Fear generalization, in which conditioned fear responses generalize or spread to related stimuli, is a defining feature of anxiety disorders. The behavioral consequences of maladaptive fear generalization are that aversive experiences with one stimulus or event may lead one to regard other cues or situations as potential threats that should be avoided, despite variations in physical form. Theoretical and empirical interest in the generalization of conditioned learning dates to the earliest research on classical conditioning in nonhumans. Recently, there has been renewed focus on fear generalization in humans due in part to its explanatory power in characterizing disorders of fear and anxiety. Here, we review existing behavioral and neuroimaging empirical research on the perceptual and non-perceptual (conceptual and symbolic) generalization of fear and avoidance in healthy humans and patients with anxiety disorders. The clinical implications of this research for understanding the etiology and treatment of anxiety is considered and directions for future research described.

**Keywords:** fear conditioning; generalization; avoidance; anxiety

Fear is a remarkably adaptive behavioral response, allowing us to predict, react, and adjust to past, present, and future threat. Fear learning may, however, go awry or become excessive and lead to the development of psychopathology. In experimental psychopathology, the behavioral mechanisms of fear learning are investigated using the fear-conditioning paradigm—perhaps the most well-established translational model of the acquisition of clinically relevant fear and anxiety (Beckers, Krypotos, Boddez, Eftting, & Kindt, 2013; Boddez, Baeyens, Hermans, & Beckers, 2014; Bouton, 2002; Craske, Hermans, & Vansteenwegen, 2006; Mineka & Zinbarg, 2006). Fear-conditioning involves an initially neutral stimulus (the conditioned stimulus or CS), such as a light or a tone, being repeatedly paired with an aversive unconditioned stimulus (US). After only a few CS-US pairings, presentations of the CS alone will elicit a
conditioned fear response (CR) assessed in humans using measures of physiological arousal, response tendencies (e.g., freezing), or expectancy ratings. This basic paradigm has proven enormously successful in advancing knowledge of the neurobehavioral mechanisms involved in the acquisition, expression, and regulation of fear specifically and emotional behavior more generally (Craske et al., 2006; LeDoux, 2014; Vervliet & Raes, 2013). As an illustrative example, consider an individual diagnosed with panic disorder defined by recurrent and unexpected panic attacks, heightened emotional arousal and an intense fear of future or anticipated attacks. Such individuals may have previously experienced an attack (the US) in a public place (the CS) and, as a result, come to experience derivatives of these conditioned fear responses and avoid situations associated with panic attacks.

Fear responses are considered excessive or indicative of psychopathology when the continuum of intensity increases; that is, when the CR occurs at such a high level and with escalating frequency that ongoing behavior is impaired and results in avoidance of the fearful stimuli. The underlying neurobehavioral mechanisms supporting this transition from adaptive fear learning to clinical fear, avoidance, and anxiety are now well understood thanks to the fear-conditioning paradigm (Beckers et al., 2013; Lissek et al., 2005; Vervliet & Raes, 2013). But fear learning is rarely (if ever) limited to those specific instances or events in which conditioning occurred. Instead, a range of related objects, situations, and events that resemble the aversive learning experience come to elicit fear as well. Generalization of conditioned fear shows that CRs are often elicited by stimuli not associated with the aversive event but which resemble the CS along a formal, perceptual dimension (Honig & Urcuioli, 1981). The study of stimulus generalization processes like this has a long history of research in both Pavlovian and instrumental/operant conditioning (e.g., Hull, 1943; McLaren & Mackintosh, 2002; Pavlov, 1927), but has only recently been extended to fear generalization in humans (Dunsmoor, Mitroff, & LaBar, 2009; Lissek et al., 2008, 2010; Vervliet, Vansteenwegen, Baeyens, Hermans, & Eelen, 2005; Vervliet, Vansteenwegen, & Eelen, 2004).

The relevance of fear generalization to experimental psychopathology research on fear and anxiety is illustrated by returning to our example of panic disorder. An individual who experiences a panic attack while riding a train may associate a host of information with the experience that is not limited to trains. For instance, he or she may express fear and avoidance when confronted with perceptually related cues or situations, such as all enclosed spaces (e.g., a bathroom stall, being stuck in traffic, or inside an MRI scanner), other forms of transportation (e.g., buses, trams, subways, cars), or situations in which mechanical sounds and movement are experienced (e.g., on an escalator). Neither is fear generalization confined to external stimuli: if the individual experienced a racing heartbeat when experiencing the panic attack, then other internal sensations like feeling a tight chest, a numb feeling in the arms, or spells of dizziness might each acquire the potential to evoke fear (Bouton, Mineka, & Barlow, 2001; Lissek et al., 2010). In addition to perceptually related stimuli, generalization might involve classes of symbolically related cues. For instance, a person suffering from obsessive-compulsive disorder might initially avoid chemicals like herbicides and pesticides for fear of contaminating their child. The set of perceived threats may gradually expand to include knives, syringes, dirty objects, and cigarette butts (Kaczkurkin & Lissek, 2013). Although none of these stimuli share strong perceptual relatedness, they are all part of an idiosyncratic category of cues of potential danger to the health of the sufferer’s young child for whom she experiences excessive responsibility.

Fear generalization is a defining feature of anxiety disorders such that the focus of fear becomes excessive and unrestricted, extending to a whole class of objects, persons, and situations (American Psychiatric Association, 2013; Craske et al., 2009, Lissek, 2012). This perceptual and nonperceptual generalization means that aversive experiences with one exemplar may lead people to infer that classes of related cues are fearful, dangerous, and need to be avoided, despite their physical differences. As a result, the clinically anxious individual comes to fear, and avoid, all potentially threatening objects and situations, leading to impairment in social functioning and diminished quality of life (Craske et al., 2009). This broadening of stimuli that come to elicit fear and anxiety can involve mechanisms of stimulus generalization as traditionally described in the conditioned learning literature, as detailed in this review.

In the present article, we review existing behavioral and neuroimaging empirical research on the perceptual and nonperceptual (conceptual and symbolic) generalization of fear and avoidance, and its relevance to anxiety disorders. We also evaluate the current status and future direction of research on fear and avoidance generalization and describe clinical and therapeutic implications. It is important to note that the present review is only concerned with research conducted with humans using the fear-conditioning and generalization paradigm. As a result, related nonhuman neurophysiology research
on generalization (e.g., Ciocchi et al., 2010; Kheirbek, Klemenhagen, Sahay, & Hen, 2012), human work on generalization of nonfear learning (e.g., Schechtman, Lauf, & Paz, 2010) or on the role of environmental factors in generalization (e.g., Pace-Shott et al., 2009) will not be considered here. Further, learning theory models of stimulus generalization and discrimination constitute a substantial literature that cannot be covered here due to space constraints (see, for example, McLaren & Mackintosh, 2002).

**Perceptual Fear Generalization**

In perceptual generalization, conditioned responses established to a sensory stimulus (CS+; e.g., a specific light, sound, or shape) are elicited to some degree by other stimuli tested in a similar way. Early animal demonstrations of what we would today regard as fear generalization involved an appetitive (i.e., rewarding) original CS. Many early studies of stimulus generalization incorporated an appetitive (i.e., rewarding) US, like food, but an early animal demonstration of fear generalization along perceptual similarity was by Hovland (1937). First, in acquisition, one of four equidistant tones was contingently paired with an electric shock presented to the left wrist. Then, in the generalization test, the four tones were presented again and skin conductance responses (SCRs) were recorded. In line with Pavlov’s observation of generalization in an appetitive preparation, Hovland observed that the three tones that were never contingently paired with the US, but were perceptually similar to the CS, also elicited fear responses. Moreover, the responses to these generalization stimuli decreased as a function of perceptual similarity (higher or lower in tone frequency). This response decrement is visualized as a logarithmic curve and is now widely known as a generalization gradient. The generalization response gradient is generally centered at the reinforced stimulus (CS+) and diminishes as a function of perceptual similarity to the CS+ along the sensory dimension (Ghirlanda & Enquist, 2003; Honig & Urcuioli, 1981).

Hovland (1937, p. 136) first noted that the form of the gradient is an important indicator of learning and generalization, with a broad gradient taken to reflect widespread generalization or poor stimulus control, whereas a sharp gradient centered on the CS+ can be taken to reflect discrimination or strong stimulus control. As we shall see, the shape of the gradient can also provide individual difference measures (e.g., clinical status of the participants) and reflect the impact of different experimental manipulations (e.g., verbal instructions).

