A mechanism-oriented approach to psychopathology: The role of Pavlovian conditioning

Frauke Nees, Angela Heinrich, Herta Flor *

Department of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

A R T I C L E   I N F O

Article history:
Received 22 December 2014
Received in revised form 1 May 2015
Accepted 6 May 2015
Available online 12 May 2015

Keywords:
Classical conditioning
Mental disorder
Research domain criteria
Brain imaging
Psychophysiology
Behavior
Learning
Extinction

A B S T R A C T

The Research Domain Criteria Project suggests to base the classification of mental disorders on dimensions of observable behavior and neurobiological measures of these functions rather than on symptom-based descriptive categorical diagnoses. We suggest a mechanismic approach that focuses on the role of learning as a core mechanism that can be studied in animals and humans. We review human studies on neurobiological, psychophysiological, and behavioral correlates of Pavlovian associative learning and delineate commonalities and differences across disorders. In addition to the hedonic value, the learning phase (i.e., habituation, acquisition, extinction, extinction recall), the role of stimulus properties (i.e., cue and context), and event timing (e.g., delay and trace conditioning) were considered. We address how core behavioral and psychophysiological indicators of conditioning, such as contingency ratings and skin conductance responses or startle modulation, respectively, are altered. We also discuss plastic changes in core brain regions and the interaction of brain regions in inhibitory and excitatory circuits. We also address the translation of findings pertaining to classical conditioning and its affiliated processes into the development of new behavioral and pharmacological treatments for mental disorders, and discuss productive avenues for future studies.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

According to the Research Domain Criteria (RDoC) initiative of the National Institute of Mental Health (Cuthbert and Insel, 2010, 2013), common mechanisms that contribute to differing forms of psychopathology and their associated symptomatology may provide the basis for a new classification framework for research on mental disorders. For example, in anxiety disorders, shared key features may not only range along the anxiety spectrum, but there may also be subgroups within anxiety disorders that share common mechanisms (Flor and Nees, 2014; McTeague et al., 2009, 2010, 2012). RDoC focuses on new ways of classifying psychopathology based on dimensions of observable behavior or biological measures instead of traditional categorical and symptom-oriented diagnostic criteria (Insel et al., 2010).

In this context, classical conditioning processes are among the best candidates to produce solid and reliable results that translate into advanced research and in consequence better treatments, because learning mechanisms represent a behavioral approach to psychopathology. A broad range of animal studies has provided a detailed and solid base of knowledge about psychobiological mechanisms that permits characterization also in humans (cf. Flor and Meyer-Lindenberg, 2014; Sehlmeyer et al., 2009).

2. Basic mechanisms of classical conditioning

Classical or Pavlovian conditioning is a learning mechanism that involves the acquisition and storage of emotionally significant information and the related change in behavior. An originally neutral stimulus (the conditioned stimulus, CS) is presented in conjunction with an aversive (for example, fear-eliciting) or appetitive event (the unconditioned stimulus, US). The CS acquires aversive or appetitive properties and elicits a response (conditioned response, CR) that is often but not always similar to the response that individuals exhibit to the US (see compensatory conditioning) (Braverman, 1979). Besides the acquisition of CS–US associations, classical conditioning also involves extinction processes, a learning mechanism that is characterized by a decrease of the CR when the CS that was previously paired with the US, is now repeatedly presented alone, without this US (e.g., Hermans et al., 2006; Myers and Davis, 2002).

Deficits in extinction might additionally account for the development or maintenance of mental disorders entailing, for example, an inability to inhibit acquired maladaptive fear responses (e.g., Bouton, 2004; Vervliet et al., 2013) or an inability to associate the conditioned stimulus with the extinction context (e.g., Maren et al., 2013). It should...
be noted that extinction does not reflect an “unlearning” of the CR, i.e., the CS–US association is not completely eradicated based on evidence that the CR can, for example, be spontaneously recovered, following the mere passage of time (Pavlov, 1927; Rescorla, 2004), due to a context change (renewal, e.g., Bouton and Ricker, 1994), or due to presentations of the US alone following extinction (reinstatement, e.g., Rescorla and Heth, 1975). Therefore, extinction along with behavior following extinction (for example, extinction memory) is context-dependent (Bouton, 2004). As indicated by the renewal effect, extinction leaves the CS especially sensitive to manipulations of context in that the presence of the extinction context retrieves or sets the occasion for a CS/no US association, and thus leaves the CS under a contextually modulated form of inhibition (see Bouton, 1993; Vervliet et al., 2013).

Moreover, for associative learning, prediction error signals are fundamental: Pavlovian conditioning depends on the discrepancy between an actual and expected outcome, based on CS–US associations. Such prediction processes characterize learning and play a role during both the acquisition and the extinction of conditioned responses, and are driven by changes in value proportional to the difference between the actual and predicted outcomes (Rescorla and Wagner, 1972). Additionally, in this context, associability plays an important role as it gates the amount of future learning based on whether a conditioned stimulus has been a reliable predictor of reinforcement in the past, and thus dynamically accelerates learning to cues whose predictions are poor and decelerates learning, when predictions become reliable (cf. Pearce and Hall, 1980). Both processes have been combined in hybrid learning models (e.g., LePelley and McLaren, 2004) that assume that prediction error drives learning but that learning is also dynamically based on the cue’s associability. Thus, abnormalities in the encoding of predictions and associability could result in dysfunctional learning such as a biased estimation of outcomes or altered attribution of salience to aversive or appetitive events. In addition, the timing of the CS and the US is important in classical conditioning. In delay conditioning, the CS is presented in contiguity with the US whereas in trace conditioning there is a temporal gap between the CS and US. Whereas delay conditioning does not depend on conscious awareness of the CS–US contingency, trace conditioning requires CS–US contingency awareness (e.g., Clark and Squire, 1998). Contingency awareness affects the proportion of learning, but autonomic fear conditioning can also occur without conscious awareness of the CS–US contingency, trace conditioning requires CS–US contingency awareness (e.g., Schultz and Helmstetter, 2010). Previous studies found that only aware individuals showed differential skin conductance responses during fear conditioning (Tabbert et al., 2006; Weike et al., 2007), whereas conditioned startle potentiation as well as enhanced brain responses in the amygdala, the orbitofrontal, and the occipital cortex, on the other hand, could also be observed in unaware participants (e.g., Tabbert et al., 2006; Weike et al., 2007).

The majority of research on classical conditioning has focused on cue conditioning; however, in situations where the US is presented without a cue, the context becomes associated with the US (Phillips and LeDoux, 1994). In terms of fear conditioning, conditioned cues will evoke phasic fear responses, while contexts will lead to sustained anxiety responses (Marks, 1987). Contexts not only involve environmental characteristics but also internal states, cognitive sets, or social and cultural settings (e.g., Maren, 2001; Maren et al., 2013; Rudy et al., 2002). Conditioning is not a steady state phenomenon, but a dynamic process varying as a function of the contingencies between the CS and the US (Grillon, 2002, 2008; Grillon et al., 2006). In particular, in contrast with cue conditioning in which the CR quickly subsides after the offset of a short-lasting CS, context conditioning is characterized by the absence of such a clear signal. Instead context conditioning is based on the presence of multisensory, diffuse, and continuously present internal or external environments that involve a possible occurrence of the US without signaling the exact time of its onset or its non-occurrence (i.e., safety periods). Hence, contexts involve more unpredictability than discrete cues (cf. Grillon et al., 2006).

