Look at Me!

The socio-neuro-endocrinology of eye-contact, dominance and status

Kijk naar Mij!

De socio-neuro-endocrinologie van oogcontact, dominantie en status

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 13 juni 2012 des middags te 2.30 uur

door

David Terburg

geboren op 17 juli 1980

te Naarden

Promotoren: Prof.dr. E. J. van Honk
Prof.dr. H. Aarts

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) aan Jack van Honk (NIHC #056-24-010) en Henk Aarts (VICI #453-06-002).

Look at Me!

The socio-neuro-endocrinology of eye-contact, dominance and status

David Terburg

ISBN/EAN: 978-90-6464-563-1

Printed in the Netherlands by GVO drukkers & vormgevers B.V. | Ponsen & Looijen
Cover: Joppe Rovers / 40rovers

Table of contents

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Chapter 8

Chapter 9

Chapter 10

Chapter 11

Chapter 12

Voor mijn ouders

Chapter 1

General introduction and outline

"I have known a vast quantity of nonsense talked about bad men not looking you in the face. Don't trust that conventional idea. Dishonesty will stare honesty out of countenance any day in the week, if there is anything to be got by it."

Mr. Sampson, in 'Hunted Down' by Charles Dickens (1859)

Introduction

When people take part in discussions, card games, or any other social interaction, they regularly make eye-contact and follow each other's gaze to learn more about the interests, goals, and thoughts, of the persons they are dealing with. Generally, we are also aware that other individuals do the same thing, and therefore can easily use gaze and eye-contact to convey meaning and communicate non-verbally. The importance of gaze and eye-contact in human social environment is therefore evident. A vivid example of its impact on human culture is the phenomenon 'Evil Eye', which describes a look of envy or jealousy that is presumed to evoke bad luck or even injury. References to this phenomenon range from modern society back to classical antiquity, and people go as far as to protect themselves with the help of charms and talismans (Dundes, 1992). Humans are indeed extremely sensitive to gaze and eye-contact, and as will be described in this thesis, this is a largely unconscious and automatic mechanism that is rooted in evolution. Evidently, nonverbal social communication greatly benefits survival of humans, and other primates, as a group. Eye-contact is however also an important mediator in the formation of dominance hierarchies within the group. Dominant alpha-males use eye-contact with subordinate conspecifics to settle conflicts without resorting to violence. These displays of social dominance can be described in terms of a staring contest. Interruption of eye-contact, or 'gaze-aversion' during such staring contests, is thereby a signal of submissiveness and subordination (Mazur & Booth, 1998). Nonverbal gaze communication is therefore not only beneficial to survival of the social group, but nonverbal displays of dominance also promote survival of the individual within that group as

it can increase the access to resources and mating. As will be shown in this thesis, this mechanism of dominant eye-contact can also be observed in humans, and its implicit and automatic properties indeed suggest an evolutionary origin. Following the strong interrelation between on the one hand dominance and submissiveness and on the other hand social aggression and anxiety, the first goal of this thesis is to develop the experimental tools to study these automatic mechanisms of dominant eye-contact, and thereby provide new methods to study, diagnose and treat these psychopathologies.

This thesis will furthermore provide a biological framework of dominance behavior, which will be formulated with particular attention to brain function and the hormone testosterone. Crucially, testosterone has been argued to be involved in several forms of dominance behavior. Testosterone promotes on the one hand automatic, reflexive and nonverbal behaviors that underlie aggressive tendencies (Hermans, Putman, Baas, Koppeschaar, & van Honk, 2006a; Hermans, Putman, & van Honk, 2006b; van Honk, Peper, & Schutter, 2005; van Honk et al., 2001b; Wirth & Schultheiss, 2007), but the hormone can also promote social cooperative decision making (Eisenegger, Naef, Snozzi, Heinrichs, & Fehr, 2010). In this thesis these seemingly opposing behaviors will be combined into a biological framework, and it will be argued that testosterone promotes, depending on the social context, any behavior that is beneficial to social status. As reflected in the title of this thesis, these forms of status seeking behavior can both be described with the phrase "Look at Me!", which either refers to the status enhancing effects of doing good for the community, but also to nonverbal displays of dominance in the form of staring contests and reactive aggression.

Social factors of dominance: Eye-contact and facial expressions

Humans are predisposed to non-verbal communication through gaze (Emery, 2000; Frith, 2008), particularly because the human eye has uniquely evolved properties that facilitate detection of gaze-direction, i.e. elongated width and extreme whiteness of the sclera (Kobayashi & Kohshima, 2001). Already in early infancy humans detect gaze-direction (Farroni, Csibra, Simion, & Johnson, 2002; Farroni, Johnson, & Csibra, 2004a; Grossmann & Johnson, 2007), and newborns as young as two days old are able to follow the gaze of others

(Farroni, Massaccesi, Pividori, & Johnson, 2004b). Such an early onset of gaze communication suggests an evolutionary origin, and can indeed also be observed in our closest relatives, the primate. Already at very young age macaques follow each other's gaze (Emery, 2000; Ferrari, Paukner, Ionica, & Suomi, 2009; Ferrari et al., 2006). Moreover, macaques also have the ability to follow human gaze (Deaner & Platt, 2003), and are only aggressive towards humans that make eye-contact (Kalin & Shelton, 1989), which suggests that macaques and humans share a similar neural circuitry of gaze-perception (Deaner & Platt, 2003). Although the eyes of most primates do not share all of the profound adaptive features human eyes have (Kobayashi & Kohshima, 2001), nonverbal gaze communication is evidently an important part of their social skills (Emery, 2000). Communication through gaze seems therefore to be beneficial to survival. This is easily to conceive since it is an efficient means of communicating impending danger, and provides a simple and automatic mechanism to act as a homogenous group in the search for resources.

Gaze communication has apparently adaptive properties that promote the survival of social groups, but it has also evolved crucial features that promote survival of the individual within a social group (Emery, 2000). Humans automatically, and rapidly, detect social threat as expressed with facial features. This basal form of social threat-vigilance promotes survival of individuals and groups, and involves activation of the sympathetic nervous system that prepares the mind and body for action, e.g. fight or flight. We process for example the emotional content of most basic facial expressions, i.e. anger, disgust, fear, happiness, sadness and surprise, even before we are aware of them (Vuilleumier, 2002). This is also the case for the direction of someone else's gaze (Langton, Watt, & Bruce, 2000). A sole gaze-shift, however, only reveals the location of something potentially interesting. When it is accompanied with a facial expression, a gaze-shift acquires relevance and meaning with regard to mental state and social environment (Itier & Batty, 2009). The interaction of face and gaze perception acts even further, for eye-contact facilitates awareness of faces that are masked from conscious perception (Stein, Senju, Peelen, & Sterzer, 2011). This is especially the case for facial expressions of fear and anger. Angry faces are more easily recognized and perceived as more intense during eye-contact, while for fearful faces the same is true when gaze is averted (Adams & Kleck, 2003, 2005; Milders, Hietanen, Leppänen, & Braun, 2011). Indeed, from a social signaling perspective the fearful face with averted gaze signals a threat

in the environment, while fearful faces that make eye-contact could also be afraid of you, and are therefore a more ambiguous threat signal. A similar concept holds for the angry face with averted gaze, as its anger is most likely directed to something or someone else, which renders it not directly threatening.

Angry eye-contact is apparently such a highly relevant social signal, that it has evolved to be processed automatically. It is indeed also a means of aggressive communication in several rodent species (Brecht & Freiwald, 2011), and a main factor in the formation of social hierarchies within primate groups (Mazur & Booth, 1998). The angry stare is however also important for survival and reproduction of the individual within the group. First of all, dominant primates have first access to the best resources, and often exclusive mating rights (Archer, 2006). Crucially, dominance hierarchies in primates are mostly maintained without physical aggression. For instance, subordinate individuals in primate groups tend to always keep track of the location of dominant conspecifics, but refrain from looking directly at them, which limits the chance on aggressive confrontation (Mazur, 1985; Setchell & Wickings, 2005). When two primates do establish eye-contact, a staring contest may arise. They maintain eye-contact and exchange expressions of anger and intimidation until eventually the subordinate will avert gaze in order to prevent a violent confrontation (Mazur & Booth, 1998). We can thus distill two distinct forms of subordinate gaze behavior. First, reluctance to make eye-contact with the dominant alpha-male (gaze-avoidance). Second, in the event that such eye-contact does occur, the submissive animal will rapidly avert gaze (gaze-aversion), thereby signaling subordination which prevents a violent confrontation.

In other words, primates with a high rank in the dominance hierarchy tend to out-stare lower ranked conspecifics, which might accumulate to violence when the subordinate fails to act submissively. In this thesis it will be shown that similar mechanisms are observed in humans, and that these are highly implicit and automatic. Given that most of human psychopathology involves dysfunctional anxiety or aggression, unraveling such mechanisms of dominance and submission is of high societal relevance. Indeed, similar to primates, human conflict is usually resolved without resorting to violence. Although our societal structure is of course increasingly complex, people usually adhere to the hierarchy and do as told by their supervisor, parents, or teacher. Some individuals are, however, repeatedly involved in aggressive conflicts and violence. Others, are chronically anxious, and resort to avoidance of

any potential conflict and deter into social phobia or depression. The first goal of this thesis is therefore to find out how eye-contact and gaze behavior can be used as predictors of personality traits of anxiety and aggression.

Biological factors of dominance: Steroids and the brain

The second goal of this thesis is to describe the neural and hormonal factors underlying these mechanisms of eye-contact and social dominance, and how these contribute to psychopathologies of anxiety and aggression. In the current literature on this subject the most studied brain area is the amygdala. The amygdala is a small, almond shaped brain structure, located in the temporal cortex of each hemisphere of the brain. The amygdalae sit between the subcortex and cortex, and are therefore ideally situated to influence and coordinate both. Indeed, part of the amygdala shares properties in neural structure with subcortical areas that are traditionally argued to be involved in basal processing and automatic, unintentional behavior, whereas other parts of the amygdala share the neural structure of cortical areas, that are more involved in higher-order processing and intentional behavior (Whalen & Phelps, 2009). This is an important feature of the amygdala, because many of human psychopathologies are argued to arise from a misbalance between the subcortex, that humans share with most other vertebrates and developed early in evolution, and the later developed mammalian cortex (MacLean, 1990; Porges, 2001). As such, the amygdala has been argued to play a pivotal role in a subcortical-cortical alarm system, whereby the amygdala relays the output from fast basal processing of threat in the subcortex to the slower, but more deliberate, cortex (Liddell et al., 2005; Phelps & LeDoux, 2005; Whalen & Phelps, 2009). This allows humans to respond rapidly to threats by using their ‘old’ reptilian instincts, but also works in the opposite direction. In other words, cortical structures and the amygdala jointly influence and regulate basal subcortical actions. In this bidirectional relationship the amygdala thus transfers threat-related information from subcortex to cortex, but also directly affects the behavioral output of the subcortical systems in response to threat (Whalen & Phelps, 2009). This might explain why the amygdala seems to be involved in nearly any process where an affective factor plays a role, but makes it hard to pinpoint what the specific role of the amygdala is in human behavior. Indeed, researchers have ascribed often contrasting affective

properties to the amygdala; e.g. conscious and unconscious processing, learning and unlearning, threat-assessment and threat-responding, fear and anger, etc. (Pessoa & Adolphs, 2010). As we will see in this thesis, this framework still largely holds, but only when the amygdala is regarded as a heterogeneous brain structure. Importantly, the amygdala is not one, but a diverse set of mutually excitatory and inhibitory structures, which may explain the diversity of roles the amygdala seems to play in both fear and aggression.

Interestingly, a similar brain network has been argued to be involved in the processing of eye-contact. As described above, both facial expression and eye-contact are processed preconsciously in subcortical areas (Senju & Johnson, 2009), and angry faces with direct gaze activate the amygdala more than when gaze is averted (Sato, Yoshikawa, Kochiyama, & Matsumura, 2004). Shifts of gaze are detected in the superior temporal sulcus (STS), but gaze-following has been argued to be hard-wired in the brain, again involving the interaction of cortical areas and the amygdala (Emery, 2000). Crucially, socially anxious individuals show profound cardiac acceleration during eye-contact (Wieser, Pauli, Alpers, & Mühlberger, 2009), avoid eye-contact with emotional faces (Mühlberger, Wieser, & Pauli, 2008), and avoid angry faces specifically when there is eye-contact (Roelofs et al., 2010). These results indeed confirm that eye-contact is processed automatically as threatening through basal subcortical processing and sympathetic arousal, and is evaluated cortically in a brain network centered around the amygdala. Furthermore, social anxiety is characterized by fear of social evaluation (Watson & Friend, 1969), and is strongly related to reduced dominance behavior or submissiveness (Trower & Gilbert, 1989; Weeks, Heimberg, & Heuer, 2011), which confirms the profound parallels between the processing of eye-contact and personality characteristics in the dominance-submission dimension.

Another important mediator of social dominance-submission behavior are steroid hormones. The steroid hormone cortisol is generally a reliable predictor of anxiety and fear (Johnson, Kamilaris, Chrousos, & Gold, 1992), and is therefore an important marker for hyper-vigilant responding to threat in general. The steroid hormone testosterone, on the other hand, is causally involved in aggression in many species (Nelson & Trainor, 2007), but crucially only in social aggression against conspecifics (Archer, 2006). Interestingly, recent evidence suggests that testosterone can also promote social cooperative behavior, which would stem from the motivation to increase social status (Eisenegger et al., 2010). Indeed,

both social aggression and the drive for social status are behaviors that shape the dominance-submission relationship between conspecifics. As described above in primates, testosterone is therefore potentially an important mediator in the mechanisms of eye-contact and aggression, but also for higher-order social decision making. Furthermore, testosterone acts on the same network of subcortex-amygdala-cortex (Hermans, Ramsey, & van Honk, 2008; van Wingen, Mattern, Verkes, Buitelaar, & Fernandez, 2010), and can therefore provide important information on the amygdala's role in social behavior, fear and aggression. The third goal of this thesis is therefore to describe the role of testosterone in status seeking behavior, with particularly attention to the distinctions and parallels between reactive aggression and social cooperation, and the brain network centered around the amygdala. Since testosterone also has important anxiolytic properties (Hermans et al., 2006a; Hermans et al., 2006b; van Honk et al., 2005), this intricate relationship of testosterone with status might help to elucidate how and when testosterone has therapeutic value for the treatment of disorders of anxiety and fear (Haglund, Nestadt, Cooper, Southwick, & Charney, 2007).

Outline of this thesis

The thesis starts in **Chapter 2** with a theoretical framework rooted in the interaction of endocrine and brain functions (Terburg, Morgan, & van Honk, 2009a). This framework draws from the concept that motivated behavior in humans, and other animals, can be described within the dimension of approach-avoidance. In general, humans and other animals will not put effort in something that is neutral to them, but when something is desirable, or aversive, they will be motivated to approach or avoid it respectively, a process which affects our goals and motives even unconsciously (Dijksterhuis & Aarts, 2010). This is a crucial aspect of evolution in general, and survival in particular, that not only motivates humans and other animals to avoid potential danger and reproduce (Ressler, 2004), but also maintains and shapes new behaviors according to very basic principles already described by Pavlov in his early conditioning studies (Pavlov, 1927). This notion has important implications for the interpretation of how anxiety and aggression develop in humans, and how these personality characteristics relate to biological factors. Crucially, aggression and anger are approach motivated goal-directed behaviors and traits, whereas fear and anxiety are related to response

inhibition and avoidance (Carver & Harmon-Jones, 2009; Ernst & Fudge, 2009; Harmon-Jones, 2003a, 2004; van Honk & Schutter, 2007b). This is especially important from a social perspective, because disorders of anxiety and aggression are often linked to a social context, e.g. fear of social evaluation and domestic violence, which is also an important factor in how these disorders are diagnosed (American Psychiatric Association, 2000). Within this social context anxious or aggressive personality styles typically dissociate in their basic reaction to facial threat. Although both anxious and aggressive individuals respond vigilantly to threat, the first will subsequently defensively avoid it (Mogg, Bradley, de Bono, & Painter, 1997), whereas the second will approach and confront it in search of a possible rewarding outcome (van Honk & Schutter, 2007b). In other words, an anxious individual will vigilantly detect a threat, but will respond with a flight reaction, or avoidance, whereas an aggressive individual also will rapidly detect it, but will respond by approaching the threat to fight it.

As will be described in **Chapter 2**, these personality styles are strongly linked to biological and neurological factors, namely steroid hormones and a balance between 'low' subcortical brain regions, which we share with other vertebrates, and 'high' cortical brain regions that later developed in mammals and humans. Importantly, these subcortical areas are mostly involved in automatic responding to threat, whereas the cortex is more involved in higher-order and conscious processing (MacLean, 1990; Porges, 2001). Furthermore, steroid hormones strongly influence both levels. Generally, high endogenous levels of the steroid hormone cortisol are an indication of enhanced automatic fear-vigilance, and an avoidant, fearful character, whereas high levels of testosterone are related to aggressive vigilance, approach motivation, and a dominant personality style. Especially when testosterone levels are high, and cortisol levels are low, humans are therefore predisposed to social aggression, e.g. they have the motivation to dominate, and are not fearful to act on it (Terburg et al., 2009a). A better understanding of these biological factors and the individuals' response to threat, might therefore provide important information for diagnoses and treatment of psychopathology in humans, with on the one hand disorders of fear and anxiety (depression, phobia, etc.), and on the other hand disorders of aggression (psychopathy, conduct disorder, etc.) (American Psychiatric Association, 2000).

As discussed above, an important brain structure for the perception of, and response to, facial threat, is the amygdala. In human neuroscience this structure is often described as a

homogeneous structure implicated in many forms of social and emotional behavior (Pessoa & Adolphs, 2010; Whalen & Phelps, 2009), and considered to be crucial for reflexive and unconscious reactions to threat (Phelps & LeDoux, 2005; Whalen et al., 2004). The amygdala is however not a homogeneous brain-region. Its several sub-nuclei are highly different in structure and connectivity and are therefore best regarded separately (Davis & Whalen, 2001; Heimer, Harlan, Alheid, Garcia, & de Olmos, 1997; McNaughton & Corr, 2004). **Chapter 3** provides a neural framework involved in rapid, automatic threat-vigilance, based on amygdala sub-region functionality. We show in a multimodal study using a combination of neuropsychological and neuroscientific techniques; e.g. cognitive and affective measures, unconscious processing, eye-tracking, and structural and functional magnetic resonance imaging (sMRI and fMRI), that patients with brain damage to the basolateral sub-region of the amygdala (BLA), but a still functional central-medial amygdala (CMA), respond hyper-vigilantly to fearful faces. Embedded in the literature on research in rodents and primates, we argue that the BLA can attenuate the CMA's output to subcortical areas, both directly and indirectly through the orbitofrontal cortex (OFC), thereby reducing threat-vigilance. We propose therefore a neural model wherein the CMA promotes threat-vigilance, and the BLA and OFC jointly regulate the CMA. Since increased threat-vigilance is a recurring symptom in the psychopathology of anxiety and aggression, this model provides an important framework for the study of these disorders.

The following three chapters of this thesis turn to a direct investigation of the mechanisms underlying eye-contact and social dominance in three newly developed eye-tracking paradigms. In these studies gaze-behavior is linked to personality characteristics based on the approach-avoidance model, and interpreted from a neural framework centered around the amygdala. The first study is a gaze-imitation experiment in **Chapter 4**. In this study we used the technique eye-tracking, but not in the conventional way as a measure of where, how often, and how long, individuals look at the presented stimuli. Instead, we developed a reactive eye-tracking paradigm, whereby participants respond to events on a computer screen by gazing as fast as possible at a predefined location on that screen. By manipulating the social context before the participant's gaze-shift we can measure how social information influences natural gaze behavior. We show in this study that an observed gaze-shift, i.e. one makes eye-contact with someone who subsequently looks to the left or right, is

reflexively followed. In other words, when you see someone shifting gaze, you look automatically in the same direction, which we call gaze-imitation (Terburg, Aarts, Putman, & van Honk, 2012a). Importantly, the gaze-imitation reflex was stronger when the observed gaze-shift was accompanied by dynamic facial expressions of fear. As discussed earlier a fearful gaze-shift is a signal of impending threat from the gaze-cued direction, thus the enhanced reflex to follow this gaze-shift confirms the general tendency of threat-vigilance in humans. Furthermore, trait anger was in this study associated with stronger reflexive gaze-imitation towards reward, as signaled by happy gaze-shifts, which confirmed that trait anger is related to approach motivation towards potential rewards (Carver & Harmon-Jones, 2009; Harmon-Jones, 2003a, 2004).

Chapter 5 describes an eye-tracking experiment explicitly designed to index how personality characteristics of on the one hand trait anxiety and submissiveness, and on the other hand trait anger and dominance, influence natural gaze behavior in a socially threatening context. The study first confirmed extensive literature showing that anxious individuals tend to remember threatening information better than non-anxious individuals (Mitte, 2008). Next, the study showed that, similar to the above described primate behavior (Mazur, 1985; Mazur & Booth, 1998; Setchell & Wickings, 2005), submissive individuals avoid making eye-contact with, and rapidly avert gaze from, angry faces (Terburg, Aarts, & van Honk, 2012b).

In the experiment described in **Chapter 6** we focused our effort specifically on such gaze-aversion as a mechanism of social dominance and submissiveness. In a paradigm that again made use of reactive eye-tracking, it was shown that dominant individuals are slower to avert gaze from unconsciously presented angry compared to happy faces. In other words, dominant individuals maintain, implicitly and reflexively, eye-contact when their social dominance is challenged (Terburg, Hooiveld, Aarts, Kenemans, & van Honk, 2011).

In the following chapters of this thesis the hormonal mechanisms that underlie social dominance behavior are studied and described. **Chapter 7** starts with an experiment that indexed the relation between the steroid hormone testosterone and the perception of how trustworthy others are. In a double-blind placebo-controlled design the study showed that testosterone administration decreases trust, but only in those participants that are normally highly trusting (Bos, Terburg, & van Honk, 2010). Thus, only people who might be

considered ‘socially naïve’ became less trusting after testosterone administration, which indicates that testosterone adaptively induces social vigilance, which is a vital aspect of dominance and leadership (Coates & Herbert, 2008; Mazur & Booth, 1998).

In **Chapter 8** the relation between testosterone and dominance was tested directly, and we show that this is an unconscious and reflexive mechanism. We applied the same reactive eye-tracking paradigm from **Chapter 6** in a double-blind placebo-controlled study design, and found that after testosterone administration participants were slower to avert gaze from unconsciously presented angry faces. Thus, testosterone works on an automatic and reflexive mechanism that promotes social dominance behavior (Terburg, Aarts, & van Honk, 2012c).

These two studies indeed provide evidence for a causal link between testosterone and dominance, and in **Chapter 9** mediating factors are discussed in relation to sex-differences in aggression. Additional to the earlier described role of cortisol in the relation between testosterone and aggression, it is now also argued that prenatal testosterone levels are an important mediator in the behavioral effects of testosterone (Terburg, Peper, Morgan, & van Honk, 2009b). Prenatal testosterone has been shown to influence brain development (Peper et al., 2009a; Peper et al., 2009b), and predict physical aggression in men (Bailey & Hurd, 2005). Furthermore, prenatal testosterone levels also influence the ratio in length of the second and fourth digits of the right hand (digit-ratio, or 2D:4D), which provides for a method to index prenatal testosterone levels of now adult individuals (Breedlove, 2010; Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004; Millet & Dewitte, 2006). Using this method in a testosterone administration study it has been shown that testosterone causally impaired cognitive empathy, but only in women that were prenatally exposed to high levels of testosterone (van Honk et al., 2011a). Using a similar method we show in **Chapter 10** that testosterone can promote social cooperative behavior, but this time only in women that were prenatally exposed to low testosterone levels.

Additional to the effects of prenatal testosterone exposure, the last study confirmed earlier evidence on a positive contribution of testosterone to social cooperative behavior (Eisenegger et al., 2010). This seems rather contradictory given all the earlier evidence that testosterone promotes social dominance and aggression. Eisenegger and colleagues however convincingly argued that social cooperation can also spawn from a desire to achieve social status (Eisenegger, Haushofer, & Fehr, 2011; Eisenegger et al., 2010). In **Chapter 11** we

adopted this argument, and provide a neural model that not only explains the effects of testosterone on dominant aggression, but also its effects on social cooperation, by linking them both to social status (van Honk, Terburg, & Bos, 2011b). We followed thereby the observation that, depending on the social context, high testosterone men can be extremely aggressive, but also extremely cooperative or even altruistic (Dabbs & Dabbs, 2000). The difference in social context we define as either a situation with no direct status threat, and a situation with immediate or direct threat to status. In the low threat context no immediate subcortical threat-vigilance is required and testosterone's effects on the decoupling of the OFC from the amygdala has the overhand. Consequently, the amygdala is not inhibited by the OFC, and contributes to the brain's 'safe-guarding' mode. The amygdala can therefore mediate other cortical structures in order to defend or increase social status, which can lead to non-social tendencies like reduced cognitive empathy (van Honk et al., 2011a), but can also induce social cooperation (Eisenegger et al., 2010; van Honk, Montoya, Bos, van Vugt, & Terburg, in press). On the other hand, when there is a direct threat for social status, subcortical threat-vigilance plays a profound role in the effects of testosterone. Testosterone upregulates gene expression of vasopressin in the amygdala, which in turn activates the hypothalamus and brainstem resulting in vigilant responding to the threat. This pathway underlies the reactive aggression related to testosterone, and most likely promotes the reflexive inhibition of gaze-aversion we described in **Chapter 6** (Terburg et al., 2012c).

In the concluding **Chapter 12** of this thesis the neural framework for threat-vigilance from **Chapter 3** will be combined with the neural framework of testosterone and dominance from **Chapter 11**, whereby the latter is interpreted from a heterogeneous perspective on the amygdala. Important in this respect is that the upregulation of vasopressin by testosterone has in animal research most profoundly been observed in the CMA, while the anxiolytic functions of the BLA-CMA pathway are independent from vasopressin. It will therefore be argued that testosterone's upregulation of vasopressin in the CMA is primarily involved in aggressive vigilance, testosterone directly inhibits the HPA stress-axis, and testosterone might upregulate the anxiolytic functions of the BLA. Thus, testosterone upregulates the amygdala, but this underlies increased motives for dominance and social aggression, while acute fear responses are reduced. In this concluding chapter we thus provide a neural framework that explains how testosterone directs the brain to always promote social status. In short, testosterone contributes

to an increase in social status by promoting either social aggression or social cooperation, depending on the social context.

Chapter 2

The testosterone-cortisol ratio: A hormonal marker for proneness to social aggression

Based on: Terburg, D., Morgan, B., & van Honk, J. (2009)
International Journal of Law and Psychiatry, 32, 216-223

Abstract

Social aggression is an escalating hazard for individuals and society. It is most frequently observed as impulsive-reactive aggression in antisocial personality disorder (APD), but in psychopathic aggressive personalities instrumental social aggression is more prominent. However, the psychobiological mechanisms underlying human social aggression are still poorly understood. Here we propose a psychobiological mechanism that may explain human social aggression wherein the steroid hormones cortisol and testosterone play a critical role. High levels of testosterone and low levels of cortisol have been associated with social aggression in several species but it seems that in those individuals wherein these hormonal markers combine social aggression is most violent. In this review we discuss fundamental and clinical research which underscores the potential of the testosterone-cortisol ratio as a possible marker for criminal aggressive tendencies.

"Stop being so testosteroney!"

Introduction

"Stop being so testosteroney", 'the guys' are being told by one of 'the girls' in the popular sitcom 'Friends' after they agree to, instead of calling a girl following a successful date, "let her dangle" and wait for her to call (Abrams, 1995). This joke works, because it relies on the common sense view that typical male behavior is linked to the steroid hormone testosterone. A more specific common sense view is the widely accepted relation between testosterone and violent or thrill-seeking behavior. For example, the popular television channel MTV describes 'The Getaway', a movie speckled with chases and violence, as "testosteroney" (Heitmueller, 2006). Of course these kinds of behavior are the result of complex processes that involve not only the hormone testosterone, but many other factors. In this review we concentrate on the effects of the hormones testosterone and cortisol on socially aggressive behavior. We argue that the ratio of the basal levels of these steroid hormones is a marker for proneness to social aggression.

Social behavior promotes the well-being of the individual and the group to which the individual belongs. Social acts of dominance and submissiveness are the core behaviors in the formation of a social hierarchy consisting of dominant and subordinate individuals. Such social hierarchies are observed in most mammalian species and greatly interact with the adrenal and gonadal hormone systems and stress-related mechanisms (Sapolsky, 2005). Unlike in most other mammals, human social structures are extremely diverse and most people are part of several social hierarchies. Family life, work environments or peer friendship-groups all have their own social structure and an individual has to play many different roles, with varying positions on the social staircase in each of these roles. Difficulties in switching between these social structures and adapting behavior to the role in the present social hierarchy can result in undesirable behavior, like social withdrawal or aggression. The hormones testosterone and cortisol seem to modulate these behaviors. High levels of testosterone have been associated with dominant aggressive behavior in both men (Dabbs Jr., Carr, Frady, & Riad, 1995; Dabbs Jr. & Morris, 1990) and women (Dabbs Jr. & Hargrove, 1997; Dabbs Jr., Ruback, Frady, Hopper, & Sgoutas, 1988). Low cortisol levels

have also been linked to aggressive social tendencies (McBurnett et al., 1991; Vanyukov et al., 1993; Virkkunen, 1985), whereas high levels of cortisol have been reported in anxious depression (Bohus, de Kloet, & Veldhuis, 1982; Johnson et al., 1992; Schulkin, 2003a), and seem to be linked to low mood (van Honk et al., 2003a), non-clinical anxiety and submissive behavior (Brown et al., 1996; Sapolsky, 1990).

The biological mechanisms of testosterone and cortisol

To specify the relation between on the one hand testosterone/cortisol and on the other hand social aggression, we will start with the biological mechanisms involved. Cortisol and testosterone are the end products of two hormonal axes, the hypothalamus-pituitary-adrenal (HPA) axis and the hypothalamus-pituitary-gonadal (HPG) axis respectively. **Figure 2.1** represents a simplified framework of these axes. See the review of Johnson and colleagues (1992) for a full representation of the axes and their interconnections.

The HPA axis is activated during a stressful event. Corticotropin-releasing hormone (CRH) and arginine vasopressin (AVP) are released in the hypothalamus. These initiate the production and secretion of adrenocorticotropic hormone (ACTH) in the pituitary gland, which facilitates the production of glucocorticoids, most importantly cortisol, in the adrenal gland. The HPG axis is involved, among other things, in the reproductive and immune systems. Gonadotropin-releasing hormone (GnRH) is secreted in the hypothalamus and transported to the pituitary gland, where it stimulates the production and secretion of luteinizing hormone (LH) and follicle stimulating hormone (FSH). These are transported to the gonads where they induce the production of testosterone (Johnson et al., 1992). As depicted in **Figure 2.1**, testosterone inhibits HPA functioning at the hypothalamic level by decreasing AVP levels (Viau, 2002), while cortisol has inhibitory effects on all three levels of the HPG axis.

The HPA axis is, together with activity of the sympathetic nervous system, part of the fight-or-flight reaction in situations of threat or stress. A major aspect of this stress reaction is preparation for physical activity, instigated by the sympathetic nervous system and maintained by HPA-produced CRH (Johnson et al., 1992). These preparations are referred to as autonomic arousal and consist of, among others, increase of heart-rate, perspiration

(measured with skin-conductance), widening of the pupils, potentiated startle reflexes and down-regulation of the gastrointestinal systems. Its end product cortisol helps to restore homeostasis after a stress-response and high endogenous levels of this hormone are therefore generally seen as a sign of stress and anxiety (Brown et al., 1996). Moreover, both testosterone and cortisol bind to steroid-responsive centers in the amygdala (Wood, 1996), a brain structure centrally involved in emotional processing (LeDoux, 2000), where approaching (e.g. fight) (testosterone) or avoidant (e.g. flight) (cortisol) behavior is facilitated (Schulkin, 2003b).

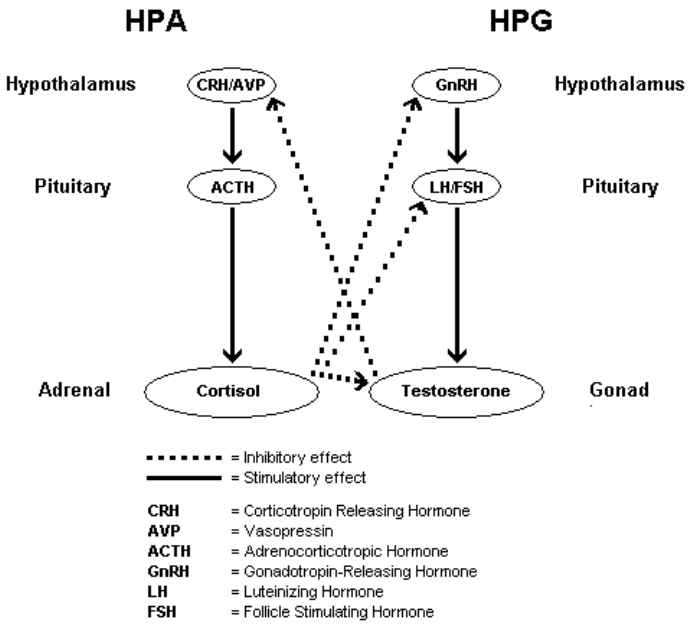


Figure 2.1 The hypothalamic-pituitary-adrenal (HPA) and hypothalamic-pituitary-gonadal (HPG) axes; structures involved, hormonal cascades and their functional interrelations.

The psychological mechanisms of testosterone and cortisol

Testosterone and cortisol are thus part of a biological balance that influences very basic and primary reactions to threat by moderating autonomic responses. On the psychological level this balance is also evident. To explain this we will start with an associative model of consciousness postulated by Ressler (2004). This model states that conscious emotions are cognitive representations of basic motivational drives mediated by punishment and reward. In line with Pavlov's (1927) classical conditioning experiments, stimuli are coupled with rewards or punishments, which results in approaching or avoidant tendencies to these stimuli.

In Pavlov's dogs this can be observed in overt behavior, anticipatory drooling following a sound recently coupled with a juicy steak. In humans, conscious emotions emerge, which help guide our overt behavior. Ressler (2004) further argued that survival depends on well-established approaching and avoidant tendencies. A balance should exist in motivational behavior, avoiding danger and approaching the right goals.

Motivational imbalance could result in psychological disorders. In his motivational imbalance model, Arnett (1997) explains psychopathy in terms of motivational imbalance and autonomic arousal. He used the behavioral inhibition and activation model of Gray (1987) and Fowles (1980) and related it to a great body of psychophysiological data on autonomic arousal, like heart-rate changes and skin-conductance responses. In short, the model defines two mutually inhibitive systems, the behavioral inhibition (BIS) and behavioral activation (BAS) systems. The BIS is punishment driven and results in not acting when the possibility of punishment is present. The BAS represents reward seeking behavior and makes one act towards a possible reward. Arnett (1997) shows that psychopaths have less autonomic arousal (primarily skin-conductance measurements) in a punishment situation, which means that they are less inclined to avoid punishment. Arnett links this to a lower BIS activity. Although based on fewer studies, he also makes a strong argument for more autonomic arousal (heart-rate responses) in rewarding situations in psychopaths, which he explains as a reward dependency and a stronger activation of the BAS.

The balance between behavioral activation and inhibition, BAS and BIS, seems to be the psychological equivalent of the biological balance observed in HPA and HPG activity, or their end-products, cortisol and testosterone. As already stated, the HPA axis is heavily involved in the instigation and maintenance of the fight-or-flight response. High levels of its end-product cortisol can therefore be a sign of high punishment sensitivity, which explains the relation between cortisol and anxious depression (Bohus et al., 1982; Johnson et al., 1992; Schulkin, 2003a) and anxiety (Brown et al., 1996). Testosterone inhibits HPA activity, thus autonomic responses to threat are less strong when testosterone levels are high and punishment sensitivity is reduced. This is further supported by decreased autonomic arousal (heart-rate and skin-conductance measurements) in stressful situations (e.g. preparing and giving a speech) in subjects with APD (Raine, Lencz, Bihrle, LaCasse, & Colletti, 2000). As stated in the BIS-BAS theory of Arnett (1997), low punishment sensitivity, i.e. low BIS, is

related to psychopathy and aggressive behavior. Moreover, testosterone stimulates vasopressin gene expression in the amygdala, which is related to heightened reward sensitivity (DeVries, DeVries, Taymans, & Carter, 1995; Szot & Dorsa, 1994). Likewise, cortisol stimulates CRH gene expression, which is related to punishment sensitivity (Schulkin, 2003b) and the sustainment of a fearful stance (Schulkin, Morgan, & Rosen, 2005). Thus, high testosterone/low cortisol ratios seem to predict approach motivation/reward sensitivity. In these motivational stances, individuals are more likely to confront threat, which could result in aggressive behavior. A high testosterone/cortisol ratio therefore predisposes for socially aggressive behaviors.

Recent findings have established two additional relations between testosterone and aggression. The first is the enhancement of attention to aggressive stimuli (such as aggressive faces) by high levels of testosterone (van Honk et al., 1999). The second is based on the idea that motivational tendencies originating in the emotional brain are mediated by cognitive processes. In humans emotional tendencies such as approach or withdrawal can be curtailed by cognitive processes, making 'response reversal' a possibility (Blair, 2004). High testosterone seems to down-regulate the interaction between cognitive and emotional systems and therefore reduces the impact of cognitive control (Schutter & van Honk, 2004). Aggressive tendencies will not be reversed by top-down processes and will more often result in impulsive aggression. Additionally, cognitive decision-making depends in this situation less on emotional bottom-up processing. When the prevailing stance is reward-sensitivity and approach-motivated, more instrumental (thought rather than feeling driven) forms of aggressive behavior can result. Thus, according to these recent findings, high levels of testosterone seem to enhance attention to aggressive stimuli and down-regulate interaction between cognitive and emotional brain systems, which both increases the possibility (attention) and probability (less cognitive control on impulses and more cognition-based decision making) of confrontation and aggressive behavior.

The triple balance model of emotion

The effects of testosterone on attention and on communication between cognitive and emotional brain systems were integrated in a triple balance model of emotion proposed by van

Honk and Schutter (2006). This model combines four existing models of psychopathy and aggression and extends these with neurobiological findings on testosterone and brain activity. The first model that was used is the low fear model of Lykken (1957), which states that psychopaths show less passive avoidance, which is a result of poor aversive conditioning. In normal life this manifests as less fear and thus less respect for social boundaries.
The second model is the somatic-marker hypothesis of Damasio (1994). Damasio's hypothesis is a model of emotional learning, which occurs after (un)conscious valenced sensations (somatic markers) are coupled with stimuli or situations. This combined with the third model used by van Honk and Schutter (2006), i.e. the previously described motivational imbalance (BIS-BAS) model of psychopathy by Arnett (1997), provides a framework of defective emotional learning in psychopaths. Motivational approach or withdrawal tendencies can be seen as basic somatic markers. When these are out of balance under influence of cortisol and testosterone, emotional learning will be impaired.

The last model integrated in the triple balance model is the violence inhibition mechanism of Blair (1995, 2003b). Blair explains that facial expressions of fear and sadness are cues to activate a violence inhibition mechanism and make the individual stop actions that might inflict harm to the other. In psychopaths this mechanism seems to be defective. Blair (1995, 2003a, 2003b) argues convincingly that both the recognition of submissive facial expressions by the emotional brain and response aversion in the cognitive brain, especially through interaction between the amygdala and orbitofrontal cortex (Blair, 2004), are impaired. Thus, psychopaths fail to recognize submissiveness and fail to respond by stopping harmful actions like aggressiveness and therefore show inappropriate social behavior.

Integrating these four models, the triple balance model of emotion (van Honk & Schutter, 2006) distinguishes three imbalances of emotional processing in psychopaths. All three of these imbalances are mediated by testosterone and cortisol. Essentially it is a brain model with two levels, the sub-cortical level for basic motivational processing and the cortical level for cognitive processing and conscious emotion perception. Both levels as well as communication between both levels can be imbalanced. Firstly, the sub-cortical level can be imbalanced, which results in an imbalance in reward and punishment sensitivity. This is mediated by the balance between the HPA (cortisol) and HPG (testosterone) axes. Secondly, the communication between sub-cortical and cortical areas can be imbalanced. The model

states that communication between sub-cortical and cortical areas, the amygdala and orbito-medial prefrontal cortex (OMPFC) respectively, is needed for the control of motivational tendencies and to provide an emotional basis for cognitive decision-making. The strength of this communication is also mediated by cortisol (more communication) and testosterone (less communication). And thirdly, the cortical level can be imbalanced. More right-sided activity in the prefrontal cortex (PFC) is associated with more fearful behavior and higher levels of cortisol (Tops et al., 2005) and more left-sided activity with approach motivation and anger (Harmon-Jones, 2003b).

A distinction can be made between instrumental, pre-meditated aggression and reactive, impulsive aggression. The first typifies aggressive behavior in primary psychopaths and is relatively rare while the second is typical of individuals with secondary psychopathy and APD and is much more common (Blair, 2004). Van Honk and Schutter (2006) use their triple balance model of emotion to explain psychopathy, but here we will argue that consideration of the testosterone/cortisol ratio renders the model applicable to both forms of aggression. According to the triple balance model, a high testosterone/cortisol ratio enhances sensitivity to reward relative to punishment. These equivalents to the somatic markers of Damasio (1994) ensure that approach motivations will prevail over avoidance reactions, or, seen from Arnett's motivational imbalance model; more BAS than BIS activity (Arnett, 1997). Then, communication between emotional (sub-cortical, amygdala) and cognitive (cortical, OMPFC) systems is diminished. The emotional information from the amygdala cannot adequately reach the cortex, and therefore cannot be used to avert the basic motivational, reward driven reactions. Thus, the violence inhibition mechanism of Blair (1995) doesn't function properly. And finally, the low levels of cortisol reduce right-sided dominance of PFC activation and, integrating Lykken's (1957) low fear model, less fear is learned and experienced. Thus, a high testosterone/cortisol ratio results in potentiated approach motivational and hence behavioural tendencies towards threat.

In particular the loss of communication between sub-cortical and cortical systems can lead to both previously described forms of aggression. Emotional information can't reach the cortex anymore and the cortex can't control sub-cortical motivational tendencies. Cognitive decision making will be based on less emotional information and could be cold and without empathy. This is a strong foundation for instrumental, pre-meditated aggression. Aggressive

urges from the sub-cortical motivation system, on the other hand, can't be reversed by cognitive control and will be executed more often resulting in reactive, impulsive aggression.

A key factor distinguishing these two forms of aggression is the degree of impulsivity present. Impulsivity is mediated by serotonin (Siever, 2008), with low serotonergic transmission shifting the balance to more impulsive behavior. In the context of our model low serotonergic transmission will facilitate bottom-up aggressive signaling from the sub-cortex so that in case of provocation, reactive aggression becomes more likely. With high serotonergic transmission, behavior is less impulsive (i.e. more cortical inhibition). As described in our model a high testosterone/cortisol ratio results in less communication between sub-cortical and cortical areas and a lateralization of prefrontal activity to the left. The resulting relative lack of emotional input from the sub-cortex, cortical imbalance towards approach motivation and anger and the heightened top-down inhibition of impulsive behavior add up to a higher probability of pre-meditated aggression. However, more research has to be done on this topic.

Evidence for the triple balance theory of emotion in psychopathy can be found in the relations between testosterone and low punishment sensitivity and reductions in fear (Boissy & Bouissou, 1994) and enhancement of reward sensitivity (Carr, Fibiger, & Phillips, 1989; van Honk et al., 2004a), as well as the relation of high levels of cortisol with fear (Rosen & Schulkin, 1998; Schulkin, 2003a; Schulkin et al., 2005) and punishment sensitivity (van Honk, Schutter, Hermans, & Putman, 2003b). However, more specific evidence on the separate balances can be provided by looking at research conducted with facial expressions.

Attention, facial expressions and social aggression

Facial expressions play an important role in human social behavior. On both conscious and subconscious levels they are able to give and receive signals of emotional state and information about the surroundings. Emotional facial expressions are used in all cultures and some of these expressions are found to be universal (Biehl et al., 1997; Darwin, 1872; Ekman, 1999a, 1999b). Both anger and fearful facial expressions are social signals that relate to threat. Either the one who expresses the emotion or perceives it is threatened and has to decide whether to face the threat or to avoid it.

The emotional Stroop task is a task that measures attentional biases using the Stroop paradigm (Stroop, 1935). An emotional stimulus, in this case a facial expression, is presented and the subject has to name the color. The affective value of the stimulus interferes with this and therefore the color-naming will be slowed-down. Findings with the emotional Stroop task suggest that, especially on a subliminal or implicit level (masked stimuli) (Toates, 2006), angry faces interfere more with attention in subjects inclined to approaching behavior than in avoidant subjects (Putman, Hermans, & van Honk, 2004; van Honk, Tuiten, de Haan, van den Hout, & Stam, 2001a).

A similar relation was found with testosterone. In both men and women a positive correlation between basal testosterone levels, as measured in saliva, and attentional interference by angry facial expressions, was found (van Honk et al., 1999). Additionally, after testosterone administration to women, an increased heart-rate response (**Figure 2.2**) to passive viewing of angry and not neutral or happy faces was observed (van Honk et al., 2001a). Thus, testosterone seems to help direct attention to angry facial expressions and promotes fight or flight autonomic responses such as acceleration of the heartbeat (Lang, Bradley, & Cuthbert, 1998) to threat.

The effect of testosterone on attention to fearful faces is reversed. The normally observed attentional interference by fearful facial expressions disappears after testosterone administration to women (van Honk et al., 2005) and the fear-related startle reflex (Davis, Gendelman, Tischler, & Gendelman, 1982) was found to be diminished after testosterone administration (Hermans et al., 2006a). For avoidant behavior and cortisol, on the other hand, the relation with attention to angry facial expressions is different. Social anxiety and high basal cortisol levels (measured in saliva), are related to attention away from angry faces (Putman et al., 2004; van Honk et al., 1998).

Van Honk and Schutter (2007b) explain these results in terms of dominance and subordination. Through interaction of dominant and subordinate individuals, social hierarchies are established and, as stated earlier, social hierarchy and levels of testosterone and cortisol interact (Sapolsky, 2005). Under influence of high testosterone levels, attention to threat and reward sensitivity are enhanced and an angry face is perceived as a challenge (Archer, 2006) and will be confronted, an act of dominance. High cortisol results in diminished attention to threat and punishment sensitivity is enhanced, angry faces will be

avoided, which is an act of subordination (van Honk & Schutter, 2007b). In this view it is easy to conceive that, also with their mutual inhibitory qualities in mind, combined high testosterone and low cortisol levels can result in an extremely dominant stance and predispose towards confrontation and in the end, aggression.

The behaviors just described are all mediated by attentional and subconscious mechanisms. Cognition and conscious emotional processing, however, obviously play a role in social behavior too. As previously described in the triple balance theory of emotion (van Honk & Schutter, 2006), sub-cortical structures like the amygdala communicate with cortical structures. Through this communication, emotions become consciously available and behavior can be cognitively controlled (Blair, 2004). Through this cognitive control, domineering and aggressive tendencies developed in the sub-cortex can be mediated and reversed by top-down processes. Additionally cognitive decision-making can be influenced by sub-cortical emotional (bottom-up) processing. Again, this communication is mediated by testosterone and cortisol. These are processes that have been studied mostly in recent years.

