

RUNNING HEAD: CONDITIONING AND ANXIETY DISORDERS

**Revisiting classical conditioning as a model for anxiety disorders:**

**A conceptual analysis and brief review**

Jan De Houwer

Ghent University, Ghent, Belgium

*In press. Behaviour Research and Therapy.*

mailing address:

Jan De Houwer  
Ghent University  
Henri Dunantlaan 2  
B-9000 Ghent  
Belgium  
email: [Jan.DeHouwer@UGent.be](mailto:Jan.DeHouwer@UGent.be)

phone: 0032 9 264 64 45  
fax: 0032 9 264 64 89

**Abstract**

For almost a century now, conditioning research has provided important insights in the etiology and treatment of anxiety disorders. Nevertheless, doubts were raised about whether anxiety disorders are related to conditioning. In this paper, I focus on distinguishing different claims about the relation between anxiety disorders and conditioning as well as ways of evaluating the merits of these claims. More specifically, a distinction is made between the claim that anxiety disorders are conditioning effects and the claim that anxiety disorders are due to a specific type of conditioning mechanism (i.e., the formation and activation of S-R associations, S-S associations, or propositions). Based on a brief review of the literature, I clarify which pieces of evidence are relevant for which claims and illustrate that different claims are differentially supported by the available evidence. Finally, I discuss two strategic reasons for conceptualizing anxiety disorders as conditioning effects rather than as effects of a particular conditioning mechanism.

*Keywords: conditioning, anxiety disorders, conceptual analysis*

## **Revisiting classical conditioning as a model for anxiety disorders:**

### **A conceptual analysis and brief review**

Ever since the infamous “little Albert” study of Watson and Rayner (1920), research on classical conditioning has been regarded as a major source of information about the etiology and treatment of anxiety disorders in psychotherapy. Highly effective treatment strategies such as exposure and systematic desensitization were derived directly from insights obtained in classical conditioning research (Eelen & Vervliet, 2006). More generally, behavior therapy was conceived of as the application of conditioning principles in the treatment of psychopathology (Wolpe & Lazarus, 1966).

Although conditioning research continued to provide ideas on how to improve the psychotherapy of anxiety disorders (see Craske, Hermans, & Vervliet, 2018, for a recent review), doubts have been voiced about the core assumption of the conditioning approach to anxiety disorders, namely the idea that anxiety disorders are actually related to conditioning. In a seminal publication, Rachman (1977) provided several arguments against this idea. Likewise, Poulton and Menzies (2002) reviewed evidence which suggests that (some) anxiety disorders are due to genetic factors rather than to conditioning.

Given the historical importance of the conditioning approach to anxiety disorders, it is not surprising that critical analyses such as those of Rachman (1977) and Poulton and Menzies (2002) evoked quite some debate. In the context of this debate, a number of highly visible papers were published that defended the merits of conditioning research for understanding the etiology and treatment of anxiety disorders (e.g., Field, 2006; Mineka & Zinbarg, 2006). The current paper aims to contribute to this debate by distinguishing different claims about the nature of the relation between anxiety disorders and classical conditioning.

The fact that there is merit in distinguishing these claims is made apparent on the basis of (1) a brief review of the relevant literature which reveals that different claims are not supported to the same extent and (2) a discussion of strategic reasons for preferring certain claims over others.<sup>1</sup>

In the first part of the paper, I start by pointing out that conditioning can be defined either as an effect (i.e., a change in behavior that is due to the pairing of stimuli) or as a mediating mechanism (e.g., the formation of associations in memory). Hence, the claim that anxiety disorders are instances of conditioning could imply either that they are instances of conditioning effects or that they are due to a specific conditioning mechanism (i.e., the formation of S-R associations, S-S associations, and/or propositions). I then discuss the way in which these different claims can be evaluated. In the second part of the paper, I briefly review evidence that is relevant for evaluating these different claims. Although this evidence has been reviewed elsewhere in a much more comprehensive manner (e.g., Craske, Hermans, & Vansteenwegen, 2006; Craske et al., 2018; Field, 2006; Mineka & Zinbarg, 2006; Poulton & Menzies, 2002), the main contribution of the current review is that it evaluates the evidence for each individual claim separately, thereby revealing that evidence against one claim is not necessarily evidence against another claim or against the clinical relevance of conditioning research in general. The third and final section presents two strategic arguments for conceptualizing anxiety disorders as conditioning effects rather than as effects of a particular conditioning mechanism and thus further highlights the importance of distinguishing between these different claims.

---

<sup>1</sup> The focus will be on classical conditioning because anxiety disorders are typically linked with this type of conditioning rather than with operant conditioning. In the present paper, I will thus often use the term “conditioning” as shorthand for “classical conditioning”.

### **Conceptual Analysis**

Any debate about the relation between classical conditioning and anxiety disorders necessarily depends on one's conception of the terms "classical conditioning" and "anxiety disorders". Although there are many different anxiety disorders that might be related in different ways to conditioning (e.g., Mineka & Zinbarg, 2006), for reasons of simplicity, I do not differentiate between the different disorders, nor do I commit myself to specific criteria that can be used to delineate anxiety disorders from other disorders or non-pathological conditions. Remaining silent about these issues allows me to focus on an analysis of the concept "classical conditioning" and the implications of this analysis.

### **Two Ways of Thinking about Classical Conditioning**

**Classical conditioning as an effect.** When conceptualized as an effect, classical conditioning corresponds to a change in behavior that is due to the prior pairing of stimuli (Bolles, 1979; De Houwer, 2007; Eelen, 2018; Rescorla, 1988). For example, an increase in the skin conductance level that is evoked by a tone (conditional stimulus; CS) qualifies as an instance of classical conditioning if the increase is caused by the pairing of the tone with an aversive electric shock (unconditional stimulus; US). As such, labeling a certain change in behavior as an instance of classical conditioning involves a hypothetical causal attribution, more precisely, the causal claim that the change is due to the pairing of stimuli rather than other elements in the environment (e.g., the mere repeated presentation of a US; De Houwer, Barnes-Holmes, & Moors, 2013; De Houwer & Hughes, in press).