2.1. Central, peripheral–physiological and behavioral indicators of classical conditioning

Classical conditioning studies in humans employ a range of behavioral, neuronal and peripheral–physiological indicators. On the neuronal level, studies examining electroencephalographic (EEG) (e.g., Keil et al., 2007; Stolarova et al., 2006) and magnetoencephalographic (MEG) measures (e.g., Kluge et al., 2011; Moratti and Keil, 2009; Weisz et al., 2007) during classical conditioning have shown that stimuli can attain access to preferred processing through associations with affective value such as danger/threat or reward. As synchronous oscillations are ideally suited for flexible formation of cell assemblies relevant to acquisition and extinction of CRs (e.g., Pare et al., 2002), and entailing rapid switches in the presence of affective stimuli, oscillatory signals can provide important information about controlling behaviors related to critical changes in stimulus associations during classical conditioning (e.g., Flor et al., 1996; Headley and Weinberger, 2011; Miltner et al., 1999). Electrophysiological studies of human classical conditioning have indicated a range of cortical changes involved in conditioned responses. Increased gamma band activity, and more importantly gamma band coherence between brain regions that receive two classes of stimuli (as required for building associations during conditioning), have been found to be involved in associative learning (Miltner et al., 1999). Moreover, different early and late components of the EEG slow potential response and gamma band activity have been shown to be differentially involved in aversive classical conditioning depending on the stage of conditioning and the affective quality of the CS (Flor et al., 1996).

In addition, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have been employed to investigate the neurobiological bases of classical conditioning. Processes related to the acquisition and extinction of classically CRs are related to brain regions including the amygdala, hippocampus, insula, anterior cingulate cortex (ACC) and regions of the prefrontal cortex (PFC), in the medial region (mPFC) (Sehmeyer et al., 2009), as shown in human lesion, PET and MRI studies (e.g., Bechara et al., 1995; Knight et al., 2004; Sehmeyer et al., 2009). In addition, activation of the striatum has been observed during both appetitive and aversive conditioning (e.g., Jensen et al., 2003). Moreover, reinforcement (direct prediction error) learning is tracked by activation in the striatum, the corticomedial amygdala, and the midbrain, whereas the basolateral amygdala may track associability (Boll et al., 2013; Li et al., 2011), and classical delay fear conditioning relies on subcortical brain structures including the amygdala while trace conditioning involves the hippocampus. Eyeblink conditioning in particular involves the cerebellum (Jirenhed et al., 2007; McCormick and Thompson, 1984). In addition, additional CS–US combinations that include, for example, the presentation of the CS before the CS–US association (latent inhibition) or involve more complex CS–US associations as in blocking designs (e.g., Arcediano et al., 1997) may be used to uncover dysfunctional learning processes that reflect, for example, insufficient inhibitory processes. Dysfunctions in learning processes of these latter types involve differing neural systems and thus translate into different mental disorder symptoms. Furthermore, activation of pain-related motor and somatosensory areas has been observed specifically when tactile stimuli were used as USs (Sehmeyer et al., 2009). While the amygdala has been identified as an important brain region for the acquisition and expression of conditioned fear responses (Maren and Quirk, 2004), the extinction of conditioned responses is thought to also depend on the PFC. The hippocampus has been identified as central for contextual modulation of fear acquisition as well as reinstatement and renewal processes that relate to extinction memory (cf. Acheson et al., 2012; Vervliet et al., 2013). Ventral striatal activation during classical conditioning has been described as representing prediction error (PE) and salient cue processing, independent of the valence (appetitive or aversive) of the stimulus (e.g., Delgado et al., 2008; Jensen et al., 2003, 2007; Li et al.,
between the CS and US captures explicit learning components with ability to detect similarities between non-identical but related stimuli. This stimulus generalization mechanism is highly adaptive in that the CR to a range of novel stimuli resembling the original UR-eliciting US. Lissek et al., 2005, 2010). Moreover, conditioned responses can extend to mental disorders (e.g., Bouton et al., 2006; Dunsmoor et al., 2011; Maren et al., 2013). In addition, and more important for the investigation of classical conditioning in PTSD and anxiety disorders, differential conditioning provides information about threat or danger (e.g., aversive or paired CS) versus safety (non-paired CS) processing, and in consequence approach and avoidance behavior, which may be of particular importance to understanding mental disorders.

3. Pavlovian conditioning in mental disorder

While simple conditioning procedures involve conditioning of a single, isolated CS, differential conditioning procedures include two distinct CSs, one of them (CS+) reliably paired with the US, and the other (CS−) never paired with the US. In human research, differential conditioning is the most widely used procedure, because it controls for possible non-associative learning effects on subsequent responses such as dishabituation or sensitization that may appear during presentation of the paired CS. Additionally, and more important for the investigation of classical conditioning in mental disorders, differential conditioning provides information about threat or danger (e.g., aversive or paired CS) versus safety (non-paired CS) processing, and in consequence approach and avoidance behavior, which may be of particular importance to understanding mental disorders.

3.1. Cue conditioning

Data on aversive cue conditioning are reported for acquisition from studies in patients with PTSD, anxiety disorders (panic disorder, generalized anxiety disorder, social phobia, specific phobia), depression, borderline personality disorder, schizophrenia, psychopathy, and addiction, and for extinction in patients with PTSD, anxiety disorders (specific phobia, panic), and schizophrenia. On appetitive conditioning, data only exist for acquisition processes in patients with schizophrenia and specific phobia.

3.1.1. Acquisition

3.1.1.1. PTSD. A number of studies have examined peripheral physiological measures of fear conditioning in PTSD patients compared to healthy controls. Several of these studies found increased SCR, heart rate, and EMG responses to a CS (e.g., colored circle) that was paired with an aversive (fear) US (e.g., painful electric stimulus or bursts of white noise) relative to an unpaired CS during acquisition (e.g., Orr et al., 2000; Peri et al., 2000), and, one study additionally reported increased SCR values to both CSs already during habituation (Orr et al., 2000). Exaggerated autonomic and facial fear responses were also found using an eyetrack conditioning procedure in PTSD patients compared to non-PTSD traumatized controls (Burris et al., 2007), with a differential response also evident on a neural level. Another study by Neylan et al. (1999) that examined PTSD patients compared to healthy controls...
showed larger P300 and N100 amplitudes and shorter P300 and N100 latencies in response to trauma-related stimuli along with an increased P50 ratio suggestive of impaired gating during learning in these patients (Neylan et al., 1999). Work by others has shown that patients with PTSD (compared to traumatized non-PTSD individuals or healthy controls) exhibit hyperactivation of the amygdala in response to fearful versus neutral CSs during acquisition (e.g., Bremner et al., 2005; Liberzon and Sripada, 2008; Shin et al., 2006). However, this enhanced CS+/CS− differentiation in PTSD patients, evidenced also by more negative evaluations of CSs associated with trauma-reminders (e.g., Wessa and Flor, 2007), has not consistently been found across previous research. For example, one study that used an eyelid conditioning procedure found similar conditioned responses during acquisition in PTSD patients and non-PTSD traumatized subjects compared to healthy non-traumatized controls (Ayers et al., 2003).

