MAINTAINING FACTORS OF FEAR-RELEVANT ILLUSORY CORRELATIONS

Inauguraldissertation

zur Erlangung der Doktorwürde der Fakultät für Humanwissenschaften

der Julius-Maximilians-Universität Würzburg

vorgelegt von Julian Wiemer

aus Schwebenried

Würzburg 2015

Erstgutachter:	Prof. Dr. Paul Pauli
Zweitgutachter:	Prof. Dr. Andreas Mühlberger
Datum des Rigorosums:	29. Januar 2015

Acknowledgments

I would like to thank Prof. Paul Pauli and Prof. Andreas Mühlberger for supervision and advise throughout this dissertation project, my parents, Eva, Marta, Evelyn, Christian, Philipp, Matthias, Max, Youssef, Heike, Stefan, Johanna, Petra, Anna, Elena, Markus, Lea, Hannah, Miriam, Doro, Bastian, Felix and probably many more nice people for sharing beverages and ideas, for their support, company and friendship, and Johanna, Enne, Marika, Jan, Chalid and all participants for their help in gathering the data for this thesis.

Table of contents

Abstract

Biased cognitive processes are very likely involved in the maintenance of fears and anxiety. One of such cognitive processes is the perceived relationship between fear-relevant stimuli and aversive consequences. If this relationship is perceived although objective contingencies have been random, it is called an (a posteriori) illusory correlation. If this relationship is overestimated before objective contingencies are experienced, it is called an (a priori) expectancy bias. Previous investigations showed that fear-relevant illusory correlations exist, but very few is known about how and why this cognitive bias develops. In the present dissertation thesis, a model is proposed based on a review of the literature on fear-relevant illusory correlations. This model describes how psychological factors might have an influence on fear and illusory correlations. Several critical implications of the model were tested in four experiments.

Experiment 1 tested the hypothesis that people do not only overestimate the proportion of aversive consequences (startle sounds) following emotionally negative stimuli (pictures of mutilations) relative to neutral stimuli (pictures of household objects), but also following highly arousing positive stimuli (pictures of erotic scenes), because arousal might be an important determinant of illusory correlations. The result was a significant expectancy bias for negative stimuli and a much smaller expectancy bias for positive stimuli. Unexpectedly, expectancy bias was restricted to women. An a posteriori illusory correlation was not found overall, but only in those participants who perceived the aversive consequences following negative stimuli as particularly aversive.

Experiment 2 tested the same hypothesis as experiment 1 using a paradigm that evoked distinct basic emotions (pictures inducing fear, anger, disgust or happiness). Only negative emotions resulted in illusory correlations with aversive outcomes (startle sounds), especially the emotions of fear and disgust. As in experiment 1, the extent of these illusory correlations was correlated with the perceived aversiveness of aversive outcomes. Moreover, only women overestimated the proportion of aversive outcomes during pictures that evoked fear, anger or disgust.

Experiment 3 used functional Magnetic Resonance Imaging (fMRI) to measure biased brain activity in female spider phobics during an illusory correlation paradigm. Both spider phobics and healthy controls expected more aversive outcomes (painful electrical shocks) following pictures of spiders than following neutral control stimuli (pictures of mushrooms). Spider phobics but not healthy controls overestimated the proportion of aversive outcomes following pictures of spiders in a trial-by-trial memory task. This a posteriori illusory correlation was correlated with enhanced shock aversiveness and activity in primary sensory-motor cortex in phobic participants. Moreover, spider phobics' brain activity in the left dorsolateral prefrontal cortex was elevated in response to spider images. This activity also predicted the extent of the illusory correlation, which supports the theory that executive and attentional resources play an important role in the maintenance of illusory correlations.

Experiment 4 tested the hypothesis that the enhanced aversiveness of some outcomes would be sufficient to causally induce an illusory correlation. Neutral images (colored geometric figures) were paired with differently aversive outcomes (three startle sounds varying in intensity). Participants developed an illusory

correlation between those images, which predicted the most aversive sound and this sound, which means that this association was overestimated relative to the other associations. The extent of the illusory correlation was positively correlated with participants' self-reported anxiety. The results imply that the previously found relationship between illusory correlations and outcome aversiveness might reflect a causal impact of outcome aversiveness or salience on illusory correlations.

In sum, the conducted experiments indicate that illusory correlations between fear-relevant stimuli and aversive consequences might persist – among other factors - because of an enhanced aversiveness or salience of aversive consequences following feared stimuli. This assumption is based on correlational findings, a neural measure of outcome perception and a causal influence of outcome aversiveness on illusory correlations. Implications of these findings were integrated into a model of fear-relevant illusory correlations and potential implications are discussed. Future investigations should further elucidate the role of executive functions and gender effects. Moreover, the trial-by-trial assessment of illusory correlations is recommended to increase reliability of the concept. From a clinical perspective, the down-regulation of aversive experiences and the allocation of attention to non-aversive experiences might help to cure anxiety and cognitive bias.

Zusammenfassung

Verzerrte kognitive Prozesse sind sehr wahrscheinlich an der Aufrechterhaltung von Furcht und Angst beteiligt. Ein solcher kognitiver Prozess ist der wahrgenommene Zusammenhang zwischen Reizen, vor denen bereits Angst besteht und unangenehmen Konsequenzen. Wenn so ein Zusammenhang wahrgenommen wird, obwohl die objektiven Kontingenzen zufällig sind, spricht man von einer illusorischen Korrelation (a posteriori). Wenn so ein Zusammenhang überschätzt wird, bevor objektive Kontingenzen erlebt werden, spricht man von einer Erwartungsverzerrung (a priori). Frühere Untersuchungen zeigten, dass angstrelevante illusorische Korrelationen existieren, aber bisher ist nur wenig darüber bekannt, wie und warum diese entstehen. In der vorliegenden Dissertation wird ein Modell vorgeschlagen, dass auf bisherigen Erkenntnissen zu angstrelevanten illusorischen Korrelationen beruht. Das Modell beschreibt, welche psychologischen Faktoren die Entstehung von Angst und illusorischen Korrelationen begünstigen könnten. Mehrere Implikationen dieses Modells wurden in vier Experimenten getestet.

Experiment 1 überprüfte die Hypothese, dass Menschen nicht nur die Häufigkeit unangenehmer Konsequenzen (Schreckgeräusche) nach emotional negativen Reizen (Bilder von Verletzungen) überschätzen, verglichen mit neutralen Reizen (Bilder von Haushaltsgegenständen), sondern auch nach sehr aufregenden positiven Reizen (Bilder von erotischen Szenen), weil die allgemeine Erregung einen Einfluss auf illusorische Korrelationen haben sollte. Das Ergebnis war eine signifikante Erwartungsverzerrung bei negativen Reizen und eine sehr viel kleinere Erwartungsverzerrung bei positiven Reizen. Unerwarteter Weise waren Erwartungsverzerrungen auf Frauen beschränkt. Eine illusorische Korrelation (a posteriori) konnte im Allgemeinen nicht festgestellt werden, sondern lediglich bei solchen Probanden, die die unangenehmen Konsequenzen nach negativen Reizen als besonders unangenehm empfanden.

Experiment 2 überprüfte die gleiche Hypothese wie Experiment 1 anhand einer Versuchsanordnung, die verschiedene Basisemotionen hervorrufen sollte (durch Bilder, die Angst, Ärger, Ekel oder Freude induzierten). Nur negative Emotionen führten zu illusorischen Korrelationen (a posteriori) mit unangenehmen Ereignissen (Schreckgeräusche), insbesondere die Emotionen Angst und Ekel. Wie auch in Experiment 1 korrelierte das Ausmaß der illusorischen Korrelation mit der wahrgenommenen Unangenehmheit der unangenehmen Ereignisse bei der entsprechenden Bildkategorie. Darüber hinaus überschätzten nur Frauen den Anteil negativer Ereignisse bei Bildern, die Angst, Ekel, oder Ärger hervorriefen.

Experiment 3 verwendete funktionelle Magnetresonanztomografie (fMRT), um verzerrte Gehirnaktivität bei Spinnenphobikerinnen während eines Versuchs zu illusorischen Korrelationen zu messen. Sowohl Spinnenphobikerinnen als auch gesunde Kontrollprobandinnen erwarteten mehr unangenehme Konsequenzen (schmerzhafte elektrische Reize) bei Bildern von Spinnen als bei neutralen Kontrollreizen (Bilder von Pilzen). Spinnenphobikerinnen, aber nicht gesunde Kontrollprobandinnen überschätzten jedoch im Nachhinein den Anteil unangenehmer Konsequenzen bei Bildern von Spinnen in einer Trial-by-Trial Gedächtnisaufgabe. Diese illusorische Korrelation (a posteriori) korrelierte mit der erhöhten Unangenehmheit der elektrischen Reize und mit Aktivierung im primären senso-motorischen Kortex der phobischen Teilnehmerinnen. Darüber hinaus wiesen Spinnenphobikerinnen in Reaktion auf die Bilder von Spinnen eine

erhöhte Aktivierung im linken dorsolateralen präfrontalen Kortex auf. Diese Aktivität sagte auch das Ausmaß der illusorischen Korrelation vorher, was die These unterstützt, dass exekutive und Aufmerksamkeitsressourcen eine wichtige Rolle in der Aufrechterhaltung illusorischer Korrelationen spielen.

Experiment 4 überprüfte die Hypothese, dass die erhöhte Unangenehmheit mancher Konsequenzen hinreichend sein würde, um kausal eine illusorische Korrelation hervorzurufen. Neutrale Bilder (geometrische Figuren in drei verschiedenen Farben) wurden mit unterschiedlich unangenehmen Konsequenzen gepaart (Schreckgeräusche in drei verschiedenen Intensitäten). Bei den Probanden entwickelte sich eine illusorische Korrelation mit der Farbe, die das unangenehmste Geräusch voraussagte und diesem Geräusch, das heißt, der Zusammenhang wurde im Vergleich zu den anderen Zusammenhängen überschätzt. Das Ausmaß der illusorischen Korrelation korrelierte positiv mit der Ängstlichkeit der Teilnehmer. Die Ergebnisse legen nahe, dass der bisher gefundene Zusammenhang zwischen illusorischen Korrelationen und der Unangenehmheit der unangenehmen Konsequenzen auf einen kausalen Einfluss der Unangenehmheit oder Salienz der Konsequenzen auf illusorische Korrelationen zurückgehen könnte.

Zusammengefasst zeigten die durchgeführten Experimente, dass illusorische Korrelationen zwischen angstrelevanten Reizen und unangenehmen Konsequenzen – neben anderen Einflussfaktoren – aufgrund einer erhöhten Unangenehmheit oder Salienz unangenehmer Konsequenzen bei gefürchteten Reizen bestehen könnten. Diese Annahme basiert auf korrelativen Ergebnissen, einem neuralen Maß der Konsequenzverarbeitung und dem gefundenen kausalen Einfluss der Unangenehmheit unangenehmer Konsequenzen auf illusorische Korrelationen. Implikationen dieser Befunde werden in ein Modell zu angstrelevanten illusorischen Korrelationen integriert und diskutiert. Zukünftige Studien sollten die Rolle exekutiver Funktionen und Geschlechtsunterschiede genauer untersuchen. Es empfiehlt sich dabei, illusorische Korrelationen Trial-by-Trial zu erfassen, um die Reliabilität des Konzepts zu erhöhen. Aus klinischer Sicht könnten die Beruhigung unangenehmer Erfahrungen und die Aufmerksamkeitsallokation auf nicht-unangenehme Erfahrungen helfen, Ängste und kognitive Verzerrungen zu vermindern.

1. Introduction

"Man is always prey to his truths"

Albert Camus

This quote from the essay “The myth of Sisyphus” by Albert Camus implies that truth is not only something objective, but rather something subjective that can vary between individuals. Moreover, we are “prey” to our truths - our feelings, our thoughts and our behavior are determined by what we think is true. In other words, our perception of reality may have more influence on us than reality itself. While Camus meant that a human being who has recognized that the world is absurd, will always be bound to this truth and can never get rid of this truth again, some subjective truths can certainly change. For instance, inaccurate truths can guide our behavior and feelings, but they can be changed to reach a more accurate or just different point of view that enables one to feel, think and act in a new way. Sometimes individuals become very fearful in the presence of humans or objects or they literally become “prey” to animals like dogs or spiders because in their “truths” these animals are more dangerous than they actually are. Similarly, in the Buddhist tradition, “right view” is considered the first step on the eightfold path to awakening, and “wrong perceptions” are considered the main cause of suffering (Epstein, 2004). Basically, the idea that thoughts affect emotions can also be found in modern cognitive behavior therapy (Beck & Dozois, 2011). On the other hand, our thoughts and perceptions are also guided by our emotions and motivations. For example, reward and punishment can determine what we perceive in ambiguous visual stimuli (Rock & Fleck, 1950). Emotions and cognitive representations of the environment are closely interconnected. The present dissertation project is an attempt to help to explain the relationship between fear and biased thoughts, in this case illusory correlations between fear-relevant stimuli and aversive consequences. A better understanding of the maintenance of fear and illusory correlations may help to cure negative emotions and psychological disorders.

2. Theoretical background

2.1. Overview

The theoretical background of the present thesis will discuss very shortly the affective states of fear and anxiety and how they might develop into anxiety disorders. Particularly, a very short overview will be given about fear conditioning, its strengths and shortcomings in explaining anxiety disorders, and why a simplistic view of conditioned fear is insufficient in this regard. In order to fully understand the emergence and maintenance of anxiety disorders, higher order cognitive processes like attention, working memory, evaluations and expectations should be taken into account. These cognitive processes can be biased in a way that promotes anxiety, presumably especially in ambiguous threatening situations. One of these cognitive biases is an illusory correlation between feared objects and aversive consequences. A review of the literature will demonstrate that fear-relevant illusory correlations exist in high fearful individuals comprising several kinds of fears, are clinically relevant, but are barely understood in terms of underlying psychological mechanisms. A model for the maintenance of fear-relevant illusory correlations will be proposed in order to empirically test its implications.

2.2. Fear learning and maintenance

Fear and anxiety

According to the German national health survey conducted by the Robert-Koch-Institute between 2008 and 2011, the one-year-prevalence of anxiety disorders is estimated 15%, in women even 21% (Jacobi et al., 2014). Therefore, anxiety disorders are among the most prevalent psychological diseases, thus producing a significant burden on the national health care system and leading to serious personal suffering. Despite impaired well-being, anxiety disorders are considered to be rooted in the emotions of fear and anxiety, which are part of normal psychological functioning and have been essential for survival and are so still today.

Many authors distinguish between fear and anxiety as two related but distinct emotional states that help dealing with threats in the environment (Epstein, 1972; Barlow, Chorpita, & Turovsky, 1996; Öhman, 2008). Fear is a high arousing state that is proposed to emerge in a situation in which the danger is concrete and the organism is ready for a fight-or-flight-response. In contrast, anxiety refers to the agitated condition when a potentially threatening stimulus might occur, but the exact time and nature of the stimulus occurrence are unknown. As a consequence, fear and anxiety are accompanied by different cognitive and physiological states to deal with the current type of situation. If a stimulus is unpredictable or unknown, it is important to get more information about it, which is why the sensory system should be more sensitive in anxiety. This is expressed for instance in a decreased pain threshold in anxiety (Rhudy & Meagher, 2000), and subjectively larger visual fields and faster eye movements due to facial fear expression (Susskind et al., 2008). In addition, attention becomes more stimulus-driven in anxiety, which might help to detect the potential threat (Eysenck, Derakshan, Santos, &

Calvo, 2007). If the organism actively copes with a specific threatening stimulus, it is helpful that the pain threshold is increased in fear (Rhudy & Meagher, 2000). Attention should then be focused on the only goal to escape or to dominate in a fight. In other words, fear should lead to a more focused state of attention than anxiety, although evidence on this assumption is still lacking to the knowledge of the author. One reason for this might be that intense fear that is arousing enough to induce the reduced sensory receptivity is difficult to evoke in humans without violating ethical considerations. However, especially animal studies support the view of a stress-induced analgesia in a state of fear (e.g. Fanselow, 1986). In the brain, fear and anxiety may be mediated by overlapping, but slightly different circuits. While fear and anxiety are both based on the central nucleus of the amygdala, anxiety also seems to involve the bed nucleus of the stria terminalis (BNST) which is located near the basolateral amygdala (Davis, Walker, Miles, & Grillon, 2009).

In very simple terms, fear and anxiety are related phenomena because they help to deal with a threat in the environment - fear with a certain threat and anxiety with an uncertain threat. This has also implications for inter-individual differences in the occurrence of these affect programs. The more certain a threat is, the less differences between individuals' behavior should occur. For example, if someone comes across a bear in the forest and that bear tries to attack, most people will probably run away. Possibly, the intensity of feeling fear will vary across individuals, but probably not very much. On the other hand, if someone walks through the forest and suddenly hears the sound of a cracking branch, the variability of potential thoughts and behaviors should be much higher. Some people might think, it was a companion stepping onto a branch, some might think it was a wild animal (but what animal?), some may just ignore it. Importantly, these inter-individual differences may determine whether someone is afraid or not. In some cases, cognitive, behavioral or physiological responses in such an uncertain situation may be exaggerated in relation to the real potential danger or in relation to the normal population. Exaggerated responses on the cognitive level (e.g. attention, interpretation) have been called cognitive biases (Mathews & MacLeod, 1994; 2005). From a clinical perspective, cognitive biases could be a maintaining factor of an anxiety disorder, and thus targets for treatment (e.g., Clark & Beck, 2010; MacLeod & Mathews, 2012). Finally, exaggerated responses are diagnostic criteria of practically all anxiety disorders and may constitute an important distinguishing feature between normal and psychopathological behavior (DSM-IV, 1994).

So, biased cognitions should be more likely to occur in ambiguous situations. An example of such biased cognitions may be the inflated expectancy of a potential threat or, in relation to this, an overestimation of the relationship between a situation or a stimulus and an aversive event. The present thesis is about such an overestimation of contingencies, so-called illusory correlations (Tomarken, Mineka, & Cook, 1989). Because of the reasons that were just described, illusory correlations should have a stronger impact on anxiety in an ambiguous situation than on fear in an unambiguous situation. However, although fear occurs in unambiguous situations, this does not mean that fear does not play a role in illusory correlations. For example, if an ambiguous situation turns into an unambiguous situation, the fear response might be amplified because of increased anxiety in the anticipation phase and this increased anxiety may have relied on an illusory correlation. In fact, expectancy of threat leads to anticipatory anxiety (Phelps et al., 2001) and anxiety enhances responses to noxious stimulation (Rhudy & Meagher, 2000). In addition, fear may play a role in the development of illusory correlations due to

enhanced memory encoding of threat, similar to the arousal induced improvement of memory (Cahill & McGaugh, 1998). One process mediating this enhanced encoding may be the strengthened association between environmental stimuli and the occurring threat, as realized in fear conditioning (Pavlov, 2003; see next section).

Fear conditioning

Fear responses can be established by the contingent pairing of a neutral conditioned stimulus (CS+) with an inertly aversive unconditioned stimulus (US) that evokes an unconditioned fear response (UCR). Following such a fear conditioning procedure, the CS+ alone has the potential to evoke a conditioned fear response (CR). Usually, a conditioning experiment is separated into three phases, a habituation phase, an acquisition phase and an extinction phase. During the habituation phase, the CS+ and a safety stimulus (CS-) are presented without a US to allow habituation to both stimuli. During the acquisition phase, the US is paired with the CS+, but not with the CS-. During the extinction phase, both CS+ and CS- are presented without the US again which leads to decreasing CRs. CRs are often measured in the form of physiological responsivity, such as skin conductance responses (SCRs), heart rate or the auditory startle response (ASR).

Some studies suggest that individual differences in this process of fear conditioning are related to high trait anxiety (e.g. Glotzbach-Schoon et al., 2013) and anxiety disorders (e.g. Lissek et al., 2009). Despite this evidence for deviant fear conditioning in anxiety disorders, the effect sizes seem to be relatively small and it is unclear which process exactly is impaired (Lissek et al., 2005). In panic disorder for example, some studies found enhanced (e.g. Michael, Blechert, Vriends, Margraf, & Wilhelm, 2007) and some reduced (e.g. Lissek et al., 2009) discriminative fear responses between the CS+ and a safety signal that is not paired with an unconditioned stimulus (CS-). A meta-analysis of fear conditioning in anxiety patients and controls suggested that stronger conditioning in anxiety disorders may be due to a general enhancement of fear responses to both safety and danger signals in patients (Lissek et al., 2005). Therefore, some authors argue that anxiety goes along with a failure to inhibit fear responses to safety signals.

Several investigators did not only use neutral CSs, but also fear-relevant CSs (e.g. snakes, angry faces) to test whether CRs to fear-relevant stimuli can be acquired more easily than CRs to fear-irrelevant stimuli (Cook, Hodes, & Lang, 1986; Dimberg & Öhman, 1996; Hugdahl & Öhman, 1977; Öhman, 2009; Öhman, Eriksson, & Olofsson, 1975). This would suggest that enhanced fear conditioning to these stimuli can explain why some fears and phobias are more common than others (Seligman, 1971). In a typical experiment testing this so-called preparedness effect on fear conditioning, the difference in CRs to a fear-relevant CS+ and a fear-relevant CS- is compared with the difference in CRs to a fear-irrelevant CS+ and a fear-irrelevant CS-. If the difference is increased using a fear-relevant CS+ and CS- (e.g. a spider and a snake), the stimuli can be considered to augment fear conditioning. The fear-relevance of both the CS+ and the CS- ensures that the effects cannot be attributed to stronger responses to fear-relevant stimuli per se. Typically, these preparedness studies show that fear-relevant and fear-irrelevant CSs evoke similar conditioned SCRs during the acquisition phase, but conditioned SCRs to fear-relevant CSs are more resistant to extinction (see Öhman, 2009, for an overview). However, some

experiments also found a preparedness effect on heart rate acceleration in the acquisition phase (Cook et al., 1986).

Considering the evidence on fear conditioning, this learning process might play a role in the onset and maintenance of anxiety disorders. Nevertheless, the simple experience of paired CS-US associations seems insufficient to explain pathological fear and additional factors need to be taken into account because of several reasons. First, effect sizes seem to be small relative to the significant behavioral and emotional impairment in anxiety disorders (Lissek et al., 2005). Second, fear conditioning refers to situations in which feared stimuli are associated with threatening stimuli, but anxiety disorders are often characterized by the absence of real dangers and traumatic experiences (Rachman, 1977). Third, in a traditional fear conditioning paradigm the threat in the situation is often contingent and very clear, while differences between patients and control participants mostly emerge in ambiguous situations (Beckers, Krypotos, Boddez, Effting, & Kindt, 2013; Lissek, Pine, & Grillon, 2006). For example, if enhanced fear conditioning is observed in pathological populations or to prepared stimuli, the effects are often manifest in increased resistance to extinction learning (e.g. Michael et al., 2007; Öhman & Mineka, 2001; Öhman, 2009). In an extinction phase, an acquired conditioned response is usually reduced by the repeated presentation of the CS+ in the absence of the US. This can also be described as a more ambiguous situation because the participant is usually not informed about the absence of the US. Fourth, higher cognitive processes are often neglected in fear conditioning. In contrast to implicit arousal and defense reactions like fear potentiated startle responses, higher cognitive processes comprise for example impaired working memory functions, declarative memory, attention and expectancies about dangerous events. Given that worry, rumination and biased cognitions are important aspects of anxiety disorders, non-associative and cognitive bias approaches seem to be necessary to explain the full picture of pathologic fear and anxiety.

Beyond simple fear conditioning

One of the most critical objections against simple fear conditioning as an origin of pathological fear was that the majority of individuals suffering from biologically relevant phobias like water phobia or animal phobia cannot recall a traumatic event that might be involved in fear acquisition (Davey, 1992; Menzies & Clarke, 1995). This pattern of results was also confirmed, when information about potential conditioning events was obtained from parents of children with fear of water (Graham & Gaffan, 1997) and in prospective research (Poulton et al., 1998; see Poulton & Menzies, 2002, for a review). Rachman (1977) suggested three pathways of fear learning: fear conditioning, vicarious fear learning and instructed fear. Since most studies about the origins of fears did neither find evidence in support of modeled nor in support of instructed fear, Poulton & Menzies (2002) proposed a non-associative approach for the acquisition of biologically relevant phobia. According to this theory, humans are born with or acquire fears of evolutionary relevant objects very rapidly. Then fearful and non-fearful individuals would mainly differ in the opportunity and/or ability to unlearn such fears. This view is in accordance with findings that biologically prepared stimuli mostly delay the extinction of CRs, although the acquisition of CRs is usually not enhanced (Öhman, 2009). However, there is evidence that dental phobia which should not have been shaped evolutionarily, is associated with traumatic experiences (Poulton et al., 1997). Moreover,

moderating variables on conditioning effects may account for the absence of differences in traumatic experiences between high and low fearful persons. Such moderating variables may be trait anxiety (e.g. Indovina, Robbins, Núñez-Elizalde, Dunn, & Bishop, 2011), latent inhibition (i.e. decreased conditioning due to non-paired pre-exposure; e.g. Lubow, 1973) or a genetic influence on conditionability (e.g. Hettema, Annas, Neale, Kendler, & Fredrikson, 2003).

Additional factors that may have a moderating influence on fear conditioning or that may account for fear learning and maintenance itself are higher-order cognitive processes. The mechanistic view that fear responses are determined by the overt pairing of a CS with a US in Pavlovian conditioning has been challenged by several findings. As an example, simply informing participants about the contingency between stimuli or about the intensity of a US can lead to conditioned responses (Brewer, 1974). Therefore, it has been assumed that an inner representation of a US is more likely to determine conditioned responses than actual CS-US pairings (Kirsch, Lynn, Vigorito, & Miller, 2004). Furthermore, contingency awareness seems to be a necessary but not sufficient prerequisite for some conditioned responses to occur (Lovibond & Shanks, 2002), especially if the CS and the US are not contiguous in time (Clark & Squire, 1998). In this circumstance, working memory plays an important role in the acquisition of conscious contingency knowledge and conditioned physiological responses (Carter, Hofstötter, Tsuchiya, & Koch, 2003). Based on the current evidence, it seems very likely that fear learning is to a great deal determined by higher order cognitive processes such as attention, working memory and expectancies. However, cognitive processes are sometimes biased and do not reflect a completely accurate representation of the environment. In the first place, behavior, cognition and perception have all evolved to help an organism to survive and reproduce, but not to mirror the world accurately.

Cognitive bias

The idea that biased cognitions are in part responsible for the onset or maintenance of anxiety disorders has been discussed in earlier reviews (Mathews & MacLeod, 1994; 2005), but only recently evidence was found for a causal contribution of biased cognitions to anxiety (MacLeod & Mathews, 2012). Cognitive processes that may be biased in high anxiety include attention to and memory of threatening stimuli, interpretation of ambiguous stimuli, associative thinking and inhibitory control (Mathews & MacLeod, 2005).

To date, most convincing evidence was found for an attentional bias in anxiety. In contrast to non-anxious individuals, highly anxious individuals process threat-related stimuli with priority (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, van Ijzendoorn, 2007 for a review). Most theorists agree that both automatic and top-down attentional functions are involved in this prioritized processing (Bar-Haim et al., 2007; Mathews & Mackintosh, 1998; Öhman, 2008). Automatic attentional orienting towards threatening stimuli was found for instance in a visual search for spiders (Öhman et al., 2001). Even if threatening stimuli are presented subliminally (Mogg, Bradley, Williams, & Mathews, 1993) or completely unexpectedly (Wiemer, Gerdes, & Pauli, 2013), they can receive enhanced psychophysiological processing. These data are consistent with the idea of a fast and automatic threat detection mechanism that helps to orient attention to a stimulus for further elaborative processing (Öhman et al., 2001). According to a common model of threat processing (Bar-Haim, 2007),

information is passed from a preattentive evaluation system to a guided evaluation system. At this latter step of higher cognitive processing, anxious and fearful individuals seem to have difficulties in disengaging from threat (Fox, Russo, & Dutton, 2002), also when additional goal-related stimuli are competing for attentional resources (Gerdes, Alpers, & Pauli, 2008). Interestingly, an attentional bias to threat in anxiety seems to depend on inter-individual differences in self-reported attentional control (Derryberry & Reed, 2002).

Highly anxious individuals are also more likely to interpret ambiguous information as negative, which is for instance exhibited in comprehension latencies, if an ambiguous sentence is followed by another sentence that is only consistent with a non-negative interpretation of the ambiguous sentence (MacLeod & Cohen, 1993). Although, some studies also found evidence for a memory bias in support of negative stimuli in generalized anxiety disorder and high trait anxiety (Friedman, Thayer, & Borkovec, 2000; Russo, Fox, Bellinger, & Nguyen-Van-Tam, 2001), most studies do not support deviant memory processing in anxiety. Considering the whole picture of the literature, reviewers concluded that an attentional bias is more relevant for anxiety disorders while memory bias is closer related to depression (Mathews & MacLeod, 1994; 2005).

Current research of cognitive bias and anxiety involves cognitive bias modification (CBM) and its causal impact on biased cognitions and anxiety symptoms (MacLeod & Mathews, 2012). Previously, it had been shown that highly anxious individuals' cognitions such as attention and interpretation were more biased than those of low anxious individuals. Critically, this does neither imply that cognitive bias leads to anxiety nor that cognitive bias and therefore anxiety can be reduced in psychological treatment. However, recent evidence suggests that CBM leads to alternated biased cognitions (Dandeneau & Baldwin, 2009), reduces subclinical anxiety symptoms (Bar-Haim, Morag, & Glickman, 2011) and can even shift symptoms of generalized anxiety disorder (Amir, Beard, Burns, & Bomyea, 2009) and generalized social anxiety (Schmidt, Richey, Buckner, & Timpano, 2009) to a subclinical degree. In practice, CBM works similar as a cognitive bias test with the difference that participants are trained to selectively process neutral or positive information instead of negative information. For example, in a modified dot-probe paradigm, neutral and negative stimuli are presented simultaneously prior to a visual response target that always follows the neutral stimulus. As a consequence, participants will learn to ignore negative and attend to neutral information. This line of research encourages further investigations of the cognitive processes underlying fear and anxiety. At the moment, more knowledge is eligible about how and why cognitive biases emerge in anxiety, and whether some individuals are more prone to biased cognitions. In addition, besides attention and interpretation bias, further cognitive processes like the appraisal and expectancy of threats could be identified as potential targets of CBM.

An expectancy bias for aversive events has been less attended in reviews about cognitive bias in emotional disorders (Mathews & MacLeod, 1995; 2005). This is remarkable because the very essential function of fear and anxiety as affect programs is supposed to be the preparation of an organism for a potentially threatening event (Oatley & Johnson-Laird, 1987). This purpose of the emotion is realized via relatively specific response patterns of the autonomous nervous system (Collet, Vernet-Maury, Delhomme, & Dittmar, 1997; Ekman, Levenson, & Friesen, 1983; Rainville, Bechara, Naqvi, & Damasio, 2006). For example fear may be associated with an increase in heart rate and a decrease in high-frequency heart rate variability that is coupled with respiratory changes (Rainville et al., 2006). In addition, fear and anxiety seem to come with changes in

sensory sensitivity. Especially when the occurrence of an aversive event is uncertain as in anxiety, a more sensitive sensory system might be helpful to detect and identify a threatening stimulus, thus preparing the organism to cope with it ('fight or flight'). The very basic level of facial emotion expression is consistent with this idea. For example, widening the eyes as observed in fear and surprise expressions leads to improved visual processing as reflected in a larger visual field and faster eye-movements (Susskind et al., 2008). It may be speculated that fear and anxiety also causally alter the cognitive set in order to prepare the organism for an aversive stimulus. The most obvious cognitive process should be increased (biased) attention to an anticipated aversive event which may be reflected in increased expectancy of the aversive event. Then, the very essence of cognitive biases in anxiety might be a consequence of the fact that biased cognitions help to prepare for an upcoming aversive event.

Previous studies have demonstrated that anxiety is associated with enhanced expectancy of negative events. In general, anxious individuals seem to estimate the probability of future negative experiences higher than non-anxious individuals (Butler & Mathews, 1983; 1987). In these studies, however, the real proportion of negative events was not experimentally controlled. In some variants of fear conditioning in which the number of USs was equal between high and low anxious individuals, highly anxious participants rated the expectancy of a US higher than low anxious participants (Boddez et al., 2012; Chan & Lovibond, 1996). In these cases, expectancy was inflated in high anxiety when a second stimulus compounded with the CS+ predicted safety (inhibition) or was in contrast to the CS+ never paired with the US alone (blocking). Again, these situations are more ambiguous than the simple pairing of a CS and a US.

Importantly, an expectancy bias for aversive events was reported to be associated with fear-relevant stimuli in specific phobia (Cavanagh & Davey, 2000), panic disorder (Wiedemann, Pauli, & Dengler, 2001), posttraumatic stress disorder (Engelhard, de Jong, van den Hout, & van Overveld, 2009) and in patients with an automatic implantable cardioverter defibrillator (Pauli, Wiedemann, Dengler, & Kühlkamp, 2001). Such findings demonstrate that highly anxious individuals are cognitively prepared to learn associations between fear-relevant CSs and aversive USs. As a consequence, anxious individuals often think fear-relevant stimuli are associated with aversive events despite the fact that they are not (e.g. Tomarken et al., 1989). In other words, anxious individuals often display an illusory correlation between fear-relevant and aversive stimuli.

Preliminary conclusions

The emotional states of fear and anxiety constitute affect programs that help to deal with threats in the environment. Threatening situations differ in their level of ambiguity. While unambiguous threats evoke fear, ambiguous threats evoke anxiety. Differences between normal behavior and pathological fear are more likely to occur in ambiguous situations. Altered fear conditioning processes may be involved in the development of anxiety disorders, but probably the simple pairing of a CS with a US is not sufficient to explain the whole picture of pathological fear. Particularly, biased higher-order cognitive processes are very likely involved in the maintenance of pathological fear via the interpretation and perception of ambiguous threatening situations. One

of such cognitive biases is the overestimation of contingencies between fear-relevant stimuli and aversive consequences, also known as fear-relevant illusory correlations.

2.3. Illusory correlations in fear and anxiety

Original experimental paradigm

Individuals suffering from anxiety disorders often overestimate the proportion of aversive events following a fear-relevant stimulus. This estimated proportion is usually enhanced relative to neutral events following fear-relevant stimuli, relative to aversive events following neutral stimuli and relative to the objective proportion of aversive events following fear-relevant stimuli. An overestimation like this has been termed an illusory correlation or a covariation bias and can be described as a cognitive bias that promotes the subjective association of a CS and a US without objective contingencies in the environment. Therefore, illusory correlations can help to understand the onset and maintenance of fear and anxiety in some situations where classical fear conditioning is not a sufficient explanation, i.e. where there is no specific association between a CS and a US (see above).

The first study of a covariation bias in fearful participants (Tomarken et al., 1989) was an adaptation of a classic paradigm used to investigate illusory correlations between clinical diagnoses and symptoms (Chapman & Chapman, 1969; see excursion chapter 2.4.). Individuals high and low in fear of spiders or snakes participated in an illusory correlation experiment that served as an example for practically all following studies that examined covariation bias in fear. The participants were exposed to fear-relevant (spiders or snakes) and fear-irrelevant pictures (mushrooms and flowers). Following each picture, one of three outcomes occurred: an aversive shock (highly unpleasant but not painful; applied to the left arm), a neutral tone or nothing. The participants were instructed to "pay close attention to what is happening because your task is to determine whether or not there is a relationship between any category of slide and any of the outcomes following the slide" (Tomarken et al., 1989; p. 383). All kinds of picture-outcome combinations were exactly equally probable (33%). After the experiment, the participants completed a covariation questionnaire in which they had to estimate on visual analogue scales the percentage of a certain outcome, given that there had been a picture of a specific category. For example: "Given that you saw *a flower* slide, on what percentage of those trials was the flower followed by a *shock*?" Highly fearful individuals overestimated the percentage of fear-relevant/shock associations (≈ 50%) relative to other associations (≈ 30%), and relative to the low-fear group (≈ 40%). The low-fear group displayed an "attenuated covariation bias" (Tomarken et al., 1989; p. 384) indicated by differences in comparison with some but not all relevant associations. In a second experiment, the authors found that the effect was probably due to the aversiveness but not due to the saliency of the electric shock. A highly salient outcome consisting of a chime and a flashing light did not enhance covariation estimates for fear-relevant pictures. Interestingly, in a third experiment an explorative analysis of potentially mediating variables indicated that fear-relevant/shock

covariation estimates were predicted by pain ratings of the US in low fear participants. The authors concluded that basic perceptual processes may contribute to covariation bias.

The clinical relevance of these findings has been supported by studies showing that untreated individuals meeting DSM (Diagnostic and Statistical Manual of Mental Disorders) criteria for simple phobia displayed a covariation bias while treated phobics did not (de Jong, Merckelbach, Arntz, & Nijman, 1992; de Jong & Merckelbach, 1993). However, it should be noted that an earlier study found a similar covariation bias in untreated phobics, successfully treated phobics and healthy controls (de Jong & Merckelbach, 1991), indicating that covariation bias is not always specific to simple phobia. In this case (de Jong & Merckelbach, 1993), the usage of additional threatening cues (weapons) instead of neutral cues (mushrooms) might have eliminated biased spider-shock estimations. Considering the complete literature on covariation bias, the finding that covariation bias is equally pronounced in phobics and controls seems to be a rather exceptional finding. Later, it was even found that the extent of a covariation bias persisting after behavioral treatment predicted the return of fear in a two year follow-up (de Jong, van den Hout, & Merckelbach, 1995). These data are based on only a modest sample size of 13 spider phobics, but are consistent with the notion that covariation bias is an important determinant of pathological fear.

Expectancy bias or encoding bias?

The findings described to this point refer to covariation bias after an illusory correlation experiment, i.e. after exposure to random contingencies. It has been argued that this *a posteriori* covariation bias is in part explainable by an *a priori* expectancy bias, i.e. a bias to expect more shocks after fear-relevant pictures before such an experiment starts (McNally & Heatherton, 1993). In general, it has been discussed in the literature whether the a posteriori illusory correlations are a residue of an a priori expectancy bias or the consequence of an encoding (computational) bias during the exposure to random associations (Davey, 1995). One study found a priori expectations (McNally & Heatherton, 1993) in high and low fear of spiders that showed the same pattern as a posteriori covariation estimates (i.e. like in Tomarken et al., 1989). This suggested that a covariation bias is present before confrontation with actual contingencies and that it is not necessary to assume that the experience of real associations is encoded in a way that promotes covariation bias. In fact, the finding that high and low fearful individuals differ already on the level of a priori expectancies has been replicated in several other experiments (Cavanagh & Davey, 2000; Davey & Dixon, 1996; van Overveld, de Jong, Huijding, & Peters, 2010; van Overveld, de Jong, & Peters, 2006; Wiedemann et al., 2001). In addition, it is possible to induce an illusory correlation between neutral stimuli and aversive consequences by starting an illusory correlation paradigm with an unproportioned number of associations that promote this bias (de Jong, Merckelbach, & Arntz, 1990). Although, contingencies were again random over the whole period of the experiment, a covariation bias was generated by this non-random beginning. In sum, these findings suggest that an a posteriori covariation bias in highly fearful individuals is strongly influenced by a priori expectancies. However, the fact that a priori expectancies have an impact on a posteriori covariation biases does not exclude the possibility that high fearful individuals process actual contingencies differently. Actually, there is also evidence in favor of an encoding bias.

Mainly two reasons lead to the assumption that altered processing of CS-US associations may at least also contribute to an a posteriori covariation bias.

First, there are often situations in which an a priori expectancy bias exists for two subgroups of participants or two categories of pictures, but an a posteriori covariation bias exists only for one of the subgroups or picture categories. If expectancies are equal between groups and only one group displays an a posteriori covariation bias, then the expectancy alone cannot account for this difference. Therefore it may be assumed that the groups differ in on-line processing of the contingencies. For example, in contrast to McNally & Heatherton (1993), de Jong (1993) found that both high and low fear participants were equally likely to expect an aversive outcome after fear-relevant pictures. It was concluded that the existence of an expectancy bias is not a sufficient explanation of an a posteriori covariation bias because the latter is only present in high fear individuals. Unfortunately, both studies did not measure a posteriori covariation estimates in the same participants, but only conducted a thought experiment in which participants had to imagine that they would take part in an experiment in which different kinds of outcomes could occur after different kinds of pictures.

However, following studies reported both a priori or on-line expectancy estimates and a posteriori covariation estimates. They often found that both fearful and non-fearful participants expected fear-relevant stimuli to be associated with aversive outcomes, but only fearful participants continued to overestimate contingencies (Amin & Lovibond, 1997; Davey & Dixon, 1996; Kennedy, Rapee, & Mazursky, 1997). Importantly, two of these studies did not find differences in biased expectancies between high and low fear (Amin & Lovibond, 1997; Kennedy et al., 1997). Consequently, it is difficult to explain the a posteriori covariation bias in high fear solely on the basis of inflated expectancies. Similarly, an expectancy bias was observed for both ontogenetically and phylogenetically fear-relevant stimuli, but an a posteriori covariation bias did often only remain for phylogenetically fear-relevant stimuli (Amin & Lovibond, 1997; de Jong, Merckelbach, & Arntz, 1995; Mühlberger, Wiedemann, Herrmann, & Pauli, 2006). Ontogenetically fear-relevant stimuli are potentially dangerous stimuli that have emerged only recently in human evolution and did probably not influence genetic preparedness to develop fear of them (e.g. weapons, electric outlets, airplanes). Phylogenetically fear-relevant stimuli on the other hand share a history with human evolution for a longer period and may more likely have affected such a genetic preparedness (e.g. spiders, snakes, angry faces; see Öhman, 2009). It should be noted that other patterns of a priori and a posteriori covariation estimates were also found, for example that ontogenetic stimuli lead to both a priori and a posteriori covariation bias (Davey & Dixon, 1996). Such differences may be explained by methodological inhomogeneities between studies like in defining the level of fear, matching visual stimuli or using on-line expectancy ratings. Davey (1995) suggested that phylogenetic stimuli are more feared in the western culture than ontogenetic stimuli, and therefore more likely to be associated with aversive consequences. However, this does not imperil the rationale that in those cases of dissociations between a priori and a posteriori covariation bias, biased information processing could have accounted for this dissociation. Moreover, even in those cases where an a priori expectancy bias could explain an a posteriori covariation bias, the influence of an encoding bias may still play an additional or mediating role.

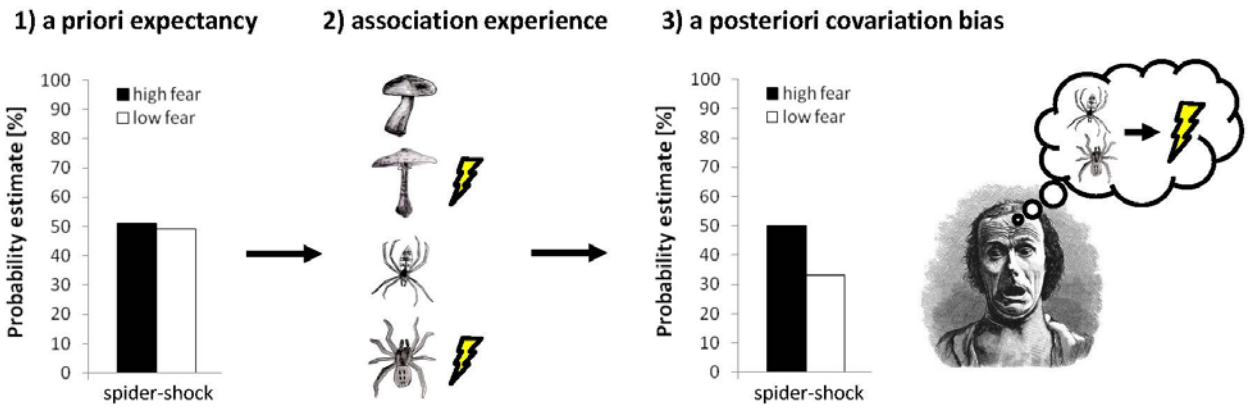

Figure 2.1. Schematic illustration of representative fear-relevant illusory correlations. Some parts of the empirical evidence suggest that both high and low fearful individuals suffer from an a priori expectancy bias (1), but high fearful individuals are more prone to an a posteriori covariation bias (3) after exposure to random contingencies (2). Therefore, fear-relevant illusory correlations may not only be a consequence of inflated expectancies, but also of biased association experiences.

The second reason why altered information processing might contribute to a posteriori covariation bias is that some studies measured on-line processes during illusory correlation experiments and found that fear-relevant/shock associations evoked different responses in comparison to fear-irrelevant/shock associations. Importantly, positive correlations between these responses and the a posteriori covariation bias were reported. Particularly, as already brought up above, in experiment 3 of Tomarken et al. (1989) especially high fear individuals and by the end of the experiment also low fear individuals reported an amplified pain experience to shocks after fear-relevant pictures relative to shocks after fear-irrelevant pictures. In low fear participants, pain ratings predicted a posteriori covariation bias. This can be considered as first evidence for a biased encoding approach to illusory correlations in fear. Unfortunately, there was no such correlation in the high fear group. It may be speculated that the high fear group reached a certain threshold of painfulness that promoted covariation bias but above this threshold painfulness and covariation bias were not correlated anymore. Another possibility could be that painfulness and covariation bias were not correlated because painfulness was measured on a self-report level. In order to identify biased encoding processes, it would be eligible to use more objective measures of physiological or brain responses. In addition, these measures would be less susceptible to demand effects of the experimental procedure. This means painfulness ratings could have been increased in the fear-relevant condition because participants were guessing that this was expected by the experimenters and tried to fulfill this hypothesis.

So far, experimenters have measured skin conductance responses (SCR), auditory startle responses (ASR) and electroencephalography (EEG) throughout illusory correlation experiments. In one of these studies, SCRs to spider images and to shocks following spider images were reported in treated and untreated spider phobics (de Jong, Merckelbach, & Arntz, 1995). While SCRs to spider images were only magnified in untreated participants, SCRs to the shocks after spider images were increased relative to control trials in both treated and untreated individuals. Since both treated and untreated spider phobics displayed similar a posteriori covariation estimates for spider trials, both groups were collapsed for a correlational analysis. Notably, only SCRs to shocks following spiders significantly predicted a posteriori covariation estimates but not SCRs to spider images

themselves. Similarly, de Jong & Merckelbach (1991) showed that SCRs to shocks following spider pictures were enhanced relative to shocks following mushroom and flower images. Although correlational analyses were not reported, the pattern of SCRs to shocks in the different groups (treated, untreated, healthy controls) was more similar to the a posteriori covariation bias than the pattern of SCRs to the images. In a later experiment the authors found a covariation bias in untreated phobics that was accompanied by elevated SCRs to spiders and shocks following spiders (de Jong & Merckelbach, 1993). In another study, biological and technological stimuli evoked comparable SCRs to cues, but different covariation bias scores (Amin & Lovibond, 1997). SCRs to outcomes were not reported. On the other hand, an expectancy bias for biological and technological threats was accompanied by similar SCRs to cues in an experimental setup without actual outcomes (Diamond, Matchett, & Davey, 1995; Honeybourne, Matchett, & Davey, 1993). Moreover, alprazomlam was found to attenuate SCRs to spider images without attenuating covariation bias, but SCRs to outcomes were not reported (de Jong & Merckelbach, 2000). Individuals suffering from generalized social phobia were found to overestimate the contingency between ambiguous social cues and angry faces (Hermann, Ofer, & Flor, 2004). Interestingly, at the beginning of the illusory correlation procedure, negative outcomes evoked enhanced SCRs in patients. In sum, there may be a relationship between covariation bias and electrodermal responding during the experience of cue-outcome associations, while SCRs to outcomes seem to be more closely connected to covariation bias than SCRs to fear-relevant cues, if both were reported.

If auditory startle stimuli (bursts of white noise) instead of cutaneous electric shocks are presented as US, an ASR can be measured. An ASR can reflect activation of a defense system in the brain and is usually modulated by hedonic valence of the affective state (unpleasant pictures potentiate the ASR) and inhibited under focused attention (Bradley, Cuthbert, & Lang, 1993). During an illusory correlation experiment using acoustic startle probes as USs, spider images evoked increased ASRs in spider phobics, while pictures of flight accidents did not induce increased ASRs in flight phobics (Mühlberger et al., 2006). In parallel, spider phobics rated startle stimuli following spiders as more unpleasant than startle stimuli following mushrooms, but flight phobics' valence ratings did not differ between conditions. Finally, only spider phobics displayed a covariation bias for spider images after the first experimental block. Although these findings are again correlational in nature, they are consistent with the idea that altered processing of outcomes may contribute to covariation bias. Particularly, from the present point of view empirical findings suggest that the US after fear-relevant stimuli is processed as more aversive and more arousing than the US after neutral trials despite both types of US are identical in physical intensity. Similar effects have been observed in the emotional modulation of pain (e.g. Kenntner-Mabiala & Pauli, 2005; Kenntner-Mabiala, Andreatta, Wieser, Mühlberger, & Pauli, 2008; Meagher, Arnau, & Rhudy, 2001; Roy, Piché, Chen, Peretz, & Rainville, 2009; Villemure & Bushnell, 2002). As a consequence, the more aversive US might be encoded preferentially and/or needs more attentional preparedness for future occurrences which might result in inflated covariation estimates.

Affective matching?

In the previous section, it was discussed why not only an expectancy bias, but also an encoding bias may lead to illusory correlations. Basically, it was outlined that the enhanced aversiveness or salience of an aversive outcome following fear-relevant stimuli relative to aversive outcomes following fear-irrelevant outcomes promotes such a covariation bias. A different explanation for biased encoding processes has been suggested by Tomarken, Sutton and Mineka (1995). According to these authors, it is not the enhanced aversiveness of the aversive outcome as such, but the similarity between the affective response to the fear-relevant picture and the aversive outcome. Therefore, it is an *affective matching* that leads to a posteriori covariation bias. Interestingly, when the belongingness of a stimulus outcome pairing was rated by the participants, the belongingness between damaged electric outlets and shocks was higher than between snakes and shocks - even in high fear participants. In contrast, when affective characteristics (e.g. negative and positive affect arousal, controllability) were rated, snakes and shocks were more similar than electric outlets and shocks (Tomarken et al., 1995). However, the authors did not find an illusory correlation between damaged electric outlets and shocks, but between snakes and shocks. It can be concluded from this study that affective matching processes are more likely to be involved in illusory correlations than semantic matching processes. Davey & Dixon (1996) had participants rate several aspects of cue and outcome qualities and found that the similarity between cues and outcomes in valence, arousal, and overall similarity predicted a priori expectancy bias and a posteriori covariation bias.

The idea of affective matching processes has also been picked up recently to explain why a cue of uncertainty leads to enhanced expectation of negative pictures although negative and neutral pictures were equally likely to occur (Grupe & Nitschke, 2011; Sarinopoulos et al., 2010). The idea is that uncertainty itself is an aversive experience and viewing negative pictures is also an aversive experience. Therefore, an affective matching takes place and promotes expectancy and covariation bias. In relation to an affective matching, it makes sense that high spider fearful individuals also expect more disgust-related outcomes (nauseating juice) following spider images than other images (de Jong & Peters, 2007b; Olatunji, Cisler, Meunier, Connolly, & Lohr, 2008; van Overveld, de Jong, & Peters, 2006). Feelings of disgust are an important aspect of fear of spiders (Mulkens, de Jong, & Merckelbach, 1996) and are probably evoked by the imagination of both spiders and nausea. However, the affective matching approach leaves some questions unanswered so far. First, why is it exactly that affective matching leads to the maintenance of a covariation bias despite random contingencies? Particularly, why does affective matching but not semantic matching lead to a covariation bias? Finally, if the affective matching between stimuli and outcomes leads to the notion of covariation, then why are associations between neutral stimuli (e.g. mushrooms) and neutral outcomes (e.g. tones) typically not overestimated? One of the hypotheses of the present dissertation is that matching processes are not necessary, if the amplified aversiveness of the aversive outcome is taken into account. Nevertheless, it should be noted that enhanced aversiveness and matching processes do not preclude each other and could contribute to illusory correlations in company. Also Tomarken et al. (1995) who were arguing for the affective matching account suggested that, an enhanced response to specific outcomes may have been a mediating process between affective matching and illusory correlations: “[…] when two stimuli elicit similar affective responses, exposure to one may potentiate responses to the second. […] Perhaps such greater responsivity to shocks occurring after snake slides rendered

snake-shock co-occurrences more distinctive and thus facilitated the development of illusory correlations." (p. 324).