**Empirical Demonstrations of Perceptual Generalization Gradients**

Contemporary generalization research in humans has been inspired by conditioning-based models of clinical anxiety (e.g., Lissek et al., 2008, Vervliet, Kindt, Vansteenwegen, & Hermans, 2010b). In the first contemporary systematic examination of fear generalization gradients in humans, Lissek et al. (2008) showed gradients of the eye-blink startle reflex to rings that varied in size from a ring paired with shock (CS+). The eye-blink (or fear-potentiated) startle reflex is an electromyographic response to a sudden stimulus that is potentiated in the presence of threat or during a state of anxiety, relative to a quiescent state (Grillon, 2008). During the acquisition phase, Lissek et al. (2008) used a small and a large sized ring as CS+ and CS−, respectively. An electric shock US followed the CS+ on 9 of the 12 trials (75% reinforcement schedule), while the CS− was never followed by shock (rings serving as CS+ and CS− were counterbalanced across participants). To test whether conditioned fear generalized from the CS+ to other perceptual stimuli, eight different unreinforced rings, ranging in size between the CS+ and CS−, were used as generalization stimuli (GSs). During the generalization test, startle responses to the GSs were tested. Fear responses were observed for stimuli that visually approximated the CS+, and the strength of the fear responses gradually decreased as the GSs became more dissimilar from the CS+, producing a decremented generalization gradient of conditioned fear responses.

In a follow-up study, Lissek et al. (2010) demonstrated that patients suffering from panic disorder (PD) displayed shallower, linear decreases in generalization than the healthy control group. Lissek et al. noted the generalization of conditioned fear of up to three increments of CS+ differentiation in PD patients, but by only one increment in controls. This pattern of responding, referred to as overgeneralization, is defined by less steep decreases in fear responses as stimuli differentiate from CS+, indicating stronger generalization and the tendency to infer threat-potential to physically similar stimuli. Overgeneralization has been observed in several clinical groups including PD (Lissek et al., 2010), generalized anxiety disorder (GAD; Lissek, Kaczkurkin, et al.,
2014), and posttraumatic stress disorder (Lissek & Grillon, 2012).

In a recent study, this fear generalization test was embedded in an instrumental computer game ("virtual farmer") where shocks could be avoided at the cost of poorer performance in the game (van Meurs, Wiggert, Wicker, & Lissek, 2014). The extent of generalization found with a measure of Pavlovian conditioning (fear-potentiated startle) correlated with generalization of the instrumental response (avoidance). To our knowledge, the only other study to have examined perceptual generalization of avoidance, as an instantiation of fear, is a study by Lommen, Engelhard, and van den Hout (2010). During acquisition, these authors used a white and a black colored circle as CS+ and CS-, respectively, of which the first was always terminated by shock. During the generalization test, circles with grey values that ranged between black and white were presented as GSs. However, before the start of this phase, participants were instructed that shocks could be avoided within a latency of 1 or 5 s. Results showed that a preselected group of participants who scored high on neuroticism avoided shocks during this phase but only on the 5 s latency trials compared to the group scoring low on neuroticism. Given that neuroticism is a risk factor for clinical anxiety, these results are in line with Lissek et al.'s (2010) demonstration of overgeneralization in PD.

Whereas most fear generalization studies have employed aversive shock USs, two recent studies used an aversive picture of a fearful screaming face as the US, with the CS+ and CS- consisting of pictures of neutral female faces. Haddad, Xu, Raeder, and Lau (2013) predicted that this CS–US combination might increase belongingness within the CS–US relation and facilitate fear learning and generalization. This "screaming lady" paradigm was tested with a group of healthy young adults. After acquisition, generalization stimuli were presented, which consisted of eight morphed pictures ranging in perceptual similarity from the CS+ to the CS-. Significant generalization was observed in fear ratings as well as fear-potentiated startle (see also Haddad, Pritchett, Lissek, & Lau, 2012). Even though the combination of a face displaying a fearful expression presented simultaneously with a 95 dB female scream might be less aversive than electric shock, this US has the advantage that it might be more suitable for testing younger age groups.

Glenn et al. (2012) tested healthy 8- to 13-year-olds with a variant of this paradigm with the only difference being that only one GS was used, which was a 50% morphed blend of the CS+ and CS- faces. Interestingly, there was a significant correlation between age and the extent of fear acquisition and generalization. Older children showed stronger acquisition effects in the fear potentiated startle response and exhibited more extensive generalization in this measure of fear. These results are important in light of fear generalization as a potential developmental or vulnerability risk factor for anxiety disorders and highlight the transition from childhood to early adolescence as a known high-risk period for both anxiety and mood disorders (Costello, Egger, & Angold, 2005).

Fear generalization has also recently been studied in the context of a voluntary movement-conditioning paradigm in which one movement was followed by shock, while a second movement was not (Meulders & Vlaeyen, 2013). Subsequently, participants were required to make novel movements that were either proprioceptively related or unrelated to the original "painful" movement CS. It was found that fear significantly generalized over this proprioceptive dimension (see also, Meulders, Vandebroek, Vervliet, & Vlaeyen, 2013).

Neural Substrates of the Perceptual Generalization Gradient

Neurobiological models of stimulus generalization can be traced back to Pavlov’s neural irradiation theory, which conceptualized generalization as a result of spreading neurophysiological activity across the cortex. Research on the neurobiology of fear generalization in laboratory animals has remained somewhat limited over the past century, especially compared to research on direct forms of fear conditioning. To date, the number of human neuroimaging investigations of fear generalization is small, but interest in the neurobehavioral mechanisms underlying the generalization of conditioned fear in humans has started to generate considerable attention.

The first fMRI study of fear generalization in humans incorporated a perceptual dimension of emotional faces and demonstrated generalized increases in activity to stimuli approximating the CS+ within regions implicated in the acquisition and expression of conditioned fear, including the insula, caudate, and thalamus (Dunsmoor, Prince, et al., 2011). The ventromedial prefrontal cortex (vmPFC), in contrast, showed generalized activity to a stimulus approximating the CS- (see the section on Intensity Generalization below for further details). Greenberg et al. (2013a) examined fear generalization using a perceptual dimension of geometric shapes found generalized activity to stimuli resembling the CS+ in the insula, caudate, anterior cingulate cortex, and supplementary motor cortex. Heightened activity in the vmPFC...
was observed to generalize test stimuli resembling the CS-

Greenberg et al. (2013b) investigated their fear generalization paradigm with a group of patients with generalized anxiety disorders (GAD). Functional connectivity analyses focused on coupling with the anterior insular (aINS), as the insula is generally implicated in anxiety-specific neurobiological deficits, including overexpression of fMRI activity to learned threats or intrinsically anxiety-provoking stimuli (Etkin & Wager, 2007; Paulus & Stein, 2006). Connectivity analyses revealed increased right aINS coupling with the right posterior insula, ACC, amygdala, and supplementary motor area, in line with the idea that the aINS is a hub where somatosensory and visceral input is integrated and relayed to other areas that guide fear-related behaviors (ACC, amygdala, SMA).

Lissek, Bradford, et al. (2014) investigated the brain mechanisms of fear generalization using a version of the "rings of varying size" paradigm. Results replicated previous fMRI investigations of fear generalization in humans (Dunsmoor, Prince, et al., 2011; Greenberg et al. 2013a,b), revealing a ("positive") gradient in bilateral insula to stimuli resembling the CS+, and a ("negative") gradient in vmPFC to stimuli more closely approximating the CS-. Positive gradients were also found in the dorsomedial prefrontal cortex and bilateral inferior parietal lobule, while negative gradients were found in bilateral hippocampus and precuneus. Functional connectivity analyses, using left and right ventral hippocampus as seed regions, revealed increased coupling with the amygdala and insula during presentation of stimuli resembling the CS+, and increased coupling with the vmPFC and the precuneus during presentation of stimuli with the least resemblance to the CS+.