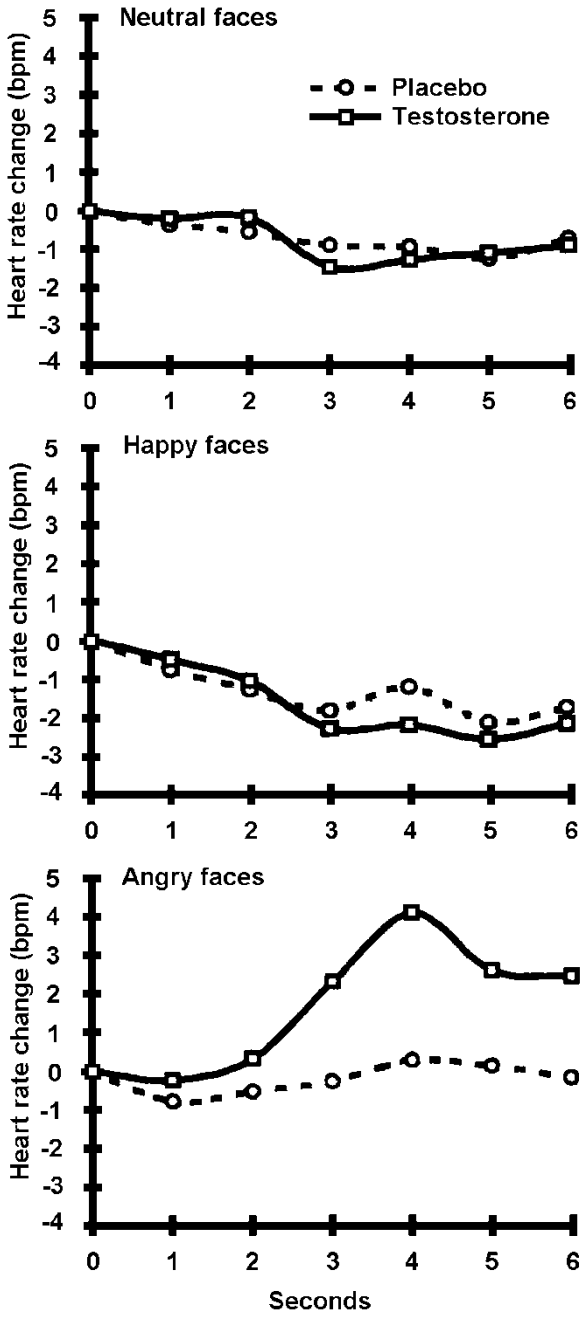


Figure 2.2 Cardiac response: Mean heart rate changes in beats per minute (bpm) after passive viewing of neutral, happy and angry faces for testosterone administration and placebo conditions. From baseline (1s pre-stimulus) to 6s post-stimulus. Heart rate increases only after viewing angry faces in the testosterone condition. Taken from: van Honk and colleagues (2001a).

Cortical emotion, facial expressions and social aggression

Conscious perception of facial expressions can be measured with an emotional recognition task developed by Montagne (Montagne, 2005). This task consists of moving facial expressions of variable affective intensities (20-100%). A neutral faces is presented, which morphs gradually (between 0.5 and 2 seconds) into an emotional facial expression. In the first round the morphing stops at 20% emotion, in the second round at 30% and so on until the ninth round where the morphing process is full and the emotion is presented in a full-blown facial expression. Subjects must name the perceived emotion, which is a cognitive process that relies on emotional and sub-cortical processes. With this task the ability in recognition of emotional facial expressions can be examined. Montagne and colleagues (2005) showed with this task that healthy subjects, with a low BIS/BAS ratio which is a sign for psychopathic personality characteristics, performed relatively poor on recognizing the emotion fear. Recently van Honk and Schutter (2007a) used this task in a testosterone administration study. They showed that after the testosterone levels of healthy young women were elevated to approximately the level of men, their performance on recognizing emotions of threat, disgust, fear and especially anger (see **Figure 2.3**), decreased. Thus, while the attention is directed to threatening facial expressions, the conscious recognition of these expressions is impaired after testosterone administration. In summary, under the influence of testosterone the prevailing stances are dominance and approach motivation, and dominance (anger) and submission (fear) in others are not recognized while attention is nonetheless directed to these very cues. This makes aggressive behavior very likely in provocative situations.

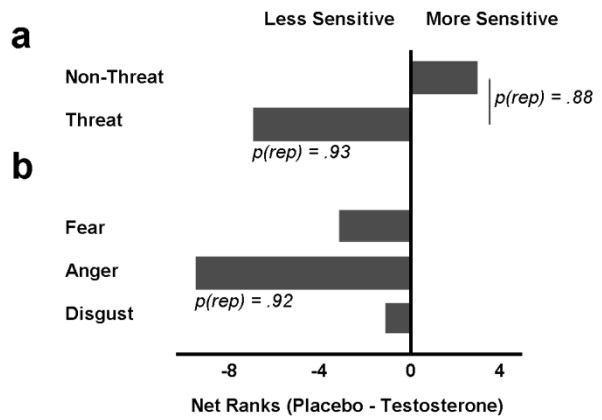


Figure 2.3 Emotion recognition: Net rank scores for placebo minus testosterone conditions. After administration of testosterone, recognition of threat emotions is impaired (upper half), which is most pronounced on the emotion anger (lower half). Taken from: van Honk and Schutter (2007a).

A possible cause for the impairment in emotion recognition is a decrease in communication between the sub-cortical and cortical emotional systems under influence of high testosterone levels. According to this view, the affective value of a perceived facial expression can't reach the cognitive systems, which impairs the conscious recognition of the expression (van Honk & Schutter, 2006). Evidence for this hypothesis is provided with an EEG/testosterone administration study by Schutter and van Honk (2004). According to an evolutionary perspective on behavioral inhibition (BIS) postulated by Knyazev and Slobodskaya (2003), the human emotional brain can be separated into three evolutionary stages. The oldest part is the reptilian brain, a sub-cortical system which promotes fight-flight motivation. The second oldest is the lower mammalian brain, a cortical system that is able to modulate and reverse motivational tendencies. And the evolutionary newest part is the human neo-cortex that is crucial in decision making and as such can modulate all underlying basic motivational behavior. All three systems communicate with each other and each system is able to inhibit the evolutionary older systems. Activity in each system is represented in EEG measurements by different oscillation bands.

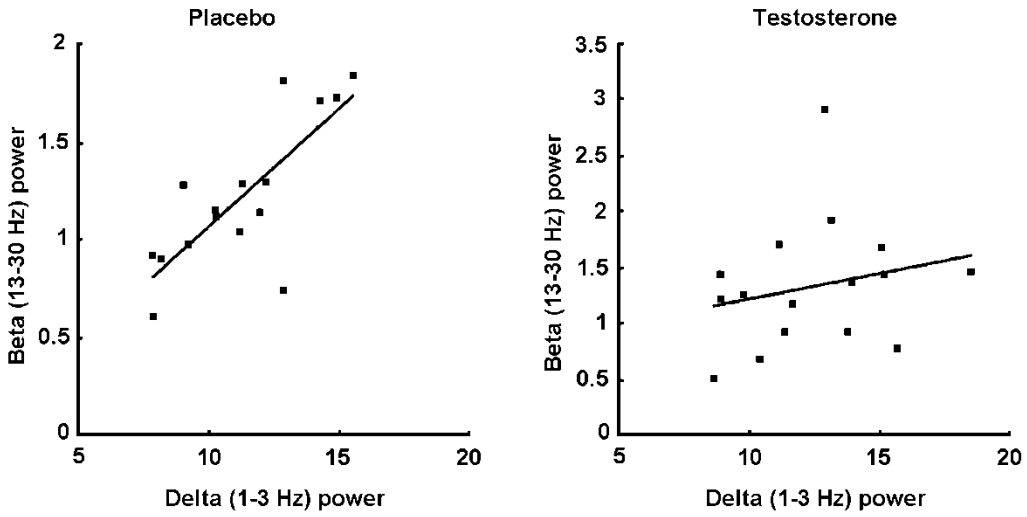


Figure 2.4 Decoupling of subcortical-cortical communication: Significant loss of midfrontal delta-beta coupling after testosterone compared to placebo administration in healthy human volunteers. Taken from: Schutter and van Honk (2004).

Schutter and van Honk (2004) showed that administration of testosterone reduced interaction between the oldest sub-cortical (fight-flight) system and the neo-cortex (decision making). After testosterone administration increased delta oscillatory activity was found. The EEG delta oscillation band is generally assumed to be related to sub-cortical activity, which is the locus of motivational processing such as fight-flight behavior. The testosterone induced increase in delta-activity constitutes causal evidence for testosterone-mediated enhancement of subcortical motivational-attentional circuits as described above. Moreover, as can be seen in **Figure 2.4**, the correlation of this delta-activity with activity in the beta oscillation band, which represents higher order decision-making in the neo-cortex, was completely abolished after testosterone administration (Schutter & van Honk, 2004). This lack of synchronization of delta- and beta-activity can be interpreted as a loss of communication between the motivational (delta) and cognitive (beta) systems. In short, testosterone seems to induce a loss of cognitive control, and thus a loss of behavioral inhibition (BIS), on motivational tendencies originating in the sub-cortex.

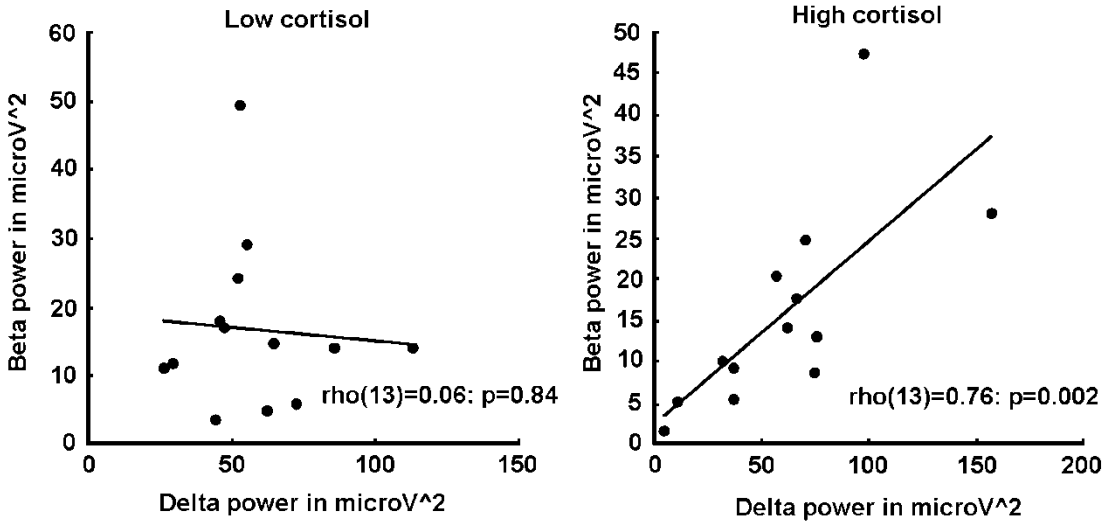


Figure 2.5 Coupling of subcortical-cortical communication: Significant midfrontal delta-beta coupling in the high cortisol group (basal levels), whereas the low cortisol group displays midfrontal delta-beta decoupling. Taken from: Schutter and van Honk (2005).

In line with these findings is a similar study linking delta and beta EEG oscillations to basal cortisol levels. Schutter and van Honk (2005) found that in subjects with high cortisol,

sub-cortical (delta) and cortical (beta) activity correlated highly, but this correlation was completely absent in the low cortisol group (see **Figure 2.5**). Using the same intepretation as the previously described study, high cortisol seems to enhance cognitive control and thus higher behavioral inhibitory activity. So, again, cortisol reverses the effects of testosterone.

The orbitofrontal cortex is generally seen as the cortical structure that mediates sub-cortical reward and punishment tendencies (Blair, 2004). Patients with lesions in this area often show aggressive and impulsive behavior (Blair & Cipolotti, 2000) and reductions in grey matter of the orbitofrontal cortex have been found in APD patients (Damasio, 2000; Raine et al., 2000). A recent study by Hermans, Ramsey and van Honk (2008) correlates basal levels of testosterone and cortisol, and testosterone administration with neuronal activity (fMRI) during passive viewing of angry versus happy faces. Firstly, they found activity in the sub-cortical areas amygdala, hypothalamus and brainstem and the orbitofrontal cortex (OFC) while viewing angry faces. Secondly, as shown in **Figure 2.6**, the observed activation in the amygdala, hypothalamus and brainstem was positively correlated with the basal testosteron/cortisol ratio. And thirdly, especially after testosterone administration activity in the amygdala and hypothalamus increased; this was also observed in the OFC, but not that strongly. In short, these findings suggest that the sub-cortical amygdala, hypothalamus and brainstem work together as a neural circuit during perception of angry facial expressions and process this information to the OFC. When the testosterone/cortisol ratio is relatively high, or when testosterone levels are increased after oral administration, angry facial expressions activate this circuit even more. The communication with the OFC however, is disturbed and information is not transferred to the OFC as efficiently as it would be with relatively normal endogenous testosterone levels.

Taking in account that testosterone diminishes the coupling of cortical and sub-cortical communication (Schutter & van Honk, 2004), it can be argued that, while activation of the neural circuit for perception of threat is stronger under influence of high testosterone and low cortisol, communication between OFC and sub-cortical areas is less efficient under these circumstances and behavioral tendencies are less cortically modulated. Thus, under high testosterone/low cortisol conditions, angry facial expressions are more attended to and confrontation motivation is facilitated, while these tendencies are less well modulated by higher cognitive processes.

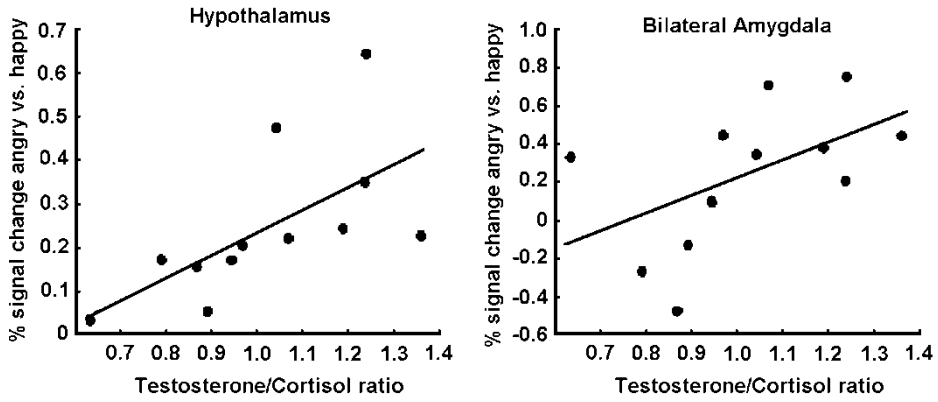


Figure 2.6 Correlations of the testosterone/cortisol ratio with brain activity: The activity found in the amygdalae and hypothalamus during passive viewing of angry faces (happy faces as control condition), correlated with the basal testosterone/cortisol levels. Taken from: Hermans, Ramsey and van Honk (2008).

Finally, according to a model of Harmon-Jones (2003b), approach-related emotion is associated with left pre-frontal activity and withdrawal-related emotions with right pre-frontal activity. Right sided dominance in the PFC is repeatedly associated with fearfulness and high levels of cortisol (Buss et al., 2003; Tops et al., 2005) and left sided dominance with approaching and aggressive behavior (Harmon-Jones, 2003b). Testosterone hasn't been related to heightened activity in the PFC, probably because testosterone diminishes input from sub-cortical affective areas and in that way down-regulates the right pre-frontal asymmetry and with that, the conscious experience of the emotion fear. Again, this is a balance influenced by testosterone and cortisol, with high testosterone and low cortisol resulting in low fear and approach related behavior, both precursors for aggressive behavior. When, under influence of a high testosterone/cortisol ratio, cognitive decision making is void of fear and less influenced by emotional information, due to the decoupling of sub-cortical and cortical structures. Decisions will be based on cold reasoning and steered by approaching and reward-seeking tendencies. Aggression becomes more and more prevalent and will be instrumental and premeditated. When, in the same situation, the motivational tendencies are in charge and these are not inhibited by cognitive control, aggression will be reactive and impulsive.

Conclusions

In this article we have tried to provide evidence for the statement that the testosterone/cortisol ratio is a hormonal marker for aggressive behavior. Following the triple balance model of emotion (van Honk & Schutter, 2006) evidence on three levels of emotion processing was reviewed and related to the testosterone/cortisol ratio and social aggression. We have illustrated that the biological balance of the HPG and HPA axes and their end products testosterone and cortisol, is also observed in a psychological balance of behavioral activation and inhibition (BIS/BAS) and approach/avoidance motivation. These psychological balances were linked to the hypothalamus and the amygdala and the role of the amygdala in attention to threat (first level in the triple balance model of emotion). A neural circuit for perception of angry facial expressions was revealed consisting of the sub-cortical amygdala, hypothalamus and brainstem, which projects to the orbitofrontal cortex. Communication between these sub-cortical and cortical areas (second balance) was shown to be necessary for conscious emotion perception and cognitive response reversal of approach/avoidance motivation. This second balance is also mediated by testosterone (less communication, less accurate recognition of threat emotions, less inhibition of the approaching stance) and cortisol (more communication, better threat recognition and more cognitive control of motivations). Finally, a high testosterone/cortisol ratio results in less consciously experienced fear on the cortical level (third balance), which was associated with diminished right sided activation in the prefrontal cortex.

On the behavioral level these imbalances, caused by a high testosterone/cortisol ratio, result in more attention to anger or cues that can evoke confrontation, more motivational tendencies to actually confront these threats, less reversal of aggressive tendencies by the cognitive system, less accurate conscious perception of dominant and dismissive facial expressions, and less experienced fear. This all combined leaves an individual with a dominant personality, incapable of empathizing with other individuals, with low fear and anxiety and prone to aggressive behavior. In normal social conduct these individuals will not have respect for the social hierarchies that exist around them, which will repeatedly result in conflict and dominance-submissiveness confrontations. They easily break the rules and use violence to fulfill their goals.

It must be noted however, that the triple balance model of emotion (van Honk & Schutter, 2006) was developed as a model explicitly for explanation of psychopathy and not for APD or social aggression per se. The authors follow Blair (2004) in the distinction of reactive and instrumental aggression, where reactive aggression is more impulsive and stress related and instrumental aggression is premeditated and cold, without emotion. Reactive aggression is observed in APD and instrumental aggression in psychopathy. This distinction, however, does not influence the assumption that the testosterone/cortisol ratio is related to aggression in general, because both forms of aggression can be explained by the triple balance model of emotion and the testosterone/cortisol ratio. In the distinction of these two forms of aggression other mechanisms possibly play a role. Van Honk and Schutter (2006) themselves point to the neurotransmitter serotonin as a possible candidate to modulate the impulsiveness and premeditation in aggressive behavior. This remains to be studied more extensively.

For now the relation between the testosterone/cortisol ratio and aggressive behavior seems evident. A recent study of Popma and colleagues (2007) showed that overt aggressive behavior of boys in a delinquency diversion program was only positively correlated with testosterone in a group with low cortisol levels. A heightened level of testosterone alone is not enough to induce violence, punishment sensitivity and fear are still of influence and behavior is inhibited. When high testosterone is combined with low cortisol however, there is nothing to stop the aggression and one could become socially aggressive and a danger to society. This knowledge could be of help in the therapeutic treatment of APD and psychopathy. Since their neuro-endocrine balance of testosterone and cortisol seems to be disturbed, their behavior could probably be adjusted by manipulating and maintaining this balance.

Chapter 3

Hyper-vigilance for fear after basolateral amygdala damage in humans

Based on: Terburg, D., Morgan, B. E., Montoya, E. R., Hooge, I. T., Thornton, H. B., Hariri, A. R., Panksepp, J., Stein, D. J. & van Honk, J.

Translational Psychiatry, in press

Abstract

Recent rodent research has shown that the basolateral amygdala (BLA) inhibits unconditioned, or innate, fear. It is however unknown whether the BLA acts in similar ways in humans. In a group of five subjects with a rare genetic syndrome, i.e. Urbach-Wiethe disease (UWD), we used a combination of structural and functional neuroimaging, and established focal, bilateral BLA damage, while other amygdala sub-regions are functionally intact. We tested the translational hypothesis that these BLA-damaged UWD-subjects are hyper-vigilant to facial expressions of fear, which are prototypical innate threat cues in humans. Our data indeed repeatedly confirm fear hyper-vigilance in these UWD-subjects. They show hyper-vigilant responses to unconsciously presented fearful faces in a modified Stroop task. They attend longer to the eyes of dynamically-displayed fearful faces in an eye-tracked emotion recognition task, and in that task recognize facial fear significantly better than control-subjects. These findings provide the first direct evidence in humans in support of an inhibitory function of the BLA on the brain's threat vigilance system, which has important implications for the understanding of the amygdala's role in the disorders of fear and anxiety.

Introduction

The human amygdala is critically involved in social and emotional behavior, and plays a vital role in the assessment of, and responding to, threat (Pessoa & Adolphs, 2010). The amygdala is however not a homogeneous brain-region, but consists of several sub-nuclei that are so different in structure and connectivity that they are best regarded separately (Davis & Whalen, 2001; Heimer et al., 1997; McNaughton & Corr, 2004). Direct evidence into the role of amygdala sub-regions in threat processing comes however predominantly from rodent lesion research (Barbas, Saha, Rempel-Clower, & Ghashghaei, 2003; Phelps & LeDoux, 2005). Although human cases of damage to the basolateral amygdala (BLA) (Hurlemann et al., 2009; Hurlemann et al., 2007), and more extended or complete amygdala damage have been described (Adolphs, 2008; Adolphs, Tranel, Damasio, & Damasio, 1994, 1995; Brand, Grabenhorst, Starcke, Vandekerckhove, & Markowitsch, 2007; Cahill, Babinsky, Markowitsch, & McGaugh, 1995; Hurlemann et al., 2010; Siebert, Markowitsch, & Bartel, 2003), there is to our knowledge no causal evidence pertaining to amygdala sub-region function in human fear processing.

The human amygdala is thought to promote fast and efficient responding to threat within a brainstem-amygdala-cortical alarm system (e.g. Liddell et al., 2005), and rodent studies suggest that the BLA has been especially implicated in such threat processing. In rodents the BLA is essential for the acquisition and extinction of conditioned fear (Amaral, 2003; Parkes & Westbrook, 2010; Phelps & LeDoux, 2005), but there is increasing evidence that the rodent BLA also inhibits unconditioned and acute fear responses (Macedo, Cuadra, Molina, & Brandao, 2005; Macedo, Martinez, Albrechet-Souza, Molina, & Brandao, 2007; Macedo, Martinez, & Brandao, 2006; Martinez, Ribeiro de Oliveira, & Brandao, 2007; Tye et al., 2011). Unconditioned fear reveals itself as acute fear or panic in humans (Graeff & Del-Ben, 2008; McNaughton & Corr, 2004), and the periaqueductal gray (PAG) is critically involved in the instigation of such fear and panic responses (McNaughton & Corr, 2004). Human neuroimaging data show that both the PAG and the central-medial amygdala (CMA) are activated when threats are imminent and unavoidable. When a threat can be avoided, however, activation shifts to the BLA and prefrontal cortex (PFC), which is thought to underlie threat estimation and response inhibition (Maren, 2007; Mobbs et al., 2009; Mobbs et al., 2007). It could be argued that the inhibitory role of the BLA on responsiveness to innate

threat cues in rodents (Macedo et al., 2005; Macedo et al., 2007; Macedo et al., 2006; Martinez et al., 2007; Tye et al., 2011) is responsible for this switch in activity from acute fear responding in the midbrain to cortical threat estimation, but it is unknown whether this evidence from rodents can be translated to humans. Nonetheless, it has been shown numerous times that the amygdala indeed responds to facial expressions of fear (Costafreda, Brammer, David, & Fu, 2008), which are the prototypical innate threat cues for humans, whereby BLA activity has been specifically linked to unconscious processing of facial fear (Etkin et al., 2004). Research addressing the question whether in congruence with rodent research the human BLA inhibits the fear response to such innate threat cues, is however lacking.

In the present study we tested a group of five women with Urbach-Wiethe disease (UWD). UWD is a rare genetic-developmental disorder characterized by focal calcifications in the bilateral amygdalae, which provides a unique window onto human amygdala function (Adolphs, 2007). Early evidence from a UWD-subject with full amygdala damage suggested a specific role for the amygdala in the recognition of static displays of facial fear (Adolphs et al., 1994, 1995), which later was shown to stem from an inability to automatically maneuver visual attention from the mouth to the emotionally critical eye-region of static faces (Adolphs et al., 2005). Indeed, neuroimaging data show that the amygdala is active when gaze is shifted from mouth to eyes (Gamer & Buchel, 2009), and is triggered specifically by fearful eyes (Whalen et al., 2004), but fear processing findings in UWD have been inconsistent (Siebert et al., 2003; Thornton et al., 2008), possibly reflecting heterogeneity in size, location, and epileptogenicity of the amygdala lesions (Hurlemann et al., 2007).

First we will show that the selective bilateral calcifications in the brains of our five UWD-subjects are limited to the BLA. We used high-resolution structural MRI to assess the relative location and extent of calcified damage, and functional MRI to assess the reactivity of the intact amygdala sub-regions. Next, we tested the crucial hypothesis that these UWD-subjects are hyper-vigilant for subliminal fear, which would support the hypothesis that the human BLA plays a role in the inhibition of acute responding to innate threat cues. UWD-subjects and a carefully matched group of healthy volunteers performed in a modified emotional Stroop paradigm that directly taps into threat-driven attentional processing (Algom, Chajut, & Lev, 2004), and can validly assess threat hyper-vigilance by using subliminally presented fearful faces as stimuli (van Honk et al., 2005; van Honk, Schutter, d'Alfonso,

Kessels, & de Haan, 2002). Finally, we assessed the ability of UWD-subjects and controls in emotion recognition of ecologically valid dynamic expressions and measured their eye-movements to assess mechanisms of visual attention.

Materials and methods

Ethics Statement

This study was approved by the Health Sciences Faculty Human Research Ethics Committee of the University of Cape Town. All participants provided written informed consent.

Participants

We tested five women, without any (history of) secondary psychopathology or epileptic insults, from a previously described UWD-cohort in South Africa (Thornton et al., 2008), where this genetic disorder is most prevalent (Siebert et al., 2003). UWD-subjects were compared against a group of healthy volunteers (n = 16) matched for gender, age, and IQ, and living in the same area of South Africa; i.e. mountain-desert villages near the Namibian border. Twelve of these participants took part in the subliminal fear-vigilance task and the dynamic emotion recognition task. Eight of them, and an additional 4 healthy volunteers, took part in the static emotion rating task that was conducted approximately two years later. Demographic data for both groups and time-points are summarized in **Tables 3.1 & 3.2**. Statistics are two-tailed non-parametric Mann-Whitney U-tests with α = .05, and effect-size (r) for significant effects, throughout the behavioral data analysis.

Neuropsychological assessment

Neuropsychological assessment of this group of South African UWD and healthy control research participants was first performed in Cape Town in May 2007. All the participants live in the remote Northern Cape mountain-desert area of Namaqualand. For many of them, coming to Cape Town for MRI scanning and neuropsychological testing was their first journey outside of Namaqualand. Namaqualand is an economically impoverished region where the quality of school education is far below Western norms. It was therefore not surprising to find that this group did not perform well on the Wechsler Adult Intelligence

Scale (WAIS-III) (Wechsler, 1997), which was developed in a First World setting according to Western cultural and educational norms. The Wechsler scale purports to measure "the global capacity of a person to act purposefully, to think rationally, and to deal effectively with his environment" (Wechsler, 1997). As can be seen in the **Tables 3.3** most of the participants in our study (in total 5 UWD-subjects and 16 control-subjects) hold jobs in a region where unemployment exceeds 70%. The problems inherent in using the WAIS-III in a transcultural setting are made starkly apparent by the fact that in May 2007 several of these participants scored in the borderline range.

This contradiction together with the progressive course of amygdala calcification in UWD made it necessary to test everyone again in 2010. This time we took note of the WEIRD (Western, Educated, Industrialized, Rich and Democratic) discussion which is currently galvanizing Transcultural Neuroscience (Henrich, Heine, & Norenzayan, 2010a, 2010b; Jones, 2010) and made several changes in the way the tests were administered.

Table 3.1 Demographic data: Age, Wechsler Abbreviated Scale of Intelligence (WASI); verbal IQ (VIQ), performance IQ (PIQ), full-scale IQ (FSIQ), and Benton Face Recognition Test short form with frontal-view (BRTF6) and side-view (BRTF21) faces, for the individual patients, and means and standard deviations (*SD*) for patients and controls, that participated in the subliminal fear-vigilance paradigm and the dynamic emotion recognition task.

	UWDs						**Controls**
	UWD 1	UWD 2	UWD 3	UWD 4	UWD 5	Mean (*SD*)	Mean (*SD*)
Age	22	29	33	47	59	38.0 (*14.9*)	35.5 (*14.5*)
VIQ	95	84	93	82	87	88.2 (*5.6*)	87.5 (*5.9*)
PIQ	98	86	85	84	82	87.0 (*6.3*)	88.9 (*9.3*)
FSIQ	97	84	87	81	83	86.4 (*6.3*)	86.8 (*6.6*)
BFRT6	6	6	6	6	6	6.0 (*0.0*)	6.0 (*0.0*)
BFRT21	15	14	16	13	12	14.0 (*1.6*)	14.8 (*1.1*)

Table 3.2 Demographic data: Age, Wechsler Abbreviated Scale of Intelligence (WASI); verbal IQ (VIQ), performance IQ (PIQ), full-scale IQ (FSIQ), and Benton Face Recognition Test with frontal-view (BRTF6) and side-view (BRTF21) faces, for the individual UWD-subjects, and means and standard deviations (*SD*) for UWD-subjects and controls, that participated in the static emotion recognition task.

	UWDs						**Controls**
	UWD 1	UWD 2	UWD 3	UWD 4	UWD 5	Mean (*SD*)	Mean (*SD*)
Age	24	31	35	49	61	40.0 (*14.9*)	38.0 (*12.5*)
VIQ	95	84	93	82	87	88.2 (*5.6*)	88.5 (*4.3*)
PIQ	98	86	85	84	82	87.0 (*6.3*)	90.9 (*7.8*)
FSIQ	97	84	87	81	83	86.4 (*6.3*)	88.3 (*4.9*)
BFRT6	6	6	6	6	6	6.0 (*0.0*)	5.9 (*0.3*)
BFRT21	15	14	16	13	12	14.0 (*1.6*)	14.9 (*1.0*)

Participants were now tested:

- In their local environment.
- By a local psychologist who speaks the same Afrikaans dialect as they do.
- Using an abbreviated test, the Wechsler Abbreviated Scale of Intelligence (WASI, which provides for a reliable IQ estimate) (Wechsler, 1999), because participants reported being overwhelmed by the burden of WAIS-III testing in 2007.
- The WASI verbal tests were translated by local linguists into the Afrikaans dialect spoken in Namaqualand

The 2010 IQ scores show a global increase of approximately 10% with everyone now falling into the low-normal range. The fact that the changes we made brought about this improvement are in line with the WEIRD discussion (Henrich et al., 2010a, 2010b; Jones, 2010). Specifically, we attribute this improvement to the fact that in 2007 participants were tested in a strange environment and by an unfamiliar person of a different race (especially problematic in post-Apartheid SA), culture, dialect and socioeconomic position. It can

Table 3.3 Social and occupational status of the participants (UWD-subjects and controls).

UWD-ID	**Social Status**
UWD 1	one child, tourism advisor
UWD 2	one child, housewife
UWD 3	own cosmetics business
UWD 4	one child, housewife
UWD 5	two children, charity work
Control-ID	
C 1	trainee nurse
C 2	two children, housewife
C 3	one child, housewife
C 4	clinic assistant
C 5	community health worker
C 6	three children, community health worker
C 7	three children, security guard
C 8	one child, factory supervisor
C 9	chamber maid
C 10	two children, store supervisor
C 11	senior nurse
C 12	clinic assistant
C 13	two children, housewife
C 14	one child, assistant nurse
C 15	three children, bank teller
C 16	one child, security guard

however be stated with confidence that the 2010 IQ scores are still an underestimate of the participants' capabilities. Firstly, although the difference in conditions between 2007 and
2010 made a significant difference, we were obviously unable to overcome all transcultural, language and educational biases inherent in the WASI (Nell, 2000). Secondly, even these improved scores are inconsistent with the participants' ability to compete very favorably for semi-skilled jobs under extremely adverse economic conditions.

Structural MRI assessment

MRI-scans were acquired with a Siemens Magnetom Allegra 3-Tesla head-only scanner at the Cape Universities Brain Imaging Centre (CUBIC) in Cape Town, South Africa. Structural whole brain T2-weighted MRI scans were obtained with 1mm isotropic resolution, TR = 3500 ms, and TE = 354 ms (see **Figure 3.1**).

T2-weighted scans of all 5 UWD-subjects were normalized to MNI-space using the unified model as implemented in SPM5 (http://www.fil.ion.ucl.ac.uk/spm), which is optimized for normalization of lesioned brains (Crinion et al., 2007).

Subsequently the extent of the calcifications was determined with the 3D volume-of-interest feature implemented in MRIcroN (http://www.cabiatl.com/mricro/mricron).

Based on MR-images the precise borders between amygdalae and neighboring structures, or between the sub-regions of the amygdala, cannot be established (Amunts et al., 2005; Solano-Castiella et al., 2010). To determine the precise location of the lesions in our UWD subjects, we therefore assigned the lesion volumes to cytoarchitectonic probability maps according to the method described by Eickhoff and colleagues (2007). In this method, that is implemented in the SPM5 anatomy toolbox (http://www.fz-juelich.de/inm/inm-1/spm_anatomy_toolbox), a volume of interest (VOI) is superimposed onto a cytoarchitectonic probability map of the amygdala and hippocampus (Amunts et al., 2005) (see **Figure 3.2**). This map is based on microscopic analyses of ten postmortem human brains and follows a generally accepted division of the human amygdala in three sub-regions. The first is the central-medial amygdala (CMA), which consists of the central and medial nuclei. The second is the basolateral amygdala (BLA), which includes the lateral, basolateral, basomedial, and paralaminar nuclei, and the third is the superficial (or corticoid) amygdala (SFA), which includes the anterior amygdaloid area, amygdalopyrifom transition area, amygdaloid-hippocampal area, and the cortical nucleus (Amunts et al., 2005). This method assigns to any given voxel a value representing the probability that it belongs to an underlying structure. These are derived from an overlap analysis of ten postmortem brains, and are therefore divided in ten separate probability classes ranging from 10% to 100% probability. For each probability-class of each structure that shares voxels with the VOI, the 'observed versus expected' class representation is computed. This value represents how much more (or less) that class is observed in the VOI compared to what could be expected from the entire probability map of that structure, and is computed with the following equation:

$$P_{observed\ versus\ expected} = \frac{P_{observed} - P_{expected}}{P_{expected}}$$

Whereby $P_{observed}$ represents the percentage of VOI voxels in that class, and $P_{expected}$ represents the percentage of voxels from that class in the whole cytoarchitectonic map of that structure.

The outcome values thus indicate which class is overrepresented in the VOI relative to the whole cytoarchitectonic map.

To estimate how well the lesion volumes fit to the underlying structure, P_{excess} values are computed using the following equation:

$$P_{excess} = \frac{P_{structure\ (VOI)}}{P_{structure\ (total)}}$$

Whereby $P_{structure(VOI)}$ represents the average cytoarchitectonic probability of the voxels that are shared by the structure and the VOI, and $P_{structure(total)}$ represents the average probability of the whole structure's cytoarchitectonic map. These values thus represent how much the average probability of the overlapping voxels exceed the overall probability distribution of that structure, and thus indicate whether the VOI overlaps with relatively high or low probability classes of that structure. In other words, P_{excess} represents how 'central' the location of the VOI is relative to that structure's cytoarchitectonic map, whereby $P_{excess} > 1$ indicates a more central, and $P_{excess} < 1$ a more peripheral location (Eickhoff et al., 2007).

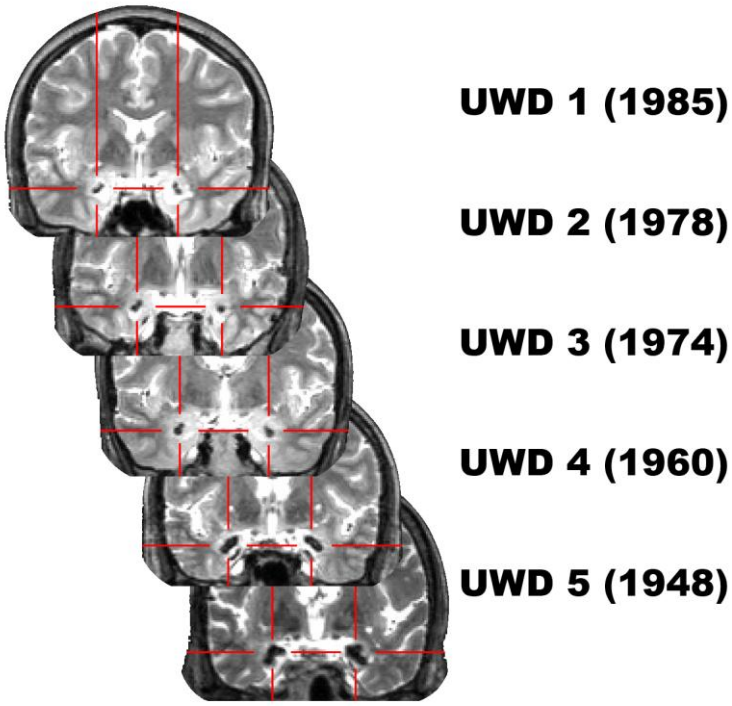


Figure 3.1 T2-weighted MR-images (coronal view) of the five UWD subjects with birth year and crosshairs indicating the calcified brain damage.

Functional MRI assessment

Functional whole brain MRI-scans were obtained with a 2D-EPI sequence with 36 slices in interleaved-ascending order, 3.5 mm isotropic resolution, Flip-angle = 70^{o}, TR = 2000 ms, TE = 27 ms, and EPI-factor = 64. The first 4 volumes were acquired prior to the start of the fMRI task, and discarded from analyses.

Participants viewed a trio of faces and matched emotional expressions by choosing one of the two lower pictures (either an angry or a fearful face) that expressed the same emotion as the picture on top. This condition was interleaved with a sensori-motor control condition involving the matching of oval shapes (Hariri et al., 2002). To increase cultural validity, gray-scaled face-stimuli included Caucasian as well as African-American actors (Tottenham et al., 2009), and the shape-stimuli were constructed from scrambled face-stimuli to match visual contrast levels.

The task was presented in a blocked design, with 5 shape-matching, interleaved with 4 emotion-matching, blocks, with six 5-second trials each, and always including faces of one gender only. All stimuli were presented equally often as target, match or non-match in randomized order. Each block was preceded by the instruction ‘match emotion’ or ‘match shape’ (in Afrikaans) for 2 seconds, making a total task duration of 288 seconds. Participants responded by a button-press with either the left or right hand, corresponding to the position of the match-stimulus.

Analyses were performed with SPM5 (http://www.fil.ion.ucl.ac.uk/spm). For each participant all volumes were realigned to the first volume, and coregistered to the structural T2-weighted volume (Friston et al., 1995). Subsequently, the resulting functional images were normalized to MNI-space using the parameters obtained from the structural analysis, and smoothed with a full-width-half-maximum Gaussian kernel of 8x8x8 mm. Contrast-maps for match-emotion > match-shape were obtained with realignment parameters and high-pass filter (cut-off 128 s) entered as regressors of no interest. For group-level statistics these were entered in a one-sample *t*-test analysis. Functional activation of the amygdala was assessed bilaterally within regions of interest (ROI’s) of the BLA and the combined CMA and superficial amygdalae (SFA) (see **Figure 3.2**). ROI’s were constructed based on the cytoarchitectonic probability maps as implemented in the anatomy toolbox for SPM5 (Amunts et al., 2005; Eickhoff et al., 2005). We applied an extent-threshold of 10 voxels, and

significance threshold was set at $p < .05$ (false-discovery-rate (FDR) corrected). This rather lenient threshold is justifiable given that we presently only assess whether the amygdala's sub-regions are responsive in general.

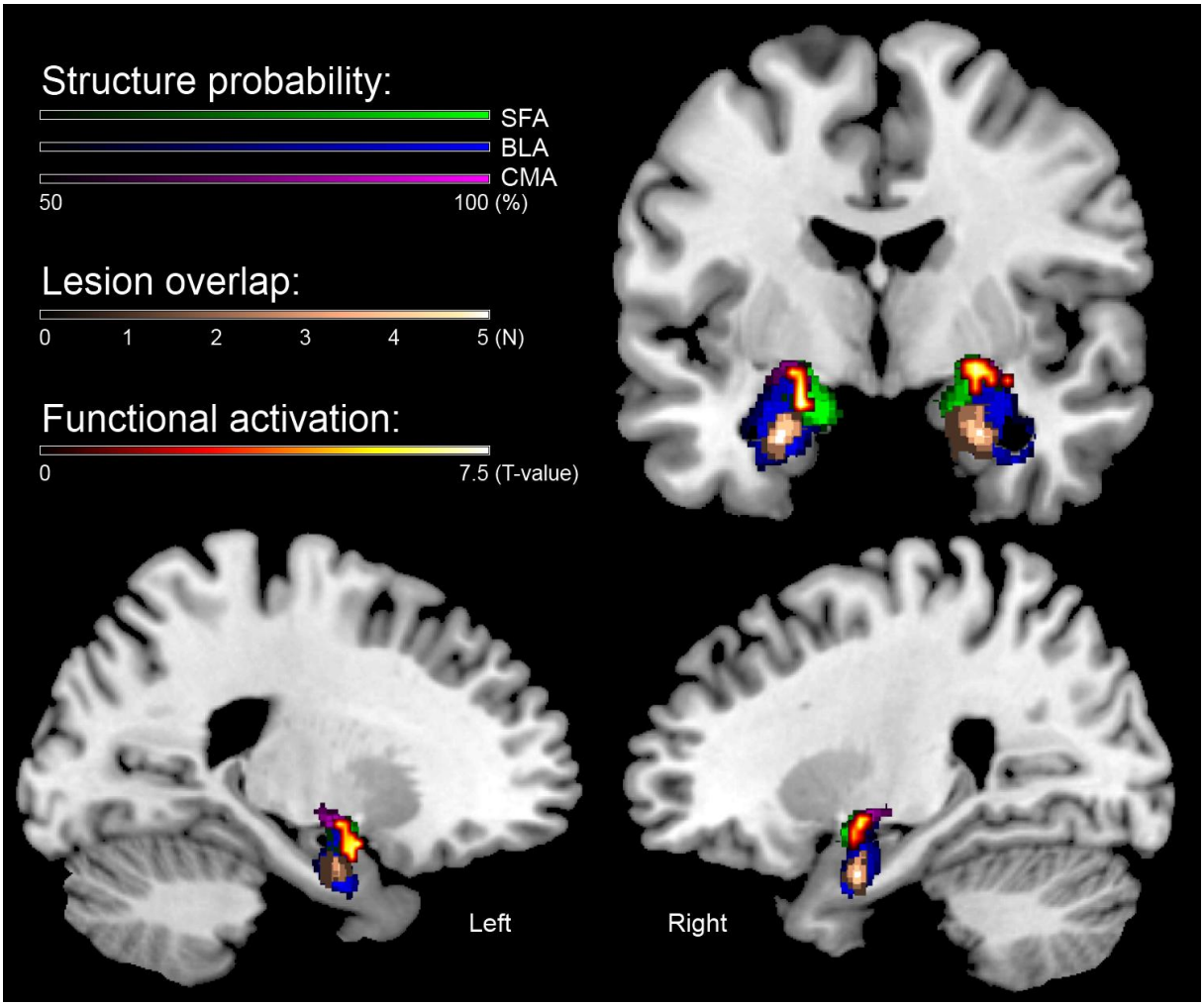


Figure 3.2 Structural and functional MRI assessment of the bilateral amygdala in our group of five subjects with Urbach-Wiethe disease. Plotted are the cytoarchitectonic probability-maps of the amygdala thresholded at 50% (Amunts et al., 2005), structural lesion overlap, and functional activation during the emotion-matching task (contrast: match emotion > match shape, $p < .05$, FDR-corrected) on a template brain. The structural method indicates that the lesions of the five UWD-subjects are located in the basolateral amygdala (BLA), while the functional method shows activation during emotion matching in the superficial amygdala (SFA) and the central-medial amygdala (CMA), but not in the BLA.

Behavioral assessment: Subliminal fear vigilance task

Participants verbally named as quickly as possible the color of backwardly-masked fearful, happy, and neutral faces (van Honk et al., 2005; van Honk et al., 2002), whereby a generic slow-down in color-naming of threat-related information is reliably associated with automatic vigilance to threat (Algom et al., 2004). After a fixation-cross (750 ms), randomly one of ninety face stimuli (Ekman & Friesen, 1976) (5 male and 5 female, 3 emotions, colored in red, green or blue) was presented for 14ms, before being replaced by a masking-stimulus. Intertrial interval was variable between 1500-2500 ms. Masking-stimuli were randomly-cut, reassembled, and rephotographed pictures of the faces. Color-naming latencies > 2*SD* from the individual means were removed from the analysis (4.6%).

Afterwards, participants performed on an objective awareness-check. This was a three-alternative forced-choice (3AFC) emotion recognition task, using the same masked stimuli from the original task, which establishes awareness of the measure of interest; emotional expression (Van Selst & Merikle, 1993; Wells & Matthews, 1994).

Behavioral assessment: Dynamic emotion task

Participants were presented with clips of faces, 2 male and 2 female actors (Ekman & Friesen, 1976; Lundqvist, Flykt, & Öhman, 1998), morphing fluidly from neutral to emotional (anger, disgust, fear, happiness, sadness and surprise), and were instructed to choose which of six emotional adjectives (Afrikaans translation of angry, disgusted, fearful, happy, sad and surprised) best described each face. The final frame remained visible until the participant responded with a button-press. Emotional intensity of the final image in the sequence ranged from 20% to 100%, in steps of 10% in consecutive blocks, with all stimuli randomized within each block. Accordingly the duration of the video-clips ranged from 0.3 s in the first to 1.7 s in the final block with full-blown emotions. The clips were presented within a visual angle of 10^{o} and eye movements were recorded from clip-start until the participant's response with a Tobii-1750 binocular infrared eye-tracker (50 Hz, 0.5^{o} accuracy).

Performance data on the full-blown (100%-morphed) trials served as a measure of emotion recognition accuracy. Additionally, for each actor and emotion the morphing percentage after which the emotion is consistently recognized was determined. These were averaged to obtain individual sensitivity scores for each emotion.

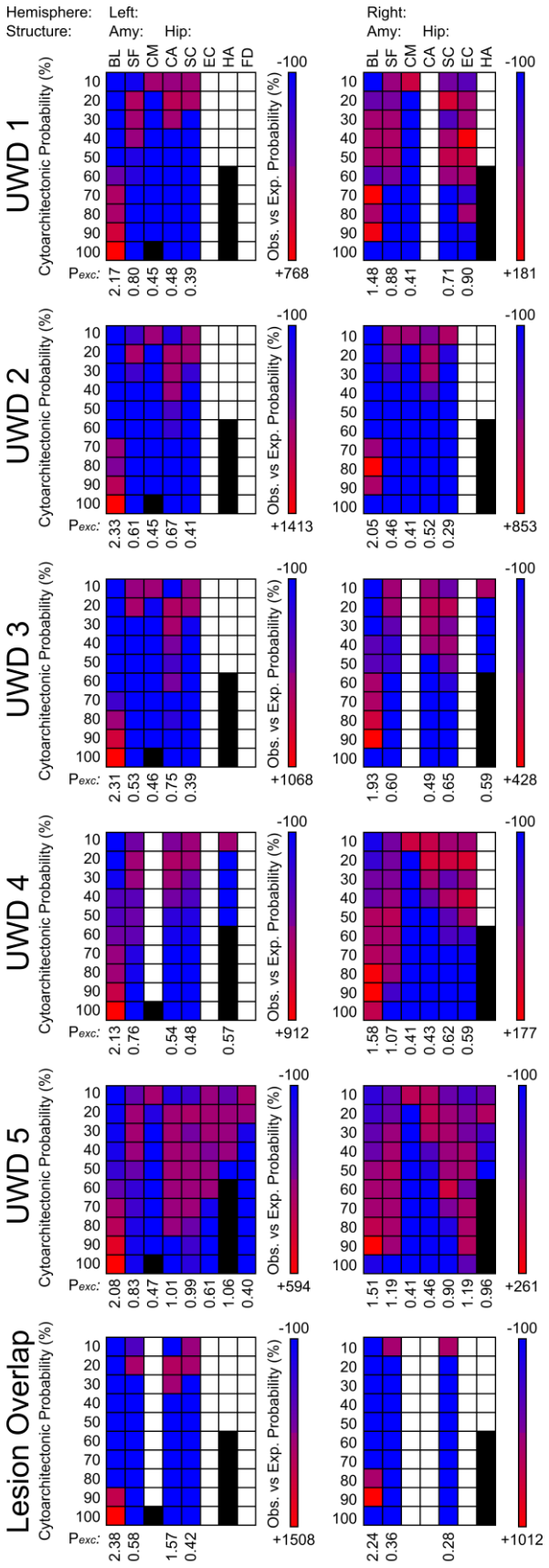
Gaze-fixations were defined as the average location of all subsequent gaze-points within 2^{o} visual angle and with a minimal duration of 60 ms (Tobii Technology, 2006). Fixations within oval areas drawn around the mouth and both eyes separately of the stimuli were used to compute average fixation duration and proportion fixations to these areas relative to all fixations on the face.