Covariation bias beyond fear of animals

Blood-injury fear. Covariation bias was not only discussed as a maintaining factor of the previously described fear of animals, but also for other anxiety disorders. One study applied the original illusory correlation paradigm (Tomarken et al., 1989) to blood-injury fear (Pury & Mineka, 1997) and found that both high and low fear participants overestimated the contingency between blood-injury pictures and aversive shocks. There was no influence of prior (undiagnosed) fear level. In another study using self-administered aversive shock and a bad tasting fluid as aversive outcomes, also reported no influence of fear of blood-injection on an expectancy bias for harm and disgust outcomes following blood donation stimuli (de Jong & Peters, 2007a). Nevertheless, a third study found an impact of prior fear on negative outcome expectations (van Overveld, de Jong, & Peters, 2010). The heterogeneous findings might be a consequence of methodological differences in the definition of prior fear and the variety of fear-relevant cues.

Social anxiety. Several studies examined illusory correlations in social anxiety. An a posteriori covariation bias was found in both (undiagnosed) high and low socially anxious individuals between angry faces and electric shocks (de Jong, Merckelbach, Bögels, & Kindt, 1998). In contrast, another study found that both high and low socially anxious individuals were able to correct an initial bias to expect negative outcomes after angry faces, but only low socially anxious participants (median split) kept a bias to expect more positive outcomes after happy faces (Garner, Mogg, & Bradley, 2006). In an experiment with formally diagnosed patients with generalized social phobia, patients but not controls overestimated the contingency between social scenes and negative emotional facial expressions after the experiment (Hermann et al., 2004). In sum, the evidence for illusory correlations in social anxiety is mixed and mostly based on subclinical samples.

Flight phobia. To date, two studies investigated illusory correlations in flight phobia. One did not result in a covariation bias for pictures of airplane crashes and aversive outcomes (Mühlberger et al., 2006), another study found that high fear individuals displayed a covariation bias, but were able to correct it within 72 trials (Pauli, Wiedemann, & Montoya, 1998). It may be concluded that the perception of covariations is indeed biased in flight phobia, but not as excessive as in animal phobia. The empirical data fit with the notion that covariation bias occurs more likely in phylogenetically relevant fears (e.g. Amin & Lovibond, 1997).

Panic disorder. Panic-prone individuals also over-associate fear-relevant stimuli (emergency scenes) with aversive outcomes while less fearful controls do not (Pauli, Montoya, & Martz, 1996). Also, patients with panic disorder display an expectancy bias for negative outcomes following emergency slides, while healthy controls do so to a significantly reduced degree (Wiedemann, Pauli, & Dengler, 2001). An a posteriori covariation bias for emergency scenes was also observed in panic disorder (Pauli, Montoya, & Martz, 2001). Interestingly, a high contingency between fear-relevant slides and aversive outcomes along with a low contingency between fear-irrelevant slides and aversive outcomes leads to a covariation bias in a subsequent block of random contingencies in both panic-prone and low fearful individuals (Pauli et al., 1996). In contrast, a high contingency

between fear-irrelevant slides and aversive outcomes along with a low contingency between fear-relevant slides and aversive outcomes does not lead to a covariation bias (Pauli et al., 2001). Again, when comparing the two studies, this is an example where expectancies were comparable (due to induction by objective contingencies here), but the ability to correct these expectancies depended on stimulus characteristics. That is, there was an induced expectancy bias for fear-relevant and fear-irrelevant slides, but the a posteriori covariation bias only remained for fear-relevant slides. This further strengthens the above outlined position of biased encoding processes during exposure to random contingencies. At least, the mere preservation of an expectancy bias seems not to be a sufficient explanation for an a posteriori covariation bias. Finally, one study did not find an a posteriori covariation bias in patients suffering from panic disorder (Amrhein, Pauli, Dengler, & Wiedemann, 2005), but an enhanced contingent negative variation (CNV) during emergency trials. The CNV is an event-related potential that can be measured electro-physiologically over midline sites during the anticipation of an event (e.g. Cui, Egkher, Huter, Lang, Lindinger, & Deecke, 2000). The extent of negativity can be considered as a function of outcome probability or intensity (Birbaumer, Elbert, Canavan, & Rockstroh, 1990; Irwin, Knott, McAdam, Rebert, 1966) and may thus be an indicator of expectancy or preparation. While in healthy controls the CNV was enhanced for both emergency and spider slides, in panic patients the CNV was specifically enhanced in emergency trials (Amrhein et al., 2005). The CNV might be useful to assess on-line expectancy during illusory correlation experiments more objectively than self-report measures. However, CNV effects even may have been attenuated because of the simultaneous presentation of emotional pictures. Emotional pictures are known to evoke a late positive potential (LPP) that might have had an effect in opposition to the CNV (Schupp et al., 2000).

Patients with cardioverter defibrillator. An automatic implanted cardioverter defibrillator is an implanted device that can stop life threatening malignant heart arrhythmias via the delivery of an electronic shock to the heart (Böcker et al., 1993). Patients carrying one of these defibrillators describe the experience of such an electronic shock as aversive and painful (e.g. Lüderitz, Jung, Deister, & Manz, 1994). In a thought experiment (Pauli, Wiedemann, Dengler, & Kühlkamp, 2001), patients who already had experienced a shock in their lifetime, expected more negative outcomes after emergency stimuli than patients who had not experienced a shock. In addition, this expectancy bias was correlated with trait anxiety. The results generalize the finding that an expectancy bias can be induced by the experience of aversive events (e.g. de Jong, Merckelbach, & Arntz, 1990) to an ecologically valid setting and suggest that both the experience of aversive events and the level of anxiety contribute to the development of an expectancy bias.

Post-traumatic stress disorder. One prospective study recruited soldiers who had come back from a deployment in Iraq 2-5 months before the experimental procedure (Engelhard, de Jong, van den Hout, & van Overveld, 2009). It was found that those participants who expected a high proportion of aversive consequences following deployment-related pictures were more likely to display symptoms of post-traumatic stress disorder 15 months after the time of deployment. These results further support the idea that fear-relevant expectancy bias is closely related to the development of anxiety disorders (de Jong, van den Hout, & Merckelbach, 1995).

Conclusions

In summary, there is enough evidence so far to assume that high fear individuals are prone to selectively associate fear-relevant stimuli with aversive outcomes even if objective contingencies are random. Probably, these selective associations are especially pronounced in phylogenetic fears like animal phobia. Illusory correlations are likely to maintain pathological fear and it is eligible to find out more about how and why illusory correlations are manifested. So far, the underlying processes are poorly understood, but some suggestions can be made based on the current literature. First, a priori expectations have an influence on illusory correlations, but also biased on-line encoding processes are necessary to explain illusory correlations. Second, one biased encoding process might be the enhanced matching of affective responses to stimuli and outcomes if both responses are aversive. Third, an alternative and perhaps simpler explanation would be that the enhanced aversiveness and/or salience of aversive outcomes following fear-relevant stimuli is sufficient to induce and maintain illusory correlations.

From a methodological point of view, the relatively high heterogeneity in the findings of illusory correlations is not satisfying. For instance, high and low fear individuals differed in expectancies and/or covariation estimates in some but not all studies of particular fears. One reason for this might be that illusory correlation parameters are usually based on only two or three self-report measures which may be relatively susceptible to error variance. For example, the knowledge or hypothesis that contingencies were random could completely eliminate a covariation bias, although a more intuitive impression of selective contingencies might still have been present in those participants. Moreover, participants could use very different strategies to estimate contingencies like trying to remember the number of associations, relying on gut feelings or even trying to fulfill the experimenter's hypotheses. Therefore, it is highly recommended to develop more reliable measures of covariation bias. One possibility would be to include more objective physiological data like SCRs, the CNV or hemodynamic responses in the brain to describe biased stimulus processing on multiple levels. Another way would be to include many different pictures per category and to retrieve a covariation estimate for each picture, i.e. ask whether an aversive outcome had been associated with a specific picture or not. Such repeated measurement should help to extinguish error variance and result in a more reliable illusory correlation parameter. Although, this measurement might describe illusory correlations on a slightly different and perhaps more implicit level than global assessment by direct contingency estimates, the relevance for fear maintenance should be very similar.

2.4. Excursion: Illusory correlations outside fear and anxiety

An extensive review of illusory correlations independent from fear and anxiety would be beyond the scope of the current thesis. Therefore, the following paragraphs reflect on the most important and influential findings on that field. This should give an impression of the general relevance of this basic cognitive phenomenon and enable to infer some psychological processes underlying fear-relevant illusory correlations.

First experimental data about 'illusory correlations' were published by Chapman (1967) and Chapman and Chapman (1967; 1969). In a first experiment, Chapman (1967) demonstrated that associations between highly related words, such as *bread* and *butter* were systematically overestimated by naïve subjects relative to less related words, such as *bread* and *foot*. Moreover, the association between semantically unrelated but long and distinctive words, such as *envelope* and *sidewalk* were overestimated. Possibly, superficial visual similarity promotes illusory correlations in addition to semantic similarity and relatedness. Alternatively, distinctiveness as such generates attentional responses that promote illusory correlations, possibly in terms of enhanced stimulus encoding. However, a further important finding was that illusory correlations for related word pairs declined as a function of repeated presentation of word pairs, suggesting that participants were able to adjust their probability estimates in the direction of the correct value.

In the following studies, Chapman and Chapman (1967; 1969) showed that their findings were relevant for typical contemporary instruments of psychodiagnosis. In the Rorschach test for example, the subject is asked to describe what he or she perceives in abstract and ambiguous visual figures. Although not very sensitive, some answers seem to be slightly associated with different kinds of psychological traits (e.g. Mihura, Meyer, Dumitrascu, & Bombel, 2012; Wood et al., 2010). However, many clinicians believed that certain answers, such as seeing feminine clothing in the figures, were signs of male homosexuality (Chapman & Chapman, 1969). Despite the conviction of experienced clinicians, previous empirical investigations suggested that most of such popular signs of homosexuality were invalid (Wheeler, 1949). In an experimental approach, participants were confronted with a random set of associations between symptoms and answers in the Rorschach test (Chapman & Chapman, 1968). Although, there were no true relationships between symptoms and answers in the set, participants displayed illusory correlations between popular pairs of symptoms and percepts, such as homosexuality and feminine clothing. Interestingly, these illusory correlations persisted, even if there were objectively negative correlations, thus preventing the participants from detecting true relationships in a set of associations.

According to Tversky and Kahnemann (1973), humans estimate probabilities and frequencies on the basis of the availability heuristic. This theory follows a simple and comprehensible logic: The more often an individual experiences the occurrence of an event or the association of two events, the easier it is to recall these events at a later time point. So, there is a natural correlation between the frequency of events and the ease of recall or availability. Now, when we want to estimate the frequency of events, we make use of this natural correlation and assess the availability of events or associations. For example, when I try to estimate the probability that my favorite football team will win the match on next Saturday, I will try to recall their last victories and losses against the opponent. The more victories are mentally available to me, the more likely I will expect

my team to win next time. However, the availability heuristic does not always turn out to be a good heuristic for frequency estimations, because the ease of recall or the availability is not only determined by the true frequency of events. For example, if my team mostly won when I was seeing the match in the stadium and mostly lost when I was seeing the match on the couch, I will probably more likely recall victories due to the more arousing experience and will be unnecessarily optimistic for the weekend. In the same vein, pairs of events come easier to our minds if the paired events are similar or in some other way related. Indeed, Tversky and Kahnemann (1973, study 10) found that there was a strong correlation between one group of participants' estimated ease of recall of word pairs and another group of participants' correct recall performance. Following this argumentation, the occurrence of illusory correlations can be explained. An individual experiences associations between similar and dissimilar items. Associations of similar items may not be more frequent than associations of dissimilar items, but they are easier to be recalled later. The individual realizes the ease of recall and concludes that similar items were paired more often than dissimilar items. This explanation is very similar to the affective matching account that has been suggested to explain fear-relevant illusory correlations (see above). However, as argued above fear-relevant associations are not only affectively similar, but should also evoke a more intense experience. This intensified experience should also affect the ease of recall.

Hamilton and Gifford (1976) elaborated on Chapman and Chapman's (1969) finding that illusory correlations emerged for two exceptionally long words. They argued that the basic mechanism behind that effect is that the statistical infrequency of a long word leads to greater attentional encoding, presumably similar to an oddball effect of infrequent stimuli on cortical processing (Sutton, Braren, Zubin, & John, 1965). Consequently, two infrequent stimuli occurring conjointly receive particularly great attentional processing, leading to an illusory correlation. Interestingly, Hamilton and Gifford (1976) demonstrated that this mechanism could generate stereotypical beliefs about minorities. Participants were presented short behavior descriptions about different members of two artificial groups A and B. One group involved fewer people. In experiment 1, behavior descriptions were either frequent and desirable or infrequent and undesirable. Although, the participants could not have stereotypic beliefs about groups A and B before the experiment, they overestimated the probability of the minority acting undesirable. In experiment 2, the authors showed that illusory correlations can also be generated between minorities and desirable behaviors if desirable behaviors were less frequent than undesirable behaviors, suggesting that these cognitive illusions were more likely consequences of distinctiveness than valence. Interestingly, such frequency-based illusory correlations increase with decreasing working memory capacity (Eder, Fiedler, & Hamm-Eder, 2011).

Alloy and Tabachnik (1984) reviewed the literature on covariation assessment by humans and animals. They came to the conclusion that a covariation perception is always formed by the joint influence of prior expectations and current situational information. An argument for the influence of prior expectations can be found for example in a study by Langer and Roth (1975). They discovered that people are more likely to expect a desired outcome in coin tossing, if they had started with desired outcomes than if they had started with undesired or random outcomes. As described above, de Jong et al. (1990) also found that illusory correlations can be induced by a non-random beginning of a series of associations, thus confirming the influence of expectations. On the other hand, humans' contingency estimates can be a quite accurate reflection of current

situational information about covariations, especially if stimuli are neutral. Non-depressed students for example accurately estimated the contingency between their own behavior and a neutral outcome (Alloy & Abramson, 1979). On the other hand, they overestimate the contingency (and therefore their control) between behavior and a positive outcome. Depressed students were accurate in both cases, but overestimated the contingency between others' behavior and positive outcomes (Martin, Abramson, & Alloy, 1984). These findings could be explained by assuming that depressed and non-depressed individuals differ in their expectations about contingencies between their own behavior and positive outcomes. If confronted with covariation information deviating from prior expectations, their contingency estimations are biased towards prior expectations. Alloy and Tabachnik (1984) distinguish between four cases of the relative strength of prior expectations and current situational information and their impact on covariation assessment (see Figure 1). In case 1, both expectations and situational information are weak. Then, the individual will not assume a covariation pattern or if so with low confidence. In case 2, the individual has strong prior beliefs, but weak situational information. Then, he or she will assume a covariation pattern depending on prior beliefs. In case 3, the prior beliefs are weak, but the situational information is strong. Then, the covariation assessment will be a relative accurate reflection of current situational information. In case 4, both expectations and situational information are strong. If these sources of information are in line with each other (case 4.1), covariation can be estimated with high confidence. If they are contradictory (case 4.2), the individual suffers from a 'cognitive dilemma'. The individual has to solve this dilemma by either relying more on situational information or prior beliefs. The authors suggest that covariation estimates are mostly biased towards prior expectations, unless disconfirming information is particularly salient.

The case of fear-relevant illusory correlations can be circumscribed as such a cognitive dilemma. The individual a priori expects spiders to be associated with shocks, but is exposed to the strong situational information that they are not. The question remains why non-fearful individuals are able to estimate contingencies more accurately than fearful individuals. One possibility might be that fearful individuals have stronger prior beliefs and are therefore more biased towards these beliefs. This explanation seems to be in conflict with the observation that both fearful and non-fearful individuals expect spiders to be associated with aversive outcomes, at least there is often no significant difference in the magnitude of contingency expectations (Amin & Lovibond, 1997; Kennedy et al., 1997). However, it might be that contingency expectations are similar in magnitude, but less flexible in fearful individuals. Then, covariation estimates would be more biased towards expectations in spider fearful subjects. An additional possibility is that fearful people suffer from distorted situational information processing. This is one of the main assumptions of the present thesis and might be added to Alloy and Abramson's (1984) model of prior beliefs and expectations (Figure 2.2). It should be noted here that the effect of distorted situational information processing could occur independently from the effect of less flexible prior expectations. That is, the two explanations do not preclude each other. Although, Alloy and Abramson's (1984) 4-case-model seems to be in accordance with the empirical evidence of covariation assessment, there might be one shortcoming of the model: It does not explicitly explain why the covariation estimates of pairs of statistically infrequent events can also be biased (Hamilton & Gifford, 1976). This further justifies the additional emphasis on that situational information processing can be distorted. In this case, infrequent events are more salient than frequent events and might provoke strengthened encoding.

<table>
<tr><th></th><th>situational information weak</th><th colspan="2">situational information strong</th></tr>
<tr><td>expectations weak</td><td>case 1:
no covariation estimate or covariation estimate with low confidence</td><td colspan="2">case 3:
covariation estimate in line with situational information</td></tr>
<tr><td rowspan="2">expectations strong</td><td rowspan="2">case 2:
covariation estimate in line with prior expectations</td><td>no conflict</td><td>case 4.1:
covariation estimate with high confidence in line with situational information and expectations</td></tr>
<tr><td>conflict</td><td>case 4.2:
‘cognitive dilemma’: covariation estimate mostly biased towards expectations or in line with distorted situational information processing</td></tr>
</table>

Figure 2.2 An adaptation of Alloy & Tabachnik's (1984) model of illusory correlations. According to this theory, an illusory correlation can occur if prior expectations and situational information are contradictory. In this case (4.2), a cognitive dilemma is often solved by estimating covariation in line with prior expectations. In this adaptation, it has been added that covariation estimates can also be influenced by distorted situational information processing. See main text for further explanations.

Taken together, previous research of illusory correlations independent from fear suggests that contingencies are overestimated if pairs of events are similar, distinctive, infrequent, or in general, if strong prior expectations about covariations exist. Such overestimations might be attenuated in individuals with high working memory capacity or under conditions with low working memory load (Eder et al., 2011). A major difference between fear-relevant and non-fear-relevant illusory correlations might be that feared stimuli and aversive consequences are not only similar, but very likely also interact with each other, so that the aversiveness and distinctiveness of the aversive outcome could be amplified in the presence of feared stimuli. At least, positive and negative emotions can modify the aversiveness of experiences (e.g. Kenntner-Mabiala & Pauli, 2005). Fear-relevant illusory correlations might be the consequence of affective similarity, amplified distinctiveness and impaired working memory performance. The emotion of fear or anxiety might even intensify illusory correlations by distorted information processing.

2.5. A model for the maintenance of fear-relevant illusory correlations

For the purpose of understanding illusory correlations in fear and anxiety, a model is proposed of how illusory correlations are developed and maintained and related to fear and anxiety. Then, the influence of specific factors on illusory correlations can be tested and ideally identified for modification to reduce biased cognitions, fear and anxiety.

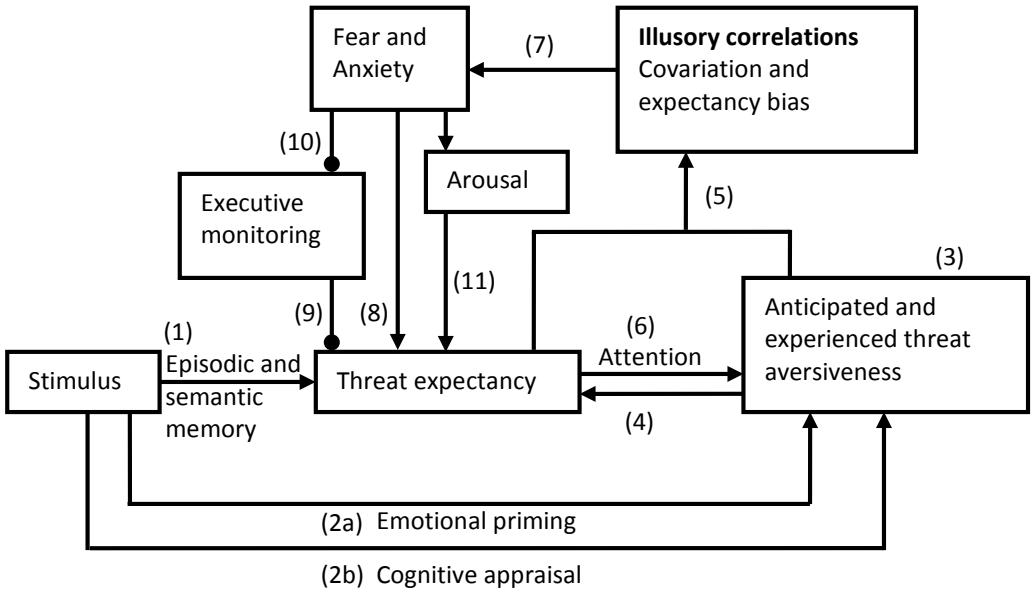

Figure 2.3 The model explains which psychological processes might be involved in the development and maintenance of illusory correlations in fear and anxiety. Dots indicate decreasing influences, arrows increasing influences. See main text for a description of the model

Model description

When a (1) stimulus appears that might or might not be associated with an aversive outcome, episodic memory is engaged to search for previous experiences of associations with aversive outcomes (availability heuristic; see chapter 2.4.). Moreover, especially if previous experience is lacking, semantic memory is involved to retrieve knowledge about the dangerousness of the stimulus (Davey & Craigie, 1997; Davey & Dixon, 1996). At this stage, the initial expectancies may also be influenced by fear of the stimulus which in turn is also based on previous experiences like conditioning, vicarious learning and instruction (Rachman, 1977). However, since commonly shared semantic knowledge should be an important heuristic for high and low fearful participants, in some cases they may not yet differ at the beginning of an illusory correlation experiment (e.g. Kennedy et al., 1997). In parallel, the (3) aversiveness of a potentially aversive outcome is (2) evaluated. This could happen in form of direct emotional priming (2a; Lang, Bradley, & Cuthbert, 1990) that leads to an experience of increased outcome intensity and aversiveness. This assumption is based on the findings of emotional modulation of pain perception (e.g. Kenntner-Mabiala & Pauli, 2005; Roy et al., 2009) and the modulation can be considered as a more hard-wired, bottom-up or impulsive process. At the same time, the outcome aversiveness can also be

appraised on a more cognitive and reflective level (2b) which can also be automatic but is already realized during the anticipation of an aversive event. Indeed, emotional experiences can be regulated by cognitive reappraisal (Gross, 1998). To be clear, this means the physically identical aversive outcome can be evaluated or experienced as more or less aversive depending on the fear-relevance of the context.

In both cases of emotional modulation and cognitive evaluation the aversive outcome will be (3) experienced or anticipated as more aversive if it follows a fear-relevant stimulus. This amplified aversiveness has two consequences. First, the expectancy of its occurrence is enhanced (4) because increased expectancy means being prepared for coping with the aversive event, and being prepared is more important for more aversive events. In other words, expectancy is not only determined by the objective probability of an event, but also by the motivational relevance, i.e. in the context of fear, the aversiveness [1] of an event. Second, the more aversive outcome is encoded preferentially relative to the physically identical but psychologically less aversive outcome (5). Indeed, increased SCRs to a US following a fear-relevant CS (de Jong, Merckelbach, & Arntz, 1995) might reflect increased arousal that promotes memory processes (Cahill & McGaugh, 1998). This preferential encoding could lead to a retrieval benefit of those CS-US associations that promote fear, and thus influence decision making in a fear confirming way when an expectancy or contingency estimation is made at a later time point. After all, the ease of information retrieval can be a powerful moderator of decision making (Tversky & Kahnemann, 1973). In addition, increased expectancy is associated with increased (6) attention to aversive outcomes, and so also contributes to the enhanced experienced aversiveness of the US. Evidence for this assumption comes for example from an experiment, in which validly cued painful stimuli were rated as more painful than invalidly cued painful stimuli (Dowman, 2001).

Together, the enhanced threat expectancy, outcome aversiveness and encoding processes lead to an illusory correlation between fear-relevant and aversive outcomes (5). This illusory correlation sustains feelings of fear and anxiety in the presence of the fear-relevant stimulus (7). As argued before, a widened attentional focus and increased expectancy of an aversive event should be a part of anxiety as an affect program. Therefore, anxiety amplifies or sustains threat expectancy (8) and fulfills a vicious circle that can be broken up by what might be called executive contingency monitoring (9). Executive monitoring is necessary to direct attention to less salient fear-disconfirming outcomes (e.g. nothing happens when a spider appears) and to compute an accurate perception of CS-US contingencies. As threat expectancy is initially increased, the monitoring process will lead to a decrease of threat expectancy (9). This process may be dependent on working memory that sustains information from previous trials and serves memory encoding. Working memory capacity in turn will be reduced or preoccupied by the emotion of fear (10).

[1] This kind of reasoning would also predict that very positive outcomes or, in the context of substance addiction for example, outcomes of high craving will also result in enhanced expectancy. However, since missing a positive outcome should be related to lower costs than missing a negative outcome (Dawkins & Krebs, 1979), human cognition may be prone to a negativity bias (Vaish, Grossman, & Woodword, 2008). Therefore a general expectancy bias for positive outcomes should also be present, but less pronounced as an expectancy bias for negative outcomes.

Finally, there is some evidence that unspecific arousal (independent of emotional valence) might play a role in the generation of an expectancy bias and illusory correlations (11). Particularly, it was found that the frequency of aversive startle probes is not only overestimated for negative but also for positive emotional pictures and imaginative scripts (vanOyen Vitvliet & Vrana, 2000; Pauli, Diedrich, & Müller, 2002). In contrast, neither physical exercise (Cavanagh & Davey, 2001; de Jong & Merckelbach, 2000) nor pharmacological dampening of physiological hyper-responsivity has an effect on US expectancy, resp. illusory correlations. Therefore, the assumption of an influence of unspecific arousal can only be made with caution.

Naturalistic example

An application of this model to real life scenarios outside illusory correlation laboratory experiments should be possible too. For instance, it can explain how biased threat expectancy hampers the extinction of fear of spiders. When seeing a spider, fear is elicited in spider fearful individuals, which automatically changes the attentional set to expecting an aversive event. One specific aversive event commonly feared by spider phobics is, for example that the spider moves quickly. Accurate or inaccurate semantic knowledge (e.g. “Spiders often move very quickly to attack their prey.”) and also episodic memory (remembering situations in which a moving spider was observed) can increase the expectation that the spider will move. On a cognitive level, the movement of the spider will be evaluated negatively (“If the spider moves, that will be very terrible.”). Thus, the anticipated aversiveness of the movement is increased and the person does not want that to happen. In fact, the person must be ready to flee if that happens and that is why attention is focused on a potential spider movement. This further enhances the expectancy of a movement. Maybe the spider will actually move and will so evoke a feeling of fear and disgust. Then, the aversiveness of the experience of seeing the spider moving will be enhanced due to an emotional modulation and the negative cognitive evaluation. This leads to a strengthened encoding of this event which will be retrieved easier when the person is confronted again with a spider. Then, expectations are raised again and the sequence of biased cognitive processing begins from the start. In case the spider will not move the next time, cognitive resources that are necessary for contingency monitoring will either be reduced due to fear and/or preoccupied by the observation of the spider and negative thoughts about the spider. As a consequence the outcome that in fact the spider did not move will not be integrated in the perception of contingencies or won’t be even registered consciously. Therefore, it needs a lot of experience until the phobic notices that spiders do not move as often as one might think and as a consequence extinction of spider fear needs prolonged exposure. Potential targets for cognitive treatment components could be the cognitive evaluation of spider movements or the voluntary direction of attention towards situation in which the spider did not move.

Why research in illusory correlations is important

Biased expectancies and illusory correlations reflect increased mental activity concerning a potentially upcoming aversive event. This increased mental activity may involve increased attention to the aversive stimulus and the imagination of its aversive nature. Given that the expectancy of an aversive event is exaggerated relative

to its actual probability, this mental activity causes unnecessary personal suffering. In addition, the cognitive resources that are occupied by attention to and imagination of potentially negative events in the future should be less available for the experience of the present moment, the engagement in other plans or the down-regulation of negative emotions. Moreover, the expectancy bias may not only maintain anxiety on the long run, but it may be a very essential cognitive part of anxiety itself. Considering that fear helps to prepare for an aversive event (Oatley & Johnson-Laird, 1987), one might also say that the expectancy bias sustains anxiety in the very moment of heightened expectancy, or in other words, the expectancy *is* part of the anxiety response Since the heightened expectancy is a cognitive phenomenon, it seems to be especially qualified to be modified in cognitive behavioral treatment. In addition, since the expectancy is a very essential part of anxiety, this should be a way to change anxiety directly.

2.6. Aim and hypotheses of the current dissertation

Four experiments were run to support the proposed model explaining illusory correlations. Testing the validity of the whole model is beyond the scope of this dissertation thesis. So, its focus lays on two main aspects. Since the evidence regarding an influence of arousal is ambiguous at the present moment, this was one factor that was tested in two studies. In addition, all experiments tested the hypothesis that outcome aversiveness is enhanced after negative emotional stimuli and leads to illusory correlations (see abstract for a short summary). In particular, the following hypotheses were tested in the present dissertation project:

1) The relationship between positive emotional stimuli and aversive consequences should be overestimated relative to neutral stimuli and aversive consequences (due to a general impact of arousal on illusory correlations).
2) The relative aversiveness of aversive consequences following emotionally negative stimuli compared to neutral stimuli should be correlated with the relative overestimation of aversive consequences following emotionally negative stimuli (due to the increased encoding depth of more aversive consequences).
3) The experimental elevation of outcome aversiveness will causally lead to an illusory correlation between neutral stimuli and the most aversive outcomes (due to the increased encoding depth of more aversive consequences).
4) Painful consequences following phobia-relevant stimuli should evoke increased brain activity in pain processing areas which should be correlated with the overestimation of these consequences relative to painful consequences following neutral stimuli (due to the increased encoding depth of more aversive consequences).
5) The overestimation of painful consequences following phobia-relevant stimuli should be associated with altered brain activity in brain areas associated with contingency monitoring (due to impairing impact of fear on this process).

3. Experiment 1: The influence of positive arousing stimuli on illusory correlations

3.1. Introduction

As described in the model for the maintenance of fear-relevant illusory correlations (chapter 2.5.), it is assumed that emotionally non-specific physiological arousal can trigger biased expectancy of aversive events and thus a covariation bias. There is no doubt that phobic stimuli induce heightened arousal in people suffering from specific phobia relative to non-fearful individuals (e.g. Aue, Hoeppli, & Piguet, 2012). Therefore, increased physiological arousal might be a potential mechanism that sustains illusory correlations in fear and anxiety. Cavanagh & Davey (2001) suggest three possible ways by which non-specific arousal could mediate heightened expectancy of negative outcomes. First, arousal may lead to attentional narrowing (Christianson, 1992) and so people focus more on typical features of fear-relevant stimuli, i.e. danger-related features. Second, decision making may become more emotional at high arousal states (Halberstadt & Niedenthal, 1997), and the intensity of harm-relevant features of fear-relevant stimuli may be increased. Third, heightened arousal may be used as a retrieval cue (Clark, Milberg, & Ross, 1983), and so it may be more likely that negative arousing information is retrieved when a decision is made about outcome expectancies.

Alternatively to these three suggestions, non-specific arousal could influence the decision making process involved in an expectancy rating by the interpretation of bodily sensations. For example, slides of seminude females are rated as more attractive by men, when they get false feedback of a heart rate increase or decrease (Valins, 1966). Moreover, false heart rate feedback induces anxiety in patients with panic disorder (Ehlers, Margraf, Roth, Taylor, & Birbaumer, 1988), and more negative appraisal of the own social performance in socially anxious individuals (Makkar & Grisham, 2013). Similarly, the state of arousal per se that phobics experience when confronted with phobia-relevant stimuli might be interpreted as a sign of danger, and so evoke negative outcome expectations. In fact, negative outcomes lead to increased conditioned arousal responses (such as increased skin conductance responses, SCRs) in fear conditioning (e.g. Lissek et al., 2005). So, there should be a natural correlation between danger signals and arousal responses that could be made use of, when judging danger based on arousal. Particularly, danger signals induce arousal responses, because they predict arousing unconditioned stimuli. If another stimulus evokes arousal, it might be misinterpreted as a danger signal. Although reward signals are also associated with elevated arousal (e.g. Delgado, Gillis, & Phelps, 2008), the coupling between arousal and negative value of visual stimuli seems to be stronger than the coupling between arousal and positive value, at least for women (Lang & Bradley, 2007). Specifically, in women, there seems to be a stronger correlation between valence and arousal for negative stimuli than for positive stimuli (Lang & Bradley, 2007). This means, arousal should be a better predictor of negative than of positive outcomes, and a better predictor of negative outcomes in women than in men. In any case, the stimuli that usually lead to illusory correlations are either paired with neutral or negative outcomes (in the laboratory as well as in nature, at least

fear-relevant stimuli). So in this context, heightened arousal should be a stronger predictor of danger than reward. By this logic, participants might reason that their elevated arousal is a sign of danger and increased probability of negative consequences.

To date, four studies have been published that examined the effect of the arousal dimension on an emotional expectancy bias (Cavanagh & Davey, 2001) or covariation bias (De Jong & Merckelbach, 2000; Pauli et al., 2002; VanOyen Witvliet & Vrana, 2000). Interestingly, a mood manipulation (watching neutral, positive or negative video clips) induced generally inflated US expectancies for both the negative and the positive condition, irrespective of the fear-relevance of the stimuli (Cavanagh & Davey, 2001; experiment 1). This led to the assumption that the effect could have been mediated by non-specific arousal. However, a manipulation of arousal level by physical exercise did not change US expectancy (Cavanagh & Davey, 2001; experiment 2). By trend, US expectancy was even reduced under arousal manipulation. In another experiment (de Jong & Merckelbach, 2000), spider phobic patients were administered alprazolam before the beginning of a traditional illusory correlation paradigm (see chapter 2.3., original experimental paradigm). Alprazolam is an anxiolytic from the pharmacological class of benzodiazepines that successfully reduced skin conductance responses to spider images in phobic women in comparison to a placebo group in this study. Despite this attenuated arousal response, both the treatment and the placebo group developed similar fear-relevant illusory correlations. So far, it can be stated that manipulating non-specific arousal by physical exercise or pharmacological inhibition of fear responses does not affect a fear-relevant expectancy bias or illusory correlations.

However, two studies found that when arousal level was varied by emotional stimuli from trial to trial, normal non-fearful samples overestimated the proportion of startle probes in the presence of both positive and negative stimuli. Startle probes are aversive acoustic or highly salient visual stimuli that provoke a sudden blink reflex which is modulated depending on the current affective state of an individual (Bradley et al., 1993). The magnitude of the startle response is increased under negative emotional experience and decreased under positive emotional experience. While illusory correlation experiments mostly worked with negative or neutral stimuli, it has been found that covariation estimates between aversive startle sounds and positive pictures are higher than for neutral pictures, although still lower than for negative pictures (Pauli et al., 2002). Similarly, when emotional imagery is used as emotional background and visual startle probes instead of acoustic startle probes, the post experimental estimation of probe frequency is primarily determined by the arousal of the conditions (VanOyen Witvliet & Vrana, 2000). That is, high arousing positive imagery led to more frequent probe estimates than low arousing negative imagery. It is interesting to note that the extent of probe frequency estimates was paralleled by a very similar pattern in the magnitudes of blink responses. The authors concluded that arousal and the response to startle probes may be important determinants of covariation bias. However, these findings may not be so easily transferrable to fear-relevant covariation bias, because the visual startle probes that have been used in this study cannot really be considered aversive, as acknowledged by the authors (VanOyen Witvliet & Vrana, 2000). Although, covariation estimates were also enhanced for positive pictures when aversive startle probes were used, response magnitude and covariation estimates were not correlating (Pauli et al., 2002), thus confirming the arousal approach but challenging the response explanation.

Taken together, directly manipulating physiological arousal does not seem to affect covariation bias. In contrast, a covariation bias can also be found between positive arousing stimuli and aversive consequences. Nevertheless, the latter findings stem from the affect modulated startle paradigm which differs from the illusory correlation paradigm that has been used to study covariation bias in fear (e.g. Tomarken et al., 1989) in two important aspects. First, in the illusory correlation paradigm, three or four concrete semantic categories of slides are used while in the startle paradigm a broad array of picture categories is used. The categorization into positive, neutral and negative pictures in the startle paradigm is therefore probably more open to personal interpretation. Second, in the illusory correlation paradigm, participants are instructed to focus their attention to stimulus-outcome-relationships, while in the affect modulated startle paradigm participants are explicitly asked to ignore the aversive outcomes. If previously found fear-relevant illusory correlations are influenced by the arousal of stimuli, then an influence of arousal should also be observable when attention is focused on the stimulus-outcome-relationships. In order to investigate the impact of positive arousing stimuli on illusory correlations with aversive outcomes, an experiment was realized with three clear picture categories (positive arousing erotic couples, negative arousing mutilations and neutral non-arousing household articles). Each picture category was equally often associated with an aversive startle sound. Startle sounds have already been used in previous illusory correlation experiments (Amrhein et al., 2005; Mühlberger et al., 2006). Participants were instructed to pay attention to the relationship between pictures and outcomes and rate their outcome expectancy in every trial. It was hypothesized that both positive and negative arousing stimuli would induce an overestimation of aversive outcomes relative to neutral stimuli.

Besides the arousal hypothesis, the model for the maintenance of fear-relevant illusory correlations predicts that negative valence would have an independent impact on illusory correlations (chapter 2.5.). Reasons for this influence might be an affective matching between aversive pictures and aversive outcomes (Tomarken et al., 1995) and the emotional modulation of aversive outcomes by the negative pictures (e.g. Kenntner-Mabiala & Pauli, 2005). More precisely, the aversive startle sound should become subjectively more aversive in negative trials and be encoded preferentially. As a consequence, these outcomes should be retrieved more easily and thus they should be overestimated in following trials and after the experiment. In addition, highly aversive outcomes should be expected more than less aversive outcomes per se, because increased anticipatory attention and preparation should be necessary (see chapter 2.5. and experiment 4). In summary, covariation estimates should be highest for negative pictures, lowest for neutral pictures and mediate for positive pictures. Moreover, aversive outcomes should be subjectively more aversive when they appear with a negative picture. This aversiveness modulation should be correlated with covariation bias. Originally, the experiment was not designed to examine gender effects on illusory correlations. However, previous experiments on illusory correlations almost exclusively relied on female participants (see Table 9.1), and very little is known about gender effects. Therefore, gender was included as an additional factor to explore differences between men and women.

3.2. Methods

Participants

In total, 48 participants (35 women, 13 men) were recruited among psychology students of the University of Würzburg and by advertisements on a local website. To compensate their efforts, they received course credit or were paid ten Euros. From this initial sample 14 participants were discarded because of a programming error in the experimental session, the ratings were missing from two participants, one aborted the experiment because of the aversive sound and another one misunderstood the experimental task. After exclusion of those participants, the final sample contained 31 individuals (18 women, 13 men) between 19 and 44 years (*M* = 24.32, *SD* = 4.64). Detailed information about further sample characteristics can be obtained from Table 3.1. There were no significant differences between men and women in age, the proportion of psychology students, state and trait anxiety (Spielberger State Trait Anxiety Inventory, STAI; Spielberger, Gorsuch, & Lushene, 1970), anxiety sensitivity (Anxiety Sensitivity Index, ASI; Reiss, Peterson, Gursky, & McNally, 1986) and concentration maintenance (Konzentrationsverlaufstest, KVTM; Abels, 1974).

Stimuli

Twenty-four pictures from three categories (eight mutilations, eight erotic couples, eight household articles) [2] were chosen from the International Affective Picture System (IAPS, Lang, Bradley, & Cuthbert, 2005). The pictures were presented in the middle of the screen with a height of 768 pixels and a width of 1024 pixels.

The startle sound consisted of a 50 ms lasting and 105 dB loud burst of white noise with an instant rise time and was applied over headphones to the participants' ears. In 50 % of the trials, it occurred at the offset of the picture presentation.

	Age	**Psychology students**	**STAI(state)**	**STAI(trait)**	**ASI**	**KVT**
Men	*M* = 23.54	n = 4	*M* = 37.08	*M* = 36.23	*M* = 14.80	*M* = 96.86
	SD = 3.38		*SD* = 6.90	*SD* = 8.67	*SD* = 8.55	*SD* = 6.40
	N = 13	*N* = 13	*N* = 13	*N* = 13	*N* = 10	*N* = 12
Women	*M* = 24.89	n = 4	*M* = 35.44	*M* = 35.06	*M* = 15.76	*M* = 99.86
	SD = 5.40		*SD* = 8.67	*SD* = 7.56	*SD* = 10.03	*SD* = 7.94
	N = 18	*N* = 18	*N* = 18	*N* = 18	*N* = 17	*N* = 18
	p = .43	*p* =.69	*p* =.58	*p* =.70	*p* =.80	*p* =.28

Table 3.1. Means (M) and standard deviations (SD) of sample characteristics in men and women. STAI = Spielberger State Trait Anxiety Inventory; ASI = Anxiety Sensitivity Index; KVT = Konzentrationsverlaufstest (concentration maintenance test; standard value [time x errors]; high values ~ good concentration).

[2] The IAPS numbers of the chosen pictures were 3010, 3030, 3060, 3062, 3071, 3100, 3101, 3130 (mutilations), 4611, 4652, 4659, 4664, 4669, 4670, 4680, 4695 (erotic couples), 7006, 7009, 7010, 7025, 7059, 7150, 7175, 7235 (household articles: plate, cup, basket, stool, carabiner, umbrella, lamp, chair).

Procedure

When the participants arrived in the laboratory, they first signed informed consent and filled out a short questionnaire about demographic data as well as the state version of the STAI. After doing so, they sat down on a comfortable chair in a dimly lit chamber where SCR and EMG sensors were applied. Next, the participants read the instructions on a computer monitor about 1.5 m in front of them. They were asked to pay attention to the relationships between three picture categories (mutilations, erotic couples, household articles) and a loud and aversive sound that would occasionally follow the pictures. Before the illusory correlation experiment started, the aversive sound was presented once and rated by the participants for aversiveness. Then they were shown a random example picture of each category and asked to give an a priori expectancy rating (0 – 100%) on the computer of how often an aversive sound would follow a picture of the given category (i.e., a priori covariation estimates). The order of example picture presentation was randomized.

In the illusory correlation experiment itself, eight pictures of each category were presented three times, resulting in a total of 72 trials. Exactly twelve of the 24 pictures of each category were followed by an aversive startle tone, twelve were followed by nothing. Six consecutive trials always contained all possible combinations between picture categories and outcomes. So, the contingency between categories and the sound was always 50% in each of these twelve blocks. The picture duration was eight seconds, the inter-trial-interval was randomly varying between eight and twelve seconds.

Throughout the experiment a continuous rating scale (0 – 100%) was presented below the pictures. During picture presentation, the participants rated the estimated probability that an aversive sound would occur at the offset of the current picture (i.e., on-line expectancy ratings). This procedure was practiced for seven trials before the experimental session with letters instead of pictures.

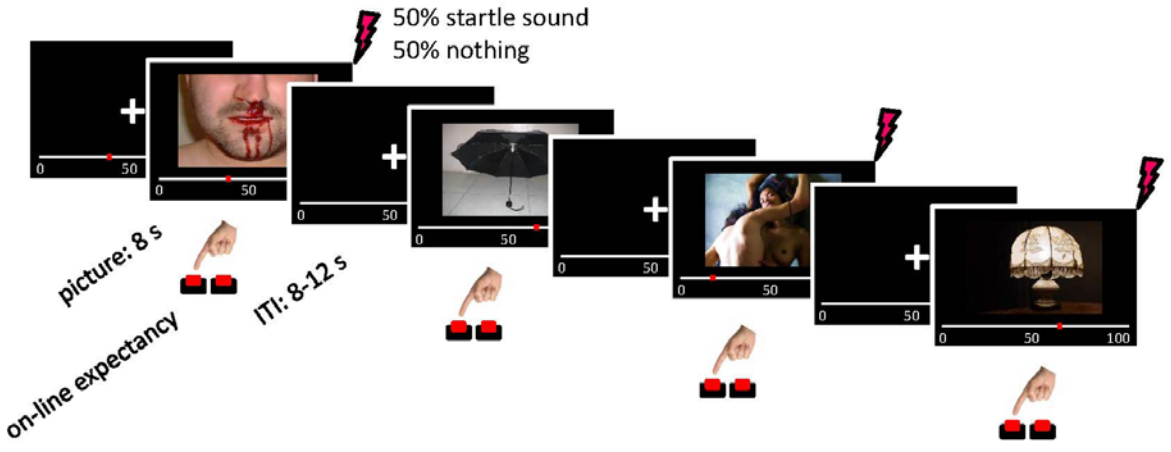

Figure 3.1. Illusory correlation paradigm. Three categories of pictures were presented (injuries, household articles, and erotic couples). Exactly 50% of the pictures of each category were followed by an aversive white noise startle sound. Participants were asked to rate the probability of a startle sound in the current trial on a continuous rating scale (0-100) that was navigated with two buttons. The displayed scale is only a representation. In the experiment, the scale was labeled in steps of 10. A variable inter-trial-interval (ITI) was used between picture presentations. The displayed pictures were found as licensed for free distribution and are only representations of the original IAPS pictures used in the study.

After the 72 trials of the experimental session, participants were asked to make the following estimations and ratings: the proportion of occurred sounds for each picture category (i.e., a posteriori covariation estimates; 0 – 100%), the aversiveness of the sound for each picture category (0 – 100), the overall proportion of pictures for each category (0 – 100%) and the overall proportion of sounds regardless of picture category (0 – 100%)[3]. In addition, all the pictures were presented again and rated for valence (1 very unpleasant – 9 very pleasant), arousal (1 not arousing at all – 9 very arousing), dangerousness (1 not dangerous at all – 9 very dangerous) and fear induction (1 not fear inducing at all – 9 very fear inducing). The pictures were presented simultaneously with the rating scales and the order of rating questions was randomized, but constant for every subject. At the end, the KVT was performed and the trait version of the STAI was filled out.

Apparatus and data analysis

The experimental presentation of pictures and sounds, as well as all ratings were realized with the software Presentation (Neurobehavioral Systems, Albany, CA, USA). The ratings were performed by moving a red cursor horizontally on a white rating scale with the help of two buttons (index and ring finger of the dominant hand). An additional third button (middle finger) served to confirm the rating with the exception of on-line covariation estimates. Here, no confirmation was necessary, because the cursor position in the moment of picture offset was always counted as the rating.

After the participants had washed their hands without soap, two surface electrodes (Ag/AgCl) were filled with isotonic electrode gel (TD-246, PAR Medizintechnik GmbH) and attached to the hypothenar eminence of the non-dominant hand for the acquisition of skin conductance responses (SCRs). Both SCRs and the electromyogram (EMG) were recorded with a V-Amp 16 amplifier and the software Vision Recorder (Brain Products, Munich, Germany). The raw data of skin conductance was down-sampled from 1000 Hz to 50 Hz and further analyzed using the Continuous Decomposition Analysis of the Matlab based software Ledalab V3.4.3. The signal was decomposed into a tonic (skin conductance level) and a phasic (SCR) driver component (Benedek & Kaernbach, 2010a, b). Therefore, the phasic driver is less biased by slow and stimulus-unrelated changes of tonic skin conductance, and served as SCR within a time window of 1 – 5 s after the onset, resp. offset of the picture. To adjust for the left-skewed distribution of SCRs, the data were logarithmized using the function ln(SCR +1).

The startle response to the aversive sounds was measured via an EMG of blink movements. To this end, two surface electrodes (Ag/AgCl) were filled with electrode crème (Parker Laboratories, Fairfield, NJ, USA) and attached directly below the left eye, onto the abraded and cleaned skin above the orbicularis oculi muscle. One electrode was positioned centrally below the pupil, the other one approximately 1 cm laterally. A reference electrode was attached behind the right ear, a ground electrode behind the left ear. The data were recorded at a sample rate of 1000 Hz with a V-Amp 16 amplifier and Vision Recorder software (Version 1.10; Brain Products,

[3] Phrasings of the questions: a priori covariation: e.g. “How often (in %) will this aversive sound follow MUTILATIONS?”; a posteriori covariation: “How often (in %) did the aversive sound follow MUTILATIONS?”; sound aversiveness: “How UNPLEASANT did you experience the sound following MUTILATIONS?”

Munich, Germany). Impedances were kept below 5 kΩ. Offline, the data were analyzed with Vision Analyzer software (Version 1.05; Brain Products, Munich, Germany). Signals from both electrodes were averaged and filtered with a 30 Hz low cut-off, a 500 Hz high cut-off and a 50 Hz notch filter. Then, the data were rectified and smoothed with a 50 ms moving average. The peak amplitude was determined between 20 and 200 ms after the startle probe onset and a baseline of 50 ms before startle probe onset was subtracted. Trials were discarded manually from the analysis if the baseline was not stable (+/- 5 mV) or blinks already occurred before 20 ms. Blink magnitudes were transformed into standard *T*-scores (*M* = 50; *SD* = 10) and averaged on an individual level for each emotion condition.

All data were analyzed with repeated-measures ANOVAs including gender as a between-subjects factor. Although the study was not designed to test gender effects, the sex of the participants turned out to have an influence when it was included on an exploratory basis. Therefore, all results are reported depending on the gender of the participants. Multivariate statistics are reported with partial η_p^2 as an effects size measure. Significant main effects or interactions are dissolved with follow-up *t*-tests (two-sided). Due to a priori hypotheses, Pearson correlation coefficients *r* (one-sided) were calculated to test the association between covariation bias and startle aversiveness.

3.3. Results

A priori covariation estimates

Only women but not men expected more aversive sounds following negative pictures than positive and neutral pictures (Figure 3.2.A). A repeated measures analysis of variances (ANOVA) with *category* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *category*, $F(2, 28) = 5.91$, $p < .01$, $\eta_p^2 = .30$, and an interaction effect of *Category x Gender*, $F(2, 28) = 5.14$, $p < .05$, $\eta_p^2 = .27$, and a marginal significant main effect of *gender*, $p = .07$. A one-factorial ANOVA for women only revealed a significant main effect of *category*, $F(2, 16) = 13.60$, $p < .001$, $\eta_p^2 = .63$. Follow up *t*-tests indicated that women gave higher a priori covariation estimates for negative ($M = 64.59$; $SD = 20.81$), than for positive ($M = 35.48$; $SD = 20.35$), $t(17) = 4.42$, $p < .001$, $d = 1.04$, and neutral pictures ($M = 29.02$; $SD = 16.94$), $t(17) = 4.91$, $p < .001$, $d = 1.16$. The difference between positive and neutral pictures was not significant, $p = .36$ [4]. In a one-factorial ANOVA for men, there was no significant main effect of category, $p = .95$.

[4] Eighteen women had been discarded, because they dropped out later in the experiment (see Methods). If those were included in this analysis, the interaction between men and women was still significant, $F(2, 45) = 6.07$, $p < .01$, $\eta_p^2 = .21$. Moreover, the difference between positive ($M = 36.34$; $SD = 19.19$) and neutral pictures ($M = 27.00$; $SD = 16.18$) turned out to be significant in women, $t(34) = 2.20$, $p < .05$, $d = .37$.

On-line expectancy ratings

During the illusory correlation experiment, the highest expectancy ratings were given for negative pictures followed by positive pictures, and lowest ratings for neutral pictures. In a repeated measures ANOVA with *category* and *block* as within-subjects factors and *gender* as a between-subjects factor, there was a significant main effect of *category*, $F(2, 28) = 5.25$, $p < .05$, $\eta_p^2 = .27$. Although the differences between categories were descriptively more pronounced in women (Figure 3.2.B), the interaction of *Category* x G*ender* was not significant, $p = .17$. Overall, the participants estimated the probability of sound occurrence higher in the presence of negative pictures ($M = 58.49$; $SD = 16.34$) than in the presence of both neutral ($M = 43.50$; $SD = 14.97$), $t(30) = 3.50$, $p < .01$, $d = .63$, and positive pictures ($M = 47.75$; $SD = 15.49$), $t(30) = 2.94$, $p < .01$, $d = .53$. In men and women combined, the difference between positive and neutral pictures was only marginal significant, $p = .08$ (two-sided).

