addition to a demographics questionnaire, we administered the Childhood Trauma Questionnaire (CTQ) (Bernstein et al., 1994), the State-Trait Anxiety Inventory (STAI) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), the Liebowitz Social Anxiety Scale (LSAS), Beck Depression Inventory (BDI) (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), as well as the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegan, 1988).

2.2.3 Procedure

2.2.3.1 Overview. After obtaining written informed consent, initial saliva samples for baseline measures of stress indicators were acquired (Figure 2.1). This was followed by the SECPT (described in more detail below) in either a stress or control condition. The three-minute stress

induction/control procedure was followed by a second saliva sample. Successful stress induction was further assessed by the administration of the SECPT questionnaire – a three-item questionnaire measuring the subjective stress response (Schwabe et al., 2008). Participants were further asked to fill out a battery of questionnaires in order to control for individual differences that could potentially influence stress responses or operant conditioning performance. In order to capitalize on effects of cortisol (delayed stress response) on behaviour, the operant task started 20 minutes after the end of the SECPT (Schwabe et al., 2008). After participants finished the task (about 60 minutes after the SECPT), the third and last saliva sample was taken. If participants did not complete all questionnaires in the 20-minute period before the learning phase, they finished them before the debriefing.

2.2.3.2 Stress procedure. In the stress condition, elevated stress levels were induced with the SECPT (Schwabe et al., 2008). First, participants were informed that their faces would be videotaped during the upcoming test for future evaluation of their facial expressions by researchers. Participants were then asked to put their non-dominant hand in ice water ($0 – 4\ ^{0}C$). They were told to keep the hand in the water for as long as possible while looking straight into the camera. The experimenter observed the participant at all times and recorded the time period during which each participant's hand remained in the water. After three minutes, participants were instructed to remove their hands from the water if they had not done so before. In the control condition the ice water was replaced with warm water ($35 – 37\ ^{0}C$) and participants were neither videotaped nor watched by the experimenter. They were also instructed to keep their hand in the water while the experimenter was present in the room.

2.2.3.3 SECPT questionnaire. To obtain a measure of subjective, psychological stress responses participants were asked to rate how stressful, painful and unpleasant the SECPT was

using a ten-point scale ranging from 1 (“not at all”) to 10 (“extremely”). The questionnaire was administered immediately after stress induction.

2.2.3.4 Salivary cortisol analysis. Saliva was collected pre SECPT, immediately post SECPT and post task (~60 min after stress induction) with a Salivette collection kit (Sarstedt AG & Co., Nümbrecht, Germany) and stored at -20 ^{0}C until the biochemical analysis of salivary levels of free cortisol. Analysis employed a luminescence immunoassay (IBL GmbH, Hamburg, Germany) performed by the lab of Prof. Dr. C. Kirschbaum, Dresden, Germany. Inter- and intra-assay variations were below 10 %.

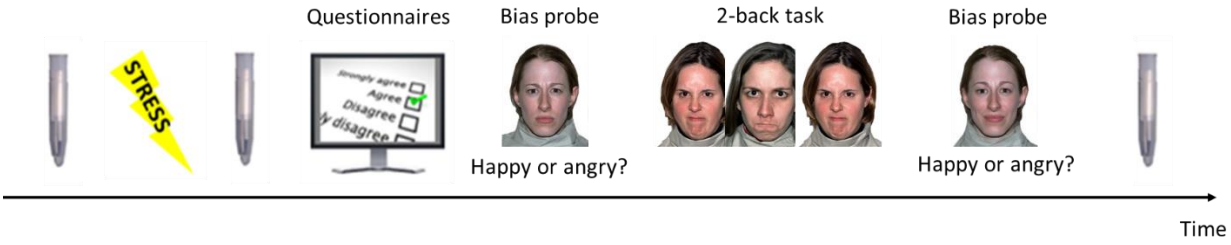

Figure 2.1. Overview of experimental procedure. Salivary cortisol samples were taken before and after stress induction by means of the socially evaluated cold pressor test (SECPT). Participants were given 20 minutes to complete several online questionnaires before starting the experimental tasks. The initial bias probe was followed by a 2-back memory task. The second bias probe task was completed before final stress measurements.

2.2.4 Bias probe

The Bias Probe task was adapted from Penton-Voak et al. (2013) . This task was performed before and after the adaptation task as an assessment of individual baseline biases in rating ambiguous faces as angry vs. happy, and the degree to which they changed with facial adaptation (Figure 2.2). Each trial began with the randomly jittered (1500 – 2000 ms) presentation of a fixation cross followed by the display of a face (1 of 15 frames taken from the continuum of emotional faces ranging from unambiguously happy to unambiguously angry). A

mask of visual noise was presented for 150 ms before participants were asked to judge whether the face just seen was happy or angry. Each participant completed a total of 90 trials (Figure 2.2). The two sets of emotional continua consisting of 15 frames were presented three times each in randomized order. Bias Probe pre- and post-adaptation was the same task but stimuli were presented in different random order. By randomizing face presentation and including many subtly varying morph frames we ensured that participants would be unlikely to remember their previous ratings of each morph frame the second time they performed the task. The participants were reminded of the instructions before completing the post-test.

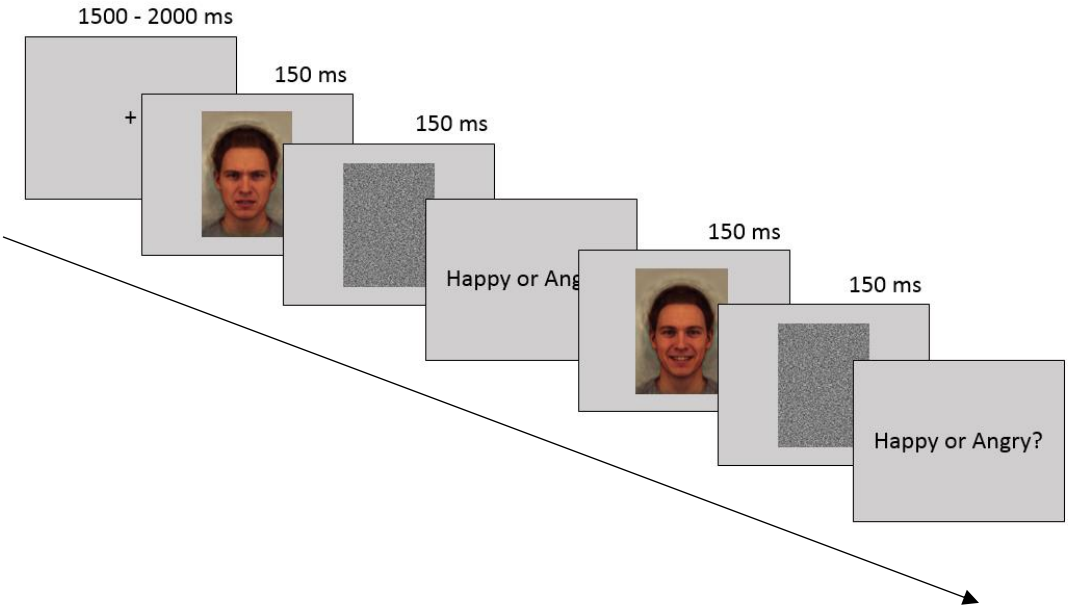

Figure 2.2. Schematic of Bias Probe task performed before and after adaptation. Adapted from Penton-Voak et al. 2013. Participants are presented with faces at different stages of parametric morphing between unambiguously happy and angry faces and are asked to make a forced-choice assessment of whether a face was happy or angry.

2.2.5 Adaptation task

The adaptation task served as a training task to shift biases towards more positive judgments of emotional expressions by capitalizing on known visual adaptation effects for facial emotion. Participants were asked to perform a 21 minute two-back working memory task where

all faces showed angry expressions (Figure 2.1). Whereas adaptation effects result from repeated viewing of the faces, the working memory task ensured participants paid attention to each individual face (Rhodes et al., 2011). The task exploited the well-documented phenomenon of visual aftereffects where, after repeated exposure to one category of visual stimulus, an ambiguous stimulus looks more like the opposite category. In the case of facial emotions, repeated exposure to an angry face shifts perception of a neutral or ambiguous facial expression to more happy (Webster et al. 1998). The two-back task was employed to ensure attention to the facial expressions: Faces were presented in random order and participants indicated via button press whether a face was the same as that two faces back. Each face was displayed for 2000 ms with 20 ms intertrial intervals. Each participant completed 13 adaptation blocks, each block consisting of 48 trials containing eight targets each.

2.2.6 Genotyping

A saliva sample (~ 1 mL) was collected from each subject in an Oragene OG-500 DNA kit (DNA Genotek, Ottawa, ON). *ADRA2b* 9bp deletion was assayed using a PCR followed by Sanger sequencing. A total of 50 ng genomic DNA was combined with 1xAmpliTaq Gold 360 buffer, 2.0 mM Magnesium Chloride, 360GC Enhancer 4ul, 200 uM dNTPs, 0.5 uM forward primer ACGAAGGTGAAGCGCTTCT and 0.5 uM reverse primer GGCCAGAAGGAGGGTGTTT, AmpliTaq Gold 360 DNA Polymerase 0.625 U/reaction for a total volume of 25ul in a 96 well plate. Initial denaturation was at 95 °C for 8 min, followed by 38 cycles at 95 °C for 50 s, 60 °C for 30 s and 72 °C for 50 s and a final extension step of 7 min at 72 °C. The PCR products were cleaned up by ExoSAP-IT Express and analyzed by sequencing (ABI 3130, Applied Biosystems).

2.3 Results

Genotype frequencies fell within the Hardy-Weinberg equilibrium ($X^2 = 4.49$, $p > .05$). For all analyses, based on previous research (de Quervain et al., 2007; Todd, Ehlers, et al., 2015; Todd, Muller, et al., 2013b; Todd et al., 2014), homozygote and heterozygote *ADRA2b* deletion carriers were treated as a single group due to the low number of homozygotes. The sample consisted of 139 deletion carriers and 109 non-carriers. Genotyping did not yield conclusive results for 18 participants. For all statistical analyses Greenhouse-Geisser corrections were applied if sphericity was violated. All analyses were performed with IBM SPSS Statistics 21.

2.3.1 Stress manipulation

The effect of stress induction was assessed by both subjective ratings and salivary cortisol. On average, participants in the stress group kept their hands in ice water for 151.8 ± 49.8 s, while all participants in the control group kept their hands in water for the full 180 s. The analysis of subjective stress ratings confirmed that, compared to the control group, participants in the stress group perceived the SECPT as more stressful, $t(204.02) = -11.26$, $p < .001$, painful, $t(169.16) = -20.68$, $p < .001$, and unpleasant, $t(243.41) = -17.36$, $p < .001$. Individual stress ratings indicated successful stress induction in all participants. As a result, all participants were included in the following analyses (Figure 2.3).

Salivary cortisol was analyzed in a mixed ANOVA with time point (pre SECPT, post SECPT, post task) as a within-subject factor and stress condition (stress vs. control) as between-subject factors. The analysis of cortisol showed a main effect of time, $F(1.47, 372.51) = 73.83$, $p < .001$. No main effect of stress condition or interaction with time was found, presumably due to the fact that the last sample was taken too long after stress induction (approx. 60 minutes). While

in some previous publications researchers chose to exclude participants based on their cortisol response (Miller, Plessow, Kirschbaum, & Stalder, 2013), others, including myself, chose to maximize sample size by including as many participants as possible (Ehlers & Todd, 2017a; Schwabe & Wolf, 2009). Importantly, the pattern of results was the same if some participants were excluded based on a more conservative subjective stress response criterion (Figure 2.3).

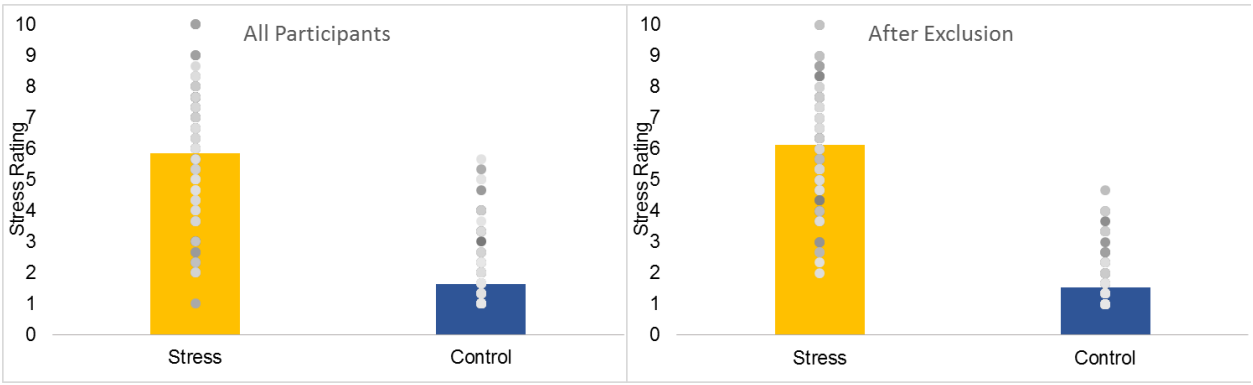

Figure 2.3. Individual subjective stress ratings for stress and control group indicate successful stress induction in all individuals. Displayed are individual averages derived from three different ratings: unpleasantness, stressfulness, painfulness. On the left, ratings of all participants are presented. On the right, a total of 14 participants was excluded due to the fact that their ratings fell within one standard deviation from the mean of the other group. The main analyses were rerun with the reduced sample in order to ensure the results are not due to individual differences in stress response. The pattern of results is the same in both samples.

2.3.2 Behavioural Results

Exploratory analyses were performed to confirm that sex (p = .432) and racial identity (Caucasian vs. Asian, p = .291) had no significant effect on behavioural measures. In addition, exploratory correlations were performed between questionnaire measures and bias scores. Only state anxiety showed robust correlations with bias, and all subsequent analyses were performed both with and without state anxiety scores as a covariate.

Moreover, exploratory analyses showed no difference between genotype or stress groups or any interactions with regard to working memory performance in the adaptation task. It should be noted that, whereas previous studies have found interactions between *ADRA2b* and emotional effects on working memory (Mammarella et al., 2016), here there was no emotion manipulation as all of the stimuli were negatively valenced.

2.3.3 Main Analysis

As stipulated at pre-registration (see Appendix A), I performed an analysis where I assessed the probability of faces being rated as happy frame by frame (that is, for each degree of morphing from 100% angry to 100% happy) pre- and post-training by *ADRA2b* genotype and stress condition. A mixed analyses of variance (ANOVA) with bias probe (pre- and post-adaptation) and frame (15 frames per continuum) as within-subject factors and *ADRA2b* genotype (deletion and no deletion) and stress condition (stress and control) as between-subject factors.

There was a main effect of frame, $F(5.57, 1357.88) = 3097.33$, $p < .001$, indicating that, unsurprisingly, participants were sensitive to the amount of emotion signal in the faces. Unambiguously happy faces were most likely to be rated as happy with decreasing probabilities as the continuum approached unambiguously angry faces. A main effect of test (pre- vs post-training), $F(1, 244) = 11.48$, $p = .001$, further revealed that the adaptation procedure led to an overall shift of affective bias toward judging faces as more positive. These main effects were qualified by a test by frame interaction, $F(8.81, 2149.94) = 17.20$, $p = .001$ (Figure 2.4a). Pairwise contrasts revealed shifts in judgment for frame 1 (the most angry face), frames 4-8, and

11-15 (the most happy faces) (*ps* < .05). Taken together, the adaptation procedure resulted in a robust shift of affective bias in a positive direction.

Importantly, there was an *ADRA2b* by frame interaction, $F(14, 3416) = 3.00$, $p < .001$, such that deletion carriers rated a greater proportion of faces as happy for frames towards the middle of the morph continuum (frames 5 and 6, *ps* < .05) (Figure 2.4b). However, I did not observe the hypothesized *ADRA2b* by frame by test interaction that would have indicated that carriers of the deletion variant shifted their bias more flexibly than non-carriers. Rather, they simply showed a stronger version of the pattern observed in all participants: A slightly positive bias pre-training that became more positive after adaptation. There was no effect of stress $F(1, 244) = 2.55$, $p = .112$, no two-way interaction between stress and test $F(1, 244) = .001$, $p = .972$ and no three-way interaction between stress, test and frame $F(14, 3416) = .716$, $p = .761$.

When controlling for state anxiety the same pattern of results reported above was found: The analysis revealed a main effect of test, $F(1, 238) = 11.44$, $p = .001$, and frame, $F(5.56, 1324.01) = 3024.71$, $p < .001$ as well as an interaction, $F(8.82, 2099.33) = 17.45$, $p < .001$. Similarly, the frame by *ADRA2b* interaction, $F(14, 3332) = 3.08$, $p < .001$, showed a stronger bias towards the positive for frames 5 and 6 for deletion carriers relative to non-carriers (*ps* < .05).

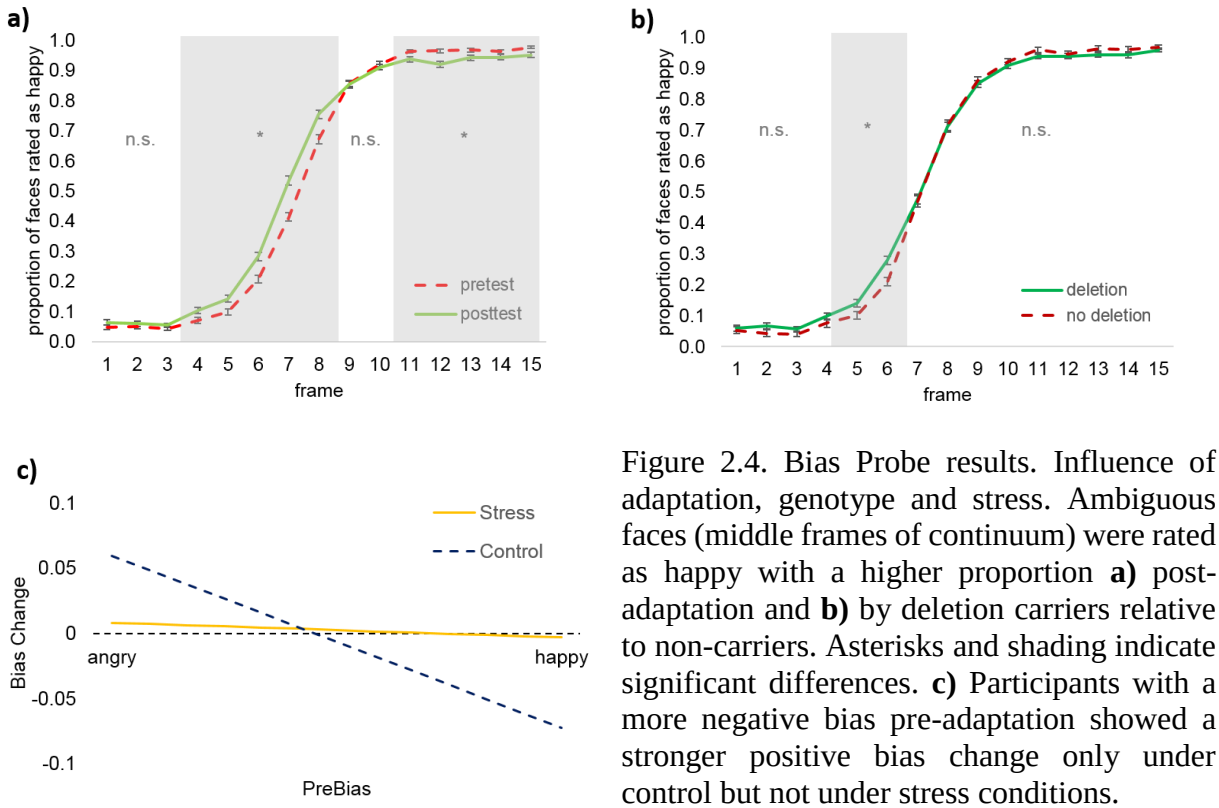

Figure 2.4. Bias Probe results. Influence of adaptation, genotype and stress. Ambiguous faces (middle frames of continuum) were rated as happy with a higher proportion **a)** post-adaptation and **b)** by deletion carriers relative to non-carriers. Asterisks and shading indicate significant differences. **c)** Participants with a more negative bias pre-adaptation showed a stronger positive bias change only under control but not under stress conditions.

2.3.4 Follow-up analysis

I reasoned that, as previously observed in rodents (Enkel et al., 2010), hypothesized effects of the stress manipulation may have been obscured because effects of stress on bias change depended on the degree of initial bias. That is, it may be that stress moderated effects of initial predisposition to rate ambiguous faces as happy or angry on bias change. In order to test the hypothesis that change in affective bias through adaptation depends on the baseline bias, and that this effect is moderated by stress, I performed a moderation analysis (Preacher & Hayes, 2004). In a regression model, pre-adaptation bias was defined as the predictor variable. This was operationalized as the proportion of trials rated as happy at frame 7, the ambiguous frame at which participants were split evenly around the median into those who were positively and negatively biased. Stress condition was included as a moderator variable. The analysis revealed

that pre-adaptation bias predicted the degree of bias change b = -.05, $t(244)$ = -2.73, p = .007, such that those with an initial bias toward judging faces as angry showed greater change in the positive direction. Importantly, this was qualified by a pre-adaptation bias by stress interaction, b = .04, $t(244)$ = 2.07, p = .04, indicating a moderation effect: Pre-training bias was a significant predictor of bias change for those who showed a more negative bias pre-adaptation changing in a positive direction in the control group, b = -.09, $t(244)$ = -3.17, p = .002, but not in the stress group, b = -.01, $t(246)$ = -.51, p = .613 (Figure 2.4c). Thus, stress diminished the effect of pre-adaptation bias on bias flexibility that allowed the more negatively biased participants to show greater change (Figure 2.4c).

2.4 Discussion

The aim of the current study was to examine the role of the NE/stress system in affective bias flexibility along a continuum ranging from stimuli signaling social punishment to those signaling social reward. The results revealed an overall robust adaptation effect in healthy young adults, such that judgments of facial emotion became more positive following repeated exposure to angry faces. Although it did not predict bias flexibility, carrying the deletion variant of the *ADRA2b* genotype was associated with a tendency to rate faces as more positive overall. Acute stress administered by means of the SECPT moderated the relation between pre-adaptation bias and bias flexibility, such that pre-adaptation bias was a significant predictor of bias change in the control group only. Specifically, participants with a more negative pre-training bias showed the strongest change in a positive direction - an effect that was abolished under stress.

Our finding that an implicit training process involving repeated exposure to angry faces effectively shifted judgments in a more positive direction replicates and extends previous

research. It has been demonstrated that humans show adaptation to various different features of faces such as gender or ethnicity (Webster & MacLeod, 2011). For example, an ambiguous face containing both female and male features is perceived as more female after being exposed to male faces (Webster, Kaping, Mizokami, & Duhamel, 2004). This adaptation effect has also been shown for emotional expressions such as disgust and surprise (Webster et al., 2004). In the present study, I used a morphing paradigm in which ambiguous facial expressions were perceived as happier (and less angry) after being exposed to a working memory task using angry facial expressions. To date, this training effect has been observed in a relatively small sample of aggressive adolescents (Penton-Voak et al., 2013) and in younger children (Picardo, Baron, Anderson, & Todd, 2016). In the current study, I was able to replicate this finding in a much larger population of healthy university students, validating and generalizing previous findings. In the adolescent sample, the change in bias remained one week after training (Penton-Voak et al., 2013). Although I did not examine long-term effects of training in this study, future research can determine how long such effects may endure.

My main focus, however, was on effects of the NE/stress system on modulation of affective judgments by this short-term implicit learning process. Based on previous research demonstrating a role of alpha2b noradrenergic receptors in emotional learning (Moriceau & Sullivan, 2004), and direct evidence that carriers of the deletion variant of the *ADRA2b* polymorphism show greater flexibility in working-memory-related affective biases (Mammarella et al., 2016), I hypothesized that *ADRA2b* deletion carriers would show more pronounced adaptation effects relative to non-carriers. However, the present study indicated no difference in bias shift between deletion carriers and non-carriers. Nonetheless, deletion carriers demonstrated a stronger overall tendency to rate ambiguous faces as more positive relative to non-carriers. An

enhanced positivity bias is consistent with previous research showing a working memory advantage for positive items in deletion carriers (Mammarella et al., 2016), and suggests that in this context deletion carriers perceive ambiguous faces as more rewarding. Thus, the current results confirm *ADRA2b*-dependent exaggeration of typically-observed affective biases, as demonstrated by multiple studies [e.g., (de Quervain et al., 2007; Todd, Ehlers, et al., 2015; Todd, Muller, et al., 2013a)] and confirmed in a recent meta-analytic review (Xie et al., 2018); however, they suggest that putative differences in NE availability do not influence the degree of flexibility or change in subjective perception of rewarding information induced by implicit learning processes. Divergence from previous studies investigating classical conditioning (Moriceau & Sullivan, 2004) and working/recognition memory (Mammarella et al., 2016) suggest that effects of alpha2b receptor activity on learning likely differ across learning processes (Xie et al., 2018).

A large body of literature has focused on the effects of acute stress on *explicit* learning. It has been suggested that when there is a delay between stress induction and a cognitive task, as in the present study, performance is typically impaired (Joels et al., 2006). Here I found no overall influence of acute stress on pre-existing biases or bias change. It should be noted that I cannot rule out that individual differences in the cortisol response might add to the variability explaining the current findings. Yet, as has been suggested by studies of non-human animals, effects of stress manipulation on adaptation can be masked by differences in initial bias (Enkel et al., 2010). Indeed, follow-up analyses showed that participants with more negative initial biases showed stronger positive bias change. Importantly, this effect was only visible in the control but not in the stress group. Thus, the results are consistent with previous research on stress and (emotional) learning in which acute stress induced with a delay impaired learning. Taken

together, the present study adds to the field by demonstrating that acute stress can also affect short-lasting perceptual effects.

Previous research has also demonstrated interactions between *ADRA2b* genotype and acute stress, such that amygdala activity was enhanced for deletion carriers only under acute stress (Cousijn et al., 2010). In contrast, I did not find any interactions between genotype and stress in the present study. In particular, I hypothesized that acute stress induction would amplify the putative difference in NE availability and hence in behaviour between *ADRA2b* deletion carriers and non-carriers. There are several possible explanations for not seeing the hypothesized effect. First and foremost, there may have been no differences to be amplified, as I observed no behavioural differences between deletion carriers and non-carriers with respect to bias flexibility. Moreover, acute stress is likely to affect the whole norepinephrine system and hence all receptor subtypes in the same way. Thus, it could be that an overall increase in NE availability affecting both inhibitory and excitatory receptors might cancel out and not result in specific enhancements in deletion carriers. Finally, due to the timing of the stressor relative to the adaptation task, the present experiment capitalized more on the effects of the slow stress response involving the release of cortisol and less on the immediate NE-driven effects (D. J. F. de Quervain, B. Roozendaal, & J. L. McGaugh, 1998; Joels et al., 2006). Thus, interactions with stress may be observed if training occurred immediately after stress induction.

While subjective stress ratings indicated successful stress induction, it should be noted that there were no significant differences in cortisol levels between stress and control group. Initial measurements were taken right before and after stress induction, where no group differences were expected (Schwabe et al., 2008). The third measurement was taken at the end of the experiment, which was about 60 minutes after stress induction. In additional studies performed

under my supervision (see following chapters), the post task measurement was taken approximately 40 minutes after stress induction and elevated cortisol levels were reliably observed. Hence, I speculate that the reason for that result is that I missed capturing peak activation in cortisol about 25-30 minutes after stress induction (Schwabe et al., 2008). Nevertheless, I am confident that the tasks were performed under the influence of acute stress due to the subjective ratings, the reliability of the induction procedure (Schwabe & Schachinger, 2018) and our own experience with it (see following chapters).

In conclusion, the current study showed that a common genetic variation putatively influencing norepinephrine availability was associated with subjective perceptions of ambiguous stimuli as more rewarding. Moreover, delayed effects of acute stress diminish the positive change in affective bias predicted by initial biases consistent with previous studies showing detrimental effects of delayed stress on appetitive learning.

In this chapter, I established that the LC-NE stress system plays a role in the perception of reward and the flexibility of the underlying subjective biases. More generally speaking, the current chapter demonstrates an important role of individual differences in the manifestation of affective biases. What remains to be answered is how those affective biases in perception, attention or memory develop in the first place. It is of great interest to identify factors of individual difference that lead to individual differences in the ability, speed or strength of developing novel emotional associations that in turn inform affective biases in different cognitive domains. Thus, in the next chapter I turn to the question how activation of the different stress systems as one factor of individual difference can facilitate or impair the formation of new affective biases via different emotional learning processes.

Chapter 3: Effects of acute stress on human appetitive operant and classical conditioning

3.1 Introduction

Implicit affective biases in attention, perception and memory do not only influence what we perceive or pay attention to, but are themselves the product of learning and memory. For example, high-arousal associative learning experiences found in certain anxiety disorders mirror effects found in controlled laboratory experiments using fear conditioning, and complement a wide literature linking fear conditioning to anxiety disorders (Lissek et al., 2008; Lissek et al., 2005; Lissek et al., 2009; Wilker et al., 2014). On the other end of the valence spectrum, attentional biases for substance-related stimuli, or cues, which predict craving in addiction can also be created through classical conditioning processes (B. A. Anderson, 2016a; Field & Cox, 2008). Thus, considerable evidence suggests that attentional biases towards specific categories of salient stimuli develop through associative learning processes. Yet outstanding questions remain about how acute stress influences emotional learning processes that can underlie the formation of attentional biases. The experiments in this chapter will examine effects of acute stress on components of learning in both classical and operant conditioning.

The ontology of addiction is often described as a series of associative learning processes (Everitt & Robbins, 2005) involving both operant and classical conditioning. Addiction (e.g. drug use) is thought to be influenced by operant conditioning in the following way: Whereas initial drug use is driven by a voluntary goal-directed process reinforced by the rewarding properties of the drug, later stages of addiction are characterized by habitual and compulsive drug use that continues despite adverse consequences (Everitt & Robbins, 2016). Pavlovian conditioning has been shown to interact with these operant conditioning processes through

simple stimulus-outcome interactions, as drug-related cues predicting reward can enhance craving and compulsive tendencies observed in addicts. Thus, identifying the role of factors that facilitate initial operant and Pavlovian learning processes, which occur before habitual behaviours are established, is crucial for understanding individual variability in vulnerability to addiction.

Stress has long been known to be a major factor in the inception and development of addictive behaviour, elevating drug self-administration and promoting relapse (Piazza & Le Moal, 1998; Sinha, 2008). Building on these studies, research in humans has focused on effects of stress on favoring habitual over goal-related behaviour. In a series of studies in human subjects, Schwabe and colleagues (2009, 2010b) exposed participants to acute psychophysiological stress or a control condition either before or after operant training tasks. Participants in the stress group showed more persistent habitual performance even in the absence of reward both when stress was induced before and after contingencies were learned (Schwabe & Wolf, 2009, 2010b). A recent study (Pool et al., 2015) further employed a Pavlovian-Instrumental Transfer (PIT) task to show that stress increases the craving for a rewarding outcome without affecting the pleasure of consuming it. Participants were exposed to an acute stress or a no-stress control condition after the learning phase. Here the stress group mobilized more effort in response to the now-unrewarded Pavlovian stimulus than the control group, which was interpreted as increased cue-triggered 'wanting' (Pool et al., 2015). As this study focused on effects of stress on transfer, outstanding questions remain about effects of stress on learning processes that precede the transfer effect, when simple associations between an action or a stimulus and a rewarding outcome are first acquired. Thus, the goal of the present study was to

examine the effects of acute stress on the initial operant conditioning and Pavlovian conditioning stages of this 3-stage PIT task.

Based on previous research, there are a number of ways in which acute stress could influence initial reward learning. First, there is research suggesting that stress may have opposing effects on different phases of learning and transfer, reducing initial associative learning while enhancing reliance on habit once a habit has been formed. For example, a body of non-human animal literature suggests that stress reduces appetitive learning (Pielock et al., 2013; Shors, 2004). Yet results in humans have been more equivocal (Schwabe & Wolf, 2009). If stress has opposing effects on learning, given previous findings that stress enhances habit formation (Pool et al., 2015; Schwabe, Tegenthoff, Hoffken, & Wolf, 2010; Schwabe & Wolf, 2011), we would expect it to impair initial associative learning processes.

One reason for inconsistent findings with regard to effects of stress on learning may be that its effects on learning and memory do not depend only on the learning phase. They are also markedly influenced by the timing of the stressor relative to learning [for review see (Joels et al., 2006)]. An acute stressor activates two stress systems: 1) Activation of the fast-acting stress system has been proposed to facilitate cognitive processes at the time of stress induction [for review see (Schwabe, Wolf, et al., 2010)]. 2) Glucocorticoids (cortisol in humans) activate a gene-mediated pathway that might lead to an elevated processing threshold for incoming information (Herman et al., 2012). For consistency with the Pool et al. (2015) study, we aimed to examine effects of delayed stress on associative learning. As activation of the glucocorticoid pathway suppresses learning, we would again expect operant and Pavlovian learning processes to be suppressed by delayed stress.

Third, stress may not only differentially affect distinct stages of habit learning, but may also have different effects on learning rate and reward sensitivity as two independent components of reward-based learning (Huys, Pizzagalli, Bogdan, & Dayan, 2013). Previous research focusing on effects of stress on depression-related anhedonia suggests a detrimental effect of stress on reward responsiveness linked to learning - at least in some participants. When used as a stressor, threat of shock has been found to reduce preference for a high probability over a low-probability reward (Bogdan & Pizzagalli, 2006). Other studies have observed such a pattern of reduced reward responsiveness under stress *only* in participants high in stress reactivity (Berghorst, Bogdan, Frank, & Pizzagalli, 2013) or behavioural inhibition (Cavanagh, Frank, & Allen, 2011). Yet, notably, the opposite pattern of improved reward responsiveness has been observed in those low in behavioural inhibition (Cavanagh et al., 2011). Thus, I also aimed to examine effects of stress on both learning rate and reward sensitivity.

Taken together, previous studies suggest that the effects of acute stress on reward learning depend on the learning phase (acquisition vs transfer), the relative timing to the stressor (immediate vs delayed) as well as the reward learning component (learning rate vs reward sensitivity). Thus, the goal of the present study was to investigate the effect of *delayed* stress on initial stages of active operant and passive Pavlovian learning using a task that allows me to assess reward sensitivity. In particular I wished to determine the effects of stress on formation of associations that are distinct from, but contribute to, habitual behaviours and Pavlovian-to-instrumental transfer effects (Pool et al., 2015; Talmi et al., 2008). For this reason, I examined effects of acute stress on behaviour in the operant and classical conditioning tasks that comprised the first two stages of the 3-stage human PIT task described above (Talmi et al., 2008). These tasks are distinct from those employed in many studies of operant conditioning in that the

associations learned are simple and learning occurs very rapidly (Pool et al., 2015; Talmi et al., 2008). For example, the association of an action and reward is learned after the first few encounters — very much as when a drug is taken for the very first time and the associated pleasurable experience is remembered immediately. Another advantage is that it allows me to investigate the willingness to exert physical effort rather than simply testing cognitive abilities. This is central to my goal of examining reward sensitivity because it allows me to measure how much work participants are willing to put into the task given a certain reward and whether this is affected by stress.