Other published studies have reported reduced discrimination between danger and safety signals in patients with PTSD. For example, Grillon and Morgan (1999) found that PTSD patients showed no significant differences in startle response during paired versus non-paired CSs, with reactivity increased for both CS− and CS− relative to a non-stimulus condition, and an additional increase of the CR (i.e., startle reactivity) to CSs of each type between two sessions of acquisition. Moreover, traumatized individuals without a PTSD diagnosis in this study exhibited increased CRs to the paired CS compared to the non-paired CS and additionally a decrease of the CR from session one to session two. The authors interpreted these findings as indicating an overgeneralization of fear responses as well as reduced safety learning as possible characteristic of PTSD. Relatedly, another study found startle potentiation to the non-paired CS in PTSD patients who were cognitively aware of CS−US pairings compared to those who were unaware (Jovanovic et al., 2010b). Thus, even if patients did learn to discriminate between danger and safety cues on a perceived experience level, they were not able to use this knowledge to transfer perceptions of safety to the conditioned inhibition trials (Jovanovic et al., 2010b).

As mentioned above, the assessment of comorbidity is a critical factor in the analyses of patients according to the RDoC approach as it may provide a perspective on commonalities and differences across disorders and additional important information about disorder-specific characteristics. However, so far, only a few studies have examined effects of comorbidity on classical conditioning in PTSD. Jovanovic et al. (2010a) compared PTSD patients with comorbid depression to patients with either PTSD or depression alone, and also healthy controls. All patient groups in this study condition less well than the healthy individuals, but the comorbid patient group showed enhanced fear potentiated startle in response to the safety cue compared to the patients with PTSD alone. Depression may thus serve as a potential augmenting factor further impairing the processing of safety signals in PTSD. However, other studies have found no significant effects of depressive symptoms on aversive eyelid conditioning in combat veterans with PTSD (e.g., Ginsberg et al., 2008), and therefore any conclusions can only be tentative.

3.1.1.2. Anxiety disorders. The inability to suppress fear under standard conditions observed on a physiological level in PTSD has also been reported in some anxiety disorders. For example, Lissek et al. (2009) reported that patients with panic disorder exhibited fear potentiated startle responses to safety cues and therefore reduced discrimination between safety and danger signals during acquisition, indicating that the safety signal was processed as the aversive event in contrast to the danger signal. Moreover, Grillon et al. (2007) found reduced rates of conditioned eyelid responses during trace, but not delay conditioning in panic patients. This might be indicative of a declarative, hippocampus-based associative learning impairment in panic disorder that disrupts cognitive processing of internal and external cues, resulting in reduced discrimination learning in such patients. This finding is corroborated by the few imaging studies that have addressed the role of brain activation patterns during fear conditioning in panic disorder. Patients with PD compared to healthy controls exhibited increased activation in the amygdala, midbrain structures, and the subgenual cingulate in response to the unpaired CS (safety signal) during acquisition (Tuescher et al., 2011). This in line with other behavioral studies on fear conditioning in PD that have found, for example, increased fear potentiated startle for safety cues as reported above (Lissek et al., 2005, 2009), again suggesting alterations in the processing of safety signals in PD. However, another study by Tuescher et al. (2011) did not find any effects for behavioral or autonomic measures which do not allow comparisons of different levels of
and the USs to the disorder needs to be considered (Lissek et al., 2008b for SAD; Schweckendiek et al., 2011 for specific phobia; Wessa and Flor, 2007 for PTSD). Moreover, the affective quality of such disorder-related stimuli might be relevant, as shown for example, as shown for example by Straube et al. (2004), who reported differences in reactivity of brain regions including the insula, amygdala, and parahippocampal gyrus in response to threatening faces in social phobics compared to healthy controls. However, there are also findings that do not support these interpretations. For example, a recent study by Tinoco-González et al. (2014) reported no significant differences in CS ratings of anxiety, valence, or arousal for SAD patients relative to non-patients during fear acquisition in a conditioning paradigm utilizing a socially relevant US (Tinoco-González et al., 2014). However, these authors compared SAD patients with individuals, who although not clinical patients, showed sub-clinically high levels of social anxiety, and the results are thus not directly comparable to those of prior studies that examined SAD patients in relation to healthy controls.

A similar pattern of results has been reported for persons with some (although not all) specific phobias. Olatunji et al. (2009) reported that patients with analogue blood injection-injury phobia, relative to healthy control individuals, showed only marginally increased differences in fear ratings following learning in an evaluative conditioning procedure in which fear and disgust evaluations of the stimuli took place before and directly after a learning in which participants focused on the stimuli for a memory recall task (Olatunji et al., 2009). These findings are in line with those of a previous study that reported increased muscle facial muscle tension in response to disgust-paired neutral pictures in high versus low blood-injury fear individuals (Schienele et al., 2005). However, this increased muscle tension was also found in response to depictions of smiling figures, making it difficult to interpret the findings (Schienele et al., 2011). Along related lines, another study by Olatunji et al. (2009) reported that acquisition-related ratings of disgust were greater in patients with blood injection-injury phobia compared to controls.

Schweckendiek et al. (2011) reported that patients with spider phobia showed increased activation in several brain regions (including the amygdala, the mPFC, the ACC, the insula, and the thalamus) in response to a conditioned stimulus that was paired with a spider picture as US, whereas no significant differences were found in response to a CS that was paired with an aversive, but non-phobia-related US (Schweckendiek et al., 2011). In addition, spider phobics in this study exhibited increased amygdala activation in response to phobia-related CSs (i.e., pictures of spiders) versus non-phobia related CSs. Findings from this study further highlight the potential importance of the disorder-relevance of CSs and USs as a moderator of conditioning effects—a parameter that needs to be more clearly examined in future studies, both within and between disorders.

Another study that compared patients with flying phobia to healthy controls (Frieden et al., 2012) found increased valence ratings for a CS that was paired with a pleasant US during acquisition relative to a CS that was paired with an unpleasant US as well as increased fear ratings to both the pleasant-paired and the unpleasant-paired CSs. These findings once more indicate differences in the learning depending on whether the CS or US is relevant for the specific disorder. In the Frieden et al. study, such complex interactions were especially apparent given that patients with flying phobia not only appeared to learn appetitive associations “better” than aversive associations, but also showed increased fear in an appetitive learning context, perhaps indicating that these patients are not able to benefit from positive events, or may in fact reverse pleasure into aversion.