Multiple stimulus pairings can also jointly influence behavior. For instance, studies on sensory pre-conditioning revealed that animals fear an originally neutral CS1 after it has been paired with another neutral CS2, provided that CS2 is afterwards paired with an aversive US

(i.e., CS1-CS2 pairings followed by CS2-US pairings; see Brogden, 1939). In this case, CS1 and the aversive US have never co-occurred but they did both co-occur with a third stimulus (CS2). The change in responding to CS1 is the joint effect of the CS1-CS2 pairings and the subsequent CS2-US pairings (i.e., neither pairings alone would produce the change in behavior). Such joint effects of different stimulus pairings are typically also regarded as classical conditioning effects (but see De Houwer & Hughes, in press, Chapter 4).

It is less clear whether changes in behavior that result from conditioning instructions (i.e., instructions about stimulus pairings) can also be considered as instances of conditioning effects. Some have argued that conditioning instruction involve the spatio-temporal pairing of words (e.g., the words “tone” and “shock”) which might as such be the cause of changes in behavior (e.g., Field, 2006). However, it seems more likely that the effects of conditioning instructions are due to the symbolic meaning of the instruction as a whole rather than the mere spatio-temporal co-occurrence of words. For instance, it is likely that the sentence “tone will be followed by shock” causes more fear for the tone than the sentence “tone will not be followed by shock” even though both sentences involve a spatio-temporal pairing of the words “tone” and “shock”. Regardless of whether effects of conditioning instructions qualify as instances of conditioning, research on the effects of conditioning instructions can be considered as an integral part of conditioning research because the instructions focus on the event that is studied in conditioning research: stimulus pairings (see De Houwer & Hughes, 2016, for an in depth discussion).

**Classical conditioning as a mechanism.** When conceived of as a mechanism, classical conditioning refers to a chain of processing steps via which stimulus pairings influence behavior. Hence, conditioning mechanisms provide a potential explanation of

conditioning effects (see De Houwer et al., 2013, and De Houwer & Hughes, in press, for in depth discussions). For instance, it is often assumed that the repeated co-occurrence of a CS and US results in the gradual formation of an association between the representation of the CS and the representation of the US in memory. Once this association has been formed, the presentation of the CS can result in the activation of the US representation, which results in changes in behavior (see Bouton, 2016, for a review of the various association formation models, and Haselgrove, 2016, for a discussion of their core assumptions). However, it has also been argued that conditioning effects are mediated by a mechanism that produces and deploys propositional beliefs (e.g., De Houwer, 2009, 2018; Mitchell, De Houwer, & Lovibond, 2009; see below for more information about propositional models of conditioning). More generally, different potential conditioning mechanisms differ with regard to assumptions about the nature of the representations that mediate conditioning effects (Step 2 in Figure 1), the conditions under which stimulus pairings result in the formation of those representations (Step 1 in Figure 1), and the conditions under which the activation of representations results in changes in behavior (Step 3 in Figure 1; see De Houwer & Hughes, in press, for more details). Because it is in principle possible that conditioning effects are produced by different mechanisms, one cannot simply equate conditioning as an effect with conditioning as a particular mechanism (also see De Houwer, 2020).<sup>2</sup>

----- Insert Figure 1 About Here -----

---

<sup>2</sup> For brevity and simplicity, we discuss only cognitive theories of conditioning mechanisms, that is, theories which focus on mental representations and mental processes as the components of mechanisms. One could also conceive of conditioning mechanisms at other levels of explanation, such as neurological or biochemical mechanisms.

**How to Determine Whether Anxiety Disorders are Instances of Classical Conditioning?**

**Anxiety disorders as instances of classical conditioning effects.** The claim that anxiety disorders are instances of classical conditioning effects implies the hypothesis that those disorders result from the pairing of certain stimuli in the past environment of that individual. For instance, arguing that a patient's fear for spiders is a classical conditioning effect boils down to the hypothesis that the fear originated from past events in which the individual saw a spider that co-occurred with other events such as a painful bite from a spider. If it turns out that the fear for spiders was not caused by stimulus pairings, it would not qualify as an instance of classical conditioning. In principle, anxiety disorders could arise also from other aspects of the environment such as genetic factors (as shaped by the ancestral environment; e.g., an inborn fear of heights, see Poulton & Menzies, 2002) or the mere repeated presence of a stimulus (e.g., sensitization effects such as the escalation of a mild fear for snakes into a phobia for snakes after being confronted with many snakes in a short period of time).

In the laboratory, researchers can implement control conditions for verifying the environmental cause of a certain change in behavior. For instance, when examining the impact of tone-shock pairings on fear responses that are evoked by the tone, researchers can add a control condition in which the tone and shock are presented equally often but in an unpaired manner (e.g., Rescorla, 1966). If the tone evokes more fear after the presentation of tone-shock pairings than after the unpaired presentation of tone and shock, one can infer that the change in fear for the tone was caused by the pairings of the tone and the shock rather than the mere repeated presentation of the tone or shock. In clinical practice, on the other hand, the anxiety disorder is already present at the time that the therapist meets the patient for

the first time. The therapist therefore cannot directly observe nor control for aspects of the past environment of the patient that might have caused the anxiety disorder. This seriously complicates the debate about whether anxiety disorders are instances of classical conditioning as an effect.

Although it is nearly impossible to determine with certainty that the anxiety disorder displayed by an individual patient is an instance of classical conditioning as an effect, it is possible to evaluate the general idea that anxiety disorders are instances of conditioning effects. A first aspect of this evaluation is to examine whether stimulus pairings are a *possible cause* of anxiety disorders (i.e., anxiety disorders can be due to stimulus pairings). Rather than trying to show that each individual anxiety disorder is due to stimulus pairings, the aim is to identify at least some instances in which stimulus pairings appear to be the cause of an anxiety disorder or behavior akin to that seen in anxiety disorders. If such instances cannot be identified, there is little merit in believing that stimulus pairings are a common source of anxiety disorders.

A second aspect of the evaluation of the merits of the conditioning approach to anxiety disorders is to assess whether stimulus pairings are a *necessary cause* of anxiety disorders (i.e., anxiety disorders can only be due to stimulus pairings). Proving that stimulus pairings are necessary is difficult to achieve because it involves a claim about all possible instances of anxiety disorders in the past, present, and future. It is possible, however, to disprove the hypothesis that stimulus pairings are a necessary cause of anxiety disorders. This involves the identification of individual instances of anxiety disorders that are not caused by stimulus pairings. If most instances of anxiety disorder appear to be due to factors other than the pairing of stimuli, it would reduce the relevance of conditioning research for clinical

practice. Note, however, that the mere demonstration of the existence of other causes of anxiety disorders as such does not imply that stimulus pairings are not also an important cause of anxiety disorders.