In the present study, two separate experiments investigated effects of acute stress on operant and Pavlovian learning as in (Pool et al., 2015). In Experiment 1a and 1b healthy undergraduate students performed a simple operant conditioning task in which they learned to squeeze a hand-grip to obtain a low (Experiment 1a) or high (Experiment 1b) monetary reward (Talmi et al., 2008). In Experiment 2 participants performed a simple Pavlovian learning task in which coloured fractal patterns were associated with monetary reward. Both procedures were performed either following acute psychophysiological stress or in a stress-free control condition. For stress induction, participants were exposed to the commonly employed socially evaluated cold pressor test (SECPT) (Pool et al., 2015; Schwabe et al., 2008). I hypothesized that the delayed effects of acute stress during the first encounter of an action-outcome contingency would a) decrease the effort and frequency with which the behaviour is performed to obtain that reward (that is reward sensitivity is reduced), and b) influence the extent of appetitive Pavlovian learning.

Experiment 1

3.2 Materials and Methods

3.2.1 Participants

Prior to data collection, a power analysis was performed in order to determine the number of subjects. Assuming an effect size of $\eta^2 = .15$ based on previous research (Pool et al., 2015) and a repeated measures ANOVA, approximately 190 participants were necessary. A sample size of at last 200 allows for attrition, hence data collection was continued until the end of the academic term in which the minimum was reached.

214 participants (155 females, mean age: 21.59 ± 3.63 years) took part in Experiments 1a and 1b (102 and 112 participants respectively). All participants were compensated for their participation by course credit. Participants were asked not to eat, consume alcohol or caffeine and exercise two hours before the experiment. Testing was completed between 9AM and 6PM. Participants were randomly assigned to stress and control conditions (103 and 111 participants respectively). The study was approved by the Human Research Ethics Board of the University of British Columbia.

3.2.2 Materials

3.2.2.1 Stimuli and apparatus. For all stimulus presentation, the MATLAB (The MathWorks, Natick, Massachusetts, USA) toolbox Cogent 2000 was used.

3.2.2.2 Operant Conditioning. The visual stimuli viewed in this experiment were images of a thermometer with a real-time changing mercury level displayed on a gray background on a computer screen to indicate grip force and an image of a Canadian quarter to indicate reward (Figure 3.1). A handgrip apparatus was connected to a grip-force transducer (Powerlab, AD

Instruments, Colorado Springs, CO, USA) that converted grip pressure into a voltage output. Variation in compression by the handgrip resulted in a voltage signal that was proportional to the force exerted. The dynamic value of the recorded signal provided participants with a real-time visual feedback that reflected the force on the handgrip, which was displayed as the "mercury" level moving up and down within the thermometer on the screen. Grip strength data (LabChart, AD instruments) was measured and stored in Newton.

3.2.2.3 Questionnaires. See Section 2.2.2.3

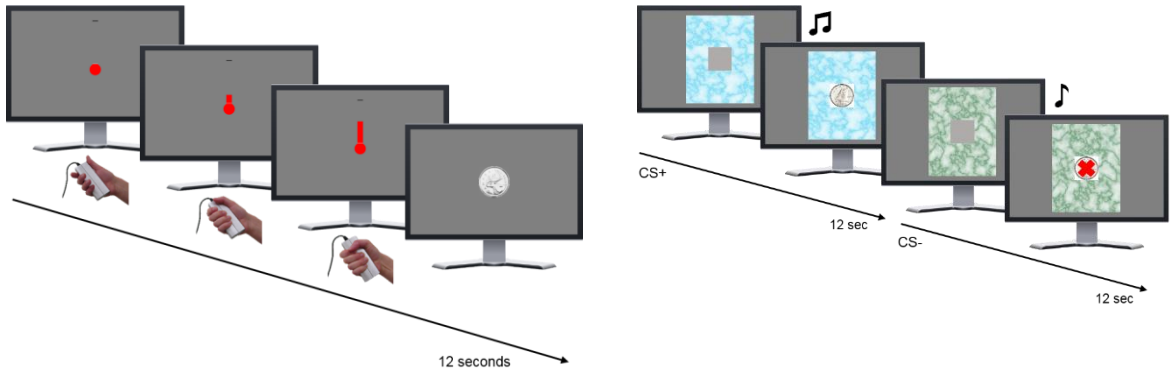

Figure 3.1. Overview of experimental design for operant and classical conditioning task. In Experiment 1, the operant conditioning task, participants squeezed a handgrip to get a monetary reward. In Experiment 2, the classical conditioning task, participants learned to associate compound stimuli (fractal pattern and tone) with reward or no reward.

3.2.3 Procedure

3.2.3.1 Overview. After obtaining written informed consent, we acquired initial saliva samples and blood pressure readings for baseline measures of physiological indicators of stress. This was followed by the SECPT in either the stress or control condition (Figure 3.2). To observe physiological reactions during stress induction we initiated continuous heart rate recording at the beginning of the SECPT. The three-minute stress induction procedure was followed by immediate blood pressure measurements and the second cortisol sample. Successful stress

induction was further assessed by the administration of the SECPT questionnaire – a three-item questionnaire measuring the subjective stress response (Schwabe et al., 2008). Participants were further asked to fill out a battery of questionnaires in order to control for individual differences that may influence stress response or operant conditioning performance. The operant task started 25 minutes after the end of the SECPT allowing cortisol to reach peak levels (Schwabe et al., 2008). Heart rate recording was stopped at this point as it is typically influenced by physical activity required for the operant task. After participants finished the task, blood pressure was measured for the last time and the third and last cortisol sample was taken. If participants did not complete all questionnaires in the 25-minute period before the learning phase, they finished them before the debriefing.

3.2.3.2 Stress procedure. For stress procedure and questionnaires see Section 2.2.3.2 and 2.2.3.3.

3.2.3.3 Heart rate. Heart rate was measured using LabChart software (AD Instruments) based on a finger pulse that was continuously measured with a pulse transducer (AD Instruments). In order to determine a baseline, heart rate was averaged within three subsequent one-minute time windows. Similarly, heart rate was measured throughout the three minute lasting stress procedure and averaged separately for three one minute time windows (Figure 3.2).

3.2.3.4 Blood pressure. Systolic and diastolic blood pressure were measured using a blood pressure monitor. Measurements were taken pre SECPT, post SECPT and post task. Data is missing for the first 30 participants.

3.2.3.5 Salivary cortisol analysis. Analysis was identical to Section 2.2.3.4.

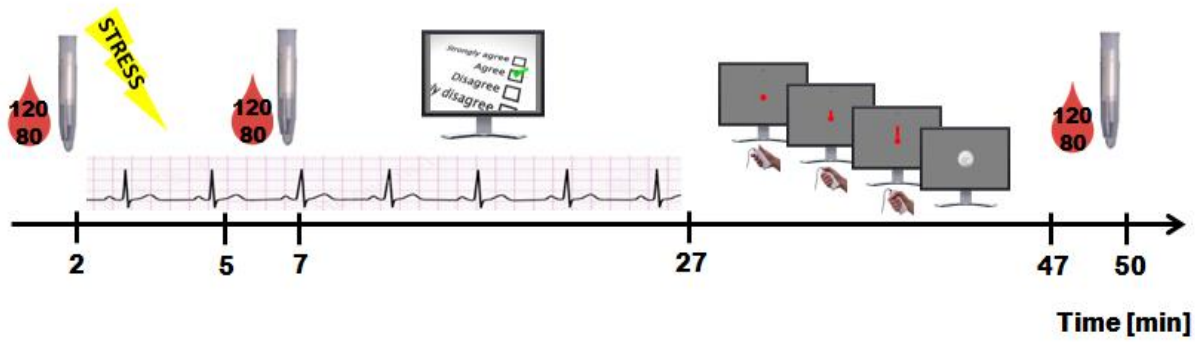

Figure 3.2. Overview of experimental procedure for Experiment 1 and 2. Blood pressure and cortisol samples were taken before and after stress induction by means of the socially evaluated cold pressor test (SECPT). Heart rate was continuously measured throughout the three minute stress test as well as while answering questionnaires. Twenty minutes after stress induction, the operant or classical conditioning task was performed followed by final blood pressure and cortisol samples.

3.2.4 Experiment 1a: Operant Conditioning Task

The operant conditioning paradigm was adapted from a Pavlovian Instrumental Transfer (PIT) task described in Talmi et al., (2008). In this procedure, participants learned to squeeze a handgrip in order to get a monetary reward (Figure 3.1). Because we wished to directly examine an earlier phase of the Pavlovian to instrumental transfer process, we designed our operant conditioning task to be equivalent to the operant conditioning task used in a previous study by Pool et al., (2015). Another advantage of this design is that it allows us to measure willingness to perform physical effort to obtain a reward. This is distinct from operant conditioning tasks that rely on learning stimulus contingencies, which largely depend on cognitive abilities. Because the task is so simple, it can be performed equally well by all participants, ensuring that differences in performance are due to effort rather than differences in cognitive ability. This allowed us to evaluate reward sensitivity as we were able to measure how much effort participants were willing to exert for the given reward.

Participants were told that they could earn CAD 0.25 per successful grip in this operant conditioning task that they would be given at end of the experiment in addition to the reimbursement for their participation. In a training trial, participants were asked to familiarize themselves with the handgrip. The grip force was visualized in real time by the mercury level displayed on the screen (Figure 3.1). Moreover, their maximum grip force was determined as criterion for their response during the main operant task by letting participants squeeze the hand grip as hard as they could. The training phase was followed by 24 operant conditioning trials each of which lasted 12 s with a 4 - 12 s fixation period as an intertrial interval (average duration 8 s). For each 12 s trial, participants were asked to squeeze the handgrip with their non-dominant hand to bring the mercury to its maximum and down again. They were told that there were up to three rewarded time windows. If they happened to reach near maximum grip force, they would gain CAD 0.25 and a coin was displayed. It was emphasized that they should decide intuitively when to squeeze the handgrip and that the displayed coins represent a real monetary reward. In fact, there were always two rewarded time windows each lasting 1 s. Participants had to reach either 50 % or 70 % of their individual maximum grip force in the rewarded time windows in order to get the reward. The criterion for the maximum force changed every second to reduce predictability.

3.2.5 Experiment 1b: Operant Conditioning with high reward

A follow-up experiment was conducted to determine whether effects of stress on operant conditioning was due to reward sensitivity. In this study we used an identical procedure to that described above, with the exception that a higher reward (CAD 1.00 per successful grip) was introduced.

3.2.6 Statistical Analysis

Two 24 x 2 mixed analyses of variance (ANOVAs) with trial as within- and stress group as between-subject factor were employed to independently test for effects of stress on operant conditioning in Experiment 1a and 1b. In a combined analysis a 24 x 2 x 2 mixed ANOVA was applied to the operant conditioning data with trial as within-subject factor and group (stress and control) and reward condition (low and high reward) as between-subject factors. Physiological data (heart rate, blood pressure and cortisol) were analyzed in a mixed ANOVA with time as within- and group (stress and control) as between-subject factors. All analyses were additionally performed with time of day - dichotomized as morning (testing between 9AM and 1PM) and afternoon (testing between 1PM and 6PM) - as a covariate. Greenhouse-Geisser corrections were applied if sphericity was violated. All analyses were performed with IBM SPSS Statistics 21.

3.3 Results

3.3.1 Control Variables

Exploratory correlations examining the relation between task performance and personality measures, state and trait anxiety, depression and childhood trauma did not reveal significant results. Furthermore, stress and control group did not differ with regard to age, sex, time of day, ethnicity and average levels of depression and anxiety.

3.3.2 Stress manipulation

3.3.2.1 Experiment 1a

The effect of stress induction was assessed by both subjective ratings and physiological measures such as heart rate, blood pressure and cortisol.

On average, participants in the stress group kept their hands in ice water for 162.64 ± 42.93 s, and participants in the control group kept their hands in water for 175.00 ± 23.98 s. Subjective stress ratings confirmed that, compared to the control group, participants in the stress group perceived the SECPT as more stressful, $t(69.07) = 8.08$, $p < .001$, painful, $t(50.18) = 14.96$, $p < .001$, and unpleasant, $t(90) = 9.84$, $p < .001$, than participants in the control group.

3.3.2.1.1 Heart rate. Analysis of heart rate (including a baseline measurement and recordings during the three minute stress induction) revealed a main effect of time, $F(1.87, 162.41) = 8.73$, $p < .001$ as well as a time by stress group interaction, $F(1.87, 162.41) = 5.48$, $p = .006$. Post hoc tests using Bonferroni correction showed that in the stress group, heart rate significantly increased in minute 1, $p < .001$, and minute 2, $p = .001$, of the stress test relative to baseline. Thus, only the stress group showed a stark increase in heart rate as a result of stress induction.

3.3.2.1.2 Blood pressure. For systolic blood pressure the analysis revealed a main effect of time, $F(2, 126) = 8.17$, $p < .001$, showing that systolic blood pressure dropped after the SECPT in both groups.

3.3.2.1.3 *Cortisol.* The analysis of cortisol showed a main effect of time, $F(1.22, 63.50) = 4.81$, $p = .010$, as well as a time by stress group interaction, $F(1.22, 63.50) = 17.12$, $p < .001$. Post-hoc comparisons revealed that cortisol levels measured 50 minutes after stress induction were significantly elevated relative to pre-stress measurements in the stress, $p = .001$, but not in the control, $p = .252$, group. The direct comparison of stress and control group further showed

that cortisol levels are significantly higher in the stress group 50 minutes after stress induction, $p = .002$. In conclusion, peak cortisol levels measured 50 minutes after stress induction were significantly elevated only in the stress group demonstrating the effectiveness of the stress induction.

3.3.2.2 Experiment 1b

Participants in the stress group kept their hands in ice water for 155.12 ± 49.05 s. All participants in the control group kept their hands in water for the maximum of 180 s. Participants in the stress group perceived the SECPT as more stressful, $t(61.38) = 9.23$, $p < .001$, painful, $t(50.96) = 16.93$, $p < .001$, and unpleasant, $t(89.03) = 5.87$, $p < .001$ than participants in the control group indicating the success of stress induction as measured subjectively.

3.3.2.2.1 Heart rate. The analysis of heart rate showed a main effect of time, $F(2.41, 195.10) = 4.76$, $p = .003$, as well as a time by stress group interaction, $F(2.41, 195.10) = 9.56$, $p < .001$. Post hoc tests using Bonferroni correction revealed that in the stress group, heart rate significantly increased in minute 1, $p < .001$, and minute 2, $p = .016$, of the stress test relative to baseline. Thus, as in Experiment 1a only the stress group showed an increase in heart rate due to stress induction.

3.3.2.2.2 Blood pressure. For systolic blood pressure the analysis revealed a time by stress group interaction $F(2, 216) = 3.07$, $p = .048$. Post hoc comparisons showed a marginal difference in the stress group between time points 2 and 3, $p = .055$. Significant differences between stress and control group were visible before stress induction, $p = .039$, as well as 50 min after, $p = .012$.

The analysis of diastolic blood pressure showed a time by stress group interaction, $F(2, 216) = 5.11$, $p = .007$. Post hoc analyses showed that in stress group there was a drop in diastolic

blood pressure from the time of the SECPT to 50 minutes after, $p = .005$. Moreover, the control group had significantly higher blood pressure than the stress group at the end of testing, $p = .001$. While the pattern of results is different from Experiment 1a, the difference in blood pressure 50 minutes after stress induction is likely to be attributed to factors other than the SECPT. It might be the result of completing the task and is not likely to reflect the activation of the fast-acting stress system.

3.3.2.2.3 *Cortisol.* As in Experiment 1a, analysis of cortisol revealed a time by stress group interaction, $F(1.54, 168.19) = 3.41$, $p = .035$. Post-hoc comparisons showed that stress and control group were marginally different at baseline, $p = .082$, as well as right after stress induction, $p = .080$. They further revealed that cortisol levels in the control group dropped (presumably due to circadian rhythm) while cortisol levels in the stress group increased 50 min after stress induction demonstrating a change in cortisol levels due to stress induction.

In summary, while not all indicators of the fast-acting stress system reflect successful stress induction, cortisol levels indicate that delayed effects of acute stress were present at the time of testing.

3.3.3 Behavioural Results

3.3.3.1 Experiment 1a: Operant Conditioning

In order to determine whether stress and control group differed in operant conditioning, the number of handgrips reaching 50 % or more of the participant's maximum grip strength (Pool et al., 2015; Talmi et al., 2008) was compared between groups.

A mixed ANOVA revealed that irrespective of experimental condition, all participants readily learned to squeeze the handgrip in the first few trials: The analysis revealed a main effect

of trial, $F(8.62,861.47) = 4.03$, $p < .001$, such that grip frequency increased with the progression of the experiment. Crucially there was a main effect of stress group, $F(1, 100) = 7.34$, $p = .008$, indicating overall fewer grips in the stress relative to the control group (Figure 3.3a). This set of findings suggests that while action-outcome relations were learned instantaneously in both groups, acute stress led to a reduction in grip rate possibly due to reduced willingness to work for the reward.

3.3.3.2 Experiment 1b: Operant Conditioning with high reward

To ensure our findings did not simply reflect lack of motivation with low levels of reward, we aimed to replicate the main findings with higher levels of reward. As a follow-up to Experiment 1a, Experiment 1b employed 4x higher reward levels with a new set of participants. Again a main effect of trial, $F(7.47,821.99) = 2.55$, $p = .011$, indicated that all participants learned how to perform the task immediately. Moreover, a main effect of stress group, $F(1, 110) = 8.52$, $p = .004$, again indicated reduced response rates under stress (Figure 3.3b). Thus, I was able to replicate the main findings from Experiment 1a in an independent sample.

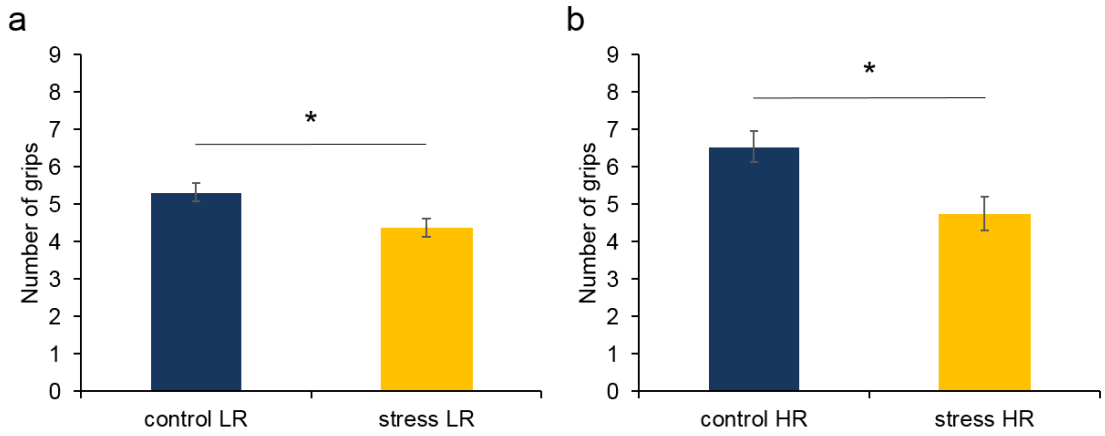

Figure 3.3. Operant conditioning results displayed separately for Experiment 1a (LR = low reward) and Experiment 1b (HR = high reward). The results show that acute stress induction reduced overall number of grips under both a) low reward and b) high reward conditions. Error bars indicate standard error of the mean. Asterisks indicate significance differences between stress and control group.

3.3.3.3 Experiment 1 a and b combined analysis

I further wished to examine whether the reduced response rate in Experiment 1a reflected reduced reward sensitivity. Because the pattern of behavioural results was equivalent across studies 1a and 1b, we combined the results from both studies and included reward level as a between-subjects factor. A mixed ANOVA with trial as within as well as stress group and reward condition as between-subject factors was employed to assess the effects of all factors and their interaction. The analysis revealed a main effect of trial, $F(8.53,1790.20) = 5.16$, $p < .001$, showing increasing grip frequency over the course of the experiment in all groups. There was a main effect of stress group, $F(1, 210) = 14.32$, $p < .001$ indicating overall fewer grips in the stress relative to the control group. Importantly, there was also a main effect reward condition, $F(1, 210) = 4.81$, $p = .029$, indicating fewer grips in the low relative to the high reward condition (Figure 3.4). There was no interaction between stress and reward level, $p > .2$. In summary, those under stress and those working for lower reward similarly demonstrated reduced willingness to

work for reward immediately following initial learning, consistent with predictions that stress reduces reward sensitivity.

In order to control for any effects of testing at different times of the day, the above reported analyses of behavioural data were also performed with time of day as a covariate. No significant interactions with time of day were observed (*p*s > .320) and the pattern of significant results did not differ from those presented above.

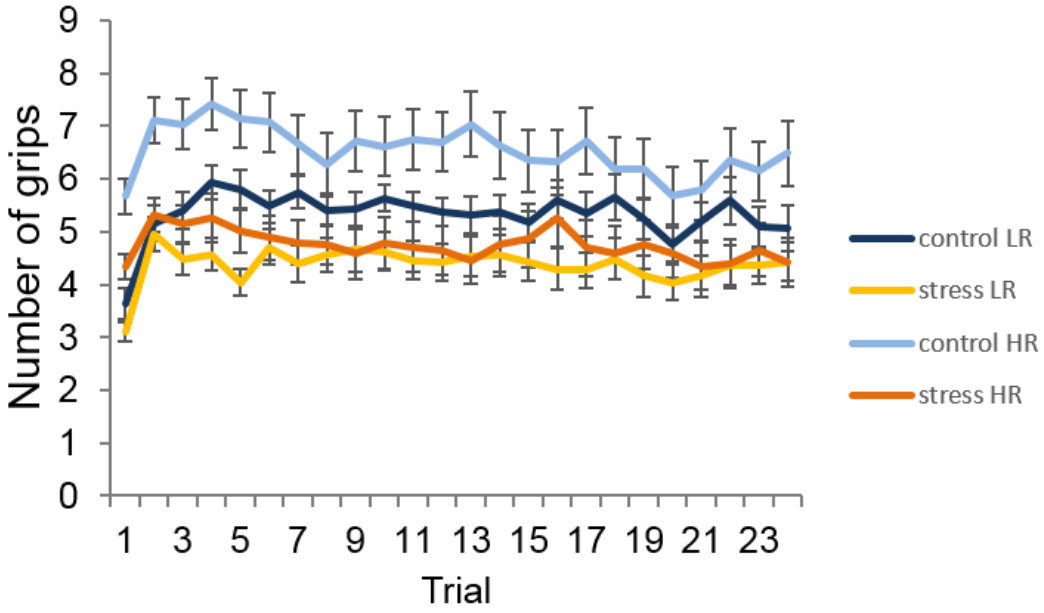

Figure 3.4. Operant conditioning (Experiment 1) results displayed trial by trial separately for control and stress group as well as for low reward (LR) and high reward (HR) groups show that mean number of grips reaching criterion force is reduced by acute stress induction and reduction of reward. Error bars indicate standard error of the mean.

Experiment 2

3.4 Materials and Methods

3.4.1 Participants

63 participants (48 females, mean age: 20.27 ± 3.04 years) completed enough trials for behavioural analyses. Nine participants were excluded due to insufficient task completion. All

participants were compensated for their participation by course credit for undergraduate psychology courses. Participants were asked not to eat, consume alcohol or caffeine and exercise two hours before the experiment. Testing was completed between 9AM and 6PM. Participants were randomly assigned to stress and control conditions (25 and 38 participants respectively). The study was approved by the Human Research Ethics Board of the University of British Columbia.

3.4.2 Materials

3.4.2.1 Pavlovian Conditioning. Stimuli were comprised of visual images of green, blue or purple fractal patterns displayed on a computer screen. These were randomly paired with sounds of cello, flute and trumpet to create three compound Pavlovian stimuli. The three compound stimuli were randomly selected to serve as CS+, CS- or baseline conditions. Monetary reward was indicated by presenting a Canadian quarter in the middle of the screen (Figure 3.1).

3.4.2.2 Questionnaires. See section 2.2.2.3.

3.4.3 Procedure

After obtaining written informed consent, we acquired initial saliva samples and blood pressure readings. This was followed by administration of the SECPT in either the stress or control condition (Figure 3.2). To observe physiological reactions during stress induction we initiated continuous heart rate recording at the beginning of the SECPT. The three-minute stress induction procedure was followed by blood pressure measurements, a cortisol sample and subjective stress ratings. The task started 25 minutes after the end of the SECPT allowing cortisol to reach peak levels (Schwabe et al., 2008). Heart rate was continuously recorded. After

participants finished the task, blood pressure and cortisol were tested one more time. For a more detailed description of the stress procedure and indicators of the stress response, see 2.2 Procedure for Experiment 1.

3.4.4 Classical Conditioning Task

Each participant completed 36 'task on' blocks with 4 s intertrial intervals or 'task off' blocks, during which the baseline stimulus was presented. The 'task off' or baseline period serves as a control condition for gathering initial likeability ratings not affected by reward expectations. Each 12 s 'task on' block was either a CS+ or a CS- trial characterized by the continuous presentation of the Pavlovian compound stimulus. Each 12 s block consisted of three 4 s time window each of which started with the random onset of the presentation of a gray patch, the cue (Figure 3.1). Participants were instructed to press a key to remove the patch in order to see whether it was hiding a reward. Participants were further told that the cue appeared three times per trial leaving to up to three possible rewards. In contrast to the operant task, participants were well aware of the fact that their action, i.e. the button press, had no influence on the outcome. No action was required during 'task off' blocks. Conditioning was assessed by reaction time in CS+ and CS- trials as well as likeability ratings of CS+, CS- and baseline stimulus.

3.4.5 Statistical Analysis

A mixed analysis of variance (ANOVA) was applied to the reaction time data with trial and CS type (CS+ and CS-) as within-subject factor and group (stress and control) as between-subject factor. Stimulus ratings were analyzed with a mixed design ANOVA with stimulus type (CS+, CS- and baseline) and stress group as factors. Physiological data (heart rate, blood

pressure and cortisol) were analyzed in a mixed ANOVA with time as within- and group (stress and control) as between-subject factors. All analyses were additionally performed with time of day - dichotomized as morning (testing between 9AM and 1PM) and afternoon (testing between 1PM and 6PM) - as a covariate. Greenhouse-Geisser corrections were applied if sphericity was violated. All analyses were performed with IBM SPSS Statistics 21.

3.5 Results

3.5.1 Control Variables

Exploratory correlations examining the relation between task performance and personality measures, state and trait anxiety, depression and childhood trauma did not reveal significant results. Furthermore, stress and control group did not differ with regard to age, sex, ethnicity and average levels of depression and anxiety.

3.5.2 Stress manipulation

The effect of stress induction was assessed by both subjective ratings and physiological measures such as heart rate, blood pressure and cortisol.

Participants in the stress condition kept their hand in ice water for 145.20 ± 54.70 s, while all participants in the control group kept their hand in water for 180 s. In addition, participants in the stress group perceived the SECPT as more stressful, $t(33.09) = 5.74$, $p < .001$, painful, $t(27.49) = 9.45$, $p < .001$, and unpleasant, $t(61) = 5.70$, $p < .001$, than participants in the control condition.

3.5.2.1 Heart rate. The analysis revealed a main effect of time, $F(3, 135) = 21.78$, $p < .001$ indicating that both groups showed an increase in heart rate as a result of the SECPT.

3.5.2.2 Blood pressure. No significant differences between stress and control group were found, $p > .2$.

3.5.2.3 Cortisol. The analysis of salivary cortisol revealed a time by condition interaction, $F(1.20, 73.29) = 10.12$, $p < .001$. Post-hoc comparisons showed that the control group had a significant drop in cortisol levels at the end of the experiment, $p = .001$, whereas cortisol levels in the stress group remain unchanged ($p = .574$). Thus, while under control conditions cortisol levels dropped presumably due to circadian rhythm, this effect was not detected in the stress group since the stress induction might have counteracted the observed drop.

Taken together, physiological indicators of acute stress do not deliver enough evidence to conclude that the fast-acting stress system was activated as a result of the SECPT, but differences in cortisol levels allow us to conclude that differences in cortisol levels were present at the time of testing, which was the intended effect.

3.5.3 Classical conditioning

In this experiment participants were asked to complete a total of 36 trials (18 CS+, 18 CS- trials in randomized order). However, most participants failed to respond in one or more trials, leaving the majority of participants with at least 14 completed trials for each condition. Thus, for the analysis, the first 14 completed trials for each condition (CS+, CS-) were taken from each individual and subjected to a mixed design ANOVA in order to compare response times in CS+ and CS- trials between participants under stress and control conditions.

The analysis revealed a main effect of trial, $F(9.10, 555.26) = 3.76$, $p < .001$, showing that reaction times decreased over the course of the experiment (Figure 3.5). Crucially, there was a CS type (CS+ and CS-) by stress interaction, $F(1, 61) = 10.67$, $p = .002$. Post-hoc comparisons

revealed that participants in the stress condition were slower to respond to CS+ relative to CS-, $p = .003$. No effect was observed in the control group ($p = .184$). Thus, appetitive classical conditioning was affected by delayed acute stress induction such that typically observed reaction time indices of conditioning were reversed by stress.

Subjective ratings of likability for experimental stimuli were also examined. Here there was a main effect of stimulus type, $F(2, 112) = 21.11$, $p < .001$, such that all participants liked CS+ stimuli better than baseline stimuli, and liked both stimuli better than the CS- fractal pattern after conditioning (Figure 3.5). This confirms that conditioning did indeed occur in both groups. There was also an effect of stress group, $F(1, 56) = 4.79$, $p = .033$, such that participants in the stress group had higher likeability ratings relative to the control group. There was no significant stimulus type by group interaction ($p = .31$). This opposing pattern of results for likeability ratings and behavioural response could suggest that these two indicators of conditioning measure different aspects of learning (e.g. outcome vs cue directed learning).

Again to control for potential effects of time of day on learning, all of the analyses reported above were also performed with time of day included as a covariate. Once again, no significant interactions between time of day and other factors were observed (ps > .692) and the pattern of significant results did not differ from that reported above.

Taken together the behavioural results suggest that despite the fact that both stress and control group did experience a conditioning effect, as evidenced by stimulus ratings, overall response times were markedly slowed under delayed acute stress. Such findings indicate a dissociation between effects of stress on implicit relative to explicit measures of Pavlovian learning.

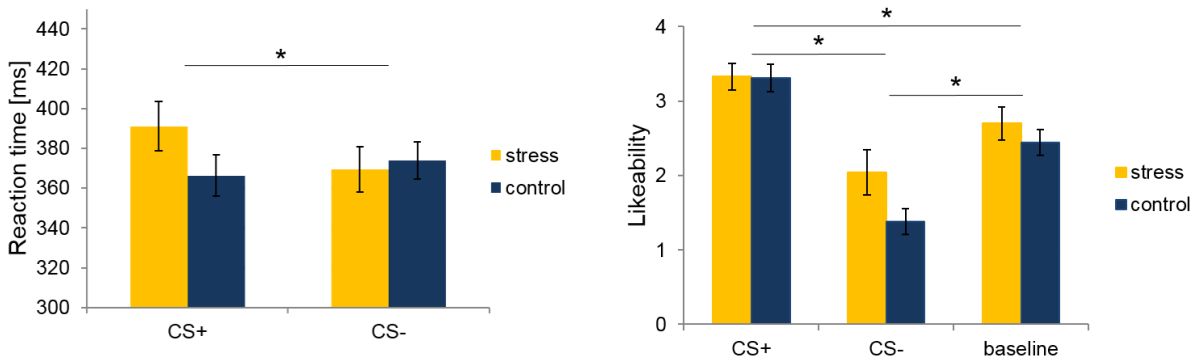

Figure 3.5. Results of classical conditioning (Experiment 2) study show reduced reaction time in CS- relative to CS+ trials under acute stress. Likeability ratings suggest successful conditioning in stress and control group with overall higher ratings under stress. Error bars indicate standard error of the mean. Asterisks indicate significance differences.

3.6 Discussion

The aim of the current chapter was to investigate the influence of delayed acute stress on simple appetitive associative learning processes in humans. Results showed that stress administered by means of the SECPT reduced operant responding as well as behavioural indices of Pavlovian learning. While the ability to learn contingencies in the operant task was unaffected by stress, following stress induction participants were overall less willing to work for a reward than they were in the no-stress control condition, and this was true regardless of whether participants received higher or lower levels of reward. In the no-stress condition, comparison of high and low reward showed that, in the absence of stress, participants were also less willing to work when the amount of reward was substantially lower. Furthermore, in the Pavlovian conditioning study likeability ratings indicated that both stress and control groups similarly developed explicit emotional associations. Yet the stress group showed an opposing behavioural pattern such that response times were faster in response to unconditioned relative to conditioned stimuli.

The operant task results revealed that overall stress reduces the willingness to work for a reward at a very early stage of habit formation, providing novel evidence that such early stages are susceptible to the detrimental effects of stress. The study was designed to assess such effects of stress in relation to findings from a previous study (Pool et al., 2015). In the study by Pool and colleagues (2015), after performing equivalent operant and classical conditioning tasks to those we employed, participants were presented with Pavlovian stimuli while performing the operant task in extinction. Results revealed that, in the stress relative to the control condition, participants were more likely to show *increased* responding (i.e. number of handgrips) when presented with the CS+. The authors concluded that under stress people are more prone to rely on habitual behaviour irrespective of the rewarding value of the outcome. That is, once habits are established, craving a reward guides participants' behaviour - an effect that is enhanced by stress. In contrast, my examination of the operant conditioning phase of the task allowed me to probe effects of stress on the establishment of instrumental responses. The current findings support the conclusion that, whereas stress may increase reliance on existing habits, initial stages of habit formation driven by the reinforcing properties of the reward are negatively affected by stress.

Another line of research has emphasized the notion that acute stress promotes the switch from goal-directed to habitual behaviour (Schwabe & Wolf, 2011) For that purpose, operant paradigms are used in which an initially rewarded action is trained until a habit is established, i.e. participants keep completing the action despite a lack of reinforcement or devaluation of the outcome. (Schwabe & Wolf, 2009). Critically, this shift from initial goal-directed or reward-oriented behaviour towards habitual responding is facilitated by acute stress (Schwabe & Wolf, 2009, 2010b). In contrast, in the present study, I measured behaviour that was not overtrained to the point that habits were strongly established. Thus, whereas previous studies provide evidence

for reduced behavioural flexibility under stress, as indicated by reduced goal-directed behaviour after devaluation, our findings further suggest that stress reduces reward-oriented behaviour or the willingness to work for a reward before habit formation can occur — at least in a simple task where learning is very rapid.

The current manipulation of reward value revealed a pattern of results consistent with research suggesting that stress reduces reward sensitivity — at least in susceptible individuals (Berghorst et al., 2013; Bogdan & Pizzagalli, 2006; Cavanagh et al., 2011). I assessed reward sensitivity by not only manipulating stress but also investigating effects of reward value. I suggest that, as the reduction in operant responding observed with stress mirrored that observed with lower levels of reward, the unwillingness to work for reward under stress may reflect reduced reward sensitivity. Theories of depression propose that stress induces an anhedonia-like state – an effect known as *learned helplessness* (Overmier & Seligman 1967; Shors & Dryver 1992). As a condition characterized by decreased reward sensitivity and motivation to pursue rewards, learned helplessness has been used as an animal model for depression (Klein, Fencil-Morse, & Seligman, 1976). While previous animal studies induced inescapable, traumatic shock, the current results are consistent with human literature showing effects that are not restricted to uncontrollable, traumatic stress (Bogdan & Pizzagalli, 2006).

It should be noted in this study I employed a very simple operant conditioning task. Here learning was instantaneous, and no stress-related differences in learning rate were observed. This had both advantages and limitations. The task not only allowed me to compare our findings to those of previous studies, but our measure of willingness to work for reward was not confounded by individual differences in the ability to learn complex reward contingencies. The simplicity of the task also effectively models common situations in which human learning is instantaneous and

the action-outcome relation is encoded after the first encounter (e.g., experiencing pleasant effects of a novel drug on the first encounter). In this way I was able observe the effects of stress on this type of salient instantaneous learning, with implications for understanding how stress may contribute to trajectories toward habitual drug taking. However, further studies should employ a more difficult learning task that manipulates reward contingencies, allowing assessment of stress on learning rates over time.