Behavioral assessment: Static emotion task

To facilitate behavioral comparison with earlier studies on UWD, our subjects also performed in a similar emotion-rating task (Adolphs et al., 2005; Adolphs et al., 1994, 1995). The static emotion task is based on the paradigm described by Adolphs and colleagues in their first studies on UWD (Adolphs et al., 1994, 1995). In each trial, participants rate how well a static face with one of seven emotional expressions (angry, disgusted, fearful, happy, sad, surprised or neutral) corresponds to one of six emotional adjectives (i.e., the Afrikaans translation of; 'How angry / disgusted / fearful / happy / sad / surprised do you think this person is?'). Stimuli were presented for 3 seconds in the center of a computer screen subtending approximately 18^{o} vertically and 14^{o} horizontally to the participant's eyes. Eye-movements were recorded with a Tobii-1750 binocular infrared eye-tracker with a sampling-rate of 50Hz, and 0.5^{o} accuracy (Tobii Technology, 2006).

Face-stimuli were 3 male and 3 female actors expressing all 7 emotions (Ekman & Friesen, 1976; Lundqvist et al., 1998), making a total of 42 stimuli. The rating task was divided in 6 blocks, one for each emotional adjective, presented in random order. In each block participants rated all 42 stimuli, in random order, on one of the adjectives. A visual-analogue scale (VAS), ranging from 'not <adjective> at all' to 'very <adjective>', was used for the rating procedure. Implicitly to the participants, the VAS was quantified in a range from -100 to +100. In the 7th and last block of the task all the stimuli were presented once again, but now participants were instructed to identify the facial expressions in a 6-alternative (the same adjectives) forced-choice design (i.e., 'Which emotion does this person display?').

Stimulus presentation commenced only after participants fixated their gaze anywhere on the screen to ensure valid eye-movement recordings without biasing the initial fixation location. After stimulus presentation the VAS appeared (subtending 28^{o} horizontally) on a touch-screen adjacent to the eye-tracker screen, and participants performed ratings by



pressing with their finger anywhere on this scale. Ratings could be adjusted until a button labeled 'next' was pressed which started the next trial. In the final forced-choice emotion recognition block, the emotional adjectives appeared as separate buttons on the touch-screen after stimulus presentation.

Figure 3.3 Observed versus expected probability matrices for the individual brain lesions and their overlap. Columns are the observed brain areas, and rows their cytoarchitectonic probability classes. Colors indicate the relative over- (red) or under- (blue) representation of a structure-class in the lesion volume. White indicates no overlap between lesion and structure probability map, and black indicates probability classes that are not represented in the cytoarchitectonic map. P_{excess} values indicate how much more likely a structure was observed in the lesion volume as could be expected from its own probability distribution, and thus reflect how central to the area the lesion volume is (Eickhoff et al., 2007). Amy = Amygdala, Hip = Hippocampus, BL = Basolateral, SF = Superficial, and CM = Central-Medial, which are all amygdala sub-regions, and CA = Cornu Ammonis, SC = Subicular Complex, EC = Entorhinal Cortex, HA = Hippocampal-Amygdaloid Transition Area, and FD = Fascia Dentata, which are all bordering- or sub-regions of the hippocampus.

Gaze-fixations were defined as the average location of all subsequent gaze-points within 2^{o} visual angle, with a minimal duration of 60 ms (Tobii Technology, 2006). Fixations within oval areas drawn around the eyes and mouth of the individual stimuli were used to compute average fixation duration and proportion fixations to these areas relative to all fixations on the face.

Results

Structural MRI assessment

As depicted in **Figure 3.1**, amygdala calcification appears to progress with age (Appenzeller et al., 2006). Calcified brain-tissue is localized in the BLA (see **Figure 3.2**), whereby the lesions in the two oldest subjects possibly extend into the borders of the right SFA. Crucially, in all subjects the CMA seems unaffected by the calcifications.

In a quantitative analysis these results are confirmed. **Figure 3.3** shows $P_{observed\ \text{versus}\ expected}$ and P_{excess} values for the individual lesions, and for the cluster of voxels where all lesions overlap. From **Figure 3.3** we can make three observations. 1) The structures that might be affected by the lesion. 2) Which probability classes of those structures are most, or least, affected. 3) How 'central' the lesions are to the probability distributions of the underlying structures, represented by P_{excess} values. For the lesioned tissue in the UWD-subjects P_{excess} reached values of 2.17, 2.33, 2.31, 2.13, and 2.08 in the left-BLA, and 1.48, 2.05, 1.93, 1.58 and 1.51 in the right, as a function of chronological age. For the lesion-overlap volumes P_{excess} reached values of 2.38 and 2.24 for the left and right BLA respectively, while P_{excess} values for all other structures was < 0.6. Thus, as can also be seen in **Figure 3.3**, the lesions are, bilaterally, most central to the BLA, whereby the left-sided lesions are centered in the area with 100% BLA-probability, and the right-sided lesions in the area with 90% BLA-probability. Moreover, for all lesion volumes P_{excess} values were highest for the BLA (all exceeding 2.0 for the left-sided and 1.5 for the right-sided lesions).

Since this method is purely based on probability distributions, it is impossible to fully exclude that other structures than the BLA are affected by the calcifications. The fact that the lesion-volumes largely overlap with high probability classes in the bilateral BLA, and that P_{excess} values greatly exceed the value of 1, can however be seen as strong support for our

claim that these UWD-subjects have bilateral damage limited to the BLA. We must however note that we cannot fully exclude that in the two oldest subjects the calcifications might extend into neighboring structures. Namely, in subject UWD 4 the right SFA (P_{excess} = 1.07), and in subject UWD 5 the left hippocampus (Cornu Ammonis: P_{excess} = 1.01), left Hippocampal-Amygdaloid Transition Area (P_{excess} = 1.06), right SFA (P_{excess} = 1.19), and right Entorhinal Cortex (P_{excess} = 1.19). Based on P_{excess} for the CMA (all < 0.5) we can however safely conclude that this structure is relatively unaffected by the bilateral calcifications found in these UWD-subjects.

Finally, visual inspection of the T2-weighted scans (MRIcroN, http://www.cabiatl.com/mricro/mricron) revealed likely age-related atrophy in the putamen of subject UWD 4. The putamen is, however, not directly involved in processing of fearful expressions (Fusar-Poli et al., 2009), and therefore not of immediate relevance to the present study.

Functional MRI assessment

ROI-analysis (p < .05, FDR-corrected, see **Figure 3.2**) revealed no significant activation clusters in the bilateral BLA, but significant activation of 40 (left), and 35 (right) voxels in ROI's consisting of the amygdala excluding the BLA. Thus, although the BLA is damaged in these subjects, the remaining amygdala tissue seems to be functional.

Behavioral assessment: Subliminal fear vigilance paradigm

Color-naming latencies on fear trials were referenced against the other emotions, to create either positive or negative attentional bias scores representing threat vigilance or avoidance respectively (van Honk et al., 2005; van Honk et al., 2002). Overall, UWD-subjects and controls were equally fast in color-naming (540 ms vs. 524 ms, U = 21, p = .383). Crucially, fear bias-scores were significantly higher for the UWD-subjects group (fear-neutral; U = 8, p = .020, r = .56, fear-happy; U = 5, p = .008, r = .64, see **Figure 3.4**), indicating that color-naming in the UWD-group was significantly slowed-down when, subliminally, a fearful face was presented. Subsequently we applied a strict neutral baseline correction by computing separate fear-neutral bias-scores for trials that were preceded by neutral trials, which eliminates trial-by-trial emotional conflict (Etkin, Egner, Peraza, Kandel, & Hirsch, 2006;

Kunde & Mauer, 2008). On this pure measure of fear hyper-vigilance the UWD-subjects again showed significant fear-interference (U = 5, p = .008, r = .64, see **Figure 3.4**).

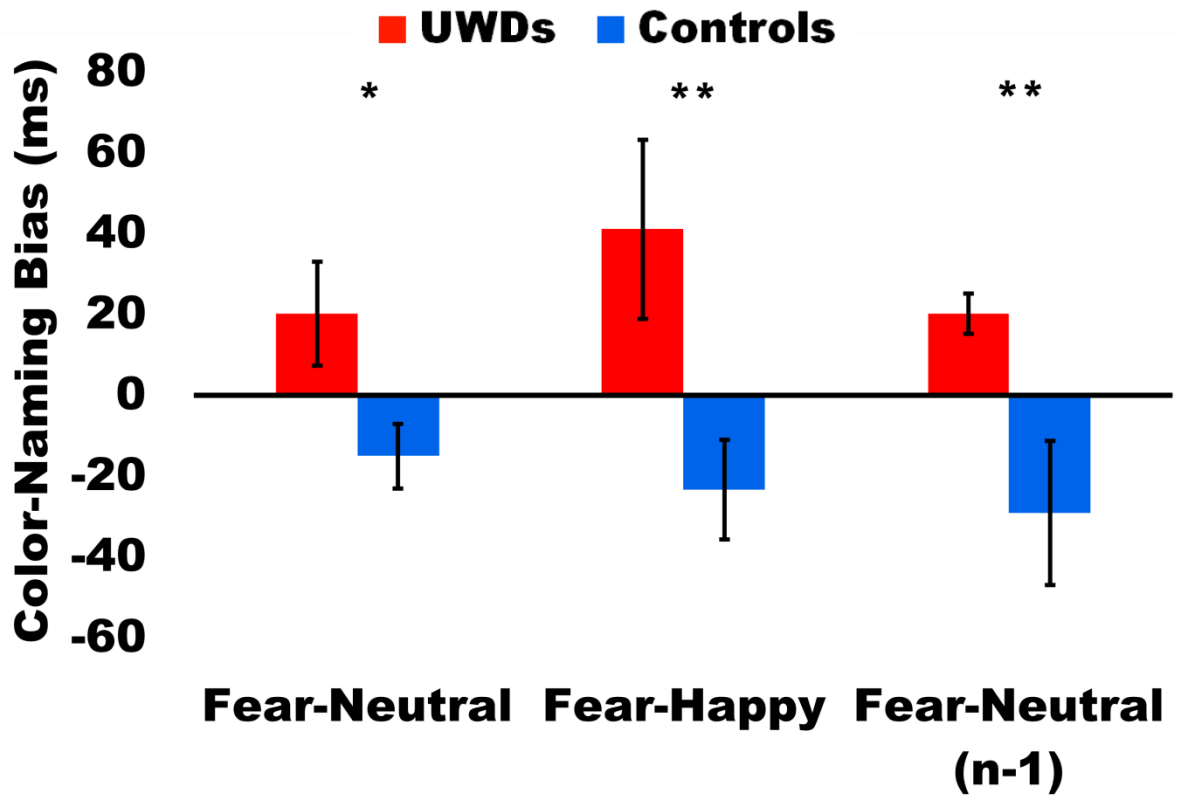


Figure 3.4 Behavioral data from the subliminal fear vigilance task. Bias scores are computed by subtracting mean latencies on neutral and happy trials from mean latencies on fear trials, and by subtracting mean latencies on neutral trials from fear trials that were preceded (n - 1) by neutral trials. Positive values represent slower color-naming responses when subliminally confronted with fearful faces compared to the control conditions, which is a reliable index of hyper-vigilance for subliminally presented threat cues (Algom et al., 2004; van Honk et al., 2005; van Honk et al., 2002). Error bars represent *SEM*. * = p < .05, ** = p < .01

None of the participants reported awareness of the facial expressions, but one control-subject scored above chance-level on the awareness-check (15 correct answers on 30 3-alternative trials, one-tailed binomial-test; p = .040). For the remaining participants emotion awareness-check performance was not different from chance-level (10.5 vs. 10.6 correct answers), and all group differences on fear hyper-vigilance remained significant after excluding this participant (fear-neutral; U = 8, p = .027, r = .55, fear-happy; U = 5, p = .011, r = .64, fear-neutral baseline corrected; U = 5, p = .011, r = .64).

Behavioral assessment: Dynamic emotion task

Table 3.4 provides performance scores (accuracy and sensitivity) on each emotion. Since our hypothesis concerns fear processing in particular, we report the eye-movement data for the fear trials, and for all emotions pooled together.

UWD-subjects and controls gazed overall equally long at the faces (U = 29, p = .916), which indicates that reaction times for both groups were similar. UWD-subjects directed 28% of their fixations to the mouth region, which was not significantly different (U = 24, p = .527) from controls (24%), and these fixations were similar in duration (376 ms vs. 409 ms, U = 25, p = .598). Percentage of fixations directed to the eyes was also similar (17% vs. 19%, U = 26, p = .673), but these were significantly longer for UWD-subjects (381 ms vs. 306 ms, U = 11, p = .045, r = .49). Thus, allocation of spatial attention was equal for both groups, but UWD-subjects exhibited more sustained visual attention to dynamically presented eyes.

Table 3.4 Emotion recognition data: Performance-scores (accuracy and sensitivity) with standard deviations and *p*-values for the non-parametric Mann-Whitney *U*-tests for patients versus controls on the dynamic emotion recognition task.

	Accuracy: %-Correct (*SD*)			**Sensitivity: %-Morph (*SD*)**		
	UWDs	Controls	*p*-value	UWDs	Controls	*p*-value
Anger	65 (*38*)	92 (*12*)	.150	76 (*26*)	60 (*13*)	.139
Disgust	45 (*27*)	60 (*38*)	.354	84 (*23*)	83 (*20*)	.916
Fear	85 (*22*)	60 (*23*)	.048*	84 (*12*)	92 (*12*)	.153
Happiness	100 (*0*)	100 (*0*)	1.00	28 (*15*)	26 (*12*)	.665
Sadness	60 (*38*)	56 (*30*)	.914	83 (*29*)	89 (*21*)	.672
Surprise	70 (*21*)	67 (*39*)	.784	80 (*16*)	71 (*28*)	.712
Total	71 (*16*)	73 (*13*)	.830	72 (*15*)	70 (*10*)	.916

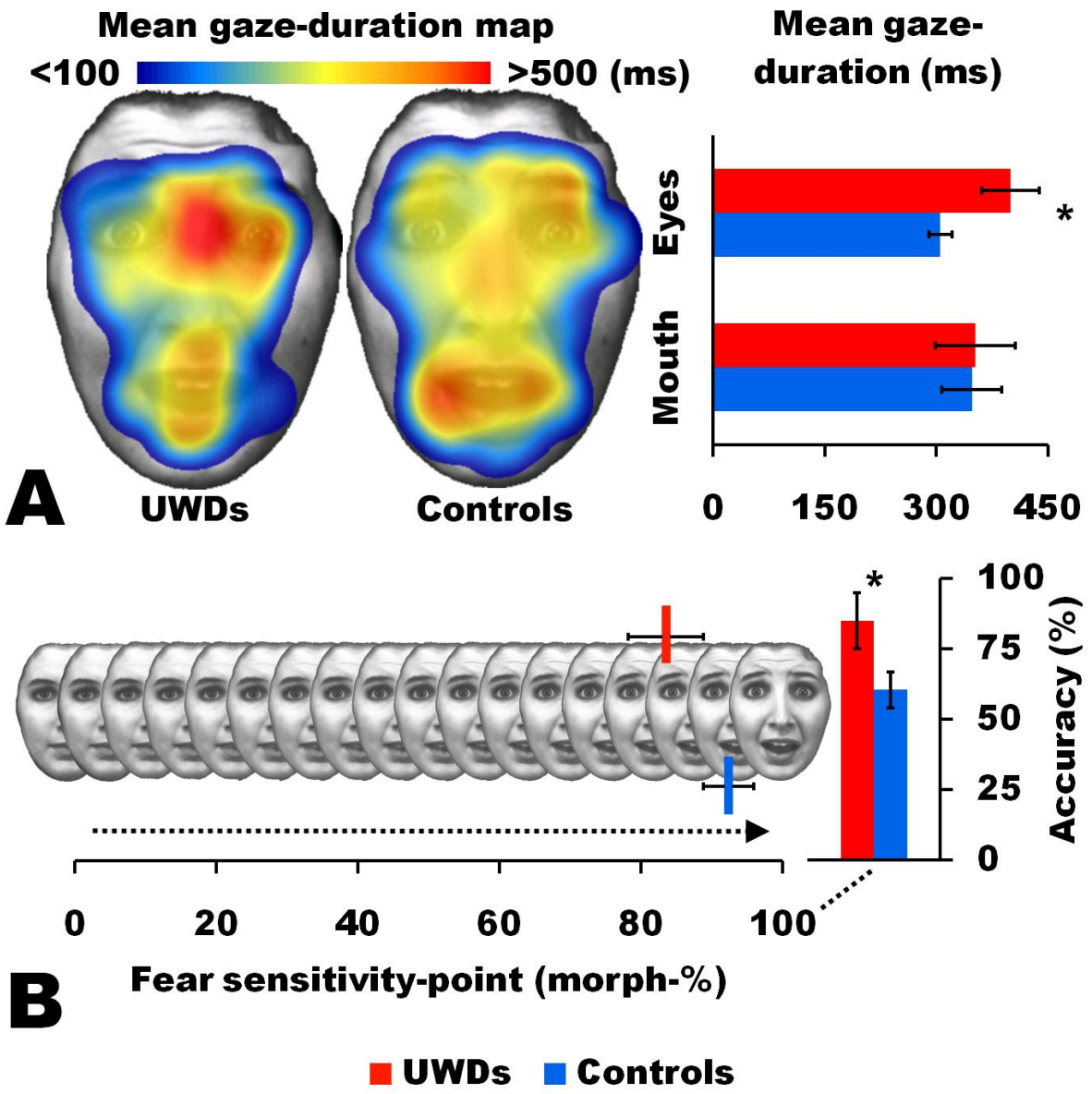


Figure 3.5 Behavioral data from the dynamic emotion task. **A)** Mean duration of gaze fixations on the fear facial expressions; mapped on one of the stimuli (Lundqvist et al., 1998), and quantified for the eye- and mouth-regions. **B)** Performance on fear recognition; sensitivity points for fear recognition indicating the average percentage of morphed fear necessary for consistent recognition, and recognition accuracy on the full-blown fear trials. Error bars represent SEM. * = $p < .05$

Total time spent looking at dynamic fearful faces was equal for both groups ($U = 27$, $p = .752$). UWD-subjects directed 25% of all face-fixations at the mouth region, which was not significantly different ($U = 24$, $p = .527$) from controls (20%). Number of fixations to the eye-

region of the fearful faces was also equal for both groups (20% vs. 23%, $U = 28$, $p = .833$). Crucially, as in the whole-task analysis, duration of fixations at dynamically presented fearful eyes was longer for UWD-subjects (400 ms vs. 305 ms, $U = 10$, $p = .035$, $r = .51$, see **Figure 3.5A**), while there was no duration difference for mouth-fixations (353 ms vs. 348 ms, $U = 27$, $p = .752$, see **Figure 3.5A**). Thus, allocation of attention was similar for both groups, but UWD-subjects exhibited prolonged attention to dynamically presented fearful eyes.

Sensitivity scores on the fear trials were not significantly different (84% vs. 92%, $U = 16.5$, $p = .153$), but in keeping with the hypothesis that visual attention to the eyes improves fear recognition ability (Adolphs et al., 2005), UWD-subjects outperformed controls on full-blown fear trials (85% vs. 60% correct, $U = 12.5$, $p = .048$, $r = .48$, see **Figure 3.5B**).

Behavioral assessment: Static emotion task

Raw ratings on the emotion rating task were normalized for each participant to control for individual differences in use of the VAS, and averaged for each presented emotion and each adjective in the 6 rating blocks. The resulting matrix of 7 (emotional expressions) by 6 (rating-questions) was compared cell-by-cell for group differences with two-tailed non-parametric Mann-Whitney *U* tests. **Figure 3.6A** is a visual representation of the normalized rating-scores for both groups and the resulting matrix of p-values for the statistical tests (not corrected for multiple comparisons). As can be seen from **Figure 3.6A** the response pattern is similar for both groups, and the only significant difference that emerged was a lower 'surprised' rating of 'sad' faces in the UWD-group ($U = 8$, $p = .020$).

To assess rating performance on each emotion we constructed weighted performance-scores by computing the average rating-difference of a facial expression on the correct adjective, with the other expressions on that adjective, and the other adjectives on that facial expression. The neutral faces were not included in this computation. Thus, for all the cells on the diagonal of the matrix figures (see **Figure 3.6A**, excluding the neutral row) the differences with the cells on the same row and in the same column were averaged. The resulting values represent the relative difference between an emotionally congruent rating with its incongruent alternatives, and as such is a measure of performance on each emotion. Because these are relative scores, non-normalized ratings were used for this computation. The resulting

performance scores are depicted in **Figure 3.6A**, but none of the group-differences reached significance (Mann-Whitney *U*-tests, all *p*'s > .4).

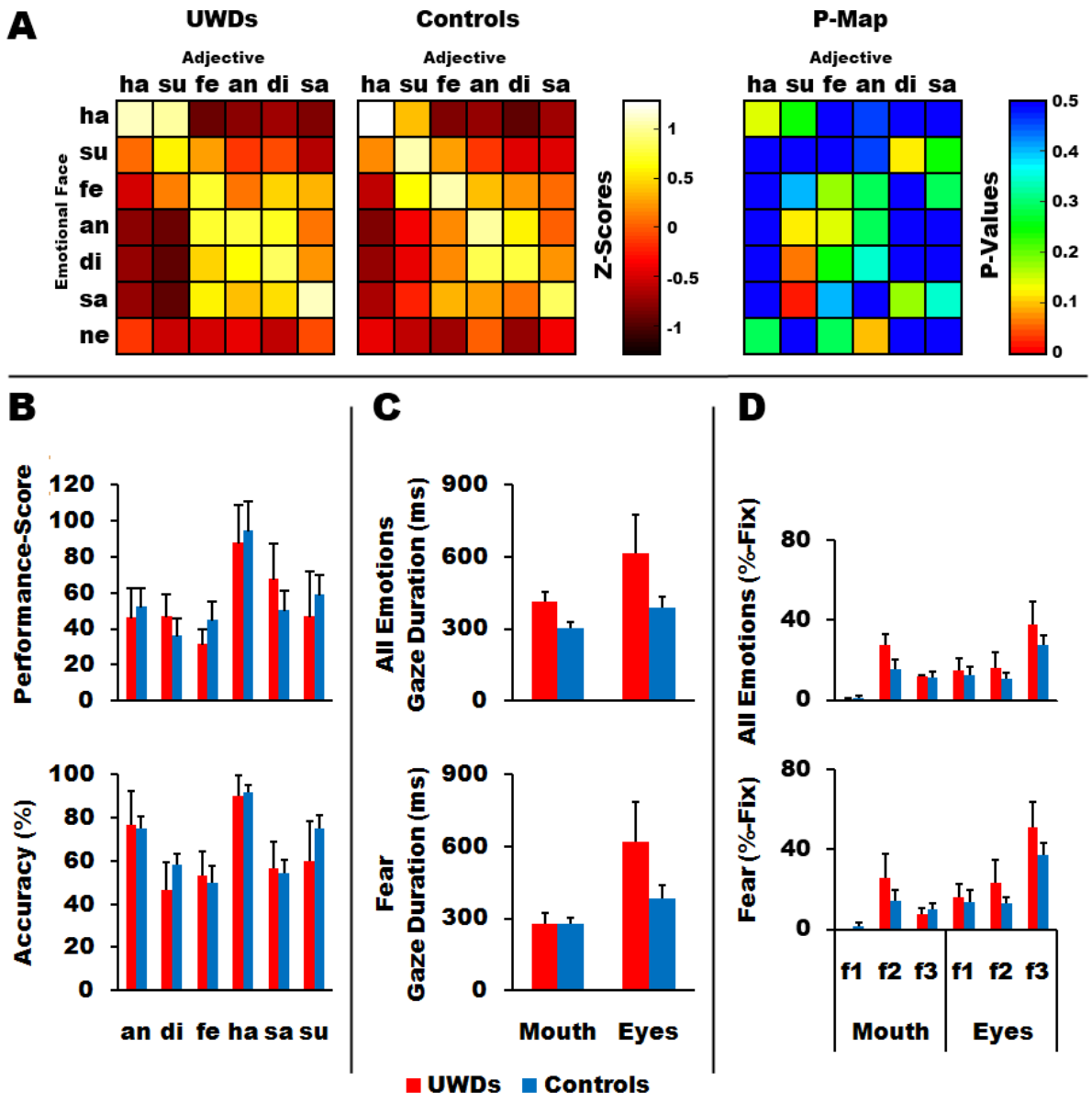


Figure 3.6 A) Average normalized ratings (Z-scores) for both groups on each adjective; happy (ha), surprised (su), fearful (fe), angry (an), disgusted (di) and sad (sa), for each emotional facial expression; happy (ha), surprised (su), fearful (fe), angry (an), disgusted (di), sad (sa), and neutral (ne), and p-values (Mann-Whitney *U*-tests) for the group-differences. **B)** Performance-scores for each emotion in the rating task and accuracy in emotion recognition. **C)** Gaze duration on the whole task and on the fear trials for fixations to the mouth and eye regions. **D)** Proportion of the first three fixation (f1, f2, f3) to the mouth and eye regions during the whole task and on the fear trials.

Lastly, we assessed emotion recognition performance in the final block of the task. In **Figure 3.6B** the average accuracy for each emotion is depicted, and again, no significant group-differences were found in performance or reaction time (Mann-Whitney *U*-tests, all *p*'s > .5). In sum, performance on intensity-rating and recognition of static emotional facial expressions was not different between UWD-subjects and healthy controls.

To limit the number of statistical comparisons, we report eye-movement data only for the fearful faces separately and for all trials combined, starting with the latter. Reported values are always UWD versus healthy control group. The average time spent looking at the faces was equal for both groups ($U = 18$, $p = .206$), which was also the case for the overall average fixation duration ($U = 19$, $p = .246$), the percentage of fixations at the mouth (13% vs. 10%, $U = 18$, $p = .206$), and eyes (20% vs. 18%, $U = 29$, $p = .916$). The duration of mouth-fixations was significantly longer in the UWD-subjects (394 ms vs. 298 ms, $U = 10$, $p = .035$, $r = .51$), which also reached a trend for the eye-fixations (616 ms vs. 381 ms, $U = 12$, $p = .058$, $r = .46$), see **Figure 3.6C**.

The average time spent looking at the fearful faces was equal for both groups ($U = 18$, $p = .206$), which was also the case for the average fixation duration ($U = 19$, $p = .246$), the percentage of fixations at the fearful mouth (11% vs. 9%, $U = 21$, $p = .342$), and fearful eyes (26% vs. 22%, $U = 28.5$, $p = .874$). Again there was a trend for longer eye-fixations in de UWD-group (625 ms vs. 379 ms, $U = 12$, $p = .058$, $r = .46$), but the duration of mouth-fixations was not significantly different (289 ms vs. 273 ms, $U = 20$, $p = .739$), see **Figure 3.6C**.

Recently it has been shown that the lack of visual attention to the eye-region of faces after complete focal bilateral amygdala damage was mainly observed in the first fixations on a newly presented static face (Kennedy & Adolphs, 2010). Therefore we analyzed the first three fixations separately. **Figure 3.6D** depicts the proportion of the first three fixations on the mouth and eyes for the whole task and the fear trials separately. None of the group-differences were significant (all *p*'s > .2). In sum, UWDs visually allocate attention equally fast and often to the eye-region of faces, but in general attend somewhat longer to the mouth and eyes, and to only the eyes when static faces are fearful.

Discussion

Using a combination of structural and functional MRI as well as eye-tracking and behavioral measures, we provide causal evidence that the human BLA acutely inhibits innate threat vigilance. Five UWD-subjects with selective bilateral damage to the BLA show hyper-vigilance to subliminally presented fearful facial expressions. Moreover, they gazed longer to the eye-region of, especially fearful, faces, while allocation of attention was similar to that of a carefully matched group of healthy controls. Following evidence that increased attention to the eye-region of faces improves fear recognition performance (Adolphs et al., 2005), UWD-subjects showed superior ability in dynamic fear recognition. These combined results establish that focal damage to the BLA makes humans hyper-vigilant to the innate cue for threat; fearful facial expressions.

To understand how BLA loss may lead to such hyper-vigilance, studies in animals provide valuable information. It is argued that the BLA inhibits the response to innate danger cues, because loss of BLA function in rodents leads to increased unconditioned fear (Macedo et al., 2005; Macedo et al., 2007; Macedo et al., 2006; Martinez et al., 2007). A possible neural pathway of such inhibition is through the CMA (Barbas, 2007; Gozzi et al., 2010; Tye et al., 2011), which projects vastly to hypothalamic and brainstem areas, that regulate emotional responding through autonomic pathways (Heimer et al., 1997). The CMA receives direct input from other parts of the amygdala (Davis & Whalen, 2001; Phelps & LeDoux, 2005; Tye et al., 2011), as well as from the PFC (Barbas et al., 2003), and is considered to be the amygdala's behavioral output center, automatically allocating attention and directing autonomic and motor responses to threat (Ciocchi et al., 2010; Gozzi et al., 2010; Haubensak et al., 2010; Heimer et al., 1997; Holland & Gallagher, 1999; LeDoux, 1993; Mosher, Zimmerman, & Gothard, 2010). Importantly, the CMA is essential for the expression of active fear behaviors, as well as freezing responses generated in brainstem areas in response to acute threats (Kalin, Shelton, & Davidson, 2004).

The BLA is often regarded as the 'sensory' amygdala. It receives input from the sensory systems via the thalamus (LeDoux, 1993), as well as highly processed polymodal sensory information from association cortices including the PFC (Barbas et al., 2003; Davis & Whalen, 2001; Solano-Castiella et al., 2010). The BLA is therefore argued to be involved in the automatic assessment of threat (Liddell et al., 2005), and acquisition and extinction of

conditioned fear (Amaral, 2003; Parkes & Westbrook, 2010; Phelps & LeDoux, 2005). Furthermore, the BLA can through mutual connections with the PFC, especially the orbitofrontal cortex (OFC), both promote and inhibit the behavioral output-functions of the CMA (Barbas, 2007; Barbas et al., 2003; Garcia, Vouimba, Baudry, & Thompson, 1999; Hampton, Adolphs, Tyszka, & O'Doherty, 2007; Quirk, Likhtik, Pelletier, & Pare, 2003; Salzman & Fusi, 2010), which might provide an explanation for the increased threat vigilance found in our UWD-subjects.

An alternative explanation might be found in the recently discovered direct inhibitory functions of the BLA on the CMA in rodents. First, an inhibitory pathway from the lateral to medial CMA (Ciocchi et al., 2010; Haubensak et al., 2010) has been shown to switch CMA functions from the promotion of basal freezing responses to active threat assessment in the presence of an acute threat (Gozzi et al., 2010), and a direct projection from the BLA to this lateral CMA was recently reported to acutely reduce fearful behaviors (Tye et al., 2011). This landmark study showed that after optogenetic stimulation of BLA terminals in the lateral CMA fearful behavior decreased, and rodents started to explore potentially unsafe surroundings. Exploring was however significantly reduced when the same projection was inhibited. Importantly, no effect was observed after glutamergic stimulation of BLA somata, possibly reflecting the direct excitatory pathways from BLA to the medial CMA that can counteract the anxiolytic effects. It therefore seems that only acute stimulation of the BLA-CMA pathway reduces fearful behavior (Tye et al., 2011), which subsequently promotes the switch away from passive fear responding in the CMA (Gozzi et al., 2010). This notion is further supported by several studies showing that BLA deactivation increases unconditioned fear behavior and acute freezing responses, while conditioned and more generalized fear is unaffected or even reduced (Macedo et al., 2005; Macedo et al., 2007; Macedo et al., 2006; Martinez et al., 2007).

Furthermore, basal fear inhibition by the BLA could also be explained in terms of parallel models of amygdala functioning. In these models, the BLA, together with the nucleus accumbens (NA), is thought to be part of a system that underlies instrumental choice behaviors, whereas CMA-NA interactions subserve reflexive behavioral responding (Balleine & Killcross, 2006; Killcross, Robbins, & Everitt, 1997). Notably, for an efficient instrumental response to threats, the option for inhibitory control of reflexive fight-flight mechanisms is

necessary, while in acutely threatening situations defensive reactivity can get priority. The latter is reflected in the switch from BLA-PFC to CMA-PAG activation when a threat becomes so proximal that it is unavoidable (Maren, 2007; Mobbs et al., 2009; Mobbs et al., 2007). Speculatively, the BLA might provide the necessary conditions for higher-order instrumental choice behaviors in mildly threatening situations. Following this model it might be expected that UWD-subjects are more impulsive in decision making, but this needs to be confirmed by future research.

Note that the above described mechanisms of acute fear regulation are independent from whether or not the amygdala has a direct role in the evaluation of threat related information. Indeed, our UWD-subjects are not impaired in emotion recognition, which suggests that the BLA does not contribute to conscious emotion recognition. Furthermore, responsivity to innate threat cues, like fearful faces, in subcortical areas is a survival reflex relatively independent from, but projecting to, the amygdala (LeDoux, 2012; Liddell et al., 2005). Downregulation of such acute fear responsivity might therefore be the BLA's default mode, thereby reducing defensive reflexes and creating the conditions for a more instrumental response. Although conscious evaluation of emotional information thus seems not affected, it might be expected that reduced inhibition of such acute threat responding will affect our UWD-subjects' ability to evaluate emotionally conflicting information correctly, but this remains to be tested.

In sum, the rodent BLA apparently can acutely inhibit fear responses to innate danger cues through its influence on the CMA, and our corresponding behavioral data in BLA-damaged subjects suggest that this BLA-CMA pathway may act in similar ways in humans. We do however have no insights into the intricate neural pathways with the present evidence, and the question whether the hyper-vigilance in our subjects with BLA damage is caused by direct disinhibition of the CMA, indirectly via prefrontal areas, or both (see **Figure 3.7**), awaits future research. Nonetheless, our data do show that damage to the BLA in humans leads to hyper-vigilant responses to innate threat cues. Such threat hyper-vigilance in humans is hypothetically related to acute fear and panic (Graeff & Del-Ben, 2008; Maren, 2007; McNaughton & Corr, 2004). Given the high prevalence of comorbid anxiety disorders, including social phobia and panic disorder observed in UWD (Thornton et al., 2008; Wiest,

Lehner-Baumgartner, & Baumgartner, 2006), our data provides important insights into the neural mechanisms of disorders of fear and anxiety.

Additionally, our UWD-subjects showed superior performance on full-blown dynamic facial fear recognition. Such counterintuitive functional improvement associated with brain damage may reflect 'paradoxical functional facilitation', which refers to the fact that brain lesions sometimes can result in improved behavioral performance (Kapur, 1996). This mechanism can be explained by considering the dynamic and active interplay of excitatory and inhibitory connections within neural circuits. When structure-A contributes to function-X, and structure-B inhibits structure-A, loss of structure-B will relieve the inhibition of structure-A resulting in improvement of function-X. Following our argument on reduced inhibition of fear vigilance in our UWD-subjects, dynamic fearful faces could also evoke hyper-vigilance. It is indeed well established that passive viewing of fearful faces evokes simultaneous autonomic responses and amygdala activity (Williams et al., 2005; Williams et al., 2006; Williams et al., 2001). We therefore argue that failure of the BLA to inhibit these basal fear responses may engender up-regulation of attentional vigilance mechanisms, and therefore hyper-vigilance to emotionally salient areas of faces, as seen in the increased fixation duration to, especially fearful, eyes. This increased processing of the eye-region, might consequently result in the here observed paradoxical functional facilitation (Kapur, 1996), in terms of improved fear recognition (Adolphs et al., 2005).

Although future research should confirm whether fearful faces also evoke stronger autonomic responses in these subjects, the high prevalence of comorbid anxiety disorders including panic disorder observed in UWD (Thornton et al., 2008; Wiest et al., 2006), might be attributable to the same disinhibition phenomenon we propose, but in these cases causing secondary psychopathology. Contrariwise, lack of and hypo-attention for fear, as seen in UWD-subject S.M. (Adolphs et al., 2005; Feinstein, Adolphs, Damasio, & Tranel, 2011), might be due to the fact that her entire amygdala is damaged. As has also been demonstrated in rodents (Davis & Whalen, 2001), and primates (Amaral, 2003; Kalin et al., 2004; Kalin, Shelton, Engeland, Haraldsson, & Marucha, 2006), full amygdala damage can result in an inability to evaluate threats as salient, which might also explain this UWD-subject's inability to automatically allocate attention to emotional salient information (Adolphs et al., 2005), while this function is fully intact in our UWD-subjects.

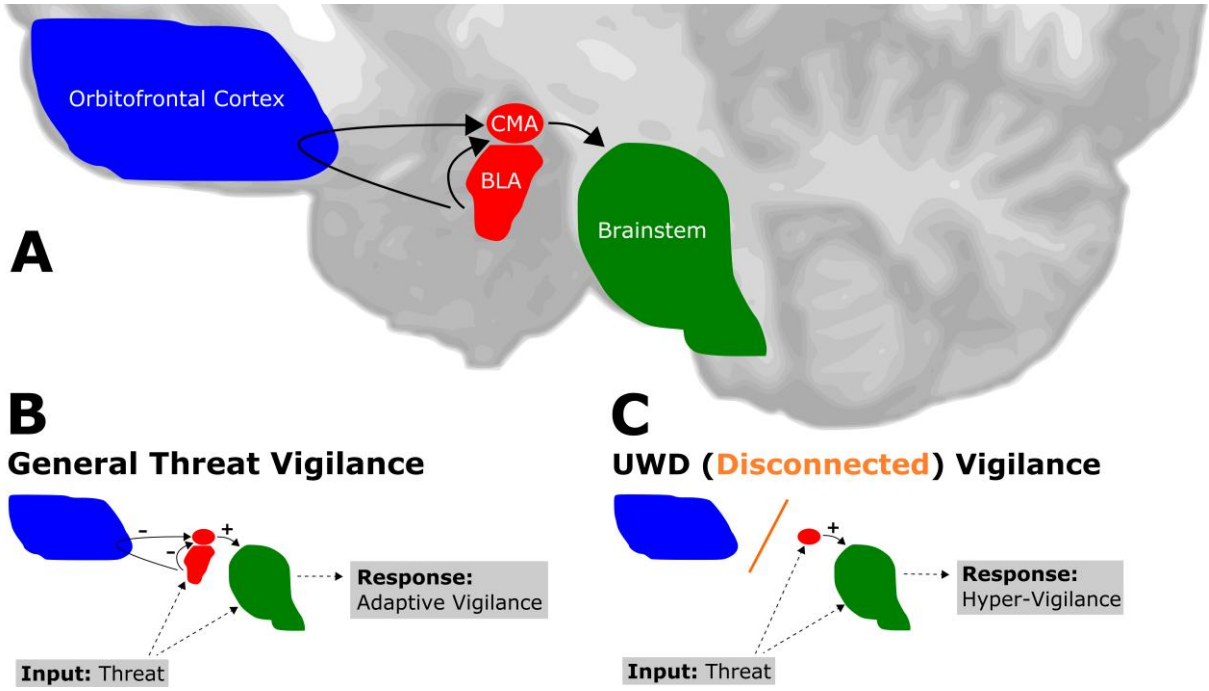


Figure 3.7 Heterogeneous amygdala model of threat vigilance in healthy humans and UWDs. **A)** Brain structures involved; Orbitofrontal cortex (OFC), central-medial amygdala (CMA), basolateral amygdala (BLA), and Brainstem. **B)** Normal threat vigilance; BLA and OFC jointly regulate the output of the CMA to the Brainstem resulting in controlled and adaptive threat vigilance. **C)** Threat vigilance in UWDs; Damage to the BLA disconnects the CMA from the regulatory functions of the BLA and OFC resulting in hyper-vigilance to threat.

Conclusion

Using a multimodal research strategy involving structural and functional MRI as well as eye-tracking and behavioral assessments, we show that five subjects with bilateral focal damage selective to the BLA, are hyper-vigilant for fearful facial expressions. Our lesion data are unique in being translational to recent rodent studies (Gozzi et al., 2010; Macedo et al., 2005; Macedo et al., 2007; Macedo et al., 2006; Martinez et al., 2007; Tye et al., 2011), and provide the first direct evidence in support of the hypothesis that the human BLA inhibits acute responses of hyper-vigilance to innate threat cues. These findings have important implications for the understanding of heterogeneous amygdala functions, and especially the role of the basolateral amygdala, in the disorders of fear and anxiety.

Chapter 4

In the eye of the beholder: Reduced threat-bias and increased gaze-imitation towards reward in relation to trait anger

Based on: Terburg, D., Aarts, H., Putman, P. & van Honk, J. (2012)
PLoS ONE, 7, e31373

Abstract

The gaze of a fearful face silently signals a potential threat's location, while the happy-gaze communicates the location of impending reward. Imitating such gaze-shifts is an automatic form of social interaction that promotes survival of individual and group. Evidence from gaze-cueing studies suggests that covert allocation of attention to another individual's gaze-direction is facilitated when threat is communicated and further enhanced by trait anxiety. We used novel eye-tracking techniques to assess whether dynamic fearful and happy facial expressions actually facilitate automatic gaze-imitation. We show that this actual gaze-imitation effect is stronger when threat is signaled, but not further enhanced by trait anxiety. Instead, trait anger predicts facilitated gaze-imitation to reward, and to reward compared to threat. These results agree with an increasing body of evidence on trait anger sensitivity to reward.

Introduction

Primate and especially human social interaction depend heavily on non-verbal communication with the eyes (Emery, 2000). The elongated width and extreme whiteness of the sclera are indeed unique features of the human eyes, argued to have evolved to facilitate such social communication (Kobayashi & Kohshima, 2001). Interestingly, following the gaze of others is reflexive and can therefore be regarded as adaptive behavior crucial to survival (Frischen, Bayliss, & Tipper, 2007; Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002). Detection of, and attending to threat are evidently adaptive behaviors. Accordingly, facial expressions (Vuilleumier, 2002) as well as gaze-direction (Langton et al., 2000) are processed automatically and preconsciously.

Humans and other primates actively follow observed gaze-shifts (Emery, 2000; Frith, 2008), but although it would provide a unique insight in reflexive and adaptive social behavior, it has not yet been experimentally studied how facial expressions influence these gaze imitations. It is however known that faces with averted gaze are labeled faster and more often as fearful, while the opposite holds for happy faces (Adams & Kleck, 2003, 2005). Moreover, facial expressions can give relevance and meaning to the gaze-shift with regard to mental state and social environment (Itier & Batty, 2009). For example, a happy gaze-shift may signal a potential reward, while a frightened gaze-shift can alert for potential threat. Although the latter is often considered to be more crucial to survival (Frischen et al., 2007), studies on attentional cueing by observed gaze-shifts, or gaze-cueing, have struggled to find general effects of facial expression (Hietanen & Leppänen, 2003). More recent studies revealed however a threat bias in gaze-cueing by fearful faces (Holmes, Richards, & Green, 2006; Tipples, 2006), but there is also evidence that this is exclusive to high anxious individuals (Fox, Mathews, Calder, & Yiend, 2007; Mathews, Fox, Yiend, & Calder, 2003). Studies using more ecologically valid dynamic facial stimuli confirmed that the threat bias in gaze-cueing is strongest in high anxious individuals, but also showed reliable general effects of facilitated gaze-cueing by fearful compared to happy facial expressions (Putman, Hermans, & van Honk, 2006; Putman, Hermans, & van Honk, 2010; Putman, Saevarsson, & van Honk, 2007).

Although these studies provide valuable information on gaze-cueing of covert attention, their generalizability to real-life social behavior is limited because participants are

instructed to refrain from making gaze-movements, and the measures of interest (e.g. button-presses and symbol-identification) are non-adaptive behavioral responses. The natural response to a gaze-shift is however to actively follow it, which is an adaptive feature of primate (Emery, 2000) and human (Frith, 2008) behavior, already observed in new-borns between 1 and 3 days old (Farroni et al., 2004b).

Studies on overt gaze-cueing, or 'gaze-imitation', in adults are scarce, but confirm that the preparation of gaze-imitation saccades is reflexive (Ricciardelli et al., 2002). Unlike reflexive covert shifts of attention, however, the actual execution of these eye-movements can be inhibited and are therefore prone to top-down modulation (Frischen et al., 2007; Koval, Thomas, & Everling, 2005). Importantly, although threat detection in gaze-cueing paradigms is enhanced in relation to anxiety (Fox et al., 2007; Mathews et al., 2003), anxiety is also strongly related to threat avoidance (Bogels & Mansell, 2004), particularly in relation to eye movement responses (Garner, Mogg, & Bradley, 2006). The anxious priority for threat in reflexive gaze-cuing might therefore not simply be applicable to the overt case.

In relation to trait anger, on the other hand, no such threat avoidance should be expected. Moreover, trait anger apparently is highly predictive for social aggression, which is marked by reduced sensitivity to the victim's fearful expression (see (Marsh & Blair, 2008) for a review). In strong agreement, trait anger is related to reduced amygdala reactivity when perceiving fearful faces (Carlson, Greenberg, & Mujica-Parodi, 2010). Additionally, trait anger has repeatedly been linked to reward-sensitivity and approach motivation (Carver, 2004; Harmon-Jones, 2004). Accordingly, the motivational drive to follow a gaze-shift might be decreased for fearful, but increased for happy cues, because the latter signals a peripheral reward.

Affective modulation of overt gaze-imitation by cues of threat and reward has not yet been experimentally studied. Therefore, we developed a new gaze-imitation task that closely resembles a situation wherein someone actively shifts gaze to a rewarding or threatening location. Participants watched video-clips of faces shifting gaze in a happy or fearful manner, and responded by gazing as fast as possible to a target appearing in the gaze-signaled, or opposite, location. This paradigm allowed us to assess whether imitative gaze-shifts are facilitated towards threat or reward and how this interacts with personality traits of anger and anxiety.

We expected faster gaze-allocation when an observed gaze-shift was imitated and further facilitation when threat was signaled with a fearful expression. Furthermore, in light of the enhanced threat detection (Putman et al., 2006) and threat-avoidance (Bogels & Mansell, 2004; Garner et al., 2006) in relation to anxiety, the positive relations between trait anxiety and covert fear-gaze cueing (Fox et al., 2007; Holmes et al., 2006; Mathews et al., 2003; Putman et al., 2006) might not be observed here. A happy gaze-shift, on the other hand, signals a potential peripheral reward. Since trait anger is associated with increased reward sensitivity (Carver, 2004; Harmon-Jones, 2004), we predict that individuals high in trait anger are relatively more motivated to follow a happy gaze-shift, which should reduce the expected priority for gaze-imitation towards threat over reward.

Methods

Ethics statement

The research reported in this article involves healthy human participants, and does not utilise any invasive techniques, substance administration or psychological manipulations. Therefore, compliant with Dutch law, this study only required, and received approval from our internal faculty board (Human Biopsychology and Psychopharmacology) at Utrecht University. Furthermore, this research was conducted, and written informed consent of each participant obtained, according to the principles expressed in the Declaration of Helsinki.

Participants and procedure

Twenty healthy volunteers (all students, age-range 18-25 years, 9 female) received course credit or a monetary reward to participate in the experiment. Stimuli and design were adapted from Putman and colleagues (2006) and consisted of video-clips of centrally presented faces changing rapidly (120 ms) from neutral to either happy or fearful, while the eyes simultaneously moved from central to peripheral gaze (left and right). The final frame was maintained for an additional 80 ms, after which the face disappeared and in 2/3 of the trials, a target appeared either to the left or right (10^{o} visual angle) of the face.

For the video-clips 8 different actors (4 female), with 2 emotions (happy and fearful) and 2 gaze-directions (left and right), were used (Ekman & Friesen, 1976; Lundqvist et al.,

1998), making 32 unique stimuli (see (Putman et al., 2006) for further details). These were presented 6 times each; twice with a target at the same location as the gaze-shift (valid trial), twice with a target at the opposite location to the gaze-shift (invalid trial) and twice with no target to avoid habitual saccade preparation (catch trial). These made a total of 192 trials, counterbalanced for emotion and condition, and presented in random order. Preceding the task, nine trials were presented for practice, using the same stimuli with gaze-, but without shift of emotion.

Participants were instructed to shift their gaze towards the target, and were explicitly, and correctly, informed that gaze-direction of the presented face did not predict target appearance or location. Responses were made with a shift of gaze to the target, which disappeared when the eye-track computer detected that the target was reached. Stimulus presentation commenced when the participant gazed at a fixation-cross, positioned where the eyes of the subsequently presented faces would appear, for a random time between 1000 ms and 1500 ms to avoid timing habituation. During the catch trials, wherein no target appeared, gaze had to be maintained at the fixation position until start of the next trial (see **Figure 4.1** for a visual representation of the task). Beforehand, participants completed trait anxiety and anger (STAI/STAS) questionnaires (Spielberger, Gorusch, & Lushene, 1970; Spielberger, Jacobs, Russel, & Crane, 1983).