Collectively, these findings suggest that fear generalization engages similar neurocircuitry implicated in the acquisition/expression (e.g., the insula and ACC) and regulation (e.g., the vmPFC) of conditioned fear. These results are generally in line with a neurobiological model of fear generalization outlined by Lissek (2012) and Lissek, Bradford, et al. (2014) (see Figure 1) that incorporates a dual-pathway hypothesis of fear learning and expression centered on the amygdala, proposed by LeDoux (1996). In this view, potentially threat-relevant generalization stimuli that closely approximate the CS+ may be relayed directly from the sensory thalamus to the amygdala, bypassing sensory cortex and thereby rapidly initiating the expression of a conditioned fear response through output connections with the insula, brainstem, and other areas involved in the psychological and physiological expression of fear. Fear generalization is also mediated by the hippocampal complex, which plays a role in pattern completion/pattern separation processes (see Gluck & Myers, 1993). Given sufficient overlap between a generalization stimulus and the CS+, the hippocampus initiates a pattern completion process that reactivates the neural representation of the CS, thereby initiating a conditioned response. Insufficient overlap, on the other hand, initiates pattern separation processes in the hippocampus, which leads to activation of the vmPFC that in turn down regulates the amygdala. Hence, fear generalization can be viewed as balancing excitation versus inhibition, which is determined by pattern completion (generalization) versus pattern separation (inhibition) processes. Evidence from existing fMRI studies lends some support for this fear generalization model, but the precise role of the amygdala, hippocampal complex, medial prefrontal cortex, and other regions in mediating fear generalization versus inhibition in humans requires further investigation. In particular, the role of the human amygdala in fear generalization remains unclear, as previous fMRI studies have not detected robust amygdala activity to GSs during generalization testing. Clearly, further neuroimaging work on generalization is warranted.

Experimental Manipulations of Fear Generalization Along Perceptual Similarity

In search of mechanisms that drive perceptual generalization, behavioral research has started to investigate the impact of experimental manipulations on generalization. For instance, Vervliet, Kindt, Vansteenwegen, and Hermans (2010a) studied the impact of prior nonfearful experiences. Participants were presented with a picture of a triangle and a picture of a parallelogram, one of which was contingently followed by shock. During the generalization test, two generalization stimuli were presented, which were sharper versions of the triangle and parallelogram that were presented during acquisition. Skin conductance responses and shock expectancy ratings were obtained during both phases. Crucial to this experiment, one group was preexposed to the generalization stimuli (without shock) prior to acquisition, whereas the control group was preexposed to two irrelevant stimuli (arrow and circle). The results showed significantly less generalization in the preexposure group, suggesting that prior nonfearful experiences can protect against fear generalization. This finding has relevance with respect to the prevention of overgeneralization.
Vervliet, Kindt, Vansteenwegen, and Hermans (2010b) designed a procedure to test the hypothesis that the extent to which the GS shares features common with the CS will determine the level of generalization observed. All participants were conditioned to a yellow triangle, and were subsequently presented with two generalization stimuli: a blue triangle and a yellow square. The first GS only shared the perceptual feature of shape (triangle), the second the perceptual feature of color (yellow). Skin conductance and US-expectancy were recorded throughout the experiment. The crucial manipulation in this study was that prior to conditioning participants were either informed that the shape of the stimuli would be relevant in this experiment (SHAPE relevant group) or that the color would be informative (COLOR relevant group). As predicted, results showed that generalization was strongest for the same shape stimulus in the SHAPE group, and strongest for the same color stimulus in the COLOR group.
A behavioral manipulation that investigates the role of stimulus features can be found in Vervliet and Geens (2014). Again, participants were conditioned to a yellow triangle (CS+) and generalization was tested for two GSs: a blue triangle and a yellow square. Crucially, the nature of the CS- was manipulated here. Depending on the nature of that stimulus, either shape or color became essential for discrimination learning during acquisition. For one group (COLOR relevant), the CS- was a red triangle, whereas for the second group (SHAPE relevant) the CS- was a yellow circle. Results of online shock expectancies and skin conductance responding showed that this manipulation significantly impacted generalization. Stronger fear responses were observed for the yellow circle GS in the COLOR-relevant group, while more fear was observed for the blue triangle GS in the SHAPE-relevant group. Again, these results demonstrate that fear generalization is influenced by feature learning.

Another intriguing finding about how the nature of the CS- can impact the shape of the generalization curve was demonstrated by Dunsmoor and LaBar (2013). In this experiment, the CS+ was a circle of which the color was in the middle of the green-blue color spectrum (near the point of subjective equality between blue and green). For one group of participants, the CS- was a clearly green circle, while for the other group this was a clearly blue circle (both were endpoints of the blue-green dimension). Following discriminatory fear conditioning between the CS+ and CS-, participants were exposed to a series of unreinforced circles that varied in color on the spectral wavelengths below and above the CS+ (more green and more blue). Opposing asymmetric generalization gradients were observed for the two groups that skewed in a direction opposite the CS-. Postdiscrimination shifts in the generalization gradient away from the CS+ is known as a gradient shift (or peak-shift effect), and has been observed in a number of laboratory and ethology studies of animal behavior (ten Cate & Rowe, 2007). Peak-shift effects can explain response biases for generalized stimuli that have never been reinforced, but contain features or stimulus elements that are more extreme or unlike an unreinforced stimulus. Peak-shift effects reveal that individuals not only learn about the CS+ during conditioning but also learn about the CS+ relative to the CS-, inducing a postdiscrimination shift in behavioral responses to stimuli that are more unlike the CS-. Together with the studies by Vervliet et al. (2010a, 2010b, 2014), these findings demonstrate how experimental manipulations can impact the generalization curve, and help unveil processes underlying perceptual fear generalization.

Individual Differences in Fear Generalization across Perceptual Similarity

A major reason for the renewed interest in conditioned fear generalization is its putative role in the etiology and maintenance of anxiety disorders. It has been proposed as a pathogenic marker of anxiety disorders, with deviant behavioral, physiological, and neural correlates (Lissek et al., 2008, 2010; Lissek, Bradford, et al., 2014; Lissek, Kaczkurkin, et al., 2014). Various studies have now shown that fear gradients vary with individual difference variables that are relevant to anxiety.

Comparing Anxiety Patients and Healthy Controls

As described earlier, an important demonstration of the clinical utility of the generalization gradient revealed that PD patients display a broader gradient of conditioned fear than healthy controls (Lissek et al. 2010). Broad generalization gradients have also been shown in GAD patients (Lissek, Kaczkurkin, et al., 2014), thereby strengthening overgeneralization as a pathogenic marker and highlighting overgeneralization as a putative transdiagnostic process cutting across various anxiety disorders.

A series of fMRI studies have compared the fear gradient between GAD patients and healthy individuals. Greenberg et al. (2013b) found a flatter slope of negative gradients in vmPFC and somatosensory activity in the GAD group, but no group differences in positive gradients in insular, ACC, SMA and caudate activity. These results suggest that fear overgeneralization results from a deficit in fear regulation during interactions of fear excitation and fear inhibition circuits along a threat-safe continuum. This conclusion was further supported by the observation that slope coefficients of patients’ individual vmPFC gradients were positively correlated with trait anxiety and depressive symptoms.

Cha et al. (2014a) found that vmPFC thickness, vmPFC functional connectivity and vmPFC structural connectivity within the corticolimbic systems predicted individual variability of vmPFC threat assessment in an independent fashion. Cha et al. (2014b) further extended these findings by comparing patients with GAD and healthy controls and found a role for the human ventral tegmental area (VTA) in the overgeneralization of fear: GAD patients showed heightened and less discriminating VTA reactivity to GS, as well as increased mesocortical and decreased mesohioocampal coupling. These findings suggest that abnormalities in the VTA and mesocorticolimbic systems may contribute to maladaptive threat processing in clinical anxiety.

While classical conditioning provides a valuable model for the etiology of anxiety disorders (e.g., Field,
Genetic Factors

Genetic influences on variability in conditioned fear generalization have recently been addressed in human genotyping research. The focus so far has been on the involvement of the brain-derived neurotrophic factor (BDNF). This factor is implicated in hippocampal-dependent learning processes, and carriers of the BDNF MET allele of the Val66Met BDNF genotype are characterized by reduced hippocampal structure and function (e.g., Hariri et al., 2003). BDNF has also been implicated in amygdala-dependent fear conditioning in rodents (Monfils et al., 2009). Moreover, variations in this polymorphism have been associated with anxiety-related traits in humans (Montag et al., 2010). BDNF is therefore a likely candidate to play a role in fear conditioning and generalization as well.