The results of the classical conditioning task further revealed a dissociation between explicit responses and behaviour: Likeability ratings indicated successful learning of reward associations in both stress and control groups. However, response times were slower for CS+ than CS- trials under stress. In contrast, no difference between CS+ and CS- was observed in controls, suggesting that only implicit measures of conditioning were influenced by acute stress. Our results are consistent with findings in non-human animals indicating that, in classical conditioning, effects of acute stress on implicit learning are dissociable from effects on explicit learning processes (Shors & Servatius, 1997). Another possible interpretation of the data can be found in the animal literature on individual differences in associative learning (Flagel, Akil, & Robinson, 2009): *Goal-trackers* prioritize rewarding outcomes without developing emotional associations with the CS+. In contrast, *sign-trackers* develop strong emotional associations with the cues signaling the reward, even at the cost of interest in the rewarding outcome (Hearst & Jenkins, 1974). In the current study, I can speculate that acute stress induction made participants more likely to act like sign-trackers, who give more weight to the associated cue and less to the rewarding outcome. Future research should be conducted to investigate sign- and goal-tracking in humans especially under the influence of environmental factors such as stress.

The pattern of results observed here (i.e. reduced operant responding) may depend in part on the timing of the associative learning tasks in relation to the acute stressor. In the present study, we employed a delay following the stress induction to capitalize on effects of glucocorticoids on behaviour. Non-human animal research has suggested that stress typically enhances learning whether training begins immediately after stress induction or with a delay (Servatius & Shors, 1994; Shors, Weiss, & Thompson, 1992), although this finding has not been found to be generalizable to all stressor types or tasks and also depends on the sex of the animal (Shors, 2004). Research in humans suggests that acute stress impairs explicit learning mediated by glucocorticoid action, while learning is enhanced when it occurs in close temporal proximity to the stressor, a process that is thought to be mostly driven by norepinephrine (NE) (Joels et al., 2006). Recently, studies demonstrated that glucocorticoid action via mineralcorticoid receptors is critical for a shift from hippocampus-based 'cognitive' to dorsal striatum-dependent 'habit' learning strategies [for review see (Vogel, Fernandez, Joels, & Schwabe, 2016)]. In line with that, the present findings suggest that goal-directed or 'cognitive' behaviours were impaired under glucocorticoid driven delayed stress effects. An important follow-up to the present study is investigating effects of stress when learning occurs directly after stress induction to differentiate the effects of glucocorticoid and NE activation and to demonstrate the involvement of the LC-NE system in more (complex) forms of reinforcement learning. Thus, in the next chapter I will present a study delineating the effects of the immediate and delayed stress response on operant conditioning.

In both experiments, for a number of different measures including psychophysiology (heart rate, blood pressure), cortisol and subjective parameters, significant stress group differences indicated that the stress manipulation was successful. Nonetheless it should be noted

that heart rate and blood pressure measurements were not available for the time of stress induction, which is the time when differences would be expected to be largest. Yet the fact that differences were observed even after the stress induction suggests that these differences were present during the SECPT. The same holds true for the cortisol samples taken right after stress induction as well as an hour after (at the end of the experimental procedure). While we did not assess peak cortisol ~25min after SECPT, elevated levels by the end of task completion indicate that cortisol levels were elevated during behavioural experiments. Moreover, heart rate and blood pressure changes due to stress induction were not visible in Experiment 2 indicating that the fast-acting stress system might not have been activated or alternatively that the measurements were not able to capture those changes due to timing. However, group differences in cortisol levels were present in all experiments suggesting that the effects of delayed stress targeted in the present study were in effect.

In conclusion, the current study showed that delayed effects of acute stress reduce operant responding presumably due to reduced reward sensitivity as one aspect of reinforcement learning. Further, stress prevented the translation of learned emotional associations into reward-oriented behaviour. Thus, consistent with what is known from stress and learning research, it seems that appetitive learning processes subsequently leading to the establishment of new habits, are suppressed for a certain period after stress induction, an effect thought to be driven by glucocorticoid processes. These findings add to our understanding of the influence of stress on early stages of habit formation that are relevant for the development of affective biases in attention and behaviour as observed in addiction or depression. A critical follow up question of the current set of studies is whether immediate, NE-driven stress effects enhance reward-based learning promoting the establishment of maladaptive habits and relapse related to addiction.

Thus, in the next chapter I will present a study directly contrasting the action of the immediate and delayed stress response on operant conditioning processes.

Chapter 4: Differential effects of the immediate and delayed stress response on operant conditioning

4.1 Introduction

Stress - a state of real or perceived threat to an organism's homeostasis (S. M. Smith & Vale, 2006) - is associated with detrimental health conditions and is often construed in a negative light. Anecdotally, most of us can remember a stressful situation such as a job interview or giving a speech, where access to previously encoded information seemed impaired. However, moderate levels of stress and arousal can also promote optimal performance in some contexts and cognitive domains (Diamond, 2005). Research involving effects of moderate stressors on cognition have found that stress can either facilitate or impair cognitive functions such as learning and memory [for example (Joels et al., 2006; Sandi & Pinelo-Nava, 2007; Shors, 2006)]. Thus, important outstanding questions concern the circumstances under which stress enhances or impairs cognition, as well as the biological processes that underlie such effects.

The biological response to an acute stressor is twofold [for review see (de Kloet, Joels, & Holsboer, 2005)]: The immediate response of the autonomic nervous system (ANS) leads to a release of catecholamines such as norepinephrine (NE) and involves rapid action of glucocorticoids. The delayed response of the hypothalamic-pituitary-adrenal (HPA) axis on the other hand involves genomic glucocorticoid actions. Many studies have investigated the effects of acute stress on learning and memory. For example, it has been shown that stress before learning can both facilitate and impair memory (Elzinga, Bakker, & Bremner, 2005). Moreover, stress immediately after learning has been shown to enhance memory (Cahill, Gorski, & Le, 2003; Roozendaal, 2000) while reduced recall performance has been reported if stress was induced right before (Buchanan, Tranel, & Adolphs, 2006; de Quervain, Aerni, Schelling, &

Roozendaal, 2009). An influential and compelling theory (Joels et al., 2006) proposes that the opposing effects of stress on learning and memory can be explained by the dualistic nature of the stress response in which the immediate stress response has been proposed to facilitate learning, while the delayed action of glucocorticoids leads to an elevated processing threshold for incoming information and hence suppresses memory formation (Herman et al., 2012). In brief, the theory suggests that if stress occurs at the same time as the learning experience, learning and subsequent memory performance will be enhanced. On the other hand, if stress is induced with a delay before learning, performance is impaired.

However compelling the theory and despite some solid evidence from animal research [for example (D. J. de Quervain, B. Roozendaal, & J. L. McGaugh, 1998)], this literature is characterized by many inconsistent findings, including those that cannot be explained by that theory. For example pharmacological manipulations with glucocorticoids have been shown to enhance memory function depending on circadian rhythm (Lupien et al., 2002). Another study failed to show reduced memory performance under the effects of cortisol in middle-aged women (Domes, Heinrichs, Reichwald, & Hautzinger, 2002), consistent with gender-specific differences in cortisol levels and memory (Wolf, Schommer, Hellhammer, McEwen, & Kirschbaum, 2001). Yet another study (Schwabe & Wolf, 2010a) demonstrated that learning under stress reduced recall and recognition memory similarly for men and women. Thus, while timing of the stressor relative to the learning experience and especially the learning or memory phase (e.g. encoding vs consolidation (Cahill et al., 2003)) [for review see (Schwabe & Wolf, 2013)] are potential and likely modulators of the effects of stress on memory, effects of stress hormones may be subject to a number of boundary conditions related to differences in time of day or sex hormone levels.

Lastly, the engagement of different memory system is likely to play a huge role in determining the effects of stress on learning (Schwabe & Wolf, 2013). While so far, I have reviewed effects of acute stress on hippocampally-mediated explicit memory, which is mostly consistent with Joels' theory of timing effects, little is known about effects of acute stress on appetitive learning processes that may be more implicit and not require the hippocampus. Previous research on aversive or fear conditioning as one form of implicit learning provides some support for the idea that stress effects also depend on the stress system activated. For example, blocking the influence of the norepinephrine system surgically or pharmacologically leads to impaired learning (Bush, Caparosa, Gekker, & Ledoux, 2010), while blockage of glucocorticoids has no immediate effects on memory acquisition (Pugh, Fleshner, & Rudy, 1997). This suggests that the neuromodulators released in the immediate NE-modulated stress system not only facilitate explicit learning but are also essential for successful fear conditioning.

Effects of stress on reward learning and decision-making processes have only recently become an important question in stress research [for reviews see [(Ehlers & Todd, 2017b; Hollon, Burgeno, & Phillips, 2015)]. Thus far, research in humans has focused on the effects of stress on habit formation – the automatization of new behaviours through reinforcement. In a series of studies, participants completed operant conditioning tasks before or after being exposed to acute stress (Schwabe & Wolf, 2009, 2010b). Participants under the influence of acute stress displayed faster or more robust habit formation both when stress was induced before and after reward contingencies were learned. More specifically, participants in the stress groups were more likely to perform actions even if those were no longer rewarded or reinforced. Another study demonstrated that under acute stress participants mobilized more effort in response to a stimulus that once signaled a reward, again indicating that a stronger stimulus-response

relationship developed in the stress group (Pool et al., 2015). Both above studies measured learning at peak cortisol levels, after a delay, thus harnessing the slower stress response. In Chapter 3 I presented my own research demonstrating that the delayed stress response can also impair simple forms of operant and classical conditioning (Ehlers & Todd, 2017a). However, it remained to be investigated how the immediate and delayed stress response affect a more complex operant conditioning task.

Thus, the goal of the current chapter was to directly test the theory of stress timing effects in the context of appetitive conditioning or implicit learning. For this purpose, participants were randomly assigned to complete an operant conditioning task under immediate or delayed stress or control condition of the commonly employed socially evaluated cold pressor test (SECPT) (Schwabe et al., 2008). I hypothesized to see facilitated operant conditioning under the effects of the immediate stress response, while I predicted the delayed stress response to have no or detrimental effects on learning.

4.2 Materials and Methods

4.2.1 Participants

In order to determine the number of subjects, a power analysis was performed prior to data collection. Assuming a moderate effect size ($\eta^2 = .095$) and power ($1 - \beta = 0.85$) and a repeated measures ANOVA with two between-subject factors indicates a required sample size of 176 participants. In order to allow for missing data, data collection was continued until the end of the term.

223 participants (164 females, mean age: 20.5 ± 2.9 years) took part in this study. Participants were compensated for their participation with course credit. 16 participants were

excluded for high baseline cortisol levels (more than 2.5 SD higher than average) and 25 for cortisol increase less than the recommended standard (Miller et al., 2013). In addition, 13 participants were excluded for missing data and 4 participants were excluded due to the fact that their average performance at the end of the learning period was below chance.

Participants were asked to refrain from consuming food, alcohol or caffeine as well as exercise two hours before the experiment due to possible interference with the stress response (Kudielka et al., 2007). Participants were randomly assigned to stress and control conditions. 51 participants completed the experiment in the delayed control conditions, 49 in the immediate control condition, 37 in the delayed stress condition and 38 in the immediate stress condition. The study was approved by the Human Research Ethics Board of the University of British Columbia.

4.2.2 Materials

4.2.2.1 Stimuli and apparatus. The MATLAB (The MathWorks, Natick, Massachusetts, USA) toolbox Cogent 2000 was used for all stimulus presentation. Stimuli consisted of six symbols presented on black background. Each stimulus had a different probability of being associated with a reward (90%, 80%, 70%, 30%, 20% and 10%).

4.2.2.2 Questionnaires. See Section 2.2.2.3

4.2.3 Procedure

4.2.3.1 Overview. After obtaining written informed consent, baseline cortisol levels were measured (Figure 4.1). This was followed by the stress or control condition of the SECPT (Schwabe et al., 2008) (described in more detail in the next section). Immediately after, cortisol

was measured again in addition to the administration of the SECPT questionnaire – a three-item questionnaire measuring the subjective stress response (Schwabe et al., 2008). Participants randomly assigned to the 'delayed' condition spent the next 25 minutes completing a series of online questionnaires, while participants in the 'immediate' condition, started the operant conditioning task immediately after completing the SECPT (Figure 3.6). 25 minutes after the SECPT both groups provided their third salivary cortisol sample. Subsequently, participants in the 'delayed' group completed the operant conditioning task, while participants in the 'immediate' condition were asked to complete online questionnaire. Before debriefing, around 50 minutes after the SECPT, all participants provided their last cortisol sample (Figure 3.6).

4.2.3.2 Stress procedure. For stress procedure and questionnaire see Section 2.2.3.2 and 2.2.3.3.

4.2.3.3 Salivary cortisol analysis. Analysis was identical to Section 2.2.3.4.

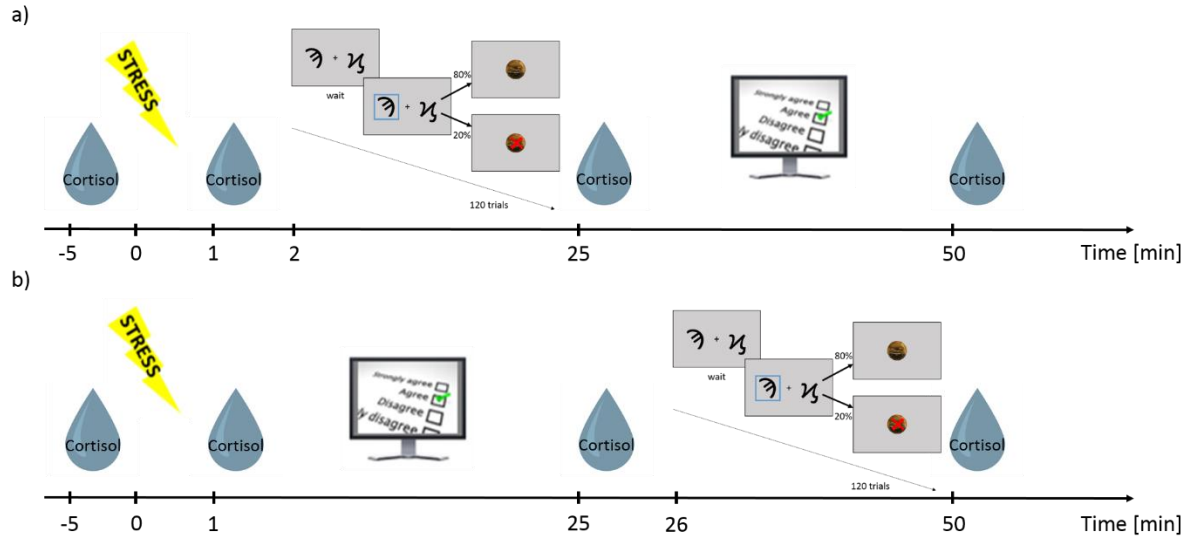

Figure 4.1. Overview of experimental procedure. Salivary cortisol samples were taken before and right after stress induction by means of the socially evaluated cold pressor test (SECPT) as well 25 minutes and 50 minutes after stress induction. a) In the immediate condition, the operant conditioning task was completed immediately after the stress or control condition of the SECPT. b) In the delayed condition, the operant conditioning was completed 25 minutes after the stress or control condition of the SECPT.

4.2.4 Operant Conditioning Task

A two-stage reinforcement learning task was employed to test effects of immediate and delayed acute stress on the acquisition phase of operant conditioning as well as the generalization of learned information (Shiner et al., 2012). This task includes both an acquisition phase that allows assessment of initial learning rate and a generalization phase that allows assessment of learning effectiveness, operationalized as the ability to generalize learned associations.

Acquisition phase. Six symbols were used as stimuli in the acquisition phase (Figure 4.2), each of which was associated with a different probability of leading to a reward (90% to 10%). For the acquisition phase, stimuli were paired in three different combinations: 90% and 10%, 80% and 20% and 70% and 30%, presented in random order. In each trial, two symbols were presented on a computer screen and participants learned to select the one that is more likely to yield reward. After a delay of 300ms, outcome (reward or no reward) was presented for 500ms. Participants were incentivized to perform to the best of their abilities by being informed that the dollars collected in the task can be exchanged for a treat at the end of the experiment. Each participant completed a total of 8 experimental blocks of 15 trials.

Generalization phase. For the generalization phase (Figure 4.3), new stimulus pairings were created by pairing the 90% and the 10% stimulus with each of the other stimuli. These were used alongside the stimulus sets used in the acquisition phase, leading to a total of 7 stimulus sets. As in the acquisition phase, participants were asked to select one of the two stimuli to maximize their reward. The difference was that no reward outcome feedback was given and participants had to base their choices on the knowledge gained in the first half of the task. All participants completed a total of 140 trials in the generalization phase.

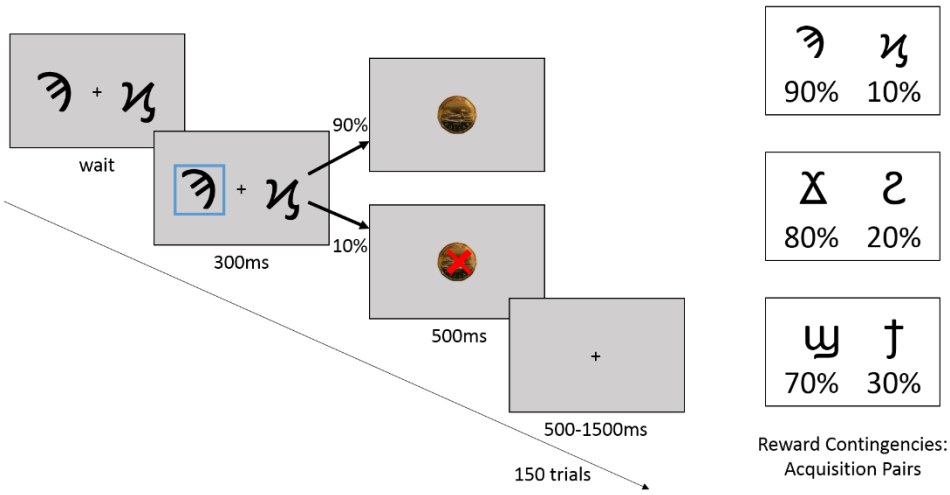

Figure 4.2. Acquisition phase of operant conditioning task. In each trial two stimuli with different reward probabilities were presented. Participants were asked to choose one of the two with the goal to maximize the reward they obtain. The reward gained was shown in each trial. A total of 120 trials were completed. Number of correct choices and reaction time were recorded.

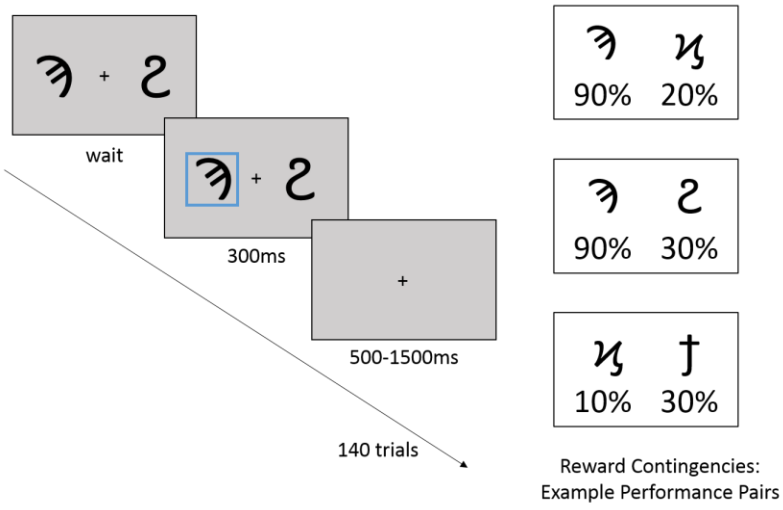

Figure 4.3. Generalization phase of operant conditioning task. Novel stimulus pairings were created from the six stimuli used in the acquisition phase. Participant were asked to choose one of two stimuli presented on the screen, and had to maximize reward based on their knowledge from the acquisition phase. Feedback about the total reward gained was only provided at the end of the task. A total of 140 were completed. Number of correct choices and reaction time were recorded.

4.2.5 Statistical Analysis

An 8 x 2 x 2 mixed analysis of variance (ANOVA) with task block as within and stress group (stress vs control) as well as timing (immediate vs delayed) as between subject factors was employed to independently test for effects of stress with different timing on operant conditioning. Cortisol was analyzed in a mixed ANOVA with time as within- and group (stress and control) as between-subject factors.

Exploratory correlations revealed no relationship between any of the individual difference measures for (social) anxiety, autism and behavioural activation/inhibition (for details see methods) and reward learning performance or stress response as indicated by cortisol levels.

4.3 Results

4.3.1 Stress Manipulation

The analysis revealed no difference between females and males with regard to the stress response. Participants in the stress group kept their hands in ice water for 168.5 ± 34.4 s, while participants in the control group kept their hands in water for 175.6 ± 16.2 s.

4.3.1.1 Subjective stress experience. Participants in the stress group experienced the SECPT as significantly more stressful, $t(90.45) = 9.37$, $p < .001$, painful, $t(76.23) = 15.38$, $p < .001$, and unpleasant, $t(171) = 14.65$, $p < .001$, confirming group differences in subjective stress experience.

4.3.1.2 Cortisol. The analysis of cortisol showed a main effect of time, $F(1.97, 288.971) = 8.48$, $p < .001$, as well as a time by stress group interaction, $F(3, 441) = 12.54$, $p < .001$. Post-hoc comparisons revealed that the stress group cortisol levels collected 25 minutes after stress induction were significantly higher than cortisol levels at baseline, $p < .001$, and right after stress

induction, $p < .001$. However, the control group did not show significantly higher cortisol levels 25 minutes after stress induction, demonstrating the effectiveness of the stress induction procedure.

4.3.2 Operant Conditioning

Acquisition Phase. Accuracy, indexed as the percentage of high reward probability choices in the eight experimental blocks, was compared between the four experimental groups (immediate stress, immediate control, delayed stress and delayed control).

A mixed ANOVA revealed a main effect of block, $F(5.24, 896.43) = 50.93$, $p < .001$, demonstrating that irrespective of experimental condition, accuracy increased over the course of learning in all participants (see Figure 4.4). The analysis further showed a block by stress group by timing interaction, $F(7, 1197) = 2.28$, $p = .026$ (see Figure 4.4). Follow-up ANOVAs performed separately in the immediate and delayed group revealed no significant block by stress interaction in the delayed group, $F(7, 602) = .614$, $p = .745$ (see Figure 4.4b). However, a significant interaction was demonstrated in the immediate group, $F(7, 595) = 2.28$, $p = .001$, demonstrating that stress immediately before an operant conditioning facilitated learning (see Figure 4.4a), while a delay after stress induction did not affect task performance.

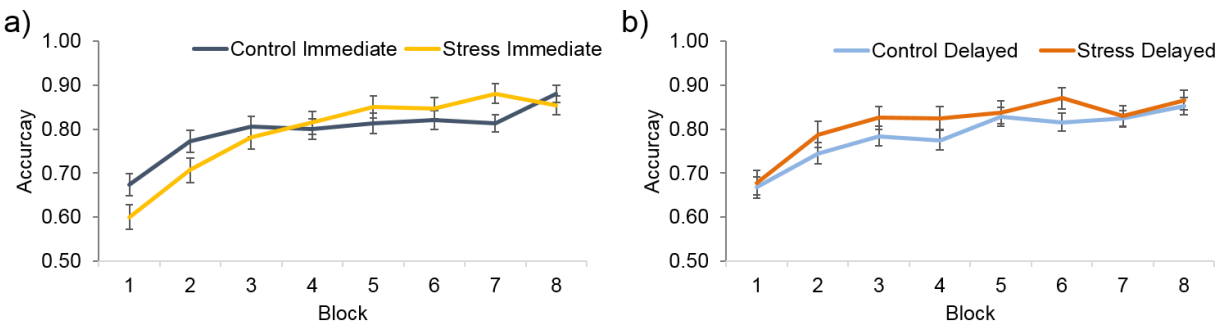

Figure 4.4 Operant conditioning results for the immediate and delayed experimental condition. The results show that acute stress induction relative to the control condition a) enhanced accuracy in the immediate stress condition but b) had no effect in the delayed condition. Error bars indicate standard error of the mean.

Generalization phase. Accuracy in the generalization phase was not modulated by stress group, $F(1, 1) = 4.11$, $p = .292$, timing of the stressor, $F(1, 1) = 29.248$, $p = .116$, or a stress by timing interaction, $F(1, 171) = .39$, $p = .844$. However, a regression analysis revealed that generalization phase accuracy was predicted by the average acquisition phase accuracy of blocks 5 to 8, $F(1, 173) = 77$, $p < .001$, with an R^2 of .308, (when learning reached asymptote). These results suggest that it is not the stress exposure that affected participants' ability to generalize learned information. However, how well information was learned in the acquisition phase was a significant predictor of participants' individual ability to generalize.

4.4 Discussion

The aim of the current chapter was to shed light on the question of how appetitive learning is differentially modulated by acute stress depending on the timing of the stressor relative to the learning experience. Results showed that stress administered by the SECPT enhanced performance in an operant conditioning task if stress was induced immediately before task completion. In contrast, operant conditioning performance was not affected by the late stress

response. Interestingly, performance in the following generalization phase that required stimulus generalization was not affected by acute stress but was solely predicted by operant conditioning performance at the end of the learning period.

The current study set out to directly test the hypothesis that the fast-acting stress response facilitates learning or memory encoding, while the slow-acting release of glucocorticoids impairs performance (Joels et al., 2006) focusing on immediate vs delayed effects of stress on operant conditioning processes.

Non-human animal research has produced mixed results regarding the delayed stress response and its effects on associative learning processes. Some early studies suggested that stress may enhance learning whether training begins immediately after stress induction or with a delay (Servatius & Shors, 1994; Shors et al., 1992). Other studies demonstrated that the pharmacological manipulation with a glucocorticoid receptor agonist may enhance learning over several session of a Pavlovian conditioning paradigm (Zorawski & Killcross, 2002, 2003). In contrast, a follow-up study reported no benefit of Pavlovian learning from exposure to glucocorticoids, but impaired instrumental learning (Pielock et al., 2013). To date, the effects of stress on reward-related processes in humans have primarily been studied in relation to habit formation as a result of overlearned rewarded actions (Schwabe & Wolf, 2011). In those studies, appetitive learning tasks are performed until a habit is established such that the initially reinforced action is performed even without a reward (Schwabe & Wolf, 2009). That shift from goal-directed to habit behaviour has been demonstrated to be facilitated by acute stress (Schwabe & Wolf, 2009, 2010b). Results from the Chapter 3 (Ehlers & Todd, 2017a) suggest that under moderate acute stress, activation of the HPA axis, as indicated by elevated cortisol levels at the time of testing, also affects emotional learning before a habit can be established. Study results

revealed reduced reward sensitivity or the willingness to work for a reward in operant conditioning as well as behavioural indices of classical conditioning under stress. In the current study I employed a more cognitively challenging task as opposed to a more effortful operant conditioning task that would allow me to assess learning rate instead of reward sensitivity. Consistent with results from Chapter 3, the speed with which action-outcome associations were developed was not affected by the slow-acting stress response. Taken together, the present findings add to the sparse literature on stress and appetitive learning in humans and are relatively consistent with the theory of time dependent effects of acute stress that has been developed mostly based on hippocampally-dependent learning and memory tasks (Joels et al., 2006).

The role of stress and especially the immediate NE-driven stress response has been neglected for a long time in the field of appetitive learning (Weinshenker & Schroeder, 2007) because reward related processing had been mostly attributed to activation in the dopamine (DA) system (K. C. Berridge, 2007; Flagel et al., 2011; Pessiglione et al., 2006). For example, the study after which I modeled the current experiment examined Parkinson's patients on and off dopamine replacement therapy (Shiner et al., 2012). That study demonstrated a role for DA only in the generalization phase but not in the actual learning phase – in contrast to the current results in which the NE-driven stress response seems to facilitate learning. A large body of literature has suggested DA to have a selective role as a mediator of incentive salience (K. C. Berridge & Robinson, 1998; Flagel et al., 2011). In other words, DA has been shown to be essential for the 'wanting' but not the 'liking' of a reward or associated stimulus. Looking beyond the role of dopamine, we have recently reviewed evidence supporting the less-studied role of NE in reward learning (Ehlers & Todd, 2017b). Consistent with the current findings of enhanced reward learning under the influence of the NE-driven stress response, non-human animal research has

shown that enhanced NE availability facilitates conditioning for reward learning processes (Zarrindast et al., 2002). Moreover, single-cell recording studies in monkeys have suggested NE plays a role in signaling reward and action associated costs and in integrating that information in decision-making processes (Bouret et al., 2012; Bouret & Richmond, 2009, 2015). Finally, convergent evidence supports the proposal that the role of the locus coeruleus norepinephrine system in reward-related processing is to signal uncertainty (Ehlers & Todd, 2017b; Sadacca et al., 2016). Thus, enhanced NE availability immediately after stress exposure may facilitate reward-oriented decision-making as a situation of uncertainty.

One limitation of the current study is that the number of trials, which was based on previous research (Shiner et al., 2012), was not sufficient to measure individual learning curves to extract more specific metrics of learning processes such as the asymptote (performance at the end of training), the amplitude (difference between baseline and asymptote) and the curvature or learning rate (Cousineau & Lacroix, 2006). Future studies could attempt to design a more complex task that allows for a much longer learning period and hence sufficient data points to perform curve fitting. One major limitation of the current study but also those presented in Chapter 3 is that activation of the NE-driven stress system is not directly tested. While the effectiveness of the stress induction procedure is assessed by elevated cortisol levels, which imply that the fast-acting stress system must have been activated before, in future studies, additional measures can be used as a manipulation check or in order to determine dose dependent effects of NE. Two popular measures of noradrenergic activity are alpha-amylase and pupil dilation. Salivary alpha-amylase has been shown to be associated with blood plasma NE levels (Thoma, Kirschbaum, Wolf, & Rohleder, 2012) and could thus be used as an overall measure of sympathetic nervous system activity before and after stress induction. In contrast, pupil dilation

has been suggested as a proxy for locus coeruleus activation (Joshi, Li, Kalwani, & Gold, 2016) that might be employed as an online measure of LC-NE activation.

Future studies should further aim to pharmacologically manipulate the noradrenergic and glucocorticoid system as it has been done in studies on habit formation (Schwabe et al., 2011; Schwabe et al., 2012) in order to clearly delineate the influence of different stress hormones on associative learning processes. In addition, besides replicating the current findings, future studies could extend the results by also contrasting effects of the immediate and delayed stress response on classical conditioning. Finally, in order to find more support for or evidence against the stress timing theory (Joels et al., 2006), future studies should strictly control for or systematically vary factors that have been shown to influence the effects of acute stress such as baseline cortisol levels (Lupien et al., 2002), the strength of the stressor (Pielock et al., 2013), sex of participants (Wolf et al., 2001) or the experimental task used (see operant conditioning tasks in Chapter 3 and 4).

In conclusion, the set of studies in the last two chapters provided some support for the theory that the immediate stress response facilitates while the delayed stress response impairs associative learning processes underlying the formation of affective biases. More broadly speaking, the current chapter suggests that exposure to acute stress in everyday life is an important contributor to individual differences in both passive and active appetitive learning processes, which in turn might lead to differences in attention, perception or memory. It can be speculated that the observed facilitation of operant conditioning in the immediate stress condition might be an adaptive response. It is likely that learning of beneficial decisions and actions performed under acute stress is facilitated in order to promote ideal behaviours in a situation of disrupted homeostasis and challenge, while cognitive processing is inhibited after initial stress

exposure. In the next chapter I will turn to the question of how the representation of an initially neutral stimulus changes its meaning after repetitive pairing with an emotionally or motivationally salient stimulus.

Chapter 5: Development of novel affective associations in the brain

5.1 Introduction

Learning which stimuli in our environment co-occur with something good or bad is critical for the survival of all organisms. The capacity to develop painful or pleasurable associations with predictive cues in our environment is highly conserved across species: It is observed in very simple organisms, such as the roundworm *Caenorhabditis elegans*, as well in much more complex creatures such as birds and mammals. It is so central to our survival and our ability to make sense of the world around us that often a single exposure to an aversive or appetitive stimulus and an associated neutral stimulus is enough for us to remember the critical information for the rest of our lives (VanElzakker, Dahlgren, Davis, Dubois, & Shin, 2014; Yamamoto, Shimura, Sako, Yasoshima, & Sakai, 1994). Many decades of research on conditioning processes in non-human animals have established basic neural and behavioural mechanisms by which animals learn what predicts the occurrence of positive or negative events (Andreatta & Pauli, 2015; J. LeDoux, 2003; Maren, 2001; Martin-Soelch et al., 2007). My own work presented in previous chapters has contributed to the current knowledge in the field, demonstrating how environmental or individual difference factors such as acute stress modulate the development of novel appetitive associations. Yet outstanding questions about the nature of the information that we learn to associate with salient events remain to be resolved.

Learning which stimuli predict the occurrence of positive or negative events engages a form of associative learning called classical or Pavlovian conditioning. In Pavlovian conditioning an initially neutral stimulus (the conditioned stimulus or CS) is paired with an inherently positive (appetitive) or negative (aversive) unconditioned stimulus (US) (Pavlov, 2010). Given the power and importance of these associative learning processes, the question arises what information

related to the unconditioned stimulus becomes associated with the CS during the learning process and how and where in the brain that information is represented. Does a conditioned stimulus (CS) reactivate brain activation patterns elicited by the unconditioned stimulus (US) in the same way that a CS produces the same response as the US behaviourally? And if so, what elements of the information of the painful or pleasurable experience are re-activated in key brain regions? Do neural responses to the CS represent the sensory properties of a painful and a pleasurable US or does it represent the more abstract concepts of hedonic valence - that is the painful and pleasurable qualities of the experience?

Traditionally, at least behaviourally, the CS has been thought of as a substitute stimulus eliciting the unconditioned response (UR), consistent with the observation that the conditioned response (CR) mimics the UR (Holland, 1990). In the mid-twentieth century, this idea appealed to supporters of the two then-prominent ways of thinking about conditioning: On the one hand, the stimulus-response view (S-R) argued that conditioning is the result of an association between a stimulus and its response, which can be a motor, physiological or hedonic response (S-R). On the other hand, the stimulus-stimulus (S-S) view (Byrne, 2003) argued that with learning the CS comes to elicit the same afferent activity initially elicited by the US (Hull, 1943) and an association between a sensory process and a motor response is formed (Spence, 1950). Defenders of the S-S view believed in greater cognitive flexibility such that the CS becomes associated with the US, which enables a wider range of possible responses (Tolman, 1932). Much of the historical debate focused on operant conditioning, in which an explicit response is required in order to receive a reward (Skinner, 1963); however, more recently, in the context of classical conditioning, questions have arisen regarding the nature of the "CS-activated US representative" - that is, about what aspect of the US the CS becomes to represent (Holland,

1990). While it is currently acknowledged that, most likely, learning cannot fully be accounted for by either S-S or S-R associations alone, let us consider the extreme: In the case of exclusive S-S driven processes, after learning, a CS such as the bell in Pavlov's experiments would make the animal literally taste and smell the food associated with it and thus evoke the same perceptual experience as the US, which in turn leads to salivation, the response. In contrast, the S-R model would predict that, although salivation occurs in response to the bell, critically — it does so without evoking the sensation of the absent US. Rather, it might evoke a feeling of pleasure associated with the food. Current research provides some evidence for both sides. On the one hand, *autoshaping* or *sign-tracking* describes a phenomenon in classical conditioning in which animals behaviourally engage with the CS in the same way as they would with the US (Brown & Jenkins, 1968; Flagel, Watson, Akil, & Robinson, 2008). Critically this effect is so far reaching that some animals completely neglect the US suggesting that the CS serves as a perfect substitute for the US and that S-S conditioning has occurred (Morrison, Bamkole, & Nicola, 2015). On the other hand, outcome devaluation experiments in the habit formation literature suggest that once a certain action is established, a set response to a stimulus is initiated irrespective of the associated outcome indicating S-R learning (Schwabe et al., 2007; K. S. Smith & Graybiel, 2016). In summary, there is some evidence for both theories presumably due to differences in the experimental setup and type of learning investigated making it impossible to solve this debate with pure behavioural methods. However, employing contemporary brain-decoding methods, in the current chapter, my aim is to address this question from a new angle by investigating what aspects of the US - the sensory sensation or hedonic response - the CS is representing in brain circuitry supporting emotional learning by associating neutral cues with both painful and pleasurable touch.