Apparatus and analyses

For the present study we are primarily interested in gaze-shifts, as this is the most natural way of overt orienting. A gaze-shift consists of an eye-movement, and a simultaneous, but small, head-movement (Guitton, 1992), which is restricted by most eye-track systems using head-fixation. The gaze-imitation task was therefore presented, and gaze-data recorded, using a Tobii-1750 binocular eye-tracker with integrated TFT-display, 8 ms response time, 50 Hz sampling-rate and 0.5^{o} accuracy (Tobii Technology, 2006). With this eye-track system head-fixation is not necessary, which allows for relatively unrestricted gaze responses.

Latency of the gaze-shifts was defined as the time between onset of, and first gaze-point within 1^{o} of the target. Trials with latencies shorter than 100 ms or longer than 1200 ms (0.4%) were removed from analysis. Mean latencies were computed for all 4 conditions (threat/reward x valid/invalid), and were used in three analysis steps. First, we assessed

overall and emotion-specific gaze-imitation effects. Thereto, mean latencies were entered as within-subject variables in a 2x2 repeated-measures ANOVA, followed by paired-samples *t*-tests.

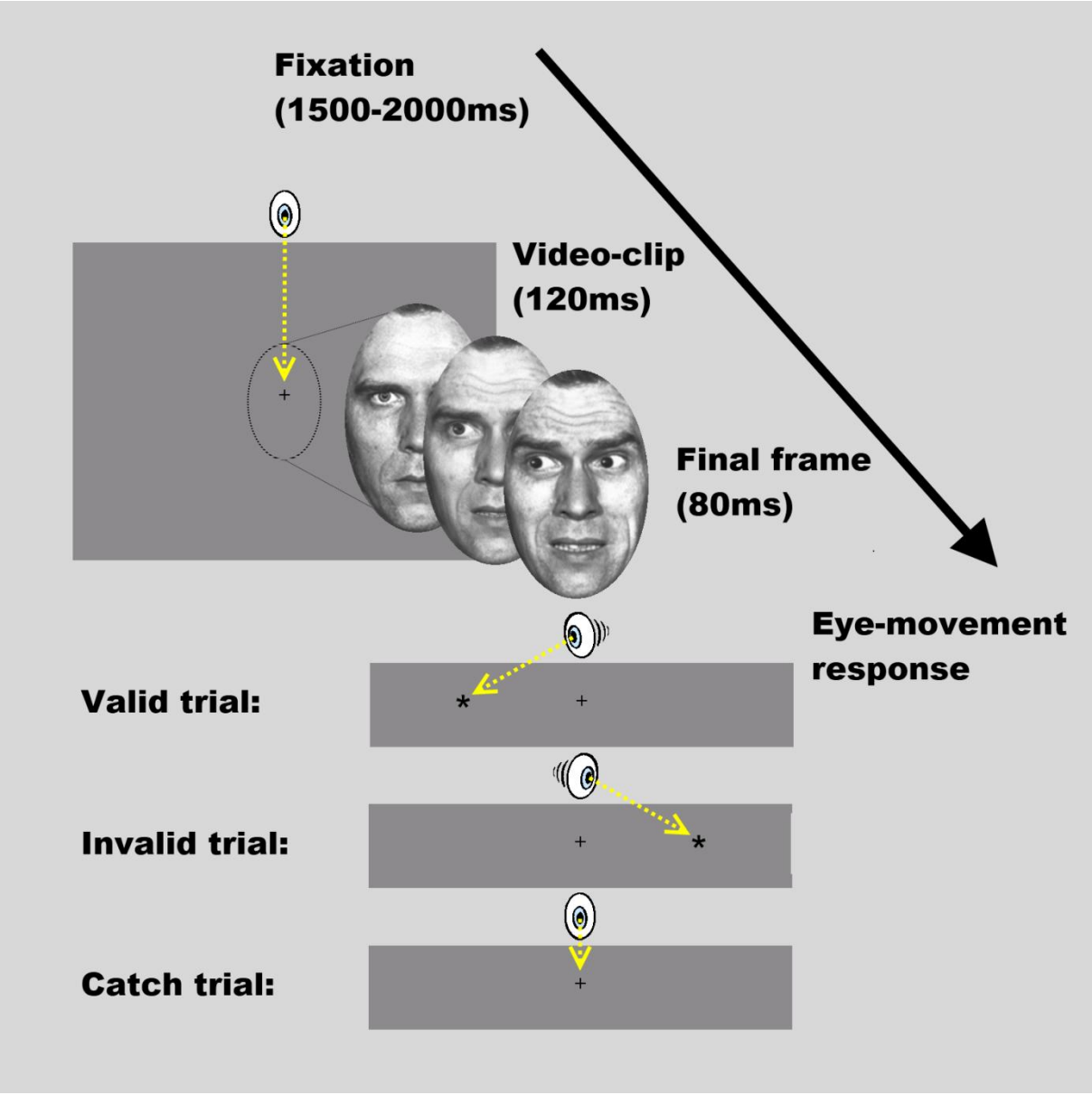


Figure 4.1 Visual representation of the gaze-imitation task. After gaze-fixation participants watched video-clips wherein faces fluently shifted from neutral to fearful or happy expressions, while the eyes shifted from center to left or right. Participants were instructed to allocate their gaze as fast as possible to the target that appeared on the left or right side of the screen when the clip ended. One-third of the trials was valid (target in same location as stimulus gaze-shift), one-third invalid (target in opposite location of stimulus gaze-shift) and one-third catch (without target, thus without eye-movement response). The example stimulus shown here was adapted from the Pictures of Facial Affect database (Ekman & Friesen, 1976).

Second, we assessed the emotion-specific influence of the personality characteristics STAI and STAS on gaze-imitation. STAI and STAS scores were correlated with gaze-imitation biases computed for both emotions separately by subtracting the average latencies on the valid from the invalid trials. These contrasts provide a reliable measure of gaze-imitation, because if gaze-imitation is a reflexive mechanism, this would affect both conditions in opposite direction; i.e. gaze-shifts will be facilitated in the valid trials, and delayed in the invalid trials. Furthermore, these bias-scores represent emotion-specific indices of gaze-imitation without confounding effects of between-subjects variability in overall reaction speed, whereby higher values represent stronger effects of gaze-imitation.

Third, we assessed how STAI and STAS influenced gaze-imitation towards reward compared to threat. In classic attentional-cueing (Frewen, Dozois, Joanisse, & Neufeld, 2008), and covert gaze-cueing experiments (Putman et al., 2006), such top-down modulation is often described in terms of engagement and disengagement. The first applies to the valid trials only, and is a measure of how fast attention is directed towards a peripheral target, whereas the second applies to the invalid trials as a measure of how fast one can disengage attention from a peripheral location. For direct assessment of the effect of personality characteristics on the difference between gaze-imitation towards threat and reward, however, we are primarily interested in how the imitative gaze-shift (i.e. the engagement component) is modulated, because this constitutes the top-down influence on actual gaze-imitation. Moreover, the disengagement component, or the shift of gaze in the opposite direction to an observed gaze-shift, involves suppression and inversion of the initial gaze-imitation reflex. In other words, while disengagement in gaze-cueing studies is a purely attentional mechanism, in a gaze-imitation task it would involve inhibition of reflexive motor-responses (Nummenmaa & Hietanen, 2006). A reliable assessment of between-emotion differences in disengagement would therefore involve in-depth saccade analysis to identify these, likely small, erroneous saccades. The gain of minimal movement restriction, provided by the use of the Tobii-1750 eye-tracker, came however with the cost of a relatively low sampling-rate of gaze-data, which does not allow for such analyses. Therefore we assessed the top-down influence on gaze-imitation towards threat compared to reward only in the valid condition. STAI and STAS were thereto correlated with threat/reward bias scores computed by

subtracting the average latencies on the valid-fear trials from the valid-happy trials. Thus, higher values represent a gaze-imitation bias for threat relative to reward.

In sum, we first assessed overall gaze-imitation and the difference between gaze-imitation towards threat and reward. Next, we assessed modulation of gaze-imitation by STAI and STAS through contrasting valid and invalid trials for each emotion. Finally, we assessed the effect of these two personality traits on the actual affective modulation of gaze-imitative gaze-shifts by computing their correlation with the contrast of threat and reward trials in the valid condition. All reported statistics are conducted with two-sided $\alpha = 0.05$.

Results

Mean latencies of gaze-allocations for all four conditions are shown in **Table 4.1**. We found a significant effect of validity ($F(1,19) = 19.253$, $p < .001$, $\eta_p^2 = .503$), and a significant interaction of validity and emotion (fear/happy) showed that the validity effect, or gaze-imitation, was reliably stronger in the fear compared to happy condition ($F(1,19) = 7.680$, $p < .05$, $\eta_p^2 = .288$, see **Table 4.1** and **Figure 4.2**). Separate paired-sample *t*-tests confirmed reliable gaze-imitation effects for both the fearful (16 ms faster in valid trials, $t(19) = 4.084$, $p < .001$) and happy (6ms faster in valid trials, $t(19) = 3.083$, $p < .01$) conditions. Furthermore, the main effect of emotion was significant for the valid condition (8 ms faster in fearful trials, $t(19) = 2.109$, $p < .05$), but not for the invalid condition (2 ms slower in fearful trials, $t(19) = -.604$, $p = .553$). This confirms that in the valid trials, where the observed gaze-shifts are imitated, gaze-shifts were faster when the observed gaze-shift was accompanied with a fearful expression.

Table 4.1 Mean latencies (with standard deviation) of gaze-allocation for each condition in the gaze-imitation task.

	Threat (Fearful Face)	**Reward (Happy Face)**
Valid	281 (*27*) ms	289 (*32*) ms
Invalid	297 (*34*) ms	295 (*31*) ms

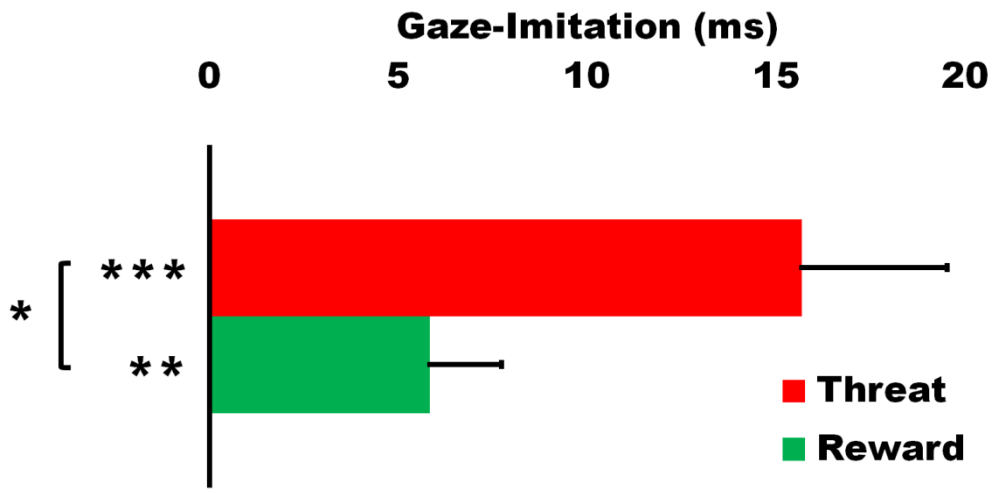


Figure 4.2 Gaze-imitation effects for the threat (fearful faces) and reward (happy faces) conditions. Values represent mean latencies of gaze-allocation in invalid minus valid trials. The gaze-imitation effect is significant in both conditions and significantly stronger in the threat condition. Error-bars represent SEM. * = $p < .05$, ** = $p < .01$, *** = $p < .001$

The correlational analysis showed that trait anxiety (STAI) and trait anger (STAS) were not significantly related in our subject sample ($R = .19$, $p = .416$). Furthermore, STAI was not related to gaze-imitation, as indexed by the contrast of invalid minus valid trials, in the fear ($R = -.08$, $p = .737$) and happy ($R = .15$, $p = .519$) conditions. STAS was as predicted significantly related to increased gaze-imitation in the happy ($R = .54$, $p < .05$), but not in the fear ($R = .23$, $p = .324$) condition. Finally, as predicted, STAS was strongly related to a reduced fear/happy bias in the valid condition ($R = -.58$, $p < .01$, see **Figure 4.3**), while for STAI there was no significant relation ($R = .16$, $p = .492$). In sum, gaze-imitation is not directly modulated by STAI, but STAS is associated with greater gaze-imitation towards reward as signaled by happy facial expressions, and with a reduced gaze-imitation bias towards threat as signaled by fearful compared to happy facial expressions.

Discussion

In this study we show that allocation of gaze is faster when the gaze-shift of someone else is imitated. Moreover, when the observed gaze-shift is accompanied with a dynamic fearful

expression, which communicates a peripheral threat, the gaze-imitation effect is stronger than when peripheral reward is signaled with dynamic happy gaze-shifts. As predicted, this threat-bias was strongly reduced in relation to heightened trait anger, but unrelated to trait anxiety.

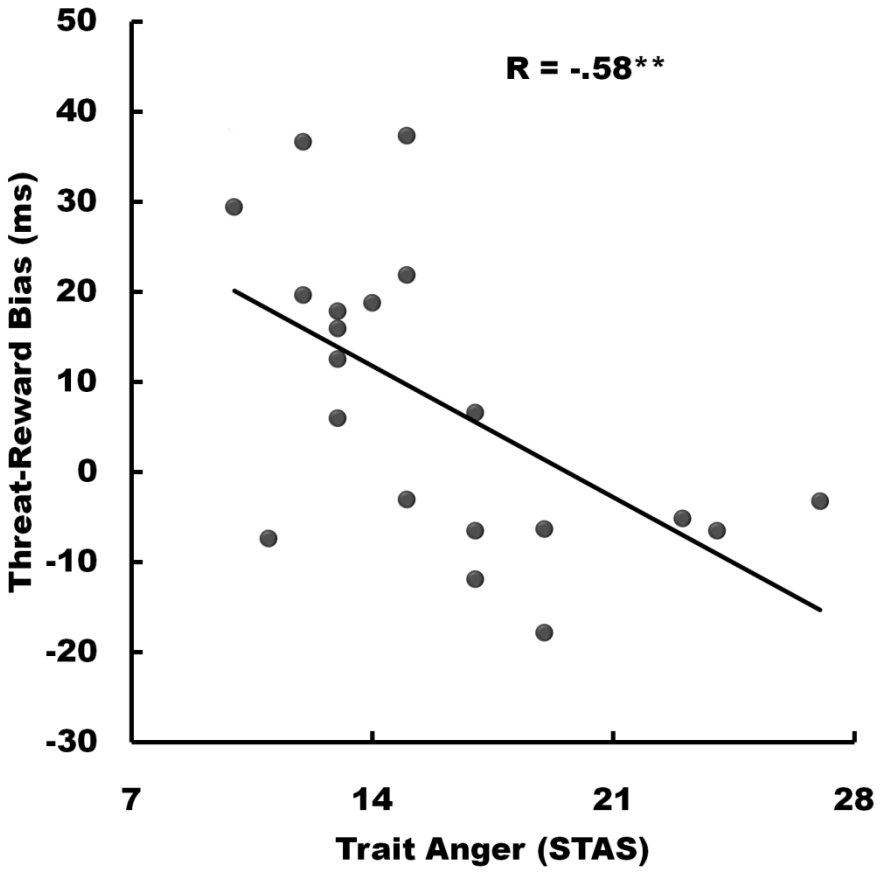


Figure 4.3 Linear relation of trait anger (STAS) with the threat-reward bias. High values represent a stronger gaze-imitation effect towards threat.

Firstly, these results replicate the findings of Ricciardelli and colleagues (2002), who found facilitated allocation of gaze in the direction of observed (neutral) gaze-shifts. Secondly, the threat-bias in gaze-imitation concurs with the literature on gaze-cueing, and is arguably an adaptive reflex. The biological underpinnings of this reflex might be found in the amygdala's involvement in the processing of both gaze and emotional expression. Direction of gaze is processed in the superior temporal sulcus (STS), which projects both to intraparietal areas for subsequent allocation of attention, as well as to the amygdala (Emery, 2000; Frischen et al., 2007). Moreover, the amygdala is automatically activated by threat, and

fearful faces in particular (Adolphs, 2008), and has both direct and indirect influence on the allocation of attention towards threat (Vuilleumier, 2005). STS-amygdala interactions might therefore underlie the integration of affective value and gaze-direction (Itier & Batty, 2009), and thus the present reflexive modulation of gaze in response to threat.

On top of this emotional modulation, we show here that the gaze-imitation bias for threat compared to reward is reduced in relation to trait anger. As mentioned in the introduction, trait anger has repeatedly been related to increased reward sensitivity (Carver, 2004; Harmon-Jones, 2004). A happy gaze-shift as signal of potential peripheral reward may therefore carry high motivational value for those high in anger, which might have resulted in the here found increase of gaze-imitation towards reward in relation to trait anger. It must however be noted that, based on the present data, we cannot entirely exclude that trait anger also reduces gaze-imitation towards threat. Indeed, trait anger is associated with reduced amygdala activity when perceiving fearful faces (Carlson et al., 2010), and reduced sensitivity for fearful facial expressions, which is argued to underlie social aggression (Marsh & Blair, 2008). The present data are however in favor of increased reward-sensitivity in relation to trait anger, and we therefore assume that the trait anger shift from threat to reward in gaze-imitation is driven by angry individuals imitating happy gaze more strongly, thereby reducing the general imitation bias for fearful gaze-shifts.

Our results furthermore show that trait anxiety has no direct relation to the emotional modulation of gaze-imitation. In the light of recent evidence that covert gaze-cueing towards threat is enhanced in relation to anxiety (Holmes et al., 2006; Putman et al., 2006; Tipples, 2006), and sometimes even exclusive to anxiety (Fox et al., 2007; Mathews et al., 2003), this is an intriguing finding. Crucially, Putman and colleagues (2006) confirmed, with the exact same stimuli and design in a study on covert gaze-cueing (i.e., with button-press on target-detection), enhancement of the threat-bias in relation to trait anxiety. Apparently anxiety facilitates target-detection when an observed gaze-shift indicates that it might be a threat, but does not facilitate overt responding towards the threat. Importantly, in the present paradigm covert target-detection always precedes the actual gaze-shift, and it therefore seems that the increase in threat-detection speed in relation to anxiety is somehow counteracted during the subsequent overt response.

These contrasting effects might be explained by the fact that overtly gazing at a threat can be distressing, which is an essential feature of the vigilance-avoidance hypothesis of anxiety. The vigilance-avoidance hypothesis (Bogels & Mansell, 2004), predicts that increased vigilance facilitates detection of threat in anxious individuals, but that this threat is subsequently defensively avoided to reduce internal distress. Indeed, as already noted, the overt gaze-imitation reflex can be inhibited (Koval et al., 2005), thus the increased speed of covert target-detection when a gaze-shift indicates that it is a possible threat , might be reflexively counteracted by anxious avoidance mechanisms. Speculatively, in the case of gaze-imitation, anxious individuals put the attentional system in reverse after a threat has been detected, in order to avoid confrontation, and reduce internal distress. A limitation of the present study is however that we did not assess target-detection and overt responding separately. Whether gaze-shifts towards threat are indeed counteracted, or maybe simply not affected by anxiety, is therefore something that should be tested in future research.

Another limitation of the present study is that the design did not allow for a neutral baseline measure. Although the correlational analysis shows us that the reduced threat-bias in relation to trait anger is most likely the result of increased gaze-imitation towards reward, we cannot exclude that gaze-imitation towards threat might also be reduced. Both interpretations are supported in the literature (Carlson et al., 2010; Carver, 2004; Harmon-Jones, 2004; Marsh & Blair, 2008), and future research on gaze-imitation should therefore address this issue.

In summary, allocation of gaze is reflexively facilitated when an observed gaze-shift is imitated. When someone gazes away fearfully, signaling a potential threat, this gaze-imitation effect is stronger. Moreover, we provide evidence that trait anger shifts this threat-bias towards relatively stronger imitation of happy facial cues; i.e. a shift in the sensitivity for threat towards reward. Additionally, in line with the vigilance-avoidance hypothesis we speculate that trait anxiety induces conflict between facilitated covert threat-detection and overt threat-avoidance. Finally, the study of actual gaze-behavior appears to be an ecologically valid method to promote the understanding of the mechanisms behind real-life gaze following behavior in relation to anxiety and anger. Taken together with the large body of work accumulated in recent years on covert attentional mechanisms, the study of

interactive overt social gaze-behavior can importantly contribute to psychology and neuroscience.

Chapter 5

Memory and attention for social threat: Anxious hypercoding-avoidance, and submissive gaze-aversion

Based on: Terburg, D., Aarts, H. & van Honk, J. (2012)

Emotion, in press

Abstract

Rivalry for dominance is a recurrent challenge in human social interaction. During these social dominance interactions some people rapidly break eye-contact, whereas others merely try to avoid such eye-to-eye confrontations. The first is an example of submissive gaze-aversion, whereas the second reflects anxious gaze-avoidance. We tested these distinct forms of gaze-behavior within a social-memory setting, and show that anxious individuals vigilantly attend to, superiorly remember, and subsequently avoid, social threats (i.e. angry faces). Furthermore, submissive individuals, as indexed by high trait anxiety and low trait anger, exhibit rapid gaze-aversion from facial anger. Mechanisms of hypervigilance-avoidance thus seem to underlie natural gaze-behavior and enhanced memory for threat in anxiety. Accordingly, we propose the term hypercoding-avoidance, which describes how anxious individuals habitually scan their immediate social environment for threat, remember its location, and subsequently avoid it. Moreover, this is the first experimental evidence showing that submissive gaze-aversion is distinct from anxious gaze-avoidance.

Introduction

Challenges for dominance are important and adaptive in human social interaction. Some people do, at the outset, seek social interaction, but readily submit to the wishes of dominant opponents. Others are too afraid to face such challenges and simply avoid social interactions for dominance. To properly function in the social system, it is crucial to know the intentions and affective states of others. Rapid detection of social hostility, and superior memory for its location, would greatly facilitate the prevention of social confrontation and subsequent aggression. This type of behavior can also be observed during the formation and maintenance of primate dominance hierarchies. Subordinate individuals tend to keep track of the location of dominant conspecifics, but refrain from looking directly at them in a sustained manner (Mazur, 1985; Setchell & Wickings, 2005). When two primates do establish eye-contact, a staring-contest may arise, wherein the subordinate will avert their gaze in order to prevent provoking an aggressive confrontation (Mazur & Booth, 1998). We can thus distill two distinct forms of subordinate gaze-behavior. First, avoidance of direct gaze to the dominant threat (gaze-avoidance), because this might prevent potentially dangerous eye-contact. Second, once such eye-contact does occur, the submissive animal rapidly averts gaze (gaze-aversion), thereby signaling subordination and preventing aggressive confrontation. In the present research, our aim is to examine human attention and memory in a hostile social context to better understand their interrelationship, and to elucidate why individuals either avoid making eye-contact, or submissively avert their gaze.

Social anxiety, which is characterized by fear of social evaluation (Watson & Friend, 1969), is strongly related to reduced dominance behavior or submissiveness (Trower & Gilbert, 1989; Weeks et al., 2011). Cognitive biases for facial expressions have been widely studied in relation to social anxiety, and although attentional and memory biases for faces have been reported, these biases are not always emotion specific, and mediated by state anxiety (Heinrichs & Hofmann, 2001; Staugaard, 2010). Nonetheless, rapid avoidance of facial threat has repeatedly been shown in social anxiety (Bogels & Mansell, 2004; Putman et al., 2004). On the other hand, there is also abundant evidence that anxious individuals rapidly detect and attend to threat (see for a review; Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007). In an attempt to clarify and combine these seemingly opposing findings, researchers proposed the hypervigilance-avoidance theory of anxiety

(Mogg et al., 1997). This theory states that anxious individuals rapidly detect threat cues because they are hypervigilant for danger, but subsequently avoid it in order to reduce internal distress and defensively prevent confrontation (Bogels & Mansell, 2004). Recent studies that directly tapped into visual attention and gaze-behavior do indeed support vigilance-avoidance mechanisms. For instance, anxiety predicted vigilant eye-movements towards threatening faces (Bradley, Mogg, & Millar, 2000; Mogg, Millar, & Bradley, 2000), but also subsequent avoidance of emotional faces compared to objects and neutral faces (Garner et al., 2006). Moreover, anxiety predicted longer visual scan-paths (vigilance), but reduced gaze to the eyes (avoidance) of especially angry faces (Horley, Williams, Gonsalvez, & Gordon, 2004). Following hypervigilance-avoidance theory we therefore predict that anxious individuals in a social group will rapidly detect angry others, but will subsequently avoid looking at them.

Although hypervigilance-avoidance theory successfully combined the cognitive attentional biases towards and away from threat, it has not yet been related to the profound memory bias for threatening information found in anxiety, which is most consistently found when facial stimuli are used, and recall rather than recognition of threatening information is measured (Mitte, 2008). This prevalence for recall of threatening information might reflect the recurring intrusive memories observed in clinical anxiety, which together with attentional biases for threat have been argued to contribute to the maintenance and increase of anxiety levels (Eysenck, 2004). Therefore we will aim in this study to relate levels of anxiety to both vigilance-avoidance mechanisms, as well as a predisposition to recall threatening information.

Crucially, the hypervigilance-avoidance theory does not predict the rapid gaze-aversion of submissive animals in the primate-example from above. Gaze-aversion from eye-contact with an angry opponent is a dominance-submissiveness social interaction, whereas the hypervigilance-avoidance theory explicitly deals with avoidance of such social interactions. In humans, submissive gaze-aversion, in terms of rapid attentional avoidance of subliminally presented angry faces, has indeed been associated with social anxiety and high levels of the stress-hormone cortisol, both markers of a submissive stance (Putman et al., 2004; van Honk et al., 1998). Contrariwise, vigilant attention to these social threats correlated positively with both self-reported and biological markers of dominance; i.e. approach motivation, trait anger, and testosterone levels (Putman et al., 2004; van Honk et al., 2001a; van Honk et al., 1999; Wirth & Schultheiss, 2007). Recently we found critical converging evidence with genuine

measures of eye-contact and gaze-aversion: Both self-reported dominance motives and biological manipulation with testosterone administration reflexively slowed down gaze-aversion from subliminally presented angry stares (Terburg et al., 2012c; Terburg et al., 2011). These effects therefore imply that rapid gaze-aversion from facial threat is related to submissiveness, which is a form of reflexive subordination not predicted by the vigilance-avoidance hypothesis.

Recently it has been argued that such gaze-aversion could indeed be a separate social mechanism from gaze-avoidance (Garner et al., 2006). Garner and colleagues proposed that socially anxious individuals will show an attentional pattern of vigilance-avoidance to threat when stimuli compete for attention, but will show defensive disengagement, or gaze-aversion, when submissiveness in a social confrontation is possible. Given the large overlap between social anxiety and submissiveness (Trower & Gilbert, 1989; Weeks et al., 2011), we therefore argue that vigilance-avoidance attention to angry faces might be related to general anxiety, whereas immediate gaze-aversion after angry eye-contact is more related to submissiveness. Following this, we propose to make a clear distinction between gaze-avoidance and gaze-aversion, and use this to establish behavioral differences in anxiety and submissiveness. Within this framework anxious individuals initially orientate vigilantly to, but subsequently remember and avoid social threats. Submissive individuals will furthermore immediately give in to social challenges by means of rapid gaze-aversion from angry faces. This model reflects every-day behavior, wherein anxious individuals will remember and avoid social threats, but will not necessarily submit to the actual challenge. Only genuinely submissive individuals will also show rapid gaze-aversion from facial threat.

We tested these assumptions in a setting analogous to the primate group behavior described above. We developed a social memory task (SMT), wherein participants have 20 seconds to memorize the location of four angry faces and their neutral equivalents (see **Figure 5.1**). Memory performance is indexed by relocation accuracy, and the task's spatial organization allows us to measure visual attention (i.e., gaze-behavior) during encoding of such social information. We can thus simultaneously measure attentional processes related to gaze-aversion and avoidance, and memory for potentially threatening information (e.g. location of the angry faces).

Given that social anxiety and submissiveness are highly intertwined personality characteristics (Trower & Gilbert, 1989; Weeks et al., 2011), this construct cannot disentangle submissiveness from anxiety, as is the goal in the present study. Moreover, memory biases in relation to social anxiety are, although often observed in relation to facial stimuli (Heinrichs & Hofmann, 2001), not emotion specific, and seem to be affected by variations in state anxiety (Staugaard, 2010). The primate dominance/submission dimension can however also be reflected within the more general human affective domains of trait anxiety and anger (Barros & Tomaz, 2002), whereby the combination of high trait anxiety and low trait anger relate to submissiveness (Russell & Mehrabian, 1974; Smith, Traupman, Uchino, & Berg, 2010). For the present study we therefore concentrated on trait anxiety and anger. We expected trait anxiety to predict attentional vigilance-avoidance to angry faces, and simultaneously improve memory for their location. Trait anger is generally not related to such a memory-bias (Owen, 2010; Wilkowski & Robinson, 2007), but following the primate analogue, we expected submissive individuals, as indexed by high anxiety and low anger (Russell & Mehrabian, 1974; Smith et al., 2010), to avert their gaze more rapidly from angry faces.

Methods

Participants

Healthy volunteers (twenty females, twenty males, mean age = 22.7, *SD* = 2.8), with (corrected to) normal visual acuity, participated for either course-credit or payment. All participants provided written informed consent.

Stimuli and procedure

Stimuli in the SMT were gray-scaled faces of four actors (two females) with neutral and emotional (anger and happy) expressions (Ekman & Friesen, 1976; Lundqvist et al., 1998). In the encoding-phase eight faces were presented randomly positioned on a black screen, and participants were instructed to memorize each face's location. Presentation time was 20 seconds, and a gradually decreasing blue bar on top of the screen indicated how much time was left. After presentation the faces disappeared and immediately reappeared in random

order on top of the screen, while their original locations remained visible as white squares. Participants relocated the faces with the computer-mouse, and could correct themselves, without time-restrictions, until all faces were relocated (see **Figure 5.1**).

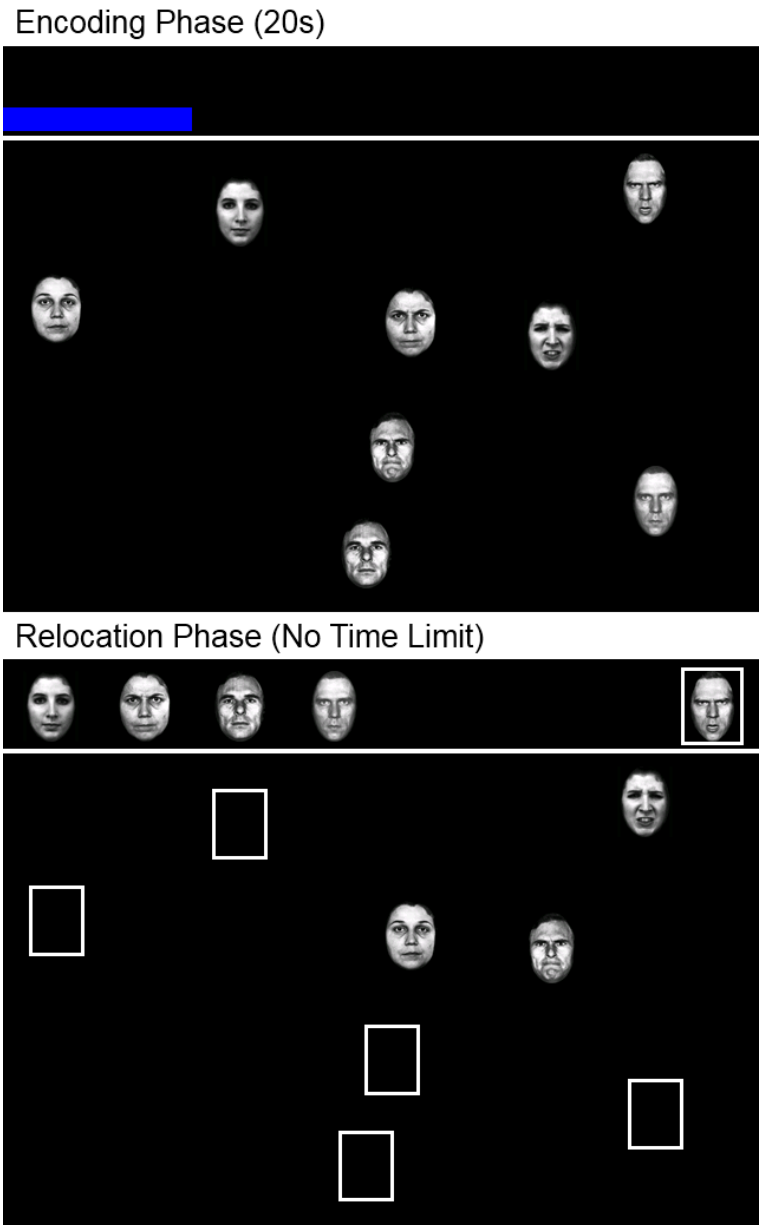


Figure 5.1 Example of a trial in the SMT. During the encoding phase (upper screenshot) participants have 20 seconds to memorize the location of the randomly presented faces, which are eight faces from four actors showing an emotional (happiness, or as in this example anger) and a neutral expression. In the subsequent relocation phase (lower screenshot) participants are instructed to relocate the faces that reappeared in random order at the top of the screen, using a computer-mouse. Eye movements were monitored during the encoding phase.

Only one emotion was used in each trial, hence each trial consisted of four angry or happy faces and their neutral equivalents. Participants performed in four trials of each emotion presented in counterbalanced blocked order, and filled-out the trait anxiety inventory

(STAI) (Spielberger et al., 1970) and trait anger scale (STAS) (Spielberger et al., 1983), beforehand.

Data acquisition and analysis

Memory performance was indexed by percentage of correctly relocated emotional and neutral faces in both conditions. Eye movements were recorded during the encoding-phase with a Tobii-1750 binocular infrared eye-tracker, sampling at 50 Hz, 0.5^{o} accuracy. Gaze-fixations were defined as the average location of all subsequent gaze-points within 1.5^{o} visual angle, with a minimal duration of 100 ms (Tobii Technology, Danderyd, Sweden).

To assess visual attention and its time-course we first divided the gaze-data into four time-blocks (0-5, 5-10, 10-15 and 15-20 seconds). Next, we filtered gaze-fixations not directed to the faces from the data (5.6%)[1], and computed emotional biases of visual attention by contrasting mean fixation duration (FD), and percentage of fixations (PF) to emotional versus neutral faces. FD and PF are independent measures of visual attention reflecting how often gaze is fixated (PF), and the average duration of these fixations (FD). Lower FD-values therefore represent faster gaze-aversion from, and lower PF-values represent gaze-avoidance of, the emotional compared to neutral faces. To test for effects of vigilance and subsequent avoidance of threat, we contrasted the PF-bias for anger in the first four fixations with the overall anger-bias in the remainder of each trial. For this analysis the first four fixations were selected because with eight faces presented in each trial, these provide a reliable amount of data, while minimizing the chance that the same face is gazed-at multiple times. Bias-scores were first standardized, and subsequently the anger-bias in the whole trial minus the first four fixations was subtracted from anger-bias in the first four fixation. Thus, positive values indicate an attentional pattern of vigilance-avoidance for angry compared to neutral faces. We first report the eye movement analyses, followed by memory performance. All reported statistics are conducted with two-tailed $\alpha = .05$.

[1] It must be noted that most of these filtered data come from the final time-block, wherein the participants often started to check how much time was left by looking at the gradually decreasing time-bar (see **Figure 5.1**).

Results

Questionnaires

Two participants (one female and one male) were excluded from analyses due to apparatus failure. For the remaining 38 participants the average STAI-score = 38.0, *SD* = 9.4, and STAS-score = 15.7, *SD* = 3.7, which correspond to average values found in healthy subject-samples (Spielberger et al., 1970; Spielberger et al., 1983). There were no gender differences on STAI, $t(36) = .24$, $p = .813$ or STAS, $t(36) = -.70$, $p = .489$, and the scales were not correlated, $r = .034$, $p = .840$. To facilitate the interpretation of the behavioral effects of trait anxiety we applied a median-split resulting in low-anxiety ($n = 20$, mean STAI = 30.7, *SD* = 2.4) and high-anxiety ($n = 18$, mean STAI = 46.0, *SD* = 7.4) groups.

SMT: Gaze (angry condition), Vigilance-avoidance (PF)

In the angry condition participants directed 49.7%, *SD* = 4.3, of their gaze-fixations to angry faces. Mean PF-values were subjected to a General Linear Model analysis with BLOCK (0-5, 5-10, 10-15 and 15-20 seconds) as within-subjects variable and STAI and STAS as continuous variables. STAI and STAS were standardized before analyses. Main effects for BLOCK, $F(3,102) = .76$, $p = .522$, and STAS, $F(1,34) = .28$, $p = .604$, were non-significant, which was also the case for all interaction effects (all p's > .14). As hypothesized, however, STAI significantly predicted PF, $F(1,34) = 4.46$, $p = .042$, $\eta_p^2 = .12$ (see **Figure 5.2**).

Next, we directly tested the hypothesis that trait anxiety predicted an attentional pattern of vigilance-avoidance towards angry faces, by correlating the measure of vigilance-avoidance (the contrast of PF in the first four fixations with PF in the rest of the trial, see **Methods**) with STAI. As predicted this resulted in a significant positive correlation, $r = .42$, $p = .008$, which indicates that STAI is associated with an attentional pattern of vigilance-avoidance to angry compared to neutral faces.

SMT: Gaze (angry condition), Gaze-aversion (FD)

In the angry condition average duration of gaze-fixations to angry faces was 432 ms, *SD* = 66, and to neutral faces 434 ms, *SD* = 62. Mean FD-values were subjected to a General Linear Model analysis with BLOCK (0-5, 5-10, 10-15 and 15-20 seconds) as within-subjects variable and STAI and STAS as continuous variables. STAI and STAS were standardized before

analyses. Main effects for BLOCK, $F(3,102) = 1.96$, $p = .124$, and STAS, $F(1,34) = .07$, $p = .891$, were non-significant. As hypothesized, however, STAI significantly predicted FD, $F(1,34) = 6.03$, $p = .019$, $\eta_p^2 = .15$ (see **Figure 5.2**), STAI interacted significantly with STAS, $F(1,34) = 4.49$, $p = .042$, $\eta_p^2 = .12$, and the three-way interaction of STAI x STAS x BLOCK also reached significance, $F(3,102) = 3.17$, $p = .027$, $\eta_p^2 = .09$.

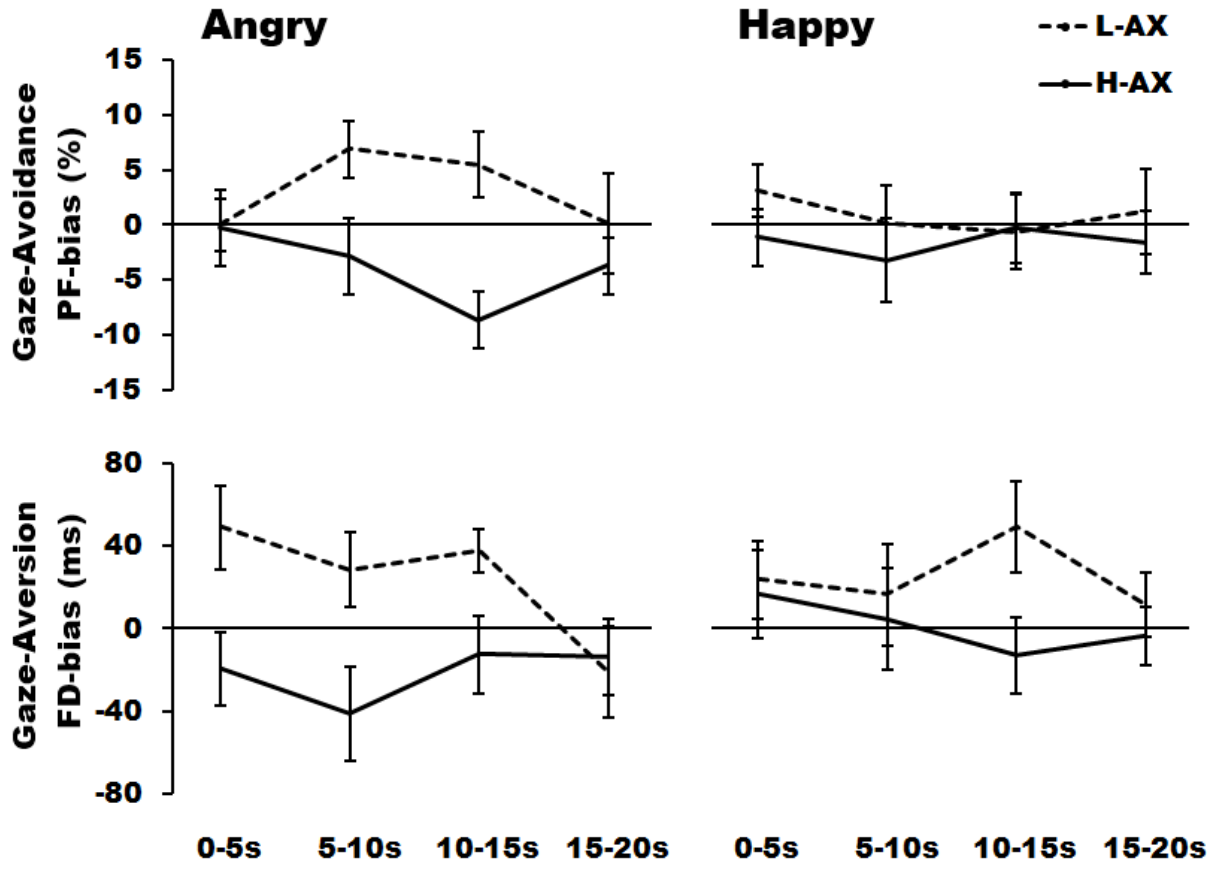


Figure 5.2 Raw eye-tracking data, emotional biases for high-anxiety (H-AX) and low-anxiety (L-AX) groups, defined by a median-split for visual purposes, for the separate time-blocks in the angry and happy condition. Lower values on the PF-bias and FD-bias represent an attentional bias away from the emotional compared to neutral faces in the number of fixations, and their average duration, respectively. Error-bars represent standard error of the mean.

To further specify these interaction effects, we tested the four time-blocks separately in a General Linear Model with standardized STAI and STAS scores as continuous variables. STAI significantly predicted FD in the first, $F(1,34) = 6.90$, $p = .013$, $\eta_p^2 = .17$, and third, $F(1,34) = 5.12$, $p = .030$, $\eta_p^2 = .13$, time-blocks, and near significantly in the second time-

block, $F(1,34) = 3.52$, $p = .069$, $\eta_p^2 = .09$. STAS was in none of the blocks a significant predictor of FD (all *p's* > .5), but the STAI x STAS interaction reached significance in the first, $F(1,34) = 7.52$, $p = .010$, $\eta_p^2 = .18$, and second, $F(1,34) = 4.21$, $p = .048$, $\eta_p^2 = .11$, blocks. Correlational analysis confirmed that in the first block the gaze-aversion effect in the high anxiety-group was strongest when trait anger was low, $r = .57$, $p = .013$ (see **Figure 5.3**), but this was not significant in the second block, $r = .31$, $p = .209$. In sum, these results confirm that anxiety predicts gaze-aversion from angry faces, which is strongest in submissive individuals as indexed by high trait anxiety and low trait anger, especially during the first five seconds of the task.

SMT: Gaze (happy condition)

In the happy-condition participants directed 49.8%, *SD* = 4.7, of their gaze-fixations to happy faces. Average duration of fixations to happy faces was 421 ms, *SD* = 70, and to neutral faces 413 ms, *SD* = 71. Analysis of gaze-avoidance and gaze-aversion revealed no significant effects for BLOCK or STAI (all *p*'s > .12, see also **Figure 5.2**), thus these data were not further analyzed.

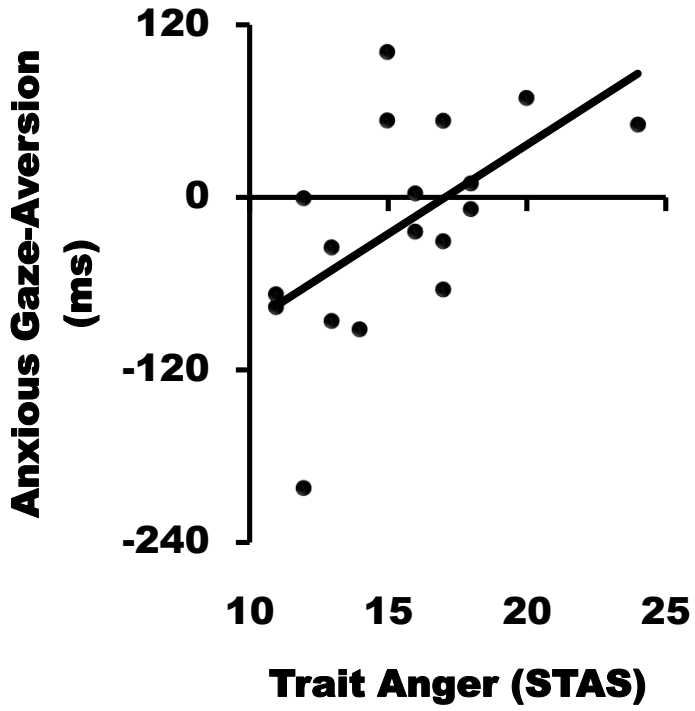


Figure 5.3 Correlation of gaze-aversion from angry faces in the high-anxiety group with scores on trait anger.

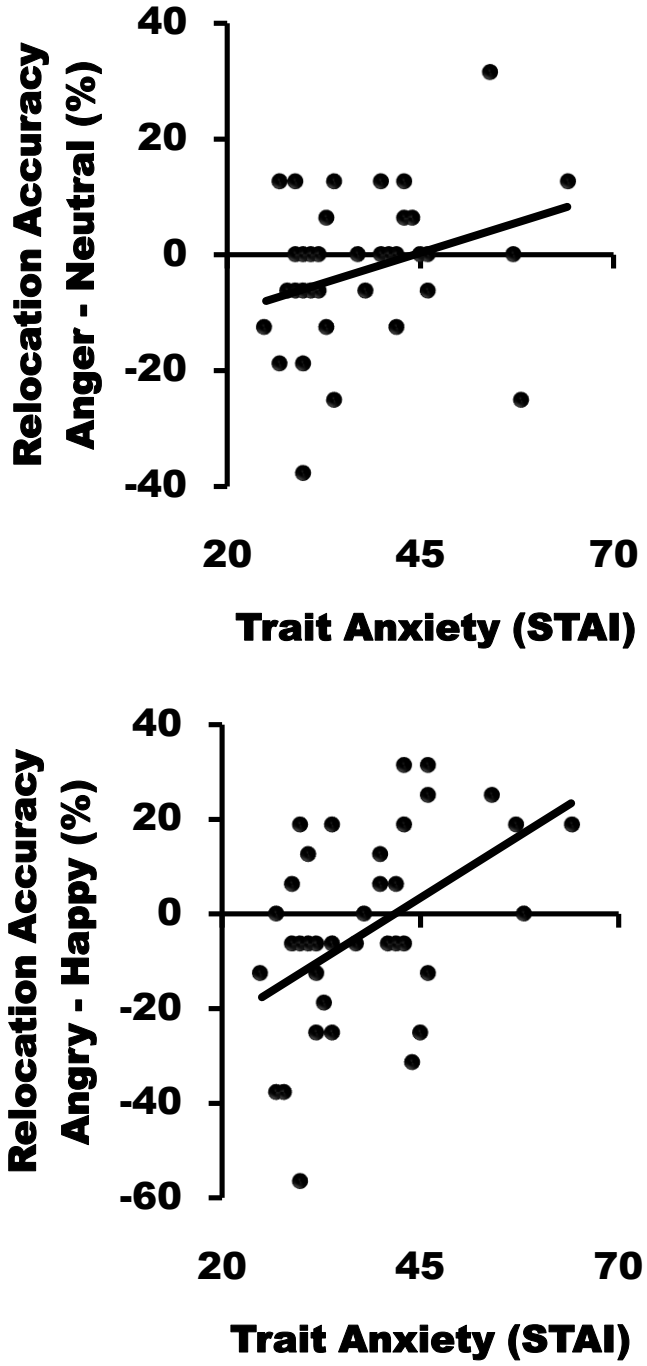


Figure 5.4 Correlation of relocation accuracy biases with scores on trait anxiety.