Finally, in patients with generalized anxiety disorder, Cha et al. (2014) reported increased activation in the ventromedial (vm)PFC in response to an aversively paired CS in GAD patients compared to healthy controls. Somewhat in contrast to this, a study of youths with GAD (Britton et al., 2013) found reduced activation in prefrontal regions including the vmPFC and the ACC three weeks following fear conditioning in response to morphed fearful faces that were paired with an aversive US during the conditioning procedure. In addition to differences in conditioning procedures used, these contrasting results for the PFC across studies might be related to the fact that the prefrontal cortex develops late and that prefrontal functions in adolescents differ from those in adults (Ernst and Müller, 2008; Ernst et al., 2006).

3.1.1.3. Depression. Although evidence exists that depression may be characterized by altered instrumental learning processes (e.g., Kuehner et al., 2011), there are also some studies that have reported deviations in classic conditioning. A study by Nissen et al. (2010) found no significant differences in classical discrimination learning between patients with major depressive disorder and healthy controls at the level of contingency awareness during fear acquisition. However, patients in this study exhibited increased SCRs to paired versus non-paired CSs during fear acquisition, whereas healthy controls showed no such difference. The finding of increased SCRs to the paired versus non-paired CS in depressed patients may indicate a prominent role for learning-related arousal in this disorder. In turn, the observed SCR increase may reflect increased general activation of the amygdala, given its status as the central neural substrate of fear acquisition, and of the expression of fear-related autonomic responses (Cacciaglia et al., in press; Phelps, 2006).

In other work, Greer et al. (2005) reported that depressed patients exhibited significantly fewer CRs in the form of blink responses, for both delay and trace eyelinking conditioning. This is in contrast to the findings of increased differential SCR responses in depressive patients reported by Nissen et al. (2010), and points to aversive conditioning deficits as a potential pathogenetic factor for major depression, with a possible contribution of cerebellar and hippocampal dysfunction for the development of the disorder given evidence for a role of these structures in delay and trace eyelinking conditioning. The contradictory findings across these two studies could be related to the differing conditioning procedures that were used, as eyelink conditioning represents a rather motor-response-driven learning process whereas classical fear conditioning more strongly reflects an affective-motivational, learning process.

3.1.1.4. Borderline personality disorder. Findings for patients with borderline personality disorder highlight the specific importance of dissociative symptoms for associative learning effects in this patient group. BPD patients with higher levels of dissociative experiences show reduced rather than enhanced acquisition of aversive delay conditioning, as evidenced by diminished differences in valence and arousal ratings for the paired versus non-paired CSs, compared to both healthy controls and BPD patients with low states of dissociative experiences (Ebner-Priemer et al., 2009). Moreover, Ebner-Priemer et al. (2009) reported reduced, fear acquisition for high-dissociative BPD patients, not only on a perceived-experiential, but also a psychophysiological level (i.e., these patients showed diminished SCR in response to paired versus non-paired CSs relative to both healthy controls and BPD patients with low dissociative symptoms). These results could reflect differences in brain structure and function that may affect learning processes. For example, dissociative states in this study were shown to be associated with reduced activation in the amygdala, a brain region that drives fear acquisition and is known to be altered in BPD.

3.1.1.5. Schizophrenia. Holt et al. (2012) reported that patients with schizophrenia compared to controls showed blunted brain activation in the posterior cingulate gyrus, precuneus, inferior parietal cortex, hippocampus, and thalamus during the acquisition phase of a fear conditioning task, but no significant differences during fear extinction. Other studies have reported impaired behavioral performance, as evidenced by reduced eyelink amplitudes (Parker et al., 2013) as well as onset and peak latencies (Marenco et al., 2003), in conjunction with decreases in rCBF specifically in the middle and medial frontal lobes, anterior
cerebellar lobules I/V and VI, and the thalamus (Parker et al., 2013). However, in contrast with these findings, Sears et al. (2000) reported increased CRs during a delay eyelink conditioning procedure in schizophrenia patients as compared to healthy controls.

3.1.1.6. Psychopathy. There are some studies that have examined classical aversive conditioning in psychopathy (e.g., Birbaumer et al., 2005; Flor et al., 2002; Hare, 1965; Hare and Quinn, 1971; Rothemund et al., 2012; Veit et al., 2002). These studies have generally found a lack of differentiation between the paired CS and the non-paired CS in electrodermal measures and startle potentiation as well as valence and arousal ratings, whereas contingency ratings appeared intact (see also Sommer et al., 2006). In addition, diminished activation in the amygdala and orbitofrontal cortex as well as the insula and ACC were observed, suggesting deficient associability of fear with intact cognitive processing of fear. This interpretation is supported by electrocortic research demonstrating intact ERPs and even higher anticipatory contingent negative variation amplitudes in high-psychopathy patients compared to controls EEG recordings. Moreover, Veit et al. (2013) suggest a role for specific subtypes of psychopathy, in that specifically more affective/interpersonal symptoms in particular showed reduced SCRs to aversive CSs. Cognitive-emotional interactions may therefore represent an important modulating factor in fear conditioning in psychopathy.

3.1.1.7. Substance addiction. For addiction, instrumental/operant conditioning may be of great importance to the development and maintenance of addictive behaviors and symptoms, yet, there are also two studies that have addressed fear conditioning in alcohol dependent individuals (Finn et al., 1994; Stephens et al., 2005). Both studies found that individuals at high risk for alcohol addiction failed to acquire associations between the CS and US, as indicated by impaired SCRs (Finn et al., 1994; Stephens et al., 2005). In addition, some other studies that have investigated delay eyelink conditioning in abstinent alcohol-dependent patients have found reduced conditioned eyelink responses in the patients compared to healthy controls (Fortier et al., 2008; McGlinchey-Berroth et al., 2002). These findings indicate a deficient evaluation of risk (perhaps reflecting a weak behavioral inhibition system) that may represent a liability for the development of alcoholism. Although these findings appear to contrast with the literature on cue reactivity, which provides evidence for increased responsiveness to neutral cues that may serve as CSs, they actually corroborate these findings, because fear conditioning and cue reactivity relate to opposing processes in terms of affective valuation. That is, whereas in fear conditioning the US is a threatening stimulus, cue reactivity entails associations with the drug itself, i.e., a positive stimulus serves as US, and thus complements appetitive conditioning. Therefore, although predominance of instrumental conditioning tends to be viewed as the important process in addictions, the above-described findings for fear conditioning may have important implications for treatment of substance problems, and should be considered carefully in this context.

3.1.2. Extinction

3.1.2.1. PTSD. For PTSD, delayed fear extinction has been reported as a relatively robust finding across studies (e.g., Norholm et al., 2011; Orr et al., 2000; Peri et al., 2000). This has been demonstrated on both an experiential level (e.g., by increased US expectancy ratings) and a physiological level. Increased SCR, heart rate, and EMG responses to a CS paired with an aversive US during extinction was found for PTSD patients compared to healthy controls (e.g., Orr et al., 2000; Peri et al., 2000). Moreover, reduced activation of the PFC during extinction was PTSD patients compared to healthy controls. However, the pattern appears to be more complex, not only with respect to the differential responses to the paired versus non-paired CSs (i.e. fear versus safety signal), but also depending on the physiological measure used (e.g., skin conductance versus heart rate versus startle response). Some studies have found dampened SCRs to the paired CS (Bremner et al., 2005) or increased startle responses and SCRs to both paired and non-paired CSs in PTSD patients versus non-PTSD individuals (Grillon and Morgan, 1999; Orr et al., 2000; Peri et al., 2000). In line with the findings from extinction, this indicates reduced differential conditioning and deficits in safety signal processing as well as an overgeneralization of fear (Peri et al., 2000; Bremner et al., 2005; Blechert et al., 2007; Wessa and Flor, 2007; Milad et al., 2008; Jovanovic et al., 2009; Lissek et al., 2009).