Traditionally, the decades old debate about what a CS is representing or whether it becomes associated with the US (S-S) or the UR (S-R) has lacked the tools to clearly answer this question, as pure behavioural or univariate brain imaging approach cannot provide sufficient information. Because of their ability to decode representations of mental content, contemporary multivariate approaches to analyzing fMRI data offer the potential to resolve this question. Rather than measuring activation in individual voxels, or *average* voxel activation in a region of interest (ROI), and comparing those between experimental conditions, such multivariate approaches as multivoxel pattern analysis (MVPA) allow us to compare patterns of activation *across* voxels in a given brain region (Kragel, Koban, Barrett, & Wager, 2018). Representational similarity analysis (RSA) (Kriegeskorte & Kievit, 2013; Kriegeskorte et al., 2008) is an MVPA approach that examines correlated patterns of voxel activation associated with cognitive states, for example responses to the presentation of a US or CS+. Characteristic patterns of correlation of the patterns of BOLD response across voxels are decoded as neural representations of the cognitive events. We can then compare how similarly or differently to each other cognitive states are represented in the brain, and how that may change over the course of learning. In the present study we employ a two-step approach to examining what information is represented with learning in a Pavlovian conditioning paradigm. Comparing the representational patterns in the US in different brain regions to theoretical model patterns allows to understand what type of information a region represents (e.g. sensory vs hedonic response to painful stimuli). In a second step, I can then determine which theoretical model or information type representation is reactivated by the CS+ allowing me to show whether the sensory or hedonic response to the US is reactivated by the CS.

An extensive body of previous research has thoroughly mapped the neural systems involved in associative learning processes and can inform hypotheses about the content that will be represented with learning. In aversive conditioning, the amygdala has been identified as a key hub for the development of new associations. Univariate neuroimaging studies have demonstrated amygdala activation during fear acquisition, with a direct correlation of neural activity and physiological measures of conditioning (Furmark, Fischer, Wik, Larsson, & Fredrikson, 1997; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). As a result, I would expect the amygdala to represent the hedonic response associated with the US. Research on non-human animals has further delineated the role of the amygdala as a central hub allowing to form new associations between CS and US (Fanselow & Poulos, 2005; J. E. LeDoux, 2000). Given the direct role in forming novel associations between CS and US, it is conceivable that the amygdala might also show evidence of stimulus-stimulus conditioning suggesting that it would also represent sensory aspects of the US. An early review of human neuroimaging findings in fear conditioning identified a brain network that includes not only the amygdala but also the ACC and the insula (Buchel & Dolan, 2000). Since then the involvement of the ACC in fear acquisition has been demonstrated by multiple studies [for review see (Etkin et al., 2011; Greco & Liberzon, 2016)]. Similarly, numerous studies have found support for the role of the insula in aversive conditioning (Bulganin, Bach, & Wittmann, 2014; Sehlmeyer et al., 2009). As two central structures in interoception and emotion regulation (Craig, 2002; Stevens, Hurley, & Taber, 2011) my hypothesis is that ACC and insula will provide evidence for stimulus-response conditioning by representing the hedonic response to the US. The vmPFC has been mostly related to the expression (Sierra-Mercado, Corcoran, Lebron-Milad, & Quirk, 2006) and inhibition (Motzkin, Philippi, Wolf, Baskaya, & Koenigs, 2015) of fear. Thus, I expect the

vmPFC to show evidence of representing the response to or hedonic attachment with the US. Recent research has also shown that even cortical sensory systems show modulation by fear conditioning such that differences in cortical response for CS+ and CS- were observed in extended visual cortex (McTeague, Gruss, & Keil, 2015; Miskovic & Keil, 2013).

The literature on aversive conditioning using painful stimuli that do not necessarily lead to a fear response as employed in the current study is sparser. However, the brain regions and networks underlying the sensory and emotional processing of painful experience are relatively well mapped. Traditionally, somatosensory cortex, ACC, prefrontal cortex, the thalamus and the insula have been related to the sensory experience of pain (Apkarian, Bushnell, Treede, & Zubieta, 2005; Jensen et al., 2016; Schweinhardt & Bushnell, 2010). Critically, brain regions mediating pain processing show strong overlap with regions involved in fear conditioning, which makes the regions of interest (ROIs) described above prime candidates to observe evidence for S-S or S-R learning.

A similar set of brain regions has been shown to be involved in appetitive conditioning. The amygdala has been demonstrated to be critical for development of CS-US associations (Martin-Soelch et al., 2007). As a central hub with connections to prefrontal and striatal regions, the amygdala is important for the integration and relay of information (Everitt et al., 2003; Everitt & Robbins, 2005) and is expected to represent sensory aspects of the US as well as hedonic attachment. The orbitofrontal cortex, which is densely connected with the amygdala, has been demonstrated to be essential for reward evaluations and outcome expectancies (O'Doherty, 2004). As a result, I expect to see a representation of negative and positive value enabling such decision-making processes. The general role of the ACC and insula in emotion regulation (Stevens et al., 2011), interoception (Craig, 2002) and subjective emotional states (Namkung,

Kim, & Sawa, 2018), make them important players in appetitive conditioning. Previous research has associated the insula with expected reward magnitude (B. W. Smith et al., 2009) and anticipatory affect related to gain and loss decisions (Knutson & Greer, 2008). As such, I hypothesize the insula to represent hedonic aspects of the US and CS. In summary, although much is known about the brain regions and networks that mediate aversive and appetitive conditioning in humans, to date it is unknown what aspects of the US is represented and reactivated by a CS in any given region.

Employing RSA allows us to go beyond previous studies using multivariate approaches to understand the development of novel affective associations. As a precedent to the current study, Visser and colleagues (2015; 2013; 2011) employed trial-by-trial RSA in order to examine how the development of associations affects how an initially neutral stimulus (CS+) is represented and categorized in the brain. The authors were able to show that before fear conditioning, neutral stimuli such as faces and houses were represented as more similar to other stimuli of the same category in several brain regions (e.g. superior and inferior frontal gyrus, occipital cortex). However, after repetitively pairing some of the category exemplars with an aversive shock, those paired with a shock, the CS+, and those without pairing, the CS-, were represented more similar to each other irrespective of their initial visual category (i.e. faces and houses) (Visser et al., 2011). That is the initial classification based on visual categories was overwritten by the emotional association. Moreover, previous research demonstrated that long-term fear memory for those stimuli was predicted by neural representations at the time of learning (Visser et al., 2013) in line with my hypothesis that emotional learning patterns inform biases in attention and memory. Another previous study employed an approach similar to the current study. Here researchers compared the neural pattern activation in response to US and

CS+ during a fear conditioning paradigm and tracked how the reactivation of US pattern by the CS+ develops over the course of learning but only in the insula (Onat & Buchel, 2015). Critically, they only found evidence for a pattern reactivation in the anterior insula and not in any other brain region.

Conditioning to appetitive stimuli such as food is equally critical for an animal's survival as is aversive conditioning to e.g. a predator [for review see (Martin-Soelch et al., 2007)]. Yet few studies have ever attempted to directly compare aversive and appetitive conditioning to compare the underlying neural circuity. One challenge to such a comparison is to establish appetitive conditioning paradigms that are comparable with aversive learning. Biologically relevant primary reinforcers such as food have to be found allowing for first order appetitive conditioning (Clark, Hollon, & Phillips, 2012). Food itself can problematic because the rewarding effects of food depend critically on an organism's hunger level. Human researchers have tried to circumvent this difficulty by using secondary reinforcers such as money (Austin & Duka, 2010) or erotic pictures (Klucken et al., 2009). However, neuroimaging studies have shown that while there is overlap in brain activation, partially different valuation system might be involved when associations with primary and secondary reinforcers are developed (Levy & Glimcher, 2011) potentially because different learning systems might be involved depending on the type of reinforcer (Delgado, Jou, & Phelps, 2011). Few human studies have used primary reinforcers such as odor (Gottfried, O'Doherty, & Dolan, 2002), water (Kumar et al., 2008) or food (Prevost, Liljeholm, Tyszka, & O'Doherty, 2012) trying to control for satiety and thirst levels in pure appetitive conditioning paradigms. Another problem arises when aiming to compare aversive and appetitive conditioning in the same study. In a recent study, food was used for appetitive conditioning, while a painful electric shock was employed for aversive

conditioning (Andreatta & Pauli, 2015). While this study confirmed behaviourally successful appetitive and aversive conditioning, the question arises again whether differences in brain activation would be observed that could be explained simply by the nature or sensory modality of the unconditioned stimulus rather than by its inherent valence.

In order to address that problem, in the current study two biologically relevant, tactile US were used. It is well established that within our cutaneous system we can differentiate between fast myelinated fibers that convey information about tactile or sensory stimulation and slower unmylenated fibers, so called C fibers, originating at nociceptors (or pain receptors) in our skin that carry hedonic or affective information about painful stimulation (McGlone & Reilly, 2010). Only recently, have researchers provided evidence that a similar system of unmyelinated fibers, the C-tactile fibers, exists in hairy skin (Vallbo, Olausson, & Wessberg, 1999). Critically, those fibers are activated by soft brush and convey affective information about pleasant touch (Loken, Wessberg, Morrison, McGlone, & Olausson, 2009). Interestingly, stimulation of those unmyelinated afferents has been associated with distinct patterns of activation in the anterior cingulate and orbitofrontal cortex, while primary somatosensory cortex showed less activation (Rolls et al., 2003). That pattern of results suggests that 1) there are direct afferents from the skin to prefrontal and midline regions and 2) the dissociation of painful and pleasurable tactile stimulation is maintained in the cortex. In summary, aversive and appetitive stimulation of the skin presents an ideal, biologically relevant US that allows for direct comparison of both types of conditioning in the brain.

Thus, in the current study participants completed an appetitive and aversive conditioning paradigm in the scanner in which an appetitive brush stroke to the forearm and painful pressure applied to the thumbnail were paired with the presentation of neutral facial expressions. CS-only

blocks were interspersed with conditioning blocks in order to extract representational patterns of CS-only stimuli and CS-US pairings separately. Similarity of US and CS representational patterns in eight regions of interest was extracted on a trial-by-trial basis. I predicted that overall CS would reactivate US representational patterns in brain regions typically associated with conditioning (i.e. amygdala, vmPFC, ACC, insula and higher order visual areas such as the ventral visual stream). Moreover, I predicted that the experimental setup would allow me to answer the question of whether CS reproduce the sensory experience of the appetitive or aversive unconditioned stimulus, or the emotional response or attached meaning to the US. I expected to see those differences in higher order associative regions but less so in primary sensory cortices.

5.2 Materials and Methods

5.2.1 Participants

Data from sixty-one young, healthy participants (age: 21.1 ± 2.8 years, 41 females) was included in the analysis. Initially 107 participants were recruited from Cornell University to participate in a brain imaging study of appetitive and aversive classical conditioning tasks. A number of participants had to be included for different reasons: 20 participants had missing data (imaging run, stimulus onset files, motion correction files) while multi-echo preprocessing described in 5.2.4 failed for 26 participants. All participants gave written, informed consent and had normal or corrected-to-normal vision. Participants were pre-screened for a history of anxiety and depression as well as other psychopathology, epilepsy and brain surgery in addition to general suitability for fMRI data collection. Pre-screening was followed up by an in person interview to ensure inclusion criteria were met. Due to the fact that this study was conducted as part of larger research program, all participants were genotyped.

5.2.2 Materials

5.2.2.1 Stimulus and apparatus. Six faces were chosen from the Karolinska directed emotional faces set, comprising three male and three female exemplars each with a neutral expression (Goeleven, De Raedt, Leyman, & Verschuere, 2008). These faces were used as the conditioned stimuli (CS) in a classical conditioning paradigm. Unconditioned stimuli (US) consisted of either a pressure pain delivered to the right thumb, or an appetitive brush stroke to the participant's forearm. Pressure-pain stimuli were delivered using a custom designed hydraulic device, similar to those used in previous studies (Giesecke et al., 2004; Lopez-Sola et al., 2010) capable of transmitting controlled pressure to 1 cm^2 surface placed on the subjects' right thumbnail. In individual calibration sessions, it was ensured, that the pressure intensity was aversive but not excessively painful. Appetitive brush strokes were manually applied to the right forearm lasting ~4s. Individual subjective responding to brush stimuli were recorded in a separate session prior to all experimental scanning, with only participants who responded positively to the manipulation invited to participate in the scanning session.

5.2.3 Procedure

5.2.3.1 Stimulus ratings. As a measure of subjective stimulus assessment and conditioning, participants were asked to rate the likeability and trustworthiness of the faces used as CS+ and CS- stimuli on a scale from 1-100 (1) before and (2) after conditioning as a measure of conditioning. Due to technical difficulties, stimulus ratings are only available for 40 of the 61 participants included in the analysis.

5.2.3.2 Experimental tasks. While undergoing functional MR scanning, participants completed two unique conditioning tasks with nearly identical structure modeled after Visser et

al. (2015). These tasks differed only in the nature of the tactile unconditioned stimulus (US; see above), and the gender of the face stimuli. In each task, participants completed seven CS-only blocks interleaved with six CS-US paired blocks (see Figure 5.1). Single blocks of either the CS-only or the CS-US pairing entailed one presentation of each face stimulus. The order of the two CS+ and the one CS- was randomized within each CS-US block. Individual trials started with an initial fixation period (19500 ms) followed by the presentation of a facial stimulus (4000 ms). The fixed and long interstimulus interval was included in the experimental design to reduce intrinsic noise correlations (Visser et al., 2013). During CS-only trials, all faces were presented without tactile stimulation (see Figure 5.2). During CS-US paired trials, two of three facial stimuli were paired with tactile stimulation, thus creating two CS+ and one CS- facial stimuli (see Figure 5.3). The US was delivered from the midpoint after the visual stimulus presentation (2000 ms post-onset), and remained for the duration of the visual presentation and after (4000 ms). Face pairings were randomly assigned for each participant but held constant across the duration of the experiment.

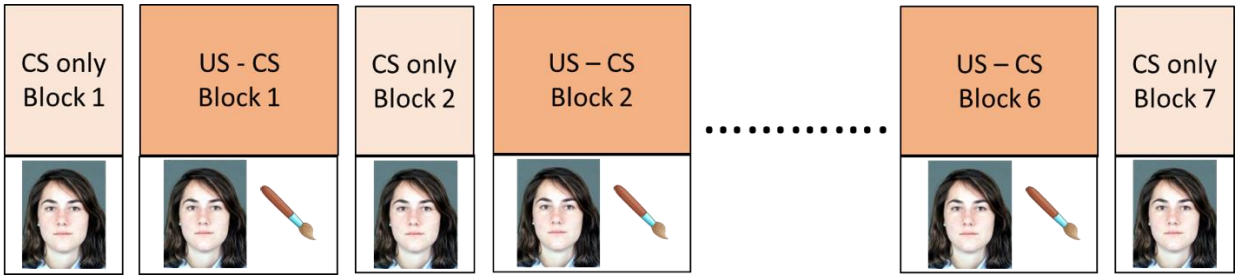

Figure 5.1. General structure of appetitive and aversive conditioning tasks. Both experimental tasks (appetitive and aversive conditioning) followed the same general structure in which 7 CS-only blocks were interleaved with 6 CS-US paired blocks. Displayed is a schematic of the appetitive conditioning task in which faces were either presented by themselves (CS-only block) or with an appetitive brush stroke (CS-US block). In the aversive conditioning paradigm, the brush stroke was replaced with pressure pain.

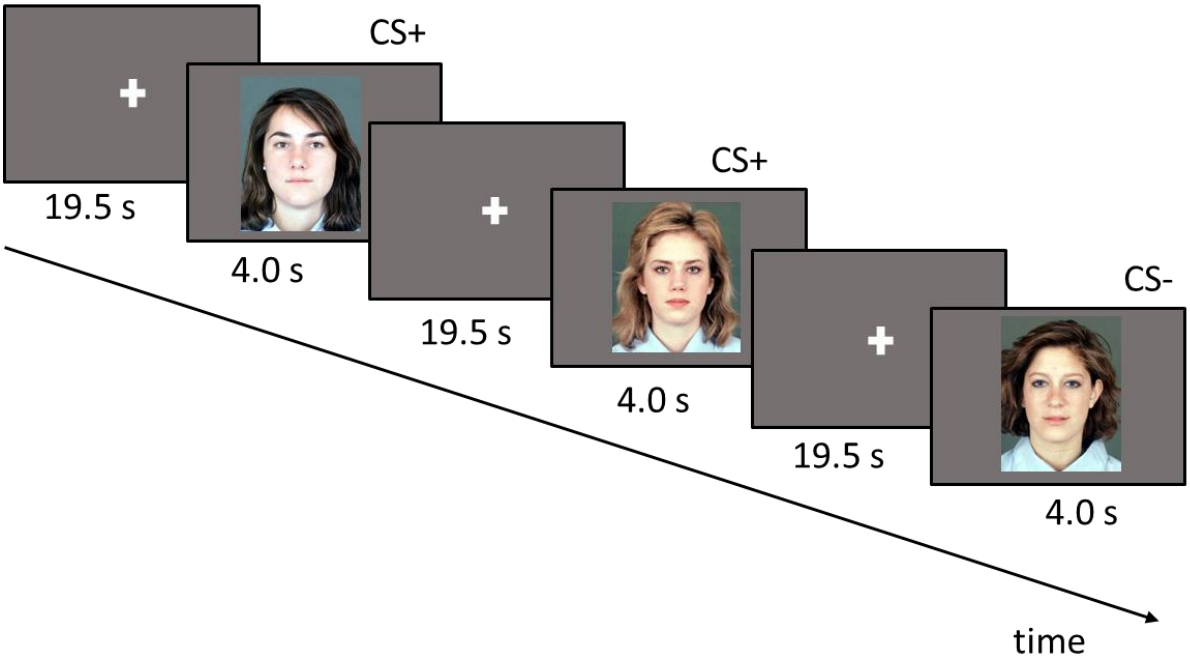

Figure 5.2. Detailed schematic of the CS-only trials. Interstimulus intervals were fixed and a 4 second presentation of the CS. Female faces were presented in the appetitive task and male faces in the aversive conditioning task.

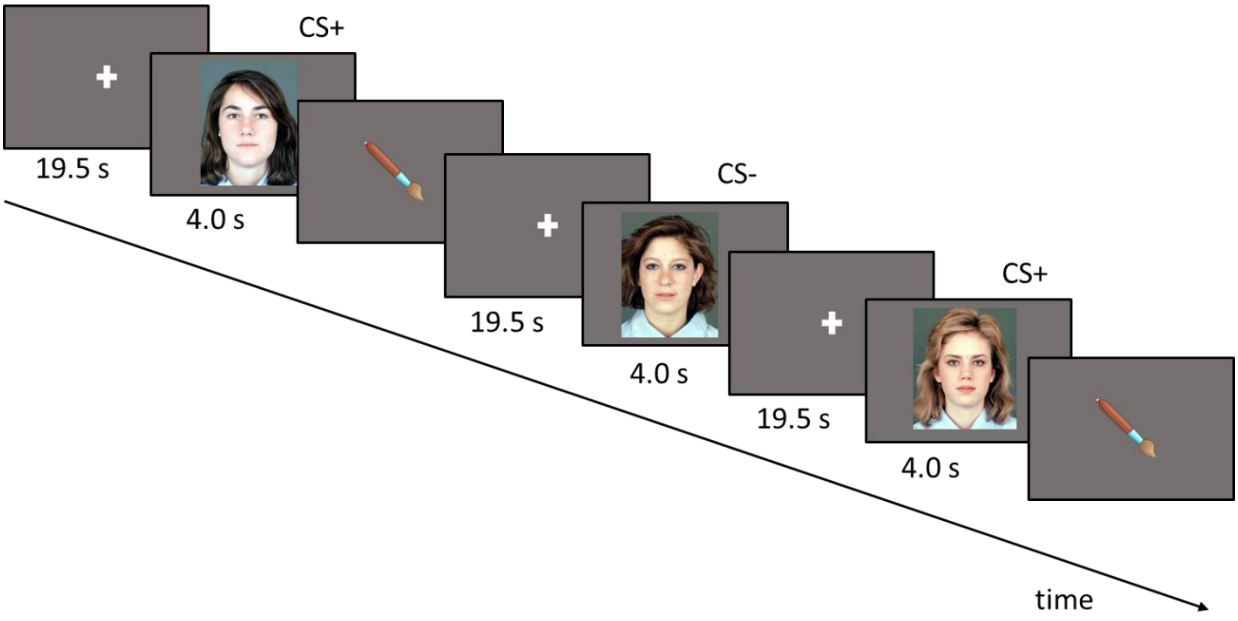

Figure 5.3. Detailed schematic of the appetitive CS-US paired trials. A 4-second presentation of a female face was accompanied by an appetitive brush stroke to the forearm. In the aversive conditioning task, male faces were paired with pressure pain to the thumbnail.

5.2.4 MRI Acquisition and Preprocessing

5.2.4.1 Acquisition. Scanning was conducted on a 3 Tesla GE Discovery magnetic resonance scanner using a 32-channel head coil. For each subject, a T1-weighted MPRAGE sequence was used to obtain high-resolution anatomical images (repetition time (TR) = 7ms, echo time (TE) = 3.42 ms, field of view (FOV) 256 x 256 mm, slice thickness 1mm, 176 slices). The functional tasks were acquired with the following multi-echo (ME) EPI sequence: TR = 2000 ms, TE1 = 11.7 ms, TE2 = 24.2 ms and TE3 = 37.1 ms, flip angle 77°; FOV 240 x 240 mm. A total of 102 slices was acquired with a voxel size of 3 x 3 x 3 mm. Pulse and respiration data were acquired with scanner-integrated devices.

5.2.4.2 Preprocessing. Multi-echo independent component analysis (ME-ICA, meica.py version 3.2 beta1) was used to denoise the multi-echo fMRI data. An optimally combined (OC) dataset was generated from the functional multi-echo data by taking a weighted summation of the three echoes, using an exponential T2* weighting approach (Posse et al., 1999). Multi-echo principal components analysis was first applied to the OC dataset to reduce the data dimensionality. Spatial independent component analysis (ICA) was then applied and the independent component time-series were fit to the pre-processed time-series from each of the three echoes to generate ICA weights for each echo. These weights were subsequently fitted to the linear TE-dependence and TE-independence models to generate F-statistics and component-level κ and ρ values, which respectively indicate blood-oxygen-level-dependent (BOLD) and non-BOLD weightings (Kundu, Inati, Evans, Luh, & Bandettini, 2012). The κ and ρ metrics were then used to identify non-BOLD-like components to be regressed out of the OC dataset as noise regressors (Kundu et al., 2013).

5.2.5 fMRI Analyses: Structurally Determined Regions of Interest

To assess tactile (pressure pain, pleasant brush) and hedonic representations in neural patterns, eight bilateral regions of interest (ROIs) were generated from the standard anatomical atlas (MNI_caez_ml_18) implemented in the Analysis of Functional NeuroImages (AFNI) software package (Cox, 1996): primary somatosensory cortex (S1) and primary visual cortex (V1) were selected as the primary sites of tactile and visual information respectively. In addition, ventral visual stream (VVS) was chosen due to its role in visual classification (Kanwisher, McDermott, & Chun, 1997; Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013). Amygdala, ventromedial prefrontal cortex (vmPFC) (Mackey & Petrides, 2014), anterior cingulate cortex (ACC) and insula were further selected for their hypothesized roles in affect processing (A. K. Anderson & Phelps, 2002; Chikazoe, Lee, Kriegeskorte, & Anderson, 2014) and pain representations (Kragel, Kano, et al., 2018; Orenius et al., 2017). The insula was further divided into an anterior and posterior portion due to its functional and anatomical subdivisions (Nieuwenhuys, 2012) for a total of eight ROIs

5.2.6 fMRI Analysis

Data analysis of the fMRI data was conducted using Analysis of Functional NeuroImages (AFNI) software (Cox, 1996). Regressor files of interest were generated for all individual trials across the experiment, modelling the time course of each stimulus presentation during each run (36 total events). The relevant hemodynamic response function was fit to each regressor to perform linear regression modeling. This resulted in a β coefficient and t value for each voxel and regressor. To facilitate group analysis, each individual's data was transformed into the standard brain space of the Montreal Neurological Institute (MNI)

In order to identify the representational pattern elicited by the experimental stimuli, representational similarity analysis (RSA) was performed using the Python package PyMVPA (Hanke et al., 2009). For each participant, in each ROI the spatial pattern of β weights in response to each experimental condition or event was correlated with the pattern of activation in response to all other events. This step was performed separately for each ROI. Thus, pairwise correlation coefficients for all experimental events of a single ROI resulted in a similarity matrix containing correlations for all CS-US combinations for all trials for each participant. Fischer transformations were performed on all similarity matrices to allow comparisons between participants.

5.2.7 fMRI Analyses: Model specification

In order to characterize representations of sensory and hedonic information associated with experience of pain/pleasure of the US, ideal representation models (IRMs) were generated to represent dissociable correlation patterns that would be observed in the experimental data if it would contain perfect representation of distinct theoretically-derived constructs (Fig 5.4). Similarity matrices were constructed for 1) Experimental Task, 2) Active Touch, 3) Specific Touch Types, 4) Pleasant Brush, 5) Pressure Pain, 6) Touch Valence, 7) Positive Events, 8) Negative Events, 9) All Valence, 10) Salience, 11) Facial Stimulus, 12) Violation of Expectation and 13) Temporal Adjacency (see Figure 5.4). 1) 'Experimental Task' contrasts the two tasks – the difference between which are male and females faces and tactile stimulation used in aversive and appetitive condition tasks respectively. 2) 'Active Touch' represents all active tactile stimulation by pressure pain or pleasant brush in the same way, while 3) 'Specific Touch Types' predicts high similarity only between the same type of tactile stimulus: pressure pain, pleasant

brush and no active touch but tactile input from the scanner environment. 4) 'Pleasant Brush' and 5) 'Pressure Pain' represent specific active touch in distinct ways, but in a nonlinear fashion, while 6) 'Touch Valence' contrasts negative and positive tactile stimulation. 7) 'Positive Events' and 8) 'Negative Event' represent the active positive and negative stimulation as well as the omission of painful and the lack of pleasant stimulation respectively. 9) 'All Valence' contrasts any positive or negative events with neutral events, while 10) 'Salience' represents any salient event: active tactile stimulation or violation of expected touch. 11) 'Facial Stimulus' represents individual facial stimuli. 12) 'Violation of Expectation' predicts high similarity for CS- stimuli when a participant might expect tactile stimulation but does not receive any. 13) 'Temporal Adjacency' represents high similarity between two temporally adjacent stimuli within the same experimental task. Of note, the order of CS+1, CS+2 and CS- pairings with US was randomized in each block.

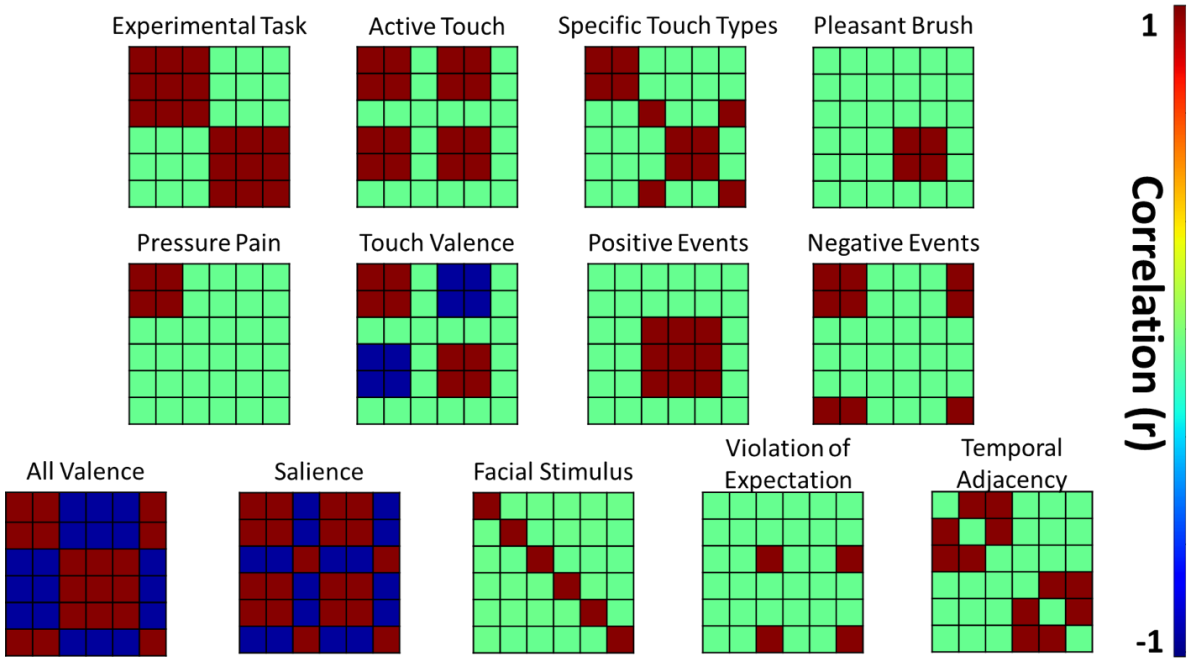

Figure 5.4. Ideal Representation Models (IRMs) demonstrate the representational pattern that would be observed in the experimental data if a region would represent the theoretically derived constructs displayed.

5.2.8 fMRI Analyses: Model fit selection and comparison

In order to determine the IRM combinations that best explain the observed correlation in the US data in each ROI a Bayesian Information Criterion (BIC) analysis was conducted using a *greedy best-first algorithm (GFBS)* (Doran & Michie, 1966) (see Figure 5.5a). BIC is a measure for model fit, with smaller values indicating better fit (Schwarz, 1978). For this analyses, correlation matrix transformations were performed using Matlab (The MathWorks, Natick, Massachusetts, USA) and BIC analysis and multiple regression were conducted in R (RCoreTeam, 2013). Multiple regression was performed on the best IRM combination in order to obtain regression coefficients and thus being able to determine each IRM's independent contribution to the observed data (see Figure 5.5a).

Using multiple regression, CS-only data was compared to the best fitted combination of IRMs identified for the US data (see Figure 5.5b). Critically, CS-only data was split up into early, mid and late conditioning in order demonstrate the temporal development of representational pattern reactivation over the course of learning (see Figure 5.5b). Regression coefficients were used to quantify the extent to which the reconstructed US patterns were reactivated by CS and the significance of individual IRM reactivations.

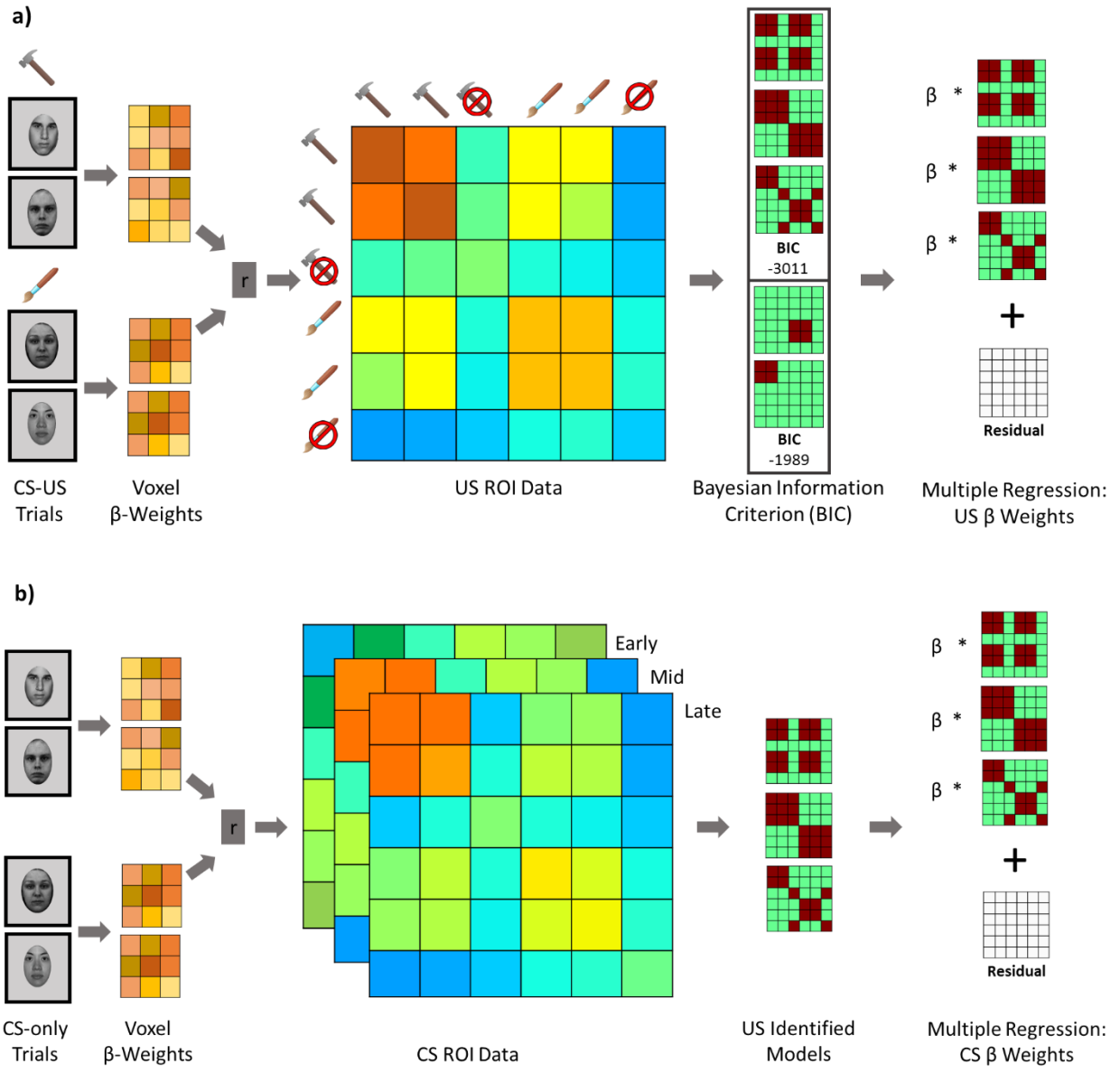

Figure 5.5. Data analysis pipeline. a) Voxel activation in each ROI, averaged over all CS-US trials for each task condition, was correlated with each other task condition by means of representational similarity analysis (RSA). A Bayesian Information Criterion (BIC) score was used to find the best combination of ideal representation models for each ROI. Multiple regression was used to obtain beta coefficients for each model. b) RSA was performed separately on early, mid, and late CS-only trials. Multiple regression was used in order to quantify the US model reactivation by CS-only data.