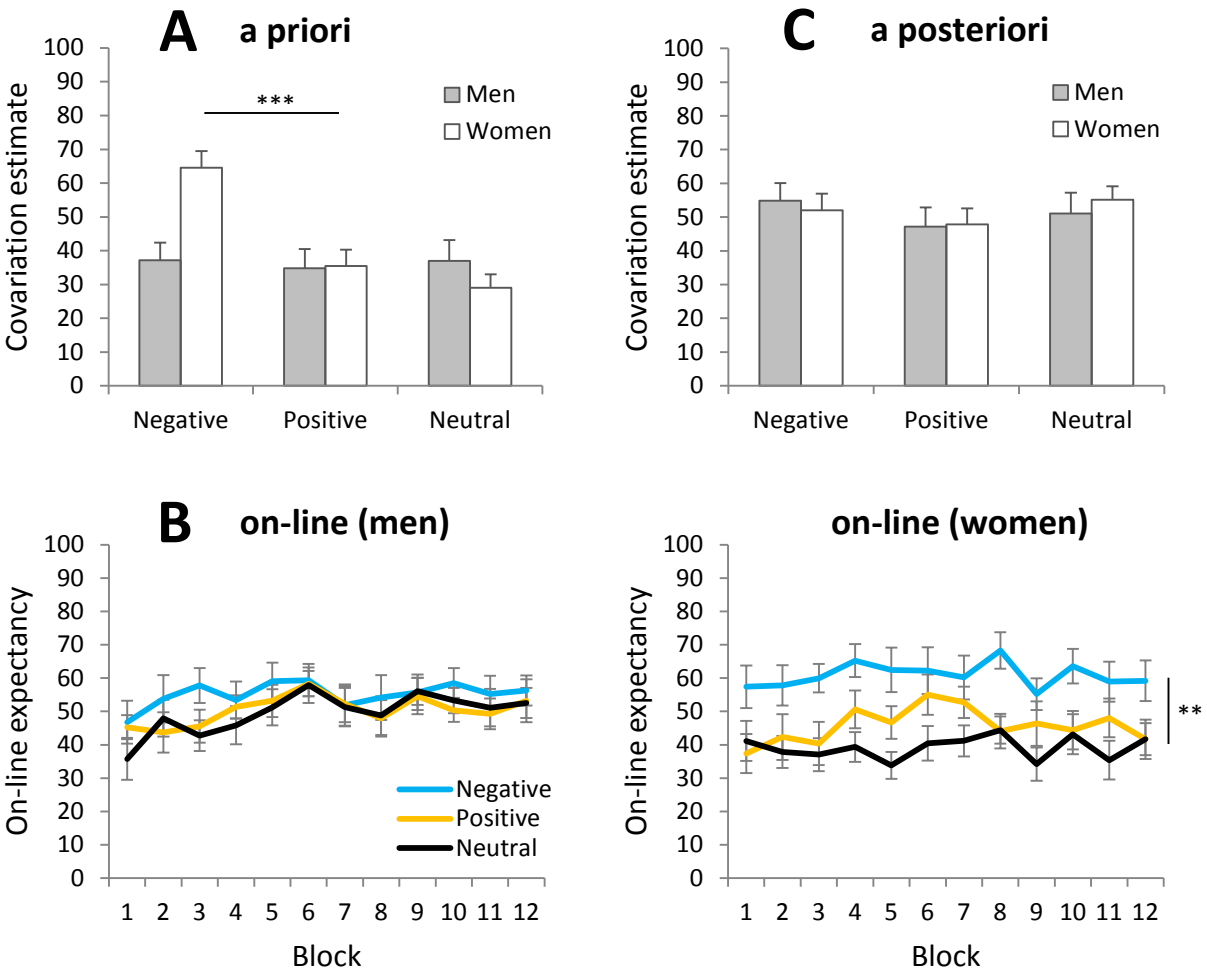

Figure 3.2. The diagrams show the expected proportion of shocks before (A a priori), the expected proportion of shocks during the experiment (B on-line) and the estimated proportion of shocks after the illusory correlation experiment (C a posteriori). Only women but not men displayed an a priori expectancy bias and marginally an on-line expectancy bias for negative images. After the experiment, no covariation bias was observed. ** p < .01; *** p < .001

A posteriori covariation estimates and sound aversiveness

After the illusory correlation session, neither men nor women displayed a significant covariation bias (see Figure 3.2.C). In a repeated measures ANOVA with *category* as a within-subjects factor and *gender* as a between-subjects factor, there were no significant main effects and no significant interaction (all *p*s > .38).

Similarly, there was no differential subjective aversiveness of the startle sounds following the different picture categories. The data of one man and one woman were missing because the aversiveness rating was only included later in the study. In a repeated-measures ANOVA with *category* as a within-subjects factor and *gender* as a between-subjects factor, there were no significant main effects and no significant interaction, all *p*s > .09.

Despite the absence of category effects on a posteriori covariation estimates and sound aversiveness ratings, the expected positive relationship between those dependent variables was found. Measures of covariation bias and aversiveness bias were calculated for negative and positive pictures relative to neutral pictures (i.e., covariation estimate for negative pictures *minus* covariation estimate for neutral pictures; sound aversiveness for negative pictures *minus* sound aversiveness for neutral pictures; covariation estimate for positive pictures *minus* covariation estimate for neutral pictures; sound aversiveness for positive pictures *minus* sound aversiveness for neutral pictures). For negative, $r = .35$, $p < .05$ (one-sided), but not positive pictures, $r = .03$, $p = .45$ (one-sided), the covariation bias correlated with the aversiveness bias (Figure 3.3).

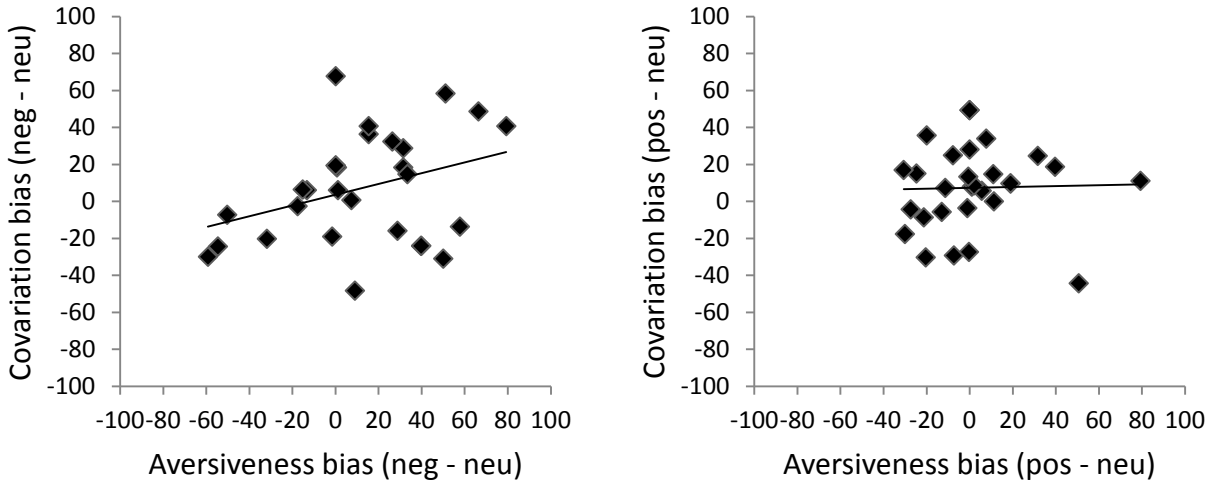

Figure 3.3. The scatterplots depict the relationships between the covariation bias and the modulation of sound aversiveness by picture content. While the covariation bias and the sound aversiveness were significantly correlated for negative images (left panel), there was no such relationship for positive images (right panel).

Startle responses

Startle responses were potentiated after negative pictures and inhibited after positive pictures relative to neutral pictures (see Figure 3.4). There was no influence of gender. Because of technical failure, two women and two men were missing in this analysis. In a repeated measures ANOVA, there was only a significant main

effect of *category*, $F(2, 24) = 7.25$, $p < .05$, $\eta_p^2 = .30$, but no interaction of *Category x Gender*, p = .83. Follow-up *t*-tests revealed that startle responses following negative pictures ($M = 51.59$; $SD = 2.66$) were significantly stronger than startle responses following neutral ($M = 49.94$; $SD = 1.77$), $t(26) = 2.20$, $p < .05$, $d = .43$, and positive pictures ($M = 48.53$; $SD = 2.26$), $t(26) = 3.45$, $p < .01$, $d = .66$. Startle amplitudes after positive pictures were lower than after neutral pictures, $t(26) = 2.34$, $p < .05$, $d = .45$.

Skin conductance responses

SCRs to outcomes and SCRs to picture onset were analyzed separately. Skin conductance responses were higher for sounds than for nothing as outcomes, but were not modulated by the category of the pictures (see Figure 3.4). For outcomes, a repeated-measures ANOVA with *category* and *outcome* as within-subject factors and *gender* as a between-subject factor revealed significant main effects of *outcome*, $F(1, 23) = 58.99$, $p < .001$, $\eta_p^2 = .72$. Neither *gender* nor *category* showed any significant impact on SCRs to the outcomes. In follow-up *t*-tests, SCRs to startle sounds ($M = .21$; $SD = .15$) were higher than to nothing ($M = .68$; $SD = .40$), $t(24) = 7.82$, $p < .001$, $d = 1.56$.

SCRs to picture onsets did not differ depending on picture category or gender. A repeated-measures ANOVA with *category* as a within-subject factor and *gender* as a between-subjects factor did not reveal any significant results, $p > .15$.

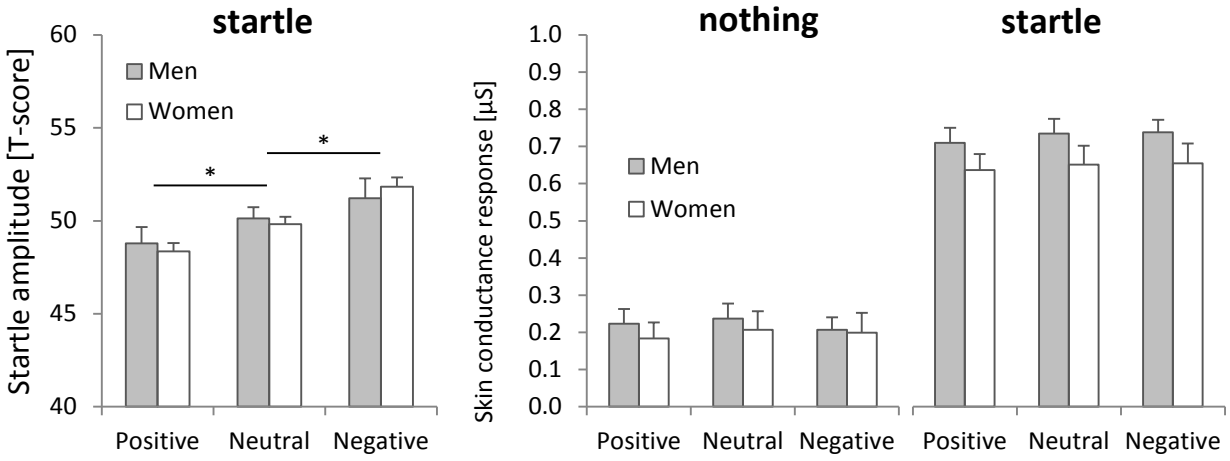

Figure 3.4. The histograms show startle and skin conductance responses (SCRs) to the outcomes depending on the preceding picture category. Startle amplitudes were measured to startle sounds, SCRs to both nothing and startle sounds. Picture valence modulated only startle amplitudes (left panel), but not SCRs (right panel). ** $p < .01$; *** $p < .001$

Picture ratings: valence, arousal, dangerousness, fear induction

The ratings of valence, arousal, dangerousness and fear induction of the pictures were analyzed in separate two-factorial repeated measures ANOVAs with *category* as a within-subjects factor and *gender* as a between-subjects factor (see Figure 3.5). The analysis of valence ratings revealed a significant main effect of *category*, $F(2, 28) = 39.67$, $p < .001$, $\eta_p^2 = .74$, and a marginal significant interaction of *Category x Gender*, $F(2, 28)$

= 3.18, p = .06, η_p^2 = .19, but no significant main effect of *gender*, p = .12. Despite the marginal significant interaction of *Category x Gender*, there were no significant differences between men and women for any picture category in follow-up *t*-tests, all *p*s > .16. Overall, positive pictures (M = 6.75; SD = 1.41) were rated as more pleasant than both neutral (M = 5.43; SD = 1.14), $t(30) = 4.87$, $p < .001$, d = .87, and negative pictures (M = 2.57; SD = 1.50), $t(30) = 9.29$, $p < .001$, d = 1.67. Negative pictures were rated as less pleasant than neutral pictures, $t(30) = 7.22$, $p < .001$, d = 1.29.

Likewise, the arousal ratings were characterized by a significant main effect of *category*, $F(2, 28) = 73.31$, $p < .001$, η_p^2 = .84, but no significant main effect of *gender* and no significant interaction (all *p*s > .18). Follow-up *t*-tests showed that negative (M = 6.15; SD = 2.02), $t(30) = 9.73$, $p < .001$, d = 1.29, and positive pictures (M = 6.03; SD = 1.44), $t(30) = 11.79$, $p < .001$, d = 1.29, were more arousing than neutral pictures (M = 2.23; SD = 1.30). There was no significant difference between positive and negative pictures, p = .75.

For dangerousness ratings, also a significant main effect of *category*, $F(2, 28) = 41.86$, $p < .001$, η_p^2 = .75, and a marginal significant interaction of *Category x Gender*, $F(2, 28) = 3.02$, $p < .07$, η_p^2 = .18, were found, but no main effect of *gender*, p = .87. Despite the marginal significant interaction, no differences were found between men and women in any of the emotion categories. Overall, negative pictures (M = 6.28; SD = 2.15) were rated as more dangerous than neutral (M = 2.02; SD = 1.45), $t(30) = 9.64$, $p < .001$, d = 1.29, or positive ratings (M = 2.00; SD = 1.64), $t(30) = 9.03$, $p < .001$, d = 1.73. No difference was found between positive and neutral pictures, p = .89.

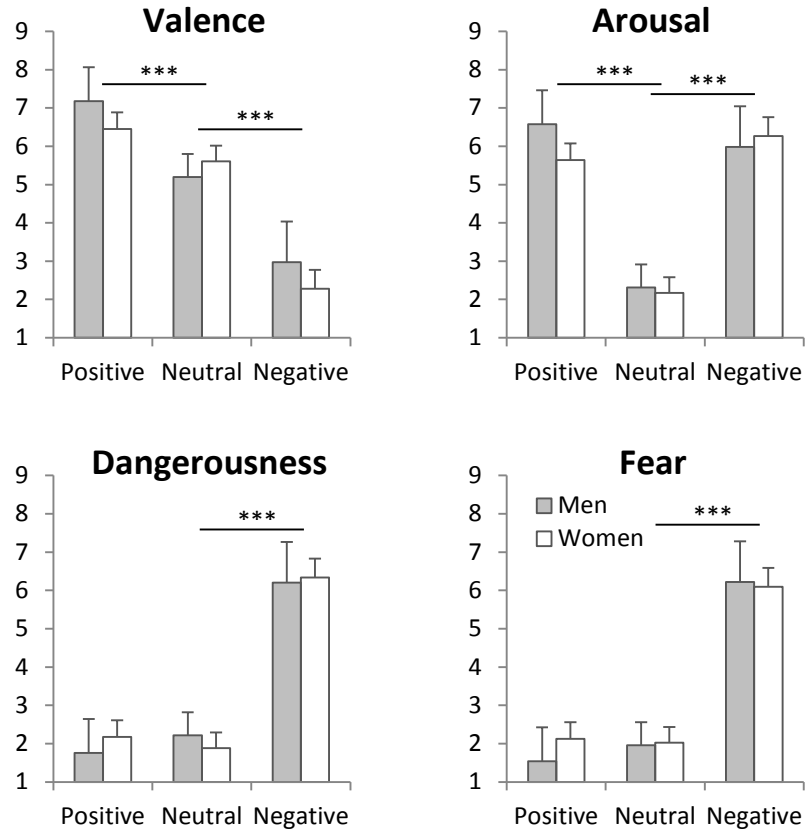

Figure 3.5. The histograms show ratings. High valance ratings indicate pleasantness, while low valence ratings indicate unpleasantness. ** p < .01; *** p < .001

Fear induction was also characterized by a significant main effect of *category*, $F(2, 28) = 44.89$, $p < .001$, $\eta_p^2 = .76$, but no main effect of *gender* and no interaction was found (all *p*s > .22). Follow-up *t*-tests indicated that negative pictures ($M = 6.15$; $SD = 2.14$) were more fear inducing than neutral ($M = 1.99$; $SD = 1.49$), $t(30) = 9.48$, $p < .001$, $d = 1.70$, or positive pictures ($M = 1.88$; $SD = 1.61$), $t(30) = 9.69$, $p < .001$, $d = 1.74$. No difference between positive and neutral pictures was found, $p = .46$.

3.4. Discussion

The first goal of experiment 1 was to find out whether high arousing positive pictures would lead to an over-association with aversive outcomes in an illusory correlation paradigm, that is when participants are instructed to pay attention to category-outcome-relationships, which was not the case in previous investigations. The assumption was confirmed partially with several restrictions. A priori, women but not men expected more aversive startle sounds after positive arousing slides than after neutral slides. This was only true, when all participants were included in the analysis of a priori expectancy ratings (inclusion of also those that were excluded later in the experiment). During the experiment, on-line expectancy ratings indicated only a marginal significant trend of higher expectancy during positive arousing trials than during neutral trials while an expectancy bias for negative pictures was still present. Finally, after the experiment, no a posteriori covariation bias was observed; neither for positive, nor for negative pictures.

The finding of an a priori expectancy bias for negative outcomes following erotic slides relative to neutral slides replicates an earlier finding (Wiedemann, Pauli, & Dengler, 2001). In this previous study, more aversive shocks were expected after erotic scenes than after mushrooms by panic patients and healthy controls. Although Wiedemann et al. (2001) did not report an effect size for this particular contrast, the mean differences suggest that the effect was more pronounced than in the present experiment. Possibly household objects are more likely to be associated with an aversive startle sound than mushrooms with electrical shocks on a semantic level. Indeed, covariations are often overestimated, if there is a strong semantic linkage between stimuli (e.g. Chapman, 1967). For example, an umbrella or a lamp (as in the pictures here) can actually make a sound (as a startle sound) while mushrooms do not produce shocks. As a consequence, the difference between erotic and neutral slides might have become smaller in the present experiment. This also exemplifies a general problem of testing the effect of arousal with emotional stimuli, which is that the arousal level is practically always confounded with the semantic content of the stimulus. Specifically, one cannot be sure whether people expect more aversive stimuli following erotic pictures merely because of heightened arousal or because their content is more easily associated with a sound or a touch. However, what can be concluded from the present and a previous study (Wiedemann et al., 2001) is that stimulus valence is not the only determiner of negative outcome expectancy because positive stimuli can also be associated with negative consequences, although probably to a somewhat lesser degree. The magnitude of such an expectancy bias might rely on inter-individual differences and may be relevant for psychological disorders. For example, patients with major depression may be more likely to expect something negative in positive contexts like a party or an exciting leisure activity. This could further explain why pleasant activities seem to be avoided in depression (Lewinsohn & Graf, 1973).

On-line, the expectancy of an aversive sound was only by trend enhanced in positive trials relative to neutral trials in men and women combined. This cannot be interpreted as a vanishing of an a priori expectancy bias because the a priori bias for positive stimuli was only present in the larger sample including all drop-outs. If anything, the trend of an on-line bias emerged during the experiment because the same sample did not display an a priori expectancy bias. It may be speculated that emotional arousal leads to impairment of learning associations between positive pictures and nothing as an outcome. In fact, associative memory during the perception of emotional arousing stimuli is impaired (e.g. Mather et al., 2006; Touryan, Marian, & Shimamura, 2007). For example, associations between central and peripheral items are remembered worse when one item is negatively arousing than when both items are neutral (Touryan et al., 2007). In the illusory correlation paradigm, this effect might lead to greater impairment of the memory of the association between nothing and positive pictures in comparison to the memory of the association between startle sounds and positive pictures, because *nothing* is less salient and might therefore rely on attentional resources to a greater extent.

In future experiments, it would be interesting to have only one positive arousing condition without a negative arousing condition to see whether the effect size for the positive condition would increase. It may be possible that the presence of a negative and arousing condition leads to diminished expectancy ratings for the positive condition. However, at present, one has to assume that if emotional arousal per se affects illusory correlations with aversive outcomes, the effect is probably relatively small and clearly the valence or non-arousal-specific effects have a more eminent impact on these illusory correlations. This can probably be concluded on the basis of the present results because arousal ratings were well matched between positive and negative pictures. So, differences between the positive and the negative condition cannot be attributed to arousal effects. In addition, there were no differences in SCRs to pictures which might also be due to the equal probability of an aversive outcome (Bradley, Moulder, & Lang, 2005). In fact, it has been shown that pleasant pictures that signal threat of shock can induce the same defensive physiological responses as negative pictures (Bradley et al., 2005). This might also explain the absence of differences between SCRs to neutral and emotional pictures. However, the startle reflex clearly discriminated between positive, neutral and negative pictures, indicating that the motivational state was still appetitive in positive trials and defensive in negative trials (Bradley et al., 1993).

A priori and on-line expectancy ratings of the present results indicate an expectancy bias for mutilation slides in a non-selected sample regarding fear of blood and injury. So far, these findings are in line with an earlier study that found a posteriori covariation biases in individuals both high and low in fear of blood-injury (Pury & Mineka, 1997). However, in the present study, no a posteriori bias was found after the experiment, which is in contrast to Pury and Mineka (1997) and Pauli et al. (2002), who also found an a posteriori covariation bias between generally negative pictures and aversive outcomes in normal participants. One possible reason for the failure to find an a posteriori covariation bias here is that on-line expectancy was rated in every trial, while other investigators asked participants to attend to stimulus-outcome-relationships at the beginning of the experiment without on-line expectancy ratings (Pury & Mineka, 1997) or to ignore aversive outcomes (Pauli et al., 2002). One might argue that on-line expectancy ratings forces participants to really pay attention to the task throughout the experiment, which may lead to more accurate contingency estimates in the end. However, previous experiments

found that the usage of on-line expectancy ratings does not affect a posteriori covariation estimates (Grupe & Nitschke, 2011) or can even enhance illusory correlations between salient stimuli (Arkes & Harkness, 1983, as cited in de Jong, Merckelbach, & Arntz, 1995). Still, there are some methodological differences between these and the current study and it cannot be excluded that on-line estimates may have had a dampening influence on a posteriori covariation estimates in the present experiment. For instance, Grupe and Nitschke (2011) were using affective pictures as outcomes, which are a more complex and perhaps more distracting feedback than startle sounds and nothing. Consequently, the proportion between negative and neutral outcomes might have been more difficult to monitor in that previous experiment, and was overestimated regardless of on-line estimations. Another feature of the present experiment that might have influenced a posteriori covariation estimates is that only two outcomes were used (nothing, aversive sound) while Pury and Mineka (1997) used three outcomes (nothing, aversive shock, tone). Using more outcomes leads to fewer associations between pictures of mutilations and aversive events. This reduced amount of critical associations might not have been sufficient to recognize that there were no differences in aversive outcomes between conditions. Finally, the modality of the aversive outcome might have had an influence on contingency estimates. Since both electrical shocks and pictures of mutilations constitute harmful events to human skin, the semantic belongingness of stimuli and outcomes might have been stronger in the previous series of experiments (Pury & Mineka, 1997) than in the present one.

An unexpected finding of the current study was that the a priori expectancy bias for aversive outcomes to follow negative (and positive) pictures was only present in women, but not at all in men. It must be admitted that the experiment was not designed to study gender differences and the proportion of men (N = 13) and women (N = 18) was not balanced. Although the sample size was very low for men, it was a striking finding that expectancy estimates did almost not even differ descriptively in men (d = 0.01), while women exhibited a very large expectancy bias (d = 1.16). Further evidence will be needed to exclude that the gender effect was a chance finding. Since it was attempted to replicate this finding in experiment 2, potential mechanisms and implications of a gender effect are discussed in the discussion section of experiment 2.

Finally and very importantly, the present results confirmed the hypothesis that covariation bias and subjective outcome aversiveness were positively correlated. This means that those participants to whom the aversive startle sound was more aversive when it followed the negative pictures, were also more likely to overestimate the contingency between negative pictures and aversive outcomes. This is in accordance with the idea that illusory correlations can be partially explained by the emotional modulation of outcome aversiveness. This means, aversive events in the context of negative pictures are perceived as even more aversive than in the context of neutral pictures. This increased aversiveness should be associated with an enhanced encoding process. In addition, stimuli of increased aversiveness should trigger higher expectancies due to a general negativity bias (Vaish et al., 2008). Similarly, Grupe and Nitschke (2011) found that the unpleasantness of negative pictures overall predicted the overestimation of negative pictures associated with uncertainty cues. Moreover, Tomarken et al. (1989) found that the painfulness of the US after fear-relevant stimuli predicted the covariation bias in the control group. The present results for the first time replicate that this specific emotional modulation predicts covariation bias, and generalizes it to pictures of mutilations and aversive startle sounds.

In sum, experiment 1 revealed that positive arousing stimuli induced a small a priori expectancy bias for negative events, but no significant on-line or a posteriori covariation bias. On an exploratory level, it was found that only women displayed an a priori expectancy bias for negative outcomes after negative pictures. The effect was relatively large and might reflect very important gender differences relevant for psychopathology. Future studies are needed to replicate the effect, to examine underlying psychological mediators and to make judgments about clinical implications. As predicted by the proposed model for the maintenance of illusory correlations (chapter 2.5), there was a significant relationship between the a posteriori illusory correlation and the emotional priming of outcome aversiveness. Experiments 3 and 4 of the present thesis will provide evidence on the questions whether this relationship is manifest in biased neural processing of the outcomes, and whether the increased aversiveness may have a causal impact on illusory correlations.

4. Experiment 2: The influence of different basic emotions on illusory correlations

4.1. Introduction

Previous studies showed that illusory correlations between emotional stimuli and aversive consequences can emerge not only if the emotional stimuli are negative, but also if they are positive and arousing (VanOyen Witvliet & Vrana, 2000; Pauli et al., 2002). In experiment 1, women expected more negative consequences after positive than after neutral slides. During the experiment, there was only a trend that aversive outcomes were more expected during positive trials and there was no a posteriori covariation bias. Taken together, the effect of positive arousing stimuli on covariation bias could not have been clearly replicated in a case where participants were instructed to focus attention on stimulus-outcome-associations. In experiment 2, it was examined whether a covariation bias for positive arousing stimuli could be replicated in an affect-modulated startle paradigm where attention is not explicitly focused on stimulus-outcome-associations. Moreover, there were two results in experiment 1, which might become clearer if replicated in experiment 2.

First, there was a positive correlation between the covariation bias for negative pictures and the aversiveness of startle sounds. This is consistent with the idea that the increased aversiveness of negative outcomes after negative pictures might be in part responsible for the phenomenon of illusory correlations. However, the evidence for this idea would be strengthened if it could be replicated in a larger sample. Moreover, a bigger sample size would allow for the conduction of a multiple regression analysis to control for the influence of picture valence on covariation bias. Particularly, it might be possible that the relationship between startle aversiveness and covariation bias is only a side-effect of a relationship between picture valence and the covariation bias. That is, the more unpleasant mutilation pictures are to individuals the more likely do they associate them with an unpleasant outcome. The increased aversiveness of the startle sound could then be no more than a consequence of the increased unpleasantness of the mutilation pictures. However, according to the model of illusory correlation maintenance (chapter 2.5), the startle aversiveness as such should have an influence on covariation bias. For example, even if someone does not find mutilation pictures very unpleasant, he or she should develop an illusory correlation if he or she finds the startle sound after mutilation pictures very unpleasant. This could be found out by running a multiple regression analysis in which picture valence is included in the model before startle aversiveness is included in the model (Bortz & Döring, 2009). Then, it could be tested whether startle aversiveness explains a unique portion of variance of the covariation bias in addition to picture valence.

The second finding of experiment 1 that should be replicated in experiment 2 was that evidence of an illusory correlation was only found in women, but not in men. Particularly, only women but not men displayed an a priori expectancy bias for negative pictures. This finding was especially interesting, because many

covariation bias studies have been conducted exclusively with female phobics (see Table 9.1). If only women were susceptible to fear-relevant illusory correlations, then many previous findings could not be generalized to men. Moreover, an increased susceptibility to illusory correlations between emotional and aversive stimuli (or even cognitive biases in general) might help to explain the heightened prevalence of anxiety disorders in women (Jacobi et al., 2014). Although experiment 2 was, just like experiment 1, not a priori designed to investigate gender differences, experiment 2 offered a chance to explore whether similar tendencies as in experiment 1 could be found.

In experiment 2, an affect-modulated startle paradigm was realized to examine the emotional reactivity in bitter-sensitive individuals, so-called supertasters (Müller & Macht, 2007). To this end, supertasters' startle responses were measured during the presentation of positive, negative and neutral pictures. These results are not reported here, but since the frequency of startle sounds was equal across emotion conditions, it could be tested whether a covariation bias occurred after the experiment (similar to Pauli et al., 2002). Therefore, participants were asked how often a startle sound had appeared and how aversive the startle sound was depending on picture type – just like in experiment 1. Also similar to experiment 1, startle reflex served as an indicator of defensive-appetitive affect during picture presentation (Bradley et al., 1993), while pupil dilation was recorded here to investigate participants' emotional arousal (Bradley, Miccoli, Escrig, & Lang, 2008). In contrast to most affect-modulated startle experiments (e.g. Bradley et al., 1993) the emotion induction was not realized in a two-dimensional manner depending on valence and arousal (i.e. positive, negative and neutral pictures), but it was attempted to evoke discrete basic emotions (i.e. fear, disgust, anger, happiness). Previously, it has been shown that discrete basic emotions can be induced by emotional pictures. For example, mutilations or rotten food mostly seem to evoke disgust, while interpersonal violence or threatening animals are more likely to induce fear (Caseras et al., 2006; Lang, Greenwald, Bradley, & Hamm, 1993; Stark et al., 2007). These picture categories might also be associated with different startle magnitude patterns depending on personality characteristics like behavioral inhibition (Caseras et al., 2006). While happiness can also be induced by presenting pictures of smiling babies for example (Lang et al., 1993), anger seems to be more difficult to induce with pictures (Lang et al., 1993; Mikels et al., 2005). However, it has been found that individuals high in trait-anger respond with increased approach-related frontal alpha-asymmetry to anger-inducing pictures with racist content for example (Harmon-Jones, 2007).

Studying covariation bias between distinct emotions and aversive outcomes in the present experiment was a novel approach in the research of illusory correlations and allowed for the investigation of fear-relevant illusory correlations in a not specifically fearful population. Most previous studies examined covariation bias between fear-relevant stimuli and aversive consequences in high fearful individuals. If generally threatening visual stimuli were able to induce a covariation bias in a normal population, fear-relevant illusory correlations could be studied without costly recruitment of patients. Moreover, it is not known whether other negative emotions such as anger and disgust are equally likely as fear to induce illusory correlations with aversive harm-relevant consequences.

However, the major goals of experiment 2 were to test if positive arousing stimuli would, besides negative stimuli, also lead to a covariation bias with negative outcomes, to find additional evidence for the

assumption that the perceived aversiveness of the aversive outcome determines covariation bias, and to see if gender was again a predictor of covariation bias.

4.2. Methods [5]

Participants

Originally, 77 participants took part in the experiment, which was, in the first place, designed to test the emotional reactivity in high and low bitter-sensitive individuals. Exclusion criteria were psychological and neurological disorders, acute illness that could have affected the ability to taste, hearing disabilities and pregnancy. From the 77 recruited participants, one aborted the experiment, one could not be classified in bitter-sensitivity, technical problems occurred in five participants, three scored high on depression (Beck Depression Inventory, BDI-II ≥ 20; Beck, Steer, Ball, & Ranieri, 1996) and eleven already had experience with an illusory correlation experiment. So, 21 participants were discarded from the original sample and therefore 56 participants (38 women, 18 men) between 18 and 36 years (M = 24.77, SD = 3.91) were remaining for the present analysis. Men were on average 1.9 years older than women (p = .09) and in a slightly more anxious state than women according to the Spielberger State Trait Anxiety Inventory (STAI, p = .03; Spielberger et al., 1970). They did not differ in the STAI (trait), the BDI-II and the Positive and Negative Affect Schedule (PANAS, p > .20; Watson, Clark, & Tellegen, 1988). Details about questionnaire data are available in Table 4.1.

	Age	Psychology students	STAI(state)	STAI(trait)	BDI-II	PANAS +	PANAS -
Men	M = 26.06	n = 1	M = 38.33	M = 37.33	M = 5.50	M = 2.91	M = 1.29
	SD = 3.73		SD = 8.61	SD = 8.97	SD = 5.93	SD = 0.52	SD = 0.29
	N = 18	N = 18	N = 18	N = 18	N = 18	N = 18	N = 18
Women	M = 24.16	n = 9	M = 33.11	M = 35.16	M = 5.40	M = 3.01	M = 1.21
	SD = 3.89		SD = 5.76	SD = 7.67	SD = 4.55	SD = 0.57	SD = 0.22
	N = 38	N = 38	N = 38	N = 38	N = 38	N = 37	N = 37
	p = .09	p =.14	p =.03	p =.35	p =.94	p =.53	p =.26

Table 4.1. Means (M) and standard deviations (SD) of sample characteristics in men and women. No significant differences were found between men and women in the number of psychology students, trait anxiety, positive and negative affect. Men were by trend older and in a more anxious state than women.

[5] Experiment 2 was originally designed to test differences in emotional reactivity between high and low bitter-sensitive individuals. Here, the reported methods and results are focused on covariation estimates between emotional stimuli and aversive startle sounds. A complete description of all methods and results including those relevant for the influence of bitter-sensitivity is available in two unpublished diploma theses at the University of Würzburg (Vogt, 2012; Brenner, 2012).

Stimuli

On the basis of a pilot study, in which 135 emotional and neutral pictures from the IAPS (Lang et al., 2005) had been rated by 18 participants, 95 pictures [6] were chosen for the manipulation of basic emotions in the present study. Each picture had been rated on a scale from 1 to 9 regarding how intensely four basic emotions were induced by it (happiness, fear, disgust, anger). In at least 19 pictures for each category, the target emotion was rated significantly higher than the other emotions, and so they were chosen for the main experiment. In addition, 19 neutral pictures, which did not induce a specific emotion, were chosen as control stimuli (neutral category). Pictures inducing happiness showed for example smiling babies, kissing couples, delicious food, cute animals or beautiful landscapes. Pictures inducing fear showed for example tornados, threatening predators or weapons directed to the viewer. Pictures inducing disgust showed for example mutilated bodies and body parts, excrements, dead animals or vermins. Pictures inducing anger showed for example air pollution, violence or sexual harassment against women. Neutral pictures included pictures of household articles and mushrooms. The pictures were presented in the mid of a computer screen (1024 x 1280 pixels) with a height of 921 pixels and a width of 691 pixels. Pictures were presented in color, the background throughout the experiment was grey (RGB values: R = 130, G = 130, B = 130). A randomly scrambled version of every picture was created with an algorithm in Matlab (The Mathworks, Natick, Massachusetts, USA). The scrambled pictures contained the same pixels as the original pictures and were used to control for luminance effects on pupil dilation (Beatty & Lucero-Wagoner, 2000). That is, a scrambled picture served as a baseline to allow for the pupil's adjustment to the luminance of the picture, and so to control for the effects of physical picture characteristics on pupil dilation.

The startle sound was a 50 ms lasting and 95 dB loud burst of white noise with an instant rise time and was applied over headphones to the participants' ears. They were presented during 12 of 19 pictures of each category.

Procedure

When the participants arrived in the laboratory, they first signed informed consent and filled out a short questionnaire about demographic data as well as the PANAS. Then EMG electrodes were attached. They sat down on a chair in front of the eye-tracking column through which they could view the visual stimuli on the monitor. They put on the headphones, laid their chin on a flexible chin rest and leaned their forehead against the eye-tracking column. The eye-tracker was calibrated before the experiment started.

The affect-modulated startle experiment was divided into two blocks with each containing all 95 emotional pictures in 95 trials. Every trial started with a white fixation cross for 2 s, followed by a scrambled

[6] The IAPS numbers of the chosen pictures were 1440, 1710, 1811, 2071, 2080, 4611, 4658, 4659, 4660, 4680, 4689, 5623, 5626, 5830, 7200, 7330, 7350, 8170, 8190 (happiness); 1050, 1052, 1120, 1301, 1302, 1525, 1930, 1931, 1932, 5970, 5972, 5973, 6230, 6250, 6260, 6370, 6510, 6550, 9611 (fear); 1274, 2730, 3000, 3010, 3030, 3060, 3062, 3063, 3100, 3130, 3150, 3170, 3400, 7380, 9042, 9140, 9300, 9405, 9570 (disgust); 2688, 2691, 2751, 4621, 6212, 6315, 6360, 6530, 6900, 9090, 9110, 9180, 9252, 9280, 9341, 9409, 9520, 9560, 9800 (anger); 5500, 5510, 5520, 5530, 7002, 7004, 7006, 7009, 7010, 7025, 7034, 7035, 7050, 7080, 7090, 7100, 7175, 7235, 7700 (neutral).

version of the to-be-presented picture for 3 s, and the actual emotional or neutral picture for 5 s. The inter-trial-interval varied randomly between 2 and 2.5 s. In both blocks, 12 random pictures of each category were combined with a startle probe. Of these 12 startle probes, 6 occurred in an early time interval (250 – 300 ms after picture onset) and 6 in a late time interval (3000 – 4500 ms after picture onset). The participants had a short break between the two blocks. Before the second block, the eye-tracker was calibrated again.

After the two experimental blocks, the participants were shown one of three subsets of the emotional pictures, and rated valence, arousal and the intensity of the target emotion. Like in experiment 1, ratings were given on scales from 1 to 9. Valence ratings were asked by "How pleasant or unpleasant is this picture to you?" (1 – very unpleasant; 9 – very pleasant), arousal ratings by "How arousing is this picture to you?" (1 – not arousing at all; 9 – very arousing) and target emotions by "How much pleasure (fear, anger, disgust) does this picture trigger in you?" (1 – no pleasure [fear, anger, disgust]; 9 – extreme pleasure [fear, anger, disgust]). Only the target emotion and no other emotion was prompted for each picture. The participants rated only one third of the pictures (one of three predefined subsets containing an equal number of pictures of each category). This was done to save time because the experimental session lasted relatively long in total (about 2 hours) due to filling out questionnaires and testing the participants' bitter-sensitivity.

Like in experiment 1, the participants also did the following estimations: the proportion of occurred sounds for each picture category (i.e., a posteriori covariation estimates; 0 – 100%), the aversiveness of the sound for each picture category (0 – 100), the overall proportion of pictures for each category (0 – 100%) and the overall proportion of sounds regardless of picture category (0 – 100%). The actual proportion of startle probes was 12/19 (~63%) for each of the five picture categories. These general ratings for the different picture categories were assessed after the target emotion was rated for the pictures to ensure that the participants had the opportunity to learn which pictures were supposed to induce the various emotions. After this, bitter-sensitivity was determined by tasting and rating differently bitter solutions of NaCl and PROP (see Macht & Müller, 2007). Finally, the participants completed the STAI (state), the STAI (trait), the BDI-II and an additional set of questionnaires that was not relevant for the current research question.

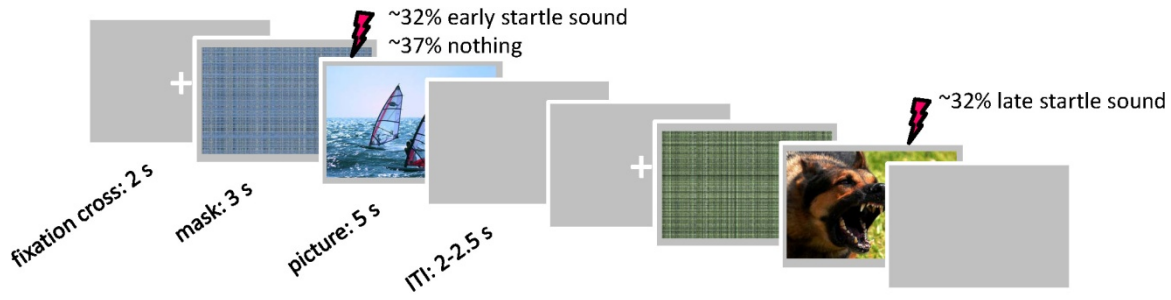

Figure 4.1. Illusory correlation paradigm. Five categories of pictures were presented to evoke different emotions (happiness, fear, anger, disgust, neutral). Approximately 63% (12/19) of the pictures of each emotion category were accompanied by an aversive white noise startle sound. A scrambled version of every picture was presented before the actual picture presentation to correct for differences in the illuminance of the pictures. The displayed pictures were found as licensed for free distribution and are only representations of the original IAPS pictures used in the study.

Apparatus and data analysis

The experimental presentation of pictures and sounds, as well as all ratings were realized with the software Presentation (Version 14.8, Neurobehavioral Systems, Albany, CA, USA). Rating data were collected with the help of scales on which participants could move a red cursor with a computer mouse. Estimates were confirmed by clicking the left mouse button.

The electromyogram (EMG) of the startle blink response was measured with two surface electrodes (Ag/AgCl) below the left eye on the orbicularis oculi muscle. One electrode was placed centrally below the pupil, another electrode was placed approximately 1 cm laterally. Before the electrodes were attached, the skin was abraded and cleaned, and the electrodes were filled with electrode gel (Parker Laboratories, Fairfield, New Jersey, USA). A ground electrode was attached to the left clavicle. The signal was amplified and recorded at 1024 Hz with a Varioport system (Becker Meditec, Karlsruhe, Germany). Offline, the data were analyzed with Vision Analyzer software (Version 1.05; Brain Products, Munich, Germany). Signals from both electrodes were averaged and filtered with a 28 Hz low cut-off, a 500 Hz high cut-off and a 50 Hz notch filter. Then, the data were rectified and smoothed with a 50 ms moving average. The peak amplitude was determined between 20 and 200 ms after the startle probe onset and a baseline of 50 ms before startle probe onset was subtracted. Startle amplitudes were transformed to standard *T*-scores (M = 50; SD = 10) and averaged on an individual level for each emotion condition.

Pupil dilation was registered at a sample rate of 240 Hz with an iView X Hi-Speed System (SensoMotoric Instruments, Berlin, Germany). It is a stationary eye-tracking column that illuminates the eye with an infrared diode and records the pupil size with an infrared-sensitive camera. An integrated chin and forehead rest stabilized the participant's head in order to minimize movements for accurate measurement. Ceiling light was dimmed to enable concentration on visual stimuli. The system was calibrated for each participant individually via a 13 point calibration procedure before the experiment started. Pupil diameter was processed by first averaging the horizontal and the vertical pupil diameters. Then, blinks were substituted by the mean from the 20^{th} (~83 ms) to the 10^{th} (~42 ms) sample before a blink. A scrambled version of each emotional picture was presented before picture onset to allow the pupil to adjust to the individual luminance of the picture. In order to correct for the effects of luminance on pupil dilation, a baseline of 0.5 s before picture onset was subtracted from the pupil diameter of five 1-s-intervals during the picture viewing period.

All data were analyzed with repeated-measures ANOVAs including gender as a between-subjects factor and the emotion induced by the pictures as a within-subjects factor (pleasure, fear, disgust, anger, neutral). Like experiment 1 experiment 2 was not designed to test gender effects. So, the number of participants was not balanced between men (N = 18) and women (N = 38). Nevertheless, due to a significant gender effect in experiment 1, all results are reported depending on the gender of the participants again. Multivariate statistics are reported with partial η_p^2 as an effects size measure. Significant main effects or interactions are dissolved with follow-up *t*-tests (two-sided). Due to a priori hypotheses, Pearson correlation coefficients *r* (one-sided) were calculated to test the association between covariation bias and startle aversiveness.

4.3. Results

A posteriori covariation estimates

In contrast to the hypothesis that also positive arousing stimuli should lead to a covariation bias, only negative emotions caused higher covariation estimates than the neutral condition. As in experiment 1, this effect only held true for women (Figure 4.2). A repeated measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 4.57$, $p < .01$, $\eta_p^2 = .26$, and a marginal significant interaction effect of *Emotion x Gender*, $F(4, 51) = 2.27$, $p = .08$, $\eta_p^2 = .15$. Based on the previous finding of a gender effect on expectancy ratings in experiment 1 and this marginal significant interaction, the emotion effect was analyzed separately for men and women. In a one-factorial repeated measures ANOVA for women, there was a highly significant *emotion* effect, $F(4, 34) = 8.52$, $p < .001$, $\eta_p^2 = .50$. Covariation estimates for fear ($M = 47.41$; $SD = 19.28$), $t(37) = 5.37$, $p < .001$, $d = 0.87$, disgust ($M = 50.70$; $SD = 22.88$), $t(37) = 5.55$, $p < .001$, $d = 0.90$, and anger pictures ($M = 41.40$; $SD = 17.33$), $t(37) = 3.86$, $p < .001$, $d = 0.63$, but not happiness pictures ($M = 29.83$; $SD = 17.15$), $p = .55$, were higher than covariation estimates for neutral pictures ($M = 28.18$; $SD = 13.28$). Covariation estimates for anger pictures were lower than for disgust, $t(37) = 3.33$, $p < .01$, $d = 0.39$, and fear pictures, $t(37) = 2.39$, $p < .05$, $d = 0.54$. There was no significant main effect of *emotion* in men, $p = .72$.

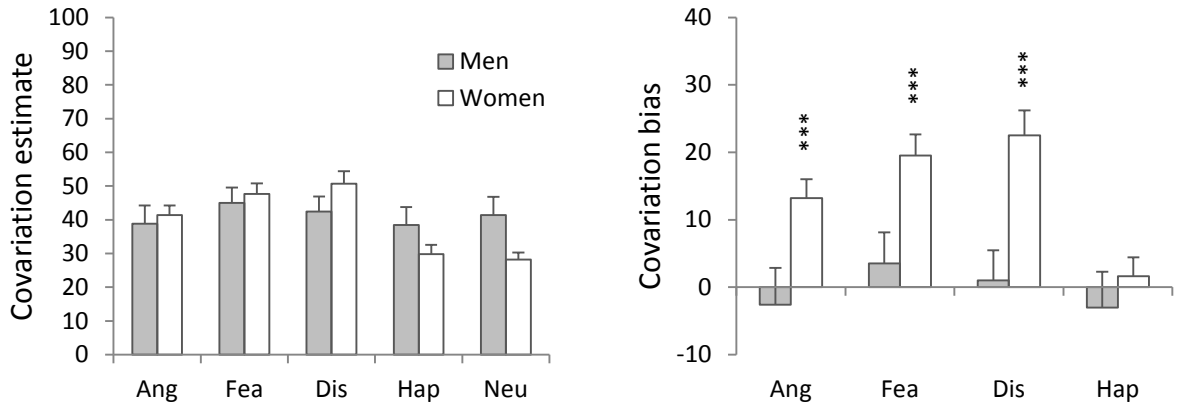

Figure 4.2. A posteriori covariation estimates (left panel) and bias scores (right panel) between picture categories and the aversive startle sound. Bias scores were calculated by subtracting covariation estimates of neutral pictures from covariation estimates of emotional pictures. In women, covariation bias was higher for disgust and fear pictures than for anger pictures. There was no covariation bias for happiness pictures and men did not display any covariation bias. Objective contingency was ~68%. dis = disgust; fea = fear; ang = anger; hap = happiness; neu = neutral. * $p < .05$; ** $p < .01$; *** $p < .001$

Startle aversiveness and correlation between covariation bias and startle aversiveness

Startles in the context of negative emotions (fear, anger, disgust), especially fear and disgust were more aversive than startles in the context of happiness and neutral pictures. A repeated measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 6.35$, $p < .001$, $\eta_p^2 = .33$, but no significant interaction of *Emotion x Gender*, $p = .35$, or main effect of

gender, $p = .21$. Startles following fear ($M = 71.22$; $SD = 20.74$), $t(55) = 5.13$, $p < .001$, $d = 0.69$, anger ($M = 61.51$; $SD = 22.12$), $t(55) = 2.98$, $p < .01$, $d = 0.40$, and disgust pictures ($M = 73.13$; $SD = 23.92$), $t(55) = 4.98$, $p < .001$, $d = 0.67$, were rated as more aversive than startles following neutral pictures ($M = 52.27$; $SD = 26.17$). Startles during fear, $t(55) = 4.33$, $p < .001$, $d = 0.58$, and disgust images, $t(55) = 4.22$, $p < .001$, $d = 0.57$, were even more aversive than startles during anger images, but fear and disgust did not differ from each other, $p = .41$. Happiness ($M = 52.04$; $SD = 31.04$) and neutral pictures did not differ significantly, $p = .93$.

Overall, correlation analyses confirmed the hypothesis of positive relationships between covariation biases and the emotional modulation of startle aversiveness (see Table 4.2 and Figure 4.3). The notion that disgust pictures were associated with the aversive startle sound was correlated with increased startle aversiveness in both men and women, ps $< .05$. Analogue correlations for fear pictures showed non-significant trends in men and women that became significant in an overall analysis, $p < .05$. Descriptively, a dissociation was observed for happiness and anger pictures with a positive relationship for anger only in men, $p < .01$, and a positive relationship for happiness only in women, $p < .01$.

	Happiness	**Fear**	**Disgust**	**Anger**
Men	$r = .045$	$r = .303$	$r = .404$	$r = .581$
	$p = .430$	$p = .111$	$p = .048$	$p = .006$
Women	$r = .486$	$r = .229$	$r = .312$	$r = -.036$
	$p = .001$	$p = .084$	$p = .028$	$p = .415$
Overall	$r = .329$	$r = .282$	$r = .348$	$r = .153$
	$p = .007$	$p = .018$	$p = .004$	$p = .131$

Table 4.2. Correlation coefficients between covariation bias (emotion minus neutral) and startle aversiveness bias (emotion minus neutral) for men and women separately and men and women combined (overall).

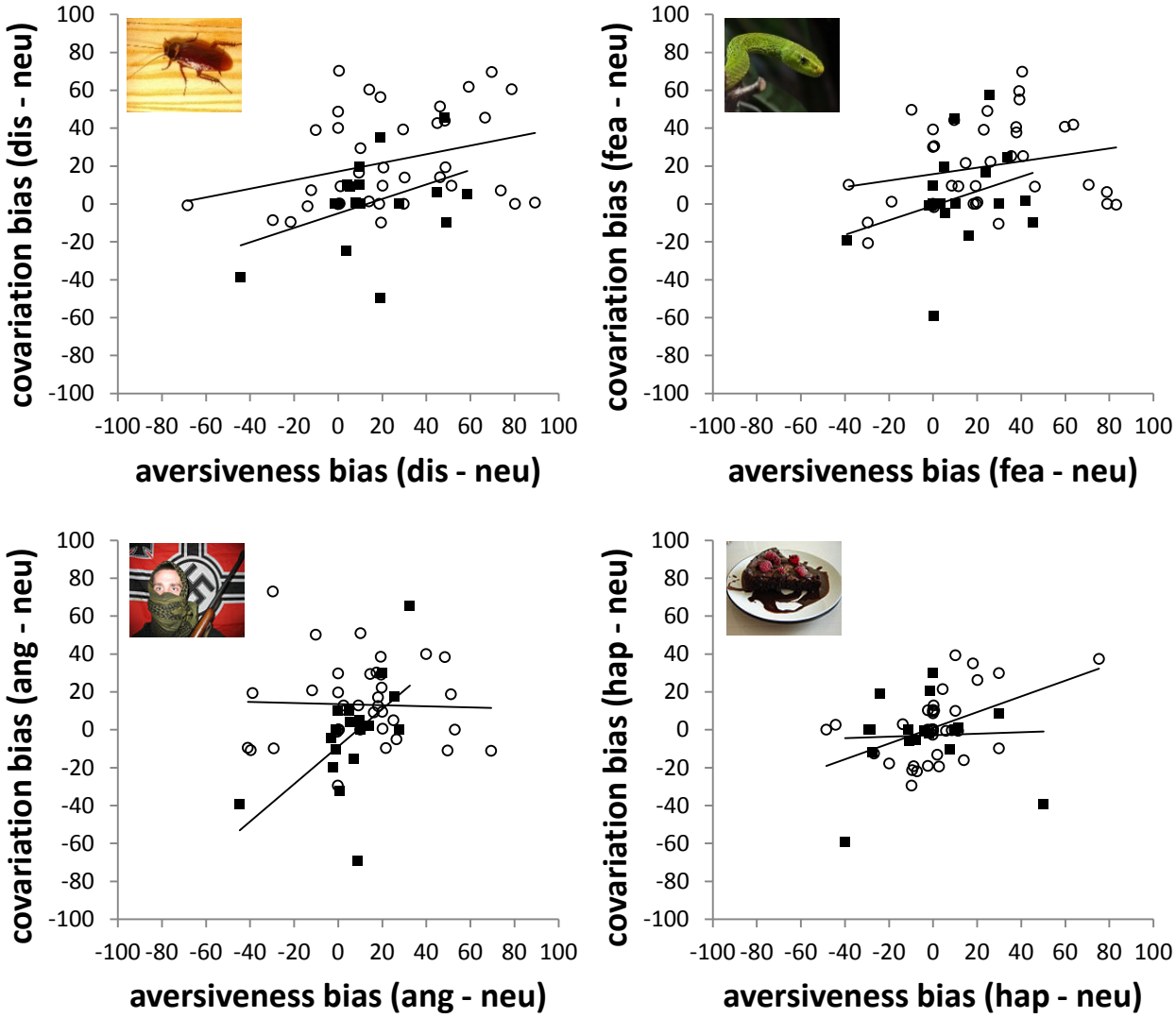

Figure 4.3. Scatterplots for correlations between covariation bias (emotion minus neutral) and aversiveness bias (emotion minus neutral). See table 4.2 for correlation coefficients. Women = white circles; Men = black squares; dis = disgust; fea = fear; ang = anger; hap = happiness; neu = neutral. The displayed pictures were found as licensed for free distribution and are only representations of the original IAPS pictures used in the study.