The presence of second order conditioning with a specific failure to extinguish conditioned responses was reported by Wessa and Flor (2007) for PTSD patients compared to healthy controls and traumatized persons without PTSD. Patients with PTSD in this study exhibited delayed fear extinction as evidenced by increased US expectancy ratings following extinction not only compared to healthy controls without any traumatic experiences, but also compared to individuals with traumatic exposure but without a PTSD diagnosis (Mineka and Oehlberg, 2008). In addition, extinction-related processes such as extinction retention or recall of fear extinction are important mechanisms that may aid in recovery from a psychologically traumatic event, and as such might constitute a possible resilience factor (Rauch et al., 2006; Milad et al., 2006; Davis et al., 2006; Sotres-Bayon et al., 2004; Maren and Quirk, 2004). In line with these suggestions, Milad et al. (2008) investigated pairs of monozygotic twins discordant for combat exposure, using a two-day fear conditioning and extinction procedure to clarify the status of reduced extinction retention as a pre-existing versus acquired sign in PTSD. Half of the combat-exposed twins in this study exhibited PTSD. Whereas veterans with as compared to those without PTSD showed less extinction retention as indicated by larger SCRs to previously paired CSs, the twins of the PTSD individuals did not exhibit such impairments. Extinction retention deficits may therefore represent a disorder-related consequence resulting from combat trauma rather than a predisposing factor (Milad et al., 2008).

In the context of associative learning and psychopathology, the discussion of whether disturbances in conditioning are pre-existent vulnerability factors or rather a consequence of the disorder is a subject of ongoing debate. So far, the majority of studies are cross-sectional, but some longitudinal studies exist that provide support for the prediction of psychopathology from conditioning characteristics. One study by Guthrie and Bryant (2006) found that extinction processes were not only impaired in PTSD patients following the development of the disorder, but also operated as a risk factor for the later development of PTSD symptoms (i.e., the emergence of PTSD was predicted by slower extinction of corrugator electromyogram responses assessed before a traumatic event). Elsewhere, Lommen et al. (2013) studied Dutch soldiers deployed to Afghanistan, and found that reduced extinction learning prior to deployment (as indexed by conditioned responding over extinction trials) predicted subsequent PTSD symptom severity, over and beyond other risk factors such as stress symptoms, exposure, and neuroticism. These data indicate an important etiological role for extinction deficits in PTSD. Moreover, these findings are in line with the assumption that perceived CS-US expectancy may contribute to the persistence of PTSD (Mineka and Oehlberg, 2008). However, other potential predictors should also be examined in future research as the prospective relationship of extinction learning with PTSD symptomatology was only modest in the study by Lommen et al. (2013). In addition, the role of impairments in brain function, for example in the inhibitory function of the medial PFC, should be investigated as a contributor to extinction learning deficits (e.g., Quirk et al., 2006).

However, there may be subgroups of PTSD with common mechanisms, and comorbidity may moderate the psychophysiological response patterns characteristic of a specific disorder (cf. McTeague et al., 2009, 2010, 2012). Findings from Orr et al. (2012) suggested an influence of comorbidity such as depression and intelligence levels on physiological responding during conditioning. They found that SCRs to loud tones, as indicator of aversive/threat processing, have an effect
on extinction learning and posttraumatic stress symptoms. These findings suggest that additional factors, in particular comorbidity, need to be considered.

3.1.2.2. Anxiety disorders. While Olatunji et al. (2009), as noted earlier, reported that acquisition-related ratings of disgust were greater in spider phobia patients compared to controls (Olatunji et al., 2009), no significant group differences were found for extinction, either for fear or for disgust ratings (Olatunji et al., 2009).

By contrast, Michael et al. (2007) reported extinction deficits for panic patients compared to healthy controls in terms of larger SCRs to previously avertively-paired CSs during extinction, but comparable conditioned SCRs during acquisition (Michael et al., 2007).

3.1.2.3. Schizophrenia. Patients with schizophrenia may be characterized by a failure to demonstrate appropriate context gating of extinction memory retrieval. For example, whereas healthy individuals exhibited lower SCRs 24 h after successful fear conditioning, schizophrenia patients successfully acquired and extinguished conditioned fear responses as indexed by increased SCRs to the previously paired CS during extinction as compared to acquisition in this condition (Holt et al., 2012). In addition, after learning, schizophrenia patients exhibited reduced activation of the vmPFC in response to the extinction (safe) context, while showing enhanced SCRs to CS presentations as noted, whereas healthy controls showed an opposing pattern of decreased SCR to CSs along with increased vmPFC responses to the safe context (Holt et al., 2012). These findings point to a failure in extinction memory retrieval in schizophrenia, although other studies of patients with schizophrenia (Holt et al., 2009; Jensen et al., 2008; Romaniuk et al., 2010) have yielded mixed results, with partly reversed patterns of differential conditioned fear responses and higher responses to the non-paired CS (Holt et al., 2009; Romaniuk et al., 2010), and/or lower responses to the paired CS (Jensen et al., 2008; Romaniuk et al., 2010). However, studies so far are scarce and further research is necessary to reach firm conclusions. Finally, Parker et al. (2013) used an eyeblink conditioning procedure and found impaired behavioral performance as well as reduced activity in both the frontal lobe and insular cortex during extinction in schizophrenia patients. But again, this eyeblink conditioning procedure is not directly comparable with the classical fear conditioning procedure used in the other studies and thus does not allow clear conclusions.

3.1.3. Comparison among disorders

Studies focusing on patients with anxiety disorders have reported marginally, but nonsignificantly increased differentiation in fear ratings following cue-related learning, for example, in patients with analogue blood injection-injury phobia (Olatunji et al., 2009). However, other studies have found enhanced discrimination between CSs in patients with spider phobia compared to healthy control individuals (Schweckendiek et al., 2011). Furthermore, phobic patients compared to healthy controls in a study by Schweckendiek et al. (2011) showed increased activation in a fear-related brain network including the amygdala, insula, mPFC, ACC, and thalamus in response to a phobia-related CS (i.e. a geometrical figure that was paired with a spider-picture US), along with increased activation of the amygdala to phobia-related versus non-phobia-related CSs (Schweckendiek et al., 2011). Interestingly, no significant differences were found in this study for non-phobia-related conditioning processes in patients compared to controls (Schweckendiek et al., 2011).