5.3 Results

5.3.1 Stimulus ratings

Likeability and trustworthiness ratings for CS+ stimuli in the appetitive and aversive task as well as CS- ratings from the aversive task were averaged and compared between pre and post conditioning as a measure of conditioning. The analysis revealed that CS+ stimuli in the appetitive task were perceived as significantly more likeable and trustworthy after relative to before conditioning, $t(39) = -3.11$, $p = .003$, indicating that the appetitive associations with the pleasant brush stroke changed the hedonic response to the CS+. The comparison of pre and post CS+ rating from the aversive task did not reveal a significant difference, $t(39) = 1.63$, $p = .112$. However, the CS- used in the aversive conditioning task was rated as significantly more likeable and trustworthy after conditioning, $t(39) = -2.31$, $p = .026$. This pattern of results suggest in the aversive task, the stimulus not paired with an aversive outcome changed its hedonic properties more as a result of being perceived as safety signal (Marshall, Acheson, Risbrough, Straus, & Drummond, 2014). Overall stimulus ratings suggest successful conditioning in both domains.

In the following sections I will present the results from the three main analysis steps. First, ideal representation models (IRMs) for the US representational patterns were identified in order to show what type of information each region of interest represents. Second, the IRMs were used to build the ideal reconstructed US (rUS) pattern, which in turn was used to show reactivation of that representational pattern by CS using multiple regression. Third, the central question of this chapter was answered by employing multiple regression to show how well each IRM identified for the US predicts the pattern of activation elicited by the CS. This allowed me to answer the question of whether more sensory or hedonic aspects of the US are reproduced by

the CS. Critically, the third step was only completed for those ROIs that showed significant reproduction of US patterns by CS.

5.3.2 US ideal model strengths

Primary Somatosensory Cortex (S1): Model comparisons and multiple regression revealed a significant relationship between the representational pattern elicited by the US in S1 and the ideal representation models 'Active Touch', 'Experimental Task' and 'Specific Touch Types'. The three predictors explained 23.8 % of the variance ($R^2 = .238$, $F(3,1277) = 132.70$, $p < .001$). For individual predictor weights refer to Table 5.1. The pattern of results indicates a representation primarily driven by tactile information.

Visual Cortices: Best model fit for representational pattern in V1 was found for the IRM 'Experimental Task'. 2% of the variance ($R^2 = .020$, $F(1,1279) = 26.21$, $p < .001$) was explained by this single predictor (see Table 5.1) suggesting that visual cortex activation is best explained by a model solely based on visual information. In contrast, the ventral visual stream represented both visual and hedonic/tactile information as demonstrated by a combination of the IRMs 'Experimental Task' and 'Pressure Pain ($R^2 = .031$, $F(2,1278) = 20.43$, $p < .001$) (see Table 5.1).

Amygdala: A combination of three IRMs best fit the representational pattern of the amygdala. 'Touch Valence', 'Active Touch' and 'Experimental Task' best predicted pattern of brain activation in the amygdala ($R^2 = .114$, $F(3,1277) = 54.56$, $p < .001$) (see Table 5.1) suggesting that the amygdala represents tactile information and the hedonic information associated with it as well as the salience of facial stimuli.

Ventromedial Prefrontal Cortex (vmPFC): 'Pressure Pain', 'Pleasant Brush' and 'Temporal Adjacency' were all significant predictors of the US representational pattern of the

vmPFC (R^2 = .083, $F(3,1277) = 38.77$, $p < .001$) (see Table 5.1). The ventromedial Prefrontal Cortex thus represents the hedonic value attached to the tactile stimulation, and represents positive and negative value in a nonlinear fashion.

Anterior Cingulate Cortex (ACC): 'Pressure Pain', 'Active Touch', 'Experimental Task' were identified as the best fitting combination of IRMs for the ACC (R^2 = .156, $F(3,1277) = 78.28$, $p < .001$) (see Table 5.1) pointing towards a representation of (negatively valenced) tactile information.

Insula: The pattern of activation in the anterior insula was best explained by a combination of 'Pressure Pain', 'Active Touch' and 'Negative Events' (R^2 = .143, $F(4,1276) = 53.01$, $p < .001$), suggesting the representation of hedonic, especially negative information. A similar pattern of results was observed for the posterior insula with 'Active Touch', 'Pressure Pain' and 'Experimental Task' being the strongest predictors (R^2 = .360, $F(3,1277) = 239.10$, $p < .001$).

5.3.3 US representational pattern reactivation by CS

The goal of the next step was to determine the extent of US pattern reactivation by CS. To do so, CS-only matrices were fitted to the ideal reconstructed US (rUS) pattern for each ROI, and performed separately for early, mid and late conditioning (see Figure 5.5). rUS patterns were developed by summing all contributing IRMs, scaled to the associated β, as identified by BIC and multiple regression analyses performed on the US data. rUS were used in these analyses to help minimize noise in US representational patterns.

Visual Cortices: V1 CS representational patterns were significantly predicted by V1 rUS during early (R^2 = .004, $F(1,1279) = 4.81$, $p < .05$) and late (R^2 = .005, $F(1,1279) = 6.42$, $p <$

.05) conditioning, with β-coefficients of β = .721 and β = .933 respectively. V1-rUS marginally predicted V1 CS representation pattern during mid conditioning (R^2 = .003, $F(1,1279) = 3.67$, $p < 0.10$), with β = .661. Across VVS regions, the rUS pattern was a significant predictor of representational patterns at all conditioning time points (*Early:* β = .697, R^2 = .006, $F(1,1279) = 7.09$, $p < .05$; *Mid:* β = .856, R^2 = .007, $F(1,1279) = 9.36$, $p < .05$; *Late:* β = .733, R^2 = .004, $F(1,1279) = 5.65$, $p < .05$). This indicates minimal effect of affective pairings on representational patterns in these visual regions.

Primary Somatosensory Cortex (S1): Multiple regression did not reveal a significant relationship between the S1-CS representational patterns and S1-rUS patterns during early conditioning (R^2 = .002, $F(1,1279) = 2.62$, $p > .10$), but did for both mid (R^2 = .006, $F(1,1279) = 7.82$, $p < .05$) and late (R^2 = .004, $F(1,1279) = 4.91$, $p < .05$) conditioning. β-coefficients of S1-rUS predictors for early, mid and late conditioning were β = .186, β = .353 and β = .286 respectively. The pattern of results indicates that effects of conditioning are visible in early somatosensory regions.

Amygdala: CS patterns were not found to be significantly related to Amygdala rUS patterns at any conditioning time point (*Early:* β = -.190, $R^2 < .001$, $F(1,1279) = 0.786$, $p > .10$; *Mid:* β = .312, R^2 = .002, $F(1,1279) = 2.04$, $p > .10$; *Late:* β = .346, R^2 = .002, $F(1,1279) = 2.30$, $p > .10$) potentially due to limited spatial resolution masking effects of conditioning.

Ventromedial Prefrontal Cortex: CS representational patterns in vmPFC were not significantly predicted by vmPFC rUS pattern during early conditioning (β = -.258, R^2 = .001, $F(1,1279) = 1.30$, $p > .10$). As conditioning progressed, both mid (R^2 = .007, $F(1,1279) = 8.48$, $p < .05$) and late (R^2 = .005, $F(1,1279) = 6.37$, $p < .05$) CS representational patterns were predicted by the vmPFC rUS, with β-coefficients of β = .738 and β = .629 respectively. As the vmPFC rUS

contains both pleasant brush and pressure pain information, this pattern of results suggests that activation patterns in vmPFC carries information representative of successful conditioning for both appetitive and aversive associations.

Anterior Cingulate Cortex: Similar to vmPFC, ACC rUS was not significantly related to ACC CS representational patterns at early conditioning mid ($R^2 < .001$, $F(1,1279) = 0.71$, $p > .10$), but highly predictive at both mid ($R^2 = .011$, $F(1,1279) = 14.59$, $p < .001$) and late conditioning mid ($R^2 = .013$, $F(1,1279) = 16.93$, $p < .001$). β-coefficients of ACC rUS predictors for early, mid and late conditioning were $\beta = -.128$, $\beta = .606$ and $\beta = .674$ respectively. These results indicate that ACC may represent affective associations acquired though classical conditioning.

Insula: A significant relationship between the anterior insula CS representational patterns and anterior insula rUS patterns was not detected for early conditioning ($R^2 < .001$, $F(1,1279) = 0.23$, $p > .10$), but for both mid ($R^2 = .016$, $F(1,1279) = 20.62$, $p < .001$) and late ($R^2 = .017$, $F(1,1279) = 21.74$, $p < .001$) conditioning. β-coefficients of rUS predictors for early, mid and late conditioning in anterior insula were $\beta = .058$, $\beta = .605$ and $\beta = .674$ respectively. This suggests that activation patterns in anterior insula are highly representative of conditioned affective associations. The posterior insula rUS pattern was not significantly predictive of the CS representational pattern during early conditioning ($\beta = .058$, $R^2 < .001$, $F(1,1279) = 0.49$, $p > .10$), was significantly predictive at mid conditioning ($\beta = .291$, $R^2 = .010$, $F(1,1279) = 12.48$, $p < .001$), and marginally predictive at late conditioning ($\beta = .159$, $R^2 = .002$, $F(1,1279) = 3.08$, $p < .10$) providing evidence for CS reactivation of US representational patterns.

5.3.4 US component reactivation by CS

The primary goal of this study was to examine what information associated with experiencing the US becomes represented with learning. Thus, after having established that, at least in certain brain regions, the US representational pattern is reactivated by the CS, the next goal was to determine what type of information is carried forward in conditioning, i.e. tactile or hedonic information. A reactivation of models 'Pressure Pain', 'Pleasant Brush', 'Touch Valence', 'Negative Events' and 'Positive Events' particularly *without* reactivation of general 'Active Touch' representational patterns is interpreted as reactivation of hedonic information. While it cannot be ruled out that tactile aspects are reproduced as well, the differential reactivation of pure tactile and hedonic/tactile models strongly suggests hedonic reactivation. I investigated the components of US reactivation elicited by the CS by performing multiple regression analysis for those ROIs that showed significant overall reactivation of the US representational pattern. Here the combinations of best fitting IRMs identified for each ROI in the previous step were predictors and the CS representational pattern was the outcome variable. Critically, this step was performed separately for early, mid and late conditioning in order to evaluate temporal development of reactivation. For full results see Table 5.3.

Visual Cortices: 'Experimental Task' was a significant predictor for the V1 CS representation pattern early ($\beta = .04$, $p = .029$, $R^2 = .004$, $F(1,1279) = 4.81$, $p = .004$) and late ($\beta = .052$, $p = .009$, $R^2 = .005$, $F(1,1279) = 6.42$, $p = .011$), and a marginal predictor for mid ($\beta = .037$, $p < .10$, $R^2 = .003$, $F(1,1279) = 3.67$, $p = .056$) conditioning (see Figure 5.6a), indicating that the representation of visual information does not depend on learned associations. In contrast, the ventral visual stream showed significant component reactivation for 'Experimental Task' only in early conditioning ($\beta = .046$, $p = .011$, $R^2 = .006$, $F(2,1278) = 4.04$, $p = .018$), while

reactivation of 'Pressure Pain' was only visible mid conditioning (β = .04, p = .021, R^2 = .005, $F(2,1278) = 5.25$, p = .005). This pattern of results suggests that in the VVS, CS reactivate visual information before, but hedonic information after affective associations have developed (see Figure 5.6b).

Primary Somatosensory Cortex: For CS activation in S1, significant fit of the US-identified component 'Experimental Task' was found for the mid conditioning period only (β = .081, p = .0001, R^2 = .011, $F(3,1277) = 4.94$, p = .002), which could indicate residual representation of tactile information from interspersed CS-US blocks. 'Touch Type' was a significant predictor of CS activation only late in conditioning (β = .061, p = .022, R^2 = .012, $F(3,1277) = 5.34$, p = .001), pointing towards a representation of distinct types of touch by the CS (see Figure 5.6c).

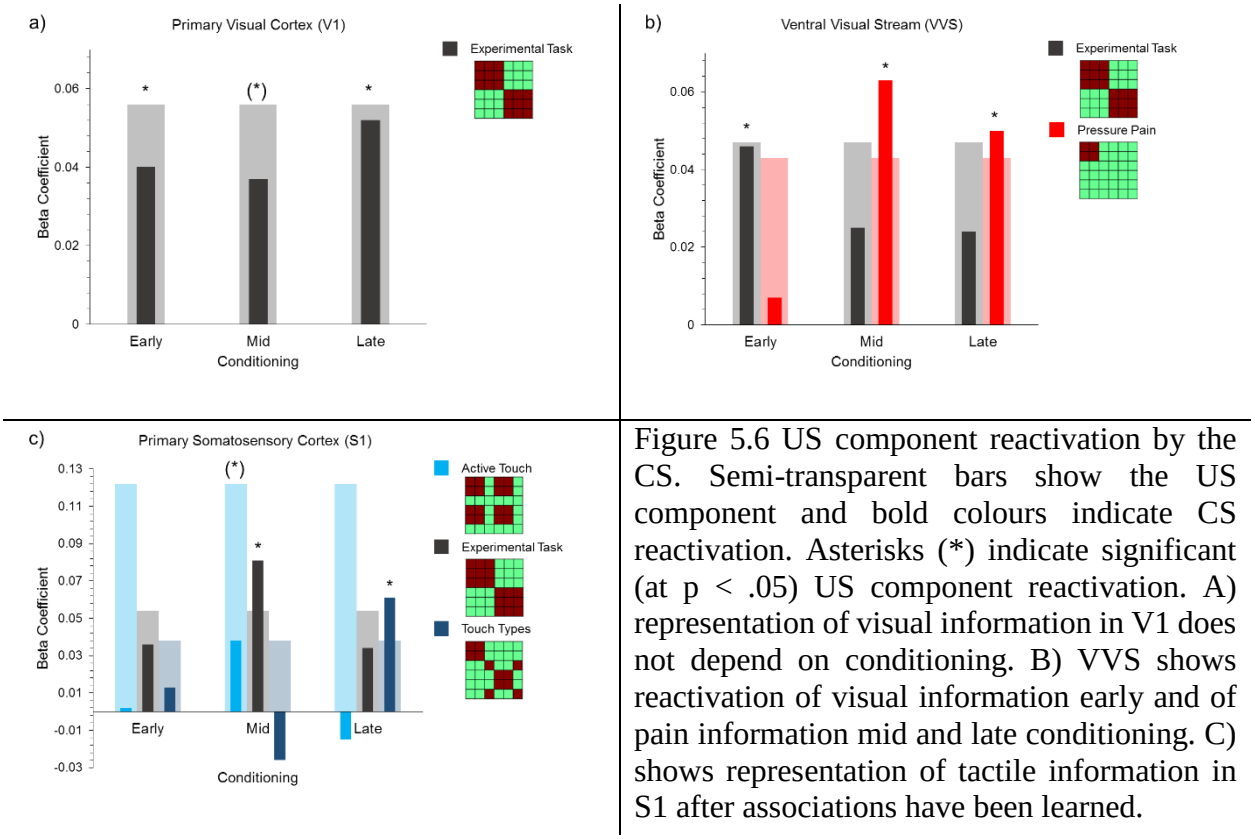

Figure 5.6 US component reactivation by the CS. Semi-transparent bars show the US component and bold colours indicate CS reactivation. Asterisks (*) indicate significant (at p < .05) US component reactivation. A) representation of visual information in V1 does not depend on conditioning. B) VVS shows reactivation of visual information early and of pain information mid and late conditioning. C) shows representation of tactile information in S1 after associations have been learned.

Ventromedial Prefrontal Cortex: Early in conditioning the CS representational pattern in the vmPFC was not significantly predicted by the US identified combination of IRMs (R^2 = .001, F(3,1277) = .61, p = .609). However, mid conditioning, both 'Touch Valence' (β = .05, p = .036) and 'Active Touch' (β = .06, p = .010) were significant predictors of the CS activation pattern. 'Touch Valence' was also a significant predictor of the CS representational pattern late in conditioning (β = .07, p = .003). Overall, the CS reactivates a significant portion of activity initially elicited by the US both in mid (R^2 = .008, F(3,1277) = 3.29, p = .020) and late (R^2 = .007, F(3,1277) = 3.00, p = .030) conditioning (see Figure 5.7a). Both 'Pressure Pain' and 'Pleasant Brush' were reactivated once conditioned associations were established, suggesting that in the vmPFC, aversive and appetitive information is carried forward in conditioning in a nonlinear fashion.

Anterior Cingulate Cortex: Significant US component reactivation by CS was observed mid (R^2 = .015, F(3,1277) = 6.26, p < .001) and late (R^2 = .016, F(3,1277) = 6.80, p < .001) conditioning. The effect was driven by the significant predictor models 'Pressure Pain' (mid: β = .060, p = .020, late: β = .059, p = .033) and 'Experimental Task' (mid: β = .038, p = .023, late: β = .044, p = .013) (see Figure 5.7b). The analysis reveals a strong conditioning effect in the ACC, where significant component reactivation is only observed after learning. The pattern of results further suggests a representation of hedonic but not tactile or sensory information by the CS.

Insula: The anterior insula showed clear conditioning effects, with significant US component reactivation mid (R^2 = .02, F(4,1276) = 6.91, p < .001) and late (R^2 = .023, F(4,1276) = 7.48, p < .001) but not early (R^2 = <.001, F(4,1276) = .22, p = .925) in conditioning (see Figure 5.6c). Critically, only 'Pressure Pain' (mid: β = .145, p < .001, late: β = .150, p < .001) and 'Negative Events' (mid: β = -.068, p = .004, late: β = -.053, p = .035) showed

component reactivation, indicating that negative hedonic information is carried forward during conditioning in the anterior insula. The posterior insula on the other hand, only showed significant prediction by 'Experimental Task' mid ($\beta = .047$, $p = .005$, $R^2 = .019$, $F(3,1277) = 8.11$, $p < .001$) and late conditioning ($\beta = .049$, $p = .001$, $R^2 = .014$, $F(3,1277) = 5.96$, $p < .001$), which might represent the differences between appetitive and aversive conditioning task, i.e. tactile and facial stimuli. Overall, the current data indicates distinct patterns of reactivation in anterior and posterior insula supporting the notion that these subregions are functionally different (see Figure 5.7d).

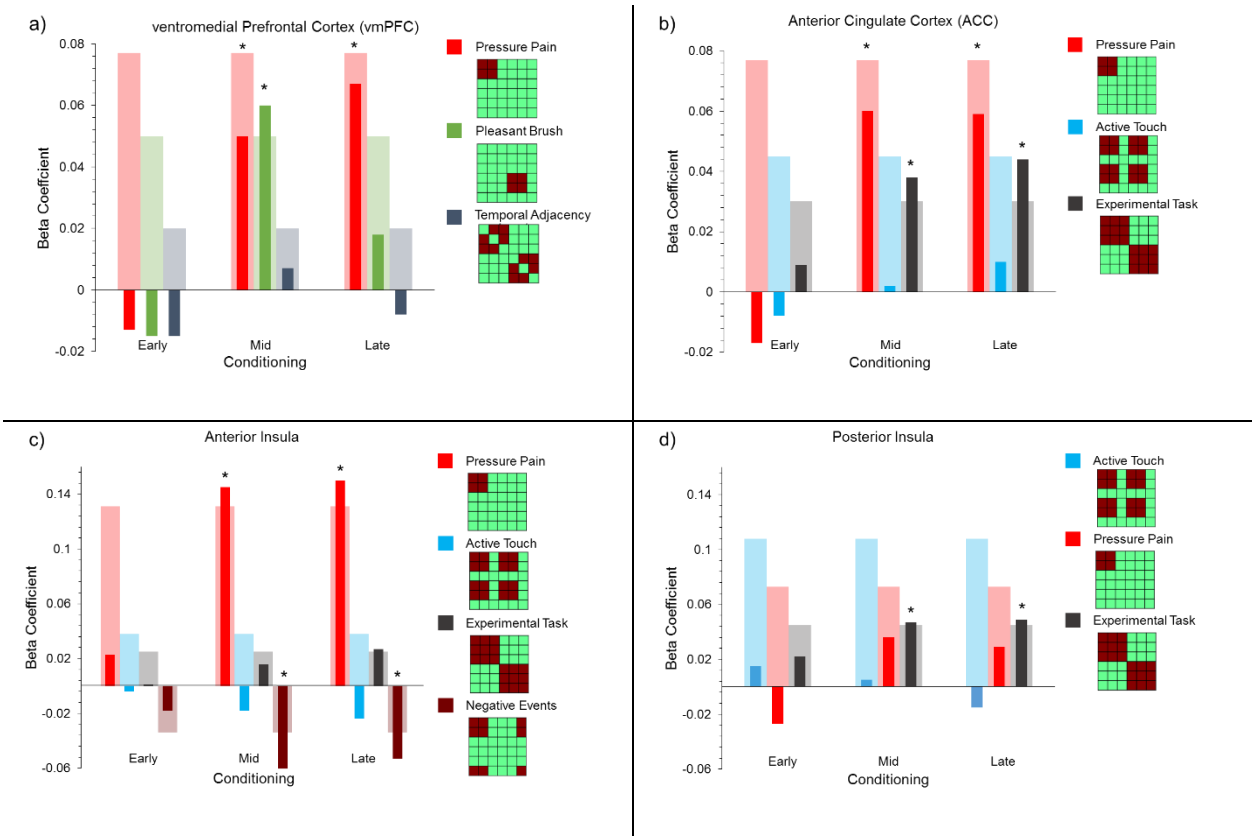

Figure 5.7 US component reactivation by CS. Semi-transparent bars show US component, bold colours indicate CS reactivation. Asterisks (*) indicate significant (at $p < .05$) US component reactivation. A) vmPFC, B) ACC and C) anterior insula show strong conditioning effects where hedonic information is represented after learning only. D) Posterior insula shows a distinctively different pattern of results from anterior insula.

Table 5.1 US Ideal Representation Models (IRMs) as identified by multiple regression

			US Ideal Representation Models (IRMs) – β Coefficients												
ROI	**R^2**	**P value**	**ET**	**AT**	**STT**	**PB**	**PP**	**TV**	**PE**	**NE**	**AV**	**S**	**FS**	**VE**	**TA**
V1	.020	3.19e-7	.056*	-	-	-	-	-	-	-	-	-	-	-	-
VVS	.031	1.84e-9	.047*	-	-	-	.043*	-	-	-	-	-	-	-	-
S1	.238	2.2e-16	.054*	.122*	.038*	-	-	-	-	-	-	-	-	-	-
Amy	.114	2.2e-16	.033*	.037*	-	-	-	.020*	-	-	-	-	-	-	-
vmPFC	.083	2.2e-16	-	-	-	.050*	.077*	-	-	-	-	-	-	-	.020*
ACC	.156	2.2e-16	.030*	.045*	-	-	.077*	-	-	-	-	-	-	-	-
aIns	.143	2.2e-16	.025*	.038*	-	-	.131*	-	-	-.034*	-	-	-	-	-
pIns	.360	2.2e-16	.045*	.108*	-	-	.073*	-	-	-	-	-	-	-	-

Table 5.2. US reactivation by CS representation patterns in early, mid and late conditioning

	US Reactivation by CS Conditioning								
	Early			**Mid**			**Late**		
ROI	**R^2**	**P value**	**β**	**R^2**	**P value**	**B**	**R^2**	**P value**	**B**
V1	.004	.028	.721	.003	.056	.661	.005	.011	.933
VVS	.006	.008	.697	.007	.002	.856	.004	.018	.733
S1	.002	.106	.186	.006	.005	.353	.004	.027	.286
Amy	<.001	.375	-.19	.002	.153	.312	.002	.13	.346
vmPFC	.001	.255	-.258	.007	.004	.738	.005	.012	.629
ACC	<.001	.401	-.128	.011	<.001	.606	.013	<.001	.696
aIns	<.001	.631	.058	.016	<.001	.605	.017	<.001	.674
pIns	<.001	.486	.054	.010	<.001	.291	.002	.079	.159

V1 = primary visual cortex; VVS = ventral visual stream; S1 = primary somatosensory cortex; Amy = amygdala; vmPFC = ventromedial prefrontal cortex; aIns = anterior insula; pIns = posterior insula; ACC = anterior cingulate cortex

* indicates p <.05; m indicates p < .10

Table 5.3 US Ideal Representation Models (IRMs) predicting CS representation pattern

				CS Ideal Representation Models (IRMs) – Beta Coefficients												
ROI	**Time**	**R^2**	**P value**	**ET**	**AT**	**STT**	**PB**	**PP**	**TV**	**PE**	**NE**	**AV**	**S**	**FS**	**VE**	**TA**
	early	.004	.029	.040*	-	-	-	-	-	-	-	-	-	-	-	-
V1	*mid*	.003	.056	.037[m]	-	-	-	-	-	-	-	-	-	-	-	-
	late	.005	.011	.052*	-	-	-	-	-	-	-	-	-	-	-	-
	early	.006	.018	.046*	-	-	-	.007	-	-	-	-	-	-	-	-
VVS	*mid*	.008	.005	.025	-	-	-	.063*	-	-	-	-	-	-	-	-
	late	.005	.047	.024	-	-	-	.050[m]	-	-	-	-	-	-	-	-
	early	.005	.101	.036	.002	.013	-	-	-	-	-	-	-	-	-	-
S1	*mid*	.011	.002	.081*	.038[m]	-.026	-	-	-	-	-	-	-	-	-	-
	late	.012	.001	.034	-.015	.061*	-	-	-	-	-	-	-	-	-	-
	early	.001	.609	-	-	-	-.015	-.013	-	-	-	-	-	-	-	-.015
vmPFC	*mid*	.008	.020	-	-	-	.060*	.048*	-	-	-	-	-	-	-	.007
	late	.007	.030	-	-	-	.017	.067*	-	-	-	-	-	-	-	-.007
	early	.001	.706	.009	-.008	-	-	-.017	-	-	-	-	-	-	-	-
ACC	*mid*	.015	<.001	.038*	.002	-	-	.060*	-	-	-	-	-	-	-	-
	late	.016	<.001	.044*	.010	-	-	.059*	-	-	-	-	-	-	-	-
	early	<.001	.925	.001	-.004	-	-	.023	-	-	-.018	-	-	-	-	-
aIns	*mid*	.021	<.001	.016	-.018	-	-	.145*	-	-	-.068*	-	-	-	-	-
	late	.023	<.001	.027	-.024	-	-	.150*	-	-	-.053*	-	-	-	-	-
	early	.003	.266	.022[m]	.015	-	-	-.027	-	-	-	-	-	-	-	-
pIns	*mid*	.019	<.001	.047*	.005	-	-	.036[m]	-	-	-	-	-	-	-	-
	late	.014	<.001	.049*	-.015	-	-	.029	-	-	-	-	-	-	-	-

V1 = primary visual cortex; VVS = ventral visual stream; S1 = primary somatosensory cortex; vmPFC = ventromedial prefrontal cortex; ACC = anterior cingulate cortex; aIns = anterior insula; pIns = posterior insula; * indicates p <.05; [m] indicates p < .10

5.4 Discussion

The principle underlying classical conditioning is that an initially neutral stimulus, the CS, produces the same response behaviourally as an inherently positive or negative US after repetitive paired exposure to the two stimuli. The central question of this chapter was whether a CS also reactivates the pattern of neural activity that was initially elicited by the US and what information about the CS is represented. In contrast to pure behavioural investigations, a neuroimaging approach allowed me to go beyond demonstrating that the basic principle of classical conditioning is preserved in the brain. Instead, it enabled me to ask what component of the US representation is reactivated – the sensory experience or the affective attachment with the stimulus. In other words, do we observe evidence of S-S or S-R associations? Using representational similarity analysis on aversive and appetitive classical conditioning data revealed that for most regions of interest included in the current analysis a significant amount of variance in the CS data could be explained by the US neural activity, indicating US pattern reactivation by CS. Critically, this effect was only found mid to late conditioning but not early in the learning process. By creating ideal representation models of what a brain region could represent and finding the best combination of models for each ROI, I was further able to show the distinct type of information that is represented in each brain region and which type of information is being reactivated by the CS. In summary, it appears that primary sensory regions reactivate components of the US that do not rely on learning while the other brain regions included in the analysis show a relatively clear reactivation of those components representing hedonic value of the US – providing evidence for a dominance of stimulus-response learning. Critically, this effect is only observed mid and late but not early in conditioning suggesting that what the CS is representing is markedly changed by its conditioned associations.

US representational pattern reactivation by CS. The first central question of this chapter was whether CS reactivate the representational pattern elicited by the US once conditioned associations are established. In more general terms, this allowed me to answer the question whether as a result of the learning process, representation of CS becomes highly similar to that of US - suggesting that through conditioning the initially neutral stimulus is represented in the same way as something inherently positive or negative. The results from the current study suggest that indeed overall pattern reactivation is prominent in most regions of interest included in the analysis. Taking a closer look at different ROIs shows a more nuanced pattern of results. Primary visual and ventral visual cortices show significant reactivation of US representational patterns early, mid and late conditioning indicating that reactivation is not observed because of conditioned associations but rather because of the visual properties of the stimuli that did not change as a result of learning. In contrast, primary somatosensory cortex, vmPFC, ACC as well as anterior and posterior insula show significant prediction of CS representational pattern by US data only *after* conditioned associations have developed. To my knowledge this is the first study to demonstrate that the basic mechanisms observed behaviourally in conditioning - that is that the CS elicits the same response as the US (Maren, 2001; Martin-Soelch et al., 2007; Pavlov, 2010) - is also observed in the brain.

Finding evidence for a representation of conditioned associations in vmPFC, ACC and insula is in line with previous research. The vmPFC has been related to expression (Sierra-Mercado et al., 2006) and inhibition (Motzkin et al., 2015) of fear during conditioning as well as reward evaluations and outcome expectancies (O'Doherty, 2004) consistent with studies showing vmPFC lesions are accompanied by insensitivity for positive or negative future consequences (Bechara, Tranel, & Damasio, 2000). Thus, it is not surprising to see that the vmPFC shows US

reactivation by CS once conditioning is established. The ACC has been shown to be involved in fear acquisition [for reviews see (Etkin et al., 2011; Greco & Liberzon, 2016)] and in normal contingency learning during conditioning (Jackson et al., 2016) making it a critical player in associative learning processes, which means it should share representations of the US allowing for conditioning to occur. Last, the insula has been shown to play a role in aversive conditioning (Sehlmeyer et al., 2009) as well as in general in interoception (Craig, 2002). Thus, as an important region involved in subjective emotional states (Namkung et al., 2018) it was expected to see reactivation of the representation of the unconditioned stimuli that are accompanied by positive and negative sensory experience and presumably emotions. In the next section, we will take a closer look at which specific aspects of the US are reactivated by the CS in these different ROIs with the goal to shed light on the debate about S-S and S-R associations.

In summary, several brain regions show reactivation of US representational patterns by CS consistent with their involvement in conditioning processes, with the exception of the amygdala. At first glance, that finding might seem surprising given the central role of the amygdala especially in aversive conditioning (J. E. LeDoux, 2000). One possible explanation however is that while the amygdala is a central hub in conditioning that integrates cortical and subcortical information allowing the formation of new associations between CS and US (Fanselow & Poulos, 2005; J. E. LeDoux, 2000; Martin-Soelch et al., 2007), previous research has suggested that only subnuclei of the amygdala (e.g. the central nucleus) may be involved in the behavioural expression of conditioned associations (Duvarci, Popa, & Pare, 2011; J. E. LeDoux, 2000; Zimmerman, Rabinak, McLachlan, & Maren, 2007). The rather coarse approach of treating the amygdala as a homogenous region due to limits in spatial resolution with

conventional fMRI employed in the current study might have masked reactivation of the US pattern as a marker of conditioned response.

US ideal representational models. In order to determine what aspect of the US and in turn the CS is represented in each region of interest, 13 ideal representation models were developed that display dissociable correlation patterns that would be observed if the US was perfectly represented by the underlying theoretically-derived constructs.

US activation in primary and ventral visual cortices was best explained by the IRM 'Experimental Task' presumably due to the differences in visual input, i.e. male and female faces used in the aversive and appetitive conditioning tasks respectively. This result was expected especially for VVS which includes the fusiform face area – a face-selective region in our brain (Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher et al., 1997). More variance of the US representation in VVS is explained when 'Pressure Pain' is included as an additional model. While a representation of pressure pain, which might represent sensory or hedonic information, in the visual cortex might seem surprising, visual cortex activation has been shown to be modulated by emotional information through connections with the amygdala [e.g., (Furl, Henson, Friston, & Calder, 2013)]. The best-fitting combination of IRMs identified for primary somatosensory cortex are less surprising. Consistent with its role as the primary site for processing and perception of tactile information [see for example (Preusser et al., 2015)], US representation in S1 was best explained by models representing and differentiating different types of touch, indicating that S1 is primarily representing tactile aspects of the US. A representation of mostly hedonic aspects of the US was observed in amygdala, vmPFC, insula and ACC. More specifically, US representation in the amygdala is best captured by a mix of valenced and neutral tactile information in line with the view that the amygdala processes

information that is salient or relevant for an organism at any given point rather than being a pure detector and processor of threatening stimuli (Cunningham & Brosch, 2012). Nonlinear representation of pressure pain and pleasant touch in the vmPFC is consistent with its role in association learning and previous research showing representation of primary reinforcers such as somatosensory stimuli as well as the distinct response to pleasant and painful touch [for review see (Rolls, 2004)]. Anterior and posterior insula as well as ACC show a very similar pattern of IRMs. Those regions seem to represent both tactile information in general ('Active Touch') as well as valenced information ('Experimental Task'), with a more prominent representation of negative information ('Pressure Pain'). This pattern of results is consistent with the role of the insula and ACC in emotional awareness in general (Craig, 2002; Gu, Hof, Friston, & Fan, 2013) and in pain processing in particular (Jensen et al., 2016).

In summary, the current results suggest that primary visual and somatosensory cortices primarily represent visual and tactile aspects of the US, while the other regions included in the analysis show a more hedonic-value driven representation.

US component reactivation by CS. As established above, most regions of interest included in the current analysis showed reactivation of the US representational pattern by CS. In this next step, I will discuss results pertaining to the exact pattern of reactivation by the CS in order to answer the question whether a CS reproduces the representation of the US as a whole or just certain characteristics, e.g. tactile vs hedonic information. In other words, do we observe S-S or S-R conditioning? Primary visual and ventral visual cortices showed a presumably visually-driven component reactivation that did not depend on conditioning consistent with the fact that visual input did not change over the course of learning. Critically, however, VVS shows reactivation of the component 'Pressure Pain' once conditioned associations have developed in

line with previous findings that indices of conditioning can be detected in extended visual cortex (McTeague et al., 2015; Miskovic & Keil, 2013) and with the notion that S-R learning might have occurred. For S1 most components, especially general 'Active Touch', were not reproduced. This is in agreement with the general assumption that the neural basis of conditioning cannot be found in primary sensory but higher order emotional brain regions (J. E. LeDoux, 2000; Martin-Soelch et al., 2007). One reason for the sporadic reactivation of some tactile components could be that activation from the interleaved US blocks might bleed into the CS-only task blocks. The vmPFC, anterior and posterior insula as well as the ACC similarly show clear effects of conditioning, in that component reactivation is only observed mid and late conditioning but not early in conditioning. In the vmPFC, only 'Pressure Pain' and 'Pleasant Brush' are reactivated, suggesting that the CS only represents the hedonic information that became associated with it, again suggesting S-R and not S-S associations are formed. Of note, significant reactivation of pleasant brush is not observed late in conditioning which might indicate that habituation to the appetitive brush is faster than to pressure pain (Triscoli, Ackerley, & Sailer, 2014). Interestingly, while the US representation in anterior and posterior insula showed large overlap, a distinctive role for the two regions was observed during conditioning. The anterior insula showed reactivation of components representing negatively valenced information while the posterior insula activation is best represented by 'Experimental Task' – a model that contrasts appetitive and aversive conditioning. Much research has been done on the role of the insula and its subregions with somewhat inconsistent results. A recent meta-analysis (Kurth, Zilles, Fox, Laird, & Eickhoff, 2010) revealed that most functions investigated have been associated with activation in the anterior portion, especially emotional processing (Craig, 2002). Only sensorimotor processing has been exclusively mapped onto the posterior insula while pain

processing has been shown to involve the entire insula. In the light of the meta-analysis findings, the current results of reactivation of negative information representation in the anterior insula might indicate reactivation of hedonic or affective US components. In contrast, the 'Experimental Task' reactivation in the posterior insula could be explained by the differences in sensory input between appetitive and aversive tasks, i.e. female faces paired with pleasant brush and male faces paired with pressure pain respectively. In the ACC 'Pressure Pain' and 'Experimental Task' were both reactivated while similar to anterior and posterior insula 'Active Touch' showed no reactivation suggesting again that the CS becomes associated with the hedonic value associated with the painful stimulation. Thus, insula and ACC provide additional evidence for the idea that S-R conditioning is observed.