SMT: Memory performance

Mean relocation accuracy in the angry condition was 63.5%, *SD* = 17.5, for neutral, and 60.0%, *SD* = 19.2, for angry faces. In the happy condition 65.6%, *SD* = 22.6, of the neutral,

and 64.8%, *SD* = 20.5, of the happy faces were correctly relocated. Mean relocation accuracies were subjected to a General Linear Model analysis with CONDITION (anger and happy), and TYPE (emotional and neutral faces) as within-subjects variables, and standardized STAI and STAS as continuous variables. The main effect for STAI was not significant, $F(1,34) = 0.18$, $p = .894$, but the interaction for CONDITION and STAI was, $F(1,34) = 6.00$, $p = .020$, $\eta_p^2 = .15$. All main and interaction effects for STAS were non-significant (all p's > .3), thus STAS was removed from further analyses. To further specify the effects of STAI on relocation performance, we computed performance bias scores by subtracting relocation performance on neutral from emotional faces within each condition, and subtracting performance in the happy from the angry condition for both emotions. Correlational analysis showed that relocation performance on neutral faces was not different across conditions in relation to STAI, $r = .17$, $p = .312$, which was also the case for the happy-neutral bias score , $r = -.12$, $p = .468$. Crucially, as shown in **Figure 5.4**, both the anger-neutral, $r = .35$, $p = .032$, and anger-happy biases, $r = .47$, $p = .003$, correlated significantly with STAI, which indicates that STAI predicted better performance on relocation of angry compared to happy, as well as neutral faces.

Finally, to assess whether the vigilance-avoidance pattern of attention that we found in relation to anxiety also underlies the memory bias for anger, we correlated the vigilance-avoidance bias (first four fixations versus the rest of the trial, see **Methods**) with relocation performance on angry faces. The resulting positive correlation, $r = .32$, $p = .047$, indeed indicates that vigilance-avoidance attention underlies the memory bias for the location of angry faces.

Discussion

In this study we tested the influence of hostility in the social environment on gaze-behavior and memory. Following the hypervigilance-avoidance theory of anxiety (Bogels & Mansell, 2004), we predicted that anxious individuals would rapidly attend to socially threatening angry faces (vigilance), but subsequently avoid looking directly at them (avoidance). Furthermore, we expected superior memory for the location of such social threats in anxious individuals, since anxiety is reliably associated with increased memory for threatening

information (Mitte, 2008). Finally, on the basis of primate behavior in social hierarchies (Mazur & Booth, 1998), and evidence from human gaze-movement studies (Garner et al., 2006; Terburg et al., 2012c; Terburg et al., 2011), we made a clear distinction between gaze-avoidance and gaze-aversion, and hypothesized that submissive individuals would exhibit faster gaze-aversion from the socially threatening angry faces.

On the basis of our results we can confirm all three hypotheses. Anxious individuals indeed showed an attentional pattern of vigilance-avoidance, whereby they initially gazed more often to (vigilance), but subsequently avoided looking directly at (avoidance), angry faces. Furthermore, both vigilance-avoidance attention and trait anxiety significantly predicted improved memory for the location of angry faces. Finally, high anxious, and especially submissive participants, as indexed by high trait anxiety and low trait anger (Russell & Mehrabian, 1974; Smith et al., 2010), averted their gaze more rapidly from angry faces. Thus, submissive individuals showed more rapid gaze-aversion from social threat.

We can thus make two important observations. First, we confirm the hypervigilance-avoidance theory of anxiety (Bogels & Mansell, 2004; Garner et al., 2006) for natural gaze-behavior in a social setting. Anxious individuals rapidly detect, but subsequently avoid social threat. Importantly, we provide evidence that these mechanisms go hand-in-hand with superior memory for socially threatening information. In everyday life, anxious individuals might habitually scan their immediate surroundings for unfriendly and potentially dangerous people and remember their location in order to avoid further social confrontation. We can therefore speak of hypercoding-avoidance, whereby initial vigilance facilitates detection of, and memory for threat, which can subsequently be avoided.

Second, submissiveness is clearly distinct from anxiety on the basis of social gaze-behavior. Anxiety makes one alert for, but avoidant from social threat, whereas submissiveness is also reflected in rapid gaze-aversion from dominant eye-contact, reminiscent to a staring-contest (Mazur & Booth, 1998). Therefore, our results confirm the observations of submissive gaze-behavior in primates, translate these to humans, and extend our understanding on how anxiety and submissiveness are distinctly reflected in human social gaze-behavior in a hostile context. Interestingly, in a recent neuroimaging study trait anger predicted activity in the dorsal amygdala in response to angry faces, but only in high trait anxious individuals (Carré, Fisher, Manuck, & Hariri, 2010). Thus, activity in the dorsal

amygdala when viewing angry faces was relatively decreased in submissive individuals as indexed by high anxiety and low anger. The dorsal amygdala has indeed been implicated in social approach and avoidance behaviors (Ernst & Fudge, 2009), and is generally regarded to be involved in attention to emotionally relevant information (Davis & Whalen, 2001; Phelps & LeDoux, 2005). Increased responding in this area might therefore help to direct and maintain attention to, and thus inhibit submissive avoidance from, social signals of dominance. Speculatively, dorsal amygdala reactivity to angry facial expressions might distinguish submissiveness from anxiety. This should however be confirmed in future research. Additionally, given the limited ecological validity of our experimental setting, future research should establish whether our findings extend to real-life situations, and confirm the evolutionary origin of social gaze-behavior and its relation to personality traits of anxiety and submissiveness.

In light of the disorders of (social) anxiety and fear, it is furthermore important to make a functional distinction between on the one hand vigilance-avoidance and memory-bias for threat, and on the other hand submissive gaze-aversion. As our results show, the first two are restricted to anxiety in general, whereas the third involves a strong social component. Indeed, where the first relates to how anxious individuals rapidly detect threat, remember its location, and subsequently avoid it to reduce stress (Bogels & Mansell, 2004), the second is an overt social signal indicating subordination to the angry other (Mazur & Booth, 1998). It might therefore be argued that in social anxiety, which is characterized by a strong fear of social evaluation, but also by submissiveness (Trower & Gilbert, 1989; Watson & Friend, 1969; Weeks et al., 2011), both components are active. Vigilance-avoidance reduces the chance on explicit social evaluation, and gaze-aversion reflects submissiveness. Therefore, while we in the present study focused on trait anxiety and anger in order to disentangle anxiety and submissiveness, future research can also focus on the here reported social gaze-behaviors in relation to social anxiety and social phobia.

In sum, by assessing natural gaze-behavior within a hostile social context we show that submissive individuals avert their gaze more rapidly from social threats, and anxious individuals detect such threats more rapidly and subsequently avoid them. On top of this attentional pattern of anxious hypervigilance-avoidance, anxiety also predicts superior memory for the location of social threats. We propose that anxiety is reflected by

hypercoding-and-avoidance, whereby the environment is habitually scanned for possible threats, and the location of threat is subsequently memorized and avoided.

Chapter 6

Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces

Based on: Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011) Psychological Science, 22, 314-319

Abstract

In primates, dominance/submission relationships are generally automatically and nonaggressively established in face-to-face confrontations. Researchers have argued that this process involves an explicit psychological stress-manipulation mechanism: Striding with a threatening expression, while keeping direct eye contact, outstresses rivals so that they submissively avert their gaze. In contrast, researchers have proposed a reflexive and implicit modulation of face-to-face confrontation in humans, on the basis of evidence that dominant and submissive individuals exhibit vigilant and avoidant responses, respectively, to facial anger in masked emotional Stroop tasks. However, these tasks do not provide an ecologically valid index of gaze behavior. Therefore, we directly measured gaze responses to masked angry, happy, and neutral facial expressions with a saccade-latency paradigm and found that increased dominance traits predict a more prolonged gaze to (or reluctance to avert gaze from) masked anger. Furthermore, greater non-dominance-related reward sensitivity predicts more persistent gaze to masked happiness. These results strongly suggest that implicit and reflexive mechanisms underlie dominant and submissive gaze behavior in face-to-face confrontations.

Introduction

A typical bar brawl often starts with two individuals in a face-to-face dominance contest. Overt social aggression may be prevented when one of them communicates submission by word or gesture. These mechanisms in humans seem to share commonalities with frequently observed behavior during dominance contests in primate social systems. In primates, dominance/submission relationships are established primarily by individuals staring at one another (staring endurance) until one averts the eyes (gaze aversion) to signal submission and avoid aggression (Mazur & Booth, 1998). It has been argued that a psychological stress-manipulation mechanism is operative in these face-to-face competitions between group members: Opponents who are "outstressed" by the exchange of threats and the endurance of staring may relieve their discomfort by submissive gestures, such as gaze aversion (Mazur & Booth, 1998).

The angry facial expression serves as an important threat signal in these dominance encounters (Öhman, 1986). In humans, dominance/submission behaviors have not yet been investigated using genuine staring endurance and gaze aversion. However, an extensive line of research with pictorial emotional Stroop tasks has shown that self-reported and hormonally indexed traits of dominance and submission predict vigilant and avoidant responses, respectively, to angry faces (van Honk & Schutter, 2007b). For example, the behavioral activation system (BAS), trait anger, and basal testosterone levels are strongly associated with vigilant responses to (masked) angry facial expressions (Putman et al., 2004; van Honk et al., 2001a; Wirth & Schultheiss, 2007). Furthermore, avoidant responses to masked anger have been demonstrated in socially anxious subjects and in subjects with high levels of cortisol (Putman et al., 2004; van Honk et al., 1998, 2000). On the basis of these data, van Honk and Schutter (2007b) proposed that vigilant and avoidant responses to angry faces in emotional Stroop tasks index motives of dominance and submission. Moreover, these findings were predominantly obtained in backward-masking conditions, which suggests that the mechanisms are implicit and reflexive, and therefore not part of the explicit psychological stress-manipulation mechanism that is thought to operate in social dominance encounters (Mazur & Booth, 1998).

However, rapid color naming, the dependent variable in the emotional Stroop task, is a rather indirect and ecologically weak measure of dominant behavior. The hypothesis put

forward by van Honk and Schutter (2007b) can be truly confirmed only by measuring interactive gaze behavior directly. For that reason, in the present research we replaced verbal color-naming responses with ecologically valid behavioral responses—gaze aversion in face-to-face confrontations.

To facilitate direct comparison with previous studies, we devised a new task akin to emotional Stroop paradigms. The required response, however, was rapid aversion of gaze from subliminally presented angry, happy, or neutral facial expressions (see **Figure 6.1**). The anger gaze is a signal of dominance (Mazur & Booth, 1998), and characteristics of dominance or submissiveness, respectively, should inhibit or facilitate aversion of gaze from facial anger (van Honk & Schutter, 2007b). Compared with an angry expression, however, a happy facial expression is a nondominant gesture (Ellis, 2006). Although a smile is mimicked reflexively, women do so more than men (Hess & Bourgeois, 2010), and indeed this difference is argued to be rooted in males' dominance motivation and higher levels of testosterone (Dabbs Jr., 1997; Ellis, 2006). Furthermore, particularly after subliminal presentation, happy faces evoke positive evaluations of pictures (Murphy & Zajonc, 1993) and promote appetitive motivation (Winkielman, Berridge, & Wilbarger, 2005). Unlike an angry face, a happy face is thus an automatic and nondominant cue for reward.

The Behavioral Activation Scale (Carver & White, 1994) may well tap into both dominant and reward-sensitive behavior. It consists of three subscales: Fun Seeking (BASF), Drive (BASD), and Reward Responsiveness (BASR). BASF (e.g., "I often act on the spur of the moment") indexes willingness to engage in novel rewarding situations and is a measure of reward sensitivity unrelated to dominance or anger. BASD and BASR index affective response to rewards. Although anger or dominance are never explicitly mentioned in these subscales (e.g., BASD: "I go out of my way to get things I want"; BASR: "It would excite me to win a contest"), they are linked to susceptibility to anger-evoking scenarios (Carver, 2004), self-reported anger (Harmon-Jones, 2003a), expression of anger (Smits & Kuppens, 2005), and vigilance toward masked angry faces (Putman et al., 2004). Moreover, neuroimaging studies have shown that higher scores on these subscales predict increased responding to angry facial expressions in neural regions implicated in aggression (Beaver, Lawrence, Passamonti, & Calder, 2008), and that this effect occurs within 200 ms after stimulus presentation (Bediou, Eimer, d'Amato, Hauk, & Calder, 2009).

Neuroeconomic research has confirmed the implicit relation of BASD and BASR to dominance: Higher scores on these two subscales, but not higher BASF scores, predicted larger offers in an ultimatum game (Scheres & Sanfey, 2006), and researchers have argued that larger offers in this game originate from an increased concern for social status (Eisenegger et al., 2010). The combined BAS subscales thus support a motivational interpretation of behavioral activation in which both dominance- and non-dominance-related reward sensitivity have their place (Carver, 2004; Harmon-Jones, 2004).

We hypothesized that increasing levels of dominance-related reward sensitivity (BASD and BASR) would predict slower gaze aversion from masked facial anger (relative to masked facial happiness). Given the experimental results described earlier, we expected that both inhibition of gaze aversion in high-dominant individuals and facilitation of gaze aversion in low-dominant individuals would underlie this effect (see van Honk & Schutter, 2007b, for a review). Furthermore, because BASF measures non-dominance-related reward sensitivity, we hypothesized that higher BASF scores would predict a bias for positive reward cues that would be reflected in slower gaze aversion from masked happy faces (relative to masked angry faces). We tested these hypotheses with the newly developed gaze-aversion task.

Methods

Forty healthy volunteers (20 female; 20 male; mean age = 22.7 years, *SD* 2.8) participated for either course credit or payment. The face stimuli for the gaze-aversion task were colorized (blue, green, and red) faces of 10 actors (5 female, 5 male), each expressing three emotions (angry, happy, and neutral; Ekman & Friesen, 1976). The 90 stimuli were presented once in random order. Each trial commenced with a fixation screen, which was followed by a 33 ms presentation of a colorized face stimulus and then a mask of the same color. The mask remained on the screen until the participant responded. Mask stimuli were cut-up and randomly reassembled faces. So that contrast and luminance levels would be constant over the whole trial, a (gray) mask stimulus was presented with the fixation cross during the gaze-fixation phase. As shown in **Figure 6.1**, three gray dots were presented below the stimulus. During presentation of the mask, each gray dot was replaced by a dot of a different color (blue, green, and red; colors randomly assigned in each trial). Participants responded by

looking, as quickly as possible, away from the mask to the dot that was the same color as the preceding face stimulus. Before the task, participants completed 10 practice trials with neutral faces only.

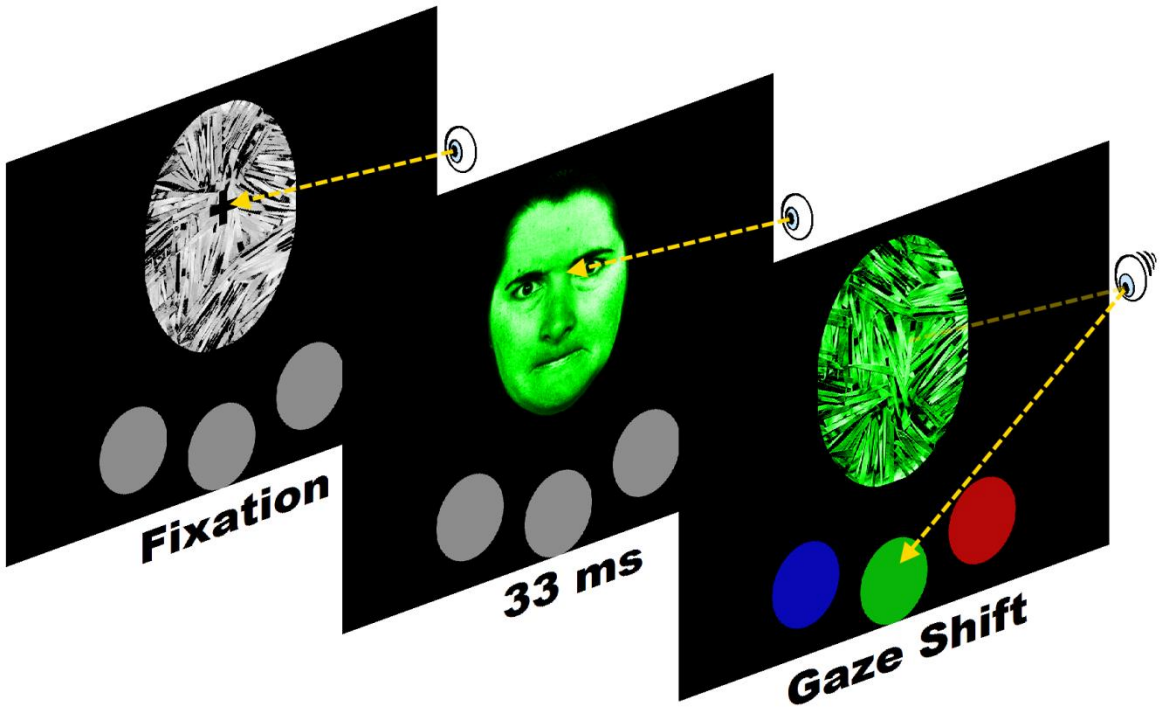


Figure 6.1 Illustration of the gaze-aversion task. Participants fixated the screen, were presented with a backward-masked colorized angry, happy, or neutral facial expression, and responded by making a saccadic eye movement to the dot the same color as the face.

The rationale behind this task is that the presentation of masked emotional expressions can evoke dominance (angry faces) and reward-seeking (happy faces) responses that can delay looking away from the mask (the location of the face). Thus, the difference between latencies on angry-face trials and happy-face trials can serve as a measure of implicit dominance or nondominance approach motives.

Eye movements were recorded with a Tobii-1750 binocular infrared eye tracker (sampling at 50Hz, 0.5° accuracy; Tobii Technology, Danderyd, Sweden) with an integrated LCD display (8ms response time). We used this system because it was recently shown that restriction of body movement greatly reduces approach-motivational responses (Harmon-Jones & Peterson, 2009), and this system is not head mounted, leaving participants fairly free

in their movements. Stimulus presentation commenced when participants fixated the fixation cross for a randomly determined interval (1000-1500 ms, to avoid timing habituation), and saccade latency was estimated as the time between onset of the stimulus and the first gaze at the correctly colored dot. Latencies less than 100 ms or more than 2 standard deviations above or below the overall mean (2.5%) were removed from the analysis, which was conducted with two-tailed tests ($\alpha = .05$).

After giving informed consent, participants filled out the BAS questionnaire (Carver & White, 1994) and then performed the gaze-aversion task. Finally, they performed an objective awareness check intended to establish whether the emotion in the facial stimuli was masked successfully. In this awareness check, all 30 faces were shown once again with a mask. Colors were randomly assigned, but each color appeared 10 times. Participants had to report the presented emotion, choosing from three options (angry, happy, or neutral). Thus, with this awareness check, we tested not for awareness of the faces, but rather for awareness of the stimulus quality of interest: emotional expression.

Results

First, we assessed performance on the awareness check. An individual score of 15 or higher was significantly above the chance level of 10 correct responses (binomial upper limit with one-tailed α of .05 for $n = 30$ and an expected proportion of correct answers of 1/3). Thirteen subjects scored 15 or higher and were considered to have (some) explicit awareness of the presented emotions. Moreover, a negative correlation between performance on the awareness check and average saccade latency ($r = –.42$, $p < .01$) indicated that the face stimuli interfered most with performance on the gaze-aversion task when the faces' emotions were processed implicitly. Because several emotional Stroop studies have shown that the interaction between angry expressions and motivational traits occurs exclusively when the stimuli are masked (van Honk & Schutter, 2007b), we created two separate groups of subjects. The implicit group ($n = 27$) scored at chance level on the individual awareness check, and the explicit group ($n = 13$) scored significantly above chance level. Average saccade latency was 448 ms ($SD = 64$ ms) for the implicit group and 422ms ($SD = 51$ ms) for the explicit group. Repeated

measures analysis of variance did not reveal any main effects of emotion (angry, happy, or neutral) on saccade latency in either the implicit group, $F(25) = 0.283$, $p > .7$, or the explicit group, $F(11) = 0.016$, $p > .9$.

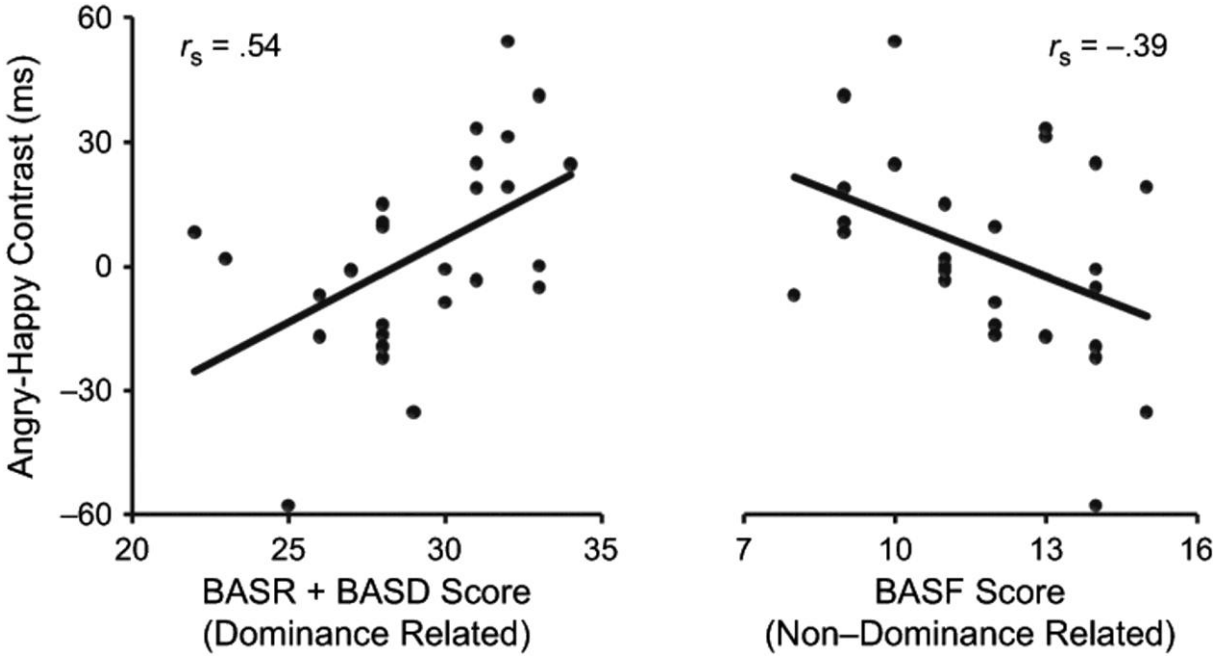


Figure 6.2 Scatter plots showing the correlation between the implicit group's angry-happy contrast scores and their scores on the Behavioral Activation Scale. Angry-happy contrast scores were calculated by subtracting mean latencies on happy-face trials from mean latencies on angry-face trials; consequently, a high score indicates longer saccade latencies for angry-face trials (i.e., inhibition of submissive gaze aversion). Results are shown separately for the dominance-related questionnaire subscales pooled together (Behavioral Activation Scale Drive and Reward Responsiveness, or BASD and BASR, respectively) and for the non-dominance-related subscale (Fun Seeking, or BASF).

To test our hypotheses regarding dominance- and non-dominance-related reward sensitivity directly, we computed angry-happy contrast scores by subtracting mean latencies on happy-face trials from mean latencies on angry-face trials. High contrast scores thus represent longer gaze toward angry faces than toward happy faces. In the implicit group, the angry-happy contrast scores were significantly correlated with all three BAS subscale scores; the correlations were positive for BASD ($r_s = .39$, $p < .05$) and BASR ($r_s = .44$, $p < .05$) and negative for BASF ($r_s = –.39$, $p < .05$). There were no significant relations between contrast

scores and BAS subscale scores in the explicit group (r_s = –.10, p = .75; r_s = .04, p = .91; r_s = .38, p = .20, respectively). Furthermore, BASF was not significantly related to BASR (r_s = .26, p = .11) or BASD (r_s = .06, p = .73), which indicates that the negative effect of BASF on gaze aversion in the implicit group was independent of effects of BASR and BASD. Because BASD and BASR scores were positively correlated (r_s = .34, p < .05), and both predict dominance motives (see the introduction), we pooled them into a single scale; scores on this scale were significantly correlated with angry-happy contrast scores in the implicit group (r_s = .54, p < .01; see **Figure 6.2**).

To compare the implicit and explicit groups directly, we conducted a linear regression analysis with angry-happy contrast score as the dependent variable and group, Group × Dominance-Related BAS Score (BASR + BASD), and Group × Non-Dominance-Related BAS Score (BASF) as regressors. The overall model was significant (n = 40, R = .574, p < .05), and both interactions, indicated a stronger relation between BAS score and angry-happy contrast score in the implicit group than in the explicit group, t(38) = 2.15, p < .05, and t(38) = –2.9, p < .01, respectively,. Separate regression analyses for the two groups confirmed that dominance- and non-dominance-related BAS scores were significant predictors of angry-happy contrast scores in the implicit group (n = 27, R = .688, p < .001), with dominance-related BAS scores making a positive contribution to contrast scores (β = 0.56, p < .01) and non-dominance-related BAS scores making a negative contribution (β = –0.47, p < .01). The model for the explicit group was not significant (n = 13, R = .15, p = .44).

Discussion

We have shown that slower gaze aversion from masked facial anger is significantly predicted by the dominance-related BAS subscales, BASD and BASR. Accordingly, the present data provide direct support for the hypothesis that speed of gaze aversion from masked facial anger depends on motives of dominance and submission (van Honk et al., 2001a; van Honk et al., 1998, 2000; Wirth & Schultheiss, 2007). Additionally, the third BAS subscale, BASF, independently predicted relative engagement with masked happy faces, a result confirming that this subscale represents non-dominance-related reward sensitivity. Thus, our eye-tracking

gaze-aversion task not only successfully provoked modulation of implicit face-to-face gaze behavior related to motives of dominance and submissiveness, but also revealed theoretically grounded dissociated responses to angry and happy expressions within the construct of behavioral activation (Carver, 2004; Harmon-Jones, 2004).

Crucially, the relation between dominance motives and gaze aversion was observed only in subliminal conditions. We cannot, however, exclude the possibility that humans use gaze contesting to consciously outstress opponents (Mazur & Booth, 1998), especially because genuine face-to-face confrontations generally persist long enough to initiate complex psychological mechanisms. Our findings do, however, support an extensive line of psychobiological research on rapid and reflexive initiation of dominance/submission behavior (van Honk & Schutter, 2007b).

The neural mechanism underlying dominance/submission behaviors in primates was extensively described by Emery and Amaral (2000). They ascribed a vital role to the amygdala, which integrates sensory information, such as facial expressions, with social context and connects to endocrine and autonomic systems to facilitate appropriate behavior. Similar mechanisms have been described in humans. It is thought that when conscious evaluation of facial expressions is prevented, the sensory information is still crudely evaluated for threat in subcortical structures and relayed to prefrontal areas via the amygdala (Vuilleumier, 2002). Researchers have argued that this mechanism is an adaptive implicit alarm system that serves to direct attention and potentiate responding to threat (Liddell et al., 2005). The responses to masked facial threat in our high-BASD/high-BASR subjects possibly reflected enhanced responding of this implicit defense system, mediated by the motivation to stand one's ground and (if necessary) fight rather than flight. When confronted with an explicit threat, this fight-or-flight mechanism can be inhibited by higher-order cortical processes that maintain executive control (Nomura et al., 2004). Such inhibition might explain the lack of motivational modulation of responses in the group that explicitly processed the emotional stimuli in the present experiment, as well as the lack of such modulation in several previous emotional Stroop studies with angry faces (van Honk & Schutter, 2007b).

Because both excessive reactivity of this implicit defense system and lack of prefrontal inhibition of such reactivity are often associated with aggressive behavior (Siever, 2008), an

alternative explanation of our results might be that the reflexive fight response is not as effectively inhibited in dominant as in nondominant individuals. Recent evidence showing reduced white matter connections between the amygdala and frontal areas in psychopathy, possibly resulting in poor impulse control (Craig et al., 2009), seems to point in this direction. However, given that this higher-order mechanism requires conscious evaluation of threat, impulse control may largely concern the explicit mechanisms proposed by Mazur and Booth (1998). To further unveil the biological mechanisms of reflexive dominance, researchers will need to focus on interactions between subcortical and cortical brain structures.

In sum, the present data provide direct evidence for implicit and reflexive modulation of dominance/submission behaviors in humans. Furthermore, we have shown a dissociation between dominance- and non-dominance-related reward sensitivity within the BAS. High-BASF individuals implicitly track positive social signals, whereas high-BASR/high-BASD individuals persist in implicit angry face-to-face confrontations. Because success in dominance contests is achieved only when eye contact is never interrupted, such persistence may reflect an adaptive mechanism to ensure advantage in social dominance confrontations (Putman et al., 2004; van Honk et al., 2001a). Likewise, facilitation of gaze aversion among individuals with submissive characteristics is adaptive, as it reduces the chance of injury and saves valuable resources (van Honk et al., 1998, 2000). Finally, utilizing saccade latencies as a social behavioral measure may have great potential for future social and motivational research.

Chapter 7

Testosterone decreases trust in socially naïve humans

Based on: Bos, P. A., Terburg, D., & van Honk, J. (2010)
Proceedings of the National Acedemy of Sciences, 107, 9991-9995

Abstract

Trust plays an important role in the formation and maintenance of human social relationships. But trusting others is associated with a cost given the prevalence of cheaters and deceivers in human society. Recent research has shown that the peptide hormone oxytocin increases trust in humans. However, oxytocin also makes individuals susceptible to betrayal as under influence of oxytocin subjects perseverate in giving trust to others they know are untrustworthy. Testosterone, a steroid hormone associated with competition and dominance is often viewed as an inhibitor of sociality, and may have antagonistic properties with oxytocin. The following experiment tests this possibility in a placebo-controlled, within-subjects design involving the administration of testosterone to 24 female subjects. We show that compared to the placebo, testosterone significantly decreases interpersonal trust, and, as further analyses established, this effect is determined by those who give trust easily. We suggest that testosterone adaptively increases social vigilance in these trusting individuals to better prepare them for competition over status and valued resources. In conclusion, our data provide unique insights into the hormonal regulation of human sociality by showing that testosterone down regulates interpersonal trust in an adaptive manner.

Introduction

The hormonal regulation of human social relationships has recently been approached by several disciplines, including psychology, economics and neuroscience (Baumgartner, Heinrichs, Vonlanthen, Fischbacher, & Fehr, 2008; Delgado, 2008; Hermans et al., 2008; Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005; van Honk, 2009). Of the many important findings, the discovery that oxytocin increases interpersonal trust (Kosfeld et al., 2005), as well as the perseveration of trust toward the untrustworthy (Baumgartner et al., 2008), has been of considerable interest given current debates about the evolution of prosocial behavior in humans and other animals (Delgado, 2008). Here we investigate whether testosterone, a hormone associated with success in competition for resources and dominance (Archer, 2006), and an alleged inhibitor of sociality (van Honk, 2009), may counteract the role of oxytocin in interpersonal trust. More specifically, we investigate whether, and in what way, testosterone administration in humans decreases interpersonal trust with unfamiliar others.

Humans are highly social and cooperative animals, whose social relationships importantly rely upon trust. Without trust, suspicion spreads through human social interaction, allowing fear to threaten relationships by instilling vigilance for treachery and betrayal. Compared with other animals, humans are much more likely to trust and cooperate with genetically unrelated and unfamiliar others, and these differences might constitute social adaptations that underlie their evolutionary success (Kosfeld et al., 2005). Trust has, however, a downside: naïve, trusting humans run a much greater risk of being misguided and deceived by others. In the same way that we have evolved capacities to help others, we have also evolved capacities to deceive and cheat. Thus, those who are willing to believe what others say, or fail to probe the motivations underlying their actions, may fall prey to considerable economic and social costs.

Although humans are essentially social animals (Adolphs, 2009), competition for resources also underlies the evolution of our species. It is thus critical to understand both the evolutionary and moment-to-moment dynamic between competition and trust, as both have played a critical role in both the construction and destruction of society (Diamond, 1997).

Recent research in humans using an economic exchange task has shown that administration of oxytocin, a peptide hormone known for its role in attachment and bonding (Insel & Young, 2001), increases interpersonal trust in an economic game as evidenced by

higher monetary allocations to unfamiliar others (Kosfeld et al., 2005). Other studies have also shown, however, that oxytocin induces perseverative trust: following oxytocin administration, subjects continue to allocate substantial amounts of funds to untrustworthy others, despite being told that their opponents had repeatedly violated their trust (Baumgartner et al., 2008). These results highlight the janus-face of trust: high levels of interpersonal trust are beneficial in social interactions, but may place individuals at great personal risk (Diamond, 1997).

Testosterone, a steroid hormone with potentially toxic consequences for human sociality (van Honk, 2009), might counteract the maladaptive aspects of trust. Testosterone has been associated with social dominance and success in competition (Archer, 2006), and may restrain interpersonal trust to ensure social scrutiny for status and economic concerns. Indeed, testosterone levels in humans correlate positively with financial gain on the stock market, and as such, appear predictive of economic shrewdness (Coates & Herbert, 2008). These results, however, are only correlational and thus do not clarify testosterone's relation with interpersonal trust. To explore the possible causal role of testosterone in trusting behavior, and in particular, to test whether testosterone decreases interpersonal trust in humans, we investigated the effect of a single administration of testosterone to healthy volunteers in a trust experiment. In a double blind, counterbalanced design, we sublingually administered either 0.5 mg of testosterone or a placebo to 24 adult females on two separate days (72 hr interval between treatments). Only women participated because the parameters (quantity and time course) for inducing neurophysiological effects after a single sublingual administration of 0.5 mg of testosterone have been established in women (Tuiten et al., 2000), but are unknown in men (for details; see **Methods and Materials**).

We used facial trustworthiness evaluations as a measure of interpersonal trust to control for the inherent rewarding properties of economic exchange tasks. The association of testosterone with reward and risk-taking is very strong, and could potentially interfere with the measure for trust in an economic exchange task (Kosfeld et al., 2005; van Honk, 2009). Importantly, trustworthiness judgments of non-familiar faces is not only a highly validated procedure (Adolphs, Tranel, & Damasio, 1998; Todorov & Duchaine, 2008) unconfounded by reward, but these judgments are also highly correlated with investments in an economic-trust task (van 't Wout & Sanfey, 2008). A recent study showed higher trustworthiness ratings

to unfamiliar others after oxytocin administration compared to placebo, demonstrating the validity of using a comparable paradigm for measuring trustworthiness (Theodoridou, Rowe, Penton-Voak, & Rogers, 2009). For these reasons the trustworthiness task is our method of choice for measuring the effect of testosterone administration on subjects' interpersonal trust levels.

Methods and materials

Subjects

The Ethics Committee of the University Medical Centre Utrecht approved the protocol of our experiment wherein 24 healthy young women (mean age 20.2) participated. All women received testosterone and placebo, in randomized order, with a 72 hr latency between sessions. Subjects had no (history of) psychiatric disorders, neurological or endocrine abnormalities. They did not smoke and used no medication other than contraceptives. We controlled for influences of hormonal change due to menstrual cycle by only including women who used single-phase contraceptives, and testing them during the 3-week period they were on these contraceptives and not during menstruation (see also Aarts & van Honk, 2009). In this 3-week contraceptive period menstrual-cycle influences are virtually absent**.** Moreover, any effects of the contraceptives would be equal during the placebo or testosterone condition.

Substance administration

The drug samples consisted of 0.5 mg of testosterone, 5 mg of (the carrier) cyclodextrine, 5 mg of ethanol, and 5 ml of water. Testosterone was omitted from the placebo samples, and both testosterone and placebo were administered sublingually. Previous experimental research established the time course of changes in blood levels of testosterone and physiological responsiveness in typical young women after a single sublingual administration of 0.5 mg of testosterone (Tuiten et al., 2000). A 10-fold increase in total testosterone was observed 15 min after intake with testosterone levels returning to baseline within 1.5 hr (Tuiten et al., 2000). It was also shown that this single administration of testosterone significantly elevated vaginal

pulse amplitude in healthy young women which peaks around 4 hours. Thus, physiological effects after single sublingual administrations of 0.5 mg testosterone peak 2.5 hours after the testosterone level in the blood has returned to baseline. Note, that vaginal pulse amplitude, a centrally-driven response evoked by erotic material, is the only physiological measure known to possess a non-habitual nature, thus allowing multiple measures throughout the day (Tuiten et al., 2000; van der Made et al., 2009b). There is no method available to assess the time course of effects of testosterone in human males, while in females the present time-course method may have unique applicability in the treatment of sexual dysfunction (van der Made et al., 2009a; van der Made et al., 2009b). Crucially, the reliability and generalizability of behavioral effects after a 4-hr delay has been successfully established in more than 20 studies, addressing both sexual, social and emotional behaviors in young typical women (e.g. Bos, Hermans, Montoya, Ramsey, & van Honk, 2010; Eisenegger et al., 2010; Hermans et al., 2006b; Hermans et al., 2008; van der Made et al., 2009b; van Honk et al., 2005; van Honk & Schutter, 2007a; van Honk et al., 2001b). Therefore, in the present protocol, a 4-hour delay between testosterone administration and measurement of mood and the trustworthiness ratings was again used.

Physiological levels and potential neuroendocrine mechanisms

The 10-fold increase in testosterone levels that our method induces (Tuiten et al., 2000) seems rather high in the light of increases seen in treatment studies. However, it is important to note that there are important differences between the chronic treatments, which do not consider a time course of effects, and our single administration approach. Our single sublingual administration of 0.5 mg testosterone produces an increase in absolute levels of testosterone in most cases higher than that seen with chronic treatment, but within and during a very short period. Crucially, it is conjectured by van der Made et al. (2009b) that this increase will not produce a proportional increase in the free fraction of testosterone; the amount of testosterone reaching the brain will be much less. A sex hormone binding globulin (SHBG) saturation threshold mechanism has been postulated: The increase of testosterone into the body will first bind to SHBG (and to albumin, to a smaller extend), before being able to produce an increase in the free fraction (van der Made et al., 2009b). The increase of testosterone produced by the sublingual 0.5 mg administration method does not compare to the 10-fold increase in total

testosterone in the blood, but would be large enough to pass this putative SHBG threshold, resulting in a short increase in the free testosterone fraction. This short increase, however, is responsible for cognitive, affective, and behavioral effects observed a few hours later, which have been reported in numerous studies in human females, as noted above.

Generalizability of effects to males

The parameters (quantity and time course) for inducing neurophysiological effects after a single-sublingual administration of 0.5 mg of testosterone are thus known in women, but not in men. Nonetheless, based on findings from our correlational research on testosterone and human social behavior in which we used males and females, we expect the effects of testosterone administration to be similar for males and females (van Honk et al., 2001b; van Honk et al., 1999). Moreover, we have repeatedly shown that testosterone administration in females results in more male-typical social behavior (Hermans et al., 2007; van Honk & Schutter, 2007a). Finally, others have shown that testosterone administration in females (Eisenegger et al., 2010) seems to increase status seeking behavior, and this finding agrees with correlations between endogenous testosterone levels and status-related behaviors shown in men (for a review see Mazur & Booth, 1998), and women (Cashdan, 1995; Dabbs Jr. & Hargrove, 1997; Josephs, Newman, Brown, & Beer, 2003; Josephs, Sellers, Newman, & Mehta, 2006). This adds to the growing evidence that testosterone plays an important role in female social behavior (Mehta, Jones, & Josephs, 2008; Mehta, Wuehrmann, & Josephs, 2009; Newman, Sellers, & Josephs, 2005; Wirth & Schultheiss, 2007). In sum, the relation between testosterone and social behavior apparently has much communality in human males and females.

Behavioral experiment

The stimuli in the trustworthiness task consisted of 150 grayscale frontal pictures of unfamiliar faces with neutral emotional expressions, of which 100 were adapted from Adolphs et al. (Adolphs et al., 1998) and 50 were taken from the Psychological Image Collection at Stirling (PICS: http://pics.psych.stir.ac.uk/). For our within-subject design we created two sets of 75 stimuli which were matched based on trustworthiness ratings in a previous study with 36 healthy adult subjects (Baas et al., 2008).

On each test day all stimuli of one set were presented once, in random order, both sets being counterbalanced with administration order. Pictures were presented in the middle of a 17" LCD display subtending a visual angle of approximately 8^{o} on a gray background. Directly below the stimulus a visual-analogue-scale was presented ranging from (left to right) 'very untrustworthy' to 'neutral' to 'very trustworthy'. For each stimulus, subjects were presented with the question 'How trustworthy do you think this person is?' and answered by clicking on the scale with a mouse cursor. After the response to each trial, a button appeared with the description 'next'; the subject's response to the scale could be adjusted until this button was clicked, and then disappeared. For each presentation trial, the scale was reset to the 'neutral' position. The stimuli were presented using software written in E-prime (Psychology Software Tools, inc). Subjects performed trustworthiness ratings once on each set, counterbalanced with order of administration.

For data analysis, the scale positions were coded from -100 (very untrustworthy) to 0 (neutral) to +100 (very trustworthy) in steps of 1. These scores were averaged for each subject and both test sessions to obtain individual measures of trustfulness in testosterone and placebo conditions.

Testosterone saliva measurement

Salivary sampling was chosen to obtain baseline testosterone levels. Salivary testosterone has proven to be a reliable noninvasive biomarker not only in the social (Dabbs Jr. & Hargrove, 1997; Josephs et al., 2003; Mehta et al., 2008; Newman et al., 2005; van Honk et al., 1999) and clinical sciences (Arregger, Contreras, Tumilasci, Aquilano, & Cardoso, 2007; van der Made et al., 2009a), and has also been successfully applied in economic research (Coates & Herbert, 2008; Sapienza, Zingales, & Maestripieri, 2009). Salivary sampling avoids possible confounding influences induced by (anticipation on) blood sampling procedures, which in humans are known to induce substantial stress, and increases in stress hormones such as cortisol (Hubbard, Kalimi, & Liberti, 1997). Our sampling method was based upon Granger et al. (2004), which has been successfully applied in several previous studies (e.g. Coates & Herbert, 2008; Hermans et al., 2008).

Testosterone in saliva was measured after diethylether extraction using a competitive radio-immunoassay employing a polyclonal antitestosterone-antibody (Dr. Pratt AZG 3290).

[1,2,6,7-3H]-Testosterone (TRK402, Amersham Nederland B.V.) was used as a tracer following chromatografic verification of its purity. The lower limit of detection was 10pmol/l and inter-assay variation was 16.1; 11.5; and 5.1% at 21; 100 and 230 pmol/l respectively (*n* = 4,5,5). Samples of two subjects were contaminated and showed out of normal range levels and were therefore not included in further analysis.

Our analyses showed that testosterone levels measured from saliva before administration did not differ between the testosterone and placebo administration condition in the complete group ($F(1,21) = 2.19$; *NS*), and also not in the high trusting subject group, which was accountable for our effects ($F(1,11) = 1.27$; *NS*). Furthermore, differences in baseline testosterone levels between subjects' placebo and testosterone condition (entered as a covariate in the original analyses) did not explain any variance in the effects of testosterone administration on trust, neither in the complete group ($F(1,20) = 0.42$; *NS*), nor in the high trusting group ($F(1,10) = 1.24$; *NS*). Finally, low trusting subjects compared to high trusting subjects did not show higher baseline testosterone levels in their placebo ($F(1,20) = 0.72$; *NS*) or testosterone condition ($F(1,20) = 0.87$; *NS*) in the experiment. Thus, our findings on testosterone administration cannot be attributed to variation in baseline testosterone levels in subjects between conditions, or to differences in baseline testosterone between conditions in general. Finally, testosterone baseline levels can also not account for our exclusive effect in the high trust group.

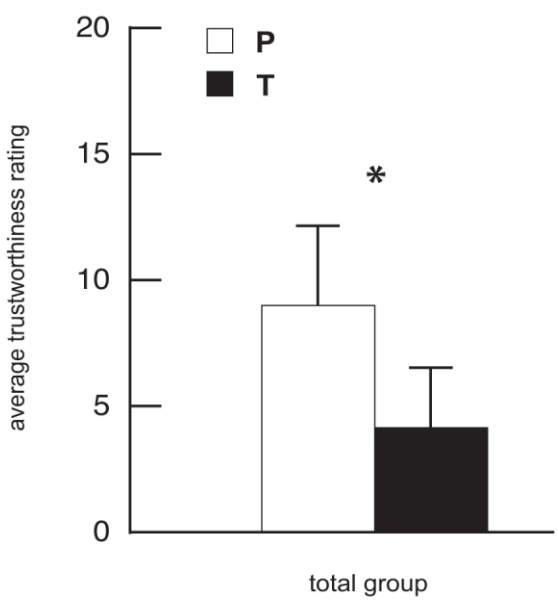


Figure 7.1 Testosterone induced a significant decrease in interpersonal trust in the total group (*n* = 24). A repeated-measures ANOVA (testosterone-placebo) showed ($F(1,23) = 4.56$, * $p = 0.044$). White bars represent placebo (P), black bars represent testosterone (T), and error bars represent standard error of the mean (*SEM*).

Results

In agreement with our hypothesis, we show a significant overall reduction in trustworthiness ratings after testosterone compared to placebo ($F(1,23) = 4.56$, $p = 0.044$) (**Figure 7.1**). This significant reduction has an effect size of Cohen's $d = .36$. To address the possibility of individual differences, we applied a linear regression to examine whether subjects' individual basic trust levels (observed from their ratings in the placebo condition) predicted testosterone-induced changes in trustworthiness. This analysis yielded a statistically significant correlation ($r = -0.66$, $p = 0.001$; **Figure 7.2**), with individuals' baseline trust levels explaining 43% of the variance in the effect of testosterone on interpersonal trust. To better qualify this effect, we applied a median split on the 24 basic-trust levels to create groups of 12 high and 12 low-trusting subjects. Analyses (**Figure 7.3**) showed no effects in low-trusting subjects ($F(1,11) = 0.79$, *NS*), but high-trusting subjects presented a substantial reduction in interpersonal trust following testosterone administration ($F(1,11) = 10.89$, $p = 0.007$). The effect sizes of testosterone's effect in the low and high trust group are respectively $d = .08$ and $d = .92$. Note that the absence of an effect in the low trust group is not caused by a floor effect. That is, the dependent measure could range from -100 to 100, which did not restrict the ratings of the faces in the low trust group, since the average scores ranged from -13.5 to 8.5 and were normally distributed. In sum, testosterone administration reduced interpersonal trust, but only in subjects who were generally trusting, and therefore more at risk for deceit.

To control for potential secondary mood-generated effects of testosterone on interpersonal trust, we administered the shortened-version of the profile-of-mood-states (POMS) (Shacham, Reinhardt, Raubertas, & Cleeland, 1983) prior to the trustworthiness task for both the placebo and testosterone conditions; the POMS includes the subscales tension-anxiety, depression, anger, fatigue and vigour. Paired *t*-tests for the subscales showed non-significant effects (all p's > 0.24, two-tailed). Furthermore, subjects were asked after the experiment to indicate or guess the day they received testosterone. Subjects' scores were at chance (binomial = 0.84, two-tailed) and there was no statistically significant relationship between the subjective guess of the day of testosterone administration and trustworthiness ratings; an ANOVA that used testosterone-induced change in trust as a between-subject factor and correct-versus-incorrect guesses as a between-subject factor was not significant ($F(1,22) = 0.18$, *NS*). In sum, the effects of testosterone on interpersonal trust are not mediated by

either mood or subjective preconceptions (Eisenegger et al., 2010); they are pure effects of the hormone on behavior. Furthermore, as can be seen in **Methods and Materials**, testosterone levels which were measured from saliva before the experiment did not predict the trustworthiness scores, and the individual variance in these baseline testosterone levels between conditions did also not explain the effect of testosterone administration on interpersonal trust.

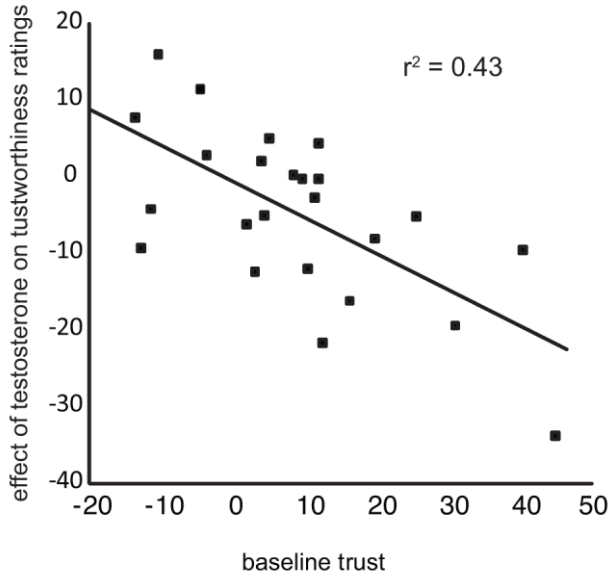


Figure 7.2 Plot of the baseline trust ratings, correlated against the effect of testosterone on trust judgments. The points on the left side of the graph, representing subjects who displayed low interpersonal trust in baseline measures, are clustered around zero for an effect of testosterone, indicating that their behavior was not affected by hormone treatment. In contrast, in the subjects displaying high interpersonal baseline trust, represented by the points on the right side of the graph, testosterone significantly decreased interpersonal trust.