Thus, groups in this study did not differ in experiential ratings or SCRs, and no clear differential SCRs were observed for healthy controls and in the spider phobics for responses to non-phobia related stimuli. These findings are in line with previous studies that also used pictures as US and did not observe significant differential SCRs (e.g., Klucken et al., 2009; Wessa and Flor, 2007). Taken together, these results point to the importance of phobia-related emotional learning rather than general fear conditionability for the development and maintenance of phobia.

Dysfunctions in conditioning are therefore not common for each type of phobia. Moreover, the finding of enhanced amygdala reactivity to phobia-relevant CSs supports the view that in phobia as for example compared to PTSD (Rauch et al., 2000), a hyperactivation of the amygdala may occur only in response to disorder-relevant cues (cf. Wright et al., 2003), while an enhanced activation of the insula during fear conditioning may represent a more general mediator also evident in patients with other anxiety disorders such as SAD (Veit et al., 2002).

However, findings on aversive (fear) conditioning from the various studies reviewed in this section and the preceding one on acquisition learning are not directly comparable, because these studies used different conditioning procedures. For example, the study by Vriends et al. (2012) on flying phobia used a modified version of Olson and Fazio’s associative learning paradigm (Olson and Fazio, 2001) in which participants viewed a series of distractors interspersed with pairings of novel objects as counterbalanced CSs, with frightening and pleasant stimuli that served as USs. Comparisons were made between the CS that was paired with the aversive US and the CS that was paired with the appetitive US, rather than between a CS that was paired with an affective US and one that was never paired with this US, as is usually the case in classical conditioning procedures. The authors interpreted their findings as evidence for a stronger conditioning effect in flying phobia that may contribute to the etiology of specific phobias (Vriends et al., 2012); however, their finding of an increased rated fear response to both pleasant-paired and unpleasant-paired CSs further (and perhaps more importantly), indicates difficulties in processing a specific kind of safety information in order to distinguish CSs of the two types. This safety information is linked by contiguity to the affective quality of the pleasant US, and thus may add a positively reinforcing element to the CS beyond its safety-signal quality. Notably, the findings of Vriends et al. (2012) add to previous suggestions that not only stronger conditioning per se, but also reduced safety learning and generalization of fear may be a distinct characteristic of focal fear conditions such as flying phobia (Mineka and Zinbarg, 1996). Both discrimination and safety learning mechanisms might interact and contribute to the development or maintenance of specific phobias.

In comparison to findings from other anxiety disorders or studies of patients with PTSD, the finding of increased CRs to both the paired and non-paired CS fits well with other associative learning studies that have reported larger CRs (indicating enhanced fear) in patients with anxiety and stress-related disorders relative to healthy controls (Grillon and Morgan, 1999; Orr et al., 2000; Peri et al., 2000; Wessa and Flor, 2007). However, it is also apparent that different mechanisms may account for diagnosed subgroups of specific phobias, with a larger proportion of commonalities found across different anxiety disorders than across different types of specific phobia. Nonetheless, it seems clear that fear conditioning is an important mechanism that plays an important role in the success of therapy based on extinction learning (e.g., Hamm, 2009). This point further underscores the need for and importance of the RDoC approach.

Based on classical conditioning results from experiential-report and behavioral as well as physiological measures, enhanced generalization may be important in PTSD and may be indicative of increased acquisition and reduced extinction of fear as an important factor in this condition. This suggestion is in line with the majority of findings pertaining to conditioning differences in anxiety disorders. However, although studies on fear conditioning in SAD have yielded results similar to those for PTSD, investigations of SAD (e.g., Lissek et al., 2008b) are again not directly comparable to PTSD or other anxiety disorders such as PD. This is because various CSs with differing affective qualities have been used in studies of these other conditions, and paired with an aversive US, so that no “real” safety signals were included to denote the absence of an aversive US. Notably, findings from studies on classical
conditioning in SAD that have used “classical”, socially-irrelevant aversive USs such as unpleasant odors or painful pressure are mainly in line with those from studies of other anxiety disorders, showing no increase in fear conditioning for SAD patients, either in SCR (Hermann et al., 2002; Veit et al., 2002), heart rate (Schneider et al., 1999), or fear potentiated startle responses (Hermann et al., 2002). However, this lack of fear learning related increases in patients may be related to group differences already present in the habituation phase.

One study in panic disorder found increased discrimination learning in patients compared to controls, indicated by increased fear potentiated startle responses to the paired versus non-paired CS during acquisition. This would link PD to impaired discrimination rather than reduced safety learning — at least on a very reflexive, non-conscious level represented by the startle reflex (Lissek et al., 2009). Interpretation of these contradictory results is difficult, given that studies that have reported conflicting results have used many common design parameters. This may indicate an involvement of other factors, for example, variations in anxiety or dissociative symptoms that might at least partly explain the different findings in patients diagnosed with the same principal diagnosis (e.g., Grillon et al., 2008; Ebner-Priemer et al., 2009). Accordingly, we encourage careful and systematic consideration of these additional factors.

Moreover, for some disorders, especially depression or BPD, details regarding the processing of safety signals and results pertaining to extinction learning are so far missing and need to be addressed in future studies. Doing so will help to advance our understanding of anxiety-related disorders and shed more light on possible commonalities and differences in relation to other mental disorders.

Studies on individuals with psychopathy (e.g., Birbaumer et al., 2005; Flor et al., 2002; Veit et al., 2002) point toward a lack of electrodermal and rated anticipatory fear responses to CSs previously paired with an aversive event. However, this reduced emotional learning tends to be accompanied by intact contingency ratings and brain ERP responses, suggesting no deficits on a cognitive level (Sommer et al., 2006). In addition, deficient amygdala and orbitofrontal responses suggest an inability to form a basic emotional association. This is a relatively unique pattern that has not been reported for other disorders, but which may vary across differing subtypes of high-psychopathy individuals (e.g., Drislane et al., 2014; Hicks et al., 2004).

There are also some recent studies that have directly compared different anxiety patient groups. One study by Tinoco-González et al. (2014) found no significant differences in CS ratings of anxiety, valence, and arousal during fear acquisition for SAD patients compared to individuals with high but sub-clinical social anxiety and/or PD patients. Another study by Otto et al. (2014) points to an important involvement of other factors that may modulate outcomes in fear conditioning studies. These authors investigated patients with anxiety and mood disorders, and found similar response patterns in PTSD, panic disorder and depression, marked by lower SCR indicative of reduced fear acquisition for PTSD and depression (especially compared to controls) and also slower extinction among the patients, in particular those with panic disorder (Otto et al., 2014). Across these targeted disorders, fear conditioning effects were moderated by cognitive processes and deficits, medication use, and comorbidity with mood disorder (Otto et al., 2014). This suggests the need for more comparative studies that also include subgroup analyses.

3.2. Context conditioning

Data on aversive context conditioning exist for acquisition in patients with PTSD only, and for extinction in patients with PTSD, anxiety disorder (phobia), and schizophrenia. For appetitive context conditioning, we did not find any published work.