Overall the analysis of US component reactivation by CS demonstrated clear effects of conditioning: All but primary sensory regions included in the analysis, showed component reactivation only once conditioned associations were established. Moreover, all brain regions that have been previously associated with conditioning show reactivation of components that carry hedonic information with them instead of pure tactile information. Thus, the pattern of results observed here suggests that what is carried forward in conditioning and becomes associated with the CS is the affective attachment with the stimulus rather than the pure sensory experience. In other words, when we are exposed to a conditioned stimulus we re-experience the pleasant or unpleasant feelings elicited by the US rather than the sensory stimulation. Taken together, the current analysis provides strong support for the notion of stimulus-response conditioning in which a CS reproduces the unconditioned response rather than the perceptual experience of the US.

The present study comes with several limitations that should be acknowledged. As touched on above, imaging certain brain regions such as the amygdala comes with a number of problems that influence how the results can be interpreted. Due to the small size of the brain region and the relatively large voxel size (3-4mm) often used in research studies, data from the amygdala relies on a pattern of activation obtained from a limited number of data points leading to reduced signal-to-noise ratio (S. Robinson, Windischberger, Rauscher, & Moser, 2004). Moreover, studies have shown that signal presumably measured in the amygdala is likely confounded by nearby blood vessels (Boubela et al., 2015). Another limitation comes with the, albeit hypothesis-driven, selection of ROIs in contrast to a more data driven approach (Elliott, Cheruvelil, Montgomery, & Soranno, 2016; Shih & Chai, 2017). In order to address this problem, in a follow-up analysis a whole-brain parcellation approach could be used to verify the current findings and extend the results by including more parts of the brain. To date, only subjective likeability and trustworthiness ratings are available as a measure of conditioning. However, physiological measures of conditioning such as skin conductance or heart rate variability would provide an additional measure of conditioning that would enhance the interpretability of the results (Liu, Wei, Kuang, Zhao, & Tsien, 2013; Pineles, Orr, & Orr, 2009).

Two additional future approaches to the dataset involve testing for individual differences and investigate extinction. I am planning to investigate individual differences in at least two different ways. First, I would like to extract ideal representation models for each brain region for each participant individually and then look whether the individual US components predict CS reactivation in some participants but not others or whether certain participants show a more sensory biased reproduction and others a more hedonic reproduction. Second, I would like to use individual differences in the ability to condition such as likeability ratings and physiological data

to classify participants as conditioners and non-conditioners and compare the ability and pattern of US reactivation on a group level. For extinction it will be of interest to see whether we see a reversal of the learning that is described in the current chapter. It would be expected that by the end of the extinction period, the pattern of activation elicited by the CS no longer resembles that of the US; we do not see a behavioural response to a CS after extinction after all. A critical question however is whether the representation of the CS is updated to its initial neutral representation or whether the representation after extinction is significantly different from the before conditioning state which would indicate that the stimulus is now associated with an extinction memory (Dunsmoor, Niv, Daw, & Phelps, 2015).

In conclusion, in the current chapter I demonstrated for the first time that in conditioning a conditioned stimulus reactivates the pattern of activation initially elicited by the unconditioned stimulus. This pattern of results is observed in a number of different brain regions typically associated with emotional learning processes such as vmPFC, ACC and insula. The results further show that it is primarily the hedonic components of the initial US experience that is reproduced when being exposed to a CS suggesting stimulus-response learning. Or in other words that what the CS carries forward in conditioning is the meaning we attach to a US not the sensory experience. Thus, this study is to my knowledge the first to provide novel evidence for the decades old debate about S-S and S-R conditioning by employing contemporary multivariate analysis approaches to neuroimaging data.

Chapter 6: General Discussion

The aim of the current dissertation was to investigate the role of the stress system as an individual difference factor in emotional learning processes that can underlie affective biases, and the neural mechanisms underlying the development of novel affective associations. In Chapter 2, I established that genetically and environmentally based differences in the norepinephrine and stress system influence affective bias manifestation. Here putatively higher NE levels associated with a common genetic variation were associated with a tendency to perceive ambiguous stimuli as more rewarding overall. In a series of studies presented in Chapter 3 and 4, I demonstrated that acute stress influences different types of appetitive emotional learning underlying the formation of such affective biases. More specifically, I was able to show that acute stress can have detrimental effects on both classical and operant conditioning. In addition, my research has shown that depending on the timing of the stressor relative to the learning task, acute stress can have opposing effects on operant conditioning. In Chapter 5, I investigated the neural mechanisms underlying the development of novel associations through classical conditioning. Using contemporary multivariate approaches to analyzing neuroimaging data, I was able to show, for the first time, that in classical conditioning the conditioned stimulus reactivates the pattern of neural representation elicited by the unconditioned stimulus. Importantly, study results provide evidence that the conditioned stimulus represents the hedonic response to or attachment with the unconditioned stimulus rather than the perceptual features of the stimulus itself. Taken together, the set of studies that comprises this dissertation provide novel evidence for the role of stress in appetitive biases and their formation through emotional learning, as well as the neural mechanisms through which those associations are developed.

6.1 Individual differences in appetitive perceptual biases and conditioning

6.1.1 Effects of the norepinephrine/stress on affective bias flexibility

The *ADRA2b* deletion variant (Small et al., 2001), a common genetic variation, putatively associated with increased NE availability (de Quervain et al., 2007) has been reliably shown to be associated with enhanced affective biases in memory (de Quervain et al., 2007), attention (Todd, Muller, et al., 2013b) and perception (Todd, Ehlers, et al., 2015). Moreover, previous studies have suggested that emotional learning depends on the alpha2b noradrenergic receptors *ADRA2b* is coding for (Moriceau & Sullivan 2004). Here I presented results showing that deletion carriers show stronger affective biases than non-carriers; however, bias flexibility was only dependent on stress exposure (Ehlers, Ross, & Todd, 2018).

The current findings suggest that genotype-dependent differences in the NE system create differences in perceptual biases consistent with previous findings of differences in affective biases (de Quervain et al., 2007; Todd, Ehlers, et al., 2015; Todd, Muller, et al., 2013b). Notably, carriers of the deletion variant actually showed an overall more positive bias for ambiguous information consistent with the recently proposed role of norepinephrine in the processing of appetitive stimuli (Ehlers & Todd, 2017b; Weinshenker & Schroeder, 2007). However, critically, the study results also suggest that once affective biases have been established, tonic differences in NE availability do not make them more malleable. In contrast, acute stress still seems to have an effect on the short-term shift in perceptual biases providing evidence that tonic, genetically based and phasic or acute differences in the norepinephrine and stress system might play distinct roles as individual difference factors. Genetically based differences in NE availability might lead to chronic differences in neural gain (Aston-Jones & Cohen, 2005). A recent study has suggested that increased gain associated with greater NE availability narrows attention to stimuli that an

individual is predisposed to attend to and strengthens underlying neural connections (Eldar et al., 2013). Deletion carriers might experience higher gain at all times which might reinforce their existing enhanced affective biases but reduces bias flexibility. It is certainly conceivable that it is adaptive to ensure that stimuli that have been tagged as emotionally or motivationally relevant are less susceptible to short-term influences especially in a situation of high NE availability.

6.1.2 Stress and arousal in appetitive learning

Different researchers favour different definitions of emotion. Some define emotion as a collection of basic emotional states such as fear or anger (Ekman, 1992), whereas others support the view that emotions elicit states of approach or avoidance (Elliot, Eder, & Harmon-Jones, 2013; Lang & Bradley, 2008; Rolls, 2000). Many theories agree, however, that emotion and motivation can be organized in a two-dimensional space along the two axes of *direction* and *intensity* [for example (Hebb, Martinez, & Glickman, 1994; Schneirla, 1959)] or *valence* and *arousal* (Kensinger & Corkin, 2004). On the axis of valence, we find threat on one end of the spectrum and reward on the other end (Rolls, 2000). Notably, relatively isolated programs of research [for reviews see (Chiew & Braver, 2011; Pessoa, 2009)] have investigated motivation in decision-making and threat processing in affective science. Behavioural research in those isolated domains is paralleled by a divide in brain research. On the one hand, research on reward related processes (Pessiglione et al., 2006) has focused on dopaminergic striatal action [for example (de la Fuente-Fernandez et al., 2002; Haber, 2011)], while investigations concerning fear and threat put their emphasis on the role of the norepinephrine and arousal system mediated by the amygdala [for example (Delgado et al., 2011; J. LeDoux, 2003; Onur et al., 2009)].

However recent advances have been made by considering and more closely examining the role of stress and arousal in appetitive or motivational processes [for reviews see (Ehlers & Todd, 2017b; Weinshenker & Schroeder, 2007)]. The current thesis greatly adds to this area of research by demonstrating that stress can influence different types of appetitive learning. In Chapter 3, I demonstrated that at least certain aspects of both classical and operant conditioning can be negatively affected by exposure to acute stress (Ehlers & Todd, 2017a). In the series of studies presented very simple forms of associative learning were investigated in order to close the gap in human literature by demonstrating that not only habit formation (Schwabe et al., 2011; Schwabe, Tegenthoff, et al., 2010; Schwabe et al., 2012; Schwabe & Wolf, 2009) but also the emotional learning processes underlying it are influenced by acute stress. Together with data presented in Chapter 4, the experiments further inform the debate about timing effects of stress on learning processes. A compelling theory suggested that the effects of stress on learning and memory depend on the timing of the stressor relative to the learning experience (Joels et al., 2006). Informed by hippocampally-based memory or fear conditioning studies, but not by research from reward learning, the theory suggests the immediate, NE-driven stress response enhances (McGaugh, 2013) while the late stress response mediated by glucocorticoids impairs learning and memory (D. J. de Quervain et al., 1998). The current thesis provides some evidence in support for the theoretical framework and it extends it by adding findings from appetitive learning: I found that the immediate stress response can facilitate operant conditioning processes while only reward sensitivity but not accuracy in a forced-choice operant conditioning task were reduced under the influence of the late stress response. Overall my dissertation contributes to the integration of research in the domains of motivation and emotion with the goal to no longer treat them as two different entities but rather as two extremes on a continuum.

The findings presented in the current thesis raise the question for potential underlying mechanisms especially in the domain of neurotransmitters and neural mechanisms. As mentioned above, previous research has established norepinephrine and the amygdala as key players in processing fear and threat (Onur et al., 2009). However, the locus coeruleus has been shown to respond with norepinephrine release to both direct reward and punishment (Sara, 2009) and hence to salient stimuli in general rather than to threatening information only. This view is consistent with the updated view of the amygdala as a detector of salience in general rather than one of threat in particular (Cunningham & Brosch, 2012; Peck & Salzman, 2014b) as well as the findings presented in Chapter 4 of facilitated reward learning under the influence of NE. It has been further suggested that prioritization of salient information by the LC-NE system is accomplished by increasing the signal-to-noise ratio or selectivity during information processing (Aston-Jones & Cohen, 2005; Markovic et al., 2014; Mather, Clewett, Sakaki, & Harley, 2016b). A more nuanced view of NE in appetitive learning, coming from single-cell recording studies in monkeys, suggests that NE is required for signaling action associated costs (Bouret et al., 2012) and the integration of cue and outcome information (Bouret & Richmond, 2015). The recent turn to noradrenergic and stress action in motivation and reward learning is paralleled by research on dopaminergic influences on fear conditioning [for example (Abraham, Neve, & Lattal, 2014; Darvas, Fadok, & Palmiter, 2011; Fadok, Dickerson, & Palmiter, 2009)].

6.2 Neural mechanisms underlying development of novel affective associations

In the current thesis I presented results showing that once a conditioned association is established, a CS reactivates the representational pattern initially elicited by the US, thus neural representations mirror the basic principle underlying classical conditioning. However, a

multivariate neuroimaging approach allowed me to go beyond simply showing CS reactivation. Instead, I also provided novel evidence speaking to the debate of S-S and S-R learning in classical conditioning (Holland, 1990) or whether a CS reproduces the hedonic response or sensory stimulus features of the US. I was able to show that those brain regions that are typically involved in conditioning and interoception such as vmPFC, insula and ACC (Craig, 2002; Kim & Jung, 2006; Martin-Soelch et al., 2007) reproduce hedonic aspects of the US but not sensory stimulus properties thus supporting S-R learning.

How can these results help us understand how environmental factors such as stress can enhance or impair such associative learning processes? Let us consider the conditioning paradigm used in the scanner, in which faces were passively associated with pain or pleasure by simultaneous presentation. I can imagine several mechanisms through which learning could be impaired or enhanced by stress or other factors and hence create individual differences in performance: On the one hand, glucocorticoid action as a result of HPA axis activation, might elevate the processing threshold for information not related to the stressful event (Joels et al., 2006), which might impair learning ability, which could reduce or slow down the development of the full reactivation of US representations by CS. This reduced processing of information after a stressful event might be adaptive since it allows an organism to restore homeostasis (S. M. Smith & Vale, 2006). Another way in which stress could impair the ability to learn novel associations properly is in a differential conditioning paradigm as used in the current study, in which one stimulus is never paired with the US, the CS-. Difficulties in differentiating CS+ and CS- can result in generalization of learned associations. Such patterns of fear generalization are characteristic of anxiety disorders (Duits et al., 2015). Indeed NE has been shown to increase fear generalization in context conditioning (Asok, Kandel, & Rayman, 2018). On the other hand,

NE might enhance associative learning processes. It is well established that NE can enhance fear conditioning (Giustino & Maren, 2018; Tully & Bolshakov, 2010). The current results could indicate that under the influence of NE, an aversive or appetitive US appears more salient (Mather, Clewett, Sakaki, & Harley, 2016a), which could in turn evoke a stronger unconditioned or hedonic response speeding up the development of the CS representation in the brain.

In summary, the current findings about the neural mechanisms underlying the formation of novel associations might add to our understanding of how individual differences in the speed, robustness and specificity of the learning process arise. In the next section I will touch on how those differences can in turn shape our attentional biases.

6.3 Implications for affective bias formation

Biases in attentional prioritization do not only influence what we perceive, encode and remember [for reviews see (Awh, Belopolsky, & Theeuwes, 2012; Jiang, Swallow, & Rosenbaum, 2013; Shomstein & Gottlieb, 2016; Todd & Manaligod, 2018)] but are themselves the product of learning and memory. For example, response times in a vocalization task were markedly slowed down if task material was superimposed on a CS+, which was explained in terms of the attention grabbing properties of the conditioned association (Merckelbach, van Hout, de Jong, & van den Hout, 1990). Research from my lab has demonstrated that conditioned associations with a highly arousing traumatic experience result in strong attentional biases for trauma-related stimuli and predict anxiety symptoms (Lee et al., 2013; Todd, MacDonald, et al., 2015). Those results are in line with an extensive body of literature showing enhanced fear conditioning in patients with anxiety disorders (Lissek et al., 2008; Lissek et al., 2005; Lissek et al., 2009; Wilker et al., 2014). Likewise, it has been shown that reward learning creates

attentional biases for reward-related stimuli (B. A. Anderson, 2016a), which in turn have been shown to be elevated in addiction (Field & Cox, 2008). Thus, considerable evidence suggests that affective biases for specific stimuli can develop through emotional learning processes and that individual differences therein present itself in differences in vulnerability for certain psychopathological conditions.

While not directly investigated in the current thesis, the question for possible underlying mechanisms arises that could explain how individual differences in emotional learning translate into affective biases. As discussed above noradrenergic action as a result of the immediate stress response might facilitate associative learning processes by enhancing the salience of the US (Mather et al., 2016b). On the one hand, we can imagine enhanced learning of and resultant affective biases for rewarding stimuli providing potential for addictive tendencies (B. A. Anderson, 2016a). On the other hand, heightened response to e.g. fearful stimuli and increased generalization can promote the development of biases characteristic of anxiety disorders or trauma (Duits et al., 2015). But not only enhanced learning can become problematic. Difficulties developing emotional associations under stress can impair our ability to establish positive behaviours or habits (Schwabe & Wolf, 2011).

6.4 Emotional learning and affective biases: Relevance for psychopathology

The research discussed so far is of direct relevance to psychopathology since anxiety, depressive and addictive disorders are characterized by abnormalities in emotional learning and affective biases.

Aversive conditioning. Alterations in the neural circuit mediated by the LC-NE system are thought to underlie maladaptive patterns of fear learning expressed as fear and anxiety

disorders such as PTSD (Etkin & Wager, 2007; Liberzon et al., 1999). Fear learning is of course highly adaptive and critical for animals' well-being and survival. In situations of potential or actual threat or danger, rapid fear and defense mechanisms - including the release of NE and stress hormones – are activated (Lupien & McEwen, 1997; Morilak et al., 2005). However, fear and stress responses are adaptive only when the timing and level of their activation is appropriate to the situation. A dysregulation of fear response or defensive behaviour can develop into a fear or anxiety disorder (Rosen & Schulkin, 1998). For example, posttraumatic stress disorder (PTSD) is an anxiety disorder characterized in part by attentional biases to mild stressors or cues related to the traumatic event that gave rise to the disorder as well as intrusive memories of the traumatic event (Rau, DeCola, & Fanselow, 2005; Todd, MacDonald, et al., 2015). Pavlovian fear conditioning has been widely used as an animal model for PTSD contributing to the current understanding of the disorder (Mahan & Ressler, 2012). Animal models of fear conditioning and human studies with PTSD patients and healthy controls provide evidence for a critical role of NE in this example of disordered fear learning. For example, patients with PTSD show greater baseline cerebral spinal fluid NE concentrations (Geracioti et al., 2001). Much research on NE and PTSD has focused primarily on symptoms of the disorder or the fear response. As a result, many studies have demonstrated that pharmacological manipulation of the NE system can either worsen (Bremner et al., 1997; Olson et al., 2011) or alleviate symptoms of PTSD (Davidson, 1992; Petrakis et al., 2012; Taylor et al., 2006) depending on the site of action. Moreover, carriers of the *ADRA2b* deletion variant, putatively associated with enhanced NE availability, showed greater susceptibility to intrusive traumatic memory than non-carriers (de Quervain et al., 2007).

Thus, there is substantial convergent evidence linking PTSD, as an example of a disorder thought to be the result of disrupted fear learning, to altered noradrenergic transmission in fear learning and possible memory modulation. I further speculate that NE modulated alterations in fear learning observed in patients with PTSD may give rise to robust attentional biases for trauma-related cues observed in patients (Todd, MacDonald, et al., 2015), demonstrating that specific affectively biased attentional sets develop as a result of individual differences in associative learning. Future research should test this hypothesis directly.

Appetitive conditioning. A prevailing view in the addiction literature is to characterize addiction as a disorder of appetitive learning (Everitt & Robbins, 2016): On the one hand, drugs act as reinforcers, such that the rewarding effect of the drug leads to enhanced drug-taking. On the other hand, environmental stimuli that become associated with the drug effects can acquire incentive salience through Pavlovian conditioning (Everitt & Robbins, 2005). An important component of addiction is an imbalance of goal-directed and habitual behaviour. In the beginning, drug taking or substance use is a goal-directed process guided by the reinforcing properties or the 'liking' of the drug. However, over time behaviour can shift towards the habitual. That is, 'wanting' or craving for the substance develops irrespective of the rewarding outcome and often despite accompanying negative consequences – a process shown to be dependent on dopaminergic action (K. C. Berridge, 2007). Thus, instead of relying on action-outcome relations, addicts show a high degree of stimulus-response instrumental responding. Support for this idea can be found in both human and non-human animal research [for review see (Everitt & Robbins, 2016)]. These findings raise the question of what determines whether behaviour shifts from goal-directed to habitual and what may make some people more prone to experience the shift. I propose that the stress and norepinephrine system contributes to this shift,

and that individual differences therein may underlie differences in vulnerability to addictive habits. As described earlier, in some contexts high NE levels have been associated with more rigid, less flexible behaviour (Eldar et al., 2013). Thus, either transient elevation of NE levels (e.g. by acute stress) or altered NE availability based on genotype (e.g. *ADRA2b* polymorphism) may explain greater predisposition to maladaptive habit formation observed in some individuals. In fact, both human and non-human studies have revealed that chronic or acute stress can bias behaviour towards the habitual (Dias-Ferreira et al., 2009; Graham, Yoon, & Kim, 2010; Schwabe & Wolf, 2009) adding to the literature showing that acute stress — and resultant NE and corticosteroid action — elevate drug self-administration and promote relapse (Piazza & Le Moal, 1998; Sinha, 2008). Pavlovian learning has also been shown to be a factor in drug addiction since environmental and drug-related cues can promote craving, drug-taking and relapse (Everitt & Robbins, 2016). As described earlier, associative learning can largely modulate attentional biases - for example to drug-related cues - which in turn guide or control our behaviour. Biases to those reward-related cues (B. A. Anderson, 2016a) can in turn inform instrumental behaviour through Pavlovian Instrumental Transfer (PIT), in which an initially neutral cue that becomes associated with the drug may elicit instrumental behaviour such as drug-taking. Critically, PIT has likewise been demonstrated to be promoted by acute stress (Pool et al., 2015).

While addiction is characterized by attentional biases associated with increased approach motivation, the opposite picture is present in patients with major depressive disorder (MDD). Anhedonia – the inability to experience pleasure – is a cardinal symptom of depression (Hasler, Drevets, Manji, & Charney, 2004; Kasch, Rottenberg, Arnow, & Gotlib, 2002). Importantly anhedonia is characterized by reduced attentional biases to reward (Wang, Brennen, & Holte,

2006). This again is thought to be due to altered patterns of associative learning observed in depression (Kumar et al., 2008; Must et al., 2006; Vrieze et al., 2013). A number of studies have suggested that patients with depression display a deficit in approach motivation, are less responsive to rewards and show reduced activation in reward circuitry [for review see (Bogdan, Nikolova, & Pizzagalli, 2013)]. A recent study employed a computational meta-analysis to formalize the relation between anhedonia and reinforcement learning and to answer the question of whether MDD patients simply show reduced reward sensitivity or whether the ability to learn from a reward signal is impaired (Huys et al., 2013). Results suggested that the actual learning rate - that is, the speed with which the action-outcome association is established - is not affected in patients with depression. However, patients show overall reduced effort and willingness when working for the same reward as controls, suggesting that their reward sensitivity is reduced (Huys et al., 2013). Consistent with the proposed link between attentional biases and associative learning processes, patients with anhedonia display altered reward learning as well as reduced attentional biases (Brailean, Koster, Hoorelbeke, & De Raedt, 2014; Pizzagalli, Iosifescu, Hallett, Ratner, & Fava, 2008). This suggests that altered learning processes indeed give rise to differences in attentional prioritization related to psychopathology. There is additional evidence that acute stress affects reward sensitivity (Berghorst et al., 2013; Bogdan & Pizzagalli, 2006; Cavanagh et al., 2011).

In summary, a large body of literature suggests that stress mediated alterations and individual differences in the appetitive associative learning system give rise to specific patterns of biased attention. Attentional biases can both be strengthened (e.g. addiction) and weakened (e.g. depression) through reward learning and can develop into deeply habitual patterns of orienting to the world that underlie the etiology and maintenance of psychopathology.

6.5 Limitations and future directions

While the current work represents an important step towards understanding the influence of stress and arousal on appetitive conditioning and the neural mechanisms through with emotional associations develop, the work also comes with some limitations and opportunities for future research. One major limitation of the stress studies is that we did not directly measure noradrenergic activity. In the future, a combination of salivary alpha-amylase levels and pupil dilation could be used to get measures of sympathetic nervous system activity before and after stress induction (Thoma et al., 2012) and locus coeruleus activation (Joshi et al., 2016). Behavioural experiments could be extended by testing timing effects of stress on classical conditioning as well as generalization and discrimination in conditioning. Individual differences in classical conditioning could further be investigated by examining the phenomenon of sign- and goal-tracking (Flagel et al., 2009; Flagel et al., 2008; T. E. Robinson & Flagel, 2009) that has been sparsely investigated in humans (Garofalo & di Pellegrino, 2015) but shows some overlap with the here discussed debate of S-S and S-R learning (Holland, 1990). A main follow up of the neuroimaging data analysis will be to investigate individual differences and extend the results to extinction. In the context of individual differences, I would like to see whether the individual representation of the US predicts CS reactivation and subsequently whether different subgroups of associative learning styles become apparent such that some participants show a more sensory biased reproduction and others a more hedonic reproduction. A critical question with regard to extinction is whether the representation of the CS is updated to its initial neutral representation or whether the representation after extinction is significantly different from the before conditioning state which would indicate that the stimulus is now associated with an extinction memory (Dunsmoor et al., 2015)

6.6 Conclusion

In conclusion, the results presented here provide important advancement towards our understanding of the influence of individual differences in the norepinephrine and stress system on emotional learning processes and the underlying neural mechanisms. I demonstrated that genetically and environmentally based differences in the norepinephrine and stress system influence perceptual biases for reward. Moreover, I was able to show that acute stress affects different types of appetitive learning providing novel evidence for a role of the norepinephrine and stress system in reward processing and the subsequent development of affective biases. I further provide evidence that in classical conditioning the CS reactivates the pattern of representation elicited by the US, especially those aspects that represent the hedonic response to the US. I discuss potential mechanisms through which those novel finding can help us to better understand emotional learning processes behaviourally and how it informs how we prioritize aspects of the world around us. Last, I discuss how those individual differences in the stress and norepinephrine system create differences in vulnerability for psychopathological conditions such as PTSD, addiction and depression. Taken together, this thesis provides some of the first evidence demonstrating an effect of acute stress on appetitive learning processes as well as for the mechanistic understanding of how novel affective associations develop in the brain.

Bibliography

Abraham, A. D., Neve, K. A., & Lattal, K. M. (2014). Dopamine and extinction: A convergence of theory with fear and reward circuitry. *Neurobiology of Learning and Memory, 108*, 65-77. doi:10.1016/j.nlm.2013.11.007

Anderson, A. K., & Phelps, E. A. (2002). Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *J Cogn Neurosci, 14*(5), 709-720. doi:10.1162/08989290260138618

Anderson, B. A. (2016a). The attention habit: how reward learning shapes attentional selection. *Ann N Y Acad Sci, 1369*(1), 24-39. doi:10.1111/nyas.12957

Anderson, B. A. (2016b). What is abnormal about addiction-related attentional biases? *Drug Alcohol Depend, 167*, 8-14. doi:10.1016/j.drugalcdep.2016.08.002

Anderson, B. A., Faulkner, M. L., Rilee, J. J., Yantis, S., & Marvel, C. L. (2013). Attentional bias for nondrug reward is magnified in addiction. *Exp Clin Psychopharmacol, 21*(6), 499-506. doi:10.1037/a0034575

Anderson, B. A., Leal, S. L., Hall, M. G., Yassa, M. A., & Yantis, S. (2014). The attribution of value-based attentional priority in individuals with depressive symptoms. *Cogn Affect Behav Neurosci, 14*(4), 1221-1227. doi:10.3758/s13415-014-0301-z

Andreatta, M., & Pauli, P. (2015). Appetitive vs. Aversive conditioning in humans. *Front Behav Neurosci, 9*, 128. doi:10.3389/fnbeh.2015.00128

Apkarian, A. V., Bushnell, M. C., Treede, R. D., & Zubieta, J. K. (2005). Human brain mechanisms of pain perception and regulation in health and disease. *Eur J Pain, 9*(4), 463-484. doi:10.1016/j.ejpain.2004.11.001

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat Rev Neurosci, 10*(6), 410-422. doi:10.1038/nrn2648

Asok, A., Kandel, E. R., & Rayman, J. B. (2018). The Neurobiology of Fear Generalization. *Front Behav Neurosci, 12*, 329. doi:10.3389/fnbeh.2018.00329

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci, 28*, 403-450. doi:10.1146/annurev.neuro.28.061604.135709

Austin, A. J., & Duka, T. (2010). Mechanisms of attention for appetitive and aversive outcomes in Pavlovian conditioning. *Behav Brain Res, 213*(1), 19-26. doi:10.1016/j.bbr.2010.04.019

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: a failed theoretical dichotomy. *Trends in Cognitive Sciences, 16*(8), 437-443. doi:10.1016/j.tics.2012.06.010

Balleine, B. W. (2011). Sensation, Incentive Learning, and the Motivational Control of Goal-Directed Action. In J. A. Gottfried (Ed.), *Neurobiology of Sensation and Reward*. Boca Raton (FL).

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*(4-5), 407-419.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van, I. M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological bulletin, 133*(1), 1-24. doi:10.1037/0033-2909.133.1.1

Bechara, A., Tranel, D., & Damasio, H. (2000). Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *Brain, 123 (Pt 11)*, 2189-2202.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Arch Gen Psychiatry, 4*, 561-571.
Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Front Hum Neurosci, 7*, 133. doi:10.3389/fnhum.2013.00133
Bernstein, D. P., Fink, L., Handelsman, L., Foote, J., Lovejoy, M., Wenzel, K., . . . Ruggiero, J. (1994). Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am J Psychiatry, 151*(8), 1132-1136.
Berridge, C. W., & Waterhouse, B. D. (2003). The locus coeruleus-noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. *Brain research. Brain research reviews, 42*(1), 33-84.
Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl), 191*(3), 391-431. doi:10.1007/s00213-006-0578-x
Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res Brain Res Rev, 28*(3), 309-369.
Bogdan, R., Nikolova, Y. S., & Pizzagalli, D. A. (2013). Neurogenetics of depression: a focus on reward processing and stress sensitivity. *Neurobiol Dis, 52*, 12-23. doi:10.1016/j.nbd.2012.05.007
Bogdan, R., & Pizzagalli, D. A. (2006). Acute stress reduces reward responsiveness: implications for depression. *Biol Psychiatry, 60*(10), 1147-1154. doi:S0006-3223(06)00471-9 [pii]10.1016/j.biopsych.2006.03.037
Boubela, R. N., Kalcher, K., Huf, W., Seidel, E. M., Derntl, B., Pezawas, L., . . . Moser, E. (2015). fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions. *Sci Rep, 5*, 10499. doi:10.1038/srep10499
Bouret, S., Ravel, S., & Richmond, B. J. (2012). Complementary neural correlates of motivation in dopaminergic and noradrenergic neurons of monkeys. *Front Behav Neurosci, 6*, 40. doi:10.3389/fnbeh.2012.00040
Bouret, S., & Richmond, B. J. (2009). Relation of locus coeruleus neurons in monkeys to Pavlovian and operant behaviors. *J Neurophysiol, 101*(2), 898-911. doi:10.1152/jn.91048.2008
Bouret, S., & Richmond, B. J. (2015). Sensitivity of locus ceruleus neurons to reward value for goal-directed actions. *J Neurosci, 35*(9), 4005-4014. doi:10.1523/JNEUROSCI.4553-14.2015
Brailean, A. M., Koster, E. H., Hoorelbeke, K., & De Raedt, R. (2014). Attentional modulation by reward and punishment cues in relation to depressive symptoms. *J Behav Ther Exp Psychiatry, 45*(3), 351-359. doi:10.1016/j.jbtep.2014.03.003
Bremner, J. D., Innis, R. B., Ng, C. K., Staib, L. H., Salomon, R. M., Bronen, R. A., . . . Charney, D. S. (1997). Positron emission tomography measurement of cerebral metabolic correlates of yohimbine administration in combat-related posttraumatic stress disorder. *Arch Gen Psychiatry, 54*(3), 246-254.
Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *J Exp Anal Behav, 11*(1), 1-8. doi:10.1901/jeab.1968.11-1
Buchanan, T. W., Tranel, D., & Adolphs, R. (2006). Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learn Mem, 13*(3), 382-387. doi:10.1101/lm.206306

Buchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Curr Opin Neurobiol, 10*(2), 219-223.
Bulganin, L., Bach, D. R., & Wittmann, B. C. (2014). Prior fear conditioning and reward learning interact in fear and reward networks. *Front Behav Neurosci, 8*, 67. doi:10.3389/fnbeh.2014.00067
Bush, D. E., Caparosa, E. M., Gekker, A., & Ledoux, J. (2010). Beta-adrenergic receptors in the lateral nucleus of the amygdala contribute to the acquisition but not the consolidation of auditory fear conditioning. *Front Behav Neurosci, 4*, 154. doi:10.3389/fnbeh.2010.00154
Byrne, J. H. (2003). *Learning & memory* (Vol. 2). New York, USA: Macmillan Reference.
Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learn Mem, 10*(4), 270-274. doi:10.1101/lm.62403
Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends Neurosci, 21*(7), 294-299. doi:S0166-2236(97)01214-9 [pii]
Cardinal, R. N., Parkinson, J. A., Marbini, H. D., Toner, A. J., Bussey, T. J., Robbins, T. W., & Everitt, B. J. (2003). Role of the anterior cingulate cortex in the control over behavior by Pavlovian conditioned stimuli in rats. *Behav Neurosci, 117*(3), 566-587.
Cavanagh, J. F., Frank, M. J., & Allen, J. J. (2011). Social stress reactivity alters reward and punishment learning. *Soc Cogn Affect Neurosci, 6*(3), 311-320. doi:10.1093/scan/nsq041
Chelazzi, L., Estocinova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., & Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. *J Neurosci, 34*(25), 8594-8604. doi:10.1523/JNEUROSCI.0277-14.2014
Chiew, K. S., & Braver, T. S. (2011). Positive affect versus reward: emotional and motivational influences on cognitive control. *Front Psychol, 2*, 279. doi:10.3389/fpsyg.2011.00279
Chikazoe, J., Lee, D. H., Kriegeskorte, N., & Anderson, A. K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nat Neurosci, 17*(8), 1114-1122. doi:10.1038/nn.3749
Clark, J. J., Hollon, N. G., & Phillips, P. E. (2012). Pavlovian valuation systems in learning and decision making. *Curr Opin Neurobiol, 22*(6), 1054-1061. doi:10.1016/j.conb.2012.06.004
Cousijn, H., Rijpkema, M., Qin, S., van Marle, H. J., Franke, B., Hermans, E. J., . . . Fernandez, G. (2010). Acute stress modulates genotype effects on amygdala processing in humans. *Proc Natl Acad Sci U S A, 107*(21), 9867-9872. doi:10.1073/pnas.1003514107
Cousineau, D., & Lacroix, G. L. (2006). Getting parameters from learning data. *Tutorials in Quantitative Methods for Psychology, 2*(2), 77-83.
Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res, 29*(3), 162-173.
Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci, 3*(8), 655-666. doi:10.1038/nrn894
Cunningham, W. A., & Brosch, T. (2012). Motivational Salience: Amygdala Tuning From Traits, Needs, Values, and Goals. *Current Directions in Psychological Science, 21*(1), 54-59. doi:10.1177/0963721411430832
Darvas, M., Fadok, J. P., & Palmiter, R. D. (2011). Requirement of dopamine signaling in the amygdala and striatum for learning and maintenance of a conditioned avoidance response. *Learning & Memory, 18*(3), 136-143. doi:10.1101/lm.2041211