Hajcak and colleagues (Hajcak et al., 2009) investigated the influence of BDNF on conditioned fear generalization in humans. Fifty-seven participants were genotyped and classified as carriers of the Val/Val BDNF allele (n = 44; 25 female), carriers of the Val/Met allele (n = 4; 3 female) or carriers of the Met/Met allele (n = 3; 1 female). As noted above, Met allele carriers are characterized by relatively poor memory functioning as well as reduced hippocampal functioning during memory tasks (Hariri et al., 2003). Val/Val-carrying groups displayed fear-potentiation of their startle reflexes to the CS+ and generalized potentiation to GSs along the perceptual dimension. Met-carriers, on the other hand, had deficient fear-potentiated startle to the CS+, indicating impaired fear conditioning (see also Lonsdorf et al., 2010). At the same time, however, they showed a fear-potentiated startle reflex to the closest generalization stimulus. A deficient ability to elicit defensive responses to appropriate stimuli in Met-carriers may contribute to patterns of overgeneralization. However, Hajcak et al.'s (2009) findings have yet to be replicated with a larger sample (Greg Hajcak Proudfit, personal communication, July 27, 2014).

Torrents-Rodas et al. (2012) included higher numbers of Met-allele carriers and indexed generalization using the uninstructed “rings of varying size” procedure developed by Lissek et al. (2008). Genotyping a group of 141 participants revealed 50 Met-allele carriers versus 91 carriers of the Val/Val- or Val/Met-genotype. Unlike the Hajcak et al. study, however, the results did not reveal significant differences between these groups in terms of the acquisition or generalization of conditioned fear (potentiated startle reflex, SCR, risk ratings). These results are in line with another study failing to show fear acquisition deficits in Met-allele carriers (Soliman et al., 2010), although that study did reveal delayed fear extinction effects. Taken together, the currently available studies show highly variable effects of the BDNF polymorphism on fear acquisition and generalization in humans. Further work on the role of genetic factors in fear conditioning and generalization is needed before any strong conclusions can be drawn from these small sample sizes.

Personality Traits

There is evidence that personality traits, as measured via questionnaires, modulate fear-conditioning processes (Joos, Vansteenwegen, & Hermans, 2014). The strongest emphasis is on trait-anxiety, the level of general anxiousness of an individual, which is typically assessed with the trait scale of the State-Trait Anxiety Inventory (STAI; Spielberger, 1983). For example, high levels of STAI-T have been associated with increased amygdala responsivity during fear conditioning, and impoverished prefrontal cortical recruitment during conditioned fear extinction (Indovina et al., 2011; but see Dunsmoor, Ahs, & LaBar, 2011).

As described earlier, Haddad et al. (2012) investigated fear conditioning in a specific procedure where a face picture served as CS+ and an angry face plus a scream as US, and a similar face (CS1-) as well as a grey oval (CS2-) that were never paired with the US. Fearfulness ratings, potentiated startle reflexes and skin conductance reactivity showed the expected fear gradient from the CS+ over CS1- to CS2-. Interestingly, this gradient was not immediately present but developed over blocks of conditioning trials (decreasing to CS1- and CS2-). Moreover, the startle reflex data also revealed faster development of the falling gradient in participants with lower levels of STAI-T (based on median-split analysis). Participants with higher levels of STAI-T, on the other hand, displayed persistent startle responses to CS1- (the similar safety stimulus), but not to CS2- (the dissimilar safety stimulus). Eventually, startle responses to CS1- also dropped in this group, reaching similar levels of differential fear responding as the low-STAI-T group. Together, these data suggest a
weakened ability to discriminate safety cues from threat cues in individuals with higher levels of STAI-T, probably driven by an overgeneralization of fear from CS+ to CS1-.

Torrents-Rodas et al. (2013) investigated the effect of trait-anxiety level on the conditioning and generalization of fear. They first assessed the STAI-T in 992 undergraduate students and selected 126 participants based on their scores, in order to create a low-, middle-, and high-anxiety group. They replicated the basic fear conditioning and generalization patterns typically found (e.g., Lissek et al., 2008), but found no robust influence of STAI-T levels on the various fear measures (risk ratings, fear-potentiated startle, skin conductance reactivity). Risk ratings revealed a significant interaction between anxiety group and stimulus, which was due to elevated risk ratings to the CS- and adjacent generalization stimuli, suggesting impaired safety learning and generalization. Torrents-Rodas et al. (2014) examined the temporal stability of this generalization effect across an 8-month interval between sessions and concluded that generalization did not differ across sessions and that a significant proportion of individual differences in these processes were stable over time.

Kaczkurkin and Lissek (2013) focused on the potential relation between obsessive-compulsive traits and overgeneralization of fear. They pre-screened a large sample of undergraduates with the Obsessive Compulsive Inventory–Revised (OCI-R) and selected a group of 28 high scorers and a group of 31 low scorers for participation. Kaczkurkin and Lissek also measured the Threat Estimation scale of the Obsessive Beliefs Questionnaire, as they hypothesized specifically a relation between high threat estimation and fear overgeneralization. Potentiated startle reflexes and risk ratings replicated the basic fear gradient (Lissek et al., 2008). Although there was no effect of high versus low obsessive-compulsive traits (OCI-R), Threat Estimation did modulate the fear gradient. The high Threat Estimation group showed greater fear-potentiated startle to ring sizes up to two units of dissimilarity from the CS+ while the low Threat Estimation group did not generalize the conditioned fear response beyond the CS+. No group differences were found in the risk ratings. This study provides the first indication that fear generalization may play a role in obsessive-compulsive disorder as well.

Nonperceptual-Based Fear Generalization

Whereas perceptual-based fear generalization is determined by the physical similarity between stimuli, humans (in particular) are adept at generalizing from past experience based on regularities that go beyond physical resemblance. How nonperceptual forms of generalization integrate with conditioned fear learning processes is an emerging area of investigation with implications for understanding broad overgeneralization of fear characteristic of anxiety disorders. Here, we review the existing research on nonperceptual (intensity based) fear generalization, associative forms of generalization, such as sensory preconditioning, category based and semantic fear generalization.

Intensity Generalization

Much of the stimulus generalization literature focuses on similarity that can be determined by distance between increments along a perceptual continuum. Classic studies of stimulus generalization in pigeons, for instance, revealed graded responses to unreinforced colors as a function of similarity to the CS+ along the color spectrum (Guttman & Kalish, 1956). The shape of the generalization gradient along neutral sensory dimensions like color, tone frequency, or shape tends to be symmetrical and peak at the trained stimulus (CS+). In contrast, the shape of the generalization gradient along dimensions that vary in physical intensity—like brightness or volume—is often asymmetrical and skewed towards generalized stimuli of high intensity (Ghirlanda & Enquist, 2003). Although generalization along intensity dimensions has received far less attention than along neutral sensory dimensions, the bias to generalize conditioned responses to stimuli of high intensity may provide an especially useful account for behavior in the threat domain. Take, for example, a terrifying encounter with a particular dog. While such an experience could lead to a general fear of all dogs, it is unlikely that every dog encountered after the negative experience would elicit the same degree of fear; to an individual with a generalized fear of dogs, a Chihuahua is probably less frightening than a Rottweiler. Likewise, for a combat veteran who experienced a roadside bombing, the startle reaction to sudden noises in a safe context may be exaggerated to extremely loud sounds. The intensity factor of a generalized stimulus can be a large determinant of fear expression, and stimuli of high intensity may elicit even more fear than the actual threat stimulus.

The contribution between perceptual similarity and emotional intensity on fear generalization was examined by Dunsmoor et al. (2009) using a stimulus dimension of morph increments of the same facial identity but displaying varying intensity of emotional expression. A face with a high degree of fearful expression intensity and a face with a neutral expression served as endpoints along a fear expression intensity continuum. The intermediate morph
between fear and neutral expression intensity (a moderately intense fearful expression) served as the CS+ paired with an electric shock. Prior to conditioning, there were no baseline differences in SCRs as a function of the emotional expression intensity of the face (see Figure 2, white circles). A generalization test conducted after the intermediate value was paired with shock, however, revealed a linear gradient of SCRs with a strong response bias to generalized morphed cues of higher emotional expression intensity (see Figure 2, black circles). Thus, unlike fear generalization along nonintensity dimensions like color or shape (e.g., Dunsmoor & LaBar, 2013; Lissek et al., 2008), conditioned fear responses along a dimension of emotional intensity are asymmetrical and characterized by a shift in maximal responses to generalized stimuli of high intensity. Put another way, stimulus intensity may in some cases control fear generalization more than perceptual similarity.