3.2. Acquisition

3.2.1. PTSD. Besides enhanced cued fear conditioning and delayed cued extinction, a contextual processing deficit, and thus impaired contextual learning, have been proposed for PTSD (e.g., Acheson et al., 2012; Brewin et al., 2010; Flor and Wessa, 2010; Gilbertson et al., 2007; Kremen et al., 2012; Rougemont-Bücking et al., 2011). This hypothesis is based in part on findings indicating reduced hippocampal volume as both a vulnerability factor for the development of PTSD and as a consequence of the disorder (Gilbertson et al., 2002; Bremner, 2001), and also on the fact that the hippocampus is involved in contextual conditioning in humans as well as animals (Alvarez et al., 2008; Lang et al., 2009; Marschner et al., 2008). Patients with PTSD may thus be unable to use or profit from contextual information to control their fear responses to a previously aversive CS that no longer predicts an aversive outcome, and thus show impaired safety learning (e.g., Liberzon and Sripada, 2008).

We tested whether this effect might be related to deficient context or enhanced cue conditioning, and found that PTSD patients compared to traumatized individuals without PTSD showed reduced discrimination between the paired (danger) and the non-paired (safe) context, as evidenced by lesser differences in contingency ratings for danger versus safe contexts (Steiger et al., unpublished data). PTSD patients in this study showed improved discrimination, to the level of both the traumatized and healthy individuals, during a subsequent context-cue acquisition phase that included learning to cues presented together with the previously conditioned aversive versus safe contexts (Steiger et al., unpublished data). These results suggest that PTSD patients, compared to traumatized non-PTSD individuals and healthy controls, appear less able to discriminate contexts and show concurrently more hippocampal activation and enhanced differential left insula activation during context-cue acquisition compared to non-PTSD groups. Moreover, deficient contextual acquisition and extinction appear to be related to PTSD symptoms. These findings suggest that deficient contextual learning may contribute specifically to PTSD. However, there is also one study that found no significant difference in SCRs during contextual fear acquisition for PTSD patients compared to either traumatized individuals or healthy controls (Rougemont-Bücking et al., 2011), pointing to a need for further studies.

3.2.2. Extinction

3.2.2.1. PTSD. Levy-Gigi et al. (2015) reported context overgeneralization after conditioning in PTSD patients compared to controls in conjunction with significantly reduced hippocampal volumes such that smaller hippocampal volume predicted overgeneralization. The implication is that patients with PTSD may continue to show fear responses to trauma cues in contexts in which these cues no longer predict danger. In line with this, Milad et al. (2009) reported that PTSD patients displayed robust CRs in the form of increased SCRs to a previously extinguished CS presented within the extinction context, indicating deficits in extinction retention in PTSD (Milad et al., 2009). However, another study by Rougemont-Bücking et al. (2011) found no significant differences in SCRs during the extinction of context in PTSD patients compared to both traumatized individuals and healthy controls (Rougemont-Bücking et al., 2011). However, the findings appeared more consistent on a neural level: In PTSD patients, compared to healthy controls, showed impaired activation in the hippocampus and the vmPFC along with enhanced activation in the dorsal ACC during extinction recall (Milad et al., 2009), and reduced vmPFC activation during extinction (Milad et al., 2009; Rougemont-Bücking et al., 2011) were evident. This result points to increased fear renewal in PTSD, and suggests deficient maintenance of extinction and failure to identify safety signals as a key characteristic of PTSD.
3.2.2.2. Anxiety disorders. One study of spider phobics by Dibbets et al. (2013) investigated effects of a context switch or exposure to differing contexts as methods for preventing the context-related return of fear (renewal) following exposure therapy, and found increased arousal and anxiety ratings in response to a context switch compared to no switch in patients with spider phobia (Dibbets et al., 2013). Another study by Shiban et al. (2013) found that extinction in multiple contexts compared to extinction in one single context reduced the renewal effect in spider phobics (Shiban et al., 2013). Findings from this latter study appear to contrast with those of Bouton (e.g., 2004), who characterized extinction conditioned fear as context-specific. However, in the study by Shiban et al. (2013) extinction had already taken place in different contexts, and thus context switch occurred not only following extinction. From this perspective, data from this study also corroborates findings from and assumptions made by Bouton (e.g., 2004), providing further evidence that extinction is specifically contextually bound. The presentation of differing contexts already in the extinction phase allowed participants to interrupt renewal effects by blocking the normally strong renewal effect evident when only a single context is used during extinction. Moreover, on a physiological level, significantly reduced CRs were observed for spider phobics in Shiban et al.’s study, as indicated by decreased SCRs. Such context manipulations may help to improve the generalizability of extinction to a new context and thereby overcome the overgeneralization during cue conditioning often reported for patients with anxiety disorders.

3.2.3. Comparisons among disorders

Although dysfunctional processing of contextual information has been discussed as a critical mechanism in several forms of psychopathology, including anxiety disorders, PTSD, schizophrenia, depression, and drug addiction (Maren et al., 2013), only a few studies to date have reported results from context conditioning tasks in patient groups. While results from cue conditioning studies have been reported for many disorders, contextual fear conditioning studies have been conducted only for PTSD and phobic disorders. Deficits in contextual learning in PTSD and specific phobia may result in rigid and inflexible behavior and thus inappropriate adaptations, because the context plays an important role in the retrieval of information and the flexible representation of information including resolving ambiguity.

4. Commonalities and differences across disorders

The above reviewed studies indicate that findings from cue and context conditioning may corroborate each other. However, it should be noted that a large range of classical conditioning procedures were used across studies, making it difficult to draw direct conclusions, particularly for those mental disorders where only a few studies exist. Studies to date differ in the use of simple versus differential conditioning procedures, motor (eyeblink) versus classical fear conditioning, and stimuli with different qualities (e.g., generally aversive affective stimuli versus aversive disorder-related stimuli). Nevertheless, the results of these studies provide evidence that differential fear acquisition and reduced extinction and thus reduced safety learning may represent the most important mechanisms in the development and maintenance of mental disorders (see also Fig. 1 and supplementary Table 1 for an overview). For patients with PTSD, there are a relatively large number of studies that have used similar and therefore comparable fear conditioning paradigms, which renders methodological variations rather unlikely as a possible explanation for the somewhat differing findings reported across these studies. However, in studies of PTSD, subgroups based on time and complexity of trauma or comorbidity have rarely been considered. In addition, differences may have emerged depending on whether PTSD patients and healthy controls or patients and non-PTSD individuals were compared, indicating that trauma exposure per se may exert a distinctive effect. This is important in terms of considering risk versus resilience factors in PTSD as it is still not clear which mechanisms directly promote the development of PTSD symptoms following trauma exposure. Some longitudinal studies have provided tentative evidence that reduced extinction learning in particular may predict the development of PTSD following a traumatic event (e.g., Lommen et al., 2013). Moreover, comorbidity also plays an important role and PTSD is the only disorder for which the presence of additional symptoms, for example of depression, has been shown to result in stronger fear conditioning (e.g., Otto et al., 2014). Last, the severity of PTSD symptoms may affect classical conditioning as indicated by above-mentioned findings demonstrating associations of PTSD symptom severity with increased fear responses to safety cues and the severity of current PTSD symptoms (Jovanovic et al., 2009), and with reduced context and context-cue conditioning and a failure to extinguish (Steiger et al., unpublished).