Linear multiple regression analysis

The correlation between startle aversiveness and covariation bias may be only the consequence of a correlation between picture valence and covariation bias. In order to find out whether the emotional modulation of startle aversiveness describes a unique portion of variance of the covariation bias that is independent from picture valence and gender, a multiple linear regression was computed. To achieve more statistical power for the analysis, all participants from experiment 1 and 2 were included (56 women, 31 men) [7]. The dependent variable was the difference in a posteriori covariation estimates between negative and neutral images. The

[7] A power analysis with GPower 3.1.2 (Franz Faul, University of Kiel, Germany) revealed that a multiple linear regression with 3 predictors would require $N = 59$ participants to achieve a power of 90% to find a medium-sized effect of $f^2 = .15$ (Cohen, 1992).

disgust category in experiment 2 was chosen as the dependent variable because it was most similar to the negative pictures in experiment 1 (mutilations) and led to the descriptively strongest covariation bias. Predictors were the difference between negative (disgust/mutilations) and neutral pictures in mean *picture valence*, in *startle aversiveness* and the *gender* of the participants. They were included in the model stepwise with *gender* first, *picture valence* second and *startle aversiveness* in the last step. *Startle aversiveness* was still a significant predictor of covariation bias, *β* = .27, *p* < .05, when gender and picture valence had been included in the model before [8].

Step 1	**(constant)**		***t* = 0.55**	***p* = .587**
	gender	*β* = .27	*t* = 2.56	*p* = .012
Step 2	(constant)		*t* = -1.31	*p* = .195
	gender	*β* = .23	*t* = 2.24	*p* = .028
	picture valence	*β* = -.28	*t* = -2.73	*p* = .008
Step 3	(constant)		*t* = -1.29	*p* = .200
	gender	*β* = .21	*t* = 2.14	*p* = .036
	picture valence	*β* = -.19	*t* = -1.84	*p* = .069
	startle aversiveness	*β* = .27	*t* = 2.61	*p* = .011

Table 4.3. Results of stepwise multiple regression analysis. Startle aversiveness explained an additional independent proportion of variance after including gender and picture valence in the model. All participants from experiment 1 and 2 were included in the analysis (56 women, 31 men). Predicted variable was the covariation bias for disgust/mutilation pictures relative to neutral pictures.

Valence, arousal and emotion ratings

The emotional pictures effectively manipulated the valence dimension in both men and women equally (see Figure 4.4). A repeated-measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 137.97$, $p < .001$, $\eta_p^2 = .92$. No main effect *gender*, $p = .39$, and no interaction was found, $p = .71$. Happiness pictures ($M = 7.21$; $SD = 0.79$) were rated as more pleasant than neutral pictures ($M = 5.52$; $SD = 0.91$), $t(55) = 11.33$, $p < .001$, $d = 1.52$. Fear ($M = 3.62$; $SD = 0.94$), $t(55) = 10.81$, $p < .001$, $d = 1.44$, and anger pictures ($M = 3.48$; $SD = 0.82$), $t(55) = 11.60$, $p < .001$, $d = 1.55$, were rated as more unpleasant than neutral pictures, and disgust pictures ($M = 2.58$; $SD = 1.16$) were rated as more unpleasant than both fear, $t(55) = 7.50$, $p < .001$, $d = 1.01$, and anger pictures, $t(55) = 5.88$, $p < .001$, $d = 0.78$.

In men and women, the arousal ratings were higher for emotional pictures than for neutral pictures (see Figure 4.4). A repeated-measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 47.02$, $p < .001$, $\eta_p^2 = .79$. No main effect *gender*, $p =$

[8] If only data from experiment 2 were taken into account, the results were very similar. All three predictors contributed significantly to the model in the last step of the regression (*gender*, $p = .004$; *picture valence*, $p = .009$; *startle valence*, $p = .04$).

.66, and no interaction was found, p = .88. Happiness (M = 5.43; SD = 0.99), $t(55) = 7.50$, $p < .001$, d = 1.89, fear (M = 5.76; SD = 1.20), $t(55) = 12.64$, $p < .001$, d = 1.69, anger (M = 5.25; SD = 1.44), $t(55) = 10.44$, $p < .001$, d = 1.40, and disgust pictures (M = 6.12; SD = 1.73), $t(55) = 11.54$, $p < .001$, d = 1.54, were rated as more arousing than neutral pictures (M = 2.71; SD = 1.63). Disgust pictures were also rated as more arousing than fear, $t(55) = 2.18$, $p < .05$, d = 0.29, anger, $t(55) = 4.65$, $p < .001$, d = 0.63, and happiness pictures, $t(55) = 2.76$, $p < .01$, d = 0.37. Fear pictures were rated as more arousing than anger pictures, $t(55) = 3.00$, $p < .01$, d = 0.41. There were no significant differences between happiness pictures and fear, p = .11, or anger pictures, p = .40.

The emotion induction was successful for all emotion categories (as indicated by significant differences from 1 [no emotion], all ps < .001), but most effective for pleasure and disgust. A repeated-measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(3, 52) = 32.34$, $p < .001$, $\eta_p^2 = .65$. No main effect *gender*, p = .71, and no interaction was found, p = .37. Disgust pictures (M = 6.93; SD = 1.73) received higher emotion induction ratings than fear (M = 5.07; SD = 1.53), $t(55) = 8.21$, $p < .001$, d = 1.09, and anger pictures (M = 6.32; SD = 1.64), $t(55) = 2.36$, $p < .05$, d = 0.32. Also, happiness pictures (M = 6.73; SD = 0.99) were rated higher than fear, $t(55) = 8.01$, $p < .001$, d = 1.07, and anger, $t(55) = 2.05$, $p < .05$, d = 0.27. Anger was rated higher than fear, $t(55) = 5.05$, $p < .001$, d = 0.68. There was no significant difference between disgust and happiness pictures, p = .41.

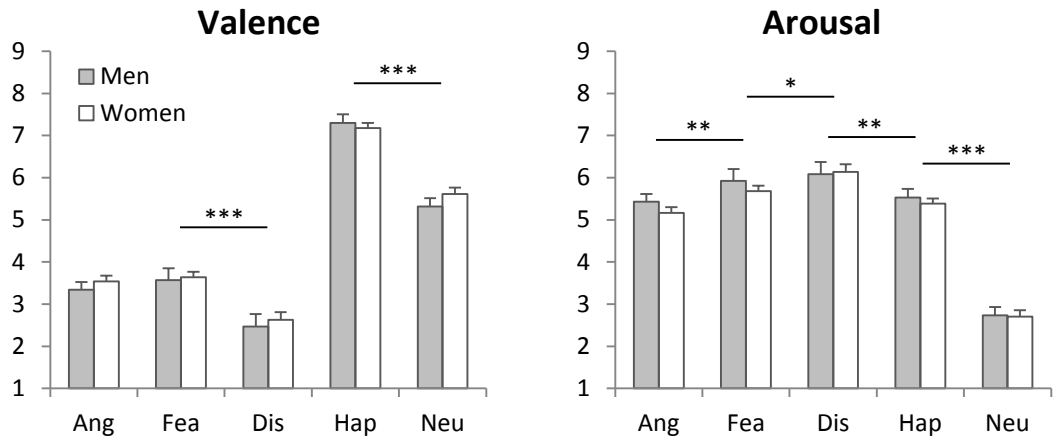

Figure 4.4. Valence (left panel) and arousal ratings (right panel) of emotional pictures. Men and women did not differ in the subjective emotional experience. High valance ratings indicate pleasantness, while low valence ratings indicate unpleasantness. Ang = Anger; Fea = Fear; Dis = Disgust; Hap = Happiness; Neu = Neutral. * $p < .05$; ** $p < .01$; *** $p < .001$

Pupil dilation

Happiness, fear and disgust pictures, but not anger pictures evoked enhanced pupil dilations relative to neutral pictures. A repeated-measures ANOVA with *emotion* and *time* as within-subjects factors, and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 8.74$, $p < .001$, $\eta_p^2 = .41$, a significant main effect of *time*, $F(4, 51) = 27.14$, $p < .001$, $\eta_p^2 = .68$, and a significant interaction of *Emotion* x *Time*, $F(4, 51) = 4.81$, $p < .001$, $\eta_p^2 = .66$. Separate repeated-measures ANOVAs for each of the 5 seconds post-stimulus with *emotion* as the only factor revealed a significant influence of *emotion* in all time points (all ps < .01), but descriptively the

emotional impact was less pronounced at the beginning of stimulus presentation (Figure 4.5). Follow-up *t*-tests showed that happiness (M = 0.09; SD = 0.17), $t(55) = 2.59$, $p < .05$, $d = 0.32$, fear (M = 0.12; SD = 0.13), $t(55) = 5.00$, $p < .001$, $d = 0.64$, and disgust pictures (M = 0.10; SD = 0.15), $t(55) = 3.17$, $p < .01$, $d = 0.39$, but not anger pictures (M = 0.04; SD = 0.12), $p = .50$, evoked stronger pupil dilations than neutral pictures (M = 0.03; SD = 0.12). There were no significant differences among happiness, fear and disgust, all $ps > .11$.

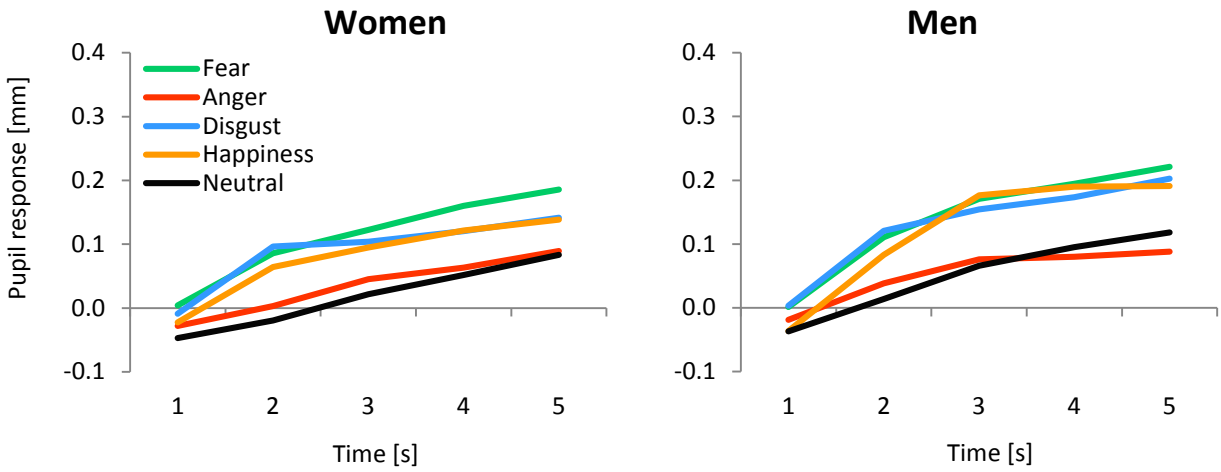

Figure 4.5. Time course of the pupil response during the five seconds of picture presentation. Fear, disgust and happiness pictures but not anger pictures evoked larger pupil dilations than neutral pictures. There was no influence of gender on pupil dilation.

Startle response (late time interval) [9]

In men, only disgust pictures potentiated startle responses relative to neutral images. In women, only fear pictures potentiated startle responses relative to neutral pictures. A repeated-measures ANOVA with *emotion* as a within-subjects factor and *gender* as a between-subjects factor revealed a main effect of *emotion*, $F(4, 51) = 6.46$, $p < .001$, $\eta_p^2 = .34$, and a significant interaction of *Emotion* x *Gender*, $F(4, 51) = 3.41$, $p < .05$, $\eta_p^2 = .21$. In women, only fear pictures (M = 51.76; SD = 3.09), $t(37) = 2.59$, $p < .05$, $d = 0.42$, but not disgust (M = 50.27; SD = 2.24), $p = .21$, and anger pictures (M = 49.77; SD = 2.46), $p = .70$, evoked higher startle amplitudes than neutral pictures (M = 49.47; SD = 3.20). There was no significant inhibition of startle amplitudes in happiness pictures relative to neutral pictures, $p = .23$. In men, however, only disgust pictures (M = 51.91; SD = 2.88) reached significance when compared to neutral pictures (M = 49.11; SD = 2.83), $t(17) = 2.44$, $p < .05$, $d = 0.57$; but not fear (M = 49.70; SD = 2.18), $p = .58$, or anger pictures (M = 50.70; SD = 3.01), $p = .11$. There was also no significant inhibition of startle amplitudes in happiness pictures relative to neutral pictures, $p = .58$.

[9] The early startle response is not reported here, because it did not show the typical pattern of prepulse inhibition. That is, unpleasant emotions did not inhibit, but potentiate the startle reflex in the early time interval. Therefore, the early startle response may be difficult to be meaningfully interpreted in terms of prepulse inhibition. Methodological characteristics of the present study like the usage of a baseline mask might be responsible for this untypical finding.

4.3. Discussion

Experiment 2 investigated illusory correlations between pictures inducing distinct emotions (happiness, fear, disgust, anger) and aversive outcomes. The research questions were (1) whether positive arousing stimuli do also induce a covariation bias with aversive events, (2) whether the aversiveness of outcomes was correlated with covariation bias, (3) whether covariation bias was enhanced in women, and (4) whether covariation bias would be differently pronounced in various emotions.

Unlike experiment 1, the experimental paradigm was primarily an affect-modulated startle paradigm. So, the participants were not attending to picture-startle-associations during the experiment. The explicitly instructed attention to these associations might have been one reason why experiment 1 did not reveal an a posteriori covariation bias for positive arousing stimuli. However, clearly, positive arousing happiness pictures did not lead to a covariation bias in experiment 2. Although the ratings might not be directly comparable between experiment 1 and 2, positive pictures were rated as slightly less arousing in experiment 2 (M = 5.43 vs. M = 6.03). Nevertheless, it is not very likely that this difference in arousal ratings is responsible for the lack of a covariation bias since arousal ratings for positive pictures were still lower in a previous study, in which participants were overestimating the proportion of startle sounds during positive slides (Pauli et al., 2002). A critical difference between this study and experiment 2 here might be that overall the proportion of happiness pictures was relatively low among all pictures (1/5 vs. 1/3). It is possible that participants were overestimating startles in negative slides with priority and then not overestimating startles in positive slides anymore to balance the overall proportion of startle sounds. Considering all the evidence from the present and previous studies (Cavanagh & Davey, 2001; de Jong & Merckelbach, 2000; Pauli et al., 2002; VanOyen Witvliet & Vrana, 2000), it seems that the unspecific arousal of emotional experiences does not have a large and robust effect on illusory correlations. The valence of the emotional stimulus and of the aversive consequence seem to be more important in the emergence of a covariation bias than the arousal of the emotional stimulus. This is in accordance with an affective matching account (Tomarken et al., 1995) and/or the influence of a magnified aversiveness of the outcome.

The latter assumption was also supported by the present finding of a correlation between the increased subjective aversiveness of outcomes after negative pictures relative to neutral pictures and the appropriate covariation bias. This means that those who overestimated the association between negative pictures and aversive outcomes, were also likely to be those who perceived the startle sound as more aversive if it followed a negative picture. This was especially true for disgust and fear pictures, i.e. those pictures that also provoked the largest covariation bias. At this point, one might argue that these correlations are only a side-effect of a relationship with picture valence, meaning that startle aversiveness is a third variable correlating with picture valence and actually irrelevant for the covariation bias. However, a stepwise multiple regression analysis revealed that subjective startle aversiveness explained an independent proportion of covariation bias variance if gender and picture valence effects were already partialled out. In this circumstance, one limitation has to be admitted which is that startle valence and covariation bias were collected at the same time point while picture valence was rated some minutes later. This might also explain an independent relationship between covariation bias and

startle aversiveness. Nevertheless, so far the data support the idea that the modulation of startle aversiveness plays a unique and important role in the emergence of illusory correlations.

Interestingly, a marginal significant interaction between gender and picture category indicated that an a posteriori covariation bias only existed in women. This finding was confirmed in a multiple regression analysis for disgust pictures, in which gender was a significant predictor of covariation bias. It was expected that women displayed a stronger covariation bias, because in experiment 1, too, only women displayed an expectancy bias for mutilation pictures.

It has to be acknowledged that both experiment 1 and experiment 2 were originally not conducted to examine gender effects on covariation bias. The sample sizes between men and women were not equal and this analysis has to be considered as exploratory. However, there are several reasons why the present finding should reflect a real effect that is warrant to be further tested in future studies. First, in total 56 women and 31 men were tested in both experiments, which suggests that although sample sizes were not balanced, at least a substantial proportion of both genders were included in the analysis. Second, men and women were matched relatively well regarding anxiety and emotional responding to the pictures. There was only a significant difference in state anxiety in experiment 2, but in terms of women being less anxious than men were. Therefore, it is unlikely that heightened anxiety in women is the mediator of the present gender effect. Although women were by trend younger than men in experiment 2, age is probably not a critical mediator, too, because the gender effect emerged in experiment 1 despite the absence of age differences between men and women. Also, men and women did not differ in any measures of emotional responding including valence and arousal ratings, startle response, skin conductance response and pupil dilation. Third, the effect sizes involved in the influence of gender on illusory correlations suggest a very large difference between men and women. For instance, in experiment 2, women displayed a large covariation bias for disgust pictures (d = 0.90; Cohen, 1992) while virtually a complete absence of a covariation bias for disgust pictures was observed in men (d = 0.05).

The difference between men and women in illusory correlations should be replicated in experiments containing large and equal samples of both genders and controlling for confounding variables in the outset of the study. Still, the present exploratory analysis already delivers quite convincing evidence that nourishes the assumption that such future experiments will reveal a substantial gender effect. However, the question remains why exactly women are more prone to a covariation bias or expectancy bias for negative emotional stimuli. The present absence of differences in subjective and physiological emotional responding suggests that this is not a mediating factor of sex differences in illusory correlations. Although, some previous studies found differences in physiological and brain responses to emotional stimuli (Kring & Gordon, 1998; Domes et al., 2010), there were no differences in startle reflex and pupil dilation here. When comparing responses to generally positive and negative pictures, physiological differences between men and women seem to be mainly restricted to facial expressions as indicated by corrugator and zygomaticus activity, but not observed in heart rate, startle reflex or skin conductance (Bradley, Codispoti, Sabatinelli, & Lang, 2001). The finding that the subjective experience of emotional reactivity does not differ between genders is in accordance with many previous investigations (Kring & Gordon, 1998; McRae, Ochsner, Mauss, Gabrieli, & Gross, 2008; Domes et al., 2010), but in contrast with some others (e.g. Bradley et al., 2001).

Taken together, the result of no differences in emotional responding between men and women is not unusual and suggests that the present sex differences in illusory correlations more likely emerged due to higher-order cognitive processes. In line with this assumption, stronger sex differences were observed in dependent variables under voluntary control (i.e. subjective reports, facial expression) than in more automatic variables (i.e. startle reflex, heart rate, skin conductance; Bradley et al., 2001). Moreover, there is evidence that enhanced emotional arousal in women occurs only at a relatively late stage of emotional processing (Gard & Kring, 2007). Other investigators found that men and women differ in emotional regulation processes (McRae et al., 2008; Domes et al., 2010) and evaluative processes (Lee, Liu, Chan, Fang, & Gao, 2005) that might take place at such a later processing stage. For example, men were found to recruit the prefrontal cortex to a higher (Domes et al., 2010) or lesser (McRae et al., 2008) degree than women during the down-regulation of negative affect while subjective experience was similar as in women. Moreover, a previous study found that women remember negative stimuli better than men do and might use different neural circuits to encode the emotional information (Canli, Desmond, Zhao, & Gabrieli, 2002). This effect has already been replicated (Mackiewicz, Sarinopoulos, Cleven, & Nitschke, 2006). Interestingly, other authors found that anticipatory brain activity (EEG) predicted later emotional memory performance in women, but not in men (Galli, Wolpe, & Otten, 2011). In a PET-study, women showed stronger connectivity in an 'emotional-arousal-network' than men when anticipating aversive visceral stimulation. This sex difference was especially pronounced during the expectation of the aversive event (Labus et al., 2008). One might speculate that the previously found sex differences in the anticipation phase of emotional events contributed to the fact that women were expecting more aversive outcomes after negative pictures in experiment 1 and overestimated this relationship in experiment 2. In sum, some previous findings suggest that men and women might differ in the linkage between 'emotional' and 'cognitive' processes, such as regulatory, evaluative and anticipatory processes. Likewise, the present investigation implies that women are more susceptible to emotionally induced cognitive biases than men are. If this result can be replicated in future studies, this could be a so far relatively unexplored and promising explanation for the fact that women are more likely to suffer from affective and anxiety disorders (Gater et al., 1998; Jacobi et al., 2014).

Besides the explanation that women display a different emotion-cognition-interaction than men, one might also speculate that women answer more socially desirable than men do and show a stronger demand effect than men. However, men and women do usually not differ on a social desirability scale (Reynolds, 1982). Furthermore, in experiment 1, there was no a posteriori covariation bias in spite of an a priori expectancy bias. If women only displayed this expectancy bias due to a demand effect, the question remains why they did not display an a posteriori covariation bias due to a demand effect. Moreover, the aforementioned findings of different neural mechanisms involved in the anticipation of aversive events suggest that biased expectancy is actually measurable on a more objective level and does not solely reflect demand effects. Future experiments examining gender effects on covariation bias could also make use of the contingent negative variation (CNV) as a physiological measure of biased expectancy (Walter, Cooper, Aldridge, McCallum, & Winter, 1964). To avoid temporal confounding of anticipatory brain potentials (CNV) and picture processing potentials (e.g. LPP), a cueing picture could be presented briefly and followed by a short anticipation period without a picture before an aversive stimulus is applied. Then, women should indicate a more negative CNV after negative pictures than after neutral pictures.

It is especially warrant to examine gender differences in illusory correlations considering the fact that the vast majority of past illusory correlation experiments was based on female participants (see Table 9.1). Only two illusory correlation experiments so far reported analyses considering the gender of participants (Davey, Cavanagh, & Lamb, 2003; Muris, Huijding, Mayer, den Breejen, & Makkelie, 2007). There was no influence of gender on the expectancy of aversive outcomes following predatory and non-predatory animals (Davey et al., 2003). Like the present experiments, this study was not explicitly designed to examine gender differences and the proportion of men and women was imbalanced. Another thought experiment found an enhanced expectancy bias in girls between nine and 16 years, although this might have been a consequence of elevated fear of spiders (Muris et al., 2007). Future experiments should consider equal numbers of both high and low fearful male and female participants, and match the groups in respect of age and fear level. However, the present results suggest that fear-relevant illusory correlations might be completely irrelevant in the maintenance of anxiety in men. Moreover, some previous results may have to be interpreted differently in the light of such substantial gender differences. For example, several studies suggest that an expectancy bias for aversive events after fear-relevant stimuli is present in both low and high fear participants (de Jong, 1993; Kennedy et al., 1997; Amin & Lovibond, 1997) while an a posteriori covariation bias is only observed in high fear individuals (Kennedy et al., 1997; Amin & Lovibond, 1997). Now, some studies found a difference already in an expectancy bias, in terms of a smaller expectancy bias in low fear subjects (e.g. McNally & Heatherton, 1993). With potential sex differences in mind, this deviant finding might have been the result of imbalanced gender proportions between the high and the low fear group. That is, the proportion of men was higher in the low fear group, which should already lead to an attenuated expectancy bias considering the present findings.

Finally, there was some evidence that the covariation bias differed between different basic emotions. While fear and disgust evoked the strongest bias, the bias for anger was less pronounced. This might be the consequence of lower arousal induced by anger pictures. Although the evidence for an impact of arousal on illusory correlations is inconsistent, arousal might still play a role if it is directly evoked by a negatively valenced stimulus. On the other hand, an attenuated effect of anger on covariation bias was also accompanied by reduced startle aversiveness relative to fear and disgust pictures. This is in line with the idea that threat cues (as present in fear and disgust images here) are associated with increased sensitivity to prepare for upcoming dangers. In contrast, anger images mostly depicted signs of social injustice and environmental drawbacks where the viewer is not threatened so much. Therefore, aversive events during fear and disgust might have been modulated as more aversive and encoded more deeply than aversive events during anger. As a consequence, the covariation bias is stronger in fear and disgust.

In summary, experiment 2 revealed that the arousal of an emotional stimulus does not influence covariation bias independent from stimulus valence. Moreover, the finding that the emotional modulation of startle aversiveness predicted covariation bias was replicated here. In addition to experiment 1, further evidence for sex differences in illusory correlations was found, in terms of this cognitive bias was restricted to women. This effect was not explainable by increased anxiety or emotional responding in women, and may be an explanation for increased prevalence of anxiety disorders in women.

5. Experiment 3: Brain activity associated with illusory correlations in spider phobia

5.1. Introduction

So far, experiment 1 and experiment 2 have demonstrated that there is a relationship between the subjective aversiveness of an outcome and illusory correlations. As the model described in chapter 2.5. predicts, the more aversive an outcome is perceived, the more likely it is associated with a preceding stimulus. This correlation was stable across both studies and still present if the subjective aversiveness of the preceding stimuli was partialled out. However, the correlation depended on subjective reports following the experiment. Therefore, it might be possible that those who overestimated the contingencies between negative pictures and aversive outcomes did not actually perceive the outcomes as more aversive than after neutral pictures. In contrast, it might be that those participants only reported as though they had perceived it as more aversive due to demand effects of the experiment or memory failure. Thus, it would be eligible to show this relationship via a more objective measure of outcome processing at the moment of outcome application. To this end, an fMRI-study was conducted to test whether there is actually biased processing of outcomes associated with illusory correlations. For this purpose, electrical stimuli were used instead of startle probes, because auditory stimuli are less suitable for the noisy environment in an fMRI-scanner and brain responses to painful electrical stimulation has been studied more extensively than responses to startle sounds (e.g. Roy et al., 2009). In this circumstance, the following paragraphs outline a short summary of previous relevant findings of the neural processing of phobia-relevant stimuli and painful stimulation.

Brain imaging of phobic fear

To date, almost 40 studies have been published that reported brain activity provoked by confrontation with phobia-relevant stimuli in specific phobia. Brain imaging was performed using functional magnetic resonance imaging (fMRI), positron emission tomography (PET) and single-photon emission computed tomography (SPECT) and has already been reviewed by several authors (Etkin & Wager, 2007; Linares et al., 2012; Shin & Liberzon, 2010). Apparently, most reliable activation has been found in a network comprising the amygdala, the dorsal ACC and the insula. In addition, the striate and extrastriate visual cortical areas, the thalamus and the fusiform gyrus are commonly found regions. Diverging evidence was reported regarding rostral ACC and dlPFC. Especially activity in the dlPFC was enhanced in some studies, but reduced in others. Findings regarding the dlPFC are summarized in more detail further below.

One study was using MRI to measure cortical thickness in specific phobia and healthy controls (Rauch et al., 2004). Phobic patients exhibited greater cortical thickness in rostral ACC, posterior cingulate cortex, insular cortex and left visual cortex. The amygdala and the hippocampus were not assessed in this study. Interestingly,

there seem to be differences in the neural correlates of subtypes of specific phobia. In one study for example, snake phobia was characterized by increased activity in the ACC, the thalamus and the insula, while dental phobia led especially to increased activation in the orbitofrontal cortex and superior frontal gyrus (Lueken et al., 2011). Moreover, in comparison to dental phobia, snake phobia is characterized by increased amygdala and hippocampus activation during anticipation of phobia-relevant stimuli and a shift to midbrain areas (Lueken et al., 2013). Blood-injection-injury phobia was found to be associated with enhanced activity of the thalamus and visual areas in comparison to spider phobia (Caseras et al., 2009). In the following, the evidence for the most relevant regions in animal phobia is described in more detail.

Amygdala. Activation of the amygdala has been reported in Pavlovian fear conditioning in humans (e.g. Andreatta, Mühlberger, Yarali, Gerber, & Pauli, 2010; Büchel, Morris, Dolan, & Friston, 1998) and plays an important role in the acquisition and output regulation of conditioned fear responses (LeDoux, 2000). In general, emotionally arousing and especially negative emotional stimuli evoke amygdala activity. In response to phobia-relevant stimuli, this seems to occur relatively fast (Larson et al., 2006) and in some cases even when phobics are distracted from phobia-relevant stimuli (Straube, Mentzel, & Miltner, 2006). Also, subliminally presented spiders and snakes have the potential to induce amygdala activation in the absence of conscious awareness (Carlsson et al., 2004; Lipka, Hoffmann, Miltner, & Straube, 2013). Although many studies showed increased amygdala activation in response to phobia-relevant stimuli, some experiments failed to do so (e.g. Paquette et al., 2003; Straube et al., 2004). One potential reason for the absence of differential amygdala activity could be habituation processes during repeated presentation of threatening stimuli (Fischer et al., 2003). It should be noted that the majority of studies using emotional stimuli found lateralized amygdala activity, mostly to the left amygdala, but the question is unresolved whether and how this traces back to differential functionality of the amygdalae (Baas et al., 2004). In sum, it is plausible to assume that the amygdala plays an important role in the rapid evaluation of phobia-relevant stimuli, and precedes more elaborative processing in other brain areas.

ACC. The ACC can be subdivided into the rostral and the dorsal ACC. This breakdown is supported by both neuroimaging (Bush, Luu, & Posner, 2000) and cytoarchitectonic studies (Vogt, Nimchinsky, Vogt, & Hof, 1995). The rostral part seems to be more relevant for emotion regulation and emotional evaluation, and is connected to subcortical regions like the amygdala, the nucleus accumbens, the hypothalamus and the periaqueductal grey, as well as the orbitofrontal cortex and the insula (Bush et al., 2000). On the other hand, the dorsal ACC – which is found more commonly and almost invariably in imaging studies of specific phobia – is considered to be more relevant to 'cognitive' processes and exhibits connectivity with dlPFC, parietal cortex, premotor cortex and the supplementary motor area (SMA). It is important for attentional and executive functions, monitoring errors and response selection. Recently, it has been suggested that the dorsal ACC may serve as a link between emotional information and appropriate response output (Shackman et al., 2011) which may be in the case of animal phobia the urge to flee. The dorsal ACC which has also been termed anterior mid cingulate cortex (aMCC) seems to be specifically involved in fear processing, while other parts of the cingulate cortex are more related to the processing of other emotional material (Vogt, 2005). In line with this, the dorsal ACC receives input from the amygdala (Vogt & Pandya, 1987). In contrast to the amygdala, the dorsal ACC responds to phobic stimuli only if the stimuli are processed directly without distraction (Straube et al., 2006). In

addition, the ACC is active during the anticipation of negative emotional and phobia-relevant stimuli, which also correlates with symptom severity (Straube, Mentzel, & Miltner, 2007). After exposure (Goossens, Sunaert, Peeters, Griez, & Schruers, 2007) and cognitive behavioral therapy (Straube, Glauer, Dilger, Mentzel, & Miltner, 2006; Lipka et al., 2013), the hyper-responsiveness of the dorsal ACC to phobia-relevant stimuli vanishes.

Insula. The insular cortex is active during interoception (Craig, Chen, Bandy, & Reiman, 2000) and awareness of the own body (Tsakiris, Hesse, Boy, Haggard, & Fink, 2007). It possibly plays a role in consciousness and awareness in general (Craig, 2009). Moreover, visceral stimulation leads to insula activity (Derbyshire, 2003). Phobia-relevant stimuli evoke activity in the human insula (Shin & Liberzon, 2010), and also general emotional stimuli if cognitive processes are involved (Phan, Wager, Taylor, & Liberzon, 2002). Patients with lesions in the insular cortex exhibit attenuated valence and arousal ratings of emotional pictures (Berntson et al., 2011). Taken together, one might speculate that a conscious cognitive representation of emotional arousal and/or valence is reflected in enhanced insula activity during phobic stimulation. Like ACC activity, insula activity normalizes after treatment (Straube et al., 2006; Goossens et al., 2007; Schienle, Schäfer, Hermann, Rohrmann, & Vaitl, 2007).

dlPFC. The evidence regarding dlPFC activity in specific phobia is equivocal with some studies showing hyperactivity (Paquette et al., 2003; Straube, Mentzel, Glauer, & Miltner, 2004; Schienle, Schäfer, Walter, Stark, & Vaitl, 2005; Åhs et al., 2009) and others showing reduced activity to phobia-relevant stimuli (Carlsson et al., 2004; Schienle et al., 2007). The dlPFC is involved in working memory processes (Curtis & D'Esposito, 2003) and is active in response to threatening or emotional stimuli in general (Phan et al., 2002; Wright et al., 2001). It has been suggested that increased dlPFC activity to phobia-relevant stimuli reflects attempts of emotional down-regulation (Paquette et al., 2003). This view is supported by the finding that hyperactivity in the dlPFC is associated with successful prevention of panic during phobic fear (Johanson et al., 1998), and that this region is important in emotion regulation (e.g. Goldin, Manber, Hakimi, Canli, & Gross, 2008). Alternatively, dlPFC activity may be associated with catastrophizing thoughts, similar to rumination effects in major depression (Cooney, Joorman, Eugène, Dennis, & Gotlib, 2010).

Brain imaging of painful stimulation

Brain imaging studies of painful stimulation revealed a reliable activation pattern in the ACC, the insula, and the primary and secondary somatosensory cortices, which have been considered as part of the 'pain matrix' (e.g. Tracey & Mantyh, 2007). However, recent evidence suggests that this activation pattern is not specific to pain perception per se, but most of the regions are responsive to salient stimuli in general (Legrain, Iannetti, Plaghki, & Mouraux, 2011). Nociceptive-specific activity may be located in the frontal operculum and the medial prefrontal cortex (Mouraux, Diukova, Lee, Wise, & Iannetti, 2011). In general, the functional role of the individual 'pain matrix' components is so far not completely understood. Activity within the somatosensory cortex could be specific to somatosensory stimuli, while activity in the secondary somatosensory cortex seems to be multimodal (Mouraux et al., 2011). Mouraux et al. (2011) further suggest that the insula is involved in non-pain-specific cognitive and emotional processes, while activity in the rostral ACC may be related to pain unpleasantness (Vogt, Derbyshire, & Jones, 1996) and the dorsal part to attentional orienting to pain (Peyron,

Laurent, Garcia-Larrea, 2000). Nevertheless, many studies exhibited that mood and emotions can alter pain experience and stimulus-related brain activity. Unpleasant odors or pictures for example lead to increased unpleasantness ratings of pain stimuli without affecting pain intensity ratings (Loggia, Mogil, & Bushnell, 2008; Villemure, Slotnick, & Bushnell, 2003). Recently, such findings were confirmed using fMRI and spinal nociceptive reflex modulation as more objective measures of pain processing under negative affect (Roy et al., 2009): When participants were viewing unpleasant emotional pictures, spinal nociceptive responses to painful electric stimuli were enhanced. Moreover, pain-related activity in the contralateral insula, paracentral lobule (primary sensory-motor area), parahippocampal gyrus, thalamus and amygdala was increased during negative emotion relative to positive emotion. Emotional and cognitive modulation of pain may be realized by descending signal transmission originating in the prefrontal cortex via amygdala, insula and the periaqueductal grey (PAG), thus inhibiting ascending nociceptive input to the brain (Wiech & Tracey, 2009).

To the best knowledge of the author, the modulation of responses to painful stimuli by phobic material has not yet been investigated. The only study, which examined the processing of aversive stimuli during the presentation of phobia-relevant pictures, used acoustic startle probes in a PET scanning paradigm (Pissiota et al., 2003). The authors found that fear-potentiated startle responses in animal phobia were reflected in enhanced activation of the rostral ACC and the left amygdaloid-hippocampal region. Strictly speaking, whether this activity traces back to anticipation of the startle probes, reaction to the startle probes or a modification of picture processing by the startle stimuli cannot be differentiated because the experimental conditions had to be presented in blocks.

The present experiment

In order to illustrate the main rationale of experiment 3, the reader should be reminded briefly of the theoretical background. As stated previously, fear conditioning serves as a model for the onset and maintenance of anxiety disorders, such as specific phobia (Mineka & Oehlberg, 2008). Indeed, several studies suggest abnormal conditioning processes in clinical and highly anxious populations (Glotzbach-Schoon et al., 2013; Lissek et al., 2009), but overall effect sizes differentiating patients from healthy controls seem to be relatively small (Lissek et al., 2005; Beckers et al., 2013). One reason may be that most conditioning studies realized a clear contingency between the conditioned stimulus (CS) and the aversive unconditioned stimulus (US). This is problematic because differences between differently anxious individuals have often emerged particularly in ambiguous situations (Lissek et al., 2006), which is not the case if contingencies are clear. Alternatively, it might be assumed that individuals with phobia specifically suffer from a biased perception of the contingency between a fear-related CS (e.g., a spider) and an aversive US (e.g., an electrical shock). In line with this approach, an illusory correlation or covariation bias has been found in fearful and phobic individuals between phobia-relevant stimuli and aversive events (Davey, 1995; de Jong & Peters, 2007; Mühlberger et al., 2006; Tomarken et al., 1989). In a classic illusory correlation experiment (e.g., Tomarken et al., 1989), participants are exposed to neutral and phobia-relevant pictures that are followed by different outcomes (i.e. aversive shock, neutral tone or nothing). Although the relationship between pictures and outcomes is random, individuals suffering from animal phobia

were found to overestimate specifically the association between phobia-relevant pictures and aversive outcomes. Interestingly, before such an experiment even non-fearful individuals are more likely to expect a shock following fear-relevant stimuli relative to neutral stimuli. Still, only phobic participants associate spiders and shocks after confrontation with random contingencies (Mühlberger et al., 2006; Davey & Dixon, 1996). Importantly, such a covariation bias likely contributes to the development or maintenance of phobia as it was found that illusory correlations persisting immediately after exposure treatment were a significant predictor of relapse two years later (de Jong, van den Hout, & Merckelbach, 1995).

It was assumed that abnormal neural processing of CS-US contingencies could in part explain why phobia patients have difficulties to learn that spiders are not systematically related to negative consequences. Therefore, functional magnetic resonance imaging (fMRI) was used to measure the brain activation of spider phobics and healthy controls during an illusory correlation experiment.

First, the dorsolateral prefrontal cortex (dlPFC) was expected to be a potential structure to be involved in illusory correlations. This assumption is based on the finding that enhanced dlPFC activity is related to contingency awareness during fear conditioning (Carter, O'Doherty, Seymour, Koch, & Dolan, 2006). Moreover, the dlPFC is less active in a working memory task when participants are distracted by emotional stimuli (Dolcos und McCarthy, 2006). Finally, previous studies showed that phobic stimuli provoked enhanced (Aupperle et al., 2009; Paquette et al., 2003; Schienle et al., 2007; Straube et al., 2004) or reduced (Carlsson et al., 2004) activity in dlPFC. Its role in an illusory correlation experiment is not known so far. Altered executive control and/or working memory performance (Miller & Cohen, 2001; Curtis & D'Esposito, 2003) may be reflected in altered dlPFC activity in the presence of phobia-relevant stimuli and account for impaired contingency monitoring in phobic individuals. Therefore, deviant dlPFC activity to spider images should prevent spider phobics from correcting biased contingency estimates, and should thus be correlated with the covariation bias.

Second, increased responses to shocks following phobia-relevant stimuli were expected within typical pain-processing areas, mainly comprising primary sensory-motor cortex, secondary sensory cortex, anterior cingulate cortex (ACC) and insula (Peyron et al., 2000). Despite several studies on the emotional modulation of pain (Kenntner-Mabiala & Pauli, 2005; Kenntner-Mabiala et al., 2008; Roy et al., 2009; Reicherts, Gerdes, Pauli, & Wieser, 2013), the impact of phobic stimuli on the neural processing of painful stimuli has not yet been investigated. Since already the first covariation bias study in animal phobia observed that the reported pain elicited by the shocks following spiders predicts covariation bias (Tomarken et al., 1989), it was further expected that hyperactivity in these pain-processing areas should predict covariation bias. Moreover, a connectivity analysis was performed to reveal additional regions potentially involved in the phobic modulation of pain.

Taken together, amplified shock-related activity was expected in response to spider-shock associations and aberrant dlPFC activity in response to spider images to be involved in phobia-relevant illusory correlations. In addition, the contribution of other brain areas typically involved in the processing of phobia-relevant stimuli to illusory correlations was examined (i.e., amygdala, ACC, insula).

5.2. Methods

Participants

Thirty-eight participants (20 patients with spider phobia and 18 non-spider-fearful females) were recruited via local advertisements in newspapers and websites. Individuals with high (≥ 20 of max. 24) or low (≤ 8) scores in a short screening for fear of spiders (SAS; Rinck et al., 2002) were contacted and invited to participate if they were female, right-handed, did not report a history of psychiatric or neurological disorder, did not take psychoactive medication, and had not been treated because of specific phobia or other psychological disorders. The presence of specific phobia was confirmed by a trained psychologist using the structured clinical interview for DSM-IV (SKID; Wittchen, Zaudig, & Fydrich, 1997). Two of the 20 patients with spider phobia had to be excluded from data analysis (acute illness, resp. stimulation electrode came off). Another patient was included in fMRI analysis, but was excluded from the analysis of ratings and correlations between fMRI and ratings because of a technical failure during data recording.

	Spider Phobia	***N***	**Healthy Control**	***N***	
Age	**21.4** ± 4.2	18	**22.2** ± 2.2	18	$p = .461^{t}$
Education	**17** A/1 M	18	**16** A/2 M	18	$p = 1.00^{x}$
SPQ	**23.2** ± 2.8	17	**4.6** ± 2.4	16	$p < .001^{t}$
FSQ	**76.7** ± 4.8	17	**4.8** ± 7.0	16	$p < .001^{t}$
STAI -state	**43.6** ± 7.3	18	**36.2** ± 8.2	18	$p = .008^{t}$
STAI-trait	**42.4** ± 5.6	17	**37.2** ± 8.4	16	$p = .044^{t}$
BDI-II	**9.2** ± 5.5	17	**6.2** ± 4.0	16	$p = .082^{t}$
Pain thr. (obj.) [mA]	**1.3** ± 0.6	18	**2.0** ± 0.9	18	$p = .008^{t}$
Pain thr. (subj.)	**4.4** ± 0.7	18	**4.8** ± 1.4	18	$p = .363^{t}$

Table 5.1. Mean and standard deviation of demographic and psychometric sample characteristics. Age is specified in years; Education is stated in the number of participants with higher education entrance qualification (A) and secondary education (M). P-Values according to ^{t}two-sided t-tests and ^{x}Fisher's exact test for cross tables. *N* = number of participants (some participants had to be discarded due to missing data).

Spider phobics and non-spider-fearful participants were matched in terms of age and education (see Table 1). The two groups significantly differed in the Spider Phobia Questionnaire (SPQ; Watts & Sharrock, 1984), the Fear of Spiders Questionnaire (FSQ; Szymanski & O'Donohue, 1995), and the state version of the State Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970). In addition, the spider phobia group was slightly more trait-anxious according to the STAI. On the Beck Depression Inventory the groups were not significantly different (BDI-II; Beck, Steer, Ball, & Ranieri, 1996).

Stimuli

Ninety different color pictures from a custom-made picture set were used in the illusory correlation experiment (30 spider pictures for phobia-relevant trials; 30 mushroom pictures for neutral trials; 30 puppy pictures for filler trials). The pictures were matched regarding complexity, that is, each picture showed only one object centered in the foreground. After the illusory correlation experiment, the pictures were rated on Likert-like scales with verbal labels for the endpoints regarding valence (1 = "very unpleasant" to 9 = "very pleasant") and arousal (1 = "not arousing at all" to "9 = very arousing").

Aversive electric stimuli (400 V for 767 ms) were generated by a current stimulator (Digitimer DS7A, Digitimer Ltd, Welwyn Garden City, UK) and applied via two steel surface electrodes (9 mm diameter; gvb-gelimed, Bad Segeberg, Germany) to the inner side of the left calf. Individual pain thresholds were identified via painfulness ratings (0 = no sensation to 10 =maximal painful sensation) of two increasing and two decreasing series of electric stimuli in steps of 0.5 mA. Participants were instructed that a rating of 4 indicated the beginning of painful sensation. This procedure resulted in a mean intensity threshold of M = 2.02 mA (SD = 0.94) for healthy controls and M = 1.27 mA (SD = 0.62) for spider phobics. This difference was statistically significant (p < .01, two-sided), but the subjective painfulness did not differ between groups (healthy controls: M = 4.78, SD = 1.35; spider phobics: M = 4.44, SD = 0.71; p = .36, two-sided). To ensure that stimuli were perceived as aversive, stimuli presented during the experiment consisted of electrical stimulation at the individual pain threshold plus 20%.

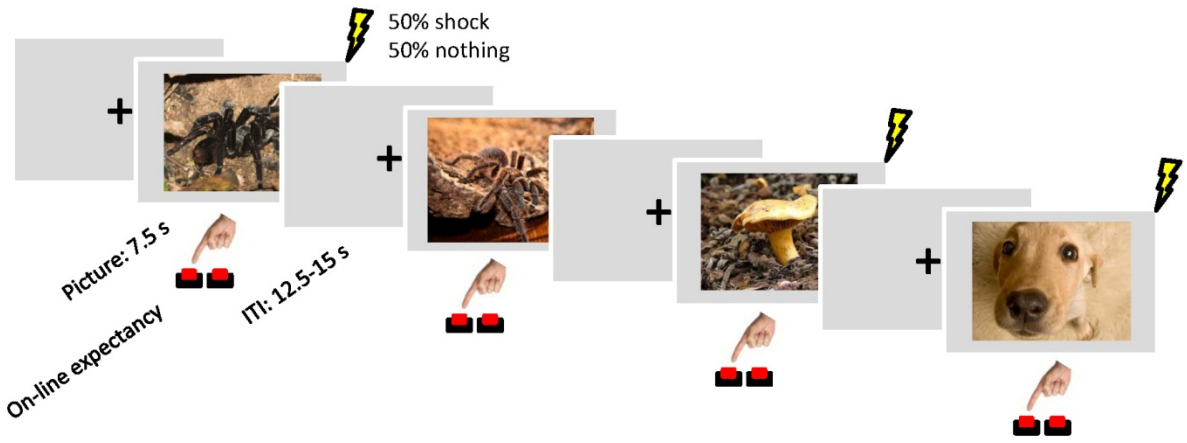

Figure 5.1. Illusory correlation paradigm. Three categories of pictures were presented (spiders, mushrooms, and puppies). Exactly 50% of the pictures of each category were followed by a painful electrical shock. Participants were asked to indicate with a button press whether they expected a shock at the end of the picture or not. A jittered inter-trial-interval (ITI) enabled the measurement of different brain regions at the same point in time.

Procedure

Ethical approval for this study was obtained from the ethics committee of the medical faculty of the University of Würzburg and in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Signed informed consent was obtained for participants after the procedure was explained. Next, all participants filled out the STAI state. Then, individual pain thresholds were determined while participants were lying in the scanner room (see 'stimuli' section). In addition, skin conductance electrodes were attached to left index and middle finger, but results are not reported due to technical failure affecting a large proportion of participants. Before the illusory correlation-experiment started, they were shown a random example picture of each category (spiders, mushrooms, puppies), and were asked to estimate the expected probability at which a specific category would be followed by stimulus shock. These expectancy ratings were given on a continuous visual scale below the picture (0–100%).

The illusory correlation-experiment lasted for about 35 minutes. Each of the 90 pictures was presented for 7.5 s. Of the 30 pictures of each category, 15 were followed by an electrical stimulus. Pictures were separated by a varying inter-trial-interval ranging from 12.5 to 15 s, showing a black fixation cross on a grey background (see Figure 1). The onset of every picture was jittered in 500-ms-steps relative to the beginning of a scan interval (2.5 s). The five jitter intervals were equally distributed to the six conditions (i.e., 3 [categories] X 2 [outcomes]) and randomly distributed within each condition. Stimuli were presented in one of six predefined pseudo-randomized orders. Across these orders, the condition of the starting stimulus was counter-balanced. Moreover, every picture was equally associated with a shock or nothing as an outcome across the six orders. In every six consecutive trials, all picture-outcome-combinations occurred. To keep attention focused on the task, on-line expectancy was rated throughout the experiment by pressing one of two buttons depending on which outcome (shock/nothing) was expected during picture presentation. The participants used the index and middle finger of the right hand (button-assignment was counter-balanced across participants). This procedure was practiced in three trials with letters instead of pictures and without electrical stimuli.

Outside the scanner, after the illusory correlation experiment, participants rated the aversiveness of the shocks and the a posteriori covariation estimates (e.g. "Given that there was a spider, how often (in %) did an electrical stimulus occur?") for each picture category (in randomized order) on visual scales (0–100[%]). To obtain a presumably more reliable measure of illusory correlation, post-experimental covariation estimates were also assessed on a trial-by-trial basis. Therefore, each picture was presented to the participants again and they were asked to decide whether a shock had been associated with this picture or not, using a scale ranging from -4 to +4 (without 0). The participants were instructed to select a positive number if they thought that the picture had been associated with a shock and a negative number if they thought that there had not been a shock. Values from 1 to 4 (or -1 to -4) served to express the certainty of their decision. Finally, the STAI-trait, the BDI-II, the SPQ and the FSQ were completed.

Apparatus and data analysis

Analysis of ratings. A priori covariation estimates were analyzed with 3x2 repeated-measures ANOVAs comprising the within-subjects-factor *picture* (spiders, mushrooms, puppies) and the between-subjects-factor *group* (phobics, controls). Trial-by-trial a posteriori covariation estimates were obtained by calculating the proportion of positive answers to the question whether a shock had been associated with a particular picture within each category depending on the actual outcome of the pictures. These proportions were analyzed with a 3x2x2 repeated-measures ANOVA comprising the within-subjects-factors *picture* (spiders, mushrooms, puppies) and *outcome* (shock, nothing) and the between-subjects-factor *group* (phobics, controls).

Scanning parameters. fMRI data were obtained using a 1.5 Tesla Siemens Avanto MRI Scanner. Functional data included whole-brain T2*-weighted single-shot gradient echo-planar images (EPI) recorded with a repetition time of 2.5 s (echo time = 30 ms, flip angle = 90°, field-of-view = 200 mm, acquisition matrix = 64 x 64, voxel-size = 3.1 x 3.1 x 5 mm). Each volume contained 25 axial slices parallel to the AC-PC-line (from the anterior to the posterior commissure) that were acquired in interleaved order. Slices were overlapping and 5 mm wide with a 1 mm gap. A total of 840 EPIs were recorded in every participant. The experimental procedure started only after the first eight EPIs to allow for a stabilization of the magnetic field. A high resolution structural image of the brain was created via T1-weighted magnetization-prepared rapid gradient-echo imaging (MP-RAGE; repetition time = 2250 ms, echo time = 3.93 ms, flip angle = 8°, field-of-view = 256 mm, acquisition matrix = 256 x 256, voxel size = 1 x 1 x 1 mm). If the magnetic field is inhomogeneous, EPI images are often spatially distorted (Hutton, Bork, Deichmann, Ashburner, & Turner, 2002). Therefore, a gradient echo (GRE) field mapping (TR: 1000ms, TE: 10 ms, slices: 25, slice thickness: 5mm, FOV: 240 mm, matrix size: 64 x 64) was performed prior to the acquisition of the functional MRI data to compensate for inhomogeneities of the magnetic field.

fMRI preprocessing. fMRI data were analyzed using Statistical Parametric Mapping software (SPM8; Wellcome Department of Imaging Neuroscience, London, UK). First, functional images were slice time corrected and realigned (second degree b-spline interpolation) using an individual voxel displacement map on the basis of GRE field mapping. The individual structural images were then coregistered to the mean individual functional image and segmented. Then functional images were spatially normalized into standard Montreal Neurological Institute (MNI) space using a voxel size of 2 x 2 x 2 mm, and smoothed with an 8 mm full-width-half-maximum (FWHM) Gaussian kernel. Unless indicated otherwise, these steps were performed with SPM8 default settings.