In a follow-up fMRI investigation, intensity biases were revealed in brain regions implicated in the acquisition of conditioned fear (CS+ versus CS-) including the insula, thalamus, and striatum (Dunsmoor, Prince, et al., 2011). That is, activity in these regions was enhanced to generalized stimuli of high intensity after, but not before, fear conditioning to a CS+ of intermediate intensity. Functional connectivity between the amygdala and a face-selective region in the fusiform gyrus was also enhanced in response to a generalized stimulus of higher intensity than the CS+, and a significant majority of subjects misidentified this high-intensity stimulus as the CS+ postexperimentally. This finding suggests that implicit response biases may occur through enhanced perceptual processing of high-intensity approximations of a learned threat following fear learning, which may in turn lead to an inflated estimate of the threat stimulus following the aversive experience. These response biases may be an important characterization of fear expression in anxiety disorders, particularly specific phobias, wherein fear behaviors are elevated in response to stimuli with exaggerated threat features and perception of feared stimuli is biased. For instance, spider phobics show a cognitive bias to misjudge a

FIGURE 2 Mean skin conductance responses (SCR) for preconditioning and the generalization test in Dunsmoor et al. (2009) showing undifferentiated SCRs to the stimuli before fear conditioning, but an asymmetrical linear generalization gradient emerges during the generalization test. Reprinted by permission.
previously encountered live spider as larger than it actually was (Vasey et al., 2012) and height phobics peering over a ledge overestimate height (Teachman, Stefanucci, Clerkin, Cody, & Proffitt, 2008).

**Sensory Preconditioning**

Fear generalization following an aversive experience is in some cases based on an association established before the experience occurred. In sensory preconditioning (SPC), for example, a prior association between neutral stimuli (e.g., a tone and a light) in the absence of aversive reinforcement enables the transfer of a conditioned response after just one of the stimuli (e.g., the light) undergoes direct conditioning with the US. Higher-order conditioning procedures such as SPC (see also second order conditioning; Gewirtz & Davis, 2000) increase the explanatory power of conditioning models of fear and anxiety to describe how stimuli indirectly related to the conditioning experience acquire the ability to evoke a threat response (Declercq & De Houwer, 2009).

While much of the empirical research using SPC is in nonhumans (Gewirtz & Davis, 2000), this procedure has been used in a limited number of human psychophysiological conditioning experiments (e.g., White & Davey, 1989). An important question raised by SPC is whether the strength of the initial stimulus-stimulus association developed during preconditioning, in the absence of reinforcement, determines the strength of subsequent generalization after one of the stimuli is reinforced. One way to address this question is to diminish the stimulus-stimulus association by repeatedly presenting one of the stimuli alone (i.e., extinction). Taking this approach, Vansteenwegen et al. (2000) showed that the transfer of conditioned fear between a stimulus pair is reduced if one item from the pair is repeatedly presented alone prior to fear conditioning. This finding suggests that SPC relies on an intact stimulus-stimulus association to facilitate the transfer of conditioned learning.

The nature of the pre-association has also been shown to enhance the effects of SPC. Dunsmoor, White et al. (2011) showed that a conceptually related stimulus pair (e.g., pictures of a spider and a spider web)—as compared to a conceptually unrelated stimulus pair (e.g., a spider and a waste barrel)—promoted the transfer of conditioned SCRs after one stimulus was paired with shock. In this case, prior conceptual knowledge on the association between spiders and webs likely strengthened the interstimulus pre-association. As a consequence of this conceptual pre-association, conditioned learning more easily transferred from the conditioned stimulus to the pre-associated stimulus. Another group of subjects who learned a pre-association between two pairs of conceptually mismatched cues (e.g., a spider paired with a wasp nest; a wasp paired with a spider web) showed generalization to both pre-associated stimuli (i.e., the spider and the wasp) after only one picture was paired with shock (e.g., only the web). In this group, generalization likely occurred via two channels: associative learning between conceptually mismatched cues acquired de novo (e.g., wasp and web), and prior conceptual knowledge of the association between cues that traditionally co-occur (e.g., spider and web). Together, these findings show that the SPC paradigm can be used experimentally to induce a transfer of conditioned fear in humans, and that the strength and nature of the prior association can determine the magnitude of fear generalization following an aversive learning experience.

It is notable that the literature on higher-order forms of fear conditioning in humans is scant overall. The lack of empirical research in this domain is somewhat surprising given the general interest in conditioning-based models of clinical anxiety and the fact that these models frequently describe overgeneralized fear expression in clinical anxiety as arising from higher-order conditioning processes. Foa, Steketee, and Rothbaum (1989), for example, invoke such processes to explain combat trauma: “Via the processes of higher order conditioning and stimulus generalization, a panoply of situations acquire fear-inducing capacities.... Thus, sudden loud noises (a car backfiring, firecrackers exploding) may trigger emotional responses that were not present during the original trauma” (p. 157). Similarly, higher-order conditioning is invoked when describing generalization following rape trauma: “Sexual activities associated with the rape trauma become conditioned stimuli for anxiety. Via generalization and higher order conditioning, other sexual activities also come to elicit fear” (p. 159). Overall, further theoretical and empirical research using higher-order conditioning procedures is warranted.

**Category-Based Fear Generalization**

The ability to consider different objects as belonging to the same category allows us to infer the properties of an unknown object based on the known properties of categorically related objects. In the absence of any direct knowledge, we can predict with reasonable accuracy that, for example, a dog we have just encountered might growl and bark to protect its food or territory, demonstrate social behaviors, and respond positively to rewards. Emotionally meaningful experiences with particular stimuli can shape the nature of these predictions.
If our experience with dogs is overwhelmingly positive, then we may be inclined to expect a positive response from other dogs. If, on the other hand, we have had a frightening experience with a dog, then we may regard similar dogs as threatening. In extreme cases, a powerful fear of a particular object may generalize to the entire category (all dogs or similar animals), and even to items (collars or leashes) and to places (parks) associated with the feared object.

Theory and research on the psychology of categories and concepts is substantial (see Murphy, 2002). Much of this literature explores how humans (mainly) form equivalence classes among perceptual categories to make inductive generalizations (Reeve & Fields, 2001; Zentall, Galizio, & Critchfield, 2002). Categorization research cuts across a number of psychological disciplines, including development, language, memory, and decision making, but historically has rarely overlapped with conditioned learning. This is perhaps not surprising given that conditioning and stimulus generalization studies tend to incorporate relatively simple CSs that exist along well-defined perceptual dimensions in order to plot gradients of CRs as a function of physical similarity. However, real-world fear learning rarely occurs to simple sensory cues like lights or tones, and instead involves complex stimuli that can be represented at the category level and within an elaborate associative or semantic network. As a consequence, meaningful experiences with a real-world object have the potential to imbue a wide variety of other objects with significance through myriad processes of inductive generalization described in the categories and concepts literature.

As described throughout this review, it is critical to infer threat potential in the absence of direct knowledge, as potentially harmful stimuli often assume multiple forms and vary considerably from one encounter to the next. The ability to create mental concepts from perceptual categories allows us to expand the range of potentially harmful stimuli beyond perceptual boundaries. That is, we can incorporate our conceptual knowledge of known objects to reason about the potential threat of novel objects. The organization of category knowledge may have particular influences on how information regarding aversive events is generalized.