The third aspect of testing the merits of the conditioning approach is to examine the extent to which stimulus pairings are a *sufficient cause* of anxiety disorders (i.e., stimulus pairings always cause anxiety disorders). Again it is difficult to prove that stimulus pairings are sufficient but possible to prove that they are not sufficient. The latter involves the search for cases in which anxiety disorders (e.g., excessive fear of spiders) are absent despite the presence of stimulus pairings (e.g., having experienced a painful bite from a spider). However, in itself, showing that these instances exist, says little about the merits of the conditioning approach. Even when stimulus pairings are not sufficient to cause the onset of anxiety disorders, they might still be responsible for anxiety disorders when these disorders do occur. Moreover, all causes are insufficient in that their effect depends on the presence of certain enabling conditions and can be moderated by variables in the environment (Mackie, 1965). This reduces the value of showing that a particular cause (e.g., stimulus pairings) is not a sufficient cause of a particular phenomenon (e.g., anxiety disorders).

It is, however, important to examine the precise way in which stimulus pairings are insufficient causes of changes in behavior. Such research reveals the moderators of classical conditioning effects. If anxiety disorders are instances of classical conditioning effects, they should be functionally equivalent to conditioning effects, that is, they should be moderated by the same environmental variables as conditioning effects. Entertaining the hypothesis that anxiety disorders are instances of conditioning effects has merit only if there are important parallels between the moderators of both phenomena. If there is substantial overlap between

the moderators of anxiety disorders (e.g., the conditions under which exposure therapy reduces anxiety) and the moderators of (certain types of) classical conditioning effects in the laboratory (e.g., the conditions under which extinction procedures reduces conditioned fear), it becomes useful to use the latter as a model of the former. If anxiety disorders and conditioning effects are functional equivalent (i.e., influenced in the same way by potential moderators) then more can be learned about the moderators of anxiety disorders by studying the moderators of classical conditioning effects in the laboratory. This brings us to one of the main points of this paper: The merits of the idea that anxiety disorders are conditioning effects can be evaluated by examining the match between the moderators of classical conditioning effects that are observed in the laboratory and the moderators of anxiety disorders.

**Anxiety disorders as caused by classical conditioning mechanisms.** One could also claim that anxiety disorders are due to the operation of a particular conditioning mechanism (e.g., the formation and activation of CS-US associations in memory). It is important to realize that such a claim implies that anxiety disorders are instances of conditioning effects. Without reasons for arguing that anxiety disorders are due to the pairings of stimuli, it does not make sense to examine claims about the mechanism via which stimulus pairings result in anxiety disorders. Verifying the role of conditioning mechanisms in anxiety disorders therefore requires verification of the assumption that anxiety disorders are instances of conditioning effects. On top of that, one also has to determine the specific mechanism via which stimulus pairings have this effect. Because there are multiple mechanisms via which stimulus pairings can cause changes in behavior, it is risky to infer the operation of a specific mechanism simply on the basis of establishing conditioning as an effect.

So how can one determine whether anxiety disorders are due to a specific conditioning mechanism? Because of their informational nature, neither mental representations such as associations or propositional beliefs, nor the processes via which such representations are formed and activated, are directly observable by an outside researcher (Neisser, 1967). However, theories about conditioning mechanisms can be evaluated on the basis of their heuristic and predictive value (De Houwer, 2011). Good conditioning theories are able to explain existing findings about the conditions under which conditioning effects occur (i.e., heuristic value) and to successfully predict new findings (i.e., predictive value). Research about the moderators of conditioning effects thus provides the input for the construction and evaluation of theories about conditioning mechanisms (see De Houwer & Hughes, in press).

Hence, the question of whether anxiety disorders are mediated by a specific conditioning mechanism should be examined by gathering information about the moderators of anxiety disorders. The idea that anxiety disorders are due to a specific conditioning mechanism has merit only if it allows one to account for the known moderators of anxiety disorders (heuristic function) and to predict new moderators (predictive function). Moreover, because each conditioning theory makes different predictions, each theory needs to be evaluated separately. Evidence against the involvement of one particular conditioning mechanism does not allow one to exclude the involvement of other conditioning mechanisms.

3

---

3 Our proposal for evaluating the merits of conditioning research bears some resemblance to the criteria of predictive validity and construct validity as proposed by Vervliet and Raes (2013). Predictive validity can be seen as referring to the match between the moderators of conditioning effects and the moderators of anxiety disorders. Construct validity refers to a match between the causal factors operating in conditioning effects and those operation in anxiety disorders. However, Vervliet and Raes do not distinguish between conditioning as an effect and conditioning as a mechanism. Hence, it is unclear whether their criteria refer to conditioning effects

### **A Brief Review of the Relevant Literature**

In this second section, I briefly review the available evidence regarding the different claims that can be made about the role of conditioning in anxiety disorders. I first discuss evidence that is relevant for deciding whether anxiety disorders are conditioning effects. Afterwards, I review evidence that is relevant for the claim that anxiety disorders are due to specific conditioning mechanisms. As noted above, the literature review is not meant to be comprehensive, nor to provide the final word on the merit of the different claims (for more comprehensive reviews, see Craske et al., 2006, 2018; Field, 2006; Mineka & Zinbarg, 2006; Poulton & Menzies, 2002). Instead, the review primarily aims to clarify what type of evidence is relevant for the different claims that can be made. By evaluating each claim separately, it also becomes clear that different claims are supported to a different extent, which in turn illustrates the merits of separating the different claims.

#### **Are Anxiety Disorders Instances of Classical Conditioning Effects?**

As argued above, in order to evaluate the claim that anxiety disorders are instances of conditioning effects, one needs to look for evidence that provides information about whether stimulus pairings are a *possible*, *necessary*, and *sufficient* cause of anxiety disorders. Each of these questions will be examined separately in this section. A structured overview of the evidence can be found in Table 1.