Davidson, J. (1992). Drug therapy of post-traumatic stress disorder. *Br J Psychiatry, 160*, 309-314.
de Kloet, E. R., Joels, M., & Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nat Rev Neurosci, 6*(6), 463-475. doi:10.1038/nrn1683
de Kloet, E. R., Oitzl, M. S., & Joels, M. (1999). Stress and cognition: are corticosteroids good or bad guys? *Trends Neurosci, 22*(10), 422-426. doi:S0166223699014381 [pii]
de la Fuente-Fernandez, R., Phillips, A. G., Zamburlini, M., Sossi, V., Calne, D. B., Ruth, T. J., & Stoessl, A. J. (2002). Dopamine release in human ventral striatum and expectation of reward. *Behavioural brain research, 136*(2), 359-363.
de Quervain, D. J., Aerni, A., Schelling, G., & Roozendaal, B. (2009). Glucocorticoids and the regulation of memory in health and disease. *Front Neuroendocrinol, 30*(3), 358-370. doi:10.1016/j.yfrne.2009.03.002
de Quervain, D. J., Kolassa, I. T., Ertl, V., Onyut, P. L., Neuner, F., Elbert, T., & Papassotiropoulos, A. (2007). A deletion variant of the alpha2b-adrenoceptor is related to emotional memory in Europeans and Africans. *Nat Neurosci, 10*(9), 1137-1139. doi:nn1945 [pii]10.1038/nn1945
de Quervain, D. J., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394*(6695), 787-790. doi:10.1038/29542
de Quervain, D. J. F., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394*(6695), 787-790.
Delgado, M. R., Jou, R. L., & Phelps, E. A. (2011). Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Front Neurosci, 5*, 71. doi:10.3389/fnins.2011.00071
Delgado, M. R., Olsson, A., & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biol Psychol, 73*(1), 39-48. doi:10.1016/j.biopsycho.2006.01.006
Derryberry, D., & Reed, M. A. (1994). Temperament and attention: orienting toward and away from positive and negative signals. *Journal of Personality and Social Psychology, 66*(6), 1128-1139.
Diamond, D. M. (2005). Cognitive, endocrine and mechanistic perspectives on non-linear relationships between arousal and brain function. *Nonlinearity Biol Toxicol Med, 3*(1), 1-7. doi:10.2201/nonlin.003.01.001
Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., . . . Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science, 325*(5940), 621-625. doi:10.1126/science.1171203325/5940/621 [pii]
Domes, G., Heinrichs, M., Reichwald, U., & Hautzinger, M. (2002). Hypothalamic-pituitary-adrenal axis reactivity to psychological stress and memory in middle-aged women: high responders exhibit enhanced declarative memory performance. *Psychoneuroendocrinology, 27*(7), 843-853.
Doran, J. E., & Michie, D. (1966). Experiments with the Graph Traverser Program. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 294*(1437), 235-259.
Drouin, C., Darracq, L., Trovero, F., Blanc, G., Glowinski, J., Cotecchia, S., & Tassin, J. P. (2002). Alpha1b-adrenergic receptors control locomotor and rewarding effects of psychostimulants and opiates. *J Neurosci, 22*(7), 2873-2884. doi:20026237

Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., . . . Baas, J. M. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depress Anxiety, 32*(4), 239-253. doi:10.1002/da.22353

Dunsmoor, J. E., Niv, Y., Daw, N., & Phelps, E. A. (2015). Rethinking Extinction. *Neuron, 88*(1), 47-63. doi:10.1016/j.neuron.2015.09.028

Duvarci, S., Popa, D., & Pare, D. (2011). Central amygdala activity during fear conditioning. *J Neurosci, 31*(1), 289-294. doi:10.1523/JNEUROSCI.4985-10.2011

Ehlers, M. R., Ross, C. J. D., & Todd, R. M. (2018). The influence of the noradrenergic/stress system on perceptual biases for reward. *Cogn Affect Behav Neurosci*. doi:10.3758/s13415-018-00657-0

Ehlers, M. R., & Todd, R. M. (2017a). Acute psychophysiological stress impairs human associative learning. *Neurobiol Learn Mem, 145*, 84-93. doi:10.1016/j.nlm.2017.09.003

Ehlers, M. R., & Todd, R. M. (2017b). Genesis and Maintenance of Attentional Biases: The Role of the Locus Coeruleus-Noradrenaline System. *Neural Plast, 2017*, 6817349. doi:10.1155/2017/6817349

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion, 6*(3-4), 169-200. doi:Doi 10.1080/02699939208411068

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nat Neurosci, 16*(8), 1146-1153. doi:10.1038/nn.3428

Elliot, A. J., Eder, A. B., & Harmon-Jones, E. (2013). Approach-Avoidance Motivation and Emotion: Convergence and Divergence. *Emotion Review, 5*(3), 308-311. doi:10.1177/1754073913477517

Elliott, K. C., Cheruvelil, K. S., Montgomery, G. M., & Soranno, P. A. (2016). Conceptions of Good Science in Our Data-Rich World. *Bioscience, 66*(10), 880-889. doi:10.1093/biosci/biw115

Elzinga, B. M., Bakker, A., & Bremner, J. D. (2005). Stress-induced cortisol elevations are associated with impaired delayed, but not immediate recall. *Psychiatry Res, 134*(3), 211-223. doi:10.1016/j.psychres.2004.11.007

Enkel, T., Gholizadeh, D., von Bohlen Und Halbach, O., Sanchis-Segura, C., Hurlemann, R., Spanagel, R., . . . Vollmayr, B. (2010). Ambiguous-cue interpretation is biased under stress- and depression-like states in rats. *Neuropsychopharmacology, 35*(4), 1008-1015. doi:10.1038/npp.2009.204

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci, 15*(2), 85-93. doi:10.1016/j.tics.2010.11.004

Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry, 164*(10), 1476-1488. doi:10.1176/appi.ajp.2007.07030504

Everitt, B. J., Cardinal, R. N., Parkinson, J. A., & Robbins, T. W. (2003). Appetitive behavior: impact of amygdala-dependent mechanisms of emotional learning. *Ann N Y Acad Sci, 985*, 233-250.

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci, 8*(11), 1481-1489. doi:10.1038/nn1579

Everitt, B. J., & Robbins, T. W. (2016). Drug Addiction: Updating Actions to Habits to Compulsions Ten Years On. *Annu Rev Psychol, 67*, 23-50. doi:10.1146/annurev-psych-122414-033457

Eysenck, H. J., & Eysenck, M. W. (1985). *Personality and Individual Differences*. New York: Plenum Press.
Fadok, J. P., Dickerson, T. M. K., & Palmiter, R. D. (2009). Dopamine Is Necessary for Cue-Dependent Fear Conditioning. *Journal of Neuroscience, 29*(36), 11089-11097. doi:10.1523/Jneurosci.1616-09.2009
Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychon Bull Rev, 1*(4), 429-438. doi:10.3758/BF03210947
Fanselow, M. S., & Poulos, A. M. (2005). The neuroscience of mammalian associative learning. *Annu Rev Psychol, 56*, 207-234. doi:10.1146/annurev.psych.56.091103.070213
Field, M., & Cox, W. M. (2008). Attentional bias in addictive behaviors: a review of its development, causes, and consequences. *Drug Alcohol Depend, 97*(1-2), 1-20. doi:10.1016/j.drugalcdep.2008.03.030
Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology, 56 Suppl 1*, 139-148. doi:10.1016/j.neuropharm.2008.06.027
Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., . . . Akil, H. (2011). A selective role for dopamine in stimulus-reward learning. *Nature, 469*(7328), 53-57. doi:10.1038/nature09588
Flagel, S. B., Watson, S. J., Akil, H., & Robinson, T. E. (2008). Individual differences in the attribution of incentive salience to a reward-related cue: influence on cocaine sensitization. *Behavioural brain research, 186*(1), 48-56. doi:10.1016/j.bbr.2007.07.022
Fridman, A., van IJzendoorn, M. H., Sagi-Schwartz, A., & Bakermans-Kranenburg, M. J. (2012). Genetic moderation of cortisol secretion in Holocaust survivors: A pilot study on the role of ADRA2B. *International Journal of Behavioral Development, 36*(1), 79-84. doi:10.1177/0165025411406859
Furl, N., Henson, R. N., Friston, K. J., & Calder, A. J. (2013). Top-Down Control of Visual Responses to Fear by the Amygdala. *Journal of Neuroscience, 33*(44), 17435-17443. doi:10.1523/Jneurosci.2992-13.2013
Furmark, T., Fischer, H., Wik, G., Larsson, M., & Fredrikson, M. (1997). The amygdala and individual differences in human fear conditioning. *Neuroreport, 8*(18), 3957-3960.
Garofalo, S., & di Pellegrino, G. (2015). Individual differences in the influence of task-irrelevant Pavlovian cues on human behavior. *Front Behav Neurosci, 9*, 163. doi:10.3389/fnbeh.2015.00163
Geracioti, T. D., Jr., Baker, D. G., Ekhator, N. N., West, S. A., Hill, K. K., Bruce, A. B., . . . Kasckow, J. W. (2001). CSF norepinephrine concentrations in posttraumatic stress disorder. *Am J Psychiatry, 158*(8), 1227-1230. doi:10.1176/appi.ajp.158.8.1227
Giesecke, T., Gracely, R. H., Grant, M. A., Nachemson, A., Petzke, F., Williams, D. A., & Clauw, D. J. (2004). Evidence of augmented central pain processing in idiopathic chronic low back pain. *Arthritis Rheum, 50*(2), 613-623. doi:10.1002/art.20063
Giustino, T. F., & Maren, S. (2018). Noradrenergic Modulation of Fear Conditioning and Extinction. *Front Behav Neurosci, 12*, 43. doi:10.3389/fnbeh.2018.00043
Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition & Emotion, 22*(6), 1094-1118. doi:10.1080/02699930701626582

Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J Neurosci, 22*(24), 10829-10837.
Graham, L. K., Yoon, T., & Kim, J. J. (2010). Stress impairs optimal behavior in a water foraging choice task in rats. *Learn Mem, 17*(1), 1-4. doi:10.1101/lm.160551017/1/790 [pii]
Greco, J. A., & Liberzon, I. (2016). Neuroimaging of Fear-Associated Learning. *Neuropsychopharmacology, 41*(1), 320-334. doi:10.1038/npp.2015.255
Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci, 7*(5), 555-562. doi:10.1038/nn1224
Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior insular cortex and emotional awareness. *J Comp Neurol, 521*(15), 3371-3388. doi:10.1002/cne.23368
Haber, S. N. (2011). Neuroanatomy of Reward: A View from the Ventral Striatum. In J. A. Gottfried (Ed.), *Neurobiology of Sensation and Reward*. Boca Raton (FL).
Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics, 7*(1), 37-53. doi:10.1007/s12021-008-9041-y
Hasler, G., Drevets, W. C., Manji, H. K., & Charney, D. S. (2004). Discovering endophenotypes for major depression. *Neuropsychopharmacology, 29*(10), 1765-1781. doi:10.1038/sj.npp.1300506
Hearst, E., & Jenkins, H. (1974). Sign-tracking: The Stimulus-reinforcer Relation and Directed Action. *Psychonomic Society*.
Hebb, D. O., Martinez, J. L., & Glickman, S. E. (1994). The Organization of Behavior - a Neuropsychological Theory - Hebb,Do. *Contemporary Psychology, 39*(11), 1018-1020.
Herman, J. P., McKlveen, J. M., Solomon, M. B., Carvalho-Netto, E., & Myers, B. (2012). Neural regulation of the stress response: glucocorticoid feedback mechanisms. *Braz J Med Biol Res, 45*(4), 292-298.
Holland, P. C. (1990). Event representation in Pavlovian conditioning: image and action. *Cognition, 37*(1-2), 105-131.
Hollon, N. G., Burgeno, L. M., & Phillips, P. E. (2015). Stress effects on the neural substrates of motivated behavior. *Nat Neurosci, 18*(10), 1405-1412. doi:10.1038/nn.4114
Hull, C. L. (1943). *Principles of behavior: An introduction to behavior theory.* Oxford, England: Appleton-Century.
Huys, Q. J., Pizzagalli, D. A., Bogdan, R., & Dayan, P. (2013). Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol Mood Anxiety Disord, 3*(1), 12. doi:10.1186/2045-5380-3-12
Jackson, S. A., Horst, N. K., Pears, A., Robbins, T. W., & Roberts, A. C. (2016). Role of the Perigenual Anterior Cingulate and Orbitofrontal Cortex in Contingency Learning in the Marmoset. *Cereb Cortex, 26*(7), 3273-3284. doi:10.1093/cercor/bhw067
Jensen, K. B., Regenbogen, C., Ohse, M. C., Frasnelli, J., Freiherr, J., & Lundstrom, J. N. (2016). Brain activations during pain: a neuroimaging meta-analysis of patients with pain and healthy controls. *Pain, 157*(6), 1279-1286. doi:10.1097/j.pain.0000000000000517
Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J Cogn Neurosci, 23*(7), 1587-1596. doi:10.1162/jocn.2010.21548

Jiang, Y. V., Swallow, K. M., & Rosenbaum, G. M. (2013). Guidance of spatial attention by incidental learning and endogenous cuing. *J Exp Psychol Hum Percept Perform, 39*(1), 285-297. doi:10.1037/a0028022

Joels, M., Pu, Z., Wiegert, O., Oitzl, M. S., & Krugers, H. J. (2006). Learning under stress: how does it work? *Trends Cogn Sci, 10*(4), 152-158. doi:10.1016/j.tics.2006.02.002

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. doi:10.1016/j.neuron.2015.11.028

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci, 17*(11), 4302-4311.

Kasch, K. L., Rottenberg, J., Arnow, B. A., & Gotlib, I. H. (2002). Behavioral activation and inhibition systems and the severity and course of depression. *J Abnorm Psychol, 111*(4), 589-597.

Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: are emotional words more vividly remembered than neutral words? *Memory & Cognition, 31*(8), 1169-1180.

Kensinger, E. A., & Corkin, S. (2004). Two routes to emotional memory: Distinct neural processes for valence and arousal. *Proceedings of the National Academy of Sciences of the United States of America, 101*(9), 3310-3315. doi:DOI 10.1073/pnas.0306408101

Kim, J. J., & Jung, M. W. (2006). Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci Biobehav Rev, 30*(2), 188-202. doi:10.1016/j.neubiorev.2005.06.005

Klein, D. C., Fencil-Morse, E., & Seligman, M. E. (1976). Learned helplessness, depression, and the attribution of failure. *J Pers Soc Psychol, 33*(5), 508-516.

Klucken, T., Schweckendiek, J., Merz, C. J., Tabbert, K., Walter, B., Kagerer, S., . . . Stark, R. (2009). Neural activations of the acquisition of conditioned sexual arousal: effects of contingency awareness and sex. *J Sex Med, 6*(11), 3071-3085. doi:10.1111/j.1743-6109.2009.01405.x

Knutson, B., & Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philos Trans R Soc Lond B Biol Sci, 363*(1511), 3771-3786. doi:10.1098/rstb.2008.0155

Koch, K., McLean, J., Segev, R., Freed, M. A., Berry, M. J., 2nd, Balasubramanian, V., & Sterling, P. (2006). How much the eye tells the brain. *Curr Biol, 16*(14), 1428-1434. doi:10.1016/j.cub.2006.05.056

Kragel, P. A., Kano, M., Van Oudenhove, L., Ly, H. G., Dupont, P., Rubio, A., . . . Wager, T. D. (2018). Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nat Neurosci, 21*(2), 283-289. doi:10.1038/s41593-017-0051-7

Kragel, P. A., Koban, L., Barrett, L. F., & Wager, T. D. (2018). Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging. *Neuron, 99*(2), 257-273. doi:10.1016/j.neuron.2018.06.009

Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn Sci, 17*(1), 26-49. doi:10.1016/j.tics.2012.10.011

Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: integrating cognition, computation, and the brain. *Trends Cogn Sci, 17*(8), 401-412. doi:10.1016/j.tics.2013.06.007

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. *Front Syst Neurosci, 2*, 4. doi:10.3389/neuro.06.004.2008

Kudielka, B. M., Hellhammer, D. H., & Kirschbaum, C. (2007). Ten Years of Research with the Trier Social Stress Test--Revisited. In E. Harmon-Jones & P. Winkielmann (Eds.), *Social neuroscience: Integrating biological and psychological explanations of social behavior* (pp. 56-83). New York, NY, US: Guilford Press.

Kumar, P., Waiter, G., Ahearn, T., Milders, M., Reid, I., & Steele, J. D. (2008). Abnormal temporal difference reward-learning signals in major depression. *Brain, 131*(Pt 8), 2084-2093. doi:10.1093/brain/awn136

Kundu, P., Brenowitz, N. D., Voon, V., Worbe, Y., Vertes, P. E., Inati, S. J., . . . Bullmore, E. T. (2013). Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proc Natl Acad Sci U S A, 110*(40), 16187-16192. doi:10.1073/pnas.1301725110

Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage, 60*(3), 1759-1770. doi:10.1016/j.neuroimage.2011.12.028

Kurth, F., Zilles, K., Fox, P. T., Laird, A. R., & Eickhoff, S. B. (2010). A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Structure & Function, 214*(5-6), 519-534. doi:10.1007/s00429-010-0255-z

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nat Rev Neurosci, 7*(1), 54-64. doi:nrn1825 [pii]10.1038/nrn1825

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron, 20*(5), 937-945.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *J Neurosci, 15*(10), 6846-6855.

LaLumiere, R. T., Buen, T. V., & McGaugh, J. L. (2003). Post-training intra-basolateral amygdala infusions of norepinephrine enhance consolidation of memory for contextual fear conditioning. *J Neurosci, 23*(17), 6754-6758.

Lang, P. J., & Bradley, M. M. (2008). Appetitive and defensive motivation is the substrate of emotion. In A. J. Elliot (Ed.), *Handbook of approach and avoidance motivation* (pp. 51-65). New York, USA: US: Psychology Press.

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cell Mol Neurobiol, 23*(4-5), 727-738.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu Rev Neurosci, 23*, 155-184. doi:10.1146/annurev.neuro.23.1.155

LeDoux, J. E., Cicchetti, P., Xagoraris, A., & Romanski, L. M. (1990). The lateral amygdaloid nucleus: sensory interface of the amygdala in fear conditioning. *J Neurosci, 10*(4), 1062-1069.

Lee, D., Todd, R. M., Gardhouse, K., Levine, B., & Anderson, A. K. (2013). *Enhanced attentional capture in survivors of a single traumatic event*. Paper presented at the Society for Neuroscience, San Diego, CA.

Levy, D. J., & Glimcher, P. W. (2011). Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J Neurosci, 31*(41), 14693-14707. doi:10.1523/JNEUROSCI.2218-11.2011

Li, S., Weerda, R., Guenzel, F., Wolf, O. T., & Thiel, C. M. (2013). ADRA2B genotype modulates effects of acute psychosocial stress on emotional memory retrieval in healthy young men. *Neurobiol Learn Mem, 103*, 11-18. doi:10.1016/j.nlm.2013.03.006

Liberzon, I., Taylor, S. F., Amdur, R., Jung, T. D., Chamberlain, K. R., Minoshima, S., . . . Fig, L. M. (1999). Brain activation in PTSD in response to trauma-related stimuli. *Biol Psychiatry, 45*(7), 817-826.

Lim, S. L., Padmala, S., & Pessoa, L. (2008). Affective learning modulates spatial competition during low-load attentional conditions. *Neuropsychologia, 46*(5), 1267-1278. doi:10.1016/j.neuropsychologia.2007.12.003S0028-3932(07)00433-2 [pii]

Lim, S. L., Padmala, S., & Pessoa, L. (2009). Segregating the significant from the mundane on a moment-to-moment basis via direct and indirect amygdala contributions. *Proc Natl Acad Sci U S A, 106*(39), 16841-16846. doi:10.1073/pnas.0904551106

Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *Am J Psychiatry, 165*(1), 124-132. doi:10.1176/appi.ajp.2007.06091513

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behav Res Ther, 43*(11), 1391-1424. doi:10.1016/j.brat.2004.10.007

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. *Behav Res Ther, 47*(2), 111-118. doi:10.1016/j.brat.2008.10.017

Liu, J., Wei, W., Kuang, H., Zhao, F., & Tsien, J. Z. (2013). Changes in heart rate variability are associated with expression of short-term and long-term contextual and cued fear memories. *PLoS One, 8*(5), e63590. doi:10.1371/journal.pone.0063590

Loken, L. S., Wessberg, J., Morrison, I., McGlone, F., & Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nat Neurosci, 12*(5), 547-548. doi:10.1038/nn.2312

Lopez-Sola, M., Pujol, J., Hernandez-Ribas, R., Harrison, B. J., Ortiz, H., Soriano-Mas, C., . . . Cardoner, N. (2010). Dynamic assessment of the right lateral frontal cortex response to painful stimulation. *Neuroimage, 50*(3), 1177-1187. doi:10.1016/j.neuroimage.2010.01.031

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain Cogn, 65*(3), 209-237. doi:10.1016/j.bandc.2007.02.007

Lupien, S. J., & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: integration of animal and human model studies. *Brain Res Brain Res Rev, 24*(1), 1-27. doi:S0165017397000040 [pii]

Lupien, S. J., Wilkinson, C. W., Briere, S., Menard, C., Ng Ying Kin, N. M., & Nair, N. P. (2002). The modulatory effects of corticosteroids on cognition: studies in young human populations. *Psychoneuroendocrinology, 27*(3), 401-416.

Mackey, S., & Petrides, M. (2014). Architecture and morphology of the human ventromedial prefrontal cortex. *Eur J Neurosci, 40*(5), 2777-2796. doi:10.1111/ejn.12654

Mahan, A. L., & Ressler, K. J. (2012). Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. *Trends Neurosci, 35*(1), 24-35. doi:10.1016/j.tins.2011.06.007

Makaritsis, K. P., Johns, C., Gavras, I., & Gavras, H. (2000). Role of alpha(2)-adrenergic receptor subtypes in the acute hypertensive response to hypertonic saline infusion in anephric mice. *Hypertension, 35*(2), 609-613.

Mammarella, N., Fairfield, B., Di Domenico, A., D'Onofrio, L., Stuppia, L., & Gatta, V. (2016). The modulating role of ADRA2B in emotional working memory: Attending the negative but remembering the positive. *Neurobiol Learn Mem, 130*, 129-134. doi:10.1016/j.nlm.2016.02.009

Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annu Rev Neurosci, 24*, 897-931. doi:10.1146/annurev.neuro.24.1.897

Maren, S., & Quirk, G. J. (2004). Neuronal signalling of fear memory. *Nat Rev Neurosci, 5*(11), 844-852. doi:10.1038/nrn1535

Markovic, J., Anderson, A. K., & Todd, R. M. (2014). Tuning to the significant: Neural and genetic processes underlying affective enhancement of visual perception and memory. *Behavioural brain research, 259*, 229-241. doi:10.1016/j.bbr.2013.11.018

Marshall, A. J., Acheson, D. T., Risbrough, V. B., Straus, L. D., & Drummond, S. P. (2014). Fear conditioning, safety learning, and sleep in humans. *J Neurosci, 34*(35), 11754-11760. doi:10.1523/JNEUROSCI.0478-14.2014

Martin-Soelch, C., Linthicum, J., & Ernst, M. (2007). Appetitive conditioning: neural bases and implications for psychopathology. *Neurosci Biobehav Rev, 31*(3), 426-440. doi:10.1016/j.neubiorev.2006.11.002

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016a). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behav Brain Sci, 39*, e200. doi:10.1017/S0140525X15000667

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016b). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences, 39*. doi:UNSP e20010.1017/S0140525X15000667

McGaugh, J. L. (1966). Time-dependent processes in memory storage. *Science, 153*(3742), 1351-1358.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annu Rev Neurosci, 27*, 1-28. doi:10.1146/annurev.neuro.27.070203.144157

McGaugh, J. L. (2013). Making lasting memories: Remembering the significant. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 10402-10407. doi:10.1073/pnas.1301209110

McGlone, F., & Reilly, D. (2010). The cutaneous sensory system. *Neurosci Biobehav Rev, 34*(2), 148-159. doi:10.1016/j.neubiorev.2009.08.004

McTeague, L. M., Gruss, L. F., & Keil, A. (2015). Aversive learning shapes neuronal orientation tuning in human visual cortex. *Nat Commun, 6*, 7823. doi:10.1038/ncomms8823

Merckelbach, H., van Hout, W., de Jong, P., & van den Hout, M. A. (1990). Classical conditioning and attentional bias. *J Behav Ther Exp Psychiatry, 21*(3), 185-191.

Miller, R., Plessow, F., Kirschbaum, C., & Stalder, T. (2013). Classification criteria for distinguishing cortisol responders from nonresponders to psychosocial stress: evaluation

of salivary cortisol pulse detection in panel designs. *Psychosom Med, 75*(9), 832-840. doi:10.1097/PSY.0000000000000002
Miskovic, V., & Keil, A. (2013). Perceiving threat in the face of safety: excitation and inhibition of conditioned fear in human visual cortex. *J Neurosci, 33*(1), 72-78. doi:10.1523/JNEUROSCI.3692-12.2013
Moriceau, S., & Sullivan, R. M. (2004). Unique neural circuitry for neonatal olfactory learning. *J Neurosci, 24*(5), 1182-1189. doi:10.1523/JNEUROSCI.4578-03.200424/5/1182 [pii]
Morilak, D. A., Barrera, G., Echevarria, D. J., Garcia, A. S., Hernandez, A., Ma, S., & Petre, C. O. (2005). Role of brain norepinephrine in the behavioral response to stress. *Prog Neuropsychopharmacol Biol Psychiatry, 29*(8), 1214-1224. doi:10.1016/j.pnpbp.2005.08.007
Morrison, S. E., Bamkole, M. A., & Nicola, S. M. (2015). Sign Tracking, but Not Goal Tracking, is Resistant to Outcome Devaluation. *Front Neurosci, 9*, 468. doi:10.3389/fnins.2015.00468
Motzkin, J. C., Philippi, C. L., Wolf, R. C., Baskaya, M. K., & Koenigs, M. (2015). Ventromedial prefrontal cortex is critical for the regulation of amygdala activity in humans. *Biol Psychiatry, 77*(3), 276-284. doi:10.1016/j.biopsych.2014.02.014
Must, A., Szabo, Z., Bodi, N., Szasz, A., Janka, Z., & Keri, S. (2006). Sensitivity to reward and punishment and the prefrontal cortex in major depression. *J Affect Disord, 90*(2-3), 209-215. doi:10.1016/j.jad.2005.12.005
Muszkat, M., Kurnik, D., Solus, J., Sofowora, G. G., Xie, H. G., Jiang, L., . . . Stein, C. M. (2005). Variation in the alpha2B-adrenergic receptor gene (ADRA2B) and its relationship to vascular response in vivo. *Pharmacogenet Genomics, 15*(6), 407-414.
Namkung, H., Kim, S. H., & Sawa, A. (2018). The Insula: An Underestimated Brain Area in Clinical Neuroscience, Psychiatry, and Neurology: (Trends in Neuroscience 40, 200-207, 2017). *Trends Neurosci, 41*(8), 551-554. doi:10.1016/j.tins.2018.05.004
Nasser, H. M., & McNally, G. P. (2013). Neural correlates of appetitive-aversive interactions in Pavlovian fear conditioning. *Learn Mem, 20*(4), 220-228. doi:10.1101/lm.029744.112
Nieuwenhuys, R. (2012). The insular cortex: a review. *Prog Brain Res, 195*, 123-163. doi:10.1016/B978-0-444-53860-4.00007-6
Nikolova, Y. S., Ferrell, R. E., Manuck, S. B., & Hariri, A. R. (2011). Multilocus genetic profile for dopamine signaling predicts ventral striatum reactivity. *Neuropsychopharmacology, 36*(9), 1940-1947. doi:10.1038/npp.2011.82
Nutt, D. J., Lingford-Hughes, A., Erritzoe, D., & Stokes, P. R. (2015). The dopamine theory of addiction: 40 years of highs and lows. *Nat Rev Neurosci, 16*(5), 305-312. doi:10.1038/nrn3939
O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol, 14*(6), 769-776. doi:10.1016/j.conb.2004.10.016
Olson, V. G., Rockett, H. R., Reh, R. K., Redila, V. A., Tran, P. M., Venkov, H. A., . . . Raskind, M. A. (2011). The role of norepinephrine in differential response to stress in an animal model of posttraumatic stress disorder. *Biol Psychiatry, 70*(5), 441-448. doi:10.1016/j.biopsych.2010.11.029
Onat, S., & Buchel, C. (2015). The neuronal basis of fear generalization in humans. *Nat Neurosci, 18*(12), 1811-1818. doi:10.1038/nn.4166

Onur, O. A., Walter, H., Schlaepfer, T. E., Rehme, A. K., Schmidt, C., Keysers, C., . . . Hurlemann, R. (2009). Noradrenergic enhancement of amygdala responses to fear. *Soc Cogn Affect Neurosci, 4*(2), 119-126. doi:10.1093/scan/nsn049

Orenius, T. I., Raij, T. T., Nuortimo, A., Naatanen, P., Lipsanen, J., & Karlsson, H. (2017). The interaction of emotion and pain in the insula and secondary somatosensory cortex. *Neuroscience, 349*, 185-194. doi:10.1016/j.neuroscience.2017.02.047

Pavlov, P. I. (2010). Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Ann Neurosci, 17*(3), 136-141. doi:10.5214/ans.0972-7531.1017309

Pecina, S., Schulkin, J., & Berridge, K. C. (2006). Nucleus accumbens corticotropin-releasing factor increases cue-triggered motivation for sucrose reward: paradoxical positive incentive effects in stress? *BMC Biol, 4*, 8. doi:1741-7007-4-8 [pii]10.1186/1741-7007-4-8

Peck, C. J., & Salzman, C. D. (2014a). The amygdala and basal forebrain as a pathway for motivationally guided attention. *J Neurosci, 34*(41), 13757-13767. doi:10.1523/JNEUROSCI.2106-14.2014

Peck, C. J., & Salzman, C. D. (2014b). Amygdala neural activity reflects spatial attention towards stimuli promising reward or threatening punishment. *Elife, 3*. doi:10.7554/eLife.04478

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and anxiety, 27*(12), 1135-1142. doi:10.1002/da.20755

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, M., McDonald, S., & Munafo, M. R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychol Sci, 24*(5), 688-697. doi:10.1177/0956797612459657095679761245965709567976124596 57 [pii]

Peper, M., Karcher, S., Wohlfarth, R., Reinshagen, G., & LeDoux, J. E. (2001). Aversive learning in patients with unilateral lesions of the amygdala and hippocampus. *Biol Psychol, 58*(1), 1-23.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042-1045. doi:10.1038/nature05051

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences, 13*(4), 160-166. doi:10.1016/j.tics.2009.01.006

Petrakis, I. L., Ralevski, E., Desai, N., Trevisan, L., Gueorguieva, R., Rounsaville, B., & Krystal, J. H. (2012). Noradrenergic vs serotonergic antidepressant with or without naltrexone for veterans with PTSD and comorbid alcohol dependence. *Neuropsychopharmacology, 37*(4), 996-1004. doi:10.1038/npp.2011.283

Piazza, P. V., & Le Moal, M. (1998). The role of stress in drug self-administration. *Trends Pharmacol Sci, 19*(2), 67-74.

Picardo, R., Baron, A. S., Anderson, A. K., & Todd, R. M. (2016). Tuning to the Positive: Age-Related Differences in Subjective Perception of Facial Emotion. *PLoS One, 11*(1), e0145643. doi:10.1371/journal.pone.0145643

Pielock, S. M., Braun, S., & Hauber, W. (2013). The effects of acute stress on Pavlovian-instrumental transfer in rats. *Cogn Affect Behav Neurosci, 13*(1), 174-185. doi:10.3758/s13415-012-0129-3

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology, 46*(5), 984-995. doi:10.1111/j.1469-8986.2009.00852.x

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: evidence from a probabilistic reward task. *J Psychiatr Res, 43*(1), 76-87. doi:10.1016/j.jpsychires.2008.03.001

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2015). Stress increases cue-triggered "wanting" for sweet reward in humans. *J Exp Psychol Anim Learn Cogn, 41*(2), 128-136. doi:10.1037/xan00000522014-56010-001 [pii]

Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M. L., . . . Kiselev, V. G. (1999). Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magn Reson Med, 42*(1), 87-97.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2013). Brain mechanisms for emotional influences on perception and attention: what is magic and what is not. *Biological Psychology, 92*(3), 492-512. doi:10.1016/j.biopsycho.2012.02.007

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav Res Methods Instrum Comput, 36*(4), 717-731.

Preusser, S., Thiel, S. D., Rook, C., Roggenhofer, E., Kosatschek, A., Draganski, B., . . . Pleger, B. (2015). The perception of touch and the ventral somatosensory pathway. *Brain, 138*(Pt 3), 540-548. doi:10.1093/brain/awu370

Prevost, C., Liljeholm, M., Tyszka, J. M., & O'Doherty, J. P. (2012). Neural correlates of specific and general Pavlovian-to-Instrumental Transfer within human amygdalar subregions: a high-resolution fMRI study. *J Neurosci, 32*(24), 8383-8390. doi:10.1523/JNEUROSCI.6237-11.2012

Pugh, C. R., Fleshner, M., & Rudy, J. W. (1997). Type II glucocorticoid receptor antagonists impair contextual but not auditory-cue fear conditioning in juvenile rats. *Neurobiol Learn Mem, 67*(1), 75-79. doi:10.1006/nlme.1996.3741

Rasch, B., Spalek, K., Buholzer, S., Luechinger, R., Boesiger, P., Papassotiropoulos, A., & de Quervain, D. J. (2009). A genetic variation of the noradrenergic system is related to differential amygdala activation during encoding of emotional memories. *Proc Natl Acad Sci U S A, 106*(45), 19191-19196. doi:10.1073/pnas.0907425106 [pii]

Rau, V., DeCola, J. P., & Fanselow, M. S. (2005). Stress-induced enhancement of fear learning: an animal model of posttraumatic stress disorder. *Neurosci Biobehav Rev, 29*(8), 1207-1223. doi:10.1016/j.neubiorev.2005.04.010

RCoreTeam. (2013). R: A language and environment for statistical computing. Vienna, Austria. Retrieved from http://www.R-project.org/

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *J Comp Physiol Psychol, 66*(1), 1-5.