To explore the role of inductive generalization in human fear conditioning, Dunsmoor et al. (2012) constructed a novel trial-unique conditioning paradigm that incorporated basic level exemplars from two superordinate categories, animals and tools, as conditioned stimuli. For one group of subjects, images of animals predicted shock (CS+) while tools were safe (CS–), while another group received the reverse contingencies (animals were safe while tools predicted shock). Critically, each basic level image was presented only once and subjects were not instructed about the CS–US contingencies. Despite variations in physical form between exemplars, and in the absence of direct knowledge about each stimulus (whether or not it predicted shock), it was expected that subjects would infer threat potential in novel category exemplars based on experiences with related exemplars. Trial-by-trial ratings of shock expectancy and conditioned SCRs confirmed that subjects generalized learning to the category level, treating novel exemplars from the CS+ category as potential threats and novel exemplars from the CS– category as safe. Functional MRI results showed that activity in category-selective regions in the occipitotemporal cortex was modulated as a function of learning history (Dunsmoor, Kragel, Martin, & LaBar, 2013). That is, lateral fusiform gyrus (animal-preferring cortex) activity was enhanced in response to images of animals in subjects for whom animals predicted shock versus subjects for whom this category was safe; whereas activity in the medial fusiform gyrus (tool-preferring cortex) showed the reverse pattern of activity in these subjects (enhanced in subjects for whom tools predicted shock versus subjects for whom tools were safe). Activity in the amygdala, insula, and anterior cingulate cortex was also enhanced in response to novel exemplars from the threat category across both learning groups, suggesting that category conditioning relies on regions traditionally implicated in simpler forms of conditioning. Thus, fear conditioning allows us to associate specific items with an aversive outcome, while our ability to represent knowledge at the category level allows us to extend this newly learned property to categorically related exemplars despite considerable variations in physical form.

Another area of insight into the complexity of human fear and anxiety is category-based induction (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990), which investigates factors that promote the generalizability of information at the category level. Examples include arguments of the nature mice have sesamoid bones [premise] therefore horses have sesamoid bones [conclusion]. In this case, the subject weighs the likelihood that the conclusion follows from the premise, which is assumed to be true, in the absence of any strong preconceived association between the conclusion exemplar and the property (sesamoid bones). Research on categorical arguments shows that stimuli more representative (or typical) of their category provide better sources of generalization than stimuli that are not
representative (or atypical) of their category (Osherson et al., 1990). The example argument concerning mice and horses, for instance, would be considered strong because the premise exemplar is typical of its category (e.g., mammal). In other words, a property belonging to a representative category member is likely to belong to other members of that category as well. In contrast, categorical arguments are weakened if the premise argument contains an atypical exemplar (see Heit, 2000; Smith, Shafir, & Osherson, 1993). For example, the argument bats have sesamoid bones therefore mice have sesamoid bones is considered weaker than the argument mice have sesamoid bones therefore bats have sesamoid bones, even though the premise and conclusion exemplars are from the same category (mammals). This effect is referred to as premise-conclusion asymmetry, and provides evidence that typical exemplars provide better sources of generalization than atypical exemplars (Osherson, et al., 1990; Rips, 1975).

Dunsmoor and Murphy (2014) adopted the framework of category-based induction to investigate whether typicality effects apply to generalization of conditioned SCRs in humans. Two groups were initially trained on typical CSs (e.g., mammals including a horse, cow, and rabbit) paired with shock, or atypical CSs (e.g., mammals including an aardvark, armadillo, and otter) paired with shock. Following acquisition, generalization was then tested, without shock, to objects from the same category but of opposing typicality; subjects trained on typical mammals were tested to atypical mammals, whereas subjects trained on atypical mammals were tested to typical mammals. In the framework of a categorical argument, the premise could be conceptualized as a typical (or atypical) CS predicts shock while the conclusion would be that an atypical (or typical, respectively) CS also predicts shock. The strength of this argument is thus evaluated by the strength of the generalized SCR. While both groups acquired conditioned fear to equivalent levels, only the group trained on typical mammals generalized responses to novel members from the same category, providing evidence that typical CSs provide better sources of fear generalization than do atypical CSs.

As the details extracted from aversive experiences do not generate a number of premises that will determine future behavior, the extent to which inductive reasoning applies to the nature of human fear and anxiety merits further empirical attention. For example, an aversive event such as being mugged near your apartment may lead one to question the general safety of the neighborhood. If you were mugged while walking through your neighborhood in the middle of the afternoon, then the conclusion that it is unsafe to walk in your neighborhood at night is likely to be strengthened. This argument contains elements of premise typicality (you may frequently walk through your neighborhood during the day) and plausibility (muggings tend to occur at night rather than during the day). If, however, you were mugged while walking in your neighborhood late at night, then the conclusion that the neighborhood is unsafe in the middle of the afternoon may be weaker, since the premise is more atypical and the conclusion less plausible. We can speculate that extremely frightening or traumatic experiences likely override inductive or analytical judgments, rendering even atypical premises sufficient to induce widespread generalizations regarding the threat potential of other stimuli or situations.

**SEMANTIC FEAR GENERALIZATION**

Early studies of nonperceptual stimulus generalization in humans used semantic stimuli (words) (Branca, 1957; Maltzman, 1977; Razran, 1939). For instance, participants would be conditioned to a word (e.g., PLANT) by pairing the presentation of the word with a US, and then tested on semantically related words (e.g., STEM) or unrelated words (e.g., MUSIC) (Maltzman, Langdon, Pendery, & Wolff, 1977). In some cases, the conditioned SCR would generalize between semantically related, but not unrelated word pairs; but the results from semantic conditioning experiments tended to be noisy and were often observed only in subjects who could verbalize the experimental contingencies (see Maltzman et al., 1977). Much of the theoretical interest in semantic conditioning and generalization concerned the acquisition of language, and few contemporary studies of human fear conditioning incorporate purely semantic stimuli.

Early researchers had come to the conclusion that simple S-R mechanisms alone may not be sufficient to account for these complex generalization phenomena along nonformal continua, and they suggested a key role for “complex thought processes” in explaining these effects (e.g., Maltzman, Langdon, & Feeney, 1970). Several other researchers began to hone in language processes themselves as underlying this complex form of fear generalization (Eisen, 1954; Mednick & Wild, 1962; Mink, 1963). It is surprising that so few researchers have followed up on this interesting line of enquiry in recent years, and that no study of semantic generalization has included an instrumental response (i.e., avoidance) component.

A recent study revisited the near-forgotten semantic generalization paradigm (e.g., Eisen, 1954) in an
attempt to examine the generalization of fear and avoidance (see Feather, 1965, for a review). Boyle, Roche, Dymond, and Hermans (2014) first exposed participants to fear conditioning in which one word was designated CS+ (e.g., broth) by pairing it with the presentation of shock, and another word was designated CS- (e.g., assist) and never followed by shock. In the next phase, an avoidance response (a space bar press) was established for the CS+, which cancelled the presentation of the scheduled shock. In the generalization test phase, words semantically related to the CS+ (e.g., soup) and CS- (e.g., help) were presented in the absence of shock. Differences in levels of avoidance, SCR and US expectancy ratings were all significant across the CS+ and CS- stimuli, as well as across their semantically generalized counterparts. Importantly, Boyle et al. found that the three measures did not covary reliably. That is, higher levels of avoidance responding were not associated with higher SCRs but were associated with higher shock expectancies for the semantically related stimuli only. Moreover, rates of avoidance of the US were not correlated well with expectancies of the US. Despite these interesting covariances among the dependent measures, the generalized words reliably produced US expectancies, SCRs, and avoidance responses consistent with their semantic relation to the CS+ or CS-.

Nonperceptual Fear Generalization: The Role of Symbolic Generalization

In the work reviewed so far, the role of perceptual and nonperceptual features in the generalization of conditioned fear and avoidance has relied on preexperimentally acquired learning about representational features and intact conceptual knowledge structures to explain how potential threats come to take on multiple, generalized forms. For instance, category-based fear generalization draws on existing conceptual knowledge (e.g., of animals and tools) to modulate and explain how related object concepts come to elicit fear (Dunsmoor et al., 2013). There are several assumptions underlying research of this kind, such as an acknowledgment that modulation of generalized fear by viewing images of objects or categories activates similar representations in different participants and that the nature and extent of individual differences in learning history with object concepts may be overlooked (or is deemed irrelevant) in generating the observed effects. While these assumptions are not necessarily limitations, it may be salutary to highlight alternative, functionally oriented approaches to fear generalization that emphasize both prediction and influence over the phenomena of interest—goals met through synthesizing de novo conditioning with methods of establishing concept-like associations or relations between arbitrary stimuli.