**Are stimulus pairings a possible cause of anxiety disorders?** There is strong evidence to support the conclusion that stimulus pairings can cause anxiety disorders. The most convincing evidence comes from experimental studies in which stimulus pairings were

---

(i.e., pairings as causes of conditioned fear and anxiety) or conditioning mechanisms (e.g., association formation as the mechanism via stimulus pairings produce conditioned fear and anxiety). Also note that their criterion of face validity does not refer to causes, neither in terms of stimulus pairings, nor in terms of mental mechanisms, but only to the topographical nature of changes in behavior.

actually manipulated. The “little Albert” study of Watson and Rayner (1920) was vital in convincing researchers that the pairing of a CS (e.g., a rabbit) with an aversive US (e.g., a loud noise) can cause fear responses to CS similar to those seen in anxiety disorders. Because of ethical reasons, few subsequent studies have tried to induce anxiety disorders in humans by pairing stimuli (see Field & Nightingale, 2009, for a review). Other research in humans and non-human animals, however, clearly shows that the pairing of stimuli can lead to behaviors very similar to those that form the core of a variety of anxiety disorders (e.g., Delgado, Olsson, & Phelps, 2006). Evidence for the causal impact of stimulus pairings on anxiety disorders was also obtained in naturalistic and retrospective studies (e.g., Yule, Udwin, & Murdoch, 1990; see Field, 2006, for a review).

**Are stimulus pairings a necessary cause of anxiety disorders?** The claim that stimulus pairings are a necessary cause of anxiety disorders can be evaluated by looking for instances of anxiety disorders that are not due to stimulus pairings (even when the actual cause is not clear) and by identifying anxiety disorders that are due to environmental causes other than stimulus pairings. Poulton and Menzies (2002; also see Rachman, 1977) reviewed a large number of retrospective and longitudinal studies suggesting that anxiety disorders can develop in the absence of any aversive stimulus pairings, that is, pairings involving aversive stimuli and the stimuli that now evoke fear or anxiety. They also identify three alternative routes by which anxiety disorders could develop in the absence of aversive stimulus pairings (i.e., evolution, observation, and verbal messages; see Smoller, Block, & Young, 2009; Askew & Field, 2008; Mertens, Boddez, Sevenster, Engelhard, & De Houwer, 2018, and Muris & Field, 2010, for reviews of the relevant literature).

The relevance of this evidence, however, depends somewhat on the way in which

conditioning effects are defined. As we noted above, some researchers entertain a broad definition of conditioning effects that includes changes resulting from multiple stimulus pairings, verbal instructions about stimulus pairings, or the observation of fearful responses in others (e.g., Field, 2006; but see De Houwer & Hughes, in press, Chapter 4). Hence, although stimulus pairings in narrow sense (i.e., the actual co-occurrence of a neutral and aversive stimulus) are almost certainly not a necessary cause or the only possible cause of anxiety disorders, one could still argue that stimulus pairings in a broad sense (i.e., including multiple stimulus pairings, verbal instructions, or the observation of others) is a significant source of anxiety disorders. This conclusion, however, depends on one's definition of conditioning as an effect.

**Are stimulus pairings a sufficient cause of anxiety disorders?** The available evidence clearly shows that the co-occurrence of a neutral stimulus with an aversive event is not sufficient to cause an anxiety disorder (e.g., Rachman, 1977). As became clear from the conceptual analysis, however, it is more important to examine whether stimulus pairings are in the same way insufficient as causes of anxiety disorders as they are insufficient as causes of classical conditioning effects. In other words, it is important to examine whether anxiety disorders and classical conditioning effects are moderated in the same way by variables in the environment. Two questions can be examined. (1) Does knowledge about the variables that moderate conditioning effects shed light on the conditions under which anxiety disorders develop? (2) Do variables that moderate the disappearance of conditioning effects inform us about how anxiety disorders can be treated? For therapists, the second question is the most crucial one.

Much of the debate about the merits of the conditioning approach to anxiety disorders

can be conceptualized as a discussion about the conditions under which pairings of a neutral stimulus and an aversive event result in the development of an anxiety disorder. Before the 1960's, some researchers had suggested that aversive stimulus pairings might be a sufficient condition for the development of anxiety disorders (e.g., Rachman, 1977). When it became apparent that aversive stimulus pairings are, at least in many cases, not a sufficient cause of anxiety disorders, one possible conclusion was that classical conditioning and anxiety disorders are not functionally equivalent and thus that the conditioning approach to anxiety disorders has little merit. However, the accuracy of conclusions about the functional equivalence of anxiety disorders and conditioning effects is limited by the accuracy of our knowledge of the functional properties of conditioning effects. The assumption that stimulus pairings are a sufficient cause of changes in behavior was simply incorrect (e.g., Davey, 1997; Field, 2006; Mineka & Zinbarg, 2006). To give just one example, research has shown that the effect of pairing a CS with a US can be reduced or even prevented by repeatedly presenting the CS on its own before the CS-US pairings (i.e., CS pre-exposure effects; see Lubow, 1995, for a review). Hence, the fact that aversive stimulus pairings are not always a sufficient cause of anxiety disorders does not disprove the functional equivalence of anxiety disorders and conditioning and does not threaten the conditioning approach to anxiety disorders.

Over the past forty years, conditioning research has uncovered many moderators of the effect of stimulus pairings on behavior (for reviews, see Bouton, 2016; De Houwer & Hughes, in press; & Field, 2006). Interestingly, knowledge about these moderators provided new insights in the development of anxiety disorders by revealing possible reasons for the large variability in the development of anxiety disorders (also see Field, 2006; Mineka & Zinbarg, 2006). For instance, individuals who experience a particular aversive event (e.g.,

pain while being treated by a dentist) are less likely to develop anxiety for stimuli related to this event (e.g., the dentist) if they have prior experience with the stimuli involved in the aversive event (e.g., visits to the dentist prior to the painful treatment; see Kent, 1997; Surwit, 1972). This observation is in line with the phenomenon of CS pre-exposure effects in conditioning research. More generally, there appear to be parallels between the conditions under which classical conditioning effects occur and the conditions under which anxiety disorders develop. This important observation suggests that both phenomena are at least to some extent functionally equivalent and that it is therefore useful to consider anxiety disorders as instances of conditioning effects.