First level analysis. Twelve regressors of interest were defined regarding onsets of both pictures and outcomes of the six conditions: *spider-before-shock*, *spider-before-nothing*, *mushroom-before-shock*, *mushroom-before-nothing*, *puppy-before-shock*, *puppy-before-nothing*, *shock-after-spider*, *shock-after-mushroom*, *shock-after-puppy*, *nothing-after-spider*, *nothing-after-mushroom* and *nothing-after-puppy*. In addition, realignment parameters were included as nuisance regressors to compensate for movement artifacts during scanning. Event-related brain activation was modeled by convolving stick functions with the canonical hemodynamic response function (HRF). Parameter estimation was corrected for temporal autocorrelations using a first-order autoregressive model. The high-pass filter cutoff was set to 128 s.

Second level analysis: The main interest was focused on two kinds of brain activity: Responses to spider pictures and responses to shocks following spider pictures. To obtain fear-related modulation of shock processing, the interaction contrast [(*shock-after-spider* > *nothing-after-spider*) > (*shock-after-mushroom* > *nothing-after-mushroom*)] was calculated. This approach controls for residual picture-related activity and reveals activity in response to only the electrical shock. To obtain activity related to picture onset, the contrast [(*spider-before-shock* & *spider-before-nothing*) > (*mushroom-before-shock* & *mushroom-before-nothing*)] was created, from now on referred to as [*spider* > *mushroom*]. Since effects of fear-relevant stimuli in comparison to neutral stimuli were particularly relevant for the covariation bias, the pictures of puppies served as filler trials and were not part of this analysis. All analyses used a random-effects model for contrast maps of *t*-scores. Regions with significant activations are reported according to the automatic anatomic labeling in the Wake Forest University (WFU) PickAtlas (Maldjian, Laurienti, Kraft, & Burdette, 2003).

Effects of experimental conditions: DlPFC, ACC, amygdala and insula were chosen a priori (see introduction) as regions of interest (ROI) for the analysis of picture-related brain activation. The same regions plus the right paracentral lobule (PCL) as the primary sensory-motor area contralateral to painful stimulation (Roy et al., 2009) were chosen as ROI for the analysis of shock-related activity. ROI masks were created using predefined regions in the WFU PickAtlas. The dlPFC mask was defined as the combination of the lateral parts of Brodmann areas (BA) 9 and 46. That is the intersection of combined superior frontal gyrus and medial frontal gyrus on the one hand and combined BA 9 and BA 46 on the other hand. For ROI analyses, the alpha-error-level was set to $p < .05$, family-wise error (FWE) corrected, with a cluster threshold of $k \geq 5$ voxels. ROI analyses were carried out bilaterally with the exception of the right PCL in response to contralateral electrical stimulation and the amygdala which is mostly lateralized to the left hemisphere in response to emotional stimuli (Baas, Aleman, & Kahn, 2004). An FIR analysis was performed to validate that picture-related effects were present before the onset of outcomes (see section 'Finite impulse response (FIR) time course' for details). For whole-brain analyses, the threshold was set to $p < .001$, uncorrected, with a minimal cluster extent of $k \geq 10$ voxels. If several significant clusters were found inside one region in the whole-brain analysis, only the one with the lowest *p*-value is reported.

Psychophysiological Interaction (PPI). Functional connectivity was analyzed between the fear-modulated shock-response of individuals with spider phobia in sensory-motor area and other brain areas by running a PPI analysis. The contrast [*shock-after-spider* > *shock-after-mushrooms*] for participants with spider phobia was included as the psychological variable of interest. The right paracentral lobulus (PCL) served as the physiological variable (seed region). The principal eigenvariate within a sphere of 5 mm radius around the individual peak voxel of the right PCL was extracted. The psychophysiological interaction that is reported revealed regions which were active more synchronously in response to the shock following spiders than in response to the shock following mushrooms.

Finite impulse response (FIR) time course. For visual inspection of brain activity time-courses related to picture onset, a FIR analysis was conducted. FIR analysis does not make any assumptions about the shape of the BOLD response and reveals event-related activity estimates for single time points. To this end, the fMRI signal was re-estimated for ROIs that significantly differed in the contrast [*spider* > *mushroom*] according to

the analysis described above (see 'Effects of experimental conditions'). Using the rfxplot toolbox (Gläscher, 2009) the percent signal change of supra-threshold voxels ($p < .05$, FWE corrected, $k \geq 5$ voxels) was extracted for each repetition time from 2.5 s pre-stimulus to 15 s post-stimulus. In addition, the trial-by-trial covariation bias score was correlated with brain activity before the onset of the outcome obtained from FIR analysis. For this purpose, mean percent signal change across the three time points (2.5, 5, 7.5 s) during picture presentation was calculated for ROIs and subtracted between conditions (*spider* minus *mushroom*). This difference score was correlated with the difference between the proportions of positive answers ('There was a shock.') to spiders and mushrooms.

Brain-behavior correlation analysis. It was attempted to find out what brain activity predicted the tendency to overestimate the contingency between spider stimuli and electrical shocks. Marsbar (http://marsbar.sourceforge.net) was used to extract mean beta-values from all voxels in ROI with significant differences between conditions (see section 'Effects of experimental conditions'). Next Pearson correlations (two-tailed test with alpha = .05) were computed between these beta-values and the covariation bias score obtained from trial-by-trial covariation estimates after the experiment. This covariation bias score was calculated by subtracting the proportion of positive answers to mushroom pictures from the proportion of positive answers to spider pictures. In addition, for whole-brain analyses of brain-behavior correlations, the same covariation bias score was entered as a covariate in the second level analyses ($p < .001$, uncorrected; $k \geq 10$ voxels).

5.3. Results

Ratings

A priori covariation estimates. The main effect *picture* was significant, $F(2, 33) = 41.55$, $p < .001$, $\eta_p^2 = .72$, while there was no significant main effect of *group* ($p = .37$, $\eta_p^2 = .02$) and no significant interaction of *Picture* X *Group* ($p = .13$, $\eta_p^2 = .06$). This shows that individuals with spider phobia expected more electrical shocks after spiders ($M = 67.11$; $SD = 23.31$) than after mushrooms ($M = 29.06$; $SD = 19.26$), $t(17) = 6.29$, $p < .001$, $d = 1.48$, or puppies ($M = 23.94$; $SD = 17.89$), $t(17) = 6.72$, $p < .001$, $d = 1.59$. Healthy controls also expected more shocks following spiders ($M = 55.11$; $SD = 20.20$) than mushrooms ($M = 33.00$; $SD = 18.66$), $t(17) = 4.20$, $p < .001$, $d = 0.99$, and puppies ($M = 18.78$; $SD = 20.24$), $t(17) = 6.02$, $p < .001$, $d = 1.42$.

On-line expectancy ratings. On-line expectancy ratings were collapsed over 10 consecutive trials per category to approach normally distributed interval scale and allow for the conductance of a repeated-measures ANOVA with the within-subjects-factors *picture* (spiders, mushrooms, puppies), *block* (1-3) and the between-subjects-factor *group* (phobics, controls). The objective contingency per block was 50% for each category. The ANOVA revealed a main effect of picture, $F(2, 33) = 25.25$, $p < .001$, $\eta_p^2 = .61$, but no other main effects or interactions (all *p*s > .27). Accordingly, in patients expectancy was higher during spider trials ($M = .69$; $SD = .14$) than during mushroom ($M = .42$; $SD = .16$), $t_{17} = 6.16$, $p < .001$, $d = 1.42$, and puppy trials ($M = .39$; $SD = .22$), $t(17) = 4.07$, $p < .001$, $d = 0.94$. Also, in healthy controls expectancy was higher during spider trials ($M = .63$; $SD = .14$) than during mushroom ($M = .45$; $SD = .15$), $t(17) = 3.94$, $p < .01$, $d = 0.95$, and puppy trials ($M = .45$; $SD = .21$), $t(17) = 2.82$, $p < .05$, $d = 0.67$.

Global a posteriori covariation estimates. Assessments of the overall percentage of shocks for each picture category estimated after the experiment were analyzed in a repeated-measures ANOVA with the factors *Picture* (spiders, mushrooms, puppies) X *Group* (patients with phobia, controls). There was no significant main effect of *picture* (p = .10, η_p^2 = .14), *group* (p = .09, η_p^2 = .08) and no significant interaction (p = .28, η_p^2 = .08). Due to the many published reports of biased covariation estimates in spider phobia, differences were further explored between pictures within each group. There was only a non-significant trend in individuals with spider phobia indicating higher covariation estimates for spiders (M = 50.83; SD = 14.36) than for mushrooms (M = 41.40; SD = 19.37), $t(16)$ = 2.07, p = .06, d = 0.50, and puppies (M = 40.74; SD = 15.53), $t(16)$ = 1.89, p = .08, d = 0.46. No such trend was observable in healthy controls (all ps > .52).

Trial-by-trial a posteriori covariation estimates. There was a significant main effect of *picture*, $F(2, 32) = 5.44$, $p < .01$, $\eta_p^2 = .25$, and a marginal significant interaction of *Picture* X *Group*, $F(2, 32) = 2.69$, $p = .08$, $\eta_p^2 = .14$. Because of many previous findings of illusory correlations in spider phobia, the probability of finding a false positive effect should be low. Therefore, this marginal significant interaction was further examined. Patients with spider phobia overestimated the contingency between shocks and spiders (M = .67; SD = .16) relative to mushrooms (M = .43; SD = .17), $t(16) = 2.73$, $p < .01$, d = 0.92, and puppies (M = .46; SD = .20), $t(16) = 3.87$, $p < .05$, d = 0.68. In contrast, control participants did not overestimate the contingency between shocks and spiders (both ps > .43). In addition, there was a significant effect of *outcome*, $F(1, 33) = 5.39$, $p < .05$, $\eta_p^2 = .14$, a significant interaction of *Picture* X *Outcome*, $F(2, 32) = 4.49$, $p < .05$, $\eta_p^2 = .22$, and a significant interaction of *Outcome* X *Group*, $F(1, 33) = 5.39$, $p < .05$, $\eta_p^2 = .14$. The overall effect of *outcome* indicated that mean probability estimates were slightly higher if the picture had actually been associated with an electrical stimulus (M = .51 ± .21) than if it was not (M = .47; SD = .11), $t(34) = 2.14$, $p < .05$, d = 0.40. As stated above, this effect was further supported by interactions of *Picture* X *Outcome* and *Outcome* X *Group*. The difference between trials with shocks and trials without was significant for puppies (M = .07; SD = .19), $t(34) = 2.14$, $p < .05$, d = 0.37, marginally significant for spiders (M = .05; SD = .16), $t(34) = 1.80$, $p = .08$, d = 0.31, and not significant for mushrooms (M = -.01; SD = .14), $p = .68$. Regarding group differences, only spider patients with spider phobia were able to discriminate between outcomes (M = .07; SD = .10), $t(16) = 3.01$, $p < .01$, d = 0.70, but not healthy controls (M = .01; SD = .14), $p = .95$.

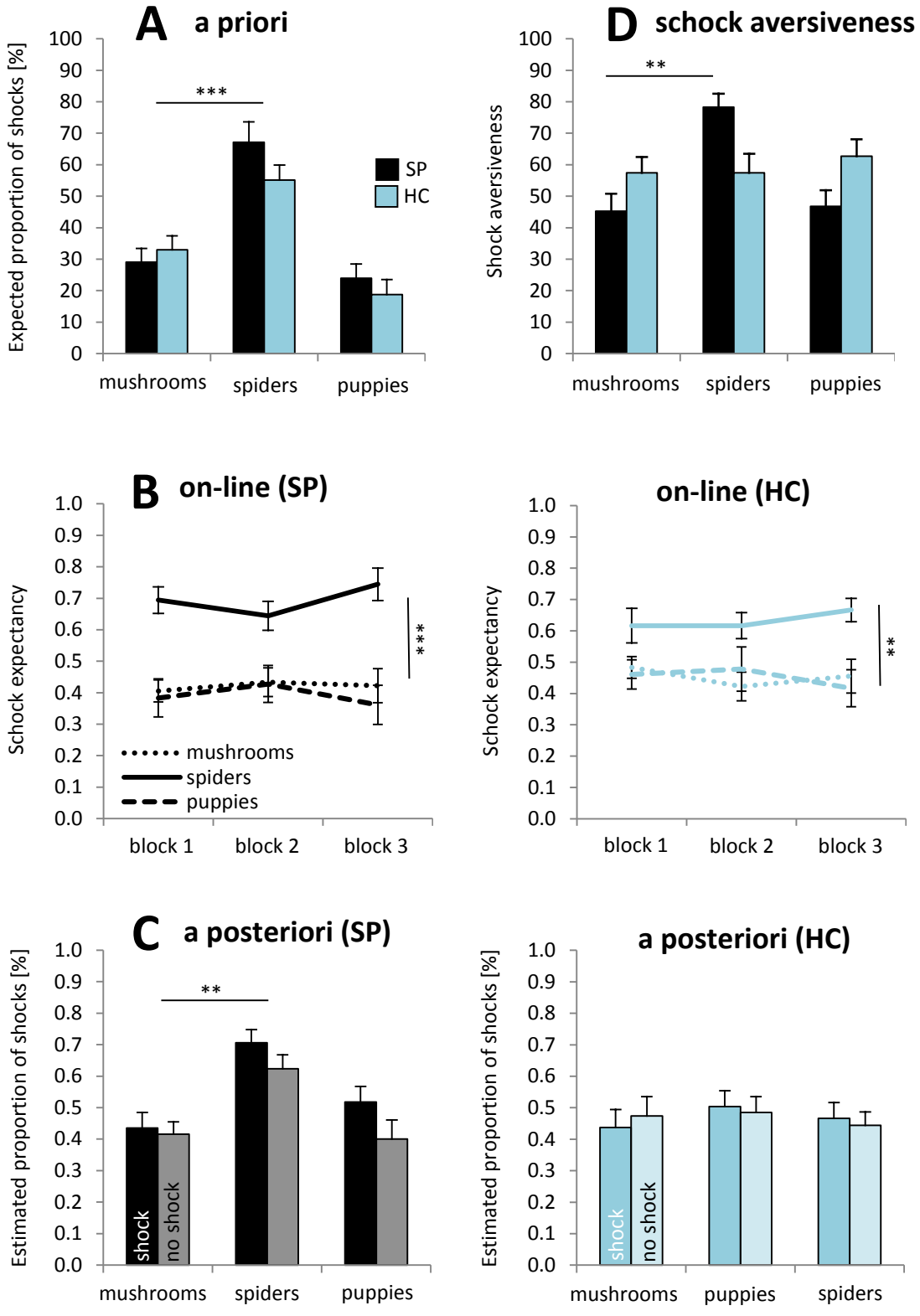

Figure 5.2. The diagrams show the expected proportion of shocks before (a priori), the expected proportion of shocks during the experiment (on-line), the estimated proportion of shocks after the illusory correlation experiment (trial-by-trial a posteriori ICs) and the perceived aversiveness of the shocks depending on the preceding picture category. For trial-by-trial a posteriori ICs, each picture was presented again and the participants were asked to indicate whether a shock had followed that particular picture. Dark (black/dark blue) and transparent (gray/light blue) bars display whether there had actually been a shock or nothing as an outcome. Both spider phobic and control participants expected more shocks following spiders before the experiment (A). Importantly, only spider phobic participants (C, left panel) but not healthy controls (C, right panel) overestimated this association after the experiment. ** $p < .01$; *** $p < .001$

Aversiveness. Aversiveness ratings were analyzed in the same way as global a posteriori covariation estimates. A significant main effect of *picture*, $F(2, 32) = 7.02$, $p < .01$, $\eta_p^2 = .31$, and a significant interaction of *Picture* X *Group*, $F(2, 32) = 10.11$, $p < .001$, $\eta_p^2 = .39$, were found which confirmed that spider phobics rated shocks following spiders ($M = 78.31$; $SD = 17.61$) as more aversive than shocks following mushrooms ($M = 45.29$; $SD = 22.59$), $t(16) = 4.00$, $p < .01$, $d = 0.97$, or puppies ($M = 46.81$; $SD = 21.09$), $t(16) = 4.21$, $p < .001$, $d = 1.02$. In the control group, aversiveness did not differ significantly depending on the preceding pictures (spiders: $M = 57.43$; $SD = 25.50$; mushrooms: $M = 57.46$; $SD = 21.23$; puppies: $M = 62.73$; $SD = 22.68$; all ps > .14).

Valence and arousal. Both kinds of picture ratings were analyzed with repeated-measures ANOVAs involving the factors *Picture* (spiders, mushrooms, puppies) X *Group* (patients with phobia, controls). For valence ratings there was a main effect of *picture*, $F(2, 32) = 107.00$, $p < .001$, $\eta_p^2 = .67$, a main effect of *group*, $F(1, 33) = 8.27$, $p < .01$, $\eta_p^2 = .20$, and an interaction of *Picture* X *Group*, $F(2, 32) = 50.68$, $p < .001$, $\eta_p^2 = .76$. Individuals with spider phobia rated spiders ($M = 1.96$; $SD = 0.69$) as more unpleasant than mushrooms ($M = 5.65$; $SD = 1.35$), $t_{16} = 12.48$, $p < .001$, $d = 3.03$, but not control participants (spiders: $M = 5.24$; $SD = 1.25$; mushrooms: $M = 5.09$; $SD = 0.91$; $p = .62$). Puppies (HC: $M = 6.84$; $SD = 1.32$; SP: $M = 7.28$; $SD = 0.97$) were rated as more pleasant than spiders and mushrooms by both groups (ps < .001). Only spider images were rated significantly more negative by patients with spider phobia than by healthy controls, $t(33) = 9.54$, $p < .001$, $d = 3.27$; both other $ps > .16$.

For arousal ratings there was a main effect of *picture*, $F(2, 32) = 59.79$, $p < .001$, $\eta_p^2 = .79$, and an interaction of *Picture* X *Group*, $F(2, 32) = 15.69$, $p < .001$, $\eta_p^2 = .50$. Individuals with spider phobia rated spiders ($M = 6.69$; $SD = 1.83$) as more arousing than mushrooms ($M = 2.07$; $SD = 1.10$), $t(16) = 9.99$, $p < .001$, $d = 2.42$. Also, the control group rated spiders ($M = 4.96$; $SD = 1.60$) as more arousing than mushrooms ($M = 3.30$; $SD = 1.24$), $t(17) = 4.88$, $p < .001$, $d = 1.15$, but spider patients with spider phobia rated the spiders as still more arousing than the control group, $t(33) = 2.97$, $p < .01$, $d = 1.01$. Both groups rated puppies (HC: $M = 4.58$; $SD = 1.82$; SP: $M = 3.11$; $SD = 1.46$) as more arousing than mushrooms (ps < .01). A significant effect of *group* for all picture types indicated that spider phobics rated all spiders as more arousing and mushrooms and puppies as less arousing (all ps < .01).

fMRI data

Brain activity during picture processing. In the spider phobia group, spider images elicited activation (*spider* > *mushroom*) in areas typically involved in the processing of phobia-relevant stimuli, including bilateral dlPFC. Moreover, activity within left amygdala, left ACC and left insula was found. Increased activation in the left dlPFC was also observed in the control group, but was still higher in spider phobics. Likewise, the left amygdala, the left ACC and the left insula were significantly stronger activated in the spider phobia group than in the control group (see Table 5.2 for ROI and Table 9.3 for whole-brain analysis).

Correlation between picture-related brain activity and covariation bias. Across spider phobia participants, correlation coefficients were calculated between mean activity in the relevant significant ROIs (*spider* > *mushroom*) and the trial-by-trial covariation bias. The correlation was significant for the left dlPFC, $r = .56$, $p = .02$, but not for the right dlPFC, the left amygdala, the left insula and the left ACC (see Table 9.4). In the

whole-brain analysis multiple additional brain regions were found to be correlated with the covariation bias (see Table 9.5). In patients with spider phobia, a cluster in the left middle frontal gyrus (BA 8) superior to our defined ROI of the dlPFC was the most prominent cluster correlating with the covariation bias score regarding cluster size and *t*-value.

Table 5.2. Significant Activations at the contrast [*spider* > *mushroom*]

	Region		MNI coordinates			k	*t*
			x	y	z		
Spider phobics	***ROI***						
	dlPFC (BA 9)	L	-20	44	42	12	6.84
		R	32	44	30	22	6.57
		R	18	56	32	8	5.56
		L	-36	32	36	8	5.41
	Amygdala	L	-18	-4	-12	22	4.84
		L	-28	-2	22	19	4.26
	ACC (BA 24)	L	-2	10	28	54	5.77
	Insula (BA 13)	L	-44	12	2	23	6.81
Control group	***ROI***						
	dlPFC (BA 9)	L	-12	48	30	14	7.10
Phobics > Controls	***ROI***						
	dlPFC (BA 9)	L	-34	30	36	13	4.86
	Amygdala	L	-28	-4	-20	6	3.62
	ACC (BA 24)	L	-2	-12	20	46	4.83
	Insula (BA 13)	L	-42	12	2	67	5.90

The table shows properties of peak voxels within a cluster. ROI threshold: $p < .05$ (FWE-corrected), $k \geq 5$; whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; BA = Brodmann area; PCL = Paracentral lobule; SMA = Supplementary motor area; dlPFC = dorsolateral prefrontal cortex; k = voxels in whole cluster

Time course analysis. Visual inspection of the finite impulse response (FIR) based analysis of the time course of activation related to picture presentation showed that the effects in all ROIs in the patient group were present prior to outcomes (see Figure 3). Inferential statistics were not applied to avoid circular testing (Kriegeskorte et al., 2009). Correlations between brain activity estimated with the FIR method prior to the end of the pictures and covariation bias estimates confirmed the result of correlations obtained with BOLD-modeled brain activity (see Table 9.4). Only left dlPFC was significantly correlated with the covariation bias.

Modulation of shock processing by phobic stimuli. For spider phobia patients, ROI analysis of the contrast [(*shock-after-spider* > *nothing-after-spider*) > (*shock-after-mushroom* > *nothing-after-mushroom*)] revealed significant activation in the right PCL. Whole-brain analysis returned an activation pattern mainly contralateral to electrical stimulation comprising right supplementary motor area (SMA), right postcentral gyrus, right fusiform gyrus, right cerebellum, right precentral gyrus, and left middle temporal gyrus (see table 5.3). For control participants, ROI analysis of the same contrast showed a significant cluster in the right insula. Whole-brain analysis indicated additional activity in right cerebellum and left cuneus. In a comparison between groups ([phobia > control]), ROI analyses did not confirm spider specific shock processing when using FWE-corrected statistics. Whole-brain analysis indicated that right dlPFC, right SMA, left PCL and left parahippocampal gyrus

were more active in the spider phobia group than in the control group (all *p*s < .001). At a more lenient threshold of $p < .005$, $k \geq 10$ (Lieberman & Cunningham, 2009), a significant difference in the right PCL ($t = 2.99$; $k = 14$; x = 8, y = -40, z = 58) was also found. There was no significant activity in the reverse contrast ([controls > phobics]), even at this less conservative threshold.

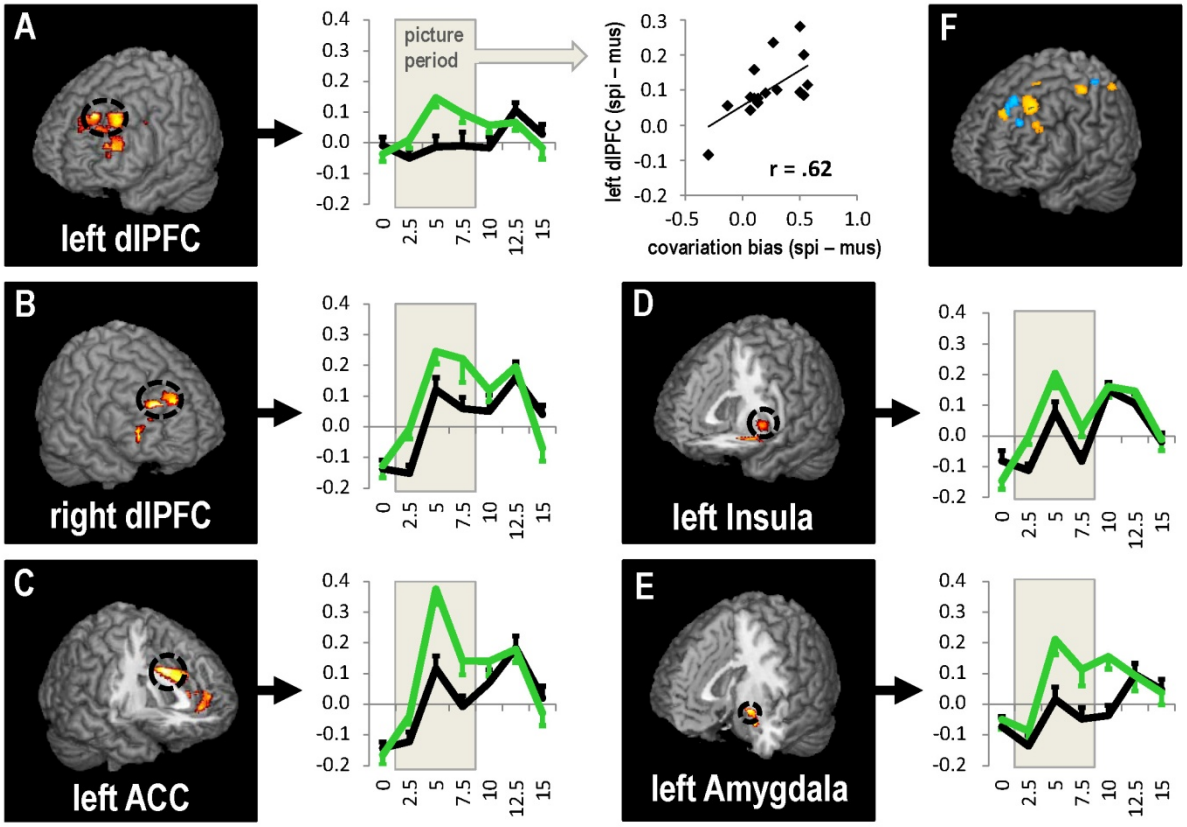

Figure 5.3. Spider phobic patients' responses to picture onsets. The five regions of interest (ROIs; A-E) with significant differences in the contrast (spider > mushroom) are depicted (ROI-analysis: p < .05, FWE-corrected; k ≥ 5 voxels; display threshold: p < .005, uncorrected; k ≥ 10 voxels). Line diagrams show the time course of percent signal change for spider (green) and mushroom (black) pictures during picture presentation (grey background from 0 to 7.5 s) in spider phobia. The scatter plot (A) visualizes the correlation between the difference in brain activity during picture presentation and the trial-by-trial IC after the experiment. In the upper right corner (F), clusters that correlated with the trial-by-trial IC can be seen mainly in left fronto-parietal regions including the dlPFC, Brodmann area 8 and the superior parietal cortex (spider phobia = yellow; healthy controls = blue; whole-brain analysis: p < .001, uncorrected; k ≥ 10 voxels).

Table 5.3. Significant Activations at the contrast [(*shock-after-spider* > *nothing-after-spider*) > (*shock-after-mushroom* > *nothing-after-mushroom*)]

	Region		MNI coordinates			k	*t*
			x	y	z		
Spider phobics	***ROI***						
	PCL (BA 6)	R	8	-26	60	13	4.59
	whole-brain						
	SMA (BA 6)	R	-12	-32	60	386	5.61
	Postcentral (BA 40)	R	50	-30	52	72	5.19
	Fusiform (BA 20)	R	28	-28	-28	40	4.48
	Cerebellum	R	10	-64	-42	44	4.46
	Precentral (BA 4)	R	36	-26	64	15	4.19
	Middle temporal (BA 37/20)	L	-52	-38	-14	12	4.13
Control group	***ROI***						
	Insula (BA 13)	R	44	-10	-4	7	6.00
	whole-brain						
	Cerebellum	R	8	-50	-36	29	5.63
	Cuneus (BA 7)	L	-14	-80	26	16	4.64
Phobics > Controls	***whole-brain***						
	Inferior frontal (BA 9/46)	R	56	22	30	59	4.85
	SMA (BA 6)	R	10	-18	56	67	4.39
	PCL (BA 6)	L	-12	-30	60	58	4.22
	Parahippocampal (BA 28)	L	-20	-20	-26	25	4.17

The table shows properties of peak voxels within a cluster. ROI threshold: $p < .05$ (FWE-corrected), $k \geq 5$; whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; BA = Brodmann area; PCL = Paracentral lobule; SMA = Supplementary motor area; dlPFC = dorsolateral prefrontal cortex; k = voxels in whole cluster

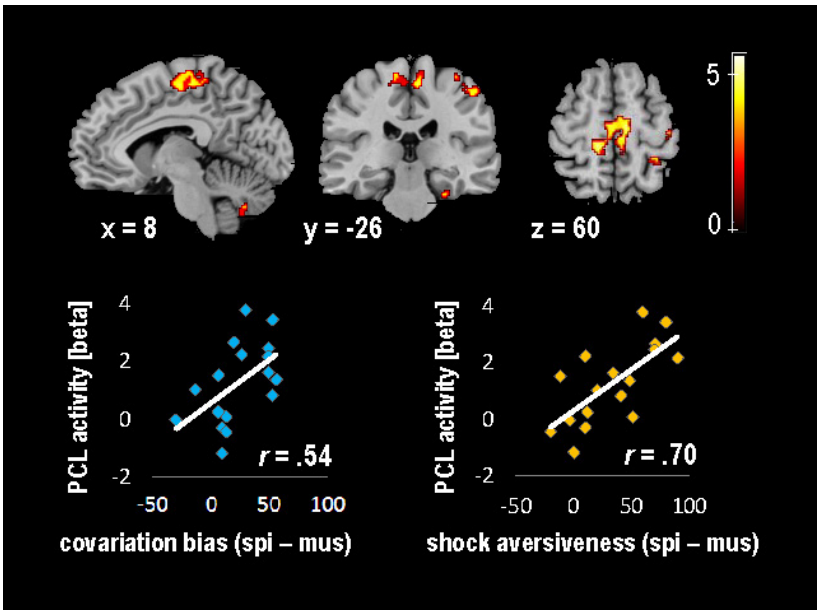

Figure 5.4. Activity in the paracentral lobule (PCL) of spider phobic patients and correlations with IC and ratings of shock aversiveness. The brain slices correspond with the peak voxel in right PCL (ROI-analysis: p < .05, k = 5 voxels) and show significant activation for the contrast [(shock-after-spider > nothing-after-spider) > (shock-after-mushroom > nothing-after-mushroom)] (p < .001, k = 10 voxels for display purpose). Scatterplots depict the association between right PCL activity (mean beta values of all voxels in the ROI) and trial-by-trial IC (left), resp. the rated aversiveness of the electrical shock (right) depending on the picture type.

Correlation between shock-related brain activity and covariation bias. In the spider phobia group, mean right PCL activity in the contrast [(*shock-after-spider* > *nothing-after-spider*) > (*shock-after-mushroom* > *nothing-after-mushroom*)] was significantly correlated with trial-by-trial covariation bias, r = .54, p = .02. Right PCL response was also correlated with the difference between subjective aversiveness of shocks following spiders and shocks following mushrooms, r = .70, p = .002. Moreover, covariation bias and the difference in aversiveness were significantly correlated, r = .67, p = .003.

In a whole-brain analysis (see Table 9.1), additional regions correlating with covariation bias were found: left precentral gyrus, left PCL, right precentral gyrus, right cerebelum and right fusiform gyrus. Again, at the more lenient threshold of p < .005, k ≥ 10, a significant cluster in the right PCL emerged (t = 3.68; k = 26; x = 4, y = -24, z = 68).

Functional connectivity analysis. An analysis of psychophysiological interactions (PPI) revealed that the heightened activity in contralateral primary sensory-motor area (PCL) in response to shocks following phobia-relevant stimuli was accompanied by an enhanced connectivity to a large network of several brain regions. Particularly, in reaction to the electric stimulus subsequent to spider pictures relative to mushroom pictures the connectivity was stronger in the right midbrain (expanding to right hippocampus), left and right cerebellum, left and right supramarginal gyrus, left ACC, right frontal inferior orbital gyrus, left and right mid cingulum, left and right insula, right lingual gyrus, right SMA, right rolandic operculum, left and right calcarine gyrus, middle temporal gyrus, right superior temporal gyrus, right precentral gyrus, left middle frontal gyrus, see Table 9.2 and Figure 5.5.

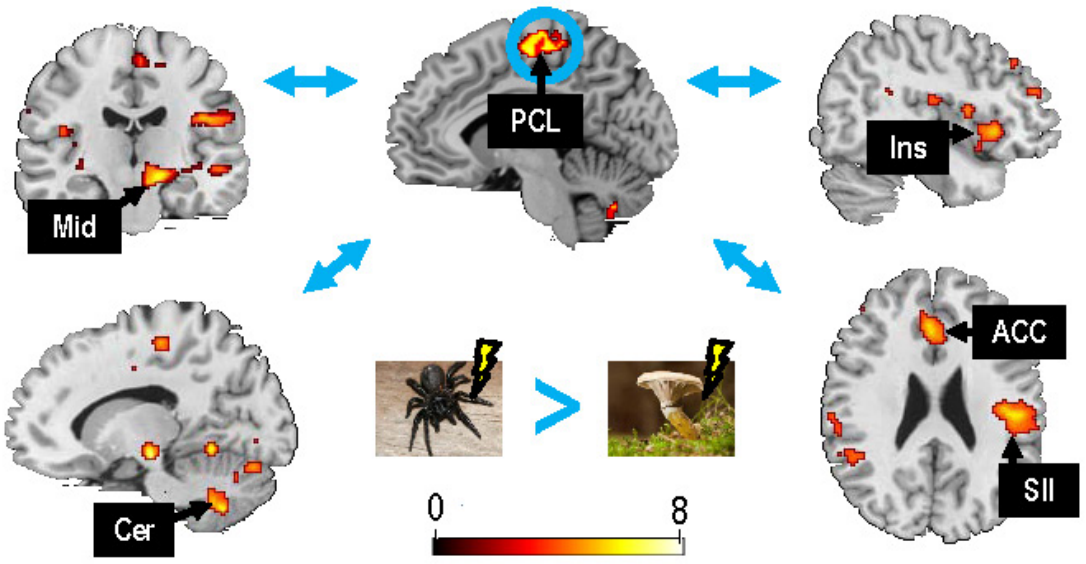

Figure 5.5. Enhanced connectivity between right paracentral lobule (PCL) and other brain regions in response to shocks following spider images relative to shocks following mushroom images in patients with spider phobia. The color bar is only valid for the four slices showing connected areas outside the seed region in PCL. Mid = Midbrain; Ins = Insula; Cer = Cerebellum; ACC = anterior cingulate cortex; SII = secondary somatosensory cortex.

5.4. Discussion

This study confirmed previous findings that both individuals with spider phobia and healthy controls expect spiders to be associated with painful shocks. Importantly, only phobics still overestimated the contingency between spider images and shocks after being exposed to a series of stimuli with random contingency between image and shock. This illusory correlation in phobic patients seems to be based on an abnormal processing of pictures and outcomes, as the present fMRI study revealed. Spider images led to increased recruitment of dlPFC, a region that has consistently been associated with executive control. Shocks following spider images were perceived as more aversive and evoked enhanced responses in primary sensory-motor area (PCL). Both PCL and dlPFC activity predicted the tendency to overestimate the contingency between spiders and shocks within spider phobics.

The finding that left dlPFC activity was enhanced in response to phobic stimuli and correlated with the covariation bias suggests that altered executive control or working memory processes may play a role in the maintenance of fear-relevant illusory correlations. Increased dlPFC activity is a common finding in phobic symptom provocation and has been proposed to reflect attempts of emotional down-regulation (Paquette et al., 2003). Considering the highly threatening impression of spiders on spider phobic patients, it is reasonable that the brain mobilizes executive control mechanisms that should help to cope with the threatening situation. An increased occupation of working memory load by phobic stimuli may be one of the factors that prevent individuals with spider phobia from correcting their notion that spider images are often followed by shocks. Previously, it has been demonstrated that the dlPFC is crucial for monitoring CS-US contingencies in fear conditioning (Carter et al., 2006). Moreover, it has been shown that a reduction in dlPFC activity in response to a working memory task can be evoked by emotional distraction and predicts task performance (Dolcos & McCarthy, 2006). Other investigators found that dlPFC was enhanced during an emotional Stroop task (Compton et al., 2003). This means that the participants showed increased activity in the dlPFC, if they had to name the color of emotionally negative words. Also, patients suffering from panic disorder display increased prefrontal activity associated with working memory, when a Stroop task is performed with panic-related words (Dresler et al., 2012). This is a somewhat similar situation as it is here, because a cognitive task has to be performed while an emotional stimulus is presented. These results further support the assumption that emotional distraction may be the reason for the hyperactivity in the dlPFC during phobia-relevant slides. Finally, dlPFC is a key structure for the overcoming of habitual strategies (Knoch, Brugger, & Regard, 2005) and may therefore also be important to abandon an a priori covariation bias.

Whole brain analyses of brain regions correlating with the covariation bias also revealed a wider definition of dlPFC as potentially involved in the generation of a covariation bias (see Table 9.5). Particularly, the left BA 8 or frontal eye-field (Paus, 1996) was correlated with the covariation bias in both spider phobics and healthy controls. Moreover, the superior parietal gyrus was correlated with the covariation bias in both groups. These regions are part of a fronto-parietal attention network (FPAN) that serves to identify high priority stimuli in the environment (Ptak, 2012). For example, shifting attention in space to visual and auditory stimuli activates the posterior parietal cortex and the frontal eye field (Smith et al., 2010). In a recent review (Ptak, 2012), it has

been suggested that the FPAN is necessary for the interaction of working memory and visual information, for example working memory load (operationalized by the number of digits that participants had to keep in mind) predicts fronto-parietal activity during visual information processing (Coull, Frith, Frackowiak, & Grasby, 1996). Moreover, it has been suggested that visual processing can be enhanced by FPAN activity (Ptak, 2012). In relation to emotional stimuli, a fronto-parietal network is involved in attentional orienting to fear-conditioned faces (Armony & Dolan, 2002) and in the viewing of emotional scenes (Moratti, Keil, & Stolarova, 2004). Regardless of stimulus modality, the FPAN seems to be involved in sustained attention, also without shifting attention focus (Pardo, Fox, & Raichle, 1991). The late positive potential (LPP), which is a commonly found electroencephalographic potential modulated by emotional arousal, is correlated with brain activity in occipital and parietal regions (Sabatinelli, Lang, Keil, & Bradley, 2007). Stimulation of the dlPFC causes attenuation of the LPP response to aversive stimuli, suggesting that this region is able to down-regulate parietal and visual emotional responses (Hajcak et al., 2010). In general, the FPAN seems to enable amplified processing of emotional stimuli in sensory cortices (Moratti et al., 2004; Vuilleumier & Huang, 2009). In the present experiment, occipital visual regions were also correlated with the covariation bias in patients and healthy controls.

Taken together, ROI and whole brain analyses of picture-related activity predicting illusory correlations implicate that a fronto-parietal attention network (comprising the dlPFC, the frontal eye field and superior parietal gyrus) was engaged while the participants were simultaneously viewing the pictures and trying to estimate the probability of a shock outcome. Based on previous findings on the FPAN, it is very likely that such activity was necessary for the task of expectancy ratings and attending to stimulus-outcome combinations. In the presence of distracting emotional images (i.e. spiders) this activity might have been enhanced due to a competition between emotional processing and the cognitive task. The degree to which the attentional network was activated relative to the neutral condition thus may have reflected the degree of emotional distraction, and this emotional distraction prevented some participants from doing the cognitive task well, i.e. estimating contingencies accurately. Therefore, this activity might have correlated with the covariation bias. Frontal and parietal activity was lateralized to the left side in both groups, possibly as a result of the cognitive task being performed with the right hand or because of the involvement of a phonological loop (Coull et al., 1996). As outlined above, one might assume that enhanced dlPFC activity reflects emotional distraction preventing an update of inflated expectations because there are less cognitive resources available to encode the real contingencies. As an alternative explanation, enhanced dlPFC activity might reflect elevated expectancy of shocks as outcomes. Then better instead of worse encoding of shocks as outcomes could have contributed to the covariation bias. A differentiation between these explanations is not possible on the basis of the present study.

The result that phobic images amplified the aversiveness of painful electrical stimuli replicates earlier findings (Tomarken et al., 1989), and is in line with increased pain perception under negative affect (Kenntner-Mabiala & Pauli, 2005; Kenntner-Mabiala et al., 2008; Roy et al., 2009). To the authors' knowledge, this is the first study to analyze the modulation of brain responses to painful stimuli by phobic fear. Similar to the modulation by negative affective images (Roy et al., 2009) increased activity in primary sensory-motor area (PCL) was found. In this circumstance, it was a novel approach here to carefully disentangle picture and pain responses

by subtracting the sole emotional experience (pictures without shocks) from the emotional modulation of pain (pictures with shocks).

Importantly, in phobics the PCL activity was positively correlated with the aversiveness of shocks and the covariation bias after the experiment. This demonstrates that the relationship between the enhanced aversiveness of shocks and covariation bias, as previously found and replicated here, is more than just a demand effect at the moment of the rating. In fact, this relationship is manifested in biased sensory processing at the moment of shock application. Increased PCL activity may reflect increased attention to the shock. Previous studies showed that activity in primary sensory cortex is reduced when attention is directed away from a painful stimulus (Bushnell et al., 1999). If increased PCL activity reflects attentional engagement, increased attention to shocks following spiders may lead to deeper encoding which in turn causes the overestimation of shocks following spiders. In consistence with this idea, the primary somatosensory cortex is not only important for sensory on-line processing, but probably also serves as a storage site for tactile information (Harris, Miniussi, Harris, & Diamond, 2002; Pasternak & Greenlee, 2005). It cannot be ruled out that enhanced PCL activity (also) reflects a suppressed motor response. This notion is supported by accompanying activity in SMA and the cerebellum. Aversive stimuli following phobic images may trigger a flight or avoidance tendency that is correlated with covariation bias. Again, this reaction could have initiated deeper encoding. In any case, the heightened aversiveness of the shock can be reason enough for the maintenance of illusory correlations. After all, overestimating the occurrence of aversive outcomes could be a consequence of preparing for the worst case, and the necessity of being prepared should be higher when the potential outcome is more aversive. In experiment 4 of the present thesis, it was showed that the experimental manipulation of the aversiveness of outcomes is sufficient to induce a covariation bias between neutral cues and aversive outcomes (Wiemer, Mühlberger, & Pauli, 2014; see experiment 4).

The finding that outcome aversiveness was associated with the illusory correlation is in accordance with previous investigations showing that an uncertainty cue leads to enhanced aversiveness of negative pictures relative to negative pictures cued with certainty (Grupe & Nitschke, 2011). In an fMRI experiment, researchers from the same laboratory demonstrated that uncertain negative pictures evoked stronger amygdala and insula responses than certain negative pictures (Sarinopoulos et al., 2010). Interestingly, the frequency of uncertain negative pictures was overestimated by most participants. This kind of uncertainty-related covariation bias was predicted by high anticipatory activity in the ACC, and low insula activity during picture processing. None of these brain regions was correlated with the illusory correlation in the present experiment, suggesting that different brain mechanisms might be involved in the generation of uncertainty-related illusory correlations with pictures and fear-related illusory correlations with pain. One important aspect for example is that pictures in the present experiment might also be considered as uncertainty cues because they were followed by an aversive outcome in only half of the trials, but ACC activity was very likely also involved in fear responses to the spider images. Therefore, the ACC does not solely reflect anticipatory processes in the present experiment which might be a reason why it did not predict the covariation bias. However, the advantage of this approach is that brain mechanisms involved in the processing of fear-relevant stimuli and correlating with fear-relevant illusory correlations can be identified.

Furthermore, patients' PCL response to shocks after spiders was strongly synchronized with a large bilateral network of brain regions involving key components of a 'pain matrix' or salience detection system, such as insula, ACC, secondary somatosensory cortex and midbrain (Legrain et al., 2011). In phobics, this connectivity was more pronounced for shocks after spiders relative to mushrooms, suggesting that the salience system was stronger involved in the processing of fear-relevant outcomes. A very prominent cluster that was revealed in this connectivity analysis was one with peak activation in the right midbrain extending to the right hippocampus. Some regions in the midbrain, such as the ventral tegmental area (VTA), have also been found to be involved in fear conditioning (Fadok, Dickerson, & Palmiter, 2009; Pezze & Feldon, 2004). The VTA supplies the meso-cortico-limbic pathway with dopamine and was found to play an important role in reward processing and learning, but is also important for the generation of conditioned and unconditioned fear responses. For example, when dopamine neurons in the VTA are activated, cats produce fear-like startle reactions (Sevens, Wilson, & Foote, 1974), and when dopaminergic activity is prevented in the VTA, cue-dependent fear conditioning is not possible anymore in mice (Fadok et al., 2009). In the light of the cluster involving both the midbrain and the hippocampus, it is noteworthy that a neural loop has been identified between these structures that is thought to support memory formation of novel and salient information (Lisman & Grace, 2005). Particularly, a novelty signal is carried from the hippocampus to the VTA and via the other arm of the loop the VTA enhances long term potentiation in the hippocampus via dopamine release, thus promoting memory processes. There is also growing evidence that the hippocampus and the midbrain are important for the generation of prediction error signals in associative learning (Delgado, Li, Schiller, & Phelps, 2008; Goosens, 2011). Considering that the participants were asked to give expectancy ratings throughout the experiment, it may be speculated that shocks after phobia-relevant pictures produced enhanced prediction error signals in the brain and so contributed to biased memory encoding. In other words, biased midbrain-hippocampus interactions in response to aversive outcomes may subserve the formation of fear-relevant illusory correlations in addition to or as a potential mediating process of the effect of the increased aversiveness of the outcome. Future studies using fMRI with higher spatial resolution to differentiate between midbrain regions and a systematic manipulation of prediction errors are necessary to support these speculations.

The following limitations of the present study should be considered when interpreting the results. First, patients with spider phobia and healthy controls did not only differ in fear of spiders, but also in state anxiety, trait anxiety, by trend in depressiveness and also in the objective intensity of the electrical stimulus. So, strictly speaking it cannot be ruled out that the current results were mediated or caused by one of these factors and not only specifically by fear of spiders. However, elevated state anxiety should be expected, when individuals with spider phobia anticipate to take part in an experiment involving the viewing of spider pictures and can just be recognized as a manipulation check that ensured that spider phobics were more anxious than healthy controls. Obviously, this elevated anxiety might not have been restricted to spider trials, but was perhaps sustained throughout the experiment. Higher trait anxiety and slightly higher depressiveness might also be the consequence of elevated state anxiety in that situation. Although, these questionnaires are thought to measure an invariable personality trait, it is not unlikely that the perception of one's own personality characteristics is influenced by the fact that one has just experienced an episode of heightened fear. Indeed, there is wide agreement that there are situational effects on the assessment of personality traits (Deinzer et al., 1995). Future

investigations should assess trait measures independent from the recruitment for an experiment about spider phobia. Besides that, since the results of the present study were associated with stimuli specifically associated with participants' fear, it should be unlikely that general anxiety can explain the results. Also, the fact that spider phobics had a lower pain threshold than healthy controls fits within that pattern of elevated anxiety. This result is interesting with respect of its own and suggests that anxiety leads to increased pain perception, which was also found by previous investigators (Rhudy & Meager, 2000). Moreover, such a systematic difference is exactly one of the reasons why the pain threshold is adjusted for each individual. The individual determination of the pain threshold ensures that despite variable pain sensitivity, all participants perceive the stimulus as comparably equal on a subjective level. Indeed, the groups did not differ on subjective pain ratings of the electrical stimulus. Finally, in all measures of shock processing (i.e. aversiveness, brain activity), spider trials were always compared with mushroom trials within each group. So, the same objective intensity of the electrical stimulus always served as its own control condition, which should attenuate or even eliminate the influence of objective pain intensity. As a second limitation, the a posteriori covariation bias was not significant in global covariation estimates. This may be because used 90 trials were used to enable reliable measurement of the BOLD-response. Previous studies typically used 72 trials. Moreover, participants were asked to rate outcome expectancy throughout the experiment to ensure that they were attending the contingencies. This may have led to more accurate global covariation estimates. However, on a trial-by-trial basis, an illusory correlation was significant in individuals with spider phobia. This conceptualization of an illusory correlation should be more reliable due to repeated measurement than a global estimation based on two questions. Besides, the assumption that both variables reflect very similar concepts is supported by the fact that trial-by-trial and global illusory correlations were correlated within both groups [10]. Finally and importantly, it should be similarly relevant for fear maintenance, because it may reflect the tendency to expect an aversive stimulus when confronted with a spider again. Third, in the present traditional illusory correlation paradigm, it cannot be disentangled whether the enhanced dlPFC activity in phobia-relevant trials can be attributed to the processing of spider images or to anticipatory processes related to outcomes. Alternative experimental procedures are needed to further reveal the exact function of the dlPFC in illusory correlations.

In summary, illusory correlations in spider phobia were found to be associated with biased neural processing of pictorial cues and aversive outcomes. In particular, enhanced aversiveness and sensory-motor processing of negative outcomes following spider images and enhanced activation of the dlPFC by phobic stimuli predicted covariation bias in spider phobia. These results further contribute to a neurobiological explanation of fear maintenance besides classic fear conditioning. On a clinical level, the findings suggest that the change of the evaluation or experience of feared consequences and the direction of executive resources away from phobic stimuli to non-feared consequences should give rise to a less dangerous cognitive representation of phobic stimuli and thus help to reduce fear and anxiety.

[10] Within healthy controls, the correlation between the two illusory correlation measures was $r = .79$, $p < .001$, within spider phobics $r = .56$, $p < .01$.

6. Experiment 4: Causal impact of outcome aversiveness on illusory correlations

6.1. Introduction

Experiment 1, 2 and 3 have shown that a covariation bias between fear-relevant or emotionally negative stimuli and aversive consequences correlates with the enhanced subjective aversiveness of aversive consequences following negative stimuli. In addition, experiment 3 demonstrated that such a correlation is manifest in biased sensory processing of aversive consequences following phobia-relevant stimuli during the illusory correlation procedure. So far, these results remain correlational in nature and it is not clear, whether the subjective aversiveness is really causally involved in the maintenance of fear-relevant illusory correlations, as the model of fear-relevant illusory correlations would predict (chapter 2.5). If the aversiveness has a causal influence on illusory correlations, it would be reasonable point of intervention to help anxious individuals to get a more accurate picture of the dangerousness of feared objects and finally to fight fear and anxiety. Therefore, differently aversive outcomes were used in experiment 4 and paired with neutral visual stimuli in an illusory correlation experiment, to test if this would lead to illusory correlations. The following experiment 4 has been published in the journal *Cognition and Emotion* (Wiemer et al., 2014). The following report on this experiment is an almost identical version of this publication.

Learning associations between temporally related stimuli and threatening events is essential for physical health and survival. However, inferring the objective likelihood that a specific cue predicts an aversive event on the basis of contingencies is not easy and errors may be fatal. The underestimation of the likelihood of a threatening event may have negative consequences while overestimating the occurrence of aversive events mostly has little costs and therefore may even be beneficial, a principle known as adaptive conservatism (Hendersen, 1985). From this point of view, people should be more likely to expect an aversive event in a situation in which the aversiveness of the event is especially high. This may also lead to an overestimation relative to objective contingencies or relative to situations with less aversive outcomes. Such a cognitive bias could be called an *illusory correlation* or a *covariation bias*, and may confirm the fear of the aversive outcome. For example, experiencing turbulences on an airplane may be highly aversive for someone who is afraid of flying. Overestimating the occurrence of turbulences will help to avoid them because next time when s/he will decide to get on an airplane s/he will be more likely to expect them. Since we all want to avoid highly aversive situations, heightened expectancies should be induced by outcome aversiveness. Generally speaking, if the costs of an outcome are high, it is necessary to be prepared for it, and this should be reflected in inflated expectancies. In other words, the aversiveness of an outcome may induce an illusory correlation between a neutral stimulus and the aversive event, meaning that the likelihood of the aversive event given the presence of this stimulus is overestimated relative to the objective contingency.

Illusory correlations would explain why anxiety disorders are often characterized by irrational cognitions about the occurrence of aversive consequences given the presence of fear-relevant stimuli. Moreover, they may serve as a model for the onset and maintenance of fear in addition to classical fear conditioning. Such theoretical approaches are particularly desirable for the explanation of pathological fear since conditioned fear responses do not seem to clearly discriminate between anxious and non-anxious individuals (Beckers et al., 2012).

Previously, illusory correlations have been observed in animal phobia (e.g., Tomarken et al., 1989; de Jong, Merckelbach, & Arntz, 1995), flight phobia (e.g., Pauli et al., 1998; Mühlberger et al., 2006), social anxiety disorder (Hermann et al., 2004), and panic-prone individuals (e.g., Pauli et al., 1996; 2001). A close relationship between covariation bias and phobia maintenance has been shown. Spider phobic patients who displayed a covariation bias after treatment were more likely to suffer from a relapse (de Jong, van den Hout, & Merckelbach, 1995) and the treatment significantly reduced both harm-related and contamination-related expectancy bias in spider phobics (van Overveld et al., 2010).