According to this approach, a potentially promising means of further understanding fear generalization involves verbal relations, such as stimulus equivalence relations, as a way of preemptively establishing the relevant “concepts” (Dymond & Roche, 2009). Research on verbal relations has shown that when language-able humans are taught a series of interrelated discriminations involving physically dissimilar (arbitrary) stimuli, the stimuli involved often become related to each other in ways not explicitly trained (Sidman, 1994). To illustrate, if choosing Stimulus X in the presence of Stimulus A is taught (i.e., A-X), and choosing Stimulus Y in the presence of Stimulus A (i.e., A-Y) is also taught, it is likely that untrained relations will emerge between X and A, Y and A (“symmetry”), X and Y, and Y and X (“equivalence”), in the absence of any feedback. When this occurs, a stimulus equivalence relation is said to have formed among the relata (Dymond & Roche, 2013; Hayes & Hayes, 1992). The emergence of untrained verbal relations such as these may, it is claimed, help explain the patterns of indirect, nonperceptually based fear generalization often seen in anxiety disorders and which arise in the apparent absence of a conditioning history with the feared object or event (Dymond & Roche, 2009; Hermans & Baeyens, 2013; Hermans et al., 2013). Indeed, the challenge in accounting for complex patterns of generalized fear present in clinical anxiety means that the “the powerful human capacity for abstract representation creates special problems in translating the rules of stimulus generalization worked out in animals. Fear and avoidance spread in animals from one context to another based on simple sensory cues. In humans, this spread may be on the basis of complex feelings” (Marks, 1987, p. 234). This spread or symbolic generalization of fear along arbitrary features or stimulus relations may account for instances of fear and avoidance that occur when, for instance, snake-phobic clients hear the word “snake” and other unrelated stimuli, such as the word “reptile,” pictures of snakes, names of different types of snakes, a real snake, and places where snakes might be found, all of which may occasion fear and avoidance (Foa & Kozak, 1986; Lang, 1985). In effect, while the original fearful CS may have been directly conditioned, the generalization of fear and avoidance responses often occurs along verbal (i.e., symbolic) dimensions and can include a host of indirectly related, arbitrary stimuli.
symbolic fear generalization

Research on the symbolic generalization of fear is growing, with early demonstration work tending to employ single-case research designs with small sample sizes (e.g., Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Valverde, Luciano, & Barnes-Holmes, 2009). Recently, Vervoort, Vervliet, Bennett, and Baeyens (2014) adapted the procedures of Dougher et al. (1994) to examine symbolic generalization of fear and extinction using contemporary-between-groups design standards and statistical analyses. Vervoort et al. (2014) first trained and tested participants to form two, four-member stimulus equivalence relations consisting entirely of arbitrary line drawings (i.e., A1-B1-C1-D1 and A2-B2-C2-D2; labeled with alphanumerics for purposes of clarity—participants were not exposed to these labels). Next, one member of the first relational class (B1) was established as CS+ by pairing it with shock, and a member of the second class (B2) was established as CS-. Symbolic generalization was then tested with presentations of the indirectly related members of each class, in the absence of shock. As predicted, conditioned fear, measured via SCR and online shock-expectancy ratings, generalized to the other members of class 1 (C1, D1) but not to class 2 (C2, D2). Extinction of conditioned fear was also shown to generalize to the other, indirectly related members, while extinguishing a generalized member did not reduce fear of the original CS+. Vervoort et al.’s findings are significant because they demonstrate the generalization and extinction of conditioned fear responses to stimuli symbolically associated with the CS+ via relational classes established intra-experimentally using unfamiliar, arbitrary stimuli (Dymond & Roche, 2009; Hermans et al., 2013).

symbolic avoidance generalization

Avoidance behavior is an instantiation of fear (Beckers et al., 2013) and plays a key role in the maintenance of anxiety (Barlow, 2002; Mineka, 1979). Avoidance may become debilitating and lead to impaired social functioning and, because of this, it is often the sole target of therapeutic change. While there is evidence for perceptual generalization of avoidance (Lommen et al., 2010; van Meurs et al., 2014), extending the analysis of symbolic generalization to include instances of avoidance is important in developing contemporary accounts of the emergence of clinical anxiety (Field, 2006; Friman, Hayes, & Wilson, 1998; Mineka & Zinbarg, 2006; Rachman, 1977).

The first supporting evidence for this approach comes from Augustson and Dougher (1997), who trained and tested participants for the formation of two, four-member stimulus equivalence relations (A1-B1-C1-D1 and A2-B2-C2-D2) and then used a differential fear conditioning procedure to establish B1 as CS+ and B2 as CS-, respectively. During a subsequent avoidance-training phase, shock could be avoided by completing a fixed-ratio (FR) 20-response requirement in the presence of B1, while shock was never scheduled to follow presentations of B2. Symbolic generalization of avoidance was then tested, in the absence of shock, with presentations of the indirectly related stimuli not present during avoidance training. Findings showed that all eight participants emitted the avoidance response to C1 and D1 but not to C2 and D2, indicating that they transferred the directly trained avoidance schedule response from B1 to the symbolically related C1 and D1 stimuli without further training.

Augustson and Dougher (1997) were the first to show the symbolic-based generalized avoidance of stimuli that had no direct relational history with aversive events. Since then, Dymond, Schlund, Roche, Whelan, Richards, and Davies (2011) have extended these findings with a larger sample size to a similar operant avoidance paradigm involving aversive images and sounds as USs, while Dymond, Schlund, Roche, De Houwer, and Freegard (2012) showed that levels of symbolic generalization resemble those seen when avoidance is acquired indirectly, such as through verbal instructions (Rachman, 1977). Dymond, Roche, Forsyth, Whelan, and Rhoden (2007, 2008) showed that the symbolic generalization of avoidance may be transformed in accordance with relations of "sameness" (i.e., equivalence) and "opposition." After training to establish two abstract shapes as contextual cues for same and opposite, respectively, the cues were presented with arbitrary nonsense stimuli and participants taught the following relations: same cue-A1-B1, same cue-A1-C1, opposite cue-A1-B2, and opposite cue-A1-C2. These relations lead to the following untrained relations: B1-C1 are same, B2-C2 are same, B1-C2 are opposite, and B1-C1 are opposite. Fear and avoidance responses conditioned to one of these stimuli may then alter or transform the functions of other, untrained relations in terms of "same" and "opposite." Dymond et al. (2007, 2008) exposed participants to a signaled avoidance task, during which responding in the presence of the stimulus B1 cancelled a scheduled US presentation. Another stimulus, B2, was never followed by the US. Then, participants were tested with presentations of C1 and C2, in the absence of the US. Findings showed that consistent avoidance responses were made in the presence of C1 but not C2 (because C1 is the
same as B1, whereas C2 is the opposite), thus demonstrating the symbolic generalization of avoidance in accordance with complex relational networks of same/opposite (see also Bennett, Hermans, Dymond, Vervoort, & Baeyens, 2014; Gannon, Roche, Kanter, Forsyth, & Linehan, 2011; Roche, Kanter, Brown, Dymond, & Fogarty, 2008).

These studies illustrate how verbal relations of equivalence, sameness and opposition, established intra-experimentally, come to act as conduits by which fear and avoidance responses so readily generalize and go on to alter or transform the functions of other, indirect stimuli or situations. In so doing, the process of symbolic generalization is said to mimic the relational basis by which clinically relevant fear and anxiety comes to be acquired and maintained. It shows how semantic-like, symbolic categories may be readily formed in language-able humans and how “meaning” or, in this instance, conditioned fear and avoidance functions, alter and spread to networks of related stimuli. The relational basis of this view of psychopathology forms the basis of relational frame theory (RFT; Dymond & Roche, 2013; Dymond, Roche, & Bennett, 2013), a functional contextual account of human language and cognition that informs several third-wave behavior therapies such as acceptance and commitment therapy (Hayes, Levin, Plumb-Vilardaga, Villate, & Pistorello, 2013).

According to RFT, excessive avoidance occurs in the presence of a wide range of stimuli and situations based on the actual and, more often than not, inferred presence of the aversive event. It follows that groups that already differ on the basis of individual differences variables may show different levels of symbolic generalization of avoidance. Dymond, Schlund, Roche, and Whelan (2014) showed that high spider-fearful individuals exhibited greater levels of symbolic generalization of avoidance than low spider-fearful individuals and also met avoidance learning criteria quicker. These findings highlight the symbolic basis of differential levels of clinically relevant avoidance and support an RFT understanding of the acquisition of excessive avoidance.