Research on the moderators of classical conditioning not only helps us understand the development of anxiety disorders but also sheds light on the treatment of anxiety disorders. It has long been known that presenting a CS (e.g., a tone) on its own after the CS-US pairings (e.g., tone-shock) reduces the effect that the CS-US pairings have on behavior (e.g., reduction in fear for the tone). This moderating effect is known as *extinction* (e.g., Bouton, 1993). The fact that anxiety disorders can be treated by presenting feared objects on their own (as is done in exposure treatment; e.g., exposure to a spider), suggests that classical conditioning and anxiety disorders are to some extent functionally equivalent. There are many other parallels between the conditions under which extinction occurs in classical conditioning research and the conditions under which treatment of anxiety disorders is successful. For instance, both extinction and exposure treatment effects are highly context dependent and dissipate quickly (see Bouton, 1993, and Craske et al., 2006, 2018, for reviews). Given the success of earlier generalizations from extinction research to the treatment of anxiety disorders, there is hope that this research will help clinicians to further improve their therapies (e.g., Craske et al.,

2018).

To summarize, stimulus pairings are certainly not always a sufficient cause of anxiety disorders. However, this conclusion does not undermine the conditioning approach to anxiety disorders. Conditioning research has shown that also in the laboratory, stimulus pairings do not always lead to changes in behavior. More importantly, there are some parallels between the conditions under which conditioning effects occur in the laboratory and the conditions under which anxiety disorders develop in real life. Likewise, variables that moderate the reduction of conditioned changes in behavior also seem to moderate treatment of anxiety disorders. To the extent that classical conditioning and anxiety disorders are indeed functionally equivalent, there is merit in thinking of anxiety disorders as instances of conditioning effects.

### **Are Anxiety Disorders Due to a Specific Conditioning Mechanism?**

If stimulus pairings are a significant source of anxiety disorders, then it makes sense to examine the conditioning mechanism via which stimulus pairings produce anxiety. For each specific conditioning mechanism, it needs to be examined whether it is compatible with the fact that stimulus pairings are a possible but not a necessary or sufficient cause of anxiety disorders. In the present section, I briefly discuss three possible conditioning mechanisms: the formation and activation of S-R associations, S-S associations, and propositions. I also consider the idea that multiple mechanisms might be responsible for specific instances of anxiety disorders. An overview of the conclusions can again be found in Table 1.

**S-R association models.** According to S-R association models, stimuli (e.g., a tone) that co-occur with a response (e.g., fear evoked by a shock) will be associated with that response and will therefore be able to evoke the response on subsequent encounters (e.g.,

Thorndike, 1911). Importantly, the formation of S-R associations is typically thought of as a relatively unconditional consequence of the co-occurrence of stimuli and responses. Note that when a CS and US co-occur, the only function of the US is to evoke a response that is then associated with the CS. Nothing is learned about the US as such.

Although S-R association models have now been largely abandoned by conditioning researchers because of their inability to account for key findings (see Bouton, 2016, De Houwer & Hughes, in press, and Rescorla, 1988, for reviews), they were dominant at the time when behavior therapy emerged (see Eelen, 2018, for a review) and continued to influence thinking about psychopathology later on (e.g., Carter & Barlow, 1995). S-R association models can indeed account for the fact that stimulus pairings are a possible cause of anxiety disorders (e.g., Mowrer, 1939). If an originally neutral stimulus (e.g., a spider) co-occurs with an aversive event (e.g., a painful bite from the spider) that evokes an emotional response, this results in the formation of an association between the neutral stimulus and the emotional response. Subsequent presentations of the neutral stimulus will thus lead to the emotional response. S-R association models are in line with the results of naturalistic and retrospective studies which show a positive relation between anxiety disorders and traumatic events. Stimuli encountered during the traumatic events would become associated with the negative emotional responses that typically arise in those events. As a result, the stimuli will evoke these negative emotional responses when presented after the traumatic event.

S-R association models cannot explain by themselves the occurrence of anxiety disorders in the absence of aversive stimulus pairings (i.e., that pairings are not a necessary condition). A stimulus and an emotional response need to co-occur in space and time in order for both to become associated. Hence, the only way for stimuli to acquire fear-provoking

properties is to co-occur with fear responses during a traumatic event. Any instance of anxiety disorders in the apparent absence of a traumatic event falls beyond the scope of S-R association models and raises doubts about the importance of S-R association formation as a mechanism underlying anxiety disorders. At the very least, one has to assume that it is not the only effective mechanism (see below for the section on multiple-mechanism accounts).

It is also difficult to reconcile S-R association models with the fact that anxiety disorders can result from instructions and observation. One might argue that instructions (e.g., “this animal is dangerous”) and observations (e.g., of someone showing a fearful expression in the presence of an animal) evoke emotional responses that are associated with certain stimuli (e.g., the sight or name of a particular animal). However, because only the co-occurrence of stimuli and responses matters for the formation of S-R associations, the effect of instructions and observation should depend only on *which* stimuli and responses occur rather than on *how* those stimuli and responses are related. For instance, if an observed model shows a fearful response in the presence of a novel animal, this should result in associations between the animal and the model’s fear response regardless of whether model has actually seen the animal (and thus whether the expression could provide information about the properties of the animal). Both common sense and empirical evidence (e.g., Baeyens, Vansteenwegen, De Houwer, & Crombez, 1996) suggests that the effect of observations does depend on cues about the way in which observed events and responses are related. Likewise, the instruction “you are in danger because of this animal” is likely to have an entirely different effect than the instruction “you are in danger but not because of this animal”. In both cases, the word “animal” co-occurs with a fearful response but only the first instruction is likely to install fear for the animal. Although particular instantiations of S-R models might be

able to explain the fact that not all instructions or observations are equally potent in installing fears, it seems unlikely that S-R models can fully capture the effect of instructions and observations on the development of fear and anxiety.

Research also showed that stimulus pairings are not always a sufficient cause of anxiety disorders (see Rachman, 1977). The fact that pairings involving a traumatic event do not always lead to anxiety disorders, clearly contradicts the old idea that the co-occurrence of stimuli and negative emotional responses always leads to formation of S-R associations and that these associations always lead to anxiety disorders (Wolpe & Rachman, 1960). One could of course develop more sophisticated S-R association formation models that incorporate assumptions about the conditions under which S-R associations are formed or altered and the conditions under which those associations influence behavior (see McAllister & McAllister, 1991, for a discussion). The value of these more sophisticated S-R association formation models depends on how well they are able to account for the known moderators of anxiety disorders (i.e., heuristic function) and to predict new, previously unseen moderating effects (i.e., predictive function; De Houwer, 2011). Although it is beyond the scope of the present paper to provide a comprehensive review of the successes and failures of S-R association formation models (see Dickinson, 1980, and McAllister & McAllister, 1991, for reviews), there is large consensus in the literature that the heuristic and predictive function of these models is limited.