In a classic illusory correlation experiment (e.g., Tomarken et al., 1989), phobic participants and healthy controls are exposed to phobia-relevant (e.g., pictures of spiders or snakes) and neutral pictures (e.g., pictures of mushrooms or flowers) followed by three types of outcomes: nothing, neutral tones, and mildly aversive electric shocks. Although contingencies between types of pictures and outcomes are overall random, phobic participants overestimate the covariation between phobia-relevant pictures and aversive outcomes after the experiment (a *posteriori covariation bias*) as well as during the experiment (*on-line covariation bias*). Fearful individuals also overestimate the contingency between phobia-relevant stimuli and an aversive outcome before *(a priori covariation bias or expectancy bias)* an illusory correlation experiment (McNally & Heatherton, 1993; Davey & Dixon, 1996).

Interestingly, while a priori expectancy biases for spider stimuli are present in both healthy and spider phobic participants, a posteriori covariation biases are mainly found in spider phobic participants (Davey & Dixon, 1996; Mühlberger et al., 2006). So, why are illusory correlations more persistent in fearful individuals? More knowledge about underlying mechanisms of illusory correlations could help to understand the resistance to correct the expectancy bias on the basis of objective contingencies in fearful persons and thus may help to understand how a covariation bias contributes to the development of anxiety disorders. One contributing factor may be the difference between fearful and non-fearful participants in the experienced aversiveness of the outcomes. Similar to how fear conditioning is facilitated through increased aversiveness of the unconditioned stimulus in animals and humans (Sigmundi, Bouton, & Bolles, 1980; Wolter & Lachnit, 1993), the covariation bias may be based on the exaggerated experience of aversiveness triggered by the aversive outcomes. Our assumption is based on the evidence confirming the affective priming hypothesis (Lang, Bradley & Cuthbert, 1990) that negative and fear-relevant stimuli enhance the affective responses to noxious electric stimuli (Rhudy & Meagher, 2001; Kenntner-Mabiala & Pauli, 2005; Kenntner-Mabiala et al., 2008). In fact, Tomarken et al. (1989) already reported that the same electric shock was considered as more painful when it followed spider or snake pictures than when it followed neutral pictures. In experiment 3 of this publication, participants rated the painfulness of the electric shock for each slide category at the beginning and at the end of the procedure. At the beginning, high-fear participants indicated enhanced painfulness ratings for shocks associated with fear-relevant

slides relative to neutral slides. At the end of the procedure, both low- and high-fear participants indicated enhanced painfulness ratings for shocks after fear-relevant slides. In a concomitant regression analysis, they found that the contingency estimates of the low-fear participants were significantly predicted by their pain ratings. Similarly, the illusory correlation experiment by Mühlberger et al. (2006) revealed that spider phobics rated the physically identical aversive startle tones as more unpleasant when it followed spider slides than when it followed mushroom slides. In contrast, flight-phobics rated startle tones following pictures of airplane crashes as equally unpleasant as startle tones following mushrooms. This difference may have caused that the phobia-relevant covariation bias in spider phobics was more resistant to learning than in flight-phobics indicating an association between covariation bias and experienced aversiveness of the consequence.

In accordance with these thoughts, the first three experiments of the present thesis in fact demonstrated that the covariation bias between negative stimuli and aversive outcomes were correlated. That is, participants who perceived the startle sound or the electrical stimulus after negative pictures as more aversive than after neutral pictures were also likely to be those who had a covariation bias between negative pictures and aversive outcomes. Importantly, experiment 3 showed that this relationship might be mediated by biased neural processing of outcomes during the illusory correlation paradigm and is not only present in the moment of the assessment of covariation estimates at the end of the experiment. Particularly, the representation of the aversive outcome in sensory-motor areas was enhanced after phobia-relevant spider images and correlated with the fear-relevant illusory correlation in spider phobics.

It is important to note that these findings are correlational and therefore can be explained in two ways. On the one hand, the fear-relevant pictures may induce a covariation bias and the resulting increased expectancies of aversive outcomes may amplify the experienced aversiveness of the outcomes (Dowman, 2001). In other words, expectancy would drive experience. On the other hand, the fear-relevant pictures may increase the experienced aversiveness of the outcomes itself and this increased aversiveness causes the covariation bias. In other words, experience would drive perceived causality. Although, most authors favor the former assumption the latter has not yet been experimentally investigated. It is plausible that the more aversive an outcome is experienced the more likely it is causally attributed to a preceding event. This would mean that when the potential cost of an outcome is high it demands better preparedness for its future appearance and this may be reflected in an overestimation of its occurrence. In consequence, it should be possible to induce a covariation bias between a neutral stimulus and aversive outcomes by merely manipulating the outcome aversiveness.

In order to test this hypothesis, an illusory correlation experiment was conducted with colored geometric figures as neutral stimuli that were randomly followed by either aversive outcomes or nothing. Importantly, the aversiveness of the outcomes was manipulated in a within-subjects-design, so that each of three colors of the geometric figures predicted a different level of outcome aversiveness, i.e. startle sounds of different intensities. Objectively, the probability of an aversive outcome was always 50 percent for each color. Participants were asked to rate their expectancy of an aversive outcome in each trial (on-line covariation estimates) and after the experiment (a posteriori covariation estimates). It was anticipated that the covariation bias for a stimulus should increase with the aversiveness of the associated outcome although the objective covariation between stimuli and outcomes is always 50 percent. Given the aforementioned findings of a covariation bias in anxiety

disorders, it was further assumed that there is a correlation between trait or state anxiety and the strength of the covariation bias induced by the aversiveness of the outcomes.

There is evidence that heightened arousal might also explain the emergence of a covariation bias. That is, both negative and positive cues can lead to an overestimation of aversive outcomes (vanOyen Witvliet & Vrana, 2000; Pauli et al., 2002). A positive relationship has been found between a covariation bias and skin conductance responses to aversive outcomes after fear-relevant stimuli (de Jong, Merckelbach, & Arntz, 1995). Yet, if physiological arousal is manipulated more directly through physical exercise (Cavanagh & Davey, 2001) or pharmacological suppression (de Jong & Merckelbach, 2000), no effect on covariation bias is found. Considering this empirical background, it is uncertain whether physiological arousal could be a mediator in the maintenance of illusory correlations. Therefore, pupil dilation was measured throughout the experiment as an index of arousal (Bradley et al., 2008). This physiological measure of arousal allows us to examine whether the increased expectancy of aversive events is accompanied by increased arousal during the anticipation period and/or by increased arousal during the experience of the aversive outcomes. The latter may be related to the encoding of the aversive event. Arousal during encoding has been shown to be associated with successful memory performance. This might explain why people are biased towards the more aversive event when making a posteriori covariation estimates.

6.2. Methods

Participants

Twenty-four students (19 women) from the University of Würzburg completed the experiment and received course credits for their participation (one additional participant aborted the experiment). Age ranged from 18 to 31 years (M = 21.25, SD = 3.22). The participants completed the German versions of the Beck Depression Inventory II (BDI-II; M = 4.63, SD = 4.28; Dozois, Dobson, & Ahnberg, 1998), and the State-Trait Anxiety Inventory (STAI-state: M = 35.33, SD = 7.60; STAI-trait: M = 37.41, SD = 9.30; Spielberger et al., 1970).

Stimulus material and apparatus

Nine simple geometric figures were used as visual cues. A circle, a rectangle, and a triangle were each presented in yellow (RGB: 242, 182, 0), blue (RGB: 101, 207, 255) and green (RGB: 71, 255, 71). Triangles and rectangles were 10 cm wide and 14.5 cm high and circles had a diameter of approximately 12 cm. The background was always grey (RGB: 160, 160, 160) and slightly darker than the visual cues all of which had equal luminance values (180). A horizontal line, 11 cm in length, was superimposed on the visual cues to assess on-line covariation estimates. The scale was labeled in steps of ten units from 0 to 100 and participants could continuously move a red cursor horizontally with the help of two buttons. Stimuli were presented via the software Presentation (Neurobehavioral Systems, Inc., Albany, CA, USA) on a 17” computer screen with a refresh rate of 75 Hz and a resolution of 1280 x 1024 pixels. Geometric figures lasted for 8 s on the screen with a variable inter-trial-interval of 10 to 14 s (steps of 1 s).

Bursts of white noise were presented as startle sounds binaurally over headphones for 50 ms. Depending on the cue's color, the intensity (and therefore the aversiveness) of the startle sounds varied (95 dB, 100 dB, or 105 dB). Startle probes of these intensities are commonly used in emotion research and considered not harmful (Andreatta et al., 2010; Blumenthal et al., 2005). Startle sounds occurred at the offset of the visual cues.

Pupil size was recorded with a sampling rate of 240 Hz with an iViewX Hi Speed system (SensoMotoric Instruments, Berlin, Germany). This system operates with an infrared illumination so that the dark pupil and the corneal reflex can both be tracked for compensation of movements of the head relative to the camera. An integrated chin and forehead rest stabilized the participant's head in order to minimize movements for accurate measurement. Ceiling light was dimmed so that the participants could concentrate on the stimuli. The system was calibrated for each participant individually via a 13 point calibration procedure before the experiment started. Pupil diameter was processed by first averaging the horizontal and the vertical pupil diameters and then blinks and artifacts were substituted (according to the procedure described by Partala & Surakka, 2003) by the mean pupil diameter from the 15th sample (62.4 ms) to the 10th sample (41.6 ms) before the blink or the artifact. First, the artifacts were defined as changes in diameter of more than 0.2 mm within 2 samples (8.3 ms) or more than 0.375 mm within 10 samples (41.6 ms). Then, each trial was inspected visually for remaining artifacts that were not detected by this algorithm, and such bad trials were excluded from further analysis (24 percent of the trials) [11]. Finally, the pupil diameter was transformed into mean values of one-second-intervals relative to a baseline of one second. For the responses to the cues the baseline was the second before the cues' onset, and for the responses to the outcomes the baseline was the second before the outcomes' onset.

Procedure

After giving an informed consent, participants filled out the state and trait versions of the STAI and then the BDI-II. Participants were seated on a comfortable chair and their chins rested on the eye-tracking column while the calibration procedure was performed. Instructions on the screen explained that they will see colored geometric figures (cues) on the screen which sometimes will be followed by startle sounds. They were instructed to estimate the probability of the appearance of a startle sound during cue presentation by using the expectancy rating scale. It is important to note that the participants were not informed that the startle sounds could differ in intensity or that the three colors predicted the startle sound's intensity. The association between the three colors and the three intensity levels of the startle sounds was counterbalanced across the participants. Three practice trials with letters instead of figures and no startle sounds allowed the participants to get familiar with the procedure and the expectancy scale. Before the experiment started, they were reminded to sit comfortably and move as little as possible.

The experiment consisted of 72 trials divided in 12 blocks of 6 trials. Each block included all 6 possible trial types, i.e., combinations of colors and outcomes (3 types of colors x 2 types of consequences). Therefore,

[11] The exclusion of 24 percent of the trials led to a better signal-to-noise-ratio, but did not change the significance of the results in comparison to no exclusion of trials.

the objective probability of a startle stimulus was 50 percent for all intensity levels (and thus, also for all colors) in each block. Moreover, each geometric shape was equally often associated with a startle sound throughout the experiment.

After the experiment, participants rated the frequency (in percent) and the unpleasantness of the associated startle sounds for each color on continuous rating scales from 0 to 100 (for e.g., "Given there was a blue figure – how unpleasant (0 = not unpleasant at all; 100 = very unpleasant) was the associated sound?", or "Given there was a blue figure – how often (in percent) did a sound occur?"). Moreover, valence and arousal ratings were collected for each of the nine visual cues. Each combination of color and shape was presented in the middle of the screen with a rating scale below the figure. The rating scale for valence ranged from 1 (very unpleasant) to 9 (very pleasant), and the rating scale for arousal ranged from 1 (not arousing at all) to 9 (very arousing). For all the ratings, the sequence of color-shape combinations to rate was randomized.

Data analysis

A posteriori covariation estimates and STAI scores were missing from two participants and pupil data from one participant. All other analyses were run with the total sample of 24 participants. Repeated measures analyses of variance (ANOVA) were conducted and two-sided follow-up *t*-tests were calculated to resolve significant effects. Multivariate statistics and indicators of effect sizes (partial η_p^2 for ANOVA effects and Cohen's *d* for *t*-tests) are reported throughout this study. Reported correlations are Pearson's coefficients *r* with one-sided *p*-values based on directed hypotheses. It should be noted that a priori the current sample size was chosen to measure a covariation bias and not correlations with personality indices. Therefore, the correlation analyses have been considered exploratory and exact *p*-values are reported without correction for multiple tests. For all other tests, the alpha-error-level (two-sided) was set at .05.

6.3. Results

Ratings

Unpleasantness ratings of startle sounds. After the experiment, participants rated the unpleasantness of the startle sounds associated with the different color cues. The unpleasantness ratings were analyzed with a one-factorial repeated measures ANOVA with 3 levels. The main effect of *startle intensity* was significant, $F(2, 22) = 4.78$, $p < .05$, $\eta_p^2 = .30$. The high intensity startle sound ($M = 78.69$; $SD = 19.99$) was rated as more unpleasant than both the medium intensity startle sound ($M = 62.66$; $SD = 24.35$), $t(23) = 2.91$, $p < .01$, $d = 0.60$, and the low intensity startle sound ($M = 60.00$; $SD = 24.62$), $t(23) = 2.84$, $p < .01$, $d = 0.58$, while the difference between the medium and the low intensity startle sound was not significant ($p = .60$). It can be concluded that the experienced aversiveness of the startle stimuli was successfully manipulated for the high intensity sound relative to both other intensities.

On-line covariation estimates. The pivotal hypothesis of the present study was that the aversiveness of an outcome will increase the bias of the on-line covariation estimates for the associated stimulus. This assumption was supported by the acquired data. The on-line covariation estimates for the cues followed by

outcomes of high intensity (and therefore high aversiveness; see above) were enhanced compared to the corresponding ratings for the cues followed by outcomes of medium and low intensities (see Figure 6.1A). This interpretation was confirmed by a 3 (*startle intensity*) x 12 (*block*) repeated measures ANOVA for on-line covariation estimates which revealed no significant effect of *block* (p = .27), and no interaction of *Block* x *Startle Intensity* (p = .19), but a main effect of *startle intensity*, $F(2, 22) = 6.91$, $p < .01$, $\eta_p^2 = .39$. As Figure 6.1A illustrates, participants exhibited higher expectancies for the high intensity (M = 56.82; SD = 10.68) than for the medium intensity (M = 50.74; SD = 12.01) cues, $t(23) = 2.62$, $p < .05$, d = 0.53, or the low intensity cues (M = 48.29; SD = 11.09), $t(23) = 3.67$, $p < .01$, d = 0.78. The medium intensity cues and the low intensity cues did not differ significantly (p = .31).

Moreover, it was tested whether there was an overestimation relative to the objective 50 percent probability. While the on-line covariation estimates for the high intensity cue significantly differed from 50 percent, $t(23) = 3.13$, $p < .01$, d = 0.64, this was not true for the medium intensity cue or the low intensity cue (ps > .46).

A posteriori covariation estimates. The a posteriori covariation estimates also revealed that the participants overestimated the occurrence of the most aversive outcomes (Figure 6.1B). The one-factorial repeated measures ANOVA revealed a significant effect of *startle intensity*, $F(2, 20) = 10.55$, $p < .001$, $\eta_p^2 = .51$, which was due to higher covariation estimates for the high intensity cue (M = 62.59; SD = 16.32) than for the medium intensity cue (M = 46.70; SD = 19.24), $t(21) = 4.34$, $p < .001$, d = 0.93, or the low intensity cue (M = 47.59; SD = 18.63), $t(21) = 2.91$, $p < .01$, d = 0.62. A posteriori covariation estimates did not differ between the low and the medium intensity cues (p = .87). Again, the covariation estimates were tested for a reliable difference from 50 percent and only the high intensity cue was associated with a significant overestimation, $t(21) = 3.62$, $p < .01$, d = 0.77.

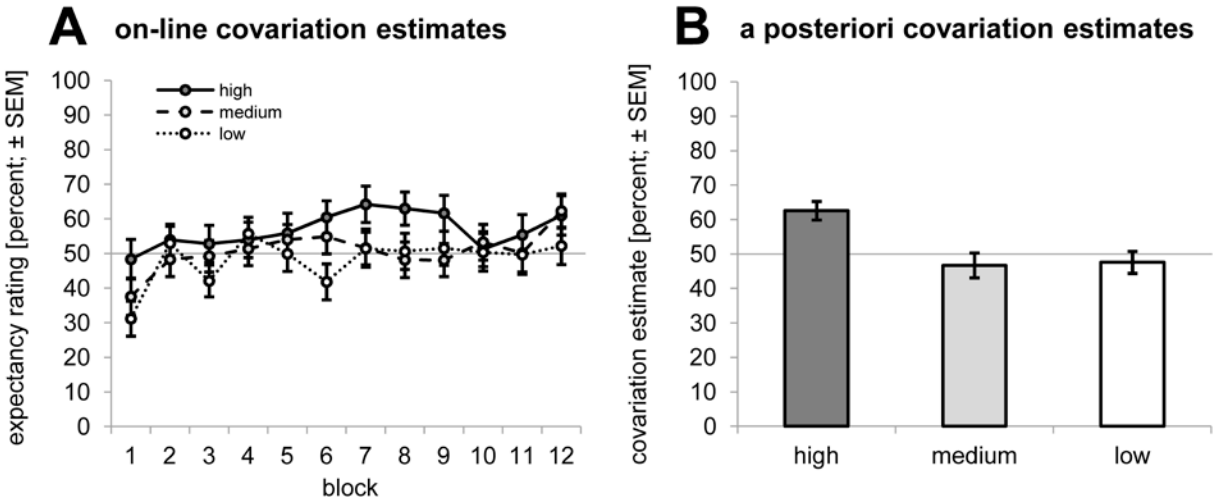

Figure 6.1 Covariation bias. On-line covariation estimates (A) for the three different cue conditions associated with three different startle intensities as outcomes, divided into twelve blocks of the illusory correlation experiment. Each block consists of every cue-outcome combination, resulting in a 50 percent frequency of startle sounds for every cue condition in every block (indicated by horizontal line). Error bars indicate standard errors of the mean. A posteriori covariation estimates (B) between each cue condition and startle sounds. The objective total frequency of startle sounds is 50 percent for each condition. Error bars indicate standard errors of the mean.

In addition to general estimations for each color, covariation estimates were collected for each visual cue (each geometric figure in each color). We examined the effects of loudness (i.e. color) and shape in a repeated measures ANOVA. Again, loudness affected covariation estimates significantly, $F(2, 22) = 5.18$, $p < .05$, $\eta_p^2 = .32$. Follow-up t-tests indicated that estimates for 105 dB cues ($M = 53.50$; $SD = 15.29$) were significantly higher than for 95 dB cues ($M = 43.26$; $SD = 20.15$), $t(23) = 2.78$, $p < .05$, $d = .57$, and marginal significantly higher than for 100 dB cues ($M = 46.84$; $SD = 18.83$), $t(23) = 1.95$, $p = .06$, $d = .40$. There was no difference between 100 dB cues and 95 dB cues, $p = .47$. Furthermore, there was no significant deviation from 50 % in any loudness condition, $p > .12$. There was also no significant interaction between of *Loudness x Shape*, but a marginal significant main effect of shape, $F(2, 22) = 3.02$, $p = .07$, $\eta_p^2 = .22$. The only significant difference was found between circles ($M = 43.68$; $SD = 18.40$) and triangles ($M = 52.14$; $SD = 14.75$), $t(23) = 2.50$, $p < .05$, $d = .50$.

Arousal and valence ratings of visual cues. Valence and arousal of each visual cue (each geometric figure in each color) were rated after the experiment and analyzed with two separate ANOVAs including the factors s*tartle intensity* and s*hape*. Most importantly, for both the ratings no effects were found for *startle intensity* (arousal: $p = .17$; valence: $p = .55$) and no significant interaction of *Startle Intensity* x *Shape* (arousal: $p = .50$; valence: $p = .50$). *Shape* did not influence the arousal ratings either ($p = .19$), but it influenced the valence ratings, $F(2, 22) = 13.34$, $p < .001$, $\eta_p^2 = .55$. Circles ($M = 5.31$; $SD = .94$) were rated as somewhat more positive than rectangles ($M = 4.49$; $SD = 1.06$), $t(23) = 4.99$, $p < .001$, $d = 1.02$, or triangles ($M = 4.43$; $SD = 1.16$), $t(23) = 3.52$, $p < .01$, $d = 0.72$. There was no difference between rectangles and triangles ($p = .82$; see Table 6.1 for details).

Loudness	**Shape**	**Valence (1-9)**		**Arousal (1-9)**		**Covariation (0-100%)**	
		M	*SD*	*M*	*SD*	*M*	*SD*
105 dB	triangle	4.25	1.89	4.62	2.26	56.81	19.70
	circle	5.54	2.32	4.20	2.14	51.12	21.40
	rectangle	4.41	2.22	4.79	2.35	52.56	22.06
100 dB	triangle	4.66	1.40	4.41	1.97	52.78	22.19
	circle	5.70	2.15	3.75	2.19	39.78	22.05
	rectangle	4.66	1.73	3.45	1.95	47.96	24.36
95 dB	triangle	4.37	1.63	4.79	2.04	46.81	23.26
	circle	4.66	2.05	4.25	2.02	40.12	25.19
	rectangle	4.37	1.99	4.41	2.01	42.84	25.74

Table 6.1. Valence ratings, arousal ratings and covariation estimates for different cue types. Means and standard deviations of valence and arousal ratings for each figure presented during the illusory correlation experiment. Objective covariation probability with startle sounds is 50 percent for each figure. High valance ratings indicate pleasantness, while low valence ratings indicate unpleasantness.

Pupil diameter

As depicted in Figure 6.2A, all visual cues at their onset evoked similar initial (first 2 s) pupillary constrictions because of the enhanced luminance. This observation was confirmed by a 3 (*startle intensity*) x 8 (*time*) ANOVA for the pupil diameter during cue presentation which returned a significant main effect of *time*, $F(7, 16) = 10.65$, $p < .001$, $\eta_p^2 = .82$, but no significant main effect of *startle intensity*, $F(2, 21) < 0.01$, $p = .99$, $\eta_p^2 < 0.01$, and no interaction of *Time* x *Startle Intensity*, $F(14, 9) = 1.96$, $p = .16$, $\eta_p^2 = .75$. In contrast, the startle sound onsets were followed by an increase in pupil diameter that reached its maximum within the first 2 s (see Figure 6.2B), and this response seemed to be modulated by the outcomes' intensity. A 2 (*outcome*) x 3 (*startle intensity*) x 8 (*time*) repeated measures ANOVA resulted in significant main effects of *time*, $F(7, 16) = 23.84$, $p < .001$, $\eta_p^2 = 0.91$, and *outcome* (startle sound vs. nothing), $F(1, 22) = 44.60$, $p < .001$, $\eta_p^2 = .67$. In addition, there were significant interactions of *Outcome* x *Time*, $F(7, 16) = 14.89$, $p < .001$, $\eta_p^2 = 0.87$, *Outcome* x *Startle Intensity*, $F(2, 21) = 3.68$, $p < .05$, $\eta_p^2 = .26$, and a significant three-way interaction of *Outcome* x *Time* x *Startle Intensity*, $F(14, 9) = 4.86$, $p < .05$, $\eta_p^2 = 0.88$.

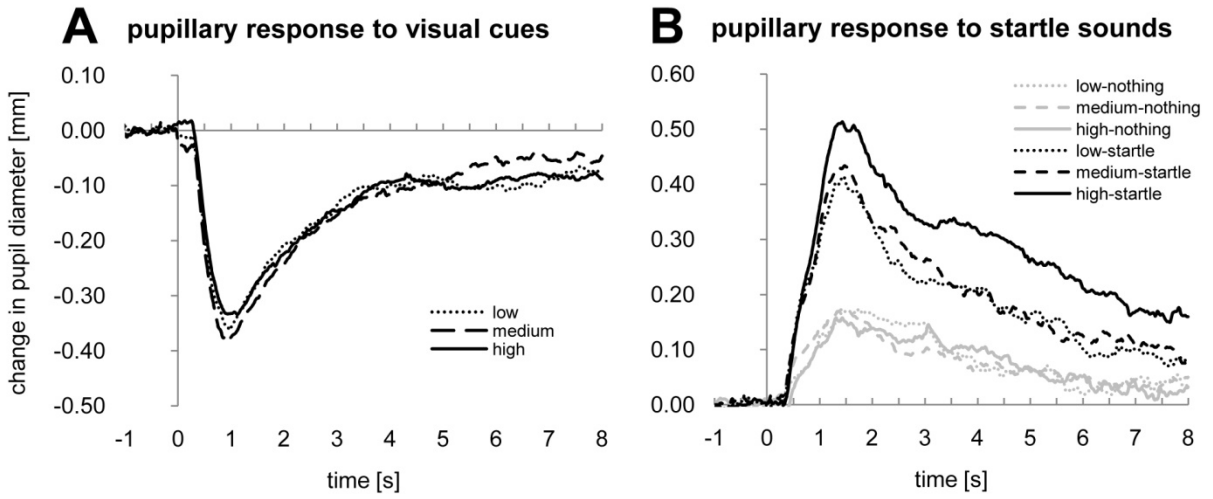

Figure 6.2. Pupil dilation. Pupillary responses to visual cues (A) and outcomes (B), i.e., startle sounds of different intensity or nothing.

To clarify these results, two separate 3 (*startle intensity*) x 8 (*time*) ANOVAs were conducted, one for each outcome. When there were startle sounds, there were significant main effects of *time*, $F(7, 16) = 33.41$, $p < .001$, $\eta_p^2 = .94$, and s*tartle intensity*, $F(2, 21) = 5.60$, $p < .05$, $\eta_p^2 = .35$, and a marginal significant interaction of *Time* x *Startle Intensity*, $F(14, 9) = 2.68$, $p = .07$, $\eta_p^2 = .81$. This *Time* x *Startle Intensity* interaction was scrutinized via separate ANOVAs for each time point. A main effect of s*tartle intensity* was found from seconds 2 to 6 (all *p*s < .05). During these intervals the high intensity startle sound elicited a larger pupillary dilation than both the medium intensity startle sound, $t(22) = 3.19$, $p < .01$, $d = 0.66$, and the low intensity startle sound, $t(22) = 3.28$, $p < .01$, $d = 0.69$. When there was no startle sound, the interaction of *Time* x *Startle Intensity* was not significant ($p = .18$) indicating that there were no differential responses to the absence of an outcome, for e.g., due to intensity dependent prediction errors.

Correlations between covariation biases and anxiety

In order to test the association between the on-line covariation estimates or the a posteriori covariation biases with state or trait anxiety, bias scores were calculated as mean on-line covariation estimates or a posteriori covariation estimates for the high intensity cue *minus* mean on-line covariation estimates or a posteriori covariation estimates for the low intensity cue. Figure 6.3 shows that the on-line bias score correlated positively with both state, r = .36, p = .04, and trait anxiety scores, r = .45, p = .02. Moreover, the a posteriori covariation bias score was significantly correlated with trait anxiety, r = .46, p = .02, and marginally significantly correlated with state anxiety, r = .34, p = .06. In contrast, there was no significant relationship between depressive symptoms as measured by BDI-II and on-line covariation bias scores (p = .13) or a posteriori covariation bias scores (p = .18).

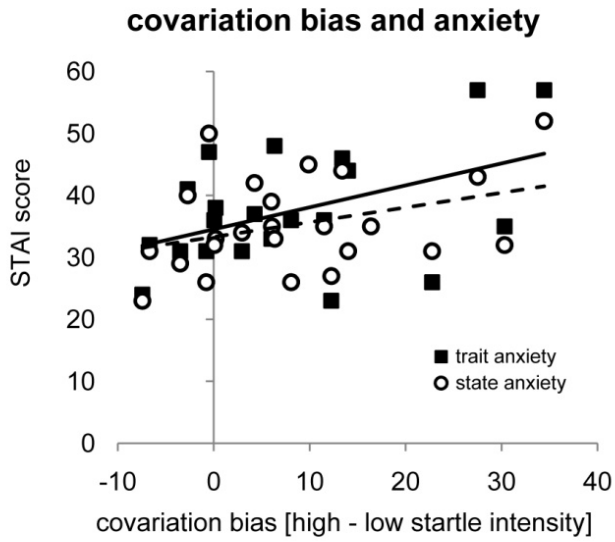

Figure 6.3. Anxiety and covariation bias. Scatter plot and regression lines for state (intersected) and trait anxiety scores (solid) vs. on-line covariation bias.

6.4. Discussion

In experiment 4, the hypothesis was tested that an illusory correlation can be induced by merely manipulating the aversiveness of an outcome. To this end, the traditional paradigm was modulated by using three neutral cues each followed randomly by either an aversive outcome or nothing. Importantly, the aversive consequence varied in intensity and therefore in aversiveness depending on the cue's color. The hypothesized illusory correlation was found on three different levels of measurement. First, on-line covariation estimates indicated that participants expected an aversive consequence with an enhanced likelihood after cues that were associated with highly aversive outcomes compared to cues that were associated with less aversive outcomes. Please note that the likelihood of outcomes was equal in all conditions. These findings based on on-line ratings suggests that the illusory correlation already emerged during the experiment and perhaps occurred very early since there was no interaction with time. Second, after the experiment, participants still overestimated the

covariation for the cues that were associated with highly aversive consequences. This implies that the illusory correlation was not only restricted to the phase when there was an actual threat of an aversive consequence or participants might have been under time pressure. Since a covariation bias was present even after the experiment, it can be assumed that it should be relevant for the participants' behavior in future. This is an important aspect considering that such a covariation bias is thought to induce or maintain anxiety (Mathews & MacLeod, 1994; de Jong, van den Hout, & Merckelbach, 1995). Finally, separate covariation estimates for every type of figure (i.e. different shapes in different colors) confirmed the biased covariation estimates. In this regard, circles were marginally less often associated with aversive sounds than triangles. Future studies might take care of this effect to reduce error variance. However, since the shape was counterbalanced between the conditions, this does not affect the effects of startle aversiveness.

What do these findings tell us about human cognition? The present results imply that the expectancy of an aversive outcome depends on the aversiveness of this outcome. This is exactly what an adaptive conservatism approach predicts. If the threat value of an outcome increases, the costs of not being prepared also increase regardless of the objective probability of that outcome. For example, if a person encounters a snake, s/he should more likely anticipate a snake attack when s/he thinks that the snake is poisonous than when s/he thinks that the snake is harmless although the objective likelihood of the snake attack is equal (given that all other variables like fear and knowledge about snakes are kept constant). Similarly, as motivation has been described as a function of expectancy and value of outcome (Wigfield & Eccles, 2000), expectancy may be a function of outcome probability and value of outcome.

The present findings may well be a consequence of a general preferential processing of negative information, also called *negativity bias* (Vaish, Grossman, & Woodword, 2008). The more aversive outcomes may be processed with priority and are consequently easier to retrieve when contingencies are estimated. There is a great deal of evidence indicating that negative information takes priority over neutral and positive information. For example, the knowledge about negative vs. positive traits of a person is prioritized in social decision making (Skowronski & Carlston, 1987) and negative pictures receive enhanced neuronal processing relative to positive pictures when arousal for both is equal (Ito, Larsen, Smith, & Cacioppo, 1998). Even three-month old infants display negativity bias in a social perception task (Kiley Hamlin, Wynn, & Bloom, 2010) suggesting that such a bias may have been adaptive in human evolution or acquired early during development. One reason for negativity bias in human cognition may be that usually higher costs are involved in not recognizing a threat than in missing positive reinforcement (Dawkins & Krebs, 1979).

In light of the existence of negativity bias, it is not surprising that the on-line covariation bias for aversive events is not limited to clinical samples, but is also observable in healthy participants like in the present study. However, the exploratory analyses revealed a positive correlation between the observed on-line covariation bias and both state and trait anxiety scores suggesting that cognitive bias in anxious individuals is a result of exaggerated responses rooted in normal cognitive functioning. These findings fit with those of other studies that also reported enhanced on-line covariation estimates for aversive stimuli in trait anxious individuals (Chan & Lovibond, 1996; Boddez et al., 2012). When highly anxious participants were unaware of the contingencies during conditioning, on-line expectancy ratings were elevated for both danger and safety cues (Chan & Lovibond, 1996).

In a blocking procedure, a stimulus usually evokes less fear when it has been "blocked" by pairing it with a clear conditioned stimulus that is followed by an electric shock. However, highly anxious participants expected an electric shock even after the blocked stimulus (Boddez et al., 2012).

These findings are in line with the idea that anxiety affects fear responses especially in ambiguous or *weak situations* (Lissek et al., 2006). In a weak situation, the relationship between stimuli is uncertain and their hedonic valence is relatively interpretable. In contrast, in a *strong situation*, unambiguous stimuli predict a clearly defined event with certainty and provoke relatively uniform reactions in all individuals. The present experimental paradigm can be better described as a weak situation and should therefore be suitable to study inter-individual differences in anxiety (Beckers et al., 2012). There was no relationship between depressive symptoms and illusory correlations. So, the findings of Chan and Lovibond (1996) were confirmed who also reported a relationship between on-line expectancy bias and anxiety but not depressive symptoms. Still, our findings may not hold true for clinical samples of depressed or anxious patients. Moreover, it should be noted that the sample size was relatively small for correlation analyses. A replication of these results should be awaited before firm conclusions can be drawn about the relationship between anxiety and an illusory correlation induced by aversive outcomes. An exaggerated covariation bias for aversive events might contribute to the maintenance of anxiety and anxiety disorders but causal relationships between illusory correlations and anxiety have not yet been investigated. However, modifying attentional bias and interpretative bias leads to a reduction of symptoms of anxiety (Hallion & Ruscio, 2011; Mac Leod, 2012). In a similar way, modifying covariation bias should lead to a reduction of anxious feelings. This could be realized by informing participants about the random contingencies between fear-relevant cues and outcomes.

The finding that expectancies are biased towards more aversive outcomes is in line with a recent study showing that cues of uncertainty can lead to inflated expectancies of aversive pictures (Grupe & Nitschke, 2011). In this study, an uncertainty cue was followed by either a neutral or an aversive picture. Two other cues predicted a neutral or an aversive picture with certainty. During the presentation of the uncertainty cue, participants displayed an expectancy bias for aversive pictures. After the experiment, the amount of aversive pictures occurring subsequent to uncertainty cues was overestimated. In addition, when aversive pictures were cued with uncertainty, they led to a more negative mood than when they were cued with certainty. The present results complement these findings by using basic aversive stimuli (i.e., startle sounds) instead of unpleasant pictures as outcomes. This suggests that the expectancy bias for aversive events under uncertainty is not restricted to aversive pictures but also counts for general aversive stimuli without any semantic content.

Moreover, Grupe and Nitschke (2011) assume that uncertainty cues may have been unpleasant a priori and conclude that covariation bias might have been the result of an affective match between unpleasant uncertainty cues and unpleasant outcomes (Tomarken et al., 1995). Yet, the present data suggest that illusory correlations can emerge without any affective matching processes since the present study used neutral geometric figures. According to the valence ratings, the figures were still considered neutral after the illusory correlation procedure, and there were no differences in valence between cues that predicted different outcome aversiveness. Nevertheless, there was a covariation bias for the most aversive outcome only that can be explained as a more general expectancy bias for aversive events. Therefore, the covariation bias between

uncertainty cues and aversive pictures (Grupe & Nitschke, 2011) could also be the result of this general expectancy bias for aversive events because the aversive events themselves were more aversive in terms of heightened negative mood. Affective matching may also contribute to illusory correlations but it is probably not a necessary prerequisite.

However, why were the cues associated with the most aversive outcome not rated as more negative or more arousing than others? One has to keep in mind that even the low intensity sound could be considered aversive and the contingency was equal across conditions. The difference in outcome aversiveness in the present study is probably not enough to induce evaluative conditioning but it is sufficient to bias expectancy ratings. Heightened contingency estimates may be more relevant for efficient coping with an aversive event than a conditioned evaluation. Therefore, the biased contingency estimates may be more readily established in the learning process.

The present results imply that a general expectancy bias for aversive events could explain a covariation bias between unpleasant stimuli because a preceding unpleasant stimulus can amplify the aversiveness of an outcome. In a similar way, this effect may also explain an *encounter expectancy bias* for phobia-relevant stimuli in people suffering from specific phobias (Aue & Hoeppli, 2012). Since for spider phobics encountering a spider is a more aversive event than encountering a fear-irrelevant animal like a bird, such an encounter expectancy bias may very well be the result of a more general expectancy bias for aversive events.

In addition, a general expectancy bias for aversive events could be in part a causal sustaining factor of covariation bias between fear-relevant and aversive outcomes. Previous findings suggest that covariation bias is not only determined by a priori expectancy bias. There seems to be, for instance, an a priori expectancy bias for technological (e.g., gun) and biological threats (e.g., snake), but only biological threats are illusory correlated with shocks at a later stage of the experiment (Amin & Lovibond, 1997; Kennedy et al., 1997; Mühlberger et al., 2006). Maybe divergent on-line processing of associations could explain why covariation bias is sustained for fear-relevant and biological stimuli. Fear and negative affect can enhance pain perception (Rhudy & Meagher, 2001; Kenntner-Mabiala & Pauli, 2005, Kenntner-Mabiala et al., 2008) and there is some evidence that there might be an association between covariation bias and perceived aversiveness of outcomes (Tomarken et al., 1989; Mühlberger et al., 2006). It may be that biological threats lead to a more aversive outcome experience than technological threats (Mühlberger et al., 2006). The sustained covariation bias for biological threats could be the consequence of this enhanced aversiveness.

The mechanisms mediating the effect of outcome aversiveness on illusory correlations remain to be determined. The aforementioned adaptiveness of expecting aversive events or enhanced arousal produced by more aversive outcomes could be plausible mechanisms. The enhanced arousal, in turn, could support encoding memory processes for more aversive associations similar to the facilitating effect of arousal on emotional memory (Canli, Zhao, Brewer, Gabrieli, & Cahill, 2000; Dolcos & Cabeza, 2002). Later on, when a decision about expectancy or frequency would be made, these aversive associations would be easier to retrieve and influence decision making. In line with this theory, skin conductance responses to shocks following fear-relevant cues are higher than following neutral cues and predict covariation estimates (de Jong, Merckelbach, & Arntz, 1995). Moreover, the present study reported pupil dilation responses to aversive outcomes that paralleled the pattern

of on-line and a posteriori covariation estimates. Pupil dilation is influenced by locus coeruleus activity in the brainstem (Laeng, Sirois, & Gredebäck, 2012), releasing norepinephrine and affecting a series of cognitive processes like attention and memory retrieval (Sara, 2009). In addition, locus coeruleus activity is associated with hippocampal long term depression and enhanced memory encoding (Lemon, Aydin-Abidin, Funke, & Manahan-Vaughan, 2009). Therefore, enhanced pupil dilation in response to the most aversive outcome could be a sign of strengthened encoding processes.

In previous studies, evidence was found for a relationship between arousal during cue presentation and a posteriori covariation bias (vanOyen Witvliet & Vrana, 2000; Pauli, et al., 2003). In the current study, pupil data during cue presentation and arousal ratings did not differ between conditions, rendering it very unlikely that the influence of outcome aversiveness on illusory correlations were mediated by enhanced arousal during the anticipation of outcomes. This does not imply that arousal during cue presentation cannot have an impact on covariation bias between emotionally arousing stimuli and aversive events and could be an additional factor independent of the effect of outcome aversiveness. Increased arousal during anticipation may be interpreted as a warning signal and thus lead to increased expectancy of an aversive event. Another possibility may be that increased arousal during cue presentation facilitates memory encoding for a following aversive outcome.

So far, one cannot say whether it is necessary to be aware of the relationship between the cues and the loudness for the covariation bias to occur. In other words, the present results do not allow to conclude whether more implicit or explicit learning processes lead to the illusory correlations. However, a follow-up study that is not part of the present thesis suggests that it is an explicit learning process. In a shortened version of the present experiment (9 instead of 12 blocks), there was no overall effect of outcome aversiveness on covariation estimates anymore. In contrast to the present experiment, participants did also not learn which cue was associated with the most aversive sound, at least according to a posteriori aversiveness ratings. However, the difference in outcome aversiveness (high minus low) was highly correlated with the covariation bias (see Figure 6.4). This suggests that only those who learned the association between colors and aversiveness developed an illusory correlation between the appropriate color and the aversive sounds. Moreover, those participants who rated the sound associated with the low-intensity-color as more aversive than the sound associated with the high-intensity-color developed an illusory correlation between the low-intensity-color and the aversive sound. This indicates that for the influence of outcome aversiveness on illusory correlations, it may be irrelevant which cue was associated with the most aversive outcome in reality. In contrast, the assumption which cue was associated with the most aversive sound is sufficient to evoke illusory correlations. One might conclude that top-down processes mediate the effect of outcome aversiveness on illusory correlations, but still further research measuring the awareness of cue-outcome-associations is eligible.

One limitation of the present study is that the experienced aversiveness of consequences was experimentally manipulated by manipulating their physical intensity, and not specifically their subjective intensity. This might affect comparability to former covariation bias experiments in which aversiveness only differed subjectively. One might argue that the difference in subjective aversiveness with constant physical intensity should be much smaller as when there are actually differences in physical intensity. However, selective manipulation of subjective aversiveness is difficult and especially difficult to quantify, although unconditioned

stimulus inflation (Hosoba, Iwanaga, & Seiwa, 2001) may be one possibility for future studies. The current effect size of the aversiveness manipulation was *d* = .58. Therefore, the difference in outcome aversiveness in the current study may still have been smaller than the difference in aversiveness between outcomes following phobic and neutral cues (*d* = 1.46; Mühlberger et al., 2006). This suggests that the difference in outcome aversiveness in the present study was not too large to simulate typical effects of fear-relevant stimuli on outcome aversiveness.

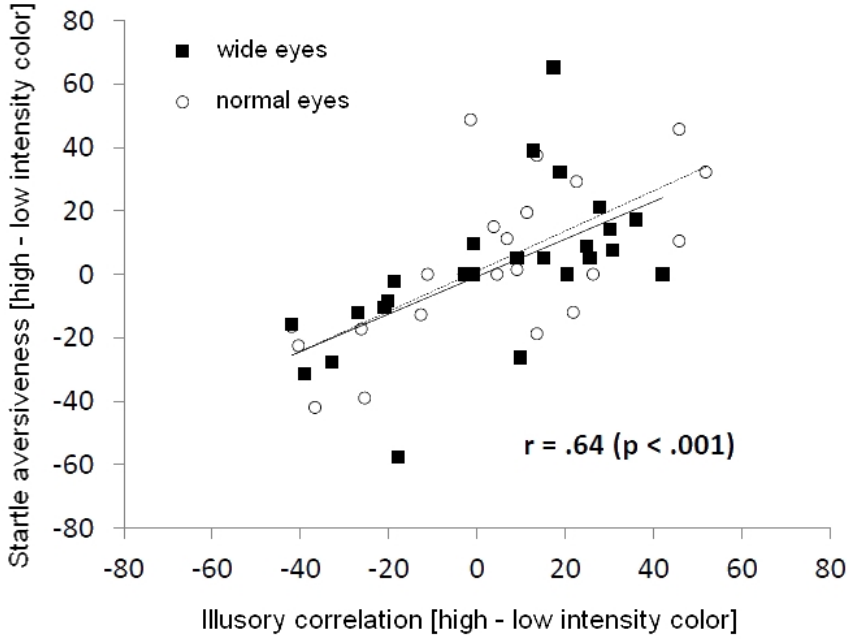

Figure 6.4. Illusory correlation and startle aversiveness in an unpublished replication experiment. The illusory correlation procedure was comparable to the present experiment 4, but contained less trials and in contrast to the present experiment, participants did not rate the high intensity sound as more aversive and more frequent after the experiment, when they gave these ratings for the different cue colors. Nevertheless, there was a strong correlation between these variables. Possibly, learning the relationship between the cue and the aversiveness is a prerequisite for the impact of aversiveness on frequency estimations. The experimental manipulation of widening the eyes like in a fear expression did not modulate this relationship.

Another limitation of the present study regarding outcome aversiveness is that its subjective impact was rated retrospectively after the experiment and not right after stimulus experience. It cannot be ruled out that different results may have been obtained when aversiveness ratings would have been collected on-line. However, on-line pupillary responses to the outcomes showed the same pattern as retrospective aversiveness ratings. Finally, the current data set does not allow strict differentiation between the effects of aversiveness per se and the salience of the outcome. Like in many studies on pain perception (Legrain et al., 2011), these two components might have been confounded. Future studies should elaborate whether a salient and not aversive outcome is sufficient to induce such an illusory correlation. However, highly salient but non-aversive outcomes have failed to induce enhanced probability estimates in a previous illusory correlation experiment (Tomarken et al., 1989).

In summary, evidence was found to support the notion that heightened outcome aversiveness can be a causal and sufficient factor to induce an illusory correlation in a nonclinical sample though the effect was

positively associated with the participants' anxiety level. These results could explain previous findings of sustained illusory correlations between fear-relevant and aversive events. Considering growing evidence for a potential causal impact of cognitive bias on anxiety (MacLeod & Mathews, 2012), it can be assumed that specifically modifying the appraisal of feared consequences may be a useful technique to reduce illusory correlations and anxiety.

7. General discussion

7.1. Summary and discussion of the present experiments

The aim of the present dissertation was to find evidence for an explanation of fear-relevant illusory correlations. A fear-relevant illusory correlation occurs when an individual overestimates the relationship between feared objects and aversive consequences (e.g. Tomarken et al., 1989). Theoretically, this cognitive bias maintains fear and anxiety (e.g. de Jong, van den Hout, & Merckelbach, 1995). Therefore, it is important to identify psychological mechanisms that conjointly lead to the emergence of fear-relevant illusory correlations. Ideally, these mechanisms can be modified to reduce biased cognitions about the environment and feelings of fear and anxiety.

On the basis of the empirical literature fear-relevant illusory correlations were explained via a model of interacting psychological processes sustaining fear-relevant illusory correlations. Mainly, two critical aspects of this model, which are not or not unequivocally supported by empirical findings so far, were tested with four experiments. The first two experiments were designed to test the influence of arousal on illusory correlations. All four experiments tested the relationship between the perceived aversiveness of an outcome and illusory correlations; experiments 1-3 on a correlational basis, experiment 4 with a causal approach.

So far, it was unclear whether arousal promotes fear-relevant illusory correlations. Direct manipulation of physical arousal level did not influence a covariation bias (Cavanagh & Davey, 2001; de Jong & Merckelbach, 2000). When arousal was manipulated within subjects on the level of the visual cues, also positive arousing stimuli were more likely to be associated with aversive consequences than neutral stimuli (vanOyen Vitvliet & Vrana, 2000; Pauli, Diedrich, & Müller, 2002). The latter effect should have been replicated in experiments 1 and 2, and transferred to the traditional illusory correlation paradigm in experiment 1. Particularly, former studies had instructed participants to ignore the aversive outcomes while fear-relevant illusory correlations usually persist even if attention is focused on associations during an illusory correlation paradigm. Since only a trial-by-trial manipulation of the arousal level by arousing cues affected illusory correlations in previous studies (vanOyen Vitvliet & Vrana, 2000; Pauli et al., 2002), positive arousing pictures were used in experiments 1 and 2 to investigate the impact of arousal on illusory correlations.

Experiments 1 and 2 failed to clarify the role of arousal in illusory correlations. Although women expected more aversive sounds after positive arousing stimuli in experiment 1, the effect was small and a remaining subsample did not show significantly biased on-line expectancies for positive arousing stimuli. In experiment 2, there was clearly no a posteriori covariation bias for happiness pictures in an affect modulated startle paradigm. From a theoretical perspective, it is plausible that arousing stimuli support memory and learning processes which in turn increase the availability of associations at a later time point and thus bias frequency estimates. However, on a trial-by-trial basis, arousal is difficult to manipulate and its effect might have been attenuated by counter expectancies and outcome aversiveness. In sum, it is still not clear whether positive

arousing stimuli can lead to illusory correlations, but if so, the effect is definitely smaller than that of negative arousing stimuli (arousal was matched very well between positive and negative pictures in experiment 1). Clearly, additional factors have to be assumed to explain fear-relevant illusory correlations.

As one important factor, the modulated aversiveness of outcomes following fear-relevant stimuli was investigated in the present thesis. Aversive shocks after spider images were perceived as more painful by phobic individuals than after neutral images, and it has been supposed that this enhanced outcome aversiveness might play a role in covariation bias (Tomarken et al., 1989). The present thesis includes the first experiments to test this hypothesis more systematically. Overall, strong evidence was found for the theory that negative emotional stimuli amplify the subjective aversiveness of aversive stimuli, and that this enhanced aversiveness plays an important role in the maintenance of illusory correlations. In all experiments 1-3, the outcome aversiveness was correlated with the covariation bias for emotionally negative pictures. This was true for mutilation pictures (experiment 1), disgust and fear pictures (experiment 2) and phobia-relevant spider pictures (experiment 3). One might argue that this correlation is only a demand effect at the end of the experiment, but experiment 3 unequivocally demonstrated that shocks after spider pictures evoked increased shock-related activity, dominantly in the primary sensory-motor cortex. This activity predicted both shock aversiveness and shock probability estimates after the experiment. Finally, in experiment 4, outcome aversiveness was manipulated experimentally by pairing neutral geometric figures with startle sounds of different intensity. As expected, this manipulation resulted in an illusory correlation for the neutral figures that predicted the most aversive sound.

The enhanced aversiveness of aversive outcomes after fear-relevant stimuli could lead to a strengthened encoding process due to increased attention and/or arousal (Cahill & McGaugh, 1998). This strengthened memory trace should be accessed more easily in the retrieval process at the moment when one estimates the number of previous occurrences and the probability in the future. This explanation is in line with earlier explanations of illusory correlations in social psychology, suggesting that members of minorities and undesirable behavior are illusory correlated because both minorities and undesirable behavior are infrequent, and therefore attract the most attention (Hamilton & Gifford, 1976). This increased salience may again lead to better memory encoding and influence decision making at a later time point via an availability heuristic (Tversky & Kahnemann, 1973). That is, better encoded memory traces are more available than weaker encoded memory traces. Based on this availability the frequency is overestimated because usually frequency leads to availability - but salience and aversiveness should do so, too. For example, the encoding of salient meanings of words leads to better recollection performance in comparison to unusual meanings of words (Rajaram, 1998). Moreover, the recall of the total amount of experienced pain in a painful medical treatment is better explained by the peak of the experienced pain than the total amount of experienced pain (Redelmeier & Kahneman, 1996).

The aversiveness explanation of fear-relevant illusory correlations does by no means exclude other explanations such as the influence of a priori expectancies (e.g. de Jong et al., 1990) and the affective matching account (Tomarken et al., 1995). A priori expectancies and the enhanced aversiveness of outcomes may conjointly influence decision making, and the aversiveness explanation can also be considered as an advancement of the affective matching hypothesis. The enhanced aversiveness may in part be a consequence of the affective similarity of the emotional stimulus and the outcome: Since defensive responses are already primed

by the emotional stimulus, the defensive response to the aversive outcome is amplified. This amplified response might mediate the effect of affective matching on the covariation bias via the previously described mechanisms. There are some arguments why it is probably the outcome aversiveness (or salience) itself that affects illusory correlations and that this relationship is not just a side effect of stimulus aversiveness on illusory correlations. First, SCRs to shocks were more similar to covariation bias data than SCRs to images in a previous study (de Jong & Merckelbach, 1991). Second, in a multiple regression analysis of experiment 1 and 2, the correlation of outcome aversiveness and covariation bias was independent from picture valence ratings. Third, in experiment 4, outcome aversiveness affected covariation bias, but not picture valence. Finally, the aversiveness explanation overcomes one shortcoming of the affective matching account: If it is affective matching that determines covariation bias, then why is the association between neutral slides and neutral outcomes not overestimated? These combinations should also be affectively similar. According to the outcome aversiveness account, these combinations should not be overestimated because neutral pictures should not amplify the aversiveness and salience of an outcome.

The theory that outcome aversiveness is enhanced after fear-relevant stimuli and promotes encoding processes might be interpreted as contradictory to the theory that contingency monitoring is complicated by increased attention to fear-relevant stimuli. The assumption that phobics have a deficit in contingency monitoring when confronted with phobia-relevant stimuli was a hypothesis of the illusory correlations model (chapter 2.5) and has been endorsed by the involvement of the dorsolateral prefrontal cortex (dlPFC) in the covariation bias in experiment 3. However, a deficit in contingency monitoring on the one side and increased processing of outcomes on the other side seems to be a conflict at first sight. Nevertheless, it is possible that the reduced contingency monitoring only takes effect in trials with nothing as outcomes. In other words, participants might focus their attention on the spider images and not notice the fact that nothing happened because 'nothing' is not very attention capturing. In contrast, the bottom-up noxious stimulation may have the potential to automatically draw attention and to put the encoding process into operation. Thus, there may be an asymmetry in the contingency monitoring deficit that specifically affects those trials that disconfirm the illusory correlation. Alternatively, it is still possible that enhanced dlPFC activity reflects improved contingency monitoring which has a selective effect on trials with aversive outcomes. This further differentiation cannot be made yet on the basis of the current experiments.