Not only for PTSD, but also for other anxiety disorders, mood disorders, schizophrenia, and personality disorders such as BPD and psychopathy, it has been demonstrated that not only acquisition but also extinction and extinction memory are altered in patients compared to healthy controls, and that extinction deficits may even be the more important mechanism for symptom development and maintenance. This was seen across differing measures of conditioning including experiential ratings, peripheral physiology, and central (i.e., brain response)
measures. However, some conflicting results have been reported for fear acquisition. Within and across disorders, findings range from no discrimination in responding to paired (danger) versus non-paired (safety) CSs to reduced discrimination learning in patients compared to controls. Such differences indicate the need for (a) comparative studies across disorders and the differentiation of subgroups within disorders, and (b) different treatment strategies depending on the occurrence of specific symptoms or additional modulators such as comorbidity of cognitive status. In turn, these considerations clearly underline the importance of the RDoC approach, in that common mechanisms contribute to differing forms of psychopathology and their associated symptomatology (see also McTeague et al., 2009, 2010, 2012).

Another important aspect of associative learning that has not been widely examined in mental disorders is the processing of prediction errors. An exception is the study of addiction, but in this context prediction errors have mainly been highlighted for instrumental behavior (e.g., Garrison et al., 2013). Reduced safety learning and overgeneralization might indicate deficits in prediction error processing that in turn underlie changes in approach and avoidance behaviors often observed in anxiety disorders (Compton et al., 2007; McNally et al., 2011), depression, schizophrenia (e.g., Gradin et al., 2011), and psychopathy (e.g., Blair, 2007).

Some studies have additionally addressed delay versus trace conditioning in mental disorders, particularly anxiety disorders, thus allowing some conclusions to be advanced regarding event timing and associated cognitive processes in the symptomatology of these disorders. Deficient responses to both delay and trace eyeblink conditioning have been observed especially in PTSD and depression, while in panic disorder only impaired trace, but not delay conditioning has been found.

Moreover, for almost no disorder, except schizophrenia, do studies exist on appetitive conditioning. While, this may be a very important mechanism for addictions, it has only been examined for operant procedures (which are beyond the scope of the present review), with studies on classical appetitive conditioning lacking.

5. Summary and outlook

Impairments in differing aspects of classical conditioning seem to characterize various mental disorders, and conditioning may be a mechanism that could greatly contribute to understanding of such disorders.

It is also apparent that there are differences within and between disorders and that not all disorders seem to be characterized by distinct aversive learning mechanisms (see Fig. 1 for an overview). As mentioned before, dysfunctional discrimination learning may be the mechanism that plays a role in the greatest number of disorders discussed in this review. However, it has also been the most investigated mechanism to date, along with extinction of conditioned fear, and therefore future studies focusing on mechanisms other than these are needed to draw firm and specific conclusions. Moreover, several possible moderators need to be taken into account, including comorbidity, symptom severity, cognitive factors such as dissociation, and gender.

In addition, in conditions such as anxiety disorders, determining how to minimize the return of fear after therapeutic interventions is important for ensuring the effectiveness of such treatments. Conducting exposure therapy using a procedure with multiple contexts (i.e. treatment that is performed across a variety of environments), may reduce fear relapse (Dunsmoor et al., 2014), although this needs to be clearly tested in future studies. Furthermore, the timing of extinction-related, exposure-based intervention is another important factor that may determine treatment success (e.g. Alvarez et al., 2007; Schiller et al., 2008). For example, Mystkowski et al. (2006) found that individuals who had mentally reinstated the treatment context before they encountered a phobic stimulus in a new context, showed less return of fear than individuals who did not.

Using neuroimaging methods, future research should also address more closely, structural and resting state brain anomalies that relate to associative learning impairments in patients with mental disorders. For example, differences in brain activation during fear memory consolidation and experiential fear ratings may serve as indices of homeostatic processes in the vmPFC and the ACC as well as the functional resting brain fluctuation (Feng et al., 2013).

In addition, stress has been shown to affect learning and memory processes involving neural correlates (e.g., van Stegeren, 2009). Since many mental disorders including anxiety disorders, PTSD, and depression may be characterized by alterations in stress systems and networks (e.g., Rhebergen et al., 2015) and related to stress-related experiences such as early life adversity (Mouhaan et al., 2014), a central role for stress in anomalous fear learning processes in mental disorders may be assumed. For example, Jovanovic et al. (2010b) demonstrated that impaired fear inhibition in PTSD patients was associated with alterations in HPA axis feedback related in turn to amygdala hyperactivity. Future studies addressing return of fear in these patients, and also in patients with other mental disorders, may provide further important information, given that acute stress has notably been shown to reduce rather than promote the return of fear (Merz et al., 2014), and may enhance the consolidation of extinction memory (Hamacher-Dang et al., 2013). Findings on such mechanisms in patient samples may have potential applications in extinction-based therapeutic approaches.

Moreover, the identification of pharmacological targets or epigenetic mechanisms that may improve the efficacy of exposure therapy is of further significance. Additionally, exposure-based treatment might benefit from a specific focus on instructions related to processing safety signals, given evidence for heightened fear following treatment in phobic patients instructed to focus on safety signals during exposure sessions, as opposed to focusing on experienced fear (Sloan and Telch, 2002), along with evidence that perceptions of safety may reduce positive treatment outcome (Craske et al., 2014; Powers et al., 2004).

Finally, genetic variations may modulate associative fear learning in patients with mental disorders. In healthy individuals, it has been shown, for example, that genetic variants related to modulation of the HPA axis and neuroendocrine stress circuits are associated with cue conditioning, and variants related to calcium signaling and memory processes and the regulation of the stress response are associated with context conditioning (Ridder et al., 2012; Pohlack et al., 2012; Pohlack et al., in press). These gene variants also play a role in mental disorders such as PTSD. Further research is needed to identify the predictive nature of learning processes and behavioral, psychophysiological, and plastic brain changes for the transition into specific disorders such as PTSD as well as its maintenance (Flor and Nees, 2014). Moreover, the identification of the interaction of these processes with genetic characteristics in this prediction is of further importance. Finally, along these lines, mechanisms should be compared both within and across disorders (Flor and Nees, 2014). In this context, cross-validations between animal and human genetic models and between preclinical and clinical work may be important to take into account, particularly in evaluating the potential role of candidate genes, as well as their neural and psychophysiological pathways, for mental disorders (e.g., Almli et al., 2014). The integration of epigenetic data and gene and environment interactions are further promising approaches that may additionally provide an avenue for new, also pharmacologically based, treatment approaches (e.g., Zovcic and Sweat, 2013). Finally, not only aversive but appetitive conditioning should be considered as an important mechanism, and further learning-related mechanisms such as reconsolidation should be addressed, as such processes may for example, be used to prevent the return of fear in humans (Schiller et al., 2010).

Conflict of interest

The authors have no conflicts of interest.