In sum, S-R association formation models can be evaluated by examining whether they can account for the way in which stimulus pairings produce anxiety disorders. Although they can account for the fact that stimulus pairings are a possible cause of anxiety disorders, they have problems dealing with the fact that stimulus pairings are neither a necessary nor a

sufficient cause of anxiety disorders. Hence, there seems to be little value in assuming that anxiety disorders are typically mediated by such a mechanism.

**S-S association formation models.** According to S-S association formation models, the pairing of stimuli (e.g., tone and shock) results in the formation of an association between the representations of those stimuli in memory. Once a CS representation (e.g., of the tone) has become associated with a US representation (e.g., of the shock), the presentation of the CS can activate the representation of the US, which in turn results in anticipatory responses (e.g., fear of the painful shock; but see Jozefowicz, 2018). Different S-S association formation models differ in their assumptions about when stimulus pairings lead to S-S associations and when S-S associations influence behavior (see Figure 1; for reviews, see Bouton, 2016; Pearce & Bouton, 2001).

Traumatic events could lead to anxiety disorders because they involve the pairing of initially neutral stimuli (CSs) and highly aversive stimuli (USs). This results in the formation of associations between the neutral and aversive stimuli, thus allowing the neutral stimuli to afterwards activate representations of the USs and the negative emotional responses that are part of or linked to those representations. S-S association formation models can thus explain why naturalistic and retrospective studies reveal a correlation between traumatic events and anxiety disorders.

Although S-S association formation models postulate that the pairing of a neutral and aversive event are a possible cause of anxiety disorders, such aversive stimulus pairings are not assumed to be a necessary cause of anxiety disorders. Even the pairing of neutral stimuli (CS1-CS2) can eventually lead to anxiety disorders, for instance, when the second of those neutral stimuli is subsequently paired with an aversive US (CS2-US). In this case, the initial

pairings result in an association between the first and the second stimulus representation whereas the subsequent pairings create an association between the second stimulus representation and the US. Because of these representations, activation of the representation of the first stimulus (CS1) can spread to the representation of the second stimulus (CS2) and subsequently to the representation of the US. S-S association formation models can thus account for phenomena like sensory pre-conditioning. This broadens their explanatory power to those instances of anxiety disorders in which the feared object did not appear to have been paired with aversive events. For instance, a child that first sees spiders in the garden and afterwards is bitten by a spider at school, might start fearing and avoiding gardens. It also points at new ways of treating disorders, for instance, by devaluing the threat value of the US that is thought to underlie the disorder (e.g., reducing fear for a spider bite to reduce fear of going in the garden; see Davey, 1997, for an insightful discussion).

Evidence for the hereditary nature of anxiety disorders can be accommodated by S-S association formation models that postulate selectivity in the formation of S-S associations. Note that neither S-R nor S-S association formation models inherently predict (genetically determined) selectivity in the formation of associations. Nevertheless, whereas the formation of S-R associations is sometimes thought to be an unconditional consequence of stimulus-response co-occurrences, proponents of S-S association formation models typically acknowledge the fact that not all stimulus pairings are equally potent in creating associations (e.g., Mineka & Zinbarg, 2006; Öhman & Mineka, 2001).

S-S and S-R association formation models face similar limitations in the way they account for the effects of instructions and observation. Most crucially, they struggle with the fact that the effects of instructions depend on the symbolic meaning of the whole sentence

rather than merely on the spatio-temporal co-occurrence of words. S-S associations are very limited in their capacity to code symbolic meaning, more specifically the way that stimuli are related (Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Mitchell, De Houwer, & Lovibond, 2009). For instance, instructions such as “the animal causes pain” and “the animal has pain” should lead to the same animal-pain association and thus the same changes in behavior. Although I do not exclude the possibility that S-S association formation models will be able to deal with this issue in the future, existing S-S models do not seem to provide a full account of the effect of instructions and observation on anxiety disorders (see De Houwer, 2009, 2018).

S-S association formation models are, however, compatible with the observation that stimulus pairings with aversive events do not always result in anxiety disorders. Most models incorporate explicit assumptions about the variables that moderate the effect of stimulus pairings. In fact, most of the research on classical conditioning during the past 40 years has been directed at testing the predictions of different S-S models about the variables that moderate classical conditioning effects (e.g., see Bouton, 2016, De Houwer & Hughes, in press, and Rescorla, 1988, for reviews). Many of the moderators that were identified in this research also seem to influence anxiety disorders. Others have reviewed these successes of S-S association formation models in detail (see Craske et al., 2006, 2018; Field, 2006; Mineka & Oehlberg, 2008; Mineka & Zinbarg, 2006). For instance, S-S based models have helped to shed new light on the conditions under which extinction and exposure treatments result in a reduction of fear (e.g., Bouton, 1993, and Craske et al., 2006, 2018, for reviews).

Note that different S-S models differ in their heuristic and predictive value, that is, in their ability to account for and predict the effect of moderators of classical conditioning and

anxiety disorders. For instance, the well-known Rescorla-Wagner model (Rescorla & Wagner, 1972) is unable to explain relapse after successful exposure treatment. Other S-S models, however, can account for this important observation and inform us about the conditions under which relapse might be prevented. These more recent models postulate that extinction and exposure effects are due to (context dependent) additional learning rather than forgetting. Hence, extinction and exposure effects could be made less context dependent and more long lasting by making this additional learning stronger and less context dependent (e.g., Bouton, 1993; Craske et al., 2018).

In sum, S-S association models are quite specific about the way in which stimulus pairings might produce anxiety disorders. At least some of their predictions seems to be in line with the available evidence. As noted by a reviewer, however, some arguments in support of S-S association models of anxiety disorders (e.g., the idea that multiple stimulus pairings could jointly produce anxiety disorders) are based primarily on speculation rather than actual data. Moreover, S-S models do not fit well with evidence showing that anxiety disorders can result from and be affected by verbal instructions and the observation of others.