7.2. Reappraisal of the fear-relevant illusory correlations model

In chapter 2.5, a model of psychological factors was proposed that explains how fear-relevant illusory correlations emerge and why they persist despite the exposure to disconfirming evidence. Some of these psychological factors have been tested here experimentally and on an exploratory basis. In addition, the measurement of brain activity in experiment 3 allows in some regards to hypothesize the neural underpinnings of some factors.

First, the present work confirms the assumption that aversive outcomes after negative emotional stimuli can become more aversive themselves and that this enhanced aversiveness predicts covariation bias. In addition,

experiment 4 demonstrated that the enhanced aversiveness of an outcome can be a causal factor that is sufficient to create an illusory correlation. If we can generalize from the experimental manipulation of the objective outcome aversiveness to the emotional modulation of outcome aversiveness, we can assume that the enhanced outcome aversiveness is one of the reasons why patients with specific phobia continue to associate fear-relevant stimuli with negative consequences. The brain activity that is associated with this process seems to take place in the bilateral primary sensory-motor cortex (precentral gyrus and paracentral lobulus), the cerebellum and the fusiform cortex. A priori, the 'pain matrix' was expected to play a role in the emergence of fear-relevant illusory correlations, that is the primary and secondary sensory cortices, the anterior cingulate cortex and the insula. Among these regions, the primary sensory-motor cortex was the only region to be associated with the illusory correlation.

Second, the influence of arousing stimuli on covariation bias could not have been confirmed without any doubts in the present thesis, at least as tested with positive arousing stimuli. Although, women expected more aversive sounds after erotic pictures than after household objects and a remaining subsample of participants that was exposed to random contingencies tended to expect more aversive sounds after erotic pictures, the latter finding was only marginally significant. In addition, no overestimation for positive pictures was found in experiment 2. Overall, the impact of negative arousing slides on covariation bias was clearly more pronounced than that of positive arousing slides (arousal ratings were well matched in experiment 1), indicating that valence plays an important role beyond arousal. Considering previous negative results of the influence of directly manipulated physiological arousal, one may assume that arousal might play a role if it varies from category to category, but probably it is of minor importance in the explanation of fear-relevant illusory correlations.

Third, it may be added to the model that the involvement of the dlPFC and potentially a larger fronto-parietal attention network (Ptak, 2012) is associated with illusory correlations. The dlPFC was expected to be deviantly activated in spider phobics because it was known to predict contingency awareness (Carter et al., 2003). However, before the conductance of experiment 3, it was not clear whether dlPFC activity would be enhanced or reduced in response to phobic images. One could have assumed that dlPFC activity would be reduced and therefore less executive resources would be available for contingency monitoring in spider trials. Thus, reduced dlPFC activity should have been correlated with covariation bias. However, the opposite was true and the dlPFC activity was enhanced in response to the spider images. This enhanced activity predicted covariation bias after the experiment. In both spider phobics and healthy controls, increased activation of a left fronto-parietal network in spider trials predicted covariation bias. At the moment, it seems to be most reasonable to assume that an over-engagement of a fronto-parietal attention network in fear-relevant stimuli prevents some individuals from using these resources for contingency monitoring and therefore to adjust their inflated expectancies to the real predictive value of feared objects. Another possibility is that this activity reflects enhanced contingency monitoring which only takes effect on trials with aversive outcomes.

Finally, the first two experiments suggest that there might be a significant difference between men and women in emotionally induced a priori expectancy bias (experiment 1) and a posteriori covariation bias (experiment 2). On average, only women but not men were prone to the biased contingency estimates. Further

studies explicitly designed to study gender differences are warrant to clarify these first results. If gender plays an important role in fear-relevant illusory correlations, it is not clear so far, what factors precisely are mediating this effect. Experiment 1 suggests that a difference between men and women was already present in a priori expectancies, while genders did not differ in self-reported valence and arousal of the emotional stimuli, and the emotional modulation of aversive outcomes. While women did not differ from men in emotional responsiveness, they might be different in the cognitive consequences of emotional responses.

All of these aspects were integrated in an updated model of fear-relevant illusory correlations that can be seen in Figure 7.1.

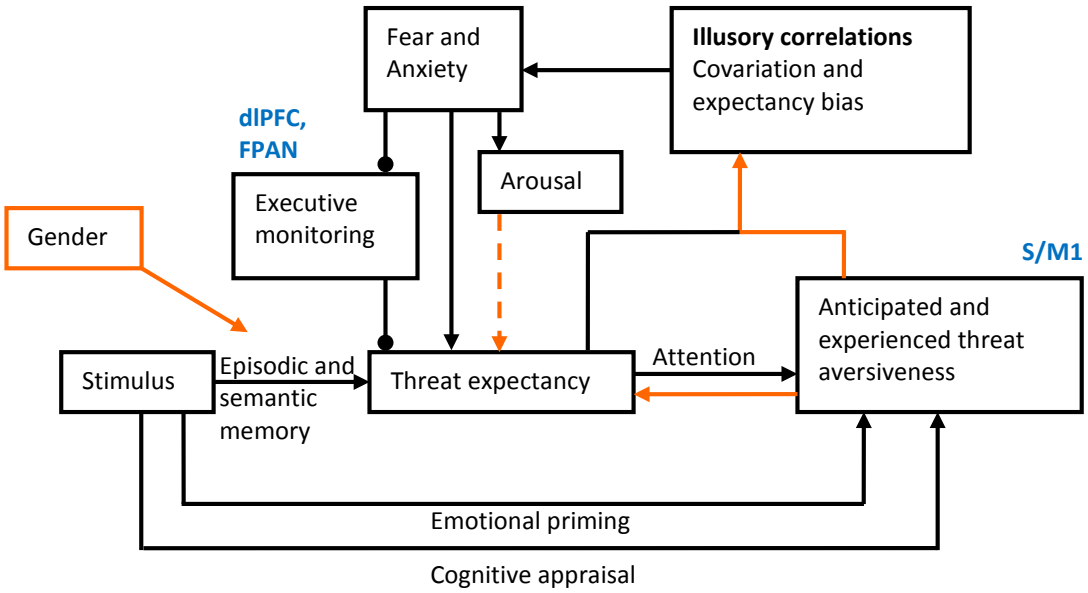

Figure 7.1. A revised version of the fear-relevant illusory correlations model. The orange arrows represent relationships that were tested in the present thesis. While outcome aversiveness has a causal impact on threat expectancy and covariation bias, the role of arousal could not be clarified in the experiments (dashed arrow). Unexpectedly, gender may be important in the generation of an expectancy bias, but future studies will have to confirm this finding. The most important brain regions underlying contingency monitoring (dlPFC, FPAN) and threat aversiveness (S/M1) are added in blue color.

S/M1 = primary sensory-motor area; dlPFC = dorsolateral prefrontal cortex; FPAN = fronto-parietal attention network

7.3. General limitations

Several limitations have already been discussed in the discussion sections of the single experiments. Here, some limitations should be mentioned which are valid for general concepts of the present thesis. One of these limitations of the present experiments and the traditional illusory correlation paradigm in general is that the concept of illusory correlations is mostly measured by asking for relatively complex global estimations like "Given that you saw a spider, how frequently [in %] did a shock occur?". Questions like these may be difficult to answer because much information has to be integrated for an appropriate answer. The participant has to understand that only the total amount of spider trials should be considered as 100% and not the total amount

of all trials. For example, the participant might not read the question very attentively and misunderstand it as "How many of all shocks were distributed to the spider trials?". If all participants understand the question correctly, they may still apply different strategies to solve the problem. Some may try to remember concrete pictures with shocks, some may try to remember concrete pictures with no shocks as outcomes, some others may just guess without thinking much about it. All this variability of possible strategies may influence the answers to a very limited number (two to six) of global questions to specify the extent of an illusory correlation. Moreover, if the participant makes an error, this error will have a great influence on the illusory correlation because it cannot be compensated by the repeated measurement. That is why the illusory correlation was measured trial-by-trial in experiment 3. Every picture occurred only once in the experiment and afterwards, participants were asked to remember whether a shock was there or not. This question is very easy to understand and asks the participant exactly what to do. In addition, the illusory correlation measure was based on much more single points of measurement per participant, i.e. 60 instead of two. Although, this procedure may measure some different aspect of an illusory correlation than a global question, it is probably less susceptible to error variance and should be equally relevant for the return of fear, maybe even more relevant because it is assessed in the presence of the feared stimulus. In experiment 3, the global and the trial-by-trial index were significantly correlated which shows that the two indices are not independent from each other but probably describe very similar phenomena.

One major problem in investigating the influence of arousal that varies from category to category is the challenge to manipulate arousal without confounding arousal with stimulus meaning and stimulus valence. In the present thesis, arousal was manipulated by including positive arousing stimuli (e.g. erotic couples, sport scenes, babies, food) and comparing their impact on contingency estimates with neutral stimuli (household objects). Future experiments should improve the semantic matching between positive and neutral categories by using for example surprised arousing faces and neutral faces. Surprise is not clear in regard of valence (Tomkins & McCarter, 1964) and so arousal, valence and content should be less confounded. Another possibility would be to use picture categories of the same valence but of differential arousal, but studies of emotional experience show that valence and arousal are correlated, meaning that highly arousing negative images are normally perceived as more negative, too (Lang & Bradley, 2007).

Finally, a main assumption of the present work is that emotions modulate outcome aversiveness and that this aversiveness is one maintaining factor of illusory correlations. Empirically, the present experiments showed that self-reported outcome aversiveness was correlated with covariation bias and outcome aversiveness also causally induced covariation bias. However, the findings do not allow for an inference whether aversiveness per se is responsible for this effect or the increased salience of a more aversive outcome. Given the widely recognized fact that emotional stimuli attract attention (e.g. Bradley et al., 1993; Nummenmaa, Hyönä, & Calvo, 2006), it is very likely that the enhanced aversiveness leads to increased attentional engagement in the outcomes. To be clear, the proposed model of illusory correlations does not hypothesize an effect of outcome aversiveness which is completely independent from outcome salience. In fact, salience might be an important mediator between aversiveness and illusory correlations. Moreover, the pleasantness of positive outcomes might also be enhanced in some circumstances and become more salient. There is one study showing that low socially anxious individuals display a larger covariation bias between happy faces and positive pictures (e.g. a

butterfly) than high socially anxious individuals (Garner, Mogg, & Bradley, 2006). Considering interpretation biases in social anxiety (Frankling, Huppert, Langner, Leiberg, & Foa, 2005) it may be speculated that positive outcomes after happy faces may be interpreted as more positive by low socially anxious individuals, thus increasing pleasantness and salience of these outcomes and promoting covariation bias. In sum, outcome aversiveness and outcome salience cannot be disentangled on the basis of the present experiments, but future studies may identify salience as a mediator of outcome salience or pleasantness.

7.4. Outlook and clinical implications

Trial-by-trial illusory correlation index

As mentioned before, one general limitation of the present and previous illusory correlation experiments was that the illusory correlation index was based on a very limited number of relatively difficult and global questions. Future studies should use every picture only once in an experiment and ask for each picture after the experiment whether an aversive outcome had occurred after a given picture or not, like it was done in experiment 3. The assumption that this index is relevant for the return of fear after successful intervention just as a more global index (de Jong, van den Hout, & Merckelbach, 1995) should be tested in future treatment studies. A global illusory correlation index can still be used to compare the predictability of both measures.

Cognitive therapy for at-risk patients

The covariation bias as a psychological marker to predict relapse after treatment may be a useful tool to decide whether a patient needs additional exposure sessions, in which he or she is confronted with a feared object. Moreover, the presence of a covariation bias might be a sign that additional cognitive restructuring techniques are necessary to overcome fear more enduringly. Patients with remaining illusory correlations should be assigned to randomized controlled trials to receive either a prolonged treatment as usual or additional cognitive therapy to see if patients would profit from cognitive training and be free from fear for a longer time period. In addition, patients without remaining illusory correlations could receive the same procedure to check if cognitive therapy would specifically help phobic individuals with biased contingency estimates.

Such a cognitive intervention may target at the perceived aversiveness of feelings associated with phobic stimuli and attention training during exposure sessions. Although, exposure training is often considered as a learning period in which the unconditioned stimulus (US) is absent, it is often unclear what the US exactly is. Even in the absence of obvious harmful behavior of a feared animal, the phobic individual might suffer from his or her own subjective interpretation of the situation that determines the aversiveness of the experience. The individual might learn to accept or reappraise aversive feelings or experiences. Cognitive emotion regulation techniques have been shown to successfully alter the experience of emotional responses (Gross, 1998) and also phobic fear (Hermann et al., 2009). For example, acceptance and reappraisal of emotional experiences can reduce physiological arousal during a public speech relative to the suppression of the emotional experience (Hofmann, Heering, Sawyer, & Asnaani, 2009). The present experiments suggest that, if the aversiveness of

emotional experiences is altered during the exposure to a phobia-relevant stimulus, patients should be less likely to associate the phobic stimulus with aversive experiences and the unlearning of fear should be more successful.

Another target of the cognitive intervention might be the attention to non-feared outcomes. We know that working memory affects illusory correlations (Eder et al., 2011) and is related to contingency awareness in fear conditioning (Carter et al., 2003). The present experiment 3 indicated that brain regions (dlPFC, FPAN) typically involved in working memory, executive functions and attention are over-occupied in the presence of phobia-relevant stimuli (Curtis & D'Esposito, 2003; Ptak, 2012). This may be interpreted as that phobic experiences bind cognitive resources that are less available to recognize the absence of non-feared consequences and to use emotion regulation strategies. Psychotherapists may help patients to shift attention from feared consequences to non-feared consequences such as that the spider does not move as much as expected or the heart is not running as fast as expected. In addition, the improvement of working memory performance might be helpful to realize the availability of cognitive resources for emotion regulation and attention shifting. One possibility might be expressive writing before an exposure session. Recently, studies showed that expressive writing can improve working memory performance (Klein & Boals, 2001) and math test performance in test anxiety (Ramirez & Beilock, 2011). Typically, in expressive writing, one writes down his or her current feelings and worries for about ten minutes. This procedure is assumed to reduce on-going rumination and to eliminate the impact of anxiety on cognitive performance (Ramirez & Beilock, 2011).

Causal effect of fear and anxiety on biased cognitions

So far, the relationship between anxiety and cognitive biases has been supported by findings of correlations between these variables. Recently, the research of cognitive bias modification demonstrated that the modification of a cognitive bias can in fact causally influence anxiety (MacLeod & Mathews, 2012). Until now, only few studies investigated the backward path of this relationship and tested whether experimentally induced anxiety also causally changes cognitive processes. For example, it was shown that fear leads to reduced electrophysiological correlates of attention allocation to errors (Moser, Hajcak, & Simons, 2005). Moreover, conditioned fear leads to an attentional bias to conditioned stimuli (Armony & Dolan, 2002). However, the author of the present thesis is unaware of any studies showing that experimentally induced anxiety causes inflated expectancies of aversive events as predicted by the illusory correlations model (chapters 2.5 and 7.2). A major challenge in the experimental induction of fear is to induce fear without introducing a second feared stimulus that draws attention away from the expected aversive event, as realized in a threat of shock paradigm (Grillon & Ameli, 1998) or the presence of a spider (Moser et al., 2005). On the other hand, modulating fear of the expected aversive event by, for instance, cognitive appraisal of the aversive event would not allow for a discrimination between the effect of this cognitive appraisal and fear or anxiety as such. Maybe the causal effect of fear can be studied by manipulating one basic component of fear to see if biased expectancies would then be triggered. A previous experiment worked with the voluntary facial expression of fear and found an effect on visual performance (Susskind et al., 2008). Similarly, the facial expression of fear might be sufficient to induce cognitive changes such as inflated expectancies of aversive events. Particularly, widened eyes might be an adaptive evolved response to prepare the organism for the occurrence of a motivationally relevant event.

Memory and memory consolidation

One opportunity of the before mentioned method of presenting every stimulus only once and to ask participants to indicate whether a shock had been associated with a given picture is to measure the accuracy of this task. In particular, the traditional global assessment of a covariation bias only measures the ability to estimate the proportion of outcomes, but not whether participants can in fact remember single episodes of stimulus-outcome associations. Indeed, experiment 3 suggests that spider phobics were better at discriminating between trials with shocks and trials without shocks, although they were not explicitly asked to memorize the associations. Future studies may use larger samples and additional non-fear-relevant outcomes to see if this memory advantage is specific for phobia-relevant associations. Importantly, this memory advantage may also explain the return of fear and may be an important predictor of relapse. Moreover, it would be interesting to examine the consolidation process of this memory advantage. Sleep for example supports the memory consolidation of learned experiences (e.g. Diekelmann & Born, 2010). One experiment, for instance, demonstrated that emotionally negative objects were remembered better after a sleep period relative to a wake period of the same duration (Payne, Stickgold, Swanberg, & Kensinger, 2008). In addition, the memory accuracy for negative objects remained unchanged after sleep while neutral objects were forgotten. If the covariation bias is a useful predictor of the return of fear, it may be even more valid when assessed one day after the illusory correlation procedure.

Sex differences

Experiment 1 and 2 revealed sex differences in expectancy bias and covariation bias with women expecting more aversive sounds after negative pictures while men on average did not show any sign of biased expectancies or covariation estimates. This unexpected finding was based on exploratory analyses and should be replicated in future experiments. Nevertheless, the magnitude of the effect size and the fact that the sex difference observed in experiment 1 was replicated in experiment 2, along with the fact that the vast majority of previous covariation bias experiments included predominantly female participants (see Table 9.1) justifies a further examination of sex differences in biased contingency estimates. So far, only two covariation bias studies reported the analysis of sex differences. One study found stronger covariation bias in girls than in boys, but gender might have been confounded with fear level (Muris et al., 2007). Another study reported no differences between men and women (Davey et al., 2003). In contrast to these and the present investigations, future experiments should be explicitly designed to examine the role of gender and carefully match men and women regarding sample size, age, experience with psychological experiments and anxiety. Importantly, sex differences in cognitive biases might explain why women are more prone to anxiety disorders than men (Jacobi et al., 2014).

Illusory correlations as an approach to delusions?

A delusion is a “false belief based on incorrect inference about external reality that is firmly sustained despite what almost everybody else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary.” (DSM-IV, 1994). Thus, delusions might be interpreted as extreme consequences of

cognitive biases and the illusory correlations model could help to further understand why delusions emerge and are maintained. Previous experiments demonstrated that paranoid schizophrenics are more prone to illusory correlations between neutral words (Brennan & Hemsley, 1984) and are over-confident in probability estimates of future events (Huq, Garety, & Hemsley, 1988). Particularly, it would be interesting to find out whether illusory correlations in schizophrenia are due to seriously impaired contingency monitoring as findings of deficits in working memory might suggest (Goldman-Rakic, 1994). Besides, delusions with negative content like jealousy delusions or persecutory delusions may be a consequence of a general expectancy bias for aversive events (see experiment 4). Although the illusory correlations model is probably not sufficient to explain the whole picture and variety of delusions, the illusory correlation paradigm may be useful to complement cognitive theories of delusions (Bell, Halligan, & Ellis, 2006).

7.5. Conclusion

The present dissertation thesis about fear-relevant illusory correlations offers a model for the maintenance of this cognitive bias. In this model, evidence from previous studies and present experiments are integrated to identify maintaining psychological factors that might help to reduce biased cognitions and anxiety. Particularly, it is argued that the increased aversiveness (along with increased salience) of aversive consequences following feared stimuli is in part responsible for the fact that anxious individuals attach to their notion that feared objects are associated with aversive outcomes. For the first time, the present work demonstrates that illusory correlations are correlated not only with elevated self-reported aversiveness but also with enhanced sensory-motoric processing of fear-relevant outcomes. Moreover, evidence was found that the enhanced outcome aversiveness can be a causal factor inducing illusory correlations. A second important determinant of illusory correlations might be impaired contingency monitoring. This thesis shows that brain regions which are important for contingency awareness and attention, are enhanced and not reduced in response to phobic stimuli and predict illusory correlations. Together, the findings imply that cognitive reappraisal of potential aversive outcomes and an attention shift away from aversive outcomes to less salient non-aversive consequences may help to create a more realistic reflection of threat and to prevent the maintenance of fear and anxiety.

8. References

Abels, D. (1974). *Konzentrations-Verlaufs-Test*. Göttingen: Hogrefe.

Åhs, F., Pissiota, A., Michelgård, A., Frans, Ö., Furmark, T., Appel, L., & Fredrikson, M. (2009). Disentangling the web of fear: amygdala reactivity and functional connectivity in spider and snake phobia. *Psychiatry Research: Neuroimaging, 172*(2), 103–108.

Alloy, L. B., & Abramson, L. Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108*(4), 441-485.

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: the joint influence of prior expectations and current situational information. *Psychological Review, 91*(1), 112-149.

American Psychiatric Association, A., Association, A. P., & others. (1994). Diagnostic and statistical manual of mental disorders.

Amin, J. M., & Lovibond, P. F. (1997). Dissociations between covariation bias and expectancy bias for fear-relevant stimuli. *Cognition & Emotion, 11*(3), 273–289.

Amir, N., Beard, C., Burns, M., & Bomyea, J. (2009). Attention modification program in individuals with generalized anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 28.

Amrhein, C., Pauli, P., Dengler, W., & Wiedemann, G. (2005). Covariation bias and its physiological correlates in panic disorder patients. *Journal of Anxiety Disorders, 19*(2), 177–191.

Andreatta, M., Mühlberger, A., Yarali, A., Gerber, B., & Pauli, P. (2010). A rift between implicit and explicit conditioned valence in human pain relief learning. *Proceedings of the Royal Society B: Biological Sciences, 277*(1692), 2411–2416.

Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112*(1), 117.

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia, 40*(7), 817–826.

Aue, T., & Hoeppli, M.-E. (2012). Evidence for an encounter expectancy bias in fear of spiders. *Cognition & Emotion*, *26*(4), 727–736.

Aue, T., Hoeppli, M.-E., & Piguet, C. (2012). The sensitivity of physiological measures to phobic and nonphobic fear intensity. *Journal of Psychophysiology, 26*(4), 154–167.

Aupperle, R. L., Hale, L. R., Chambers, R. J., Cain, S. E., Barth, F. X., Sharp, S. C., & Savage, C. R. (2009) An fMRI study examining effects of acute D-cycloserine during symptom provocation in spider phobia. *CNS Spectrums 14*, 556–571.

Baas, D., Aleman, A., & Kahn, R. S. (2004) Lateralization of amygdala activation: a systematic review of functional neuroimaging studies. *Brain Research Brain Research Review 45*, 96–103.

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), 1-24.

Bar-Haim, Y., Morag, I., & Glickman, S. (2011). Training anxious children to disengage attention from threat: a randomized controlled trial. *Journal of Child Psychology and Psychiatry, 52*(8), 861–869.

Barlow, D. H., Chorpita, B. F., & Turovsky, J. (1996). Fear, panic, anxiety, and disorders of emotion.

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

Beck, A. T., & Dozois, D. J. (2011). Cognitive therapy: current status and future directions. *Annual Review of Medicine, 62*, 397–409.

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996) Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *Journal of Personality Assessment, 67*, 588–597.

Beckers, T., Krypotos, A.-M., Boddez, Y., Effting, M., & Kindt, M. (2013). What's wrong with fear conditioning? *Biological Psychology, 92*(1), 90–96.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). Explaining delusions: a cognitive perspective. *Trends in Cognitive Sciences, 10*(5), 219–226.

Benedek, M., & Kaernbach, C. (2010a). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods, 190*(1), 80–91.

Benedek, M., & Kaernbach, C. (2010b). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology, 47*(4), 647–658.

Berntson, G. G., Norman, G. J., Bechara, A., Bruss, J., Tranel, D., & Cacioppo, J. T. (2011). The insula and evaluative processes. *Psychological Science, 22*(1), 80–86.

Birbaumer, N., Elbert, T., Canavan, A. G., & Rockstroh, B. (1990). Slow potentials of the cerebral cortex and behavior. *Physiological Reviews, 70*(1), 1–41.

Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: Guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, *42*(1), 1–15.

Boddez, Y., Vervliet, B., Baeyens, F., Lauwers, S., Hermans, D., & Beckers, T. (2012). Expectancy bias in a selective conditioning procedure: trait anxiety increases the threat value of a blocked stimulus. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(2), 832–837.

Bortz, J., & Döring, N. (2006). *Forschungsmethoden und evaluation: für human-und sozialwissenschaftler*. Springer DE.

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion, 1*(3), 300-319.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1993). Pictures as prepulse: Attention and emotion in startle modification. *Psychophysiology, 30*(5), 541–545.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607.

Bradley, M. M., Moulder, B., & Lang, P. J. (2005). When Good Things Go Bad The Reflex Physiology of Defense. *Psychological Science, 16*(6), 468–473.

Brennan, J. H., & Hemsley, D. (1984). Illusory correlations in paranoid and non-paranoid schizophrenia. *British Journal of Clinical Psychology, 23*(3), 225–226.

Brenner, J. (2012). *Sind Supertaster emotional reaktiver? Schreckreflexmodulation und Pupillenreaktion während der Präsentation emotionaler Bilder.* University of Würzburg.

Brewer, W. F. (1974). There is no convincing evidence for operant or classical conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes*. (pp.450). Oxford, England: Lawrence Erlbaum.

Büchel, C., Morris, J., Dolan, R. J., & Friston, K. J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron, 20*(5), 947–957.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215–222.

Bushnell, M. C., Duncan, G. H., Hofbauer, R. K., Ha, B., Chen, J.-I., & Carrier, B. (1999) Pain perception: is there a role for primary somatosensory cortex? *Proceedings of the National Acadamy of Sciences, 96*, 7705–7709.

Butler, G., & Mathews, A. (1983). Cognitive processes in anxiety. *Advances in Behaviour Research and Therapy, 5*(1), 51–62.

Butler, G., & Mathews, A. (1987). Anticipatory anxiety and risk perception. *Cognitive Therapy and Research, 11*(5), 551–565.

Cahill, L., & McGaugh, J. L. (1998). Mechanisms of emotional arousal and lasting declarative memory. *Trends in Neurosciences, 21*(7), 294–299.

Canli, T., Desmond, J. E., Zhao, Z., & Gabrieli, J. D. (2002). Sex differences in the neural basis of emotional memories. *Proceedings of the National Academy of Sciences, 99*(16), 10789–10794.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J. D., Cahill, L., & others. (2000). Event-related activation in the human amygdala associates with later memory for individual emotional experience. *Journal of Neuroscience*, *20*(19), RC99–1.

Carlsson, K., Petersson, K. M., Lundqvist, D., Karlsson, A., Ingvar, M., & Öhman, A. (2004) Fear and the amygdala: manipulation of awareness generates differential cerebral responses to phobic and fear-relevant (but nonfeared) stimuli. *Emotion, 4*, 340–353.

Carter, R. M., Hofstötter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences, 100*(3), 1399–1404.

Carter, R. M., O'Doherty, J. P., Seymour, B., Koch, C., & Dolan, R. J. (2006). Contingency awareness in human aversive conditioning involves the middle frontal gyrus. *Neuroimage, 29*(3), 1007–1012.

Caseras, F., Fullana, M., Riba, J., Barbanoj, M., Aluja, A., & Torrubia, R. (2006). Influence of individual differences in the Behavioral Inhibition System and stimulus content (fear versus blood-disgust) on affective startle reflex modulation. *Biological Psychology, 72*(3), 251–256.

Caseras, X., Giampietro, V., Lamas, A., Brammer, M., Vilarroya, O., Carmona, S., … Mataix-Cols, D. (2010). The functional neuroanatomy of blood-injection-injury phobia: a comparison with spider phobics and healthy controls. *Psychological Medicine, 40*(1), 125.

Cavanagh, K., & Davey, G. C. (2000). UCS expectancy biases in spider phobics: Underestimation of aversive consequences following fear-irrelevant stimuli. *Behaviour Research and Therapy, 38*(7), 641–651.

Cavanagh, K., & Davey, G. C. (2001). The effect of mood and arousal on UCS expectancy biases. *Journal of Behavior Therapy and Experimental Psychiatry, 32*(1), 29–49.

Chan, C. K., & Lovibond, P. F. (1996). Expectancy bias in trait anxiety. Journal of Abnormal Psychology, 105(4), 637.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior, 6*(1), 151–155.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72*(3), 193-204.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74*(3), 271-280.

Christianson, S. (1992). Remembering emotional events: Potential mechanisms. *The Handbook of Emotion and Memory: Research and Theory*, 307–340.

Clark, D. A., & Beck, A. T. (2010). Cognitive theory and therapy of anxiety and depression: convergence with neurobiological findings. *Trends in Cognitive Sciences, 14*(9), 418–424.

Clark, M. S., Milberg, S., & Ross, J. (1983). Arousal cues arousal-related material in memory: Implications for understanding effects of mood on memory. *Journal of Verbal Learning and Verbal Behavior, 22*(6), 633–649.

Clark, R. E., & Squire, L. R. (1998). Classical conditioning and brain systems: the role of awareness. Science, 280(5360), 77–81.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Collet, C., Vernet-Maury, E., Delhomme, G., & Dittmar, A. (1997). Autonomic nervous system response patterns specificity to basic emotions. *Journal of the Autonomic Nervous System, 62*(1-2), 45–57.

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., … Heller, W. (2003). Paying attention to emotion. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(2), 81–96.

Connolly, K. M., Lohr, J. M., Olatunji, B. O., Hahn, K. S., & Williams, N. L. (2009). Information processing in contamination fear: A covariation bias examination of fear and disgust. *Journal of Anxiety Disorders, 23*(1), 60–68.

Coull, J., Frith, C., Frackowiak, R. S. J., & Grasby, P. (1996). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory. *Neuropsychologia*, *34*(11), 1085–1095.

Cooney, R. E., Joormann, J., Eugène, F., Dennis, E. L., & Gotlib, I. H. (2010). Neural correlates of rumination in depression. *Cognitive, Affective, & Behavioral Neuroscience, 10*(4), 470–478.

Craig, A. D., Chen, K., Bandy, D., & Reiman, E. M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience, 3*(2), 184–190.

Cui, R. Q., Egkher, A., Huter, D., Lang, W., Lindinger, G., & Deecke, L. (2000). High resolution spatiotemporal analysis of the contingent negative variation in simple or complex motor tasks and a non-motor task. *Clinical Neurophysiology, 111*(10), 1847–1859.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7*(9), 415–423.

Dandeneau, S. D., & Baldwin, M. W. (2009). The buffering effects of rejection-inhibiting attentional training on social and performance threat among adult students. *Contemporary Educational Psychology, 34*(1), 42–50.

Davey, G. C. (1992). Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Advances in Behaviour Research and Therapy, 14*(1), 29–66.

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

Davey, G. C., & Craigie, P. (1997). Manipulation of dangerousness judgements to fear-relevant stimuli: Effects on a priori UCS expectancy and a posteriori covariation assessment. *Behaviour Research and Therapy, 35*(7), 607–617.

Davey, G. C., & Dixon, A. L. (1996). The expectancy bias model of selective associations: The relationship of judgments of CS dangerousness, CS-UCS similarity and prior fear to a priori and a posteriori covariation assessments. *Behaviour Research and Therapy, 34*(3), 235–252.

Davey, G. C., Cavanagh, K., & Lamb, A. (2003). Differential aversive outcome expectancies for high-and low-predation fear-relevant animals. *Journal of Behavior Therapy and Experimental Psychiatry, 34*(2), 117–128.

Davis, M., Walker, D. L., Miles, L., & Grillon, C. (2010). Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology, 35*(1), 105–135.

Dawkins, R., & Krebs, J. R. (1979). Arms races between and within species. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *205*(1161), 489–511.

Deinzer, R., Steyer, R., Eid, M., Notz, P., Schwenkmezger, P., Ostendorf, F., & Neubauer, A. (1995). Situational effects in trait assessment: The FPI, NEOFFI, and EPI questionnaires. *European Journal of Personality*, *9*(1), 1–23.

De Jong, P. J. (1993). Covariation bias in phobia: Mere resistance to preexperimental expectancies? *Behavior Therapy, 24*(3), 447–454.

De Jong, P. J., & Merckelbach, H. (1991). Covariation bias and electrodermal responding in spider phobics before and after behavioural treatment. *Behaviour Research and Therapy, 29*(4), 307–314.

De Jong, P. J., & Merckelbach, H. (1993). Covariation bias, classical conditioning, and phobic fear. *Integrative Physiological and Behavioral Science, 28*(2), 167–170.

De Jong, P. J., & Merckelbach, H. (2000). Phobia-relevant illusory correlations: the role of phobic responsivity. *Journal of Abnormal Psychology, 109*(4), 597-601.

De Jong, P. J., & Peters, M. L. (2007a). Blood-injection-injury fears: Harm-vs. disgust-relevant selective outcome associations. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(3), 263–274.

De Jong, P. J., & Peters, M. L. (2007b). Contamination vs. harm-relevant outcome expectancies and covariation bias in spider phobia. *Behaviour Research and Therapy, 45*(6), 1271–1284.

De Jong, P. J., Merckelbach, H., & Arntz, A. (1990). Illusory correlation, on-line probability estimates, and electrodermal responding in a (quasi)-conditioning paradigm. *Biological Psychology, 31*(3), 201–212.

De Jong, P. J., Merckelbach, H., & Arntz, A. (1995). Covariation bias in phobic women: The relationship between a priori expectancy, on-line expectancy, autonomic responding, and a posteriori contingency judgment. *Journal of Abnormal Psychology, 104*(1), 55-62.

De Jong, P. J., Merckelbach, H., & Nijman, H. (1995). Hemisphere preference, anxiety, and covariation bias. *Personality and Individual Differences, 18*(3), 363–371.

De Jong, P. J., Merckelbach, H., Arntz, A., & Nijmam, H. (1992). Covariation detection in treated and untreated spider phobics. *Journal of Abnormal Psychology, 101*(4), 724-727.

De Jong, P. J., Merckelbach, H., Bögels, S., & Kindt, M. (1998). Illusory correlation and social anxiety. *Behaviour Research and Therapy, 36*(11), 1063–1073.

De Jong, P. J., Van Den Hout, M. A., & Merckelbach, H. (1995). Covariation bias and the return of fear. *Behaviour Research and Therapy, 33*(2), 211–213.

Delgado, M. R., Gillis, M. M., & Phelps, E. A. (2008). Regulating the expectation of reward via cognitive strategies. *Nature Neuroscience, 11*(8), 880–881.

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1511), 3787–3800.

Derbyshire, S. W. (2003). A systematic review of neuroimaging data during visceral stimulation. *The American Journal of Gastroenterology, 98*(1), 12–20.

Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology, 111*(2), 225.

Diamond, D., Matchett, G., & Davey, G. C. (1995). The effect of prior fear levels on UCS-expectancy ratings to a fear-relevant stimulus. *The Quarterly Journal of Experimental Psychology, 48*(1), 237–247.

Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience, 11*(2), 114–126.

Dolcos, F., & Cabeza, R. (2002). Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 252–263.

Dolcos, F., & McCarthy, G. (2006). Brain systems mediating cognitive interference by emotional distraction. *The Journal of Neuroscience, 26*(7), 2072–2079.

Domes, G., Schulze, L., Böttger, M., Grossmann, A., Hauenstein, K., Wirtz, P. H., … Herpertz, S. C. (2010). The neural correlates of sex differences in emotional reactivity and emotion regulation. *Human Brain Mapping, 31*(5), 758–769.

Dowman, R. (2001). Attentional set effects on spinal and supraspinal responses to pain. *Psychophysiology, 38*(3), 451–464.

Dozois, D. J., Dobson, K. S., & Ahnberg, J. L. (1998). A psychometric evaluation of the Beck Depression Inventory–II. *Psychological Assessment*, *10*(2), 83.

Dresler, T., Attar, C. H., Spitzer, C., Löwe, B., Deckert, J., Büchel, C., … Fallgatter, A. J. (2012). Neural correlates of the emotional Stroop task in panic disorder patients: An event-related fMRI study. *Journal of Psychiatric Research*, *46*(12), 1627–1634.

Eder, A. B., Fiedler, K., & Hamm-Eder, S. (2011). Illusory correlations revisited: The role of pseudocontingencies and working-memory capacity. *The Quarterly Journal of Experimental Psychology, 64*(3), 517–532.

Ehlers, A., Margraf, J., Roth, W. T., Taylor, C. B., & Birbaumer, N. (1988). Anxiety induced by false heart rate feedback in patients with panic disorder. *Behaviour Research and Therapy, 26*(1), 1–11.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221*(4616), 1208–1210.

Engelhard, I. M., de Jong, P. J., van den Hout, M. A., & van Overveld, M. (2009). Expectancy bias and the persistence of posttraumatic stress. *Behaviour Research and Therapy, 47*(10), 887–892.

Epstein, M. (2004). Thoughts without a thinker: Psychotherapy from a Buddhist perspective. New York: Basic Books.

Epstein, S. (1972). The nature of anxiety with emphasis upon its relationship to expectancy. *Anxiety: Current Trends in Theory and Research, 2*, 291–337.

Etkin, A., & Wager, T. D. (2007). Functional Neuroimaging of Anxiety: A Meta-Analysis of Emotional Processing in PTSD, Social Anxiety Disorder, and Specific Phobia. *American Journal of Psychiatry, 164*(10), 1476–1488.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336.

Fadok, J. P., Dickerson, T. M., & Palmiter, R. D. (2009). Dopamine is necessary for cue-dependent fear conditioning. *The Journal of Neuroscience*, *29*(36), 11089–11097.

Fanselow, M. S. (1986). Conditioned Fear-Induced Opiate Analgesia: A Competing Motivational State Theory of Stress Analgesiaa. *Annals of the New York Academy of Sciences*, *467*(1), 40–54.

Fischer, H., Wright, C. I., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2003). Brain habituation during repeated exposure to fearful and neutral faces: a functional MRI study. *Brain Research Bulletin, 59*(5), 387–392.

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition & Emotion, 16*(3), 355–379.

Franklin, M. E., Huppert, J., Langner, R., Leiberg, S., & Foa, E. B. (2005). Interpretation Bias: A Comparison of Treated Social Phobics, Untreated Social Phobics, and Controls1. *Cognitive Therapy and Research, 29*(3), 289–300.

Fregni, F., Boggio, P. S., Nitsche, M., Bermpohl, F., Antal, A., Feredoes, E., … others. (2005). Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Experimental Brain Research, 166*(1), 23–30.

Friedman, B. H., Thayer, J. F., & Borkovec, T. D. (2000). Explicit memory bias for threat words in generalized anxiety disorder. *Behavior Therapy, 31*(4), 745–756.

Galli, G., Wolpe, N., & Otten, L. J. (2011). Sex differences in the use of anticipatory brain activity to encode emotional events. *The Journal of Neuroscience, 31*(34), 12364–12370.

Gard, M. G., & Kring, A. M. (2007). Sex differences in the time course of emotion. *Emotion, 7*(2), 429.

Garner, M., Mogg, K., & Bradley, B. P. (2006). Fear-relevant selective associations and social anxiety: Absence of a positive bias. *Behaviour Research and Therapy, 44*(2), 201–217.

Gater, R., Tansella, M., Korten, A., Tiemens, B. G., Mavreas, V. G., & Olatawura, M. O. (1998). Sex differences in the prevalence and detection of depressive and anxiety disorders in general health care settings: report from the World Health Organization Collaborative Study on Psychological Problems in General Health Care. *Archives of General Psychiatry, 55*(5), 405–413.

Gerdes, A., Alpers, G. W., & Pauli, P. (2008). When spiders appear suddenly: Spider-phobic patients are distracted by task-irrelevant spiders. *Behaviour Research and Therapy, 46*(2), 174–187.

Glotzbach-Schoon, E., Tadda, R., Andreatta, M., Tröger, C., Ewald, H., Grillon, C., … Mühlberger, A. (2013). Enhanced discrimination between threatening and safe contexts in high-anxious individuals. *Biological Psychology, 93*(1), 159–166.

Goldin, P. R., Manber, T., Hakimi, S., Canli, T., & Gross, J. J. (2009). Neural bases of social anxiety disorder: emotional reactivity and cognitive regulation during social and physical threat. *Archives of General Psychiatry, 66*(2), 170–180.

Goldman-Rakic, P. S. (1994). Working memory dysfunction in schizophrenia. *The Journal of Neuropsychiatry and Clinical Neurosciences, 6*(4), 348-357.

Goosens, K. A. (2011). Hippocampal regulation of aversive memories. *Current Opinion in Neurobiology*, *21*(3), 460–466.

Goossens, L., Sunaert, S., Peeters, R., Griez, E. J., & Schruers, K. R. (2007). Amygdala hyperfunction in phobic fear normalizes after exposure. *Biological Psychiatry, 62*(10), 1119–1125.

Graham, J., & Gaffan, E. A. (1997). Fear of water in children and adults: Etiology and familial effects. *Behaviour Research and Therapy, 35*(2), 91–108.

Grillon, C., & Ameli, R. (1998). Effects of threat and safety signals on startle during anticipation of aversive shocks, sounds, or airblasts. *Journal of Psychophysiology, 12*(4), 329-337.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271-299.

Grupe, D. W., & Nitschke, J. B. (2011). Uncertainty is associated with biased expectancies and heightened responses to aversion. *Emotion, 11*(2), 413-424.

Hajcak, G., Anderson, B., Arana, A., Borckardt, J., Takacs, I., George, M., & Nahas, Z. (2010). Dorsolateral prefrontal cortex stimulation modulates electrocortical measures of visual attention: evidence from direct bilateral epidural cortical stimulation in treatment-resistant mood disorder. *Neuroscience*, *170*(1), 281–288.

Halberstadt, J. B., & Niedenthal, P. M. (1997). Emotional state and the use of stimulus dimensions in judgment. *Journal of Personality and Social Psychology, 72*(5), 1017.

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, *137*(6), 940.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12*(4), 392–407.

Harmon-Jones, E. (2007). Trait anger predicts relative left frontal cortical activation to anger-inducing stimuli. *International Journal of Psychophysiology, 66*(2), 154–160.

Harris, J. A., Miniussi, C., Harris, I. M., & Diamond, M. E. (2002) Transient storage of a tactile memory trace in primary somatosensory cortex. *The Journal of Neuroscience, 22*, 8720–8725.

Hendersen, R. W. (1985). Fearful memories: The motivational significance of forgetting. *Affect, Conditioning, and Cognition: Essays on the Determinants of Behavior*, 43–53.

Hermann, A., Schäfer, A., Walter, B., Stark, R., Vaitl, D., & Schienle, A. (2009). Emotion regulation in spider phobia: role of the medial prefrontal cortex. *Social Cognitive and Affective Neuroscience, 4*(3), 257–267.

Hermann, C., Ofer, J., & Flor, H. (2004). Covariation bias for ambiguous social stimuli in generalized social phobia. *Journal of Abnormal Psychology, 113*(4), 646-653.

Hettema, J. M., Annas, P., Neale, M. C., Kendler, K. S., & Fredrikson, M. (2003). A twin study of the genetics of fear conditioning. *Archives of General Psychiatry, 60*(7), 702–708.

Hofmann, S. G., & Smits, J. A. (2008). Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry, 69*(4), 621-632.

Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: The effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour Research and Therapy, 47*(5), 389–394.

Honeybourne, C., Matchett, G., & Davey, G. C. (1993). Expectancy models of laboratory preparedness effects: A UCS-expectancy bias in phylogenetic and ontogenetic fear-relevant stimuli. *Behavior Therapy, 24*(2), 253–264.

Hosoba, T., Iwanaga, M., & Seiwa, H. (2001). The effect of UCS inflation and deflation procedures on “fear”conditioning. *Behaviour Research and Therapy*, *39*(4), 465–475.

Hugdahl, K., & Öhman, A. (1977). Effects of instruction on acquisition and extinction of electrodermal responses to fear-relevant stimuli. *Journal of Experimental Psychology: Human Learning and Memory, 3*(5), 608.

Hugdahl, K., Fredrikson, M., & Öhman, A. (1977). “Preparedness” and “arousability”as determinants of electrodermal conditioning. *Behaviour Research and Therapy, 15*(4), 345–353.

Huq, S., Garety, P., & Hemsley, D. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology, 40*(4), 801–812.

Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002) Image distortion correction in fMRI: a quantitative evaluation. *Neuroimage, 16*, 217–240.

Indovina, I., Robbins, T. W., Núñez-Elizalde, A. O., Dunn, B. D., & Bishop, S. J. (2011). Fear-conditioning mechanisms associated with trait vulnerability to anxiety in humans. *Neuron, 69*(3), 563–571.

Irwin, D. A., Knott, J. R., McAdam, D. W., & Rebert, C. S. (1966). Motivational determinants of the "contingent negative variation." *Electroencephalography and Clinical Neurophysiology, 21*(6), 538–543.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887.

Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., … others. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: the Mental Health Module of the German Health Interview and Examination Survey for Adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*.

Johanson, A., Gustafson, L., Passant, U., Risberg, J., Smith, G., Warkentin, S., & Tucker, D. (1998). Brain function in spider phobia. *Psychiatry Research: Neuroimaging, 84*(2), 101–111.

Kennedy, S. J., Rapee, R. M., & Mazurski, E. J. (1997). Covariation bias for phylogenetic versus ontogenetic fear-relevant stimuli. *Behaviour Research and Therapy, 35*(5), 415–422.

Kenntner-Mabiala, R., & Pauli, P. (2005). Affective modulation of brain potentials to painful and nonpainful stimuli. *Psychophysiology, 42*(5), 559–567.

Kenntner-Mabiala, R., Andreatta, M., Wieser, M. J., Mühlberger, A., & Pauli, P. (2008). Distinct effects of attention and affect on pain perception and somatosensory evoked potentials. *Biological Psychology, 78*(1), 114–122.

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research, 21*(3), 169–184.

Kiley Hamlin, J., Wynn, K., & Bloom, P. (2010). Three-month-olds show a negativity bias in their social evaluations. *Developmental Science*, *13*(6), 923–929.

Kim, J. J., & Jung, M. W. (2006). Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neuroscience & Biobehavioral Reviews, 30*(2), 188–202.

Kirsch, I., Lynn, S. J., Vigorito, M., & Miller, R. R. (2004). The role of cognition in classical and operant conditioning. *Journal of Clinical Psychology, 60*(4), 369–392.

Klein, K., & Boals, A. (2001). Expressive writing can increase working memory capacity. *Journal of Experimental Psychology: General, 130*(3), 520-533.

Knoch, D., Brugger, P., & Regard, M. (2005) Suppressing versus releasing a habit: Frequency-dependent effects of prefrontal transcranial magnetic stimulation. *Cerebral Cortex, 15*, 885–887.

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: expression, experience, and physiology. *Journal of Personality and Social Psychology, 74*(3), 686.

Labus, J. S., Naliboff, B. N., Fallon, J., Berman, S. M., Suyenobu, B., Bueller, J. A., … Mayer, E. A. (2008). Sex differences in brain activity during aversive visceral stimulation and its expectation in patients with chronic abdominal pain: a network analysis. *Neuroimage, 41*(3), 1032–1043.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry A Window to the Preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, attention, and the startle reflex. *Psychological Review*, *97*(3), 377.

Lang, P. J., Bradley, M. M., Cuthbert, B. N., & others. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual.* NIMH, Center for the Study of Emotion & Attention.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273.

Lang, P., & Bradley, M. M. (2007). The International Affective Picture System (IAPS) in the study of emotion and attention. *Handbook of Emotion Elicitation and Assessment, 29*.

Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality and Social Psychology, 32*(6), 951–955.

Larson, C. L., Schaefer, H. S., Siegle, G. J., Jackson, C. A., Anderle, M. J., & Davidson, R. J. (2006). Fear is fast in phobic individuals: amygdala activation in response to fear-relevant stimuli. *Biological Psychiatry, 60*(4), 410–417.

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

Lee, T., Liu, H., Chan, C., Fang, S., & Gao, J. (2005). Neural activities associated with emotion recognition observed in men and women. *Molecular Psychiatry, 10*(5), 450–455.

Legrain, V., Iannetti, G. D., Plaghki, L., & Mouraux, A. (2011). The pain matrix reloaded: a salience detection system for the body. *Progress in Neurobiology, 93*(1), 111–124.

Lemon, N., Aydin-Abidin, S., Funke, K., & Manahan-Vaughan, D. (2009). Locus coeruleus activation facilitates memory encoding and induces hippocampal LTD that depends on β-adrenergic receptor activation. *Cerebral Cortex*, *19*(12), 2827–2837.

Lewinsohn, P. M., & Graf, M. (1973). Pleasant activities and depression. *Journal of Consulting and Clinical Psychology, 41*(2), 261.

Lieberman, M. D., & Cunningham, W. A. (2009) Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience, 4*, 423–428.

Linares, I. M., Trzesniak, C., Chagas, M. H. N., Hallak, J. E., Nardi, A. E., & Crippa, J. A. S. (2012). Neuroimaging in specific phobia disorder: a systematic review of the literature. *Revista Brasileira de Psiquiatria, 34*(1), 101–111.

Lipka, J., Hoffmann, M., Miltner, W. H., & Straube, T. (2013). Effects of Cognitive-Behavioral Therapy on Brain Responses to Subliminal and Supraliminal Threat and Their Functional Significance in Specific Phobia. *Biological Psychiatry*.

Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron*, *46*(5), 703–713.

Lissek, S., Pine, D. S., & Grillon, C. (2006). The strong situation: A potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology, 72*(3), 265–270.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy, 43*(11), 1391–1424.

Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., … Grillon, C. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behaviour Research and Therapy, 47*(2), 111–118.

Loggia, M. L., Mogil, J. S., & Bushnell, M. C. (2008). Experimentally induced mood changes preferentially affect pain unpleasantness. *The Journal of Pain, 9*(9), 784–791.

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes, 28*(1), 3-26.

Lubow, R. E. (1973). Latent inhibition. *Psychological Bulletin, 79*(6), 398.

Lüderitz, B., Jung, W., Deister, A., & Manz, M. (1994). Patient acceptance of implantable cardioverter defibrillator devices: changing attitudes. *American Heart Journal, 127*(4), 1179–1184.

Lueken, U., Hilbert, K., Stolyar, V., Maslowski, N. I., Beesdo-Baum, K., & Wittchen, H.-U. (2013). Neural substrates of defensive reactivity in two subtypes of specific phobia. *Social Cognitive and Affective Neuroscience*.

Lueken, U., Kruschwitz, J. D., Muehlhan, M., Siegert, J., Hoyer, J., & Wittchen, H.-U. (2011). How specific is specific phobia? Different neural response patterns in two subtypes of specific phobia. *Neuroimage, 56*(1), 363–372.

Macht, M., & Müller, J. (2007). Increased negative emotional responses in PROP supertasters. *Physiology & Behavior, 90*(2), 466–472.

Mackiewicz, K. L., Sarinopoulos, I., Cleven, K. L., & Nitschke, J. B. (2006). The effect of anticipation and the specificity of sex differences for amygdala and hippocampus function in emotional memory. *Proceedings of the National Academy of Sciences, 103*(38), 14200–14205.

MacLeod, C. (2012). Cognitive bias modification procedures in the management of mental disorders. *Current Opinion in Psychiatry*, *25*(2), 114–120.

MacLeod, C., & Cohen, I. L. (1993). Anxiety and the interpretation of ambiguity: a text comprehension study. *Journal of Abnormal Psychology, 102*(2), 238-247.

MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology, 8*, 189–217.

Makar, A. B., McMartin, K. E., Palese, M., & Tephly, T. R. (1975). Formate assay in body fluids: application in methanol poisoning. *Biochemical Medicine, 13*(2), 117–126.

Makkar, S. R., & Grisham, J. R. (2013). The Effect of Post-Event Processing on Mood, Self-Beliefs, and Interpretive Biases. *Journal of Experimental Psychopathology*, *4*(4), 368–386.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19*, 1233-1239.

Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual Review of Neuroscience, 24*(1), 897–931.

Martin, D. J., Abramson, L. Y., & Alloy, L. B. (1984). Illusion of control for self and others in depressed and nondepressed college students. *Journal of Personality and Social Psychology, 46*(1), 125-136.

Mather, M., Mitchell, K. J., Raye, C. L., Novak, D. L., Greene, E. J., & Johnson, M. K. (2006). Emotional arousal can impair feature binding in working memory. *Journal of Cognitive Neuroscience, 18*(4), 614–625.

Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research, 22*(6), 539–560.