Conclusion

Our results license the conclusion that testosterone decreases interpersonal trust, and in an apparently adaptive manner. The hormone acted selectively on our high trusting subjects, defensibly to down-regulate their trust to a level more advantageous in the competition for resources. Our data coincide with correlational evidence showing that higher testosterone levels predict financial gain on the stock market (Coates & Herbert, 2008), but seem somewhat at odds with recent findings of more fair bargaining behavior on the Ultimatum Game after testosterone administration (Eisenegger et al., 2010). However, the down-regulation of trust after testosterone administration at present was restricted to the high trusting, thus most socially-naïve half of our subject-group, and may for that reason be adaptive in the competition for status and resources. The Ultimatum Game, on the other hand

is an economic paradigm that measures fairness and not trust (cf. see Kosfeld et al., 2005), and in the Ultimatum Game fair offers are logically more often accepted. With fair offers the proposer takes control over the game and both players make money. In sum, more fair offers by the proposer in the Ultimatum Game after testosterone administration are also adaptive for achieving status and resources (Eisenegger et al., 2010). Hence, the context, i.e. trusting behaviors against fairness behaviors, in the above cases obviously defined -at first sight- differential effects of the hormone, which ultimately have common grounds. In many mammalian species testosterone's role in social behaviors is simply confined to motivating aggression in competition for status and resources. However, in humans the hormone seems to motivate for rational decision making, social scrutiny and cleverness (Eisenegger et al., 2010), the apparent tools for success in a modern society (Archer, 2006; Mazur & Booth, 1998; van Honk, 2009; but see: Zak et al., 2009). Viewed from this perspective, testosterone's relation to risk-taking behaviors in humans (van Honk et al., 2004a) might also be re-evaluated, as success on the stock market cannot be established by unrestrained risk taking, but requests a fine-tuned grasp of the balance between financial threat and reward (Coates & Herbert, 2008).

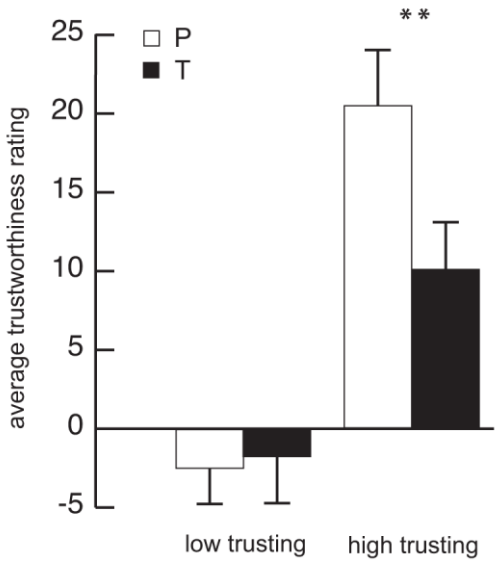


Figure 7.3 Separate repeated-measures ANOVAs for the low and high trusting subject groups showed that low trusting participants were completely unaffected by testosterone administration ($F(1,11) = 0.79$, *NS*), whereas high trusting participants showed a sizeable reduction in the evaluation of facial trustworthiness ($F(1,11) = 10.89$, ** $p = 0.007$). White bars represent placebo (P), black bars represent testosterone (T), and error bars represent standard error of the mean (*SEM*).

At present, there is little understanding of the neurobiological mechanisms by which testosterone acts on interpersonal trust. Nonetheless, animal data have shown that the amygdala is an important target of this hormone in the brain (Koolhaas, Van Den Brink,

Roozendaal, & Boorsma, 1990). Human neuroimaging studies support this finding by demonstrating the involvement of the human amygdala in the detection of facial threat (Davis & Whalen, 2001), in social evaluations of faces (Schiller, Freeman, Mitchell, Uleman, & Phelps, 2009), and specifically, in evaluations of trustworthiness from faces (Winston, Strange, O'Doherty, & Dolan, 2002). Furthermore, social evaluations of faces are impaired in patients with bilateral lesions to the amygdala, and these patients appear more trusting in their interactions with strangers (Adolphs et al., 1998). However, the amygdala does not stand alone in the social evaluation of faces; in particular, the orbitofrontal cortex (OFC), which shows strong connectivity to the amygdala, also plays an important role in these social processes. Moreover, the amygdala and OFC are thought to act in concert in the regulation of many social behaviors (Bachevalier & Loveland, 2006; Emery & Amaral, 2000), while the communication of these structures is affected by testosterone. In humans, administration of testosterone induces rapid reductions in the functional connectivity between amygdala and OFC in response to facial threat (van Wingen et al., 2010), and conversely, seems to activate the amygdala-brainstem defense circuit (Hermans et al., 2008). Interestingly, animal research shows that testosterone may induce amygdala-brainstem functional connectivity by acting on the social peptide vasopressin (Huber, Veinante, & Stoop, 2005; Koolhaas et al., 1990). Vasopressin, whose expression is regulated by testosterone (de Vries, 2008), increases outputs of the amygdala to the brainstem by acting on distinct neuronal populations within the amygdala (Huber et al., 2005).

Oxytocin, the hormone that increases interpersonal trust (Kosfeld et al., 2005), acts in a manner opposite to vasopressin, decreasing the outputs to the brainstem (Huber et al., 2005; Kirsch et al., 2005), but also increasing the involvement of frontal cortical regions, such as the OFC (Porges, 2001). Thus, testosterone and oxytocin seem to act as hormonal antagonists at the level of the amygdala, providing an adaptive balance in behavioral responses to social cues. In sum, we suggest that testosterone in the present study may have induced a prefrontal-limbic shift in social-emotional processing by regulating peptide expression in the amygdala (Huber et al., 2005; Koolhaas et al., 1990). This shift towards evolutionary older brain regions puts the brain in a defensive or vigilant mode (Huber et al., 2005; MacLean, 1990; Mobbs et al., 2007; Porges, 2001), and consequently may have down-regulated interpersonal trust. A socially vigilant stance is vital for gaining and maintaining dominance or leadership, and for

success in competition for resources (Coates & Herbert, 2008; Eisenegger et al., 2010; Mazur & Booth, 1998).

In conclusion, we show that testosterone plays a causal role in reducing interpersonal trust among unfamiliar individuals. The way in which testosterone decreased trust is consistent with its role in economic decision making and competitive interactions. The attribution of trust toward unfamiliar others was especially decreased in subjects who run the greatest risk of being misled by others, that is, those who grant trust easily. Consequently, testosterone increased social vigilance in trusting humans, presumably to better prepare them for the hard-edged competition over status and valued resources. These results provide insight into the hormonal regulation of human sociality by showing that the hormone testosterone down regulates interpersonal trust in an adaptive manner.

Chapter 8

Testosterone affects gaze-aversion from angry faces outside of conscious awareness

Based on: Terburg, D., Aarts, H. & van Honk, J. (2012)
Psychological Science, in press

Abstract

Throughout vertebrate phylogeny, testosterone has motivated animals to obtain and maintain social dominance—a fact suggesting that unconscious primordial brain mechanisms are involved in social dominance. In humans, however, the prevailing view is that the neocortex is in control of primordial drives, and testosterone is thought to promote social dominance via conscious feelings of superiority, indefatigability, strength, and anger. Here we show that testosterone administration in humans prolongs dominant staring into the eyes of threatening faces that are viewed outside of awareness, without affecting consciously experienced feelings. These findings reveal that testosterone motivates social dominance in humans in much the same ways that it does in other vertebrates: involuntary, automatically, and unconsciously.

Introduction

The notion that individual animals, including humans, pursue dominant social positions to ensure access to resources and reproductive advantage is of great scientific and societal interest (Archer, 2006; Bos, Panksepp, Bluthé, & van Honk, 2012; Eisenegger et al., 2011; Josephs et al., 2003; Josephs et al., 2006; Mazur & Booth, 1998). All the way through vertebrate phylogeny, from reptiles to mammals, the steroid-hormone testosterone has been identified as a driving force for engaging and prevailing in confrontations for social dominance (Archer, 2006), which underlie the formation of social hierarchies (Mazur & Booth, 1998). For millions of years, testosterone evidently has acted on evolutionary primordial brain mechanisms that motivate animals to increase and maintain social status and power.

In humans, though, the expanded neocortex is thought to be in control of primordial drives, and testosterone's effects on social behavior are said to have shifted to the promotion of feelings of superiority, strength, anger, and low anxiety. In turn, these consciously experienced motivational states are said to direct voluntary control of behavior dealing with social challenges and threats (Eisenegger et al., 2011; Josephs et al., 2003; Josephs et al., 2006; Mazur & Booth, 1998). This notion, however, is currently under debate because it is based on merely correlational evidence; consequently, one cannot exclude the possibility that testosterone regulates status-seeking behaviors in humans unconsciously and automatically without affecting conscious motivational states (Bos et al., 2012).

In earlier research, we showed that salivary testosterone levels were associated with attentional vigilance to angry faces (van Honk et al., 1999), and that testosterone administration increased cardiac reactivity to angry faces (van Honk et al., 2001b). A third study demonstrated that testosterone administration increases amygdala reactivity to angry (relative to happy) faces (Hermans et al., 2008). These findings converge to suggest that testosterone enhances vigilance toward social signals of dominance (i.e., angry faces). In these studies, however, the facial expressions were perceived consciously, whereas our hypothesis has been that testosterone increases vigilance, or dominance, primarily through automatic, unconscious mechanisms (van Honk & Schutter, 2007b; van Honk, Schutter, Hermans, & Putman, 2004b). Although other researchers found correlational support for this hypothesis (Wirth & Schultheiss, 2007), it has not yet been confirmed with causal

methodology. Here, we report a placebo-controlled study of the effects of testosterone administration in which we not only used infrared eye-tracking to measure a social-dominance behavior that was much more ecologically valid than the measures in our earlier studies, but also used a backward-masking technique to ensure that the facial expressions did not reach consciousness.

We administered testosterone and placebo to 20 healthy volunteers and tested effects on performance on a social dominance task, as well as self-reports on a widely used inventory that assesses conscious motivational states of anger, vigor, fatigue, anxiety and depression. In the social-dominance task, faces were presented outside of conscious awareness, and eye movements were tracked to assess participants' inclination to either gaze away from (submission) or endure (dominance) face-to-face status threats in the form of angry stares (Terburg et al., 2011). Thus, we measured genuine gaze aversion from masked angry faces and tested the causal role of testosterone in promoting social dominance unconsciously.

Methods

Participants and design

Twenty healthy volunteers (age range: 20-25 years) received sublingual testosterone and placebo in counterbalanced order, with the two tests separated by 1 week. We exclusively recruited women using single-phase contraceptives for several reasons. First, this minimized menstrual-cycle effects on basal hormone levels. Second, the magnitude and time-course of the neurophysiological effects of testosterone have been established only in women (Tuiten et al., 2000). Third, basal testosterone levels in females have been shown to correlate both with aggressive behavior and implicit measures of dominance (Cashdan, 1995; Dabbs Jr. & Hargrove, 1997; Josephs et al., 2003; Josephs et al., 2006).

Drug samples

Sublingual drug-samples consisted of 0.5 mg of testosterone, 5 mg of cyclodextrin (carrier), 5 mg of ethanol, and 0.5 ml of water. The placebo samples were the same except that testosterone was omitted. Sublingual administration of testosterone induces behavioral and physiological effects, as indexed by subjective and vaginal arousal to erotic stimuli, that peak

after 4 hr (Tuiten et al., 2000). Accordingly, experimental testing was started 4 hr after drug (and placebo) administration. Note that this method has been used successfully in more then a dozen studies on social and emotional aspects of human behavior (Bos et al., 2012).

Conscious assessment of mood state

Before performing the social-dominance task, participants completed the Profile of Mood States (Shacham, 1983), a validated 30-item questionnaire that indexes consciously experienced anger, anxiety, depression, fatigue and vigor, using visual-analogue scales.

Social-dominance paradigm

The stimuli for the social-dominance task included angry, happy and neutral faces of five men and five women. On each trial, a gray mask with a central fixation point was followed by a face that was presented in blue, green, or red for 33 ms before a mask stimulus of the same color; the masks and face had similar luminance properties. At the bottom of each face and mask display were three circles; participants were instructed that when the central stimulus turned from gray to a color, they should avert their gaze from the central fixation point to the circle with the corresponding color (see **Figure 8.1A**). The difference in gaze-aversion latency between angry and happy expressions in this task is a reliable index of dominance motives (see Terburg et al., 2011). Facial expressions were presented in a fixed sequence that was repeated five times (NxxyNyyxNNyyxNxxyN; N = neutral; x and y = angry and happy counterbalanced accross the two sessions). This order ensured that all combinations of successive trial types occurred equally often, allowing us to analyze trials following a neutral baseline separately and eliminating trial-by-trial interference of emotionally conflicting information (Etkin et al., 2006; Kunde & Mauer, 2008).

Gaze-movements were recorded with a Tobii-1750 eye tracker (Tobii Technology, Danderyd, Sweden), and gaze-aversion latency was defined as the time between face onset and first gaze on the target circle. Latencies more than 3 standard deviations from an individual's mean were excluded (2.2%).

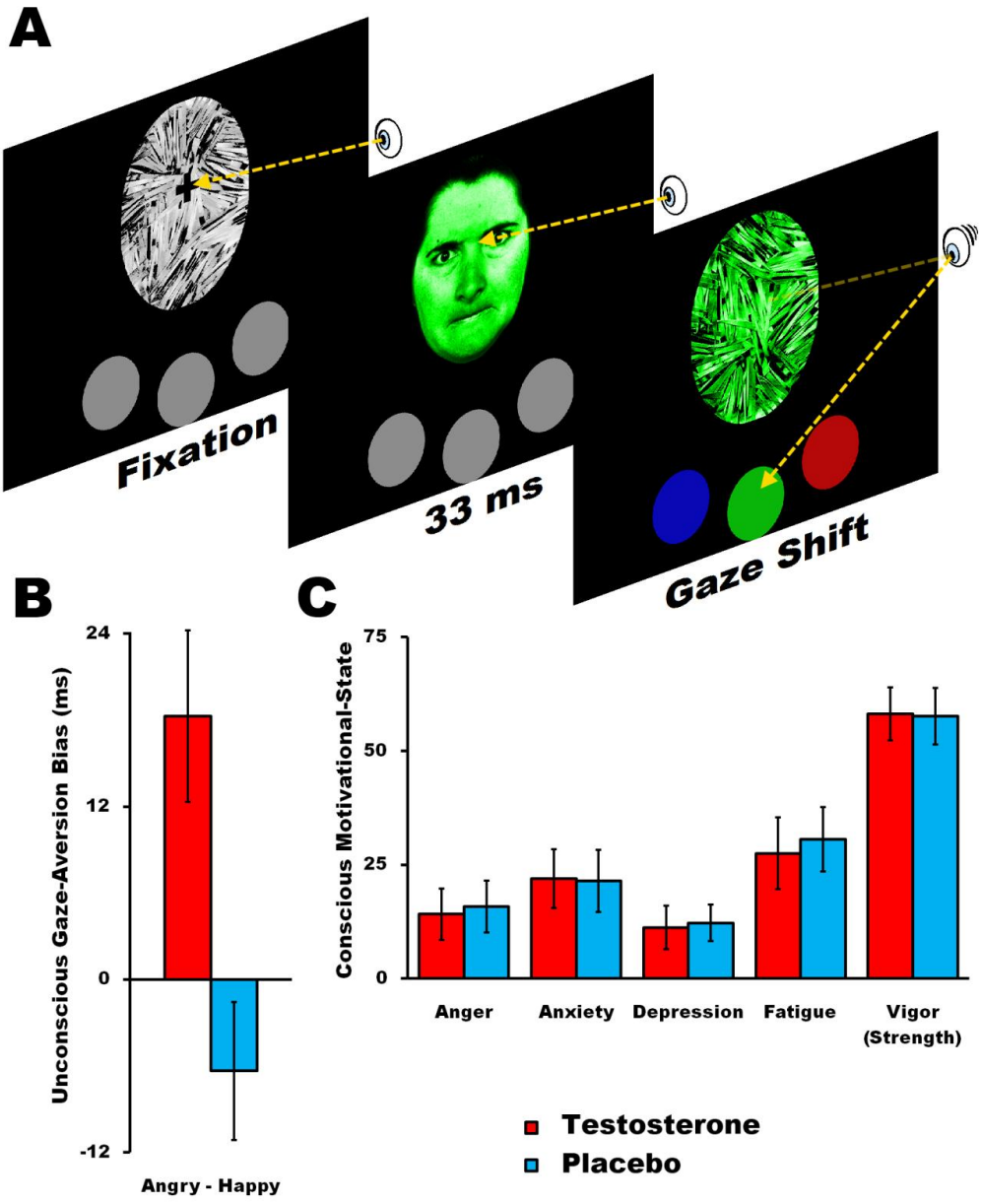


Figure 8.1 Illustration of the experimental method and results. In each trial of the social-dominance task (A), participants watched a meaningless gray picture turn blue, green, or red, at which point they were to shift their gaze downward, as fast as possible, to the circle with the corresponding color. Crucially, during the color transition, a facial expression was presented too quickly to be consciously perceived; thus, the downward gaze shift was an implicit act of gaze aversion from a social signal of reassurance (happy expression), a neutral signal (neutral expression), or a face-to-face status threat (angry expression, which rendered the gaze shift an unconscious act of submission. The graphs present (B) the mean difference in baseline-corrected gaze-aversion latency between angry and happy faces (angry – happy) and (C) the mean self-reported mood states in the two drug conditions (testosterone vs. placebo). Error bars represent standard errors of the mean.

Emotion awareness check

At the end of the final session, participants were asked whether they had seen the emotional expressions during the task. Subsequently, all 30 face stimuli (10 faces × 3 emotions) were presented again, masked, and participants were instructed to identify each facial expression as happy, angry, or neutral, in a forced-choice design.

Results

Mean latencies on angry-face and happy-face trials were baseline-corrected by subtracting the mean latency on neutral-face trials and then entered in a 2 (emotion: angry vs. happy) x 2 (drug condition: testosterone vs. placebo) repeated measures analysis of variance. The Emotion x Drug Condition interaction was significant, $F(1,19) = 8.84$, $p = .008$, $\eta_p^2 = .32$. Post hoc paired *t* tests confirmed that after testosterone administration $t(19) = 3.06$, $p = .006$, but not after placebo, $t(19) = -1.33$, $p = .201$, gaze aversion from angry faces was slower than gaze aversion from happy faces (see **Figure 8.1B**).

Next, we assessed angry- and happy-face trials that followed neutral-face trials separately. This analysis revealed a main effect of Emotion, $F(1,19) = 5.06$, $p = .037$, $\eta_p^2 = .21$, which was explained by the Emotion x Drug Condition interaction, $F(1,19) = 5.74$, $p = .027$, $\eta_p^2 = .23$ (see **Figure 8.2**). Post hoc paired *t* tests confirmed that testosterone administration slowed down gaze aversion from angry faces, $t(19) = 2.13$, $p = .046$, and not from happy faces, $t(19) = 0.10$, $p = .992$. In sum, although slower gaze aversion from angry faces compared with happy faces can be interpreted as reflecting either dominance or reduced reward-sensitivity (Terburg et al., 2011), this anger-specific effect confirms that testosterone promotes dominant-related gaze behavior.

None of the participants reported awareness of the facial expressions, but 5 scored significantly above chance level on the awareness check (i.e., > 14 correct; chance level = 10 correct; binomial test with $n = 30$, one-tailed $\alpha = .05$). Crucially, the effect of testosterone on gaze aversion remained significant (tested with one-tailed Wilcoxon signed-ranks tests because of small sample size and directed hypotheses) for both these participants ($Z = -2.02$, $p = .022$, $n = 5$), and those who were not aware of the facial expressions ($Z = -1.70$, $p = .044$, $n = 15$).

Finally, there were no effects of drug condition on self-reported mood states (all *p*'s > .5, see **Figure 8.1C**).

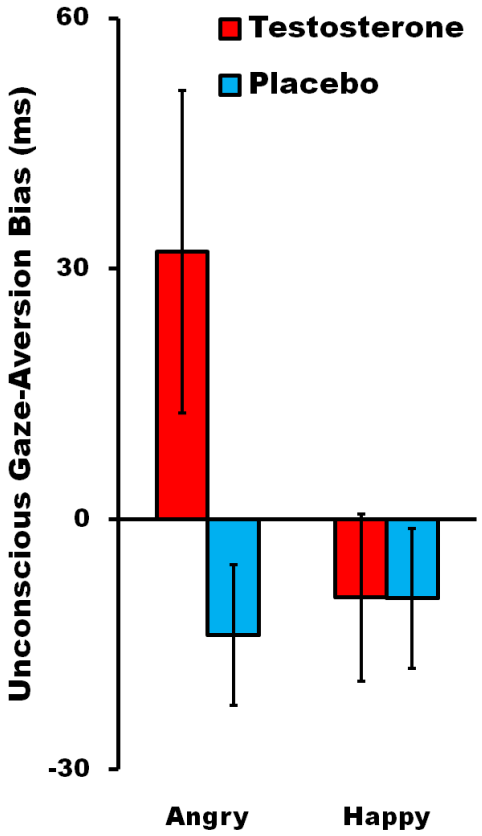


Figure 8.2 Gaze aversion latency from angry and happy faces versus the neutral baseline and controlled for trial-by-trial emotional conflict (Etkin et al., 2006; Kunde & Mauer, 2008). After testosterone administration gaze-aversion from angry faces is slower, a marker for social dominance (Terburg et al., 2011). Error bars represent standard errors of the mean.

Discussion

Our results show that after testosterone administration, participants reflexively maintain eye contact when unconsciously confronted with angry faces. Crucially, this unconscious display of dominance in face-to-face confrontations (Terburg et al., 2011) was accompanied by neither increased anger and vigor nor decreased anxiety, fatigue or depression This finding indicates that these consciously experienced motivational states do not underlie testosterone-induced social-dominance behavior.

Slower gaze aversion from angry than from happy faces has been shown to be independently related to dominance motives and reduced reward sensitivity (Terburg et al.,

2011). On the basis of these findings taken by themselves, we cannot exclude the possibility that our results are due to testosterone speeding up gaze aversion from happy faces. However, testosterone administration has previously resulted in increased reward sensitivity and appetitive motivation (Hermans et al., 2010; van Honk et al., 2004a), which makes the latter explanation unlikely. Most important, the effect of testosterone in the baseline-corrected analysis was anger-specific, which confirms our hypothesis that testosterone specifically induces dominance-related gaze behavior. Although our drug-administration method generally yields effects similar to those of endogenous testosterone in females as well as males (Bos et al., 2012; van Honk & Schutter, 2007b), future research should confirm that the results we obtained are also observed in males.

These results extend our previous findings on vigilance to consciously processed angry faces after testosterone administration (Hermans et al., 2008; van Honk et al., 2001b; van Honk et al., 1999), by showing that testosterone promotes dominance behavior towards unconsciously perceived angry faces as well. Our results add to the ongoing debate on whether testosterone promotes dominance through complex psychological mechanisms (Mazur & Booth, 1998), or reflexive biological mechanisms (van Honk et al., 2004b). Moreover, we have shown not only that testosterone vigilance to anger, but also that the hormone genuinely promotes social-dominance behavior by restraining gaze aversion when individuals are confronted with angry eye contact (Terburg et al., 2011). Although conscious psychological mechanisms unmistakably play a role in the urge for social status (Eisenegger et al., 2011; Mazur & Booth, 1998), testosterone's promotion of human social-dominance behavior evidently precedes these higher-order mechanisms. The present study thus provides compelling evidence that testosterone acts directly—involuntarily, automatically, and unconsciously—on social dominance in humans through phylogenetically ancient pathways shared with other vertebrate species (Archer, 2006; Bos et al., 2012).

Chapter 9

Sex differences in human aggression: The interaction between early developmental and later activational testosterone

Based on: Terburg, D., Peper, J. S., Morgan, B., & van Honk, J. (2009) Behavioral and Brain Sciences, 32, 290; discussion 292-311

The relation between testosterone levels and aggressive behavior is well established. From an evolutionary viewpoint testosterone can explain at least part of the sex differences found in aggressive behavior. This explanation, however, is mediated by factors such as prenatal testosterone levels and basal levels of cortisol. Especially regarding sex differences in aggression during adolescence, these mediators have great influence. Based on developmental brain structure research we argue that sex differences in aggression have a prepubertal origin and are maintained during adolescence. Evidence of prenatal, adolescent, and adult levels of testosterone in relation to aggression taken together, support Archer's (2009) argument for sexual selection as the driver of sex differences in aggression.

Archer (2009) makes a strong argument for an evolutionary basis of sex differences in aggression. His thesis is that, starting in early childhood, sex differences in aggressive behavior exist, and although these are mediated by social influences they are underlain by biological variables. One of these variables is the steroid hormone testosterone. Archer's (2009) conclusions regarding testosterone and adolescence need some refinement.

As Archer (2009) himself points out, data of self-reported aggression in male adolescents do not support findings in testosterone administration studies on aggression. He argues that although exogenous testosterone seems to enhance proneness to aggression, the rising levels of testosterone in male adolescents are not reflected in self-reported aggression measures. However, behavioral studies suggest that testosterone is a mediator of adolescent aggression. James Dabbs and colleagues showed repeatedly, in a line of studies in the nineties, associations of testosterone and violent criminal behavior. Imprisoned young males with high salivary testosterone were substantially more frequently convicted for aggressive crimes like violence and rape, and they showed more violent behavior (Dabbs Jr. et al., 1995). This was also replicated in women (Dabbs Jr. & Hargrove, 1997). Interestingly, in late adolescent males the hormone cortisol mediated the correlation between testosterone and aggressive behavior, which was found only in imprisoned adolescents with low cortisol levels (Dabbs Jr., Jurkovic, & Frady, 1991).

More recently, Popma and colleagues (2007) showed a correlation of testosterone and self-reported measures of violent behavior, but again, this was mediated by cortisol. Designed to investigate the relation among testosterone, cortisol, and aggression in early adolescence, their study pointed to an effect of testosterone on overt aggression only when cortisol levels

were low. Confirming this, Hermans and colleagues (2008) found in an fMRI study increased activity in the hypothalamus, amygdala, and orbitofrontal cortex in response to angry facial expressions. This network of brain structures is considered vital in human reactive aggression. Importantly, activity in the subcortical part of this network, namely the hypothalamus and amygdala in response to angry faces proved to be related to the ratio between testosterone and cortisol.

Based on these findings one should consider taking basal levels of cortisol into account when comparing groups on aggressive behavior. Especially during adolescence, a highly stressful period (as demonstrated hormonally by marked increases in HPA activity (Gunnar, Wewerka, Frenn, Long, & Griggs, 2009) and, behaviorally by the onset of several stress-related psychiatric illnesses (Paus, Keshavan, & Giedd, 2008)), the rise of testosterone levels (alone) in boys relative to girls will not necessarily result in a relative increase of aggressive behavior.

Another issue of consideration is testosterone in early development. Bailey and Hurd (2005) for instance, have shown that prenatal levels of testosterone, reflected in the 2D:4D digit length ratio, may mediate testosterone-aggression relationships. In males, higher prenatal testosterone levels correlated with physical aggression in adulthood. Interestingly, recent evidence from testosterone administration research in humans suggests that high prenatal testosterone levels increase sensitivity to behavioral effects of testosterone in later life (van Honk et al., 2011a). Furthermore, during early puberty (mean age 11.9 years), clear volumetric sex differences were found in brain areas mediating aggression (i.e., amygdala, striatum, rostral anterior cingulate cortex, and superior temporal gyrus; male volume > female volume) (Peper et al., 2009a). However, testosterone levels at this age could not explain these brain morphological sex differences. It might therefore be argued that a possible influence of testosterone on brain areas involved in aggression has a prenatal or early postnatal origin.

In conclusion, testosterone is unmistakably involved in human aggression and contributes importantly to sex differences in aggressive behavior. These sex differences, however, seem to originate before puberty. The relative increase of testosterone levels in adolescent boys and its relation to aggressive behavior is obscured by at least two mediators: high testosterone-sensitivity due to high prenatal testosterone levels and, especially during adolescence, levels of basal cortisol. Taking these factors into account, increased levels of

testosterone enhance aggressive behavior in both adolescent boys and girls. Thus it seems that the link between testosterone and aggression in adolescents is maintained. The here described relations between prenatal, adolescent and adult levels of testosterone, together with results found after testosterone administration, are in support of Archer's (2009) hypothesis that sex differences in aggression are a result of sexual selection.

Chapter 10

New evidence on testosterone and cooperation

Based on: van Honk, J., Montoya, E. R., Bos, P. A., van Vugt, M., & Terburg, D.

Nature, in press

In February 2010, Eisenegger et al. reported increased fair bargaining behavior after administration of testosterone in an Ultimatum Game (UG) (Eisenegger et al., 2010). Unfair offers in the UG however typically are rejected, thus not only motives for social cooperation but also threat of financial punishment may have accounted for these effects. Here with the Public Goods Game (PGG) we, unambiguously, show increased social cooperation after testosterone administration, which varies substantially with individual differences in prenatal levels of testosterone (measured by the right-hand's second-to-fourth digit-ratio). That is, testosterone promotes social cooperation exclusively among females with high 2D:4D-ratios (indicating low levels of prenatal testosterone). This finding establishes positive effects of testosterone on social cooperation, with prenatal hormonal priming providing for important individual variability.

Eisenegger et al. show increased fairness in bargaining behavior after testosterone administration in young females, and the authors suggest that this prosocial behavior is strategically driven by concerns for social status (Eisenegger et al., 2010). Indeed, in the UG such strategic concerns unmistakably play a role, and the hormone testosterone repeatedly has been associated with status concerns in humans and other animals (Bos et al., 2012; Eisenegger et al., 2011). However, unfair UG offers are typically rejected with all money being lost. Hence, threat of financial punishment may have played a role in fair bargaining behavior after testosterone administration (Eisenegger et al., 2010). We therefore tested the effects of testosterone on social-cooperative behaviors with the PGG, a game without such threat of financial punishment, wherein noncooperation can actually lead to greater profits (Van Vugt, De Cremer, & Janssen, 2007). In an experiment (approved by our ethics committee) we administered testosterone and placebo on separate days to twenty-four female students in a double-blind within-subjects design (Tuiten et al., 2000; van Honk et al., 2011b), and tested them in a three-player PGG lasting eight rounds. Each round the players received an endowment of three monetary-units (MUs), which they could either keep for themselves or contribute to the public good (Van Vugt et al., 2007); see Methods.

Using a repeated-measures Generalized Estimating Equations (GEE) analysis over all eight trials (PGG Placebo vs. Testosterone) we found no main effect of testosterone (Wald χ^2 = .048, p = .826). However, we also measured a proxy of prenatal testosterone, 2D:4D (Breedlove, 2010; Honekopp, Bartholdt, Beier, & Liebert, 2007), which has recently shown to

be a powerful predictor for effects of testosterone administration on social function (van Honk et al., 2011a); see **Methods**. With 2D:4D as covariate in the analyses the effect of testosterone on social cooperation was significant (Wald χ^2 = 9.630, p = .002) as, importantly, was the 2D:4D x testosterone interaction (Wald χ^2 = 10.140, p = .001); see **Figure 10.1A**. Next, we applied a median-split on the 2D:4D-measurements to compare individuals with relative low versus high-prenatal testosterone exposure. GEE analyses computed in both groups separately showed that subjects with low-prenatal testosterone (high 2D:4D) contributed more to the group after testosterone administration (Wald χ^2 = 7.894, p = .005), whereas subjects with high-prenatal testosterone exposure (low 2D:4D) showed no change (Wald χ^2 = 1.791, p = .181); see **Figure 10.1B**. A forced-choice test establishing that subjects were unaware of treatment condition also revealed no belief effects on public good contributions (all p's > .10). Thus, unlike in the Eisenegger et al. study, folk beliefs about testosterone did not mediate behavior in the PGG (Eisenegger et al., 2010), which is not unexpected because unfair UG offers (and not PGG non-contributions) are antisocial and risky, and fit within mainstream ideas on how testosterone affects behavior (Dabbs & Dabbs, 2000; van Honk et al., 2004a).

The present result corresponds to past research in which we also show effects of testosterone that vary strongly with prenatal testosterone exposure. In that case, high-prenatal testosterone exposure (low 2D:4D) boosted the negative impact of testosterone administration on cognitive empathy (van Honk et al., 2011a). Crucially, 2D:4D apparently is interactively shaped by testosterone and estradiol in utero, and high-2D:4D points at relative low-prenatal testosterone vs. high-prenatal estradiol (Lutchmaya et al., 2004; Manning, Scutt, Wilson, & Lewis-Jones, 1998; Zheng & Cohn, 2011). Furthermore, many effects of testosterone on social behavior are thought to arise after metabolism to estradiol (Bos et al., 2012; Eisenegger et al., 2011), but this metabolism differs between individuals (Sarachana, Xu, Wu, & Hu, 2011). Hypothetically, the balance between the sex steroids prenatally, marked by 2D:4D (Lutchmaya et al., 2004; Manning et al., 1998; Zheng & Cohn, 2011), is predictive for the rate of metabolism of testosterone into estradiol. Concretely, subjects who are prenatally more strongly primed by estradiol also metabolize more testosterone into estradiol (Zheng & Cohn, 2011), and this at present caused the selective effect in our high 2D:4D-group. Further research is necessary to test these hypotheses. Presently we challenge Eisenegger et al. by establishing positive effects of testosterone on social cooperation in which prenatal sex-

hormone priming approximated by 2D:4D conveys important individual variability. These data have strong implications for past and future hormone research.

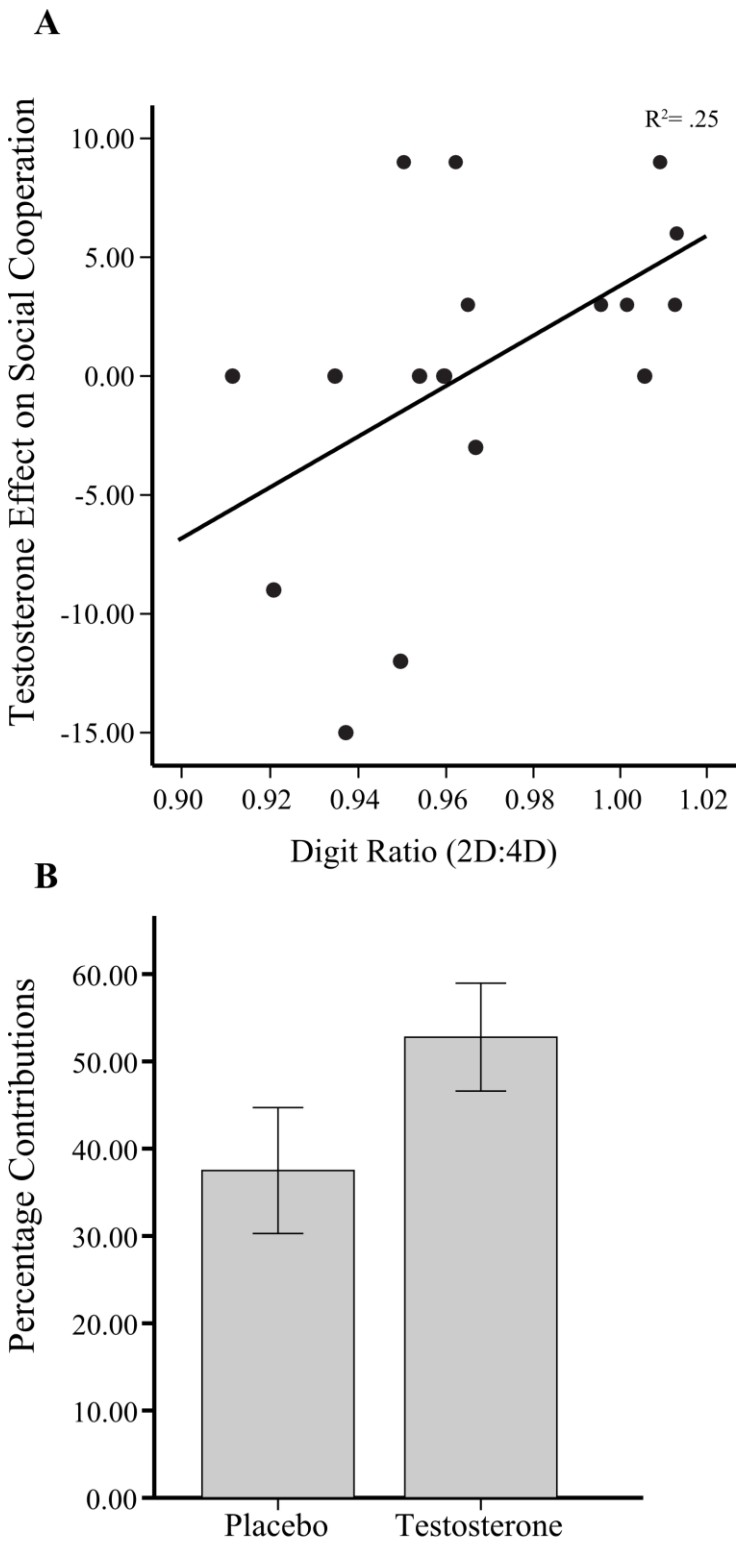


Figure 10.1 A) Individual 2D:4D measurements plotted against testosterone effect on social cooperation (mean amount of MUs contributed in the placebo condition subtracted from mean amount of MUs contributed after testosterone administration). Line depicts regression wherein digit ratio explains 25% of the variance in the overall effect of testosterone on cooperation. **B)** Mean and *SEM* of percentages overall PGG contribution after testosterone and placebo in subjects with relative high 2D:4D on basis of median split: Significantly more overall contribution to the public good after testosterone compared to placebo in high 2D:4D subjects.

Methods

PGG: All three players receive 3MU per round and can contribute all-or-nothing to the public good. Only when at least two players contribute does each player receive an extra 6MU irrespective of whether they made a contribution; thus non-contributors can profit most (9MU). To create three-person groups, confederates were used to ensure there were three players involved each time, and their decisions were randomized. It was carefully checked after the experiment that no suspicions were raised about this procedure (6 non-believers were excluded from analyses). 2D:4D: Subjects' right hands were scanned, and digit ratios were computed twice by an experienced rater (Millet & Dewitte, 2006) (correlation between measurements $p < 0.0001$).

Chapter 11

Further notes on testosterone as a social hormone

Based on: van Honk, J., Terburg, D., & Bos, P. A. (2011)
Trends in Cognitive Sciences, 15, 291-292

The hormone testosterone has a bad reputation in terms of how it influences our social behavior. According to the general public, testosterone induces violence and aggression (Eisenegger et al., 2011), and in the scientific literature the hormone is victimized as the chemical source of antisocial and immoral behavior, with high-testosterone individuals having psychological profiles that compare to sociopaths (Carney & Mason, 2010). In their passionate and insightful book, Heroes, Rogues and Lovers: On Testosterone and Behavior (2000), James and Mary Dabbs show that these views are mistaken. Dabbs and Dabbs argue that testosterone can in certain conditions motivate rebellious, aggressive and violent behavior but these conditions will mostly involve social dominance competition. However, in other conditions, testosterone can motivate behaviors that are extremely prosocial and altruistic, especially in individuals holding socially protective positions in society, such as firefighters, police officers and soldiers (Dabbs & Dabbs, 2000). The effects of the steroid hormone testosterone heavily depend on the social situation, and it is unlikely that this natural bodily fluid has instant antisocial or prosocial properties. Similarly, the peptide hormone oxytocin, popularly known as the 'love-drug', is not unconditionally a prosocial hormone. Oxytocin promotes ethnocentrism: its love is biased to the in-group and can come at the expense of out-group hate (De Dreu et al., 2010).

In their TiCS review 'The role of testosterone in social interaction' Eisenegger et al. (2011) correctly shift the discussion on testosterone and human behavior away from the simple context of social aggression. They discuss the steroid as an adaptive social hormone fulfilling a vital role in status-seeking behaviors, and the subsequent formation of social hierarchies. In this status seeking the authors distinguish between anonymous social–economical interactions, which are their main research interest, and direct face-to-face dominance contests, a focus area in our research. Our behavioral, psycho-physiological and neuroimaging data in this respect have repeatedly shown that testosterone upregulates social vigilance in response to status threats, and that in this process the amygdala seems to play a key role (Bos et al., 2012; van Honk & Schutter, 2007b). In this commentary we would like to propose two distinct neurobiological mechanisms in which interaction of the hormone and the social environment could increase social vigilance (**Figure 11.1**). Firstly, in social confrontations with low threats to status, that is most of human (economic) interaction, testosterone's upregulation of dopamine action in the orbitofrontal cortex (OFC) functionally

decouples the OFC and the amygdala (Aubele & Kritzer, 2011; Blasi et al., 2009). Accordingly, there is loss of OFC inhibition over the amygdala and the brain runs in a safeguarding mode (Bos et al., 2012; van Wingen et al., 2010) that can have antisocial but also prosocial outcomes (Dabbs & Dabbs, 2000; Eisenegger et al., 2011). Testosterone also upregulates the gene expression of vasopressin neurons in the amygdala that, when individuals encounter major threats to status or resources, results in increased activation in the social alarm pathway to the brainstem. Testosterone by way of this ultimate mechanism upholds individuals' readiness to defend status and resources with physical aggression (Bos et al., 2012).

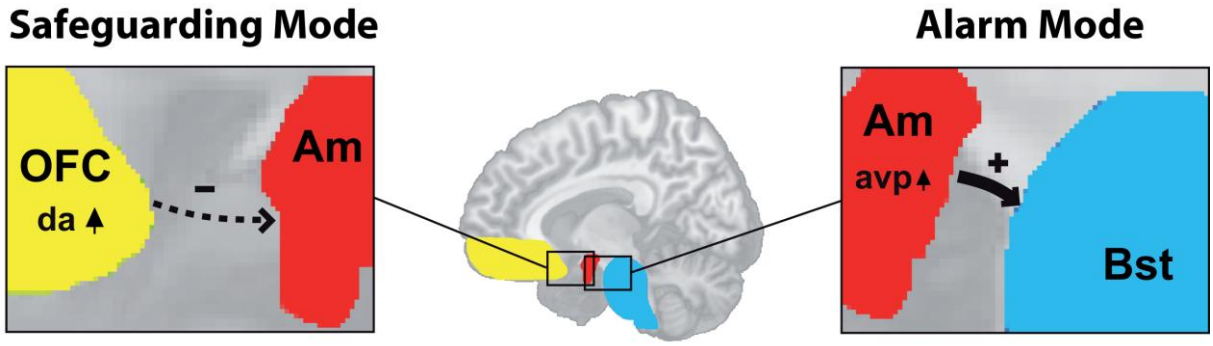


Figure 11.1 Actions of testosterone on the brain under increasing status threat. **Safeguarding Mode:** Testosterone's action on the brain during low status threat. Testosterone upregulates dopamine action in the OFC that induces decoupling of the OFC and the amygdala. Consequently, there is impaired inhibitory control of the OFC over the amygdala and the brain runs in 'safeguarding mode'. **Alarm Mode:** Testosterone's action on the brain during high status threat. In addition to the decoupling of the OFC and amygdala, testosterone, by upregulating vasopressin gene expression in the amygdala, induces hyper-coupling of the amygdala and the brainstem. Under the hyper-coupling of the amygdala and brainstem the brain is in 'social alarm mode'. AM = amygdala; Bst = brainstem; da = dopamine and avp = vasopressin.

In sum, testosterone, in a stepwise manner and under rising status threat, facilitates a processing shift from the OFC towards the brainstem. This processing shift can ultimately lead to social aggression and compares to a brain-processing shift observed in the case of fear of a proximate threat (Mobbs et al., 2007). The neurobiological processes by which testosterone modulates the social brain can be strongly comparable in rodents and humans but whereas rodents rigidly dominate with social aggression utilizing the social alarm pathway,

the twofold mechanism allows for behavioral variation in human dominance contests (Bos et al., 2012). Nonetheless, there are several other pathways by which testosterone can act on our social behavior as also discussed in the review of Eisenegger et al. (2011). That these manifold mechanisms by which the hormone can act on our social brains depend on, or are selected by, the social context, throws some light on the vast complexities of hormone–behavior relations, especially in humans. To advance our understanding of these multifaceted processes is the great challenge facing the field of social neuroendocrinology.

Chapter 12

General discussion: The socio-neuro-endocrinology of status

Abstract

The steroid hormone testosterone has traditionally been discussed as a social hormone involved in aggressive behavior. Recent evidence from testosterone administration studies shows however that testosterone profoundly reduces fear outside the social context, and can also promote cooperative behavior. Drawing from rodent and primate research we propose a neural framework wherein testosterone promotes any behavior that might defend or benefit social status. First, testosterone inhibits basal fear responsivity providing for a general fearlessness that facilitates approach oriented behavior. Second, testosterone reduces cortical control over the amygdala, resulting in general vigilance for social threat that underlies any behavior that is beneficial to social status. Third, when social status is directly challenged, testosterone promotes reactive aggression through the upregulation of vasopressin gene-expression in the central-medial amygdala. Testosterone can thus promote reactive violence, but can also produce cooperative behavior as long as it is beneficial to social status.

Testosterone and general versus social threat-vigilance

Threat-vigilance is a rather broad term for the activational effect a threat can have on the brain, body and behavior. It refers to automatic preparation for fight or flight in response to threat in general, and involves sympathetic arousal and basic processing in subcortical brain areas. There is abundant literature showing that general vigilance for evolutionary salient threat cues like spiders, snakes and facial expressions of fear and anger, is automatic and reflexive (Öhman, 2005). Interestingly, individuals predisposed to anger and aggression, but also anxious individuals, are hyper-vigilant for threat. Here an important distinction should be made between social and general threat. Fear and anxiety can be non-social. As already noted, one can be afraid of spiders, snakes, heights, or anything that might be perceived as a threat. As described in **Chapters 4 & 5**, trait anxiety is related to an attentional pattern of vigilance-avoidance when confronted with threat (Mogg et al., 1997), which can already be observed at very young age (In-Albon, Kossowsky, & Schneider, 2010). In short, anxious individuals rapidly detect threats, but will subsequently avoid it to minimize internal distress.

On the other hand, trait anger is a strong predictor of approach motivation (Carver & Harmon-Jones, 2009; Harmon-Jones, 2003a), and is related to increased vigilance toward social threat (van Honk & Schutter, 2007b; van Honk et al., 2001b), but also towards potential rewards (**Chapter 4**). For aggressive individuals a social threat is therefore an opportunity for reward, and will be approached instead of avoided, which eventually can lead to violence. Importantly, when trait anger is low in anxious individuals, anxious avoidance is coupled with a lack of angry approach motivation, and these individuals become predisposed to act submissively when confronted with a social threat. Indeed, a combination of high anxiety and low anger is related to submissive behavior (Russell & Mehrabian, 1974; Smith et al., 2010), and as shown in **Chapter 5**, these individuals avoid making eye-contact with angry conspecifics, and rapidly avert gaze in the event that they do establish such eye-contact (Terburg et al., 2012b). Moreover, as shown in **Chapter 6**, this dominance-submission mechanism of gaze-aversion is implicit and reflexive (Terburg et al., 2011). In sum, whereas trait anxiety and trait anger are both predictors of general threat-vigilance independent of the social context, the dominance-submission dimension is inherently social and associated with reflexive dominant eye-contact and approach motivation on the one hand, and submissive gaze-aversion and avoidance motivation on the other hand.

As discussed throughout this thesis, testosterone can also be considered as a social hormone. It is profoundly related to aggressive behavior (Nelson & Trainor, 2007), but crucially, only to social aggression and violence towards conspecifics (Archer, 2006). Simultaneously, testosterone also has fear reducing properties. Similar to the discussion above, these can however be considered independent from the (social) context. As discussed in **Chapter 2**, the HPG and HPA axes, with their respective end-products testosterone and cortisol, form an intricate balance. The HPA-axis plays an important role in general threat-vigilance, since it boosts sympathetic activity and arousal that acutely prepares the brain and body for action in response to any stressor or threat. Testosterone inhibits HPA activity at the very start of its hormonal cascade in the hypothalamus, which reduces acute responding to general threats (Kudielka & Kirschbaum, 2005; Terburg et al., 2009a; Viau, 2002; Viau & Meaney, 1996). Thusly, testosterone can reduce basal fear responsivity, which has been causally confirmed in testosterone administration studies in humans (Hermans et al., 2007; Hermans et al., 2006a; van Honk et al., 2005).