**Propositional models.** Propositional models are built on the assumption that classical conditioning is mediated by the formation of propositional beliefs about the relation between stimuli (e.g., De Houwer, 2009, 2018; Mitchell et al., 2009). In other words, the pairing of two stimuli (e.g., tone and shock) can lead to a change in an individual's behavior only after this individual formed the belief that the two stimuli are related in a certain way (e.g., the belief that the tone predicts the shock). Unlike S-S associations, propositions contain information not only about the fact that stimuli are related but also about *how* those stimuli are related. This has the important implication that identical stimulus pairings (e.g., the

observation that patients with a certain disease also have a certain chemical substance in their blood) can lead to different propositions (e.g., “the disease causes the substance” or “the disease is caused by the substance”) that influence behavior in different ways (e.g., whether one tries to remove the substance in order to cure the disease; see Lagnado et al., 2007, for an insightful discussion of the difference between associations and propositions). Whereas the effects of propositions on behavior can be automatic (e.g., it is difficult to stay calm when one believes that wild tiger is nearby), the formation of propositions is thought to be a largely non-automatic process that has much in common with problem-solving (see De Houwer, 2018, for more details). This implies that stimulus pairings influence behavior only when the individual has sufficient time and cognitive resources to form a conscious proposition about these pairings. Although propositional models differ in important ways from S-S and especially S-R association models, they are in line with many important findings in classical conditioning research (see De Houwer, 2009, and Mitchell et al., 2009, for reviews). For instance, the fact that classical conditioning is rarely if ever observed in the absence of awareness of the CS-US relation strongly supports the idea that conditioning is mediated by the formation of conscious propositions (see Mertens & Engelhard, in press, for a recent review).

Could the impact of stimulus pairings on anxiety disorders also be mediated by the formation of propositions (see Lovibond, 2011, for a more extensive discussion of the implications of propositional models for understanding and treating anxiety disorders)? Propositional models can certainly deal with the fact that stimulus pairings are a possible source of anxiety disorders. The co-occurrence of a neutral and a negative stimuli (e.g., the pain caused by a spider that bites you) can lead to negative beliefs about the originally neutral

stimulus (e.g., “spiders can cause a painful bite”) and thus to fear of that stimulus. Once a belief about the relation between stimuli has been formed, it can lead to all kinds of new related beliefs (e.g., “spiders are dangerous”). As such, propositional models are in line with the observation that anxiety disorders are often if not always accompanied by negative beliefs about the feared object. These beliefs can be irrational if they are based on a biased sampling of events (e.g., being exposed mainly to spiders that have painful bites rather than benign spiders) or a biased weighting of events in the world (e.g., more weight is given to one experience with a dangerous spider than to many experiences with benign spiders). Although people might not always remember the stimulus pairings that led to the formation of those propositional beliefs, at least in some cases patients should be able to retrospectively report those pairings.

The fact that anxiety disorders sometimes develop in the apparent absence of traumatic events can also be explained on the basis of propositional models. Once two neutral stimuli have co-occurred and a proposition has been formed about the relation between those neutral stimuli (e.g., “this frog has spots on its skin”), subsequent experiences (e.g., learning that amphibians with spotted skin tend to be poisonous) can dramatically alter the impact of those initial stimulus pairings on behavior (e.g., physical contact with the frog is now avoided). Hence, anxiety disorders could develop even when the feared object did not co-occur with a negative event.

The importance of genetic factors in the development of anxiety disorders can be accommodated by propositional models. Although still mere speculation at this time, it is possible that genetic factors might contribute to the development of anxiety disorders by determining the probability that conscious propositional beliefs are formed as the result of

certain stimulus pairings. For instance, throughout the evolution of mankind, spiders have posed a bigger challenge for survival than squirrels. This could have predisposed humans to form negative beliefs about spiders. For instance, someone who experiences a panic attack in a park might be more inclined to causally attribute the panic to the spider they saw crawling over the grass than to the squirrel that just passed by. As a result, he or she would be more likely to develop a fear for spiders than a fear for squirrels. An impact of genetic factors could thus be mediated by the formation of propositional beliefs (see Testa, 1974, for a precursor to this idea).

The main strength of propositional models, however, is that they are perfectly in line with the fact that anxiety disorders can result from instructions and observations. Whereas S-S and S-R associations can at best capture the co-occurrence of stimuli that part of the instructions or observation, propositions represent also the manner in which those stimuli are related and thus encode more fully the symbolic implications of instructions and observed events. Propositional models also clarify the relation between the effects of pairings, instructions, and observations. More specifically, actual pairings of stimuli are seen as one source for the formation of propositional beliefs about relations between stimuli. Those same beliefs can be formed also on the basis of the symbolic meaning of instructions about how stimuli or events are related or the observation of how others respond to stimuli. For instance, some people might believe that spiders are dangerous because they have once been bitten by a spider, others might have formed this belief because they saw someone else been bitten by a spider, and still others might have come to that belief because their parents told them that spiders can bite. Different sources of beliefs might differ in their effectiveness in producing beliefs (e.g., actual experiences might lead to more firm beliefs than observation or

instruction, e.g., Reiss, 1980) but all routes are fundamentally similar in that they can produce propositional beliefs that lead to anxiety. Note that this idea fits well with the observation that the moderators of conditioning effects are highly similar to those of the effects of conditioning instructions (Mertens et al., 2018).

Finally, propositional models also imply that stimulus pairings are not sufficient for the development of anxiety disorders. Because stimulus pairings affect behavior only after a proposition about the relation has been formed, conditioning should be moderated by variables that influence the formation of propositions. For instance, whether a stimulus pairing is noticed (i.e., a conscious proposition about the pairings is formed) depends on whether attention is directed to the stimuli. Likewise, the content of beliefs (i.e., assumptions about whether and how stimuli are related) is biased by preexisting beliefs and cues in the current environment that could disambiguate the nature of the relation between stimuli. Let us return to the example of someone who has a panic attack in a park. Whether the panic attack is causally attributed to the spider could depend on the knowledge about spiders that the person already possesses (e.g., the belief that spiders are a plausible cause of panic attacks because many other people are afraid of spiders).