Clinical Implications of Human Fear Generalization Research

Core differences between "normal/adaptive" fears and pathology such as phobias are often conceptualized in terms of fear intensity and avoidance. Pathological fear is considered to be more intense than normal fear, or at least more than the situation would reasonably require. From this perspective, dog phobia is not just being afraid of dogs, but being severely afraid, to the extent that it impairs normal functioning and prompts avoidance of dogs and of the possibility of encountering dogs. Even though intensity and avoidance might be important in discriminating both types of fear and the transition between the two, we believe that generalization is probably even more central in this respect.

Given the clinical importance of fear generalization, it is surprising that this phenomenon has, historically, received so little empirical attention. Nevertheless, the existing research already reveals elements that are of clinical importance. First, there are important implications with respect to assessment. All treatments of anxiety disorders require an in-depth assessment, which includes an extensive analysis of the crucial stimuli and their conditioning history. This provides direct input for treatment (e.g., exposure techniques, cognitive restructuring). One of the important conclusions of generalization research is that fear can be elicited by stimuli that were actually never involved in a conditioning experience. Moreover, research on symbolic generalization indicates that (a) fear can be elicited by stimuli that do not bear a perceptual relation to the original CSs, (b) fear can be elicited by stimuli that were never associated with the original US, and (c) under some circumstances fear elicited by generalization stimuli can be more intense than fear elicited by the original CSs (Dougher, Hamilton, Fink, & Harrington, 2007). This illustrates why assessment of relevant fear stimuli and their conditioning history is not always possible or fruitful: the pathway from current fear stimuli to the original conditioning events is simply too complex. Relying on obvious and salient features of fear stimuli, as reported by a client therefore, might not always be sufficient and a more in-depth and complex analysis might be necessary. In effect, the phenomenon of nonperceptual fear generalization helps make clear why conditioning experiences are not always easily traceable in the clinical context (Muris, Merckelbach, de Jong, & Ollendick, 2002; Poulton & Menzies, 2002).

A more complete understanding of nonperceptual generalization processes is required before we can draw conclusions regarding how this process should inform the use of traditional therapeutic procedures. For example, it might make a difference whether in exposure treatments we expose clients to a limited or broad array of fear stimuli. Research by Vervoort et al. (2014) indicates that extinction of the original CS spreads to the generalization stimuli, while the reverse is not necessarily true (cf. Roche et al., 2008). In the absence of concrete information regarding the original CS, use of a broad array of generalized stimuli may represent a next best option. However, Vervliet et al. (2005)
suggested that extinction with generalization stimuli might lead to more extensive return of fear, compared to extinction with the original CS, where it can be identified. These observations are relevant to clinical practice because exposure treatment is frequently conducted using generalization stimuli precisely because feared stimuli are unavailable (e.g., the actual dog involved in the biting incident) or inaccessible. Against this background, several of the authors on this review are beginning to investigate the conditions under which extinction using generalized fear stimuli might be more effective and less vulnerable to return of fear.

It is also worth considering fear generalization in the framework of the highly influential Foa and Kozak (1986) model of informational processing and fear networks. While not a direct model for stimulus generalization, per se, the notion that fear is represented within informational networks, or memory structures, is a dictum of many exposure therapies that could also explain overgeneralization in anxiety disorders. This model views fear in terms of a memory structure with interconnected informational and response elements (e.g., fight-or-flight), and information that partially overlaps with a memory within this network can be sufficient to activate a fear response. For example, a victim of an assault who later spots a person resembling their assailant will have an automatic fear response through activation in this network. In this way, the model describes fear generalization based on pattern completion, as detailed by more contemporary models of fear generalization (Lissek, Bradford, et al., 2014; Lissek, Kaczkurkin, et al., 2014).

Within the RFT tradition, one response to the seemingly autocatalytic nature of fear and avoidance generalization is to conceptualize the process of extinction as fundamentally limited for therapeutic purposes. Precisely because of the ubiquity of language and the constant social reinforcement of symbolic relations that support fear generalization (e.g., pairing of the word “poison” with “spider” in the vernacular), it may be inevitable that stimuli associated with fear and anxiety will be encountered at some point in the future. Consequently, the ACT approach (Hayes, Stroshal, & Wilson, 2011), which arose out of the relational frame perspective on human learning, and focuses on the problem of loosening the dominance of verbal relations over the generalization process itself. For example, one technique known as “defusion” (Masuda, Hayes, Sackett, & Twohig, 2004) teaches the client how to perform other instrumental responses in the presence of fear stimuli and to thereby broaden the response functions of fear stimuli rather than narrow them. In effect, the multiplicity of response functions that get established for fear stimuli in the therapeutic setting (e.g., paying attention to its color, what the word sounds like, what it rhymes with, the feelings it creates in the body, etc.) compete with the normally dominant fear and avoidance functions and reduce the probability of fear and avoidance emerging on each occasion. Moreover, clients are taught how to respond to other fear stimuli in the same way in the future (e.g., notice what the word sounds like, notice the object’s color, and so on), so the skill of defusing the literal and dominant meaning of stimuli (e.g., “a spider means I must run away”) itself becomes generalized. In this way the ACT approach parallels many of the mindfulness approaches that have emerged in recent years, but with the difference that is based on a behavioral approach to understanding how verbal processes underlie the generalization phenomenon and how undermining fear generalization in the long term will have to involve undermining the negative effects of verbal processes themselves. Put simply, generalization through exposure narrows the range of response functions (namely fear and avoidance) for an ever-narrowing range of stimuli (i.e., the client is less afraid and avoidant of what is now a very small range of stimuli). Defusion, on the other hand, broadens the range of response functions for an ever-increasing range of stimuli. The fear and avoidance functions are left perfectly intact, but the range of response options to those stimuli has been increased.

The generalization research reviewed here also points to the crucial importance of prevention (as compared to treatment). Ideally, one would be able to prevent the generalization of fear, thus averting the development of pathology. Preventative interventions can be conducted prior to exposure to probable fear conditioning experiences. This might be productive in those cases where such experiences are highly probable, such as in soldiers deployed for war interventions or during the training of fire fighters. In other cases, such interventions could be used immediately after acquisition, in the time window before (or if otherwise impossible, during) generalization. For example, certain interventions might prevent the bullied 9-year-old from developing social phobia. Given that discrimination is the flipside of generalization, such techniques could include discrimination training with relevant stimuli or situations for which it makes sense to be afraid and distinguished from other perceptually or symbolically related events (see Vervliet et al., 2010a). More research in this area, including the development of effective techniques, is warranted.
Conclusions and Future Directions

Historically, research on perceptual and non-perceptual (symbolic) generalization has been rooted in two quite separate research traditions. Most of the work on perceptual generalization has been conducted within a Pavlovian conditioning framework, which typically adopts a cognitive view. It is assumed that conditioned fear responding and generalization are based on associations between representations in memory and on feature overlap with other memory representations. Such theorizing and mental constructs are largely absent within the behavioristic-functional framework from which research on symbolic generalization originated (Dymond & Roche, 2009). Both domains have been defined within different (and often opposing) meta-theoretical traditions, and to some extent continue to exist independently, with contrasting philosophical worldviews and research published in different journals and presented at different conferences. This lack of communication is unfortunate, and only recently have researchers from both domains engaged in more extensive communication and collaboration. The present article is the result of such an endeavor. We believe there is much to be gained for a translational understanding of fear generalization and anxiety disorders from research that combines both perspectives; this also necessitates an account of how both meta-theoretical traditions can co-exist and learn from each other (see De Houwer, 2011; De Houwer, Barnes-Holmes, & Moors, 2013).

Development at the theoretical level would also help integration and further development of the clinical procedures that stem from both traditions. All too often, therapies and techniques for the treatment of anxiety disorders that come from a cognitive/Pavlovian conditioning perspective, like exposure (Craske et al., 2008) and those from a functional perspective (e.g., Eifert, Forsyth, & Hayes, 2014) are viewed as hardly commensurable or only compatible at the level of day-to-day clinical practice. Here as well, much can be gained from development and integration across theoretical levels.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

References