On the other hand, testosterone seems to increase vigilant processing of threatening and especially angry faces (Derntl et al., 2009; Hermans et al., 2008; van Wingen et al., 2009). Indeed, there is causal evidence on cardiac acceleration in response to angry faces after testosterone administration (van Honk et al., 2001b), and increased activity in the brain network of aggression in response to angry faces (Hermans et al., 2008). This research culminated into the in **Chapter 8** described study that showed more socially dominant responding to subliminally presented angry faces after testosterone administration, as indexed by slower gaze-aversion (Terburg et al., 2012c). This study confirmed that testosterone induces automatic reactive dominance behavior in humans, as has previously been observed in many other species (Archer, 2006), which suggests that phylogenetically old brain structures in the subcortex are involved (Bos et al., 2012). In sum, a combination of reduced fear reactivity, but increased reflexive social dominance behavior, seems to underlie testosterone's promotion of reactive aggression.

As described in **Chapter 11**, testosterone's promotion of reactive dominance might be the result of upregulation of vasopressin gene-expression in the amygdala (van Honk et al., 2011b). Similar to testosterone, vasopressin gene-expression is related to social, but not to predatory, aggression (Wersinger, Caldwell, Christiansen, & Young, 2007), and vasopressin

levels in the amygdala are dependent on testosterone (Goudsmit, Fliers, & Swaab, 1988; Goudsmit, Luine, & Swaab, 1990; Koolhaas et al., 1990). Moreover, social maternal aggression is specifically linked to the vasopressin system in the amygdala (Bosch & Neumann, 2011), and is argued to work through the amygdala's direct connections to subcortical structures, namely the hypothalamus and brainstem (Huber et al., 2005; Koolhaas et al., 1990; Kruk et al., 1983). Finally, both testosterone and vasopressin administration studies showed a selective decline in the ability to recognize negative emotions from faces, which is argued to underlie social aggressive tendencies (Uzefovsky, Shalev, Israel, Knafo, & Ebstein, 2012; van Honk & Schutter, 2007a). Combined these findings suggest that testosterone implicitly predisposes the individual for reactive social aggression through the upregulation of vasopressin gene-expression in the amygdala (Bos et al., 2012).

Testosterone and social aggression versus cooperation

This mechanism of reactive aggression is inherently only active when triggered by a direct social threat. Testosterone however also affects behavior in non-threatening situations. For instance, testosterone reduces cognitive empathy (van Honk et al., 2011a), reduces trust (**Chapter 7**), and can increase social cooperation (**Chapter 10**). As discussed in **Chapter 11**, a possible mechanism for these effects is through a decrease of cortical control over the amygdala, which would result in a 'safeguarding mode' that drives higher-order behavior to defend or increase social status. Endogenous testosterone levels are indeed related to more aggressive decision making and simultaneously reduced activity in the orbitofrontal cortex (OFC), a brain area consistently involved in the reduction of impulsive and aggressive behavior (Mehta & Beer, 2010). When confronted with angry faces, administration of testosterone however increased activation of the OFC, along with the amygdala and hypothalamus (Hermans et al., 2008). Follow-up research showed however that despite this increase in activation of the OFC (Hermans et al., 2008), testosterone administration reduced functional connectivity between the amygdala and the OFC (van Wingen et al., 2010). In congruence, as also discussed in **Chapter 2**, testosterone administration decreased functional crosstalk between cortical and subcortical areas as measured with EEG (Schutter & van Honk, 2004). These effects can be argued to reflect reduced prefrontal control over the amygdala's

aggressive reactivity. Indeed, psychopathy, a disorder strongly related to dominance, aggression and high testosterone levels (Blair, 2003b; Glenn, Raine, Schug, Gao, & Granger, 2010), is also characterized by reduced OFC-amygdala connectivity, both functionally and structurally (Craig et al., 2009; Motzkin, Newman, Kiehl, & Koenigs, 2011). As indicated in **Chapter 11**, one of the mechanisms underlying this reduced OFC control over the amygdala is the testosterone induced upregulation of dopamine in the OFC (Aubele & Kritzer, 2011; Blasi et al., 2009; van Honk et al., 2011b). This decoupling might drive the increasingly impulsive, risk-taking and reward-seeking tendencies associated with dopamine (Clark & Henderson, 2003), and also observed in psychopathy (Blair & Mitchell, 2008), and might contribute to reactive social aggression when status is directly challenged. When the social context is non-threatening, however, OFC-amygdala decoupling can also promote cooperative behavior as long as it is beneficial for social status (Eisenegger et al., 2010; van Honk et al., in press).

Speculatively, the direction of testosterone's effects on behavior in non-threatening situations might depend on the amount of testosterone that is converted into estradiol, which has recently been argued to show profound individual differences (Sarachana et al., 2011). As discussed in **Chapter 10**, prenatal exposure to high levels of testosterone can be indexed by a low ratio between the length of the second and fourth digits of the right hand (digit-ratio, or 2D:4D), whereas individuals with a high 2D:4D have been prenatally exposed to high levels of estradiol (Breedlove, 2010; Lutchmaya et al., 2004; Millet & Dewitte, 2006). Furthermore, prenatal testosterone levels have been shown to influence brain development (Peper et al., 2009a; Peper et al., 2009b), and predict physical aggression in men (Bailey & Hurd, 2005). Hypothetically, individuals that have been prenatally primed by estradiol (high 2D:4D) might metabolize more testosterone into estradiol (Zheng & Cohn, 2011), and estradiol action on estrogen receptors might underlie the prosocial effects of testosterone found in this group (**Chapter 10**). Accordingly, in individuals prenatally primed by testosterone (low 2D:4D), metabolization into estradiol is less prevalent, and testosterone acts more directly on androgen receptors, resulting in the negative effects on cognitive empathy found in this group (van Honk et al., 2011a). To draw definite conclusion on the testosterone-estradiol interaction this model should however be further studied in future research.

In sum, the combination of reduced HPA functioning and upregulated vasopressin in the amygdala might provide for the selective effect of testosterone on social aggression. Testosterone also reduces functional connectivity between the OFC and amygdala, which provides for the safeguarding mode described in **Chapter 11**. When a direct threat to status is present, the vasopressin upregulation in the amygdala promotes social confrontation and reactive aggression (**Chapters 5 & 8**). Conversely, when no direct threat is present, OFC-amygdala decoupling promotes more complex dominance behavior, e.g. adaptive distrust (**Chapter 7**), reduced cognitive empathy (van Honk et al., 2011a), or social cooperation (**Chapter 10**). Here, however, the effects of testosterone depend on prenatal hormone exposure, which possibly determines whether testosterone primarily acts directly on the androgen receptor, leading to reduced cognitive empathy (van Honk et al., 2011a), or indirectly on the estrogen receptor after metabolization into estradiol, leading to increased social cooperative behavior.

Testosterone and heterogeneous amygdala functioning

It is important to note that the amygdala is not a homogeneous structure. It consists of several interconnected nuclei that are mutually excitatory, but also inhibitory. In **Chapter 3** a neural framework was put forward based on a multimodal study in human subjects with selective damage to the basolateral amygdala (BLA). This study showed that damage to the BLA was associated with hyper-vigilant reflexive responding to facial threat. On the basis of these data it was argued that the control of basal threat-vigilance is most likely a joint function of the BLA and the OFC. Through attenuation of the output functions of the central-medial amygdala (CMA) the BLA and OFC can decrease sympathetic threat responding in the hypothalamus and brainstem. Additionally, based on recent rodent research it was argued that the BLA might directly inhibit the CMA, which is particularly involved in the reduction of acute fear responsivity (Macedo et al., 2005; Macedo et al., 2007; Macedo et al., 2006; Tye et al., 2011).

An interesting parallel can be found in testosterone administration studies that also show reduced acute fear responsivity, whereas conscious anxiety is generally not affected (Hermans et al., 2007; Hermans et al., 2006a; van Honk et al., 2005). As discussed above,

testosterone also upregulates vasopressin gene-expression in the amygdala, which was argued to promote reactive aggression, but vasopressin is also involved in fearful behavior (Bos et al., 2012). Thus, the question arises whether testosterone only affects output of the CMA in relation to aggression, or also affects BLA-CMA inhibition in relation to acute fear responsivity.

To answer this question we should first establish where in the amygdala testosterone upregulates vasopressin gene-expression. Interestingly, blockade of vasopressin receptors in the BLA, but not in the CMA, is anxiolytic (Salomé, Stemmelin, Cohen, & Griebel, 2006). The effects of testosterone on the BLA's vasopressin system has to our knowledge not yet been studied, but given that testosterone is also anxiolytic (Hermans et al., 2007; Hermans et al., 2006a; van Honk et al., 2005), it seems unlikely that testosterone upregulates vasopressin in the BLA. It has however been established that testosterone upregulates vasopressin gene-expression in the CMA (Goudsmit et al., 1988; Goudsmit et al., 1990; Szot & Dorsa, 1994). Moreover, this is indeed linked to the promotion of aggressive action in subcortical areas (Bosch & Neumann, 2011; Huber et al., 2005; Koolhaas et al., 1990; Kruk et al., 1983), and it has been argued that vasopressin in the CMA plays a role in the profound sex-differences in aggressive behavior (de Vries, 2008; Scordalakes & Rissman, 2004). These results combined therefore support the view that the effects of testosterone on reactive aggression and defense of social status work mainly through its effects on vasopressin in the CMA and not the BLA.

This, however, does not exclude that testosterone also affects the BLA, especially since it has been shown that pathologically aggressive dogs have increased numbers of androgen receptors in the BLA (Jacobs, Van Den Broeck, & Simoens, 2006). Given that single administrations of testosterone reduce the acute fear response (Hermans et al., 2007; Hermans et al., 2006a), it might be the case that testosterone directly promotes the inhibitive effects of the BLA on acute fear responsivity. Indeed, with exactly the same implicit fear vigilance paradigm that was used in **Chapter 3** to show that BLA damage results in fear hyper-vigilance, it was shown that testosterone administration results fear hypo-vigilance (van Honk et al., 2005). This suggests that testosterone might directly upregulate the acute anxiolytic properties of the BLA. It has indeed been shown that a similar inhibitory and anxiolytic pathway from lateral to medial CMA as described in **Chapter 3** (Ciocchi et al., 2010; Haubensak et al., 2010), which is the pathway by which the BLA inhibits CMA output

(Tye et al., 2011), is independent from vasopressin (Huber et al., 2005; Veinante & Freund-Mercier, 1997). Upregulation of vasopressin gene-expression by testosterone in the CMA thus leaves the anxiolytic effects of the BLA on the CMA intact. Hypothetically, the acute fear reducing properties of the BLA might therefore be upregulated by testosterone through its direct effects on the androgen receptor. This should however be tested in future research.

Neural framework of testosterone and social dominance

In summary, testosterone decreases fear-vigilance directly by reducing HPA functioning at the level of the hypothalamus, increases aggressive vigilance in the CMA by upregulating vasopressin gene-expression, and reduces prefrontal control over the CMA through, among other mechanisms, the upregulation of dopamine in the OFC. Additionally, testosterone might promote the inhibitive function of the BLA in acute fear responsivity. This combination of effects might explain the dual role testosterone seems to have on vigilant behavior, with on the one hand reductions in acute fear-vigilance (Hermans et al., 2007; Hermans et al., 2006a; van Honk et al., 2005), and on the other hand increasing threat-vigilance in the case of direct social dominance confrontations (Hermans et al., 2008; Terburg et al., 2012c). Furthermore, it provides a framework wherein testosterone underlies higher-order behavior that is beneficial for social status. Depending on prenatal hormone exposure, testosterone can decrease sociality by reducing cognitive empathy, but can also increase social cooperation. In other words, this framework describes several intricacies of how testosterone prepares the individual for the defense of social status. Rooted in evolutionary biology, it shows that testosterone's influence on reactive social aggression is automatic, unconscious and subcortically driven by phylogenetically ancient brain mechanisms humans share with most other vertebrates. The framework furthermore extends the influence of testosterone to higher-order human functioning, by showing how the hormone predisposes the individual to be always safeguarding its position in social hierarchy. Since aggressive violence is in our complex society not always the best option for the defense or increase of social status, these effects of testosterone on higher-order reasoning can also lead to socially cooperative, or prosocial, behaviors. Thus, testosterone adaptively promotes any social behavior that defends or increases social status and dominance.

Final remarks and future directions

The research and biological framework bundled in this thesis can be used as guideline in several lines of social neuroscientific research, and might have particular relevance for the diagnosis and treatment of psychopathologies of aggression and anxiety. First, the novel insights on human amygdala functioning transgresses the old scientific boundaries of homogeneous amygdala functions in human neuroscience. The interpretation of old and new research on the amygdala's role in human social behavior can therefore be more efficiently rooted into the underlying biology through the translation of heterogeneous amygdala functions from animal research to the human case. Second, the newly developed eye-tracking paradigms can be, and already are, used in clinical settings to increase our understanding of psychopathological aggression and anxiety, which might eventually accumulate into the development of objective tools for the diagnosis of these disorders. Third, the here described role of testosterone in status seeking behavior, with particular attention to the distinctions and parallels between reactive aggression and social cooperation, is an important step forward in our understanding on how the hormone affects the brain, and how this relates to the expression of social behavior. Testosterone is not simply a hormone of aggression, but promotes reflexive dominance contesting through eye-contact, influences highly complex social behavior, and can thereby even promote social cooperation. Thus testosterone, as reflected in the title of this thesis, "Look at Me!", promotes the status enhancing effects of contributing to the public good and community, but testosterone also promotes nonverbal dominance competitions in the form of staring contests and reactive aggression. Since testosterone also has important anxiolytic properties, and low basal testosterone levels are associated with anxiety especially within the social realm (Giltay et al., 2012), this intricate relationship of testosterone with status might benefit future research on how and when testosterone can be used in the treatment of these disorders (Haglund et al., 2007).

Samenvatting in het Nederlands

Als we een gesprek voeren, of een spelletje spelen, of op enige andere manier sociaal bezig zijn, maken we regelmatig oogcontact en volgen de blik van anderen om meer te weten te komen over elkaar en de sociale situatie. Over het algemeen zijn we ons er daarnaast goed van bewust dat anderen dit ook doen, en dus gebruiken we kijk-richting en oogcontact zelf ook om betekenis over te brengen en non-verbaal te communiceren. Mensen zijn dan ook zeer gevoelig voor kijk-richting en oogcontact, en we reageren hier dan ook onbewust en automatisch op, wat erop wijst dat er zich in de loop van de evolutie een biologisch systeem heeft gevormd om de kijkrichting van anderen snel en reflexief te verwerken. Efficiënte non-verbale sociale communicatie kan inderdaad de overlevingskansen van een sociale groep vergroten, want men kan elkaar wijzen op gevaar, of efficiënt werken in groepen tijdens het jagen en verzamelen van voedsel. Oogcontact kan echter ook een belangrijke rol spelen bij de vorming van dominantie hiërarchieën binnen de sociale groep, en draagt zo ook bij aan de overlevings- en voortplantingskansen van individuen binnen zo'n groep. In sociale primaten gebruiken dominante alfa-mannetjes bijvoorbeeld oogcontact met ondergeschikte soortgenoten om conflicten op te lossen zonder deze te laten escaleren in gewelddadigheden. Zulke non-verbale uiting van sociale dominantie kunnen worden opgevat als korte staar-wedstrijden waarbij het afwenden van de blik, en dus het verbreken van oogcontact, een signaal van onderdanigheid en ondergeschiktheid is. Aangezien dominante apen doorgaans het beste voedsel krijgen en vaak exclusieve paringsrechten hebben, wat in mindere mate ook geldt voor dominante mensen, is non-verbale communicatie dus ook gunstig voor de overlevingskansen van het individu.

In dit proefschrift maken wij gebruik van de automatische verwerking van oogcontact en gezichtsuitdrukkingen van angst en boosheid om de biologische en neurologische factoren te bestuderen die bovenstaand en ander sociaal-dominant gedrag beïnvloeden. Hierbij zal bijzondere aandacht worden besteed aan hoe het steroïde hormoon testosteron hierbij betrokken is, omdat testosteron aan de ene kant automatische, reflexieve en non-verbale gedrag dat kan leiden tot agressie bevordert, maar het kan paradoxaal genoeg mensen ook aanzetten tot meer samenwerking. In dit proefschrift worden deze schijnbaar tegengestelde gedragingen gecombineerd in één biologisch kader, en we betogen dat testosteron gedrag bevordert dat gunstig is voor de sociale status, wat afhankelijk van de sociale context tot

uiting kan komen in verschillende soorten gedrag: dominante aggressie, of sociale samenwerking.

Als basis voor dit proefschrift volgen we de theorie dat het meeste gedrag, van mens en dier, kan worden beschreven als toenaderings- of vermijdingsgedrag (**Hoofdstuk 2**). Over het algemeen steken we geen energie in dingen die geen waarde voor ons hebben, maar zodra iets wenselijk of begeerlijk is, of juist onwenselijk of gevaarlijk, zijn we gemotiveerd om actie te ondernemen. In het eerste geval zullen we toenaderingsgedrag vertonen en in het tweede geval zullen we het gevaar juist proberen te vermijden. Het evalueren van een bepaalde situatie of gebeurtenis kan grotendeels automatisch gebeuren, en dit proces beïnvloedt onze doelen en motieven dan ook niet alleen op bewust niveau, maar ook op onbewust niveau. In principe is dit dan ook weer een cruciaal aspect van evolutie dat mens en dier niet alleen motiveert om potentiëel gevaar te vermijden, maar ook om actief opzoek te gaan naar voedsel en andere levensbehoeften. Daarmee verhoogt dit mechanisme de kans op overleving en reproductie, en ligt het ook aan de basis van hoe nieuw gedrag aangeleerd wordt.

Daarnaast hebben toenaderings- en vermijdingsgedragingen ook belangrijke gevolgen voor hoe angst- en agressiestoornissen zich ontwikkelen. Cruciaal hierbij is dat agressie in principe toenaderingsgedrag is, terwijl angst een vorm is van vermijdingsgedrag. Dit is met name van belang vanuit een sociaal perspectief, omdat angst en agressie stoornissen meestal zijn gekoppeld aan een sociale context (bijvoorbeeld angst voor sociale-evaluatie, of huiselijk geweld) wat ook een belangrijke factor is in hoe deze stoornissen gediagnosticeerd worden. Mensen met angstige en agressieve persoonlijkheidskenmerken reageren echter fundamenteel anders op sociale dreiging. Hoewel zowel angstige als agressieve mensen zeer alert zullen reageren, zal een angstig persoon de sociale dreiging vervolgens defensief vermijden, terwijl een agressief iemand vaker de confrontatie aan zal gaan.

Deze persoonlijkheidskenmerken zijn sterk verbonden met neurologische en biologische factoren. Ten eerste bestaat er een evenwicht tussen de 'lage' subcorticale hersengebieden, die we delen met andere gewervelde dieren, en de 'hoge' corticale hersengebieden, die later ontwikkeld zijn in zoogdieren en mensen. Subcorticale hersengebieden zijn voornamelijk betrokken bij automatische en reflexieve reacties op dreiging, terwijl de cortex meer betrokken is bij bewuste reacties en het nemen van beslissingen. Ten tweede zijn daar de steroïde hormonen testosteron en cortisol. Steroïde

hormonen beïnvloeden zowel reflexief als bewust gedrag. Over het algemeen zijn hoge niveaus van het steroïde hormoon cortisol een indicatie van verhoogde waakzaamheid voor algemene dreiging en een vermijdend, angstig karakter. Cortisol wordt dan ook een stress-hormoon genoemd. Testosteron inhibeert de hypothalamus-pituitary-adrenal (HPA) as, welke als eindproduct cortisol heeft en vooral actief is gedurende angst en stress. Testosteron wordt dan ook een angstremmende functie toegekend. Hoge testosteron niveaus zijn ook gerelateerd aan waakzaamheid maar alleen voor sociale dreiging zoals boze gezichtsuitdrukkingen. Terwijl het stress-hormoon cortisol vooral in verband gebracht is met vermijding van dreigende situaties, is testosteron vaak gerelateerd aan agressief toenaderingsgedrag en een dominante persoonlijkheidsstijl. Vooral als testosteron niveaus hoog zijn, en cortisol niveaus laag, zijn mensen dan ook vatbaar voor sociale agressie; ze hebben de motivatie om dominant te zijn, en zijn niet bang om daarnaar te handelen.

Een belangrijke hersenstructuur voor toenaderings- en vermijdingsgedrag in het algemeen, en angst en aggressie in het bijzonder, is de amygdala. Vaak wordt de amygdala omschreven als een homogene hersenstructuur betrokken bij vele vormen van sociaal en emotioneel gedrag, en de amygdala wordt vaak als cruciaal beschouwd voor reflexieve en onbewuste reacties op dreiging. Welbeschouwd is de amygdala echter geen homogene hersenstructuur. De verschillende subkernen van de amygdala zijn dusdanig verschillend in structuur en verbindingen dat ze het best als afzonderlijke hersenstructuren beschouwd kunnen worden. In **Hoofdstuk 3** van dit proefschrift beschrijven we daarom een hersennetwerk dat betrokken is bij snelle en automatische waakzaamheid voor dreiging, op basis van de functies van subkernen van de amygdala. Met een combinatie van technieken (onbewuste waarneming, oog-bewegingsmetingen, structurele en functionele hersenscans: sMRI en fMRI), tonen we dat patiënten met hersenletsel aan de basolaterale subkern van de amygdala (BLA), maar een nog steeds functionerende centraal-mediale amygdala (CMA), hyper-alert reageren op angstige gezichten. Aangezien mensen een aangeboren neiging hebben zeer waakzaam te reageren op angstige gezichten, en een lange lijn aan onderzoek in knaagdieren laat zien dat aangeboren anst toeneemt na BLA-schade, stellen we dat een belangrijke functie van de BLA is om acute en aangeboren angst reacties in subcorticale gebieden te inhiberen. In andere woorden, de BLA werkt angstverlagend als het gaat om acute dreiging of paniek reacties. Onderdrukking van deze basale angst reacties helpt de mens meer

rationeel te reageren in geval van dreiging, en geeft mensen daardoor bijvoorbeeld ook de mogelijkheid om dominant agressief te reageren, zoals tijdens de eerder beschreven staar-wedstrijden voor dominantie.

In de volgende drie hoofdstukken van dit proefschrift beschrijven we de mechanismen die direct ten grondslag liggen aan de relatie tussen oogcontact en sociale dominantie in drie nieuw ontwikkelde oogbewegingsexperimenten. In deze studies wordt kijkgedrag gekoppeld naar toenaderings- en vermijdingsgedrag, en geïnterpreteerd vanuit wat we weten over de functies van de amygdala. De eerste studie is een kijk-imitatie experiment in **Hoofdstuk 4**. In deze studie gebruiken we een bestaande techniek voor het meten van oogbewegingen, maar niet op de gebruikelijke wijze als een maat voor waar, hoe vaak, en hoe lang mensen naar bepaalde dingen kijken. In plaats daarvan hebben we een reactieve oogbewegingstaak gemaakt, waarbij de deelnemers reageren op gebeurtenissen op een computerscherm door zo snel mogelijk naar een van tevoren vastgestelde locatie op het scherm te kijken. Door het manipuleren van de sociale context op het beeldscherm kunnen we meten hoe deze sociale informatie het natuurlijke kijkgedrag beïnvloedt. We tonen in deze studie dat mensen een waargenomen oogbeweging, dat wil zeggen als je oogcontact maakt met iemand die vervolgens wegkijkt naar links of rechts, reflexmatig volgen. Met andere woorden, als je iemand ziet wegkijken, kijk je automatisch in dezelfde richting. Belangrijk hierbij is dat deze reflex sterker wordt wanneer de waargenomen oogbeweging gepaard gaat met een gezichtsuitdrukking van angst. Een angstige oogbeweging is een signaal van op handen zijnde dreiging, dus de verhoogde reflex om deze oogbeweging te volgen bevestigt de algemene tendens van waakzaamheid in mensen. Bovendien waren in deze studie agressieve persoonlijkheidskenmerken gerelateerd aan een sterkere reflex in de richting van beloning (door het volgen van een oogbeweging die gepaard ging met een vrolijke gezichtsuitdrukking). Dit laatste bevestigt dat agressie inderdaad gerelateerd is aan toenaderingsgedrag en beloningsgevoeligheid.

Hoofdstuk 5 beschrijft een oogbewegings experiment ontworpen om persoonlijkheidskenmerken van enerzijds angst en onderdanigheid, en anderzijds agressie en dominantie, in kaart te brengen. Het is een studie waarin oogbewegingen worden gemeten als de deelnemers bezig zijn de locatie van een aantal emotionele en neutrale gezichten te onthouden. De studie bevestigt ten eerste dat angstige mensen bedreigende informatie (de

locatie van boze gezichten) beter onthouden dan niet-angstige mensen. Vervolgens laat de studie zien dat, vergelijkbaar met het eerder beschreven gedrag in primaten, onderdanige mensen snel oogcontact verbreken met boze gezichten door hun blik af te wenden. In **Hoofdstuk 6** richtten we ons daarom specifiek op het verbreken van oogcontact als een mechanisme van sociale dominantie en onderdanigheid. In een experiment dat wederom gebruik maakt van reactieve oogbewegingen, tonen we aan dat dominante mensen minder snel hun blik afwenden van onbewust gepresenteerde boze gezichten. In andere woorden, dominante mensen behouden, impliciet en in een reflex, oogcontact wanneer hun sociale dominantie wordt uitgedaagd.

In de volgende hoofdstukken van dit proefschrift worden de hormonale mechanismen die ten grondslag van zulk sociaal dominant gedrag liggen nader bestudeerd. **Hoofdstuk 7** begint met een experiment waarin we de relatie tussen het steroïde hormoon testosteron en betrouwbaarheid gemeten hebben. We tonen dat testosteron het vertrouwen in anderen doet afnemen, maar alleen in die deelnemers die normaal gesproken zeer goed van vertrouwen zijn. Dus, alleen mensen die kunnen worden beschouwd als 'sociaal naïef' worden minder goed van vertrouwen na testosteron toediening, wat erop wijst dat testosteron adaptief sociale waakzaamheid verhoogt, wat een essentieel aspect is van dominantie en leiderschap.

In **Hoofdstuk 8** testten we de relatie tussen testosteron en dominant gedrag direct, en laten we zien dat dit een onbewust en reflexief mechanisme is. We pasten dezelfde reactieve oogbewegingstaak toe als in **Hoofdstuk 6** en vonden dat na toediening van testosteron deelnemers langzamer hun blik afwendden van onbewust gepresenteerde boze gezichten. Testosteron werkt dus op een automatisch en reflexief mechanisme dat sociaal dominant gedrag bevordert.

In **Hoofdstuk 9** bespreken we de factoren die invloed hebben op sekseverschillen in agressie. Naast de eerder beschreven rol van cortisol in de relatie tussen testosteron en agressie, stellen we nu ook dat prenatale testosteron niveaus een belangrijke rol spelen in de gedragseffecten van testosteron. Testosteron niveaus in de baarmoeder tijdens de zwangerschap zijn doorgaans hoger wanneer het kind een jongen is en het is eerder aangetoond dat prenatale testosteron de ontwikkeling van de hersenen beïnvloedt, en fysieke agressie bij volwassen mannen voorspelt. Daarnaast heeft prenatale testosteron ook invloed op de verhouding in lengte van de tweede en vierde vinger van de rechterhand (vinger-ratio,

of 2D:4D), waardoor het opmeten van vingerlengte een manier is om prenatale testosteron niveaus te indexeren in volwassen mensen. Met behulp van deze methode is in een testosteron toedieningsstudie aangetoond dat testosteron cognitieve empathie vermindert, maar alleen bij vrouwen die prenataal waren blootgesteld aan hoge testosteron niveaus. Met behulp van een vergelijkbare methode laten we vervolgens in **Hoofdstuk 10** zien dat testosteron sociale samenwerking kan stimuleren, maar alleen bij vrouwen die prenataal werden blootgesteld aan lage testosteron niveaus.

Naast deze invloed die prenatale blootstelling aan testosteron heeft op hoe testosteron het gedrag later in het leven beïnvloedt, bevestigt de laatste studie eerder bewijs voor een positieve bijdrage van testosteron op sociale samenwerking. Deze resultaten lijken echter nogal tegenstrijdig met de eerder besproken evidentie dat testosteron sociale dominantie en agressie bevordert. De auteurs van die eerdere studie betoogden echter overtuigend dat samenwerking op sociaal gebied ook voor kan komen uit de motivatie sociale status te verhogen. In **Hoofdstuk 11** gebruiken we ditzelfde argument en postuleren een hersenmodel dat de effecten van testosteron niet alleen op dominante agressie verklaart, maar ook de effecten op sociale samenwerking, door ze beiden te koppelen naar sociale status. We volgen daarbij de constatering dat, afhankelijk van de sociale context, mannen met hoge testosteron niveaus zeer agressief kunnen zijn, maar ook uiterst coöperatief of zelfs altruïstisch. Het verschil in sociale context definiëren we daarbij als ofwel een situatie zonder directe statusbedreiging, ofwel een situatie met onmiddellijke of directe sociale dreiging. In de lage dreigingscontext is er geen directe waakzaamheid vereist, waardoor subcorticale gebieden zich relatief rustig zullen houden. Testosteron oefent in zulke situaties zijn invloed uit door de amygdala te ontkoppelen van corticale regulatie. Dientengevolge wordt de amygdala niet meer geremd en ontstaat er een algemene waakzaamheid voor potentiële dreiging. De amygdala is dan vrij om andere corticale structuren aan te sturen en stuurt zo het gedrag om de sociale status te verdedigen of zelfs te vergroten als daar mogelijkheden toe zijn. Dit kan leiden tot verminderde sociale vaardigheden, zoals een reductie in cognitieve empathie, maar kan ook leiden tot meer sociale samenwerking als de situatie daarom vraagt. In het geval dat er wel een directe bedreiging is voor de sociale status, speelt subcorticale waakzaamheid een grote rol in de effecten van testosteron. Testosteron versterkt genexpressie van vasopressine in de amygdala, die vervolgens de hypothalamus en hersenstam activeert wat leidt tot hyper-

alert reageren op sociale dreiging. Dit traject ligt ten grondslag aan de reactieve agressie die in verband wordt gebracht met testosteron, en veroorzaakt waarschijnlijk het reflexief vasthouden van oogcontact dat we hebben beschreven in **Hoofdstuk 6**.

In het afsluitende **Hoofdstuk 12** van dit proefschrift wordt het hersenmodel van heterogene amygdala functies en hun rol in waakzaamheid voor dreiging en acute angst uit **Hoofdstuk 3** gecombineerd met het hersenmodel van testosteron en dominantie uit **Hoofdstuk 11**. Belangrijk in dit verband is dat het stimuleren van de vasopressine genexpressie door testosteron in dieren-experimenteel onderzoek voornamelijk wordt waargenomen in de CMA, terwijl er tegelijkertijd aanwijzingen zijn dat testosteron de anstremmende activiteit in de BLA verhoogt. Daarnaast lijken deze angstremmende functies van de BLA onafhankelijk te zijn van vasopressine, waardoor testosterone in de amygdala tegelijkertijd angst kan verminderen en reactieve agressie kan verhogen.

Concluderend kunnen we stellen dat testosteron angstige waakzaamheid vermindert door het inhiberen van de HPA as, en agressieve waakzaamheid verhoogt door het versterken van vasopressine gen-expressie in de CMA en het verminderen van corticale controle over de CMA. Daarnaast heeft testosteron mogelijk een versterkend effect op de angstremmende functies van de BLA. Deze combinatie van effecten kan de dubbele rol verklaren die testosteron heeft op onbewuste waakzaamheid voor dreiging met enerzijds vermindering van acute angst reacties, en anderzijds verhoogde reactieve sociale agressie. Als er geen acute dreiging is beïnvloedt testosteron nog steeds het gedrag in status verhogende richting, maar dan zijn de effecten van testosteron sterk afhankelijk van de situatie en worden ze gemoduleerd door prenatale testosteron niveaus. Met andere woorden, de invloed van testosteron op reactieve sociale agressie is automatisch, onbewust en gedreven door fylogenetisch oude mechanismen in de subcortex welke de mens deelt met de meeste andere gewervelde dieren. Daarnaast stuurt testosteron bewust menselijk gedrag ter bescherming van de individuele positie in de sociale hiërarchie. Omdat agressief geweld in onze complexe samenleving niet altijd de beste optie is ter verdediging van sociale status, kunnen deze effecten van testosteron op bewust gedrag ook leiden tot sociale samenwerking en prosociaal gedrag. Testosteron bevordert dus elk sociaal gedrag dat sociale status beschermt of vergroot.

References

Aarts, H., & van Honk, J. (2009). Testosterone and unconscious positive priming increase human motivation separately. *Neuroreport, 20*, 1300-1303.

Abrams, D. (1995). Friends, Episode 1-20: The One with the Evil Orthodontist: NBC.

Adams, R. B., Jr., & Kleck, R. E. (2003). Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science, 14*, 644-647.

Adams, R. B., Jr., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion, 5*, 3-11.

Adolphs, R. (2007). Investigating human emotion with lesions and intracranial recording. In J. Allen & J. Coan (Eds.), *Handbook of Emotion Elicitation and Assessment* (pp. 426-439). New York: Oxford University Press.

Adolphs, R. (2008). Fear, faces, and the human amygdala. *Current Opinion in Neurobiology, 18*, 166-172.

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology, 60*, 693-716.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*, 68-72.

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature, 393*, 470-474.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669-672.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *Journal of Neuroscience, 15*, 5879-5891.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General, 133*, 323-338.

Amaral, D. G. (2003). The amygdala, social behavior, and danger detection. *Annals of the New York Academy of Sciences, 1000*, 337-347.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (4th Ed., Text Rev.)*. Washington, DC: Author.

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N. J., et al. (2005). Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: Intersubject variability and probability maps. *Anatomy and Embryology, 210*, 343-352.

Appenzeller, S., Chaloult, E., Velho, P., de Souza, E. M., Araujo, V. Z., Cendes, F., et al. (2006). Amygdalae calcifications associated with disease duration in lipoid proteinosis. *Journal of Neuroimaging, 16*, 154-156.

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews, 30*, 319-345.

Archer, J. (2009). Does sexual selection explain human sex differences in aggression? *Behavioral and Brain Sciences, 32*, 249-266; discussion 266-311.

Arnett, P. A. (1997). Autonomic responsivity in psychopaths: A critical review and theoretical proposal. *Clinical Psychology Review, 17*, 903-936.

Arregger, A. L., Contreras, L. N., Tumilasci, O. R., Aquilano, D. R., & Cardoso, E. M. (2007). Salivary testosterone: A reliable approach to the diagnosis of male hypogonadism. *Clinical Endocrinology, 67*, 656-662.

Aubele, T., & Kritzer, M. F. (2011). Gonadectomy and hormone replacement affects in vivo basal extracellular dopamine levels in the prefrontal cortex but not motor cortex of adult male rats. *Cerebral Cortex, 21*, 222-232.

Baas, D., Aleman, A., Vink, M., Ramsey, N. F., de Haan, E. H., & Kahn, R. S. (2008). Evidence of altered cortical and amygdala activation during social decision-making in schizophrenia. *Neuroimage, 40*, 719-727.

Bachevalier, J., & Loveland, K. A. (2006). The orbitofrontal-amygdala circuit and self-regulation of social-emotional behavior in autism. *Neuroscience and Biobehavioral Reviews, 30*, 97-117.

Bailey, A., & Hurd, P. (2005). Finger length ratio (2D:4D) correlates with physical aggression in men but not in women. *Biological Psychology, 68*, 215-222.

Balleine, B. W., & Killcross, S. (2006). Parallel incentive processing: An integrated view of amygdala function. *Trends in Neurosciences, 29*, 272-279.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*, 1-24.

Barbas, H. (2007). Specialized elements of orbitofrontal cortex in primates. *Annals of the New York Academy of Sciences, 1121*, 10-32.

Barbas, H., Saha, S., Rempel-Clower, N., & Ghashghaei, T. (2003). Serial pathways from primate prefrontal cortex to autonomic areas may influence emotional expression. *BMC Neuroscience, 4*, 25.

Barros, M., & Tomaz, C. (2002). Non-human primate models for investigating fear and anxiety. *Neuroscience and Biobehavioral Reviews, 26*, 187-201.

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron, 58*, 639-650.

Beaver, J. D., Lawrence, A. D., Passamonti, L., & Calder, A. J. (2008). Appetitive motivation predicts the neural response to facial signals of aggression. *Journal of Neuroscience, 28*, 2719-2725.

Bediou, B., Eimer, M., d'Amato, T., Hauk, O., & Calder, A. J. (2009). In the eye of the beholder: Individual differences in reward-drive modulate early frontocentral ERPs to angry faces. *Neuropsychologia, 47*, 825-834.

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., et al. (1997). Matsumoto and Ekman's Japanese and Caucasian facial expressions of emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21*, 3-21.

Blair, R. J. (1995). A cognitive developmental approach to morality: Investigating the psychopath. In S. Baron-Cohen (Ed.), *The Maladapted Mind: Classic Readings in Evolutionary Psychopathology* (pp. 85-113). New York: The Psychology Press.

Blair, R. J. (2003a). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*, 561-572.

Blair, R. J. (2003b). Neurobiological basis of psychopathy. *British Journal of Psychiatry, 182*, 5-7.

Blair, R. J. (2004). The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition, 55*, 198-208.

Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal. A case of 'acquired sociopathy'. *Brain, 123 (Pt 6)*, 1122-1141.

Blair, R. J., & Mitchell, D. G. V. (2008). Psychopathy, attention and emotion. *Psychological Medicine, 39*, 543.

Blasi, G., Lo Bianco, L., Taurisano, P., Gelao, B., Romano, R., Fazio, L., et al. (2009). Functional variation of the dopamine D2 receptor gene is associated with emotional control as well as brain activity and connectivity during emotion processing in humans. *Journal of Neuroscience, 29*, 14812-14819.

Bogels, S. M., & Mansell, W. (2004). Attention processes in the maintenance and treatment of social phobia: Hypervigilance, avoidance and self-focused attention. *Clinical Psychology Review, 24*, 827-856.

Bohus, B., de Kloet, E., & Veldhuis, H. (1982). Adrenal steroids and behavioral adaptation: Relationships to brain corticoid receptors. In D. Ganten & D. Pfaff (Eds.), *Progress in Neuroendocrinology* (Vol. 2, pp. 107-148). Berlin: Springer-Verlag.

Boissy, A., & Bouissou, M. F. (1994). Effects of androgen treatment on behavioral and physiological responses of heifers to fear-eliciting situations. *Hormones and Behavior, 28*, 66-83.

Bos, P. A., Hermans, E. J., Montoya, E. R., Ramsey, N. F., & van Honk, J. (2010). Testosterone administration modulates neural responses to crying infants in young females. *Psychoneuroendocrinology, 35*, 114-121.

Bos, P. A., Panksepp, J., Bluthé, R. M., & van Honk, J. (2012). Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: A review of single administration studies. *Frontiers in Neuroendocrinology, 33*, 17-35.

Bos, P. A., Terburg, D., & van Honk, J. (2010). Testosterone decreases trust in socially naïve humans. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 9991-9995.

Bosch, O. J., & Neumann, I. D. (2011). Both oxytocin and vasopressin are mediators of maternal care and aggression in rodents: From central release to sites of action. *Hormones and Behavior*.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition & Emotion, 14*, 789-808.

Brand, M., Grabenhorst, F., Starcke, K., Vandekerckhove, M. M., & Markowitsch, H. J. (2007). Role of the amygdala in decisions under ambiguity and decisions under risk: Evidence from patients with Urbach-Wiethe disease. *Neuropsychologia, 45*, 1305-1317.

Brecht, M., & Freiwald, W. A. (2011). The many facets of facial interactions in mammals. *Current Opinion in Neurobiology*.

Breedlove, S. M. (2010). Minireview: Organizational hypothesis: Instances of the fingerpost. *Endocrinology, 151*, 4116-4122.

Brown, L. L., Tomarken, A. J., Orth, D. N., Loosen, P. T., Kalin, N. H., & Davidson, R. J. (1996). Individual differences in repressive-defensiveness predict basal salivary cortisol levels. *Journal of Personality and Social Psychology, 70*, 362-371.

Buss, K. A., Malmstadt Schumacher, J. R., Dolski, I., Kalin, N. H., Goldsmith, H. H., & Davidson, R. J. (2003). Right frontal brain activity, cortisol, and withdrawal behavior in 6-month-old infants. *Behavioral Neuroscience, 117*, 11-20.

Cahill, L., Babinsky, R., Markowitsch, H. J., & McGaugh, J. L. (1995). The amygdala and emotional memory. *Nature, 377*, 295-296.

Carlson, J. M., Greenberg, T., & Mujica-Parodi, L. R. (2010). Blind rage? Heightened anger is associated with altered amygdala responses to masked and unmasked fearful faces. *Psychiatry Research, 182*, 281-283.

Carney, D. R., & Mason, M. F. (2010). Decision making and testosterone: When the ends justify the means. *Journal of Experimental Social Psychology, 46*, 668-671.

Carr, G. D., Fibiger, H. C., & Phillips, A. G. (1989). Conditioned place preference as a measure of drug reward. In J. M. Leibman & S. J. Cooper (Eds.), *Oxford Reviews in Psychopharmacology: Vol.1, Neuropharmalogical Basis of Reward* (pp. 265−319). New York: Oxford University Press.

Carré, J. M., Fisher, P. M., Manuck, S. B., & Hariri, A. R. (2010). Interaction between trait anxiety and trait anger predict amygdala reactivity to angry facial expressions in men but not women. *Social Cognitive and Affective Neuroscience*.

Carver, C. S. (2004). Negative affects deriving from the behavioral approach system. *Emotion, 4*, 3-22.

Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: Evidence and implications. *Psychological Bulletin, 135*, 183-204.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*, 319-333.

Cashdan, E. (1995). Hormones, sex, and status in women. *Hormones and Behavior, 29*, 354-366.

Ciocchi, S., Herry, C., Grenier, F., Wolff, S. B., Letzkus, J. J., Vlachos, I., et al. (2010). Encoding of conditioned fear in central amygdala inhibitory circuits. *Nature, 468*, 277-282.

Clark, A. S., & Henderson, L. P. (2003). Behavioral and physiological responses to anabolic-androgenic steroids. *Neuroscience and Biobehavioral Reviews, 27*, 413-436.

Coates, J. M., & Herbert, J. (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 6167-6172.

Costafreda, S. G., Brammer, M. J., David, A. S., & Fu, C. H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews, 58*, 57-70.

Craig, M. C., Catani, M., Deeley, Q., Latham, R., Daly, E., Kanaan, R., et al. (2009). Altered connections on the road to psychopathy. *Molecular Psychiatry, 14*, 946-953, 907.

Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., & Friston, K. (2007). Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. *Neuroimage, 37*, 866-875.

Dabbs, J. M., & Dabbs, M. G. (2000). *Heroes, Rogues and Lovers: On Testosterone and Behavior*. New York: McGraw-Hill.

Dabbs Jr., J. M. (1997). Testosterone, smiling, and facial appearance. *Journal of Nonverbal Behavior, 21*, 45-55.

Dabbs Jr., J. M., Carr, T. S., Frady, R. L., & Riad, J. K. (1995). Testosterone, crime, and misbehavior among 692 male prison inmates. *Personality and Individual Differences, 18*, 627-633.

Dabbs Jr., J. M., & Hargrove, M. F. (1997). Age, testosterone, and behavior among female prison inmates. *Psychosomatic Medicine, 59*, 477-480.

Dabbs Jr., J. M., Jurkovic, G. J., & Frady, R. L. (1991). Salivary testosterone and cortisol among late adolescent male offenders. *Journal of Abnormal Child Psychology, 19*, 469-478.

Dabbs Jr., J. M., & Morris, R. (1990). Testosterone, social class, and antisocial behavior in a sample of 4,462 men. *Psychological Science, 1*, 209-211.

Dabbs Jr., J. M., Ruback, R. B., Frady, R. L., Hopper, C. H., & Sgoutas, D. S. (1988). Saliva testosterone and criminal violence among women. *Personality and Individual Differences, 9*, 269-275.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Grosset/Putnam.

Damasio, A. R. (2000). A neural basis for sociopathy. *Archives of General Psychiatry, 57*, 128-129.

Darwin, C. (1872). *The Expressions of the Emotions in Man and Animals*. London: John Murray.

Davis, M., Gendelman, D. S., Tischler, M. D., & Gendelman, P. M. (1982). A primary acoustic startle circuit: Lesion and stimulation studies. *Journal of Neuroscience, 2*, 791-805.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6*, 13-34.

De Dreu, C. K., Greer, L. L., Handgraaf, M. J., Shalvi, S., Van Kleef, G. A., Baas, M., et al. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science, 328*, 1408-1411.

de Vries, G. J. (2008). Sex differences in vasopressin and oxytocin innervation of the brain. *Progress in Brain Research, 170*, 17-27.

Deaner, R. O., & Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Current Biology, 13*, 1609-1613.

Delgado, M. R. (2008). Fool me once, shame on you; fool me twice, shame on oxytocin. *Neuron, 58*, 470-471.

Derntl, B., Windischberger, C., Robinson, S., Kryspin-Exner, I., Gur, R. C., Moser, E., et al. (2009). Amygdala activity to fear and anger in healthy young males is associated with testosterone. *Psychoneuroendocrinology, 34*, 687-693.

DeVries, A. C., DeVries, M. B., Taymans, S., & Carter, C. S. (1995). Modulation of pair bonding in female prairie voles (Microtus ochrogaster) by corticosterone. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 7744-7748.

Diamond, J. M. (1997). *Guns, Germs and Steel: The Fates of Human Societies*. New York: W. W. Norton & Company.

Dickens, C. (1859). Hunted down. In R. Bonner (Ed.), *The New York Ledger*. New York City.

Dijksterhuis, A., & Aarts, H. (2010). Goals, attention, and (un)consciousness. *Annual Review of Psychology, 61*, 467-490.

Dundes, A. (Ed.). (1992). *The Evil Eye: A Casebook*. Madison, Wis.: The University of Wisconsin Press.

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., et al. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage, 36*, 511-521.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25*, 1325-1335.

Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences, 15*, 263-271.

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., & Fehr, E. (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature, 463*, 356-359.

Ekman, P. (1999a). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. New York: John Wiley & Sons Ltd.

Ekman, P. (1999b). Facial expressions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. New York: John Wiley & Sons Ltd.

Ekman, P., & Friesen, W. (1976). *Pictures of Facial Affect*. Palo Alto: Consulting Psychologist Press.

Ellis, L. (2006). Gender differences in smiling: An evolutionary neuroandrogenic theory. *Physiology & Behavior, 88*, 303-308.

Emery, N. J. (2000). The eyes have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews, 24*, 581-604.

Emery, N. J., & Amaral, D. G. (2000). The role of the amygdala in primate social cognition. In R. D. Lane & N. Lynn (Eds.), *Cognitive Neuroscience of Emotion* (pp. 156-191). New York: Oxford University Press.

Ernst, M., & Fudge, J. L. (2009). A developmental neurobiological model of motivated behavior: Anatomy, connectivity and ontogeny of the triadic nodes. *Neuroscience and Biobehavioral Reviews, 33*, 367-382.

Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R., & Hirsch, J. (2006). Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron, 51*, 871-882.

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., et al. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron, 44*, 1043-1055.

Eysenck, M. W. (2004). Applied cognitive psychology: Implications of cognitive psychology for clinical psychology and psychotherapy. *Journal of Clinical Psychology, 60*, 393-404.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 9602-9605.

Farroni, T., Johnson, M. H., & Csibra, G. (2004a). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience, 16*, 1320-1326.

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004b). Gaze following in newborns. *Infancy, 5*, 39-60.

Feinstein, J. S., Adolphs, R., Damasio, A., & Tranel, D. (2011). The human amygdala and the induction and experience of fear. *Current Biology, 21*, 34-38.

Ferrari, P. F., Paukner, A., Ionica, C., & Suomi, S. J. (2009). Reciprocal face-to-face communication between rhesus macaque mothers and their newborn infants. *Current Biology, 19*, 1768-1772.

Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A., & Suomi, S. J. (2006). Neonatal imitation in rhesus macaques. *PLoS Biology, 4*, e302.

Fowles, D. C. (1980). The three arousal model: Implications of gray's two-factor learning theory for heart rate, electrodermal activity, and psychopathy. *Psychophysiology, 17*, 87-104.

Fox, E., Mathews, A., Calder, A. J., & Yiend, J. (2007). Anxiety and sensitivity to gaze direction in emotionally expressive faces. *Emotion, 7*, 478-486.