Mathews, A., & MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology, 45*(1), 25–50.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annual Review of Clinical Psychology, 1*, 167–195.

McNally, R. J., & Heatherton, T. F. (1993). Are covariation biases attributable to a priori expectancy biases? *Behaviour Research and Therapy, 31*(7), 653–658.

McRae, K., Ochsner, K. N., Mauss, I. B., Gabrieli, J. J., & Gross, J. J. (2008). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. *Group Processes & Intergroup Relations, 11*(2), 143–162.

Meagher, M. W., Arnau, R. C., & Rhudy, J. L. (2001). Pain and emotion: effects of affective picture modulation. *Psychosomatic Medicine, 63*(1), 79–90.

Menzies, R. G., & Clarke, J. C. (1995). The etiology of phobias: A nonassociative account. *Clinical Psychology Review, 15*(1), 23–48.

Michael, T., Blechert, J., Vriends, N., Margraf, J., & Wilhelm, F. H. (2007). Fear conditioning in panic disorder: enhanced resistance to extinction. *Journal of Abnormal Psychology, 116*(3), 612-617.

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the comprehensive system. *Psychological Bulletin, 139*(3), 548-605.

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the International Affective Picture System. *Behavior Research Methods, 37*(4), 626–630.

Miller, E. K., & Cohen, J. D. (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Mineka, S., & Oehlberg, K. (2008) The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica, 127*, 567–580.

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102*(2), 304.

Moratti, S., Keil, A., & Stolarova, M. (2004). Motivated attention in emotional picture processing is reflected by activity modulation in cortical attention networks. *Neuroimage*, *21*(3), 954–964.

Moser, J. S., Hajcak, G., & Simons, R. F. (2005). The effects of fear on performance monitoring and attentional allocation. *Psychophysiology, 42*(3), 261–268.

Mouraux, A., Diukova, A., Lee, M. C., Wise, R. G., & Iannetti, G. D. (2011). A multisensory investigation of the functional significance of the "pain matrix." *Neuroimage, 54*(3), 2237–2249.

Mühlberger, A., Wiedemann, G., Herrmann, M. J., & Pauli, P. (2006). Phylo-and ontogenetic fears and the expectation of danger: differences between spider-and flight-phobic subjects in cognitive and physiological responses to disorder-specific stimuli. *Journal of Abnormal Psychology, 115*(3), 580-589.

Mulkens, S. A., de Jong, P. J., & Merckelbach, H. (1996). Disgust and spider phobia. *Journal of Abnormal Psychology, 105*(3), 464-468.

Muris, P., De Jong, P. J., Meesters, C., Waterreus, B., & vanLubeck, J. (2005). An Experimental Study of Spider-Related Covariation Bias in 8- to 13-Year-Old Children. *Child Psychiatry and Human Development, 35*(3), 185–201.

Muris, P., Huijding, J., Mayer, B., den Breejen, E., & Makkelie, M. (2007). Spider fear and covariation bias in children and adolescents. *Behaviour Research and Therapy, 45*(11), 2604–2615.

Nummenmaa, L., Hyönä, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion, 6*(2), 257.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1*(1), 29–50.

Öhman, A. (2008). Fear and anxiety: Overlaps and dissociations. In *Handbook of emotions (3rd ed.)* (pp. 709 – 728). Guilford Press.

Öhman, A. (2009). Of snakes and faces: An evolutionary perspective on the psychology of fear. *Scandinavian Journal of Psychology, 50*(6), 543–552.

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

Öhman, A., Eriksson, A., & Olofsson, C. (1975). One-trial learning and superior resistance to extinction of autonomic responses conditioned to potentially phobic stimuli. *Journal of Comparative and Physiological Psychology, 88*(2), 619.

Olatunji, B. O., Cisler, J. M., Meunier, S., Connolly, K., & Lohr, J. M. (2008). Expectancy bias for fear and disgust and behavioral avoidance in spider fearful individuals. *Cognitive Therapy and Research, 32*(3), 460–469.

Paquette, V., Lévesque, J., Mensour, B., Leroux, J.-M., Beaudoin, G., Bourgouin, P., & Beauregard, M. (2003) Change the mind and you change the brain: effects of cognitive-behavioral therapy on the neural correlates of spider phobia. *Neuroimage, 18*, 401–409.

Pardo, J. V., Fox, P. T., & Raichle, M. E. (1991). Localization of a human system for sustained attention by positron emission tomography. *Nature*, *349*(6304), 61–64.

Pasternak, T., & Greenlee, M. W. (2005) Working memory in primate sensory systems. *Nature Review Neuroscience, 6*, 97–107.

Pauli, P., Diedrich, O., & Müller, A. (2002). Covariation bias in the affect-modulated startle paradigm. *Journal of Behavior Therapy and Experimental Psychiatry, 33*(3), 191–202.

Pauli, P., Montoya, P., & Martz, G.-E. (1996). Covariation bias in panic-prone individuals. *Journal of Abnormal Psychology, 105*(4), 658-662.

Pauli, P., Montoya, P., & Martz, G.-E. (2001). On-line and a posteriori covariation estimates in panic-prone individuals: Effects of a high contingency of shocks following fear-irrelevant stimuli. *Cognitive Therapy and Research, 25*(1), 23–36.

Pauli, P., Wiedemann, G., & Montoya, P. (1998). Covariation bias in flight phobics. *Journal of Anxiety Disorders, 12*(6), 555–565.

Pauli, P., Wiedemann, G., Dengler, W., & Kühlkamp, V. (2001). A priori expectancy bias and its relation to shock experience and anxiety: A naturalistic study in patients with an automatic implantable cardioverter defibrillator. *Journal of Behavior Therapy and Experimental Psychiatry, 32*(3), 159–171.

Paus, T. (1996). Location and function of the human frontal eye-field: a selective review. *Neuropsychologia*, *34*(6), 475–483.

Pavlov, I. P. (2003). *Conditioned reflexes*. Courier Dover Publications.

Payne, J. D., Stickgold, R., Swanberg, K., & Kensinger, E. A. (2008). Sleep preferentially enhances memory for emotional components of scenes. *Psychological Science, 19*(8), 781–788.

Peyron, R., Laurent, B., & Garcia-Larrea, L. (2000) Functional imaging of brain responses to pain. A review and meta-analysis. *Clinical Neurophysiology, 30*, 263–288.

Pezze, M. A., & Feldon, J. (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Progress in Neurobiology*, *74*(5), 301–320.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage, 16*(2), 331–348.

Phelps, E. A., O'Connor, K. J., Gatenby, J. C., Gore, J. C., Grillon, C., & Davis, M. (2001). Activation of the left amygdala to a cognitive representation of fear. *Nature Neuroscience, 4*(4), 437–441.

Pissiota, A., Frans, Ö., Michelgård, Å., Appel, L., Långström, B., Flaten, M. A., & Fredrikson, M. (2003). Amygdala and anterior cingulate cortex activation during affective startle modulation: a PET study of fear. *European Journal of Neuroscience, 18*(5), 1325–1331.

Poulton, R., & Menzies, R. G. (2002). Non-associative fear acquisition: A review of the evidence from retrospective and longitudinal research. *Behaviour Research and Therapy, 40*(2), 127–149.

Poulton, R., Davies, S., Menzies, R. G., Langley, J. D., & Silva, P. A. (1998). Evidence for a non-associative model of the acquisition of a fear of heights. *Behaviour Research and Therapy, 36*(5), 537–544.

Poulton, R., Thomson, W. M., Davies, S., Kruger, E., Brown, R. H., & Silva, P. (1997). Good teeth, bad teeth and fear of the dentist. *Behaviour Research and Therapy, 35*(4), 327–334.

Ptak, R. (2012). The Frontoparietal Attention Network of the Human Brain Action, Saliency, and a Priority Map of the Environment. *The Neuroscientist, 18*(5), 502–515.

Pury, C. L., & Mineka, S. (1997). Covariation bias for blood-injury stimuli and aversive outcomes. *Behaviour Research and Therapy, 35*(1), 35–47.

Rachman, S. (1977). The conditioning theory of fearacquisition: A critical examination. *Behaviour Research and Therapy, 15*(5), 375–387.

Rachman, S. (1989). The return of fear: Review and prospect. *Clinical Psychology Review, 9*(2), 147–168.

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology, 61*(1), 5–18.

Rajaram, S. (1998). The effects of conceptual salience and perceptual distinctiveness on conscious recollection. *Psychonomic Bulletin & Review, 5*(1), 71–78.

Ramirez, G., & Beilock, S. L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science, 331*(6014), 211–213.

Rauch, S. L., Wright, C. I., Martis, B., Busa, E., McMullin, K. G., Shin, L. M., … Fischl, B. (2004). A magnetic resonance imaging study of cortical thickness in animal phobia. *Biological Psychiatry, 55*(9), 946–952.

Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain, 66*(1), 3–8.

Reicherts, P., Gerdes, A., Pauli, P., & Wieser, M. J. (2013) On the mutual effects of pain and emotion: facial pain expressions enhance pain perception and vice versa are perceived as more arousing when feeling pain. *Pain, 154*, 793-800.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy, 24*(1), 1–8.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, *38*(1), 119 – 125.

Rhudy, J. L., & Meagher, M. W. (2000). Fear and anxiety: divergent effects on human pain thresholds. *Pain, 84*(1), 65–75.

Rinck, M., Bundschuh, S., Engler, S., Müller, A., Wissmann, J., Ellwart, T., & Becker, E. S. (2002) Reliabilität und Validität dreier Instrumente zur Messung von Angst vor Spinnen. *Diagnostica, 48*, 141–149.

Rock, I., & Fleck, F. S. (1950). A re-examination of the effect of monetary reward and punishment on figure-ground perception. *Journal of Experimental Psychology, 40*(6), 766.

Roy, M., Piché, M., Chen, J.-I., Peretz, I., & Rainville, P. (2009). Cerebral and spinal modulation of pain by emotions. *Proceedings of the National Academy of Sciences, 106*(49), 20900–20905.

Russo, R., Fox, E., Bellinger, L., & Nguyen-Van-Tam, D. P. (2001). Mood-congruent free recall bias in anxiety. *Cognition & Emotion, 15*(4), 419–433.

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional perception: correlation of functional MRI and event-related potentials. *Cerebral Cortex*, *17*(5), 1085–1091.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*, *10*(3), 211–223.

Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., & Nitschke, J. B. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cerebral Cortex, 20*(4), 929–940.

Schienle, A., Schäfer, A., Hermann, A., Rohrmann, S., & Vaitl, D. (2007). Symptom provocation and reduction in patients suffering from spider phobia. *European Archives of Psychiatry and Clinical Neuroscience, 257*(8), 486–493.

Schienle, A., Schäfer, A., Walter, B., Stark, R., & Vaitl, D. (2005). Brain activation of spider phobics towards disorder-relevant, generally disgust-and fear-inducing pictures. *Neuroscience Letters, 388*(1), 1–6.

Schmidt, N. B., Richey, J. A., Buckner, J. D., & Timpano, K. R. (2009). Attention training for generalized social anxiety disorder. *Journal of Abnormal Psychology, 118*(1), 5.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology, 37*(2), 257–261.

Seligman, M. E. (1971). Phobias and preparedness. *Behavior Therapy, 2*(3), 307–320.

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience, 12*(3), 154–167.

Shin, L. M., & Liberzon, I. (2010). The Neurocircuitry of Fear, Stress, and Anxiety Disorders. *Neuropsychopharmacology, 35*(1), 169–191.

Sigmundi, R. A., Bouton, M. E., & Bolles, R. C. (1980). Conditioned freezing in the rat as a function of shock intensity and CS modality. *Bulletin of the Psychonomic Society*, *15*(4), 254–256.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*(1), 131.

Smith, D. V., Davis, B., Niu, K., Healy, E. W., Bonilha, L., Fridriksson, J., … Rorden, C. (2010). Spatial attention evokes similar activation patterns for visual and auditory stimuli. *Journal of Cognitive Neuroscience*, *22*(2), 347–361.

Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970) Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press.

Stark, R., Zimmermann, M., Kagerer, S., Schienle, A., Walter, B., Weygandt, M., & Vaitl, D. (2007). Hemodynamic brain correlates of disgust and fear ratings. *Neuroimage, 37*(2), 663–673.

Stevens, J., Wilson, K., & Foote, W. (1974). GABA blockade, dopamine and schizophrenia: experimental studies in the cat. *Psychopharmacologia*, *39*(2), 105–119.

Straube, T., Glauer, M., Dilger, S., Mentzel, H.-J., & Miltner, W. H. (2006). Effects of cognitive-behavioral therapy on brain activation in specific phobia. *Neuroimage, 29*(1), 125–135.

Straube, T., Mentzel, H. J., Glauer, M., & Miltner, W. H. (2004). Brain activation to phobia-related words in phobic subjects. *Neuroscience Letters, 372*(3), 204–208.

Straube, T., Mentzel, H.-J., & Miltner, W. H. (2006). Neural mechanisms of automatic and direct processing of phobogenic stimuli in specific phobia. *Biological Psychiatry, 59*(2), 162–170.

Straube, T., Mentzel, H.-J., & Miltner, W. H. (2007). Waiting for spiders: brain activation during anticipatory anxiety in spider phobics. *Neuroimage, 37*(4), 1427–1436.

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience, 11*(7), 843–850.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187–1188.

Szymanski, J., & O'Donohue, W. (1995) Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 26*, 31–34.

Tomarken, A. J., Mineka, S., & Cook, M. (1989). Fear-relevant selective associations and covariation bias. *Journal of Abnormal Psychology, 98*(4), 381-394.

Tomarken, A. J., Sutton, S. K., & Mineka, S. (1995). Fear-relevant illusory correlations: What types of associations promote judgmental bias? *Journal of Abnormal Psychology, 104*(2), 312-326.

Tomkins, S. S., & Mc Carter, R. (1964). What and where are the primary affects? Some evidence for a theory. *Perceptual and Motor Skills, 18*(1), 119–158.

Touryan, S. R., Marian, D. E., & Shimamura, A. P. (2007). Effect of negative emotional pictures on associative memory for peripheral information. *Memory, 15*(2), 154–166.

Tracey, I., & Mantyh, P. W. (2007). The cerebral signature for pain perception and its modulation. *Neuron, 55*(3), 377–391.

Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., & Fink, G. R. (2007). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cerebral Cortex, 17*(10), 2235–2244.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207–232.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology, 4*(4), 400.

Van Overveld, M., de Jong, P. J., & Peters, M. L. (2006). Differential UCS expectancy bias in spider fearful individuals: Evidence toward an association between spiders and disgust-relevant outcomes. *Journal of Behavior Therapy and Experimental Psychiatry, 37*(1), 60–72.

Van Overveld, M., de Jong, P. J., & Peters, M. L. (2010). Disgust and fear-related UCS-expectancy bias in blood-fearful individuals. *Clinical Psychology & Psychotherapy, 17*(2), 100–109.

Van Overveld, M., de Jong, P. J., Huijding, J., & Peters, M. L. (2010). Contamination and harm relevant UCS-expectancy bias in spider phobic individuals: influence of treatment. *Clinical Psychology & Psychotherapy, 17*(6), 510–518.

vanOyen Witvliet, C., & Vrana, S. R. (2000). Emotional imagery, the visual startle, and covariation bias: An affective matching account. *Biological Psychology, 52*(3), 187–204.

Villemure, C., & Bushnell, M. C. (2009). Mood influences supraspinal pain processing separately from attention. *The Journal of Neuroscience, 29*(3), 705–715.

Villemure, C., Slotnick, B. M., & Bushnell, M. C. (2003). Effects of odors on pain perception: deciphering the roles of emotion and attention. *Pain, 106*(1), 101–108.

Vuilleumier, P., & Huang, Y.-M. (2009). Emotional attention uncovering the mechanisms of affective biases in perception. *Current Directions in Psychological Science*, *18*(3), 148–152.

Vogt, A. (2012). *Aufmerksamkeitsallokation bei Supertastern: Schreckreflexmodulation (PPI) und Pupillenreaktion während der Präsentation emotionaler Bilder.* University of Würzburg.

Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience, 6*(7), 533–544.

Vogt, B. A., & Pandya, D. N. (1987). Cingulate cortex of the rhesus monkey: II. Cortical afferents. *Journal of Comparative Neurology, 262*(2), 271–289.

Vogt, B. A., Derbyshire, S., & Jones, A. K. (1996). Pain processing in four regions of human cingulate cortex localized with co-registered PET and MR imaging. *European Journal of Neuroscience, 8*(7), 1461–1473.

Vogt, B. A., Nimchinsky, E. A., Vogt, L. J., & Hof, P. R. (1995). Human cingulate cortex: surface features, flat maps, and cytoarchitecture. *Journal of Comparative Neurology, 359*(3), 490–506.

Walter, W., Cooper, R., Aldridge, V., McCallum, W., & Winter, A. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *Nature, 203*, 380–384.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063.

Watts, F. N., & Sharrock, R. (1984) Questionnaire dimensions of spider phobia. *Behaviour Research and Therapy, 22*, 575–580.

Wheeler, W. M. (1949). An analysis of Rorschach indices of male homosexuality. *Rorschach Research Exchange and Journal of Projective Techniques, 13*(2), 97–126.

Wiech, K., & Tracey, I. (2009). The influence of negative emotions on pain: behavioral effects and neural mechanisms. *Neuroimage, 47*(3), 987–994.

Wiedemann, G., Pauli, P., & Dengler, W. (2001). A priori expectancy bias in patients with panic disorder. *Journal of Anxiety Disorders, 15*(5), 401–412.

Wiemer, J., Gerdes, A. B., & Pauli, P. (2013). The effects of an unexpected spider stimulus on skin conductance responses and eye movements: an inattentional blindness study. *Psychological Research, 77*(2), 155–166.

Wiemer, J., Mühlberger, A., & Pauli, P. (2014). Illusory correlations between neutral and aversive stimuli can be induced by outcome aversiveness. *Cognition & Emotion, 28*(2), 193–207.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81.

Wittchen, H.-U., Zaudig, M. & Fydrich, T. (1997). *Strukturiertes Klinisches Interview für DSM-IV.* Göttingen: Hogrefe.

Wolter, J., & Lachnit, H. (1993). Are anticipatory first and second interval skin conductance responses indicators of predicted aversiveness? *Integrative Physiological and Behavioral Science*, *28*(2), 163–166.

Wood, J. M., Lilienfeld, S. O., Nezworski, M. T., Garb, H. N., Allen, K. H., & Wildermuth, J. L. (2010). Validity of Rorschach Inkblot scores for discriminating psychopaths from nonpsychopaths in forensic populations: A meta-analysis. *Psychological Assessment, 22*(2), 336-349.

Wright, C. I., Fischer, H., Whalen, P. J., McInerney, S. C., Shin, L. M., & Rauch, S. L. (2001). Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. *Neuroreport, 12*(2), 379–383.

9. Appendix

9.1. Table of fear-relevant illusory correlation literature

The following table summarizes studies dedicated to fear-relevant or emotionally relevant illusory correlations. The studies are listed in chronological order from the first study on fear-relevant selective associations in 1989 (Tomarken et al., 1989) to the most recent works in 2010. Studies were included in this list if they reported the expectancy and/or memory of an association between a feared or emotionally relevant stimulus (i.e. cues) and aversive consequences (i.e. outcomes). If participants were confronted with actual cue-outcome-associations, reinforcement rates are reported. Moreover, sample sizes and gender proportions of groups are recorded if possible. Abbreviations in alphabetical order: CB = covariation bias; EB = expectancy bias; ES = cutaneous electrical stimulus; DEEO = damaged and exposed electric outlet; GSP = generalized social phobia; M = number of men; N = non-anxious controls; PTSD = post-traumatic stress disorder; SCR = skin conductance response; SD = standard deviation; UCR = unconditioned response; US = unconditioned stimulus; W = number of women.

Study		Cues	Outcomes	Sample	M	W	Main results
Tomarken, Mineka, & Cook (1989)	exp. 1	snakes, spiders, mushrooms, flowers	ES (not painful); tone (33%)	high	0	23	CB in high fear
				low fear	0	22	
	exp. 2	like exp. 1	ES (not painful); tone + lightbulb (33%)	high	0	13	CB in high fear (for ES but not tone + lightbulb)
				low fear	0	12	
	exp. 3	like exp. 1	ES (not painful); tone (33% and 50%)	high + low fear	0	106	CB in high fear (33%); CB in high and low fear (50%)
de Jong & Merckelbach (1991)		spiders, mushrooms, flowers	ES (not painful); tone (33%)	healthy	0	18	CB in all groups; enhanced UCRs in spider trials
				treated	0	19	
				untreated phobics	0	20	
de Jong, Merckelbach, Arntz, & Nijman (1992)		spiders, weapons, mushrooms	ES (not painful), tone, nothing (33%)	treated	0	20	CB for spiders in untreated but not in treated phobics
				untreated spider phobics	0	18	
de Jong (1993)		flowers, spiders, weapons	harmless shock, siren, nothing (thought experiment)	high and low spider fear (median split)	33	31	EB for spiders in both low and high fear, equal confidence in estimates
de Jong & Merckelbach (1993)		spiders, weapons, flowers	ES, siren, nothing (33%)	treated	0	20	CB and on-line CB for spiders in untreated phobics; CB for weapons in treated phobics; enhanced SCRs to spiders and ES after spiders in untreated phobics
				untreated spider phobics	0	19	
Honeybourne, Matchett, & Davey (1993)		snakes + spiders, handgun + electricity outlet, landscape + flower (between subjects)	loud noise + vibration (instructed but not presented), nothing	healthy adults	18	18	equal EB for snakes, spiders, handgun and electricity outlet relative to landscape and flower; similar pattern in SCRs

Study		Cues	Outcomes	Sample	M	W	Main results
McNally & Heatherton (1993)	exp. 1	snakes, mushrooms, flowers	harmless shock, tone (thought experiment)	high fear	5	19	EB in high fear;
				low fear of snakes	8	8	smaller EB in low fear
	exp. 2	damaged and exposed electric outlets (DEEOs); mushrooms, flowers	harmless shock, tone (thought experiment)	high fear	11	14	EB in high and low fear
				low fear of DEEOs	9	11	
de Jong, Merckelbach, & Arntz (1995)		spiders, weapons, flowers	ES, siren, nothing (33%)	treated phobics	0	19	CB in untreated but not in treated group; higher SCRs to ES after spiders (UCR); correlation between UCR, CB and on-line expectancy
				untreated phobics	0	19	
de Jong, Merckelbach, & Nijman (1995)	exp. 2	flowers, spiders, weapons	ES (not painful), tone, nothing (33%)	healthy (continuous fear level)	0	19	correlation between CB and fear of spiders; correlation between imagery ability and SD of CB
de Jong, van den Hout, & Merckelbach (1995)		spiders, weapons, flowers	ES, siren (33%)	treated phobics	0	19	positive correlation between CB and relapse
Tomarken, Sutton, & Mineka (1995)	exp. 1	snakes, mushrooms, flowers	ES (not painful), tone (33%), nothing	high fear of snakes	0	16	CB for snakes only in high fear of snakes; no CB for DEEOs
				low fear of snakes	0	16	
		DEEOs, mushrooms, flowers	ES (not painful), tone (33%), nothing	high fear of snakes	0	16	
				low fear of snakes	0	16	
	exp. 2	DEEOs, snakes, flowers	ES (not painful), tone	high fear of snakes	0	24	Belongingness was highest for DEEOs and shocks, second for snakes and shocks; snakes and shocks had higher affective similarity than DEEOs and shocks in high fear group
				low fear of snakes	0	19	

Study		Cues	Outcomes	Sample	M	W	Main results
Davey & Dixon (1996)	exp. 1a	gun, chainsaw, electric plug, gasfire (between-subjects)	ES (thought experiment)	healthy adults	34	46	US expectancy predicted by prior fear, cue dangerousness and cue-outcome similarity (valence, fear, overall)
	exp. 1b	tiger, snake, spider, fire (between-subjects)	ES (thought experiment)	healthy adults	40	40	US expectancy predicted by prior fear, cue dangerousness and cue-outcome similarity (arousal, fear, overall)
	exp. 2	snakes + spiders, gun + electricity outlet (between subjects), landscape + flower	loud noise + vibration, light flash, nothing (33%)	healthy adults, phylogenetic cues	9	11	ontogenetic more dangerous than phylogenetic; EB and CB for phylogenetic-noise and ontogenetic-noise relative to fear-irrelevant-noise; similar correlations with expectancy as in exp. 1
				ontogenetic cues	9	11	
	exp. 3	snake or spider, flower, landscape	loud noise + vibration, light flash, nothing (33%)	high and low fear of snakes or spiders	12	28	EB in high and low fear but stronger EB in high fear; CB only in high fear
Diamond, Matchett, & Davey (1995)		spiders, kittens (between subjects)	loud noise + vibration (instructed but not presented), nothing	high and low fear of spiders	16	20	on-line EB and enhanced SCRs for spiders in high (but not low) fear
Pauli, Montoya, & Martz (1996)		emergencies, mushrooms, nudes	ES (painful) 1st block: 50:50:50% 2nd block: 83:17:17% 3rd block: 50:50:50%	panic-prone	2	8	1st block: CB for emergencies in panic prone individuals; 2nd block: no group differences; 3rd block: CB in both groups
				not panic-prone	4	6	
Amin & Lovibond (1997)		snakes, spiders, knifes, guns, flowers, mushrooms	ES (not painful; 50%), tone (25%), nothing	healthy adults (continuous fear level)	14	26	Equal on-line EB and SCRs for biological and technological threat stimuli; CB only for biological threat stimuli in high fear

Study		Cues	Outcomes	Sample	M	W	Main results
Davey & Craigie (1997)		snakes + spiders, gun + electricity outlet (between subjects, video-based dangerousness manipulation), landscape + flower	loud noise + vibration, light flash, nothing (33%)	healthy adults	16	24	EB for phylogenetic and ontogenetic; EB and CB enhanced by dangerousness manipulation
Kennedy, Rapee, & Mazursky (1997)		snakes/spiders, DEEOs, flowers	ES (not painful), tone, nothing (33%)	high fear of snakes/spiders and DEEOs	0	22	EB for snakes/spiders and DEEOs in high and low fear; CB only for snakes/spiders in high fear
				low fear	0	22	
Pury & Mineka (1997)	exp. 1A	surgeries, computers, engines	ES (between uncomfortable and painful), tone, nothing (33%)	high	0	18	CB for both high and low fear
				low blood-injury fear	0	22	
	exp. 1B	mutilations, babies, semi-nudes	ES (between uncomfortable and painful), tone, nothing (33%)	high	0	15	CB for both high and low fear
				low blood-injury fear	0	25	
	exp. 2	minor injuries, flowers, rabbits	ES (between uncomfortable and painful), tone, nothing (33%)	high	0	17	CB for both high and low fear
				low blood-injury fear	0	12	
de Jong, Merckelbach, Bögels, & Kindt (1998)		male or female (between-subjects) angry, happy, neutral faces	ES (not painful), siren, nothing (33%)	low	0	32	EB for angry-shock (vs. neutral- and happy shock) in high anxiety; EB for angry-shock (vs. happy-shock) in low anxiety; on-line EB and CB in high and low anxiety; no impact of cue sex and social anxiety
				high socially anxious	0	28	
Pauli, Wiedemann, & Montoya (1998)		crashed airplanes, flying airplanes, mushrooms	ES (painful, 50%), nothing	high fear of flying	2	12	CB for crashed airplanes in high fear after 1st block
				low fear	4	10	

Study		Cues	Outcomes	Sample	M	W	Main results
Cavanagh & Davey (2000)	exp. 1	spider, height view, hypodermic needle, landscape, flower, rabbit	ES (thought experiment)	high	6	27	higher expectancy for spider-shock and lower expectancy for fear-irrelevant-shock in high fear than low fear; correlations between fear and expectancy
				low fear of spiders	8	25	
	exp. 2	landscape, flower, rabbit, apple, telephone, thermometer	ES (thought experiment)	high	2	27	still lower expectancy for fear-irrelevant-shock in high fear without spider as contrast
				low fear of spiders	6	16	
de Jong & Merckelbach (2000)		spiders, weapons, flowers	ES (not painful); siren, nothing (33%)	phobics + alprazolam	0	21	lower SCR to images in alprazolam (≈ low fear); on-line bias in both phobic groups; CB in all groups
				phobics + placebo	0	22	
				low fear	0	24	
VanOyen Witvliet & Vrana (2000)		positive + negative imagery; visual startle	visual startle (67%)	healthy	24	22	enhanced CEs for arousing positive and negative stimuli
Cavanagh & Davey (2001)	exp. 1	chainsaw, gun, splintered glass, spider, snake, fire, flower, kitten, rabbit	ES (thought experiment)	healthy adults; mood induction (positive, negative, neutral; between-subjects)	24	45	EB for phylogenetic and ontogenetic; ontogenetic > phylogenetic; overall inflated US expectancy in positive and negative mood
	exp. 2	chainsaw, gun, splintered glass, spider, snake, fire, flower, kitten, rabbit	ES (thought experiment)	healthy adults; arousal manipulation (low, high, relax; between-subjects)	20	29	EB for phylogenetic and ontogenetic; no effect of arousal

Study	Cues	Outcomes	Sample	M	W	Main results
Pauli, Montoya, & Martz (2001)	emergencies, mushrooms, nudes	ES (painful) 1st block: 50:50:50% 2nd block: 17:83:17% 3rd block: 50:50:50%	panic-prone	0	9	1st block: CB and on-line EB for emergencies in panic-prone participants; 2nd block: heightened but unbiased CB and on-line EB for mushrooms in both groups; 3rd block: overall unbiased contingency estimates in both groups
			not panic-prone	1	8	
Pauli, Wiedemann, Dengler, & Kühlkamp (2001)	emergency situations, spiders, mushrooms, erotic scenes	harmless ES, neutral tone, nothing (thought experiment)	patients with ICD; discharge experience	11	1	EB for emergency-shock only with discharge experience; correlation between EB and trait anxiety
			no discharge experience	9	3	
Wiedemann, Pauli, & Dengler (2001)	emergency situations, spiders, mushrooms, erotic scenes	ES, tone, nothing (thought experiment)	panic disorder	8	21	EB for emergencies in panic disorder; smaller EB in healthy controls
			healthy controls	8	21	
Pauli, Diedrich & Müller (2002)	pleasant, unpleasant, neutral pictures	startle sound during picture presentation (95 dB; 63%), nothing	healthy	16	14	CB for unpleasant and pleasant relative to neutral pictures; unpleasant CB > pleasant CB; no correlation between CB and startle reflex
Davey, Cavanagh, & Lamb (2003)	spider, cockroach, maggot, slug (non-predatory); tiger, wolf, shark, snake (predatory); kitten, rabbit, sheep, chicken (safe)	ES (harm); nauseating juice (disgust); (thought experiment)	healthy adults	35	56	EB for non-predatory-juice and predatory-shock; no effect of gender
Hermann, Ofer, & Flor (2004)	descriptions of aversive animals, pleasant nature scenes, ambiguous social situations	angry, happy, neutral facial expression (33%)	generalized social phobia (GSP)	9	8	negative social EB in GSP, positive social EB in NAC; negative social CB in GSP; reduced (enhanced) on-line expectancy of social-negative outcomes in NAC (GSP); overall enhanced SCRs to images in GSP; in first block enhanced SCRs to negative outcomes in GSP
			non-anxious controls (NAC)	8	10	

Study		Cues	Outcomes	Sample	M	W	Main results
Amrhein, Pauli, Dengler, & Wiedemann (2005)		emergencies, spiders, mushrooms	startle sound (103 dB, 50%), nothing	panic disorder	10	20	EB for emergencies in panic disorder and healthy controls, enhanced late contingent negative variation (CNV) for emergencies in panic patients, no CB
				healthy controls	9	16	
Muris, de Jong & van Lubeck (2005)	exp. 1	spiders, weapons, Pokémons	winning, losing candy, nothing (card game; 33%)	youths (9-13 years)	70	77	CB for spider-lose and spider-win relative to spider-nothing; no correlations with fear of spiders or neuroticism within both genders
	exp. 2	spiders, weapons, Pokémons	winning, losing candy (to a "real" spider), nothing (card game; 33%)	youths (8-12 years) one half instructed to pay attention to contingencies	112	128	not instructed: in general lose/win > nothing; for Pokémons: win > lose; no correlations between CB and spider fear; instructed: in general: lose/win > nothing; for Pokémons and weapons but spiders: win > lose; no correlations with spider fear
van Overveld, de Jong, & Peters (2006)		spiders, bull terriers, maggots, rabbits	ES, disgusting taste , nothing	high	0	28	EB in high fear (for ES and taste) attenuated EB in low fear
Garner, Mogg, & Bradley (2006)		happy, angry, neutral facial expression	pleasant, unpleasant, neutral pictures	high	2	21	EB for happy-pleasant and angry-unpleasant in both groups; lasting on-line EB for happy-pleasant in low but not high social anxiety; no CB
				low socially anxious (median split)	3	20	
Mühlberger, Wiedemann, Hermann, & Pauli (2006)		spiders, flight accidents, mushrooms	startle sound (103 dB, 50%), nothing	spider phobics	1	16	EB for spiders in spider phobia, EB for flight accidents in flight phobia, CB only for spider phobia after 1st block, enhanced startle and EEG responses to spiders in spider phobia
				flight phobics	3	14	

Study		Cues	Outcomes	Sample	M	W	Main results
de Jong & Peters (2007a)		blood-donation, flowers, rabbits	self-administered ES (harm), bad tasting fluid (disgust), nothing (33%)	high	0	25	EB for harm and disgust outcome and blood-donation; on-line EB only in first block; no influence of fear and no CB
				low fear of blood-injection-injury	0	27	
de Jong & Peters (2007b)		spiders, pit bulls, maggots, rabbits	ES; disgusting taste, nothing (33%)	high + low fear	0	49	EB in high fear (for ES and taste) EB in low fear (for ES) no CB
Muris, Huijding, Mayer, den Breejen, & Makkelie (2007)	exp. 1	spiders, guns, flowers	ES, nothing (thought experiment)	youths (9-16 years)	76	74	EB for spiders and guns; stronger EB in older youths and girls; correlation between spider fear and EB for spiders
	exp. 2	spiders, guns, flowers	winning, losing candy, nothing (computer game; 33%)	youths (8-14 years)	102	118	CB for spiders-lose, guns-lose, flowers-win; correlation between spider fear and CB in older group
Olatunji, Cisler, Meunier, Connolly, & Lohr (2008)		spiders (fear), rotting foods and body products (disgust), appliances and tools (neutral)	fear, disgust, neutral facial expressions (thought experiment)	high	1	21	Equal expectancy for spider-fear, but higher expectancy for spider-disgust and lower expectancy for spider-neutral in high fear
				low fear of spiders	0	28	
Connolly, Lohr, Olatunji, Hahn, & Williams (2009)		vomit, feces (contamination fear); vicious dog, man with knife (general fear); flowers, chair (neutral)	fear, disgust, neutral facial expressions (33%)	high	8	24	CB for contamination-fear and contamination-disgust in high fear; CB for contamination-disgust in low fear
				low contamination fear	15	15	
Engelhard, de Jong, van den Hout, & van Overveld (2009)		pictures related and unrelated to troop deployment	startle sound (95 dB, once before IC paradigm and once following a deployment unrelated picture), nothing	soldiers 2-5 months back from troop deployment	171 (predomi-nantly male)		US expectancy for deployment related pictures predicted PTSD symptoms 15 months after deployment

Study	Cues	Outcomes	Sample	M	W	Main results
van Overveld, de Jong, Huijding, & Peters (2010)	spiders, dogs, maggots, rabbits	ES (harm), nauseating juice (disgust), nothing (thought experiment)	low high fear of spiders (before/ after treatment)	6 11	24 49	EB for rabbit-nothing, maggots-juice, dog-shock in high and low fear; spider-shock and spider-juice in high fear, reduction of these EBs after treatment
van Overveld, de Jong, & Peters (2010)	bloody wound, gun, maggots, growling dog, rabbit	ES (harm), nauseating juice (disgust), nothing (thought experiment)	high and low blood-fearful	8	52	comparable EB for wound and harm and disgust outcomes; inflated EB in high fear; other EBs: rabbit-nothing; maggots-juice, dog-shock, gun-shock

9.2. Supplemental data tables

Table 9.1. Activations at the contrast [(*shock-after-spider* > *nothing-after-spider*) > (*shock-after-mushroom* > *nothing-after-mushroom*)] correlating with the covariation bias

	Region		MNI coordinates			k	*t*
			x	y	z		
Spider phobics	***whole-brain***						
	Precentral (BA 6)	L	-28	-20	68	28	5.31
	PCL (BA 6)	L	0	-18	68	47	4.66
	Precentral (BA 4)	R	16	-26	68	10	4.33
	Cerebellum (BA 37)	R	46	-64	-24	14	4.13
	Fusiform (BA 19)	R	40	-66	-20	10	4.01
Control group	***whole-brain***						
	Inferior parietal (BA 40)	L	-52	-48	34	75	6.65
	Mid cingulum (BA 31)	L	-18	-32	42	30	4.75
	Pallidum	L	-20	-8	2	14	4.37
	Cerebellum	R	12	-54	-16	18	4.27

The table shows properties of peak voxels within a cluster. whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; BA = Brodmann area; PCL = Paracentral lobule; k = voxels in whole cluster

Table 9.2. Increased connectivity (psychophysiological interaction, PPI) to right PCL in spider phobic individuals for the contrast [*shock-after-spider* > *shock-after-mushroom*]

	Region		MNI coordinates			k	*t*
			x	y	z		
	whole-brain						
	Midbrain	R	14	-18	-12	223	8.00
	Cerebellum	R	14	-62	-46	258	6.54
	Cerebellum	L	-20	-66	-16	320	6.00
	Supramarginal (BA 40, SII)	R	48	-22	24	214	5.92
	Supramarginal (BA 40, SII)	L	-62	-48	18	240	5.01
	ACC (BA 24/32)	L	-4	30	22	305	5.77
	Frontal inferior orbital (BA 47)	R	48	28	-4	251	5.65
	Mid cingulum (BA 24)	L	-12	16	34	190	5.51
Spider phobics	Mid cingulum (BA 24)	R	8	-8	38	70	4.65
	Insula (BA 47)	L	-42	20	-4	112	5.31
	Insula (BA 47/13)	R	30	12	-14	33	4.99
	Lingual (BA 18)	R	22	-90	-12	188	5.20
	SMA (BA 6)	R	14	-24	56	208	5.12
	Rolandic operculum (BA 22)	R	56	0	10	126	5.11
	Calcarine (BA 17/18)	L	2	-86	8	95	5.00
	Middle temporal (BA 21/22)	L	-54	-44	-2	130	4.96
	Superior temporal	R	40	-22	-6	59	4.67
	Frontal inf. operculum (BA 44/45)	R	54	16	16	157	4.67
	Precentral (BA 8)	R	50	8	40	52	4.53
	Middle frontal (BA 9)	L	-40	30	38	10	4.07

The table shows properties of peak voxels within a cluster. whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; BA = Brodmann area; PCL = Paracentral lobule; SII = Secondary somatosensory cortex; ACC = Anterior cingulated cortex; SMA = Supplementary motor area; dlPFC = dorsolateral prefrontal cortex k = voxels in whole cluster

Table 9.3. Significant Activations at the contrast [*spider* > *mushroom*]

	Region		MNI coordinates			k	*t*
			x	y	z		
Spider phobics	***whole-brain***						
	Middle temporal (BA 37/19)	R	46	-66	-12	4237	11.23
	Middle occipital (BA 19/37)	L	-18	-74	38	6813	10.51
	Frontal inf. triang. (BA 45/46)	R	46	38	6	367	9.23
	Thalamus	R+L	0	-16	4	1037	8.22
	SMA (BA 6/24)	R+L	16	12	64	1240	7.76
	Middle frontal (BA 8/9)	L	-16	46	42	521	7.06
	Middle frontal (BA 9/10)	R	28	50	28	756	6.93
	Insula (BA 13/47)	L	-44	12	2	599	6.81
	Frontal inf. operculum (BA 9)	R	48	8	22	296	5.99
	Amygdala	L	-18	-6	-8	102	5.92
Control group	***whole-brain***						
	Frontal sup. medial (BA 9)	L+R	-4	46	24	430	8.45
	Middle occipital (BA 18/19)	L	-28	-78	-10	1475	8.43
	Middle temporal (BA 37/19)	R	18	-98	0	1345	8.13
	Fusiform (BA 37)	L	-42	-66	-22	195	5.86
	Fusiform	R	40	-46	-14	29	4.99
	Frontal superior (BA 6)	L	-12	26	64	15	5.80
	Superior temporal (BA 13)	R	50	-48	18	49	5.02
	Middle frontal (BA 46)	R	48	34	20	16	4.93
	Precentral (BA 9)	R	42	2	38	12	4.44
	Precentral (BA 6)	L	-34	-12	54	17	4.43
Phobics > Controls	***whole-brain***						
	SMA (BA 6/24)	R	16	10	66	1215	5.94
	Insula (BA 13)	L	-42	12	2	410	5.90
	Middle occipital (BA 7/37)	L	-46	-74	-10	2525	5.77
	Middle occipital (BA 19)	R	30	-76	34	143	5.36
	Inferior frontal (BA 44)	R	50	6	20	76	5.38
	Caudate	R	14	16	8	234	4.89
	Thalamus	L+R	-2	-16	2	61	4.89
	Middle frontal (BA 9)	L	-32	28	44	161	4.89
	Middle temporal (BA 37/39)	R	36	-56	-22	701	4.87
	Superior frontal (BA 8/9)	L	-18	44	42	42	4.71

The table shows properties of peak voxels within a cluster. whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; For display purposes only the 10 most significant clusters within the whole-brain analysis are shown here; BA = Brodmann area; SMA = Supplementary motor area; k = voxels in whole cluster

Table 9.4. Correlation coefficients between beta values of significantly activated ROIs at the contrast [*spider* > *mushroom*] and trial-by-trial covariation bias across spider phobic individuals (*N* = 17)

	Region		**Canonical HRF**		**FIR**	
			r	***p***	***r***	***p***
	ROI					
Spider phobics	dlPFC	L	.56*	.02	.62*	.01
		R	.42	.09	.30	.24
	Amygdala	L	.34	.18	.40	.11
	ACC	L	.27	.40	.25	.34
	Insula	L	.08	.76	.11	.67

Canonical HRF: brain activity is beta from hemodynamic response function; FIR: brain activity is percent signal change during picture presentation (finite impulse response); dlPFC = dorsolateral prefrontal cortex; ACC = anterior cingulate cortex; *r* = Pearson correlation coefficient

Table 9.5. Activations at the contrast [*spider* > *mushroom*] correlating with the covariation bias

	Region		MNI coordinates			k	*t*
			x	y	z		
Spider phobics	***whole-brain***						
	Middle frontal (BA 8)	L	-26	28	46	120	6.99
	Precentral (BA 4)	L	-32	-28	54	24	5.55
	Postcentral (BA 3)	L	-22	-36	54	12	5.22
	Superior frontal (BA 8)	L	-10	48	44	20	4.82
	Lingual (BA 30)	R	0	-62	-2	41	4.69
	SMA (BA 6)	R+L	2	-8	56	10	4.37
	Middle frontal (BA 9)	L	-40	34	34	11	4.25
	Superior parietal (BA 7)	L	-38	-60	50	10	4.09
	Superior temporal (BA 12)	R	56	-44	12	12	4.04
Control group	***whole-brain***						
	Middle occipital	L	-34	-78	6	63	6.44
	Precentral (BA 6)	R	34	-4	44	42	6.38
	Hippocampus	L	-32	-32	-10	35	5.69
	Cerebellum	L	-6	-36	-22	41	4.85
	Superior frontal (BA 8)	L	-10	32	50	35	4.83
	Lingual (BA 30)	R	8	-62	4	11	4.66
	Lingual (BA 19)	L	-10	-56	-8	44	4.58
	Superior parietal (BA 7)	L	-22	-70	46	10	4.31
	Cuneus (BA 7)	L	-16	-82	36	10	4.08

The table shows properties of peak voxels within a cluster. whole-brain threshold: $p < .001$ (uncorrected), $k \geq 10$; BA = Brodmann area; k = voxels in whole cluster

9.3. Picture references

The following pictures were used for illustration purposes. They were found as either in the public domain or as licensed for free distribution under the Creative Commons Attribution-Share Alike license [12] or Creative Commons Attribution-Non-Commercial-Share Alike license [13].

Description	Page	Link
Fearful man	19	http://commons.wikimedia.org/wiki/File:Expression_of_the_Emotions_Figure_20.png
Injury	37	http://commons.wikimedia.org/wiki/File:BrokenNose.jpg
Umbrella	37	http://commons.wikimedia.org/wiki/File:An_inside_view_of_an_Umbrella,_%E0%B4%95%E0%B5%81%E0%B4%9F%E0%B4%AF%E0%B5%81%E0%B4%9F%E0%B5%86_%E0%B4%89%E0%B5%BE%E0%B4%B5%E0%B4%B6%E0%B4%82.JPG
Erotic couple	37	https://www.flickr.com/photos/kk/3130311890/sizes/z/in/photolist-5LBEEE-asBRMK-dDFEsB-dDM3fE-aAQvut-cUb5L1-9bLimu-59fgBL-axGyVk-avnLCp-aejGuT-db7F4j-7k9ZhR-iDQm-58jXcG-3RD1z-bBF7n9-7WwFMx-9YcVXc-8TNyx-5wje8K-atakKZ-5yxVcb-7xecAN-52xnXQ-946wCN-69AWGq-7Dp47b-946wQm-52td1k-5DKVmT-489nMA-aajdaS-ataj4t-8KUaiK-38TjMu-85vh2h-gSMMv-dYhBUk-dNzLXG-dXMEVD-8V4mCr-dMEq9t-8qki95-59b4qR-aaB6eM-bmaGAi-23PfXJ-asVRBv-487kZY/
Lamp	37	http://commons.wikimedia.org/wiki/File:Nightstandlamp.jpg
Wind surfer	52	http://commons.wikimedia.org/wiki/File:Wales_Windsurfer.jpg
Military dog	52	http://commons.wikimedia.org/wiki/File:Military_dog_barking.JPG
Cockroach	56	http://commons.wikimedia.org/wiki/File:Cockroachcloseup.jpg
Green mamba	56	http://commons.wikimedia.org/wiki/File:Green_Mamba.jpg
Neo nazi	56	http://commons.wikimedia.org/wiki/File:Neo-nazi.jpg
Chocolate cake	56	http://commons.wikimedia.org/wiki/File:Chocolate_Cake_Flourless_(1).jpg
Tarantula	71	https://www.flickr.com/photos/zanthia/4330025499/sizes/m/in/photolist-5y53uV-6zKMRb-4rhdnw-4u6mb-5pFhcU-5AyueF-4JFuqP-6zTbbG-8vGKMH-8vKNRq-7ACvJ4-7ACymH-7AGjdj-7ASs1v-7AGg1Y-2wysgm-4Kn7bc-5GM2Wc-4ZJ2iY-5NvxnA-6GbTPW-4NLBbZ-6zTaKW-6kMdjy-6kH6yH-5FsADR-5FswmX-5FwRa9-5TMjqw-5HJ2MR-aHGmTx-aHGjK8-8vrLvv-aHGkxx-aHGkcc-7YywCu-aHGk3a-aHGmLB-6G7NZr-6jh6bC-6dVNj9-5CWx1U-oP4BA-ncuJh-5TMi9E-aHGkQe-5bJWbH-7u36ZV-aHGkFD-6hM5WU/
Mushroom	71	https://www.flickr.com/photos/valter/3247198425/sizes/m/in/photolist-5WWJXK-5a16uG-4PtuvT-5xAq7f-4xRjS8-oWjxx-73sdCv-41akR-6QpfuW-2TZfMJ-6NTgD7-42cNEB-5tKiw7-5YKwCQ-31Cj4-4PxDVm-7bERt7-3kqGXC-4swf8m-p4ZkA-rkQR-567gqf-73DuVS-pvB6V-7hVKPF-rojAa-4c91p1-6UMQJz-5katt-7gTif6-7aPUbN-BdRkE-84fQE2-e7vmf-98MVBz-8zgcRw-4JHhJ6-4GQ5yh-8o7mEf-6GTSTy-6nh4jV-7b3kCG-7FKq6o-7povoh-4ksRMc-ktERM-KFNWu-3ntMr9-7X2AYT-4UszGL/
Puppy	71	https://www.flickr.com/photos/ttstam/2936797299/sizes/m/in/photolist-5tvRwc-6Y1ZvP-6XeLr-7m7sXz-7mbv5w-7mboXY-ukUCT-5N222Y-4ua7av-fP7q-4qAdmZ-62uiBC-8zw1JU-826YYg-5hQhsT-2dZtue-4rkA8z-4PWR5F-ybP9S-4zFL9A-5RTB8e-6jPw81-FMLa-4xDwd-ni4mD-hhUW7-fA2EKd-3hs7sv-8zw1Y5-8zw1MY-8zsRoR-8zsR9X-8zsRwF-8zw1R7-eJEc5-5ZYfPW-5ZYcd5-5ZYwmq-7Z94E8-4Le9oq-6zmAxf-4ycMkG-4aFj3m-25ecy-aedAVg-76wNiY-4qAdie-6nYYuA-4scpBf-4gezo3/

[12] http://creativecommons.org/licenses/by-sa/3.0/de/

[13] http://creativecommons.org/licenses/by-nc-sa/3.0/de/

9.4. Curriculum Vitae

ZEITSCHRIFTEN-ARTIKEL

Wiemer, J., Schulz, S. M., Reicherts, P., Glotzbach-Schoon, E., Andreatta, M., & Pauli, P. (2015). Brain activity associated with illusory correlations in animal phobia. *Social Cognitive and Affective Neuroscience*, 10(7), 969–977.

Andreatta, M., Glotzbach-Schoon, E., Mühlberger, A., Schulz, S. M., **Wiemer, J.**, & Pauli, P. (2015). Initial and sustained brain responses to contextual conditioned anxiety in humans. *Cortex*, 63, 352-363.

Wiemer, J., Mühlberger, A., & Pauli, P. (2014). Illusory correlations between neutral and aversive stimuli can be induced by outcome aversiveness. *Cognition & Emotion, 28*, 193–207.

Herbert, C., Platte, P., **Wiemer, J.**, Macht, M., & Blumenthal, T. D. (2014). Supertaster, super reactive: Oral sensitivity for bitter taste modulates emotional approach and avoidance behavior in the affective startle paradigm. Physiology & behavior, 135, 198-207.

Wiemer, J., Gerdes, A. B. M., & Pauli, P. (2013). The effects of an unexpected spider stimulus on skin conductance responses and eye movements: an inattentional blindness study. *Psychological Research, 77,* 155-166.

AUSGEWÄHLTE TAGUNGSBEITRÄGE

Wiemer, J., & Pauli, P. (2015). Erhöhte Konnektivität zwischen senso-motorischem und visuellem Cortex sagt Covariation Bias bei Angst vorher. *40. Tagung Psychologie und Gehirn*, Frankfurt. (Poster)

Wiemer, J., & Pauli, P. (2014). Neuronale Korrelate einer illusorischen Korrelation zwischen phobierelevanten Hinweisreizen und aversiven Konsequenzen. *49. Kongress der Deutschen Gesellschaft für Psychologie*, Bochum. (Vortrag)

Wiemer, J., Floß, E., & Pauli, P. (2013). Subsequent memory effects on event-related potentials of encoding aversive stimuli under neutral and negative affect. *53rd Annual Meeting Society for Psychophysiological Research,* Florence. (Student Poster Award)

Wiemer, J., Schulz, S. M., Reicherts, P., Glotzbach-Schoon, E., Andreatta, M., Mühlberger, A., & Pauli, P. (2013). Illusory correlations between fear-relevant and aversive stimuli in spider phobia. *53rd Annual Meeting Society for Psychophysiological Research,* Florence. (Vortrag)

Wiemer, J., & Pauli, P. (2012). An expectancy bias for aversive events: the role of arousal and anxiety. *19th Annual Meeting Cognitive Neuroscience Society*, Chicago. (Poster)

Wiemer, J., Mühlberger, A., & Pauli, P. (2011). Erwartungsverzerrungen im affektmodulierten Startle-Paradigma. *37. Tagung Psychologie und Gehirn*, Heidelberg. (Poster)

9.5. Declaration of Originality

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst habe und keine anderen als die angegebenen Quellen benutzt und die aus fremden Quellen direkt oder indirekt übernommenen Gedanken als solche kenntlich gemacht habe. Die Arbeit habe ich bisher an keiner anderen Universität oder sonstigen wissenschaftlichen Einrichtung vorgelegt.

(I hereby declare that this dissertation is my own work and that all the sources that I have used or quoted have been acknowledged by means of complete references. This work has not been submitted previously for a degree at any university or other academic institution.)

_____

Julian Wiemer Würzburg, 23.07.2015