Rhodes, G., Jeffery, L., Evangelista, E., Ewing, L., Peters, M., & Taylor, L. (2011). Enhanced attention amplifies face adaptation. *Vision Res, 51*(16), 1811-1819. doi:10.1016/j.visres.2011.06.008

Robinson, S., Windischberger, C., Rauscher, A., & Moser, E. (2004). Optimized 3 T EPI of the amygdalae. *Neuroimage, 22*(1), 203-210. doi:10.1016/j.neuroimage.2003.12.048

Robinson, T. E., & Flagel, S. B. (2009). Dissociating the predictive and incentive motivational properties of reward-related cues through the study of individual differences. *Biol Psychiatry, 65*(10), 869-873. doi:10.1016/j.biopsych.2008.09.006
Roelofs, K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. *Philos Trans R Soc Lond B Biol Sci, 372*(1718). doi:10.1098/rstb.2016.0206
Rolls, E. T. (2000). Precis of The brain and emotion. *Behav Brain Sci, 23*(2), 177-191; discussion 192-233.
Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and Cognition, 55*(1), 11-29. doi:10.1016/s0278-2626(03)00277-x
Rolls, E. T., O'Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., & McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb Cortex, 13*(3), 308-317.
Roozendaal, B. (2000). 1999 Curt P. Richter award. Glucocorticoids and the regulation of memory consolidation. *Psychoneuroendocrinology, 25*(3), 213-238.
Roozendaal, B., McEwen, B. S., & Chattarji, S. (2009). Stress, memory and the amygdala. *Nature reviews. Neuroscience, 10*(6), 423-433. doi:10.1038/nrn2651
Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychol Rev, 105*(2), 325-350.
Sadacca, B. F., Wikenheiser, A. M., & Schoenbaum, G. (2016). Toward a theoretical role for tonic norepinephrine in the orbitofrontal cortex in facilitating flexible learning. *Neuroscience*. doi:10.1016/j.neuroscience.2016.04.017
Sandi, C. (2013). Stress and cognition. *Wiley Interdiscip Rev Cogn Sci, 4*(3), 245-261. doi:10.1002/wcs.1222
Sandi, C., & Pinelo-Nava, M. T. (2007). Stress and memory: behavioral effects and neurobiological mechanisms. *Neural Plast, 2007*, 78970. doi:10.1155/2007/78970
Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nat Rev Neurosci, 10*(3), 211-223. doi:10.1038/nrn2573nrn2573 [pii]
Schneirla, T. C. (1959). An Evolutionary and Developmental Theory of Biphasic Processes Underlying Approach and Withdrawal. *Nebraska Symposium on Motivation, 7*, 1-42.
Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*(2), 241-263.
Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33*(6), 890-895. doi:10.1016/j.psyneuen.2008.03.001S0306-4530(08)00064-4 [pii]
Schwabe, L., Hoffken, O., Tegenthoff, M., & Wolf, O. T. (2011). Preventing the stress-induced shift from goal-directed to habit action with a beta-adrenergic antagonist. *J Neurosci, 31*(47), 17317-17325. doi:10.1523/JNEUROSCI.3304-11.2011
Schwabe, L., Oitzl, M. S., Philippsen, C., Richter, S., Bohringer, A., Wippich, W., & Schachinger, H. (2007). Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learn Mem, 14*(1), 109-116. doi:10.1101/lm.435807
Schwabe, L., & Schachinger, H. (2018). Ten years of research with the Socially Evaluated Cold Pressor Test: Data from the past and guidelines for the future. *Psychoneuroendocrinology, 92*, 155-161. doi:10.1016/j.psyneuen.2018.03.010
Schwabe, L., Tegenthoff, M., Hoffken, O., & Wolf, O. T. (2010). Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control. *J Neurosci, 30*(24), 8190-8196. doi:10.1523/JNEUROSCI.0734-10.2010

Schwabe, L., Tegenthoff, M., Hoffken, O., & Wolf, O. T. (2012). Simultaneous glucocorticoid and noradrenergic activity disrupts the neural basis of goal-directed action in the human brain. *J Neurosci, 32*(30), 10146-10155. doi:10.1523/JNEUROSCI.1304-12.2012
Schwabe, L., & Wolf, O. T. (2009). Stress prompts habit behavior in humans. *J Neurosci, 29*(22), 7191-7198. doi:10.1523/JNEUROSCI.0979-09.200929/22/7191 [pii]
Schwabe, L., & Wolf, O. T. (2010a). Learning under stress impairs memory formation. *Neurobiol Learn Mem, 93*(2), 183-188. doi:10.1016/j.nlm.2009.09.009
Schwabe, L., & Wolf, O. T. (2010b). Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology, 35*(7), 977-986. doi:10.1016/j.psyneuen.2009.12.010S0306-4530(09)00372-2 [pii]
Schwabe, L., & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. *Behavioural brain research, 219*(2), 321-328. doi:10.1016/j.bbr.2010.12.038S0166-4328(11)00025-8 [pii]
Schwabe, L., & Wolf, O. T. (2013). Stress and multiple memory systems: from 'thinking' to 'doing'. *Trends Cogn Sci, 17*(2), 60-68. doi:10.1016/j.tics.2012.12.001
Schwabe, L., Wolf, O. T., & Oitzl, M. S. (2010). Memory formation under stress: quantity and quality. *Neurosci Biobehav Rev, 34*(4), 584-591. doi:10.1016/j.neubiorev.2009.11.015
Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*(6), 461-464.
Schweinhardt, P., & Bushnell, M. C. (2010). Pain imaging in health and disease--how far have we come? *J Clin Invest, 120*(11), 3788-3797. doi:10.1172/JCI43498
Segal, M., Disterhoft, J. F., & Olds, J. (1972). Hippocampal unit activity during classical aversive and appetitive conditioning. *Science, 175*(4023), 792-794.
Sehlmeyer, C., Schoning, S., Zwitserlood, P., Pfleiderer, B., Kircher, T., Arolt, V., & Konrad, C. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS ONE, 4*(6), e5865. doi:10.1371/journal.pone.0005865
Servatius, R. J., & Shors, T. J. (1994). Exposure to inescapable stress persistently facilitates associative and nonassociative learning in rats. *Behav Neurosci, 108*(6), 1101-1106.
Shansky, R. M., & Lipps, J. (2013). Stress-induced cognitive dysfunction: hormone-neurotransmitter interactions in the prefrontal cortex. *Front Hum Neurosci, 7*, 123. doi:10.3389/fnhum.2013.00123
Shih, W., & Chai, S. (2017). Data-Driven vs. Hypothesis-Driven Research: Making sense of big data. *Academy of Management Proceedings, 2016*(1).
Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K. P., Dayan, P., & Dolan, R. J. (2012). Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease. *Brain, 135*(Pt 6), 1871-1883. doi:10.1093/brain/aws083
Shomstein, S., & Gottlieb, J. (2016). Spatial and non-spatial aspects of visual attention: Interactive cognitive mechanisms and neural underpinnings. *Neuropsychologia, 92*, 9-19. doi:10.1016/j.neuropsychologia.2016.05.021
Shors, T. J. (2004). Learning during stressful times. *Learn Mem, 11*(2), 137-144. doi:10.1101/lm.66604
Shors, T. J. (2006). Stressful experience and learning across the lifespan. *Annu Rev Psychol, 57*, 55-85. doi:10.1146/annurev.psych.57.102904.190205
Shors, T. J., & Servatius, R. J. (1997). The contribution of stressor intensity, duration, and context to the stress-induced facilitation of associative learning. *Neurobiol Learn Mem, 68*(1), 92-96. doi:10.1006/nlme.1997.3763

Shors, T. J., Weiss, C., & Thompson, R. F. (1992). Stress-induced facilitation of classical conditioning. *Science, 257*(5069), 537-539.
Sierra-Mercado, D., Jr., Corcoran, K. A., Lebron-Milad, K., & Quirk, G. J. (2006). Inactivation of the ventromedial prefrontal cortex reduces expression of conditioned fear and impairs subsequent recall of extinction. *Eur J Neurosci, 24*(6), 1751-1758. doi:10.1111/j.1460-9568.2006.05014.x
Sinha, R. (2008). Chronic stress, drug use, and vulnerability to addiction. *Ann N Y Acad Sci, 1141*, 105-130. doi:10.1196/annals.1441.030
Skinner, B. F. (1938). The Behavior of Organisms: An experimental analysis. *The Psychological Record*, 486.
Skinner, B. F. (1963). Operant-Behavior. *American Psychologist, 18*(8), 503-515. doi:DOI 10.1037/h0045185
Small, K. M., Brown, K. M., Forbes, S. L., & Liggett, S. B. (2001). Polymorphic deletion of three intracellular acidic residues of the alpha 2B-adrenergic receptor decreases G protein-coupled receptor kinase-mediated phosphorylation and desensitization. *J Biol Chem, 276*(7), 4917-4922. doi:10.1074/jbc.M008118200M008118200 [pii]
Smith, B. W., Mitchell, D. G., Hardin, M. G., Jazbec, S., Fridberg, D., Blair, R. J., & Ernst, M. (2009). Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *Neuroimage, 44*(2), 600-609. doi:10.1016/j.neuroimage.2008.08.016
Smith, K. S., & Graybiel, A. M. (2016). Habit formation. *Dialogues Clin Neurosci, 18*(1), 33-43.
Smith, S. M., & Vale, W. W. (2006). The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. *Dialogues Clin Neurosci, 8*(4), 383-395.
Sotres-Bayon, F., Bush, D. E., & LeDoux, J. E. (2004). Emotional perseveration: an update on prefrontal-amygdala interactions in fear extinction. *Learn Mem, 11*(5), 525-535. doi:10.1101/lm.79504
Sotres-Bayon, F., & Quirk, G. J. (2010). Prefrontal control of fear: more than just extinction. *Curr Opin Neurobiol, 20*(2), 231-235. doi:10.1016/j.conb.2010.02.005
Spence, K. W. (1950). Cognitive versus stimulus-response theories of learning. *Psychol Rev, 57*(3), 159-172.
Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.
Stevens, F. L., Hurley, R. A., & Taber, K. H. (2011). Anterior cingulate cortex: unique role in cognition and emotion. *J Neuropsychiatry Clin Neurosci, 23*(2), 121-125. doi:10.1176/appi.neuropsych.23.2.12110.1176/jnp.23.2.jnp121
Stice, E., Yokum, S., Burger, K., Epstein, L., & Smolen, A. (2012). Multilocus genetic composite reflecting dopamine signaling capacity predicts reward circuitry responsivity. *J Neurosci, 32*(29), 10093-10100. doi:10.1523/JNEUROSCI.1506-12.2012
Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human pavlovian-instrumental transfer. *J Neurosci, 28*(2), 360-368. doi:10.1523/JNEUROSCI.4028-07.200828/2/360 [pii]
Taylor, F. B., Lowe, K., Thompson, C., McFall, M. M., Peskind, E. R., Kanter, E. D., . . . Raskind, M. A. (2006). Daytime prazosin reduces psychological distress to trauma specific cues in civilian trauma posttraumatic stress disorder. *Biol Psychiatry, 59*(7), 577-581. doi:10.1016/j.biopsych.2005.09.023

Thoma, M. V., Kirschbaum, C., Wolf, J. M., & Rohleder, N. (2012). Acute stress responses in salivary alpha-amylase predict increases of plasma norepinephrine. *Biol Psychol, 91*(3), 342-348. doi:10.1016/j.biopsycho.2012.07.008

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends in Cognitive Sciences, 16*(7), 365-372. doi:10.1016/j.tics.2012.06.003

Todd, R. M., Ehlers, M. R., Muller, D. J., Robertson, A., Palombo, D. J., Freeman, N., . . . Anderson, A. K. (2015). Neurogenetic variations in norepinephrine availability enhance perceptual vividness. *J Neurosci, 35*(16), 6506-6516. doi:10.1523/JNEUROSCI.4489-14.201535/16/6506 [pii]

Todd, R. M., MacDonald, M. J., Sedge, P., Robertson, A., Jetly, R., Taylor, M. J., & Pang, E. W. (2015). Soldiers With Posttraumatic Stress Disorder See a World Full of Threat: Magnetoencephalography Reveals Enhanced Tuning to Combat-Related Cues. *Biol Psychiatry, 78*(12), 821-829. doi:10.1016/j.biopsych.2015.05.011

Todd, R. M., & Manaligod, M. G. M. (2018). Implicit guidance of attention: The priority state space framework. *Cortex, 102*, 121-138. doi:10.1016/j.cortex.2017.08.001

Todd, R. M., Muller, D. J., Lee, D. H., Robertson, A., Eaton, T., Freeman, N., . . . Anderson, A. K. (2013a). Genes for emotion-enhanced remembering are linked to enhanced perceiving. *Psychological Science, 24*(11), 2244-2253. doi:10.1177/0956797613492423

Todd, R. M., Muller, D. J., Lee, D. H., Robertson, A., Eaton, T., Freeman, N., . . . Anderson, A. K. (2013b). Genes for emotion-enhanced remembering are linked to enhanced perceiving. *Psychol Sci, 24*(11), 2244-2253. doi:10.1177/09567976134924230956797613492423 [pii]

Todd, R. M., Muller, D. J., Palombo, D. J., Robertson, A., Eaton, T., Freeman, N., . . . Anderson, A. K. (2014). Deletion variant in the ADRA2B gene increases coupling between emotional responses at encoding and later retrieval of emotional memories. *Neurobiol Learn Mem, 112*, 222-229. doi:10.1016/j.nlm.2013.10.008S1074-7427(13)00203-7 [pii]

Todd, R. M., Schmitz, T. W., Susskind, J., & Anderson, A. K. (2013). Shared neural substrates of emotionally enhanced perceptual and mnemonic vividness. *Front Behav Neurosci, 7*, 40. doi:10.3389/fnbeh.2013.00040

Tolman, E. C. (1932). *Purposive behavior in animals and men.* London, England: Century/Random House UK.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res, 168*(3), 242-249. doi:10.1016/j.psychres.2008.05.006

Triscoli, C., Ackerley, R., & Sailer, U. (2014). Touch satiety: differential effects of stroking velocity on liking and wanting touch over repetitions. *PLoS One, 9*(11), e113425. doi:10.1371/journal.pone.0113425

Tully, K., & Bolshakov, V. Y. (2010). Emotional enhancement of memory: how norepinephrine enables synaptic plasticity. *Mol Brain, 3*, 15. doi:10.1186/1756-6606-3-15

Vallbo, A. B., Olausson, H., & Wessberg, J. (1999). Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. *J Neurophysiol, 81*(6), 2753-2763. doi:10.1152/jn.1999.81.6.2753

van Stegeren, A. H. (2008). The role of the noradrenergic system in emotional memory. *Acta Psychol (Amst), 127*(3), 532-541. doi:10.1016/j.actpsy.2007.10.004

VanElzakker, M. B., Dahlgren, M. K., Davis, F. C., Dubois, S., & Shin, L. M. (2014). From Pavlov to PTSD: the extinction of conditioned fear in rodents, humans, and anxiety disorders. *Neurobiol Learn Mem, 113*, 3-18. doi:10.1016/j.nlm.2013.11.014

Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J Neurosci, 35*(20), 7866-7877. doi:10.1523/JNEUROSCI.0454-15.2015

Visser, R. M., Kunze, A. E., Westhoff, B., Scholte, H. S., & Kindt, M. (2015). Representational similarity analysis offers a preview of the noradrenergic modulation of long-term fear memory at the time of encoding. *Psychoneuroendocrinology, 55*, 8-20. doi:10.1016/j.psyneuen.2015.01.021

Visser, R. M., Scholte, H. S., Beemsterboer, T., & Kindt, M. (2013). Neural pattern similarity predicts long-term fear memory. *Nat Neurosci, 16*(4), 388-390. doi:10.1038/nn.3345

Visser, R. M., Scholte, H. S., & Kindt, M. (2011). Associative learning increases trial-by-trial similarity of BOLD-MRI patterns. *J Neurosci, 31*(33), 12021-12028. doi:10.1523/JNEUROSCI.2178-11.2011

Vogel, S., Fernandez, G., Joels, M., & Schwabe, L. (2016). Cognitive Adaptation under Stress: A Case for the Mineralocorticoid Receptor. *Trends Cogn Sci, 20*(3), 192-203. doi:10.1016/j.tics.2015.12.003

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., de Boer, P., . . . Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biol Psychiatry, 73*(7), 639-645. doi:10.1016/j.biopsych.2012.10.014

Wang, C. E., Brennen, T., & Holte, A. (2006). Decreased approach motivation in depression. *Scand J Psychol, 47*(6), 505-511. doi:10.1111/j.1467-9450.2006.00525.x

Wassum, K. M., & Izquierdo, A. (2015). The basolateral amygdala in reward learning and addiction. *Neurosci Biobehav Rev, 57*, 271-283. doi:10.1016/j.neubiorev.2015.08.017

Waters, A. J., Heishman, S. J., Lerman, C., & Pickworth, W. (2007). Enhanced identification of smoking-related words during the attentional blink in smokers. *Addict Behav, 32*(12), 3077-3082. doi:10.1016/j.addbeh.2007.05.016

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557-561. doi:10.1038/nature02420

Webster, M. A., & MacLeod, D. I. (2011). Visual adaptation and face perception. *Philos Trans R Soc Lond B Biol Sci, 366*(1571), 1702-1725. doi:10.1098/rstb.2010.0360

Weinshenker, D., & Schroeder, J. P. (2007). There and back again: a tale of norepinephrine and drug addiction. *Neuropsychopharmacology, 32*(7), 1433-1451. doi:10.1038/sj.npp.1301263

Wilker, S., Elbert, T., & Kolassa, I. T. (2014). The downside of strong emotional memories: how human memory-related genes influence the risk for posttraumatic stress disorder--a selective review. *Neurobiol Learn Mem, 112*, 75-86. doi:10.1016/j.nlm.2013.08.015

Wolf, O. T., Schommer, N. C., Hellhammer, D. H., McEwen, B. S., & Kirschbaum, C. (2001). The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology, 26*(7), 711-720.

Xie, W., Cappiello, M., Meng, M., Rosenthal, R., & Zhang, W. (2018). ADRA2B Deletion Variant and Enhanced Cognitive Processing of Emotional Information: A Meta-Analytical Review. *Neurosci Biobehav Rev*. doi:10.1016/j.neubiorev.2018.05.010

Yamamoto, T., Shimura, T., Sako, N., Yasoshima, Y., & Sakai, N. (1994). Neural substrates for conditioned taste aversion in the rat. *Behav Brain Res, 65*(2), 123-137.

Yang, L., Zhao, Y., Wang, Y., Liu, L., Zhang, X., Li, B., & Cui, R. (2015). The Effects of Psychological Stress on Depression. *Curr Neuropharmacol, 13*(4), 494-504.

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat Rev Neurosci, 7*(6), 464-476. doi:10.1038/nrn1919

Yokum, S., Marti, C. N., Smolen, A., & Stice, E. (2015). Relation of the multilocus genetic composite reflecting high dopamine signaling capacity to future increases in BMI. *Appetite, 87*, 38-45. doi:10.1016/j.appet.2014.12.202

Zarrindast, M. R., Bahreini, T., & Adl, M. (2002). Effect of imipramine on the expression and acquisition of morphine-induced conditioned place preference in mice. *Pharmacol Biochem Behav, 73*(4), 941-949.

Zhang, H. F., Li, X. L., Huang, J., Li, Y., Thijs, L., Wang, Z. Z., . . . Wang, J. G. (2005). Cardiovascular and metabolic phenotypes in relation to the ADRA2B insertion/deletion polymorphism in a Chinese population. *Journal of Hypertension, 23*(12), 2201-2207. doi:DOI 10.1097/01.hjh.0000189869.48290.91

Zimmerman, J. M., Rabinak, C. A., McLachlan, I. G., & Maren, S. (2007). The central nucleus of the amygdala is essential for acquiring and expressing conditional fear after overtraining. *Learn Mem, 14*(9), 634-644. doi:10.1101/lm.607207

Zink, C. F., Pagnoni, G., Martin-Skurski, M. E., Chappelow, J. C., & Berns, G. S. (2004). Human striatal responses to monetary reward depend on saliency. *Neuron, 42*(3), 509-517.

Zorawski, M., & Killcross, S. (2002). Posttraining glucocorticoid receptor agonist enhances memory in appetitive and aversive Pavlovian discrete-cue conditioning paradigms. *Neurobiol Learn Mem, 78*(2), 458-464.

Zorawski, M., & Killcross, S. (2003). Glucocorticoid receptor agonist enhances pavlovian appetitive conditioning but disrupts outcome-specific associations. *Behav Neurosci, 117*(6), 1453-1457. doi:10.1037/0735-7044.117.6.1453

Appendix A. Pre-registration protocol

The influence of naturally occurring differences in human norepinephrine availability in emotional bias flexibility – A study protocol

Mana R. Ehlers[1], Colin J. Ross[2], Rebecca M. Todd[1]
[1]Department of Psychology, University of British Columbia, Canada
[2]Centre for Molecular Medicine and Therapeutics, Child and Family Research Institute, Department of Medical Genetics, University of British Columbia, Canada

Introduction

Typically, emotionally salient stimuli are prioritized over the mundane. Emotionally or motivationally salient events hold a special place in our memory and our attention is drawn towards emotionally relevant objects or scenes (LaBar & Cabeza 2006; Cahill & McGaugh 1998; Pourtois et al. 2013). While such affective biases in cognition can be highly adaptive, extreme biases favoring specific categories of aversive or rewarding stimuli can be symptomatic of psychopathology (Bar-Haim et al. 2007; Peckham et al. 2010). Thus it is important to understand which factors contribute to the development of such extreme biases and the degree to which existing biases can be altered.

Among other factors, previous research has found emotional enhancement of attention and memory to be linked to activity in the norepinephrine (NE) system (van Stegeren 2008; Berridge & Waterhouse 2003). Yet NE-mediated processes do not function identically across individuals. A deletion variant in the *ADRA2b* gene, which is associated with impaired function of alpha2b inhibitory adrenergic autoreceptors, is thought to be associated with increased intersynaptic NE availability (Small et al. 2001; de Quervain et al. 2007). Thus, the *ADRA2b* polymorphism, which we have observed in > 50% of Canadian participants of European and East Asian descent, provides a means to observe naturally occurring differences in NE availability. Previous research has found it to be associated with enhancled emotional modulation of attention and memory (de Quervain et al. 2007; Todd et al. 2013; Todd et al. 2014). An outstanding question concerns the role of such differences in emotional learning and flexibility.

A hypothesized role for noradrenergic alpha2b receptors in emotional learning is supported by rodent studies showing reduced emotional learning upon full development of inhibitory alpha2b receptors (Moriceau & Sullivan 2004). In line with that, we expect human deletion variant carriers, who have reduced alpha2b function, to show facilitated emotional learning indicated by greater flexibility in shifting pre-existing patterns of emotional biases.

Since several genetic polymorphisms are known to be associated with enhanced NE availability, we extend our approach to include genotyping for a total of five polymorphisms affecting the NE system. Calculating a multilocus composite score (Yokum et al. 2015) will allow us to create individual NE profile scores that will be explored in relation to bias flexibility.

Further, we will combine genotyping with stress induction in healthy undergraduates to potentiate natural occurring differences in NE availability in order to determine the influence of NE on bias flexibility. Participants will be genotyped for the *ADRA2b* deletion variant. For exploratory analyses, additional polymorphisms influencing NE function will be included. In order to potentiate naturally occurring differences in NE availability based on genotype and to explore effects of arousal on bias flexibility, we combine genotyping with acute stress induction. In the randomly assigned stress condition, stress will be induced by means of the commonly

employed socially evaluated cold pressor test (SECPT) (Schwabe et al. 2008). Degree of bias before and after training will be indexed by a so called bias probe task in which participants categorize the expressions on faces morphed along a continuum from happy to angry. A bias probe task will be performed before and after a training task. The proportion of faces rated as angry is a measure of negative bias. Bias flexibility will be operationalized as the degree to which bias moves in a more positive direction following training. Training is performed via an adaptation task, in which emotional biases to rate ambiguous emotional expressions as angry or happy will be trained in a positive direction (Penton-Voak et al. 2013). The training task exploits well-documented facial emotion adaptation effects (Webster et al. 2004) by repeatedly exposing participants to angry faces.
We hypothesize that *ADRA2b* deletion carriers will show an increased adaptation effect indicating greater emotional flexibility mediated by activity in the NE-system.

Measures

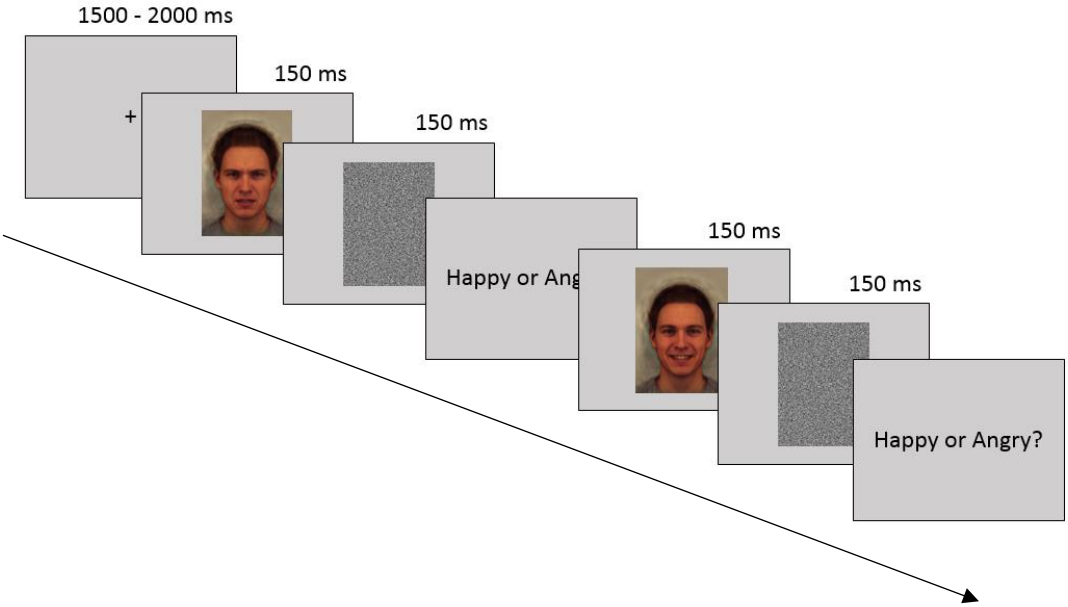

Figure 1. Schematic of bias probe performed before and after adaptation. Adapted from Penton-Voak et al. 2013. Participants are presented with ambiguously happy and angry faces and are asked to make a forced-choice assessment of whether a face was happy or angry.

Bias Probe:
The Bias Probe task will be performed before and after the adaptation task to assess individual baseline biases and assess the change by adaptation. Each trial begins with the randomly jittered (1500 – 2000 ms) presentation of a fixation cross followed by the display of a face (1 of 15 frames taken from a continuum of emotional faces ranging from unambiguously happy to unambiguously angry). A mask of visual noise is presented for 150 ms before participants are asked to judge whether the face just seen was happy or angry.
Each participant completes a total of 90 trials. Two sets of emotional continua consisting of 15 frames are presented three times each in randomized order. Facial stimuli are an East Asian and a Caucasian female taken from the NIMSTIM database.

Separately for each frame, we calculate the proportion of trials judged as "angry" in order to estimate the balance point at which a face is perceived as equally happy and angry. We can further assess the total proportion of faces rated as happy and angry. The comparison of response curves before and after adaptation will also allow us to investigate whether only one end of the face continuum is affected by adaptation or both.
Bias Probe baseline and retest follow the exact same structure. The participant is reminded of the instructions before completing the retest.

Adaptation:
Adaptation is accomplished by asking participants to perform a 21 minute 2-back memory task. This task exploits the well-documented phenomenon of visual aftereffects, in which, after repeated exposure to one category of visual stimulus, an ambiguous stimulus looks more like the opposite category. Thus repeated exposure to an angry face shifts perception of a neutral or ambiguous facial expression as more happy (Webster et al. 2004). The stimuli used in the task are 10 individual faces all displaying angry expressions taken from the NIMSTIM face database. Faces are presented in random order and upon presentation of the face, participants have to indicate via button press whether any given face is the same as that two faces back (target) or not (nontarget).
Each face is displayed for 2000 ms with 20 ms intertrial intervals. Each participant completes 13 adaptation blocks, each block consisting of 48 trials containing eight targets each.
The n-back task is employed to ensure that participants stay on track and pay attention to the facial expressions.

Stress measures:
Stress induction is added to the protocol in order to potentiate naturally occurring differences in NE availability driven be the *ADRA2b* polymorphism. Stress is induced with the socially evaluated cold-pressor test (SECPT) (Schwabe et al., 2008). In the stress condition, participants are first informed that their faces will be videotaped during the upcoming test for future evaluation of their facial expressions. Participants are then asked to put a hand in ice water (0 – 4 ºC) up to the wrist. They are told to keep the hand in the water for as long as possible while looking straight into the camera. The experimenter observes the participant at all times and records the time period during which the participant has his hand in the water. After 3 minutes participants are instructed to remove their hands from the water if they had not done so before. In the control condition the ice water is replaced by warm water (35 – 37 ºC) and participants are neither videotaped nor watched by the experimenter. They are likewise instructed to keep their hand in the water and the experimenter makes sure to look otherwise occupied.
SECPT questionnaire. To obtain a measure of subjective, psychological stress response we ask participants to rate how stressful, painful and unpleasant the SECPT was using a ten-point scale ranging from 1 ("not at all") to 10 ("extremely").
Heart rate. Heart rate is measured using LabChart software (AD Instruments) based on a finger pulse that is continuously measured with a pulse transducer (AD Instruments). Heart rate is determined at baseline and is continuously recorded after stress induction.
Blood pressure. Systolic and diastolic blood pressure are measured using a blood pressure monitor before as well as at several time points after stress induction.
Salivary cortisol analysis. Saliva is collected with a Salivette collection kit (Sarstedt AG & Co., Nümbrecht, Germany) and stored at -20 ºC until the biochemical analysis of salivary levels of

free cortisol. Analysis employs a luminescence immunoassay (IBL GmbH, Hamburg, Germany) performed by the lab of Prof. Dr. C. Kirschbaum, Dresden, Germany.

Questionnaire measures:
Questionnaires will include a custom-made demographics questionnaire, the Beck Depression Inventory (BDI), the Spielberger State-Trait Anxiety Inventory (STAI; Spielberger et al., 1983), the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegan, 1988), the Liebowitz Social Anxiety Scale (LSAS) and the Childhood Trauma Questionnaire (CTQ). We do not expect a relation between performance on these measures and *ADRA2b*. However, anxiety, depression and a history of trauma may influence degree of bias and bias flexibility and if so these measures will be used as control variables.

Genotyping:
Participants will be genotyped for the *ADRA2b* polymorphism (observed in ~50% of populations we have studied). Participants will also be genotyped for the *5-HTTLPR* polymorphism, which has been associated with emotional bias and is included as a control variable. For further exploratory analyses employing multilocus scores (Yokum et al. 2015) based on an array of genes influencing NE availability participants will also be genotyped for *ADRA2a*, *ADRA2c*, *NET* and *NET1*. Based on previous research, homo- and heterozygous *ADRA2b* deletion carriers are treated as one group due to the low number of homozygotes.

Study Objective and Hypotheses
The study aims to investigate how naturally occurring differences in norepinephrine availability (*ADRA2b* genotype) affect emotional learning and flexibility. More specifically, we want to examine whether emotional biases for threat are more flexibly altered in *ADRA2b* deletion carriers. Stress induction will be used to potentiate genotype-dependant differences in the NE and arousal system. This line of research is not only important for understanding the role in bias flexibility but has also implications for therapeutic modification of maladaptive biases.

Hypothesis 1: It is predicted that the adaptation effect, i.e. reduction of negative bias will be more pronounced in *ADRA2b* deletion carriers compared to noncarriers.

Hypothesis 2: It is predicted that the enhanced adaptation effect in deletion carriers will be potentiated by the stress induction.

Hypothesis 3: It is predicted that we will not observe differences in the degree of initial negativity bias between *ADRA2b* deletion carriers and noncarriers.

Hypothesis 4: Based on results from a pilot study examining the effect of stress on emotional bias, we predict that the initial bias will not be influenced by stress induction.

Hypothesis 5: It is predicted that higher multilocus composite scores are associated with greater bias flexibility.

Additional Research Question:
Does acute stress induction alone affect adaptation indicated by altered affective bias?

Study Design
The study will use a between-subjects design where one group will be randomly assigned to the stress condition and the other to the control condition. All participants will perform the experimental tasks and fill out the questionnaire before providing saliva samples for genotyping. Dependent variables are degree of bias measured as proportion of faces rated as angry, and the primary dependent variable of interest is the change in emotional bias between pre- and post-adaptation training. (See above for detailed task description)

Participants and Recruitment
A minimum of 252 participants will be recruited through the UBC psychology human subject pool (HSP). In addition, participants may be recruited through existing e-mail lists or poster advertisements. After reading the study description and clarifying potential questions, participants who meet the inclusion criteria will be invited to the lab to take part in the study. Participants will be reimbursed by course credit or will be paid $10 per hour.

Sample Size Determination
Power analysis for *ADRA2b* is based on effect sizes found in previous studies (ηp of .05). For sufficient power for a repeated-measures ANOVA with *ADRA2b* and stress as between-subject factors we require a sample size of 252.

Statistical Analysis
Bias Probe:
In both baseline and retest bias probe task, the bias for detecting angry faces will be calculated for each of the 15 frames averaged over repetitions. An ANOVA with genotype and stress conditions as between-subject factors will be performed in order to compare the bias between groups. Sex and ethnicity will be included as covariates.
We predict that initial bias is not affected by *ADRA2b* genotype and results from a pre-study suggest that baseline emotional biases are also not affected by acute stress induction. However, we expect that *ADRA2b* deletion carriers show stronger adaptation as indicated by reduced bias to detect angry faces.
In addition, we will calculate a multilocus composite score as it has been used previously (Yokum et al. 2015; Stice et al. 2012; Nikolova et al. 2011). Genotypes that are thought to be associated with increased NE availability will scored as 1 whereas genotypes related to low NE levels will be given a score of 0. The sum will be calculated for each participant resulting in a possible maximum of 5 and a minimum of 0. Individuals will be classified into high, intermediate and low NE profile score groups. Exploratory regression analyses will be performed to determine whether higher multilocus genetic profile scores signaling higher NE signaling predict bias flexibility.

Performance in the n-back task by means of which adaptation is achieved will also be explored. The main variables of interest in the task is the accuracy with which targets are detected. Although not a primary measure of interest, N-back performance will be analyzed with a 2 x 2 ANOVA with condition (stress and control) and genotype (deletion carrier and noncarrier) as factors.

Exploratory correlations with the BDI, STAI, LSAS, PANAS and CTW will also be explored and variables correlated with differences in bias will be included as covariates in our primary analyses.

Data Management
Data will be anonymized by a unique numeric identifier and stored on a password protected server during the course of the study. Confidential participant information such as name and contact details will be kept separately and encrypted and password protected.

When data collection is complete fully anonymized data will be made available for sharing via a university online repository or figshare.com.

Ethical Considerations
Ethics approval has been obtained from the Behavioral Research Ethics Board of the University of British Columbia (BREB Number: H15-00711). We do not foresee any risks for the participants taking part in the study. However, some participants may be upset by the stress induction. Participants will be informed that they are free to withdraw from the study at any time, without any bearing on their compensation.
Participants will be fully informed about the nature of the experiment and will be given sufficient time to ask questions before signing the consent form.

Quality Control
The investigators will be responsible for data quality.

Study Site
All experiments will be performed at the Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver BC, Canada. Questionnaires will be completed online.

Insurance
Insurance will be provided by the University of British Columbia.

Publication Policy
The study findings may be published in a scientific journal and/or presented at meetings. Study data will be shared anonymously if required by the publisher.

Study Personnel
Mana Ehlers
Department of Psychology
2136 West Mall
Vancouver, BC V6T 1Z4
Canada
Email: manaehlers@psych.ubc.ca

Dr. Colin Ross
Child & Family Research Institute
Room A3-216, 950 West 28th Avenue
Vancouver, BC V5Z 4H4
Canada
Email: cjross@cmmt.ubc.ca

Dr. Rebecca Todd
Department of Psychology
2136 West Mall
Vancouver, BC V6T 1Z4
Canada
Email: becket.todd@psych.ubc.ca