Frewen, P., Dozois, D., Joanisse, M., & Neufeld, R. (2008). Selective attention to threat versus reward: Meta-analysis and neural-network modeling of the dot-probe task. *Clinical Psychology Review, 28*, 307-337.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*, 694-724.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 2*, 165-189.

Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 363*, 2033-2039.

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., et al. (2009). Functional atlas of emotional faces processing: A voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience, 34*, 418-432.

Gamer, M., & Buchel, C. (2009). Amygdala activation predicts gaze toward fearful eyes. *Journal of Neuroscience, 29*, 9123-9126.

Garcia, R., Vouimba, R. M., Baudry, M., & Thompson, R. F. (1999). The amygdala modulates prefrontal cortex activity relative to conditioned fear. *Nature, 402*, 294-296.

Garner, M., Mogg, K., & Bradley, B. P. (2006). Orienting and maintenance of gaze to facial expressions in social anxiety. *Journal of Abnormal Psychology, 115*, 760-770.

Giltay, E. J., Enter, D., Zitman, F. G., Penninx, B. W., van Pelt, J., Spinhoven, P., et al. (2012). Salivary testosterone: Associations with depression, anxiety disorders, and antidepressant use in a large cohort study. *Journal of Psychosomatic Research, 72*, 205-213.

Glenn, A. L., Raine, A., Schug, R. A., Gao, Y., & Granger, D. A. (2010). Increased testosterone-to-cortisol ratio in psychopathy. *Journal of Abnormal Psychology*.

Goudsmit, E., Fliers, E., & Swaab, D. F. (1988). Testosterone supplementation restores vasopressin innervation in the senescent rat brain. *Brain Research, 473*, 306-313.

Goudsmit, E., Luine, V. N., & Swaab, D. F. (1990). Testosterone locally increases vasopressin content but fails to restore choline acetyltransferase activity in other regions in the senescent male rat brain. *Neuroscience Letters, 112*, 290-296.

Gozzi, A., Jain, A., Giovanelli, A., Bertollini, C., Crestan, V., Schwarz, A. J., et al. (2010). A neural switch for active and passive fear. *Neuron, 67*, 656-666.

Graeff, F. G., & Del-Ben, C. M. (2008). Neurobiology of panic disorder: From animal models to brain neuroimaging. *Neuroscience and Biobehavioral Reviews, 32*, 1326-1335.

Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T., & Schwartz, E. B. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology, 29*, 1229-1240.

Gray, J. A. (1987). *The Psychology of Fear and Stress*. New York: Cambridge University Press.

Grossmann, T., & Johnson, M. H. (2007). The development of the social brain in human infancy. *European Journal of Neuroscience, 25*, 909-919.

Guitton, D. (1992). Control of eye-head coordination during orienting gaze shifts. *Trends in Neurosciences, 15*, 174-179.

Gunnar, M. R., Wewerka, S., Frenn, K., Long, J. D., & Griggs, C. (2009). Developmental changes in hypothalamus-pituitary-adrenal activity over the transition to adolescence: Normative changes and associations with puberty. *Development and Psychopathology, 21*, 69-85.

Haglund, M. E., Nestadt, P. S., Cooper, N. S., Southwick, S. M., & Charney, D. S. (2007). Psychobiological mechanisms of resilience: Relevance to prevention and treatment of stress-related psychopathology. *Development and Psychopathology, 19*, 889-920.

Hampton, A. N., Adolphs, R., Tyszka, M. J., & O'Doherty, J. P. (2007). Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron, 55*, 545-555.

Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., et al. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science, 297*, 400-403.

Harmon-Jones, E. (2003a). Anger and the behavioral approach system. *Personality and Individual Differences, 35*, 995-1005.

Harmon-Jones, E. (2003b). Clarifying the emotive functions of asymmetrical frontal cortical activity. *Psychophysiology, 40*, 838-848.

Harmon-Jones, E. (2004). Contributions from research on anger and cognitive dissonance to understanding the motivational functions of asymmetrical frontal brain activity. *Biological Psychology, 67*, 51-76.

Harmon-Jones, E., & Peterson, C. K. (2009). Supine body position reduces neural response to anger evocation. *Psychological Science, 20*, 1209-1210.

Haubensak, W., Kunwar, P. S., Cai, H., Ciocchi, S., Wall, N. R., Ponnusamy, R., et al. (2010). Genetic dissection of an amygdala microcircuit that gates conditioned fear. *Nature, 468*, 270-276.

Heimer, L., Harlan, R. E., Alheid, G. F., Garcia, M. M., & de Olmos, J. (1997). Substantia innominata: A notion which impedes clinical-anatomical correlations in neuropsychiatric disorders. *Neuroscience, 76*, 957-1006.

Heinrichs, N., & Hofmann, S. G. (2001). Information processing in social phobia: A critical review. *Clinical Psychology Review, 21*, 751-770.

Heitmueller, K. (2006). Rewind: Movies that Break the Bank. Retrieved December 3rd, 2007, from http://www.mtv.com/movies/news/articles/1526589/03202006/story.jhtml

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Most people are not WEIRD. *Nature, 466*, 29.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61-83.

Hermans, E. J., Bos, P. A., Ossewaarde, L., Ramsey, N. F., Fernandez, G., & van Honk, J. (2010). Effects of exogenous testosterone on the ventral striatal BOLD response during reward anticipation in healthy women. *Neuroimage, 52*, 277-283.

Hermans, E. J., Putman, P., Baas, J. M., Gecks, N. M., Kenemans, J. L., & van Honk, J. (2007). Exogenous testosterone attenuates the integrated central stress response in healthy young women. *Psychoneuroendocrinology, 32*, 1052-1061.

Hermans, E. J., Putman, P., Baas, J. M., Koppeschaar, H. P., & van Honk, J. (2006a). A single administration of testosterone reduces fear-potentiated startle in humans. *Biological Psychiatry, 59*, 872.

Hermans, E. J., Putman, P., & van Honk, J. (2006b). Testosterone administration reduces empathetic behavior: A facial mimicry study. *Psychoneuroendocrinology, 31*, 859-866.

Hermans, E. J., Ramsey, N. F., & van Honk, J. (2008). Exogenous testosterone enhances responsiveness to social threat in the neural circuitry of social aggression in humans. *Biological Psychiatry, 63*, 263-270.

Hess, U., & Bourgeois, P. (2010). You smile-I smile: Emotion expression in social interaction. *Biological Psychology*.

Hietanen, J. K., & Leppänen, J. M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1228-1243.

Holland, P. C., & Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends in Cognitive Sciences, 3*, 65-73.

Holmes, A., Richards, A., & Green, S. (2006). Anxiety and sensitivity to eye gaze in emotional faces. *Brain and Cognition, 60*, 282-294.

Honekopp, J., Bartholdt, L., Beier, L., & Liebert, A. (2007). Second to fourth digit length ratio (2D:4D) and adult sex hormone levels: New data and a meta-analytic review. *Psychoneuroendocrinology, 32*, 313-321.

Horley, K., Williams, L. M., Gonsalvez, C., & Gordon, E. (2004). Face to face: Visual scanpath evidence for abnormal processing of facial expressions in social phobia. *Psychiatry Research, 127*, 43-53.

Hubbard, J. R., Kalimi, M., & Liberti, J. P. (1997). Biochemical indicators of stress. In J. R. Hubbard & E. A. Workman (Eds.), *Handbook of stress medicine: an organ system approach* (pp. 309-322). Boca Raton, FL: CRC Press.

Huber, D., Veinante, P., & Stoop, R. (2005). Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. *Science, 308*, 245-248.

Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., et al. (2010). Oxytocin enhances amygdala-dependent, socially reinforced learning and emotional empathy in humans. *Journal of Neuroscience, 30*, 4999-5007.

Hurlemann, R., Schlaepfer, T. E., Matusch, A., Reich, H., Shah, N. J., Zilles, K., et al. (2009). Reduced 5-HT(2A) receptor signaling following selective bilateral amygdala damage. *Social Cognitive and Affective Neuroscience, 4*, 79-84.

Hurlemann, R., Wagner, M., Hawellek, B., Reich, H., Pieperhoff, P., Amunts, K., et al. (2007). Amygdala control of emotion-induced forgetting and remembering: Evidence from Urbach-Wiethe disease. *Neuropsychologia, 45*, 877-884.

In-Albon, T., Kossowsky, J., & Schneider, S. (2010). Vigilance and avoidance of threat in the eye movements of children with separation anxiety disorder. *Journal of Abnormal Child Psychology, 38*, 225-235.

Insel, T. R., & Young, L. J. (2001). The neurobiology of attachment. *Nature Reviews Neuroscience, 2*, 129-136.

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience and Biobehavioral Reviews, 33*, 843-863.

Jacobs, C., Van Den Broeck, W., & Simoens, P. (2006). Increased number of neurons expressing androgen receptor in the basolateral amygdala of pathologically aggressive dogs. *Journal of Veterinary Medicine Series A, 53*, 334-339.

Johnson, E. O., Kamilaris, T. C., Chrousos, G. P., & Gold, P. W. (1992). Mechanisms of stress: A dynamic overview of hormonal and behavioral homeostasis. *Neuroscience and Biobehavioral Reviews, 16*, 115-130.

Jones, D. (2010). A WEIRD view of human nature skews psychologists' studies. *Science, 328*, 1627.

Josephs, R. A., Newman, M. L., Brown, R. P., & Beer, J. M. (2003). Status, testosterone, and human intellectual performance: Stereotype threat as status concern. *Psychological Science, 14*, 158-163.

Josephs, R. A., Sellers, J. G., Newman, M. L., & Mehta, P. H. (2006). The mismatch effect: When testosterone and status are at odds. *Journal of Personality and Social Psychology, 90*, 999-1013.

Kalin, N. H., & Shelton, S. E. (1989). Defensive behaviors in infant rhesus monkeys: Environmental cues and neurochemical regulation. *Science, 243*, 1718-1721.

Kalin, N. H., Shelton, S. E., & Davidson, R. J. (2004). The role of the central nucleus of the amygdala in mediating fear and anxiety in the primate. *Journal of Neuroscience, 24*, 5506-5515.

Kalin, N. H., Shelton, S. E., Engeland, C. G., Haraldsson, H. M., & Marucha, P. T. (2006). Stress decreases, while central nucleus amygdala lesions increase, IL-8 and MIP-1alpha gene expression during tissue healing in non-human primates. *Brain, Behavior, and Immunity, 20*, 564-568.

Kapur, N. (1996). Paradoxical functional facilitation in brain-behaviour research. A critical review. *Brain, 119 (Pt 5)*, 1775-1790.

Kennedy, D. P., & Adolphs, R. (2010). Impaired fixation to eyes following amygdala damage arises from abnormal bottom-up attention. *Neuropsychologia, 48*, 3392-3398.

Killcross, S., Robbins, T. W., & Everitt, B. J. (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature, 388*, 377-380.

Kirsch, P., Esslinger, C., Chen, Q., Mier, D., Lis, S., Siddhanti, S., et al. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. *Journal of Neuroscience, 25*, 11489-11493.

Knyazev, G., & Slobodskaya, H. R. (2003). Personality trait of behavioral inhibition is associated with oscillatory systems reciprocal relationships. *International Journal of Psychophysiology, 48*, 247-261.

Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: Comparative studies on external morphology of the primate eye. *Journal of Human Evolution, 40*, 419-435.

Koolhaas, J. M., Van Den Brink, T. H. C., Roozendaal, B., & Boorsma, F. (1990). Medial amygdala and aggressive behavior: Interaction between testosterone and vasopressin. *Aggressive Behavior, 16*, 223-229.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435*, 673-676.

Koval, M. J., Thomas, B. S., & Everling, S. (2005). Task-dependent effects of social attention on saccadic reaction times. *Experimental Brain Research, 167*, 475-480.

Kruk, M. R., Van der Poel, A. M., Meelis, W., Hermans, J., Mostert, P. G., Mos, J., et al. (1983). Discriminant analysis of the localization of aggression-inducing electrode placements in the hypothalamus of male rats. *Brain Research, 260*, 61-79.

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: A review. *Biological Psychology, 69*, 113-132.

Kunde, W., & Mauer, N. (2008). Sequential modulations of valence processing in the emotional stroop task. *Experimental Psychology, 55*, 151-156.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry, 44*, 1248-1263.

Langton, S. R., Watt, R. J., & Bruce, I. I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences, 4*, 50-59.

LeDoux, J. (2012). Rethinking the Emotional Brain. *Neuron, 73*, 653-676.

LeDoux, J. E. (1993). Emotional memory systems in the brain. *Behavioural Brain Research, 58*, 69-79.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155-184.

Liddell, B. J., Brown, K. J., Kemp, A. H., Barton, M. J., Das, P., Peduto, A., et al. (2005). A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. *Neuroimage, 24*, 235-243.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet.

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development, 77*, 23-28.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *Journal of Abnormal Psychology, 55*, 6-10.

Macedo, C. E., Cuadra, G., Molina, V., & Brandao, M. L. (2005). Aversive stimulation of the inferior colliculus changes dopamine and serotonin extracellular levels in the frontal cortex: Modulation by the basolateral nucleus of amygdala. *Synapse, 55*, 58-66.

Macedo, C. E., Martinez, R. C., Albrechet-Souza, L., Molina, V. A., & Brandao, M. L. (2007). 5-HT2- and D1-mechanisms of the basolateral nucleus of the amygdala enhance conditioned fear and impair unconditioned fear. *Behavioural Brain Research, 177*, 100-108.

Macedo, C. E., Martinez, R. C., & Brandao, M. L. (2006). Conditioned and unconditioned fear organized in the inferior colliculus are differentially sensitive to injections of muscimol into the basolateral nucleus of the amygdala. *Behavioral Neuroscience, 120*, 625-631.

MacLean, P. D. (1990). *The Triune Brain in Evolution. Role in Paleocerebral Functions*. New York: Plenum Press.

Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction, 13*, 3000-3004.

Maren, S. (2007). The threatened brain. *Science, 317*, 1043-1044.

Marsh, A. A., & Blair, R. J. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 32*, 454-465.

Martinez, R. C., Ribeiro de Oliveira, A., & Brandao, M. L. (2007). Serotonergic mechanisms in the basolateral amygdala differentially regulate the conditioned and unconditioned fear organized in the periaqueductal gray. *European Neuropsychopharmacology, 17*, 717-724.

Mathews, A., Fox, E., Yiend, J., & Calder, A. (2003). The face of fear: Effects of eye gaze and emotion on visual attention. *Visual Cognition, 10*, 823-835.

Mazur, A. (1985). A biosocial model of status in face-to-face primate groups. *Social Forces, 64*, 377-402.

Mazur, A., & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences, 21*, 353-363; discussion 363-397.

McBurnett, K., Lahey, B. B., Frick, P. J., Risch, C., Loeber, R., Hart, E. L., et al. (1991). Anxiety, inhibition, and conduct disorder in children: II. Relation to salivary cortisol. *Journal of the American Academy of Child and Adolescent Psychiatry, 30*, 192-196.

McNaughton, N., & Corr, P. J. (2004). A two-dimensional neuropsychology of defense: Fear/anxiety and defensive distance. *Neuroscience and Biobehavioral Reviews, 28*, 285-305.

Mehta, P. H., & Beer, J. (2010). Neural mechanisms of the testosterone-aggression relation: The role of orbitofrontal cortex. *Journal of Cognitive Neuroscience, 22*, 2357-2368.

Mehta, P. H., Jones, A. C., & Josephs, R. A. (2008). The social endocrinology of dominance: Basal testosterone predicts cortisol changes and behavior following victory and defeat. *Journal of Personality and Social Psychology, 94*, 1078-1093.

Mehta, P. H., Wuehrmann, E. V., & Josephs, R. A. (2009). When are low testosterone levels advantageous? The moderating role of individual versus intergroup competition. *Hormones and Behavior, 56*, 158-162.

Milders, M., Hietanen, J. K., Leppänen, J. M., & Braun, M. (2011). Detection of emotional faces is modulated by the direction of eye gaze. *Emotion, in press*.

Millet, K., & Dewitte, S. (2006). Second to fourth digit ratio and cooperative behavior. *Biological Psychology, 71*, 111-115.

Mitte, K. (2008). Memory bias for threatening information in anxiety and anxiety disorders: A meta-analytic review. *Psychological Bulletin, 134*, 886-911.

Mobbs, D., Marchant, J. L., Hassabis, D., Seymour, B., Tan, G., Gray, M., et al. (2009). From threat to fear: The neural organization of defensive fear systems in humans. *Journal of Neuroscience, 29*, 12236-12243.

Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., et al. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science, 317*, 1079-1083.

Mogg, K., Bradley, B. P., de Bono, J., & Painter, M. (1997). Time course of attentional bias for threat information in non-clinical anxiety. *Behaviour Research and Therapy, 35*, 297-303.

Mogg, K., Millar, N., & Bradley, B. P. (2000). Biases in eye movements to threatening facial expressions in generalized anxiety disorder and depressive disorder. *Journal of Abnormal Psychology, 109*, 695-704.

Montagne, B. (2005). *Perception of Facial Expressions in Psychiatric and Neurological Disorders.* Utrecht University, Utrecht.

Montagne, B., van Honk, J., Kessels, R. P., Frigerio, E., Burt, M., van Zandvoort, M. J. E., et al. (2005). Reduced efficiency in recognising fear in subjects scoring high on psychopathic personality characteristics. *Personality and Individual Differences, 38*, 5-11.

Mosher, C. P., Zimmerman, P. E., & Gothard, K. M. (2010). Response characteristics of basolateral and centromedial neurons in the primate amygdala. *Journal of Neuroscience, 30*, 16197-16207.

Motzkin, J. C., Newman, J. P., Kiehl, K. A., & Koenigs, M. (2011). Reduced prefrontal connectivity in psychopathy. *Journal of Neuroscience, 31*, 17348-17357.

Mühlberger, A., Wieser, M. J., & Pauli, P. (2008). Visual attention during virtual social situations depends on social anxiety. *CyberPsychology & Behavior, 11*, 425-430.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723-739.

Nell, V. (2000). *Cross-Cultural Neuropsychological Assessment: Theory and Practice*. New Jersey: Lawrence Erlbaum Associates.

Nelson, R. J., & Trainor, B. C. (2007). Neural mechanisms of aggression. *Nature Reviews Neuroscience, 8*, 536-546.

Newman, M. L., Sellers, J. G., & Josephs, R. A. (2005). Testosterone, cognition, and social status. *Hormones and Behavior, 47*, 205-211.

Nomura, M., Ohira, H., Haneda, K., Iidaka, T., Sadato, N., Okada, T., et al. (2004). Functional association of the amygdala and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces: An event-related fMRI study. *Neuroimage, 21*, 352-363.

Nummenmaa, L., & Hietanen, J. (2006). Gaze distractors influence saccadic curvature: Evidence for the role of the oculomotor system in gaze-cued orienting. *Vision Research, 46*, 3674-3680.

Öhman, A. (1986). Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. *Psychophysiology, 23*, 123-145.

Öhman, A. (2005). The role of the amygdala in human fear: Automatic detection of threat. *Psychoneuroendocrinology, 30*, 953-958.

Owen, J. M. (2010). Transdiagnostic cognitive processes in high trait anger. *Clinical Psychology Review, 31*, 193-202.

Parkes, S. L., & Westbrook, R. F. (2010). The basolateral amygdala is critical for the acquisition and extinction of associations between a neutral stimulus and a learned danger signal but not between two neutral stimuli. *Journal of Neuroscience, 30*, 12608-12618.

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nature Reviews Neuroscience, 9*, 947-957.

Pavlov, I. P. (1927). *Conditioned Reflexes*. New York: Dover.

Peper, J. S., Brouwer, R. M., Schnack, H. G., van Baal, G. C., van Leeuwen, M., van den Berg, S. M., et al. (2009a). Sex steroids and brain structure in pubertal boys and girls. *Psychoneuroendocrinology, 34*, 332-342.

Peper, J. S., Brouwer, R. M., van Baal, G. C., Schnack, H. G., van Leeuwen, M., Boomsma, D. I., et al. (2009b). Does having a twin brother make for a bigger brain? *European Journal of Endocrinology / European Federation of Endocrine Societies, 160*, 739-746.

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience, 11*, 773-783.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48*, 175-187.

Popma, A., Vermeiren, R., Geluk, C. A., Rinne, T., van den Brink, W., Knol, D. L., et al. (2007). Cortisol moderates the relationship between testosterone and aggression in delinquent male adolescents. *Biological Psychiatry, 61*, 405-411.

Porges, S. W. (2001). The polyvagal theory: Phylogenetic substrates of a social nervous system. *International Journal of Psychophysiology, 42*, 123-146.

Putman, P., Hermans, E., & van Honk, J. (2004). Emotional Stroop performance for masked angry faces: It's BAS, not BIS. *Emotion, 4*, 305-311.

Putman, P., Hermans, E., & van Honk, J. (2006). Anxiety meets fear in perception of dynamic expressive gaze. *Emotion, 6*, 94-102.

Putman, P., Hermans, E. J., & van Honk, J. (2010). Cortisol administration acutely reduces threat-selective spatial attention in healthy young men. *Physiology & Behavior, 99*, 294-300.

Putman, P., Saevarsson, S., & van Honk, J. (2007). Hypomanic trait is associated with a hypovigilant automatic attentional response to social cues of danger. *Bipolar Disorders, 9*, 779-783.

Quirk, G. J., Likhtik, E., Pelletier, J. G., & Pare, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience, 23*, 8800-8807.

Raine, A., Lencz, T., Bihrle, S., LaCasse, L., & Colletti, P. (2000). Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. *Archives of General Psychiatry, 57*, 119-127; discussion 128-119.

Ressler, N. (2004). Rewards and punishments, goal-directed behavior and consciousness. *Neuroscience and Biobehavioral Reviews, 28*, 27-39.

Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking: Exploring the tendency to imitate another individual's gaze. *Neuroreport, 13*, 2259-2264.

Roelofs, K., Putman, P., Schouten, S., Lange, W.-G., Volman, I., & Rinck, M. (2010). Gaze direction differentially affects avoidance tendencies to happy and angry faces in socially anxious individuals. *Behaviour Research and Therapy, 48*, 290-294.

Rosen, J. B., & Schulkin, J. (1998). From normal fear to pathological anxiety. *Psychological Review, 105*, 325-350.

Russell, J. A., & Mehrabian, A. (1974). Distinguishing anger and anxiety in terms of emotional response factors. *Journal of Consulting and Clinical Psychology, 42*, 79-83.

Salomé, N., Stemmelin, J., Cohen, C., & Griebel, G. (2006). Differential roles of amygdaloid nuclei in the anxiolytic- and antidepressant-like effects of the V1b receptor antagonist, SSR149415, in rats. *Psychopharmacology, 187*, 237-244.

Salzman, C. D., & Fusi, S. (2010). Emotion, cognition, and mental state representation in amygdala and prefrontal cortex. *Annual Review of Neuroscience, 33*, 173-202.

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 15268-15273.

Sapolsky, R. M. (1990). Adrenocortical function, social rank, and personality among wild baboons. *Biological Psychiatry, 28*, 862-878.

Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science, 308*, 648-652.

Sarachana, T., Xu, M., Wu, R. C., & Hu, V. W. (2011). Sex hormones in autism: Androgens and estrogens differentially and reciprocally regulate RORA, a novel candidate gene for autism. *PLoS ONE, 6*, e17116.

Sato, W., Yoshikawa, S., Kochiyama, T., & Matsumura, M. (2004). The amygdala processes the emotional significance of facial expressions: An fMRI investigation using the interaction between expression and face direction. *Neuroimage, 22*, 1006-1013.

Scheres, A., & Sanfey, A. G. (2006). Individual differences in decision making: Drive and reward responsiveness affect strategic bargaining in economic games. *Behavioral and Brain Functions, 2*, 35.

Schiller, D., Freeman, J. B., Mitchell, J. P., Uleman, J. S., & Phelps, E. A. (2009). A neural mechanism of first impressions. *Nature Neuroscience, 12*, 508-514.

Schulkin, J. (2003a). Allostasis: A neural behavioral perspective. *Hormones and Behavior, 43*, 21-27; discussion 28-30.

Schulkin, J. (2003b). *Rethinking Homeostasis: Allostatic Regulation in Physiology and Pathophysiology*. Cambridge, MA: MIT Press.

Schulkin, J., Morgan, M., & Rosen, J. (2005). A neuroendocrine mechanism for sustaining fear. *Trends in Neurosciences*.

Schutter, D. J., & van Honk, J. (2004). Decoupling of midfrontal delta-beta oscillations after testosterone administration. *International Journal of Psychophysiology, 53*, 71-73.

Schutter, D. J., & van Honk, J. (2005). Salivary cortisol levels and the coupling of midfrontal delta-beta oscillations. *International Journal of Psychophysiology, 55*, 127-129.

Scordalakes, E. M., & Rissman, E. F. (2004). Aggression and arginine vasopressin immunoreactivity regulation by androgen receptor and estrogen receptor alpha. *Genes, Brain and Behavior, 3*, 20-26.

Senju, A., & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences, 13*, 127-134.

Setchell, J. M., & Wickings, E. J. (2005). Dominance, status signals and coloration in male mandrills (Mandrillus Sphinx). *Ethology, 111*, 25-50.

Shacham, S. (1983). A shortened version of the Profile of Mood States. *Journal of Personality Assessment, 47*, 305-306.

Shacham, S., Reinhardt, L. C., Raubertas, R. F., & Cleeland, C. S. (1983). Emotional states and pain: Intraindividual and interindividual measures of association. *Journal of Behavioral Medicine, 6*, 405-419.

Siebert, M., Markowitsch, H. J., & Bartel, P. (2003). Amygdala, affect and cognition: Evidence from 10 patients with Urbach-Wiethe disease. *Brain, 126*, 2627-2637.

Siever, L. J. (2008). Neurobiology of aggression and violence. *American Journal of Psychiatry, 165*, 429-442.

Smith, T. W., Traupman, E. K., Uchino, B. N., & Berg, C. A. (2010). Interpersonal circumplex descriptions of psychosocial risk factors for physical illness: Application to hostility, neuroticism, and marital adjustment. *Journal of Personality, 78*, 1011-1036.

Smits, D. J. M., & Kuppens, P. (2005). The relations between anger, coping with anger, and aggression, and the BIS/BAS system. *Personality and Individual Differences, 39*, 783-793.

Solano-Castiella, E., Anwander, A., Lohmann, G., Weiss, M., Docherty, C., Geyer, S., et al. (2010). Diffusion tensor imaging segments the human amygdala in vivo. *Neuroimage, 49*, 2958-2965.

Spielberger, C. D., Gorusch, R. L., & Lushene, R. E. (1970). *State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.

Spielberger, C. D., Jacobs, G. A., Russel, S. F., & Crane, R. S. (1983). Assessment of anger: The state-trait anger scale. In J. N. Butcher & C. D. Spielberger (Eds.), *Advances in Personality Assessment* (Vol. 2, pp. 112-134). Hillsdale, New Jersey: Erlbaum.

Staugaard, S. R. (2010). Threatening faces and social anxiety: A literature review. *Clinical Psychology Review, 30*, 669-690.

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition, 119*, 307-311.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Szot, P., & Dorsa, D. M. (1994). Expression of cytoplasmic and nuclear vasopressin RNA following castration and testosterone replacement: Evidence for transcriptional regulation. *Molecular and Cellular Neurosciences, 5*, 1-10.

Terburg, D., Aarts, H., Putman, P., & van Honk, J. (2012a). In the eye of the beholder: Reduced threat-bias and increased gaze-imitation towards reward in relation to trait anger. *PLoS ONE, 7*, e31373.

Terburg, D., Aarts, H., & van Honk, J. (2012b). Memory and attention for social threat: Anxious hypercoding-avoidance, and submissive gaze-aversion. *Emotion*.

Terburg, D., Aarts, H., & van Honk, J. (2012c). Testosterone affects gaze-aversion from angry faces outside of conscious awareness. *Psychological Science*, in press.

Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces. *Psychological Science, 22*, 314-319.

Terburg, D., Morgan, B., & van Honk, J. (2009a). The testosterone-cortisol ratio: A hormonal marker for proneness to social aggression. *International Journal of Law and Psychiatry, 32*, 216-223.

Terburg, D., Peper, J. S., Morgan, B., & van Honk, J. (2009b). Sex differences in human aggression: The interaction between early developmental and later activational testosterone. *Behavioral and Brain Sciences, 32*, 290; discussion 292-311.

Theodoridou, A., Rowe, A. C., Penton-Voak, I. S., & Rogers, P. J. (2009). Oxytocin and social perception: Oxytocin increases perceived facial trustworthiness and attractiveness. *Hormones and Behavior, 56*, 128-132.

Thornton, H. B., Nel, D., Thornton, D., van Honk, J., Baker, G. A., & Stein, D. J. (2008). The neuropsychiatry and neuropsychology of lipoid proteinosis. *Journal of Neuropsychiatry and Clinical Neurosciences, 20*, 86-92.

Tipples, J. (2006). Fear and fearfulness potentiate automatic orienting to eye gaze. *Cognition & Emotion, 20*, 309-320.

Toates, F. (2006). A model of the hierarchy of behaviour, cognition, and consciousness. *Consciousness and Cognition, 15*, 75-118.

Tobii Technology. (2006). *User Manual*. Danderyd, Sweden: Tobii Technology AB.

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology, 25*, 395-410.

Tops, M., Wijers, A., Vanstaveren, A., Bruin, K., Denboer, J., Meijman, T., et al. (2005). Acute cortisol administration modulates EEG alpha asymmetry in volunteers: Relevance to depression. *Biological Psychology, 69*, 181-193.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242-249.

Trower, P., & Gilbert, P. (1989). New theoretical conceptions of social anxiety and social phobia. *Clinical Psychology Review, 9*, 19-35.

Tuiten, A., Van Honk, J., Koppeschaar, H., Bernaards, C., Thijssen, J., & Verbaten, R. (2000). Time course of effects of testosterone administration on sexual arousal in women. *Archives of General Psychiatry, 57*, 149-153; discussion 155-146.

Tye, K. M., Prakash, R., Kim, S. Y., Fenno, L. E., Grosenick, L., Zarabi, H., et al. (2011). Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature, 471*, 358-362.

Uzefovsky, F., Shalev, I., Israel, S., Knafo, A., & Ebstein, R. P. (2012). Vasopressin selectively impairs emotion recognition in men. *Psychoneuroendocrinology, 37*, 576-580.

van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108*, 796-803.

van der Made, F., Bloemers, J., van Ham, D., El Yassem, W., Kleiverda, G., Everaerd, W., et al. (2009a). Childhood sexual abuse, selective attention for sexual cues and the effects of testosterone with or without vardenafil on physiological sexual arousal in women with sexual dysfunction: A pilot study. *Journal of Sexual Medicine, 6*, 429-439.

van der Made, F., Bloemers, J., Yassem, W. E., Kleiverda, G., Everaerd, W., van Ham, D., et al. (2009b). The influence of testosterone combined with a PDE5-inhibitor on cognitive, affective, and physiological sexual functioning in women suffering from sexual dysfunction. *Journal of Sexual Medicine, 6*, 777-790.

van Honk, J. (2009). Neuroendocrine manipulation of the sexually dimorphic human social brain. In E. Harmon-Jones & J. S. Beer (Eds.), *Methods in Social Neuroscience* (pp. 45-69). New York: Guilford Press.

van Honk, J., Kessels, R. P., Putman, P., Jager, G., Koppeschaar, H. P., & Postma, A. (2003a). Attentionally modulated effects of cortisol and mood on memory for emotional faces in healthy young males. *Psychoneuroendocrinology, 28*, 941-948.

van Honk, J., Montoya, E. R., Bos, P. A., van Vugt, M., & Terburg, D. (in press). New evidence on testosterone and cooperation. *Nature*.

van Honk, J., Peper, J. S., & Schutter, D. J. (2005). Testosterone reduces unconscious fear but not consciously experienced anxiety: Implications for the disorders of fear and anxiety. *Biological Psychiatry, 58*, 218-225.

van Honk, J., & Schutter, D. J. (2006). Unmasking feigned sanity: A neurobiological model of emotion processing in primary psychopathy. *Cognitive Neuropsychiatry, 11*, 285-306.

van Honk, J., & Schutter, D. J. (2007a). Testosterone reduces conscious detection of signals serving social correction: Implications for antisocial behavior. *Psychological Science, 18*, 663-667.

van Honk, J., & Schutter, D. J. (2007b). Vigilant and avoidant responses to angry facial perceptions: Dominance and submission motives. In E. Harmon-Jones & P. Winkielman (Eds.), *Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior* (pp. 197-223). New York: Guilford Press.

van Honk, J., Schutter, D. J., Bos, P. A., Kruijt, A. W., Lentjes, E. G., & Baron-Cohen, S. (2011a). Testosterone administration impairs cognitive empathy in women depending on second-to-fourth digit ratio. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 3448-3452.

van Honk, J., Schutter, D. J., d'Alfonso, A. A., Kessels, R. P., & de Haan, E. H. (2002). 1 hz rTMS over the right prefrontal cortex reduces vigilant attention to unmasked but not to masked fearful faces. *Biological Psychiatry, 52*, 312-317.

van Honk, J., Schutter, D. J., Hermans, E. J., & Putman, P. (2003b). Low cortisol levels and the balance between punishment sensitivity and reward dependency. *Neuroreport, 14*, 1993-1996.

van Honk, J., Schutter, D. J., Hermans, E. J., Putman, P., Tuiten, A., & Koppeschaar, H. (2004a). Testosterone shifts the balance between sensitivity for punishment and reward in healthy young women. *Psychoneuroendocrinology, 29*, 937-943.

van Honk, J., Schutter, D. J. L. G., Hermans, E. J., & Putman, P. (2004b). Testosterone, cortisol, dominance, and submission: Biologically prepared motivation, no psychological mechanisms involved. *Behavioral and Brain Sciences, 27*, 160-162.

van Honk, J., Terburg, D., & Bos, P. A. (2011b). Further notes on testosterone as a social hormone. *Trends in Cognitive Sciences, 15*, 291-292.

van Honk, J., Tuiten, A., de Haan, E., van den Hout, M., & Stam, H. (2001a). Attentional biases for angry faces: Relationships to trait anger and anxiety. *Cognition & Emotion, 15*, 279 - 297.

van Honk, J., Tuiten, A., Hermans, E., Putman, P., Koppeschaar, H., Thijssen, J., et al. (2001b). A single administration of testosterone induces cardiac accelerative responses to angry faces in healthy young women. *Behavioral Neuroscience, 115*, 238-242.

van Honk, J., Tuiten, A., van den Hout, M., Koppeschaar, H., Thijssen, J., de Haan, E., et al. (1998). Baseline salivary cortisol levels and preconscious selective attention for threat. A pilot study. *Psychoneuroendocrinology, 23*, 741-747.

van Honk, J., Tuiten, A., van den Hout, M., Koppeschaar, H., Thijssen, J., de Haan, E., et al. (2000). Conscious and preconscious selective attention to social threat: Different neuroendocrine response patterns. *Psychoneuroendocrinology, 25*, 577-591.

van Honk, J., Tuiten, A., Verbaten, R., van den Hout, M., Koppeschaar, H., Thijssen, J., et al. (1999). Correlations among salivary testosterone, mood, and selective attention to threat in humans. *Hormones and Behavior, 36*, 17-24.

Van Selst, M., & Merikle, P. M. (1993). Perception below the objective threshold? *Consciousness and Cognition, 2*, 194-203.

Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological Science, 18*, 19-23.

van Wingen, G., Mattern, C., Verkes, R. J., Buitelaar, J., & Fernandez, G. (2010). Testosterone reduces amygdala-orbitofrontal cortex coupling. *Psychoneuroendocrinology, 35*, 105-113.

van Wingen, G. A., Zylicz, S. A., Pieters, S., Mattern, C., Verkes, R. J., Buitelaar, J. K., et al. (2009). Testosterone increases amygdala reactivity in middle-aged women to a young adulthood level. *Neuropsychopharmacology, 34*, 539-547.

Vanyukov, M. M., Moss, H. B., Plail, J. A., Blackson, T., Mezzich, A. C., & Tarter, R. E. (1993). Antisocial symptoms in preadolescent boys and in their parents: Associations with cortisol. *Psychiatry Research, 46*, 9-17.

Veinante, P., & Freund-Mercier, M. J. (1997). Distribution of oxytocin- and vasopressin-binding sites in the rat extended amygdala: A histoautoradiographic study. *Journal of Comparative Neurology, 383*, 305-325.

Viau, V. (2002). Functional cross-talk between the hypothalamic-pituitary-gonadal and -adrenal axes. *Journal of Neuroendocrinology, 14*, 506-513.

Viau, V., & Meaney, M. J. (1996). The inhibitory effect of testosterone on hypothalamic-pituitary-adrenal responses to stress is mediated by the medial preoptic area. *Journal of Neuroscience, 16*, 1866-1876.

Virkkunen, M. (1985). Urinary free cortisol secretion in habitually violent offenders. *Acta Psychiatrica Scandinavica, 72*, 40-44.

Vuilleumier, P. (2002). Facial expression and selective attention. *Current Opinion in Psychiatry, 15*, 291-300.

Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences, 9*, 585-594.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*, 448-457.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.

Weeks, J. W., Heimberg, R. G., & Heuer, R. (2011). Exploring the role of behavioral submissiveness in social anxiety. *Journal of Social and Clinical Psychology, 30*, 217-249.

Wells, A., & Matthews, G. (1994). *Attention and Emotion: A Clinical Perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Wersinger, S. R., Caldwell, H. K., Christiansen, M., & Young, W. S., 3rd. (2007). Disruption of the vasopressin 1b receptor gene impairs the attack component of aggressive behavior in mice. *Genes, Brain and Behavior, 6*, 653-660.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science, 306*, 2061.

Whalen, P. J., & Phelps, E. A. (Eds.). (2009). *The Human Amygdala*. New York: The Guilford Press.

Wieser, M., Pauli, P., Alpers, G., & Mühlberger, A. (2009). Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking and psychophysiology study. *Journal of Anxiety Disorders, 23*, 93-103.

Wiest, G., Lehner-Baumgartner, E., & Baumgartner, C. (2006). Panic attacks in an individual with bilateral selective lesions of the amygdala. *Archives of Neurology, 63*, 1798-1801.

Wilkowski, B. M., & Robinson, M. D. (2007). The cognitive basis of trait anger and reactive aggression: An integrative analysis. *Personality and Social Psychology Review, 12*, 3-21.

Williams, L. M., Barton, M. J., Kemp, A. H., Liddell, B. J., Peduto, A., Gordon, E., et al. (2005). Distinct amygdala-autonomic arousal profiles in response to fear signals in healthy males and females. *Neuroimage, 28*, 618-626.

Williams, L. M., Liddell, B. J., Kemp, A. H., Bryant, R. A., Meares, R. A., Peduto, A. S., et al. (2006). Amygdala-prefrontal dissociation of subliminal and supraliminal fear. *Human Brain Mapping, 27*, 652-661.

Williams, L. M., Phillips, M. L., Brammer, M. J., Skerrett, D., Lagopoulos, J., Rennie, C., et al. (2001). Arousal dissociates amygdala and hippocampal fear responses: Evidence from simultaneous fMRI and skin conductance recording. *Neuroimage, 14*, 1070-1079.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin, 31*, 121-135.

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*, 277-283.

Wirth, M. M., & Schultheiss, O. C. (2007). Basal testosterone moderates responses to anger faces in humans. *Physiology & Behavior, 90*, 496-505.

Wood, R. I. (1996). Functions of the steroid-responsive neural network in the control of male hamster sexual behavior. *Trends in Endocrinology and Metabolism, 7*, 338-344.

Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., et al. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS ONE, 4*, e8330.

Zheng, Z., & Cohn, M. J. (2011). Developmental basis of sexually dimorphic digit ratios. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 16289-16294.

List of publications

Published articles

Bos, P. A., Terburg, D., & van Honk, J. (2010). Testosterone decreases trust in socially naïve humans. *Proceedings of the National Academy of Sciences of the United States of America, 107*, 9991-9995.

Ipser, J., Terburg D., Syal, S., Phillips, N., Solms, M., Panksepp, J., Malcolm-Smith, S., Thomas, K., Stein, D. J., & van Honk, J. (*in press*). Reduced fear recognition sensitivity following acute buprenorphine administration in healthy volunteers. *Psychoneuroendocrinology*.

Montoya, E. R., Terburg, D., Bos, P. A., & van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and Emotion, 36*, 65-73.

Morgan, B., Terburg, D., Thornton, H. B., Stein, D. J., & van Honk, J. (*in press*). Paradoxical facilitation of working memory after basolateral amygdala damage. *PLoS ONE*.

Terburg, D., Aarts, H., Putman, P., & van Honk, J. (2012). In the eye of the beholder: Reduced threat-bias and increased gaze-imitation towards reward in relation to trait anger. *PLoS ONE, 7*, e31373.

Terburg, D., Aarts, H., & van Honk, J. (*in press*). Memory and attention for social threat: Anxious hypercoding-avoidance, and submissive gaze-aversion. *Emotion*.

Terburg, D., Aarts, H., & van Honk, J. (*in press*). Testosterone affects gaze-aversion from angry faces outside of conscious awareness. *Psychological Science*.

Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces. *Psychological Science, 22*, 314-319.

Terburg, D., Morgan, B. E., Montoya, E. R., Hooge, I. T., Thornton, H. B., Hariri, A. R., Panksepp, J., Stein, D. J., & van Honk, J. (*in press*). Hyper-vigilance for fear after basolateral amygdala damage in humans. *Translational Psychiatry*.

Terburg, D., Morgan, B., & van Honk, J. (2009). The testosterone-cortisol ratio: A hormonal marker for proneness to social aggression. *International Journal of Law and Psychiatry, 32*, 216-223.

Terburg, D., Peper, J. S., Morgan, B., & van Honk, J. (2009). Sex differences in human aggression: The interaction between early developmental and later activational testosterone. *Behavioral and Brain Sciences, 32*, 290; discussion 292-311.

van Honk, J., Montoya, E. R., Bos, P. A., van Vugt, M., & Terburg, D. (*in press*). New evidence on testosterone and cooperation. *Nature*.

van Honk, J., Terburg, D., & Bos, P. A. (2011). Further notes on testosterone as a social hormone. *Trends in Cognitive Sciences, 15*, 291-292.

Under consideration

Terburg, D., & van Honk, J. (*in review*). Approach-avoidance versus dominance-submissiveness: A multilevel neural framework on how testosterone promotes social status. *Emotion Review*.

Acknowledgements

De eerste die ik natuurlijk moet bedanken is Jack. Zonder jou was mijn enthousiasme voor alles wat met wetenschap en psychologie te maken heeft allang verdampt, en was deze promotie nooit gelukt. De tweede die absoluut bedankt moet worden is Henk. Dankzij jou steun werd mijn promotie uiteindelijk ook echt mogelijk en jou aanstekelijke motivatie om social neuroscience te bedrijven heeft me over menig dood punt heen geholpen.

Estrella en Peter B., wij zaten, en zitten, in ongeveer hetzelfde schuitje en dankzij jullie was dat altijd een inspirerend en bovenal gezellig schuitje. Daar gaan we dus voorlopig gewoon mee door! Dennis H., we hebben meestal geen idee wat je allemaal nou eigenlijk zegt, maar je staat altijd klaar om iedereen te helpen, hebt overal een oplossing voor en weet alles, en dan bedoel ik ook echt alles. Jiska, mijn kamergenoot van het eerste uur en baken van rust. Het was voor jou natuurlijk een goeie zet om ons te verlaten, maar ik vond het vooral erg jammer dat je wegging. Ignace, dankzij jou kan ik overweg met een eye-tracker, en sinds mijn eerste stage bij jou, ik weet niet hoe lang geleden, weet ik dat wetenschap begint met een kritische blik. Floris, jij leerde mij wat dreiging is, electro-shocks en een hoop herrie blijkbaar, waarvoor dank. Ook was onze onvergetelijke trip naar Springbok onmogelijk geweest zonder jou. Sylco, jij leerde mij wat bier is. Of nee, wacht even. Hoe zat het ook alweer?

Barak, without you all our adventures in South Africa would not have been possible. Over the years we have become good friends, and I hope we can keep working together for a long time to come. Our research in South Africa would have degenerated to a bunch of scientist in the desert trying to do fMRI and eye-tracking with the occasional springbok, if it wasn't for Mara. Mara, I have told you before, and hopefully I will be able to tell you again in the near future, you're the best! Also many thanks to Fran, Mary-Louise, Supriya, Coenie, Bruce and Lucretia, for help with scanning, and Dan, who has always supported our work in the far south.

Floor, veel dank voor je aanhoudende gastvrijheid. Elke keer weer van die gekke wetenschappers over de vloer zal geen pretje zijn, maar je geeft me altijd weer het gevoel dat ik zeer welkom en thuis ben, waarvoor ik natuurlijk ook Bowi en Tygo moet bedanken!

Verder gaat mijn dank uit naar de vele mensen die in meer of mindere mate betrokken zijn geweest bij m'n onderzoek, of echte leven, gedurende de afgelopen jaren. Leon, Frans, Albert, Dennis S., Peter P., Erno, Ivo, Stijn, Joke, Erik, Bea, Ruud, Nick, Mathijs, Do, Nicole,

Sarah, Lisette, Geert-Jan, Giannis, Linda, Corinne, Ada, Lisa, Kasper, Steven, Luuk, Daniel, Arnoud, Jasper, Laila, Ans, Veronica, Renée en Eveline allemaal ontzettend bedankt.

Ten slotte wil ik nog een paar mensen bedanken die mij zeer dierbaar zijn en mij gevormd hebben tot wie ik nu ben. Joska, eindeloos bakkeleien we over alle mogelijke zin en vooral ook onzin. Ook al zijn we inmiddels grote jongens, eenieder die dat aan wil vechten draag ik overigens een warm hart toe, we blijven dat tot in lengte van dagen lekker doen! Jona, van huisgenoot naar studiegenoot naar iemand die net als ik ook zo af en toe het squashballetje weet te raken. Al die tijd ook m'n beste maatje. Iedereen heeft ook een slechte kant zeggen ze weleens. Jij niet, en daar heb ik zo ontzettend veel van geleerd en genoten! Mijn lieve ouders die mij, beide op geheel eigen wijze, altijd gestimuleerd hebben nou eens iets te gaan doen met m'n leven. Jullie vullen elkaar daarbij ontzettend goed aan en zijn een topteam. Zonder jullie liefde en toewijding was het allemaal niks geworden met mij. Saar, mijn lieve grote zus, van badkamer-gesprek tot het maken van die schat van een kleine Xanthe, jij was en bent er altijd voor mij! Gied, broertje, rustige rots in de branding, je bent altijd jezelf en je ziet altijd dwars door me heen. Jij houdt mij met beide benen op de grond. En als laatste mijn lieve meisje Dorine. Toen ik jou leerde kennen werd ik gelukkig. Nu kan ik niet meer zonder jou. Je houdt mij gelukkig, elke dag weer en voor altijd.

Curriculum Vitae

David Terburg werd op 17 juli 1980 geboren te Naarden. In 1998 behaalde hij zijn VWO diploma aan het Goois Lyceum te Bussum. Hij begon in september 2000 met een studie psychologie aan de Universiteit Utrecht en na zijn afstuderen in 2007 ging hij als onderzoeksassistent werken voor de vakgroep psychologische functieleer aan de Universiteit Utrecht onder leiding van Prof. Dr. Jack van Honk. Daarnaast werkte hij vanaf 2008 als onderzoeksassistent bij de Registration and Imaging of Brain Systems groep in het Universitair Medisch Centrum Utrecht onder leiding van Prof. Dr. Nick Ramsey. In 2009 is hij begonnen aan zijn onderzoek naar de neuro-endocrinologie van sociaal gedrag onder begeleiding van Prof. Dr. Jack van Honk en Prof. Dr. Henk Aarts. Momenteel is hij als post-doc onderzoeker verbonden aan de Universiteit Utrecht en de University of Cape Town (Zuid-Afrika).

David Terburg was born on July 17th 1980 in Naarden. In 1998 he obtained his VWO diploma at the Goois Lyceum in Bussum. He started in september 2000 with the study of psychology at Utrecht University, and after graduating in 2007 he worked as a research assistant for the department of experimental psychology at Utrecht University under the supervision of Prof. Dr. Jack van Honk. He also worked from 2008 as a research assistant at the Registration and Imaging of Brain Systems group at the University Medical Center Utrecht, led by Prof. Dr. Nick Ramsey. In 2009 he started his research into the neuroendocrinology of social behavior under the guidance of Prof. Dr. Jack van Honk and Prof. Dr. Henk Aarts. He is currently a post-doctoral researcher at Utrecht University and the University of Cape Town (South Africa).