Although propositional models have not yet been formalized, they do provide inspiration for research on the moderators of anxiety disorders. For instance, relapse after treatment might occur when patients retrospectively reevaluate the beliefs they formed during treatment (e.g., that a spider did not hurt them during treatment because the therapist used a special kind of spider during therapy; see Raes, De Houwer, Verschuere, & De Raedt, 2011, for experimental evidence regarding this idea). Also, propositional models imply that there is no need for “treatment matching”, that is, an overlap between the type of experience that has

led to the anxiety disorder (e.g., the direct experience of stimulus pairings) and the type of experience that is used to treat the disorder (e.g., verbal instruction; see Lovibond, 2011). Of course, this does not imply that simply giving verbal information is sufficient to solve all anxiety disorders, even those that are based on stimulus pairings. The revision of beliefs is not only difficult to achieve but will be effective only if one changes those beliefs that actually underlie the feelings and behaviors one wishes to tackle. A propositional perspective on conditioning reveals that the pairing (or un-pairing) of stimuli is just one tool for belief revision, a tool that can be combined with other tools such as instructions and observation. Rather than assuming that conditioning is part of a separate realm that is driven by association formation, it can now be brought into the realm of belief revision. Nevertheless, conditioning research remains important for clinical psychology. Using conditioning procedures in the lab not only allows researchers to study an important route of belief formation but also to explore various ways of counteracting fear via belief revision. More specifically, once fear has been established via stimulus pairings, it can be examined how stimulus presentations (e.g., extinction procedures), instructions, and observations (jointly) alter fear.

In sum, also propositional models provide specific ideas about the role of stimulus pairings in anxiety disorders. Like S-S association formation models, they can account for the fact that stimulus pairings are possible but not a necessary or sufficient cause of anxiety disorders. However, they differ considerably from S-R and S-S association formation models in the way that they account for anxiety disorders that result from instructions and observation. They also fit well with the idea that beliefs play an important role in anxiety disorders. On the other hand, they are less formalized than most association formation models. Also, many of the arguments in favor of proposition formation as the mechanism via

which stimulus pairings produce anxiety disorders remain speculative.<sup>4</sup>

**Multiple mechanisms.** At least in principle, multiple conditioning mechanisms could contribute to the development of anxiety disorders. For instance, some anxiety disorders might be mediated by S-R association formation whereas others are the result of S-S association formation and still other are produced by propositional processes. Such dual- or multiple-mechanisms views are currently popular both in conditioning research (e.g., McLaren et al., 2014) and psychology in general (but see Keren & Schul, 2009, Melnikoff & Bargh, 2018, and Moors & De Houwer, 2006, for critical discussions).

The heuristic value of a multiple-mechanism view is high because it combines the explanatory power of each of the individual mechanisms. Whenever a particular empirical finding fits well with the operation of one but not another mechanism, one can argue that the observed effect is due the former but not the latter mechanism. For instance, the fact that anxiety disorders can result from instructions in the absence of traumatic events can be attributed to the operation of propositional processes whereas anxiety disorders that are resistant to the impact of instructions could be seen as the result of S-R association formation.

Although the increased heuristic value of a multiple-mechanism view is appealing, its predictive value is low when there are no constraints on the way different mechanisms interact. Without knowledge about these constraints, it is impossible to predict when a first rather than a second or third mechanism will operate. At present, little is known about these

---

<sup>4</sup> Some of the implications of propositional models for our understanding of anxiety disorders are reminiscent of the ideas put forward by Davey (1997) and Field (2006). Like proponents of propositional models, Davey and Field pointed at the importance of conscious beliefs in both conditioning and anxiety disorders, more specifically the role of conscious expectancies. They also allowed for the possibility that conditioning result from observation and instruction. Nevertheless, Davey (1997) and Field (2006) remained committed to the view that conditioning effects and anxiety disorders are mediated by the formation of S-S associations. As discussed earlier, it is difficult to see how models based on S-S associations can capture the full impact of conscious beliefs, instructions, and observation on behavior (see De Houwer, 2009, 2018; Jozefowicz, 2018).

constraints (Mitchell et al., 2009). Moreover, testing ideas about these constraints is likely to be challenging. As I noted earlier in this paper, even the hypothesis that conditioning is due to one particular mechanism can be evaluated only indirectly by examining the moderators of conditioning. Studying the way in which two or more conditioning mechanisms interact is even more challenging in that it can be achieved only by examining interactions between multiple moderators. Consider the most straightforward (but unlikely) case in which there is general agreement about the fact that the (lack of) impact of one moderator is diagnostic of the operation of one particular conditioning mechanism (e.g., the absence of US-revaluation as an indicator of S-R association formation). Studying the constraints of this mechanism requires knowledge about the variables that determine the impact of the diagnostic moderator, that is, the moderator which is assumed to be diagnostic of a particular conditioning mechanism (e.g., the conditions under which US-revaluation occurs; see Bouton, 2016). The more complex the ideas about how different mechanisms interact, the more difficult it becomes to test these ideas. But if researchers wish to move beyond the mere illusory benefits of unconstrained multiple-mechanism models, they need to engage in this type of meta-conditional research that focuses on interactions between moderators (De Houwer, 2007).

### **Are Anxiety Disorders Best Conceived of as Conditioning Effects or Products of a Conditioning Mechanism?**

The previous section highlights an important reason for distinguishing between different claims about the relation between conditioning and anxiety disorders: evidence against one particular claim (e.g., anxiety disorders as products of S-R association formation) should not be interpreted as evidence against another claim or against the conditioning

approach tout-court. This section focusses on another reason for distinguishing different claims: thinking about anxiety disorders in terms of conditioning effects versus conditioning mechanisms has important implications for the viability of the conditioning approach to anxiety disorders. First, a focus on conditioning effects improves the cumulative nature of research. Theories about conditioning mechanisms have changed in the past and will probably continue to change in the years to come. When conditioning is conceptualized in terms of one specific mechanism, evidence against that mechanism can threaten conditioning research as a whole. This point is clearly illustrated by how evidence against S-R association formation models in humans has been interpreted by many as problematic for the conditioning approach as a whole (e.g., Brewer, 1974; see Eelen, 2018). If conditioning is equated with the formation of S-R associations, and if S-R association formation does not occur in human, it no longer makes sense to conduct conditioning research in humans or to conceptualize anxiety disorders in terms of conditioning.

On the other hand, when conditioning is defined as an effect, functional knowledge about moderators can be framed independently from theories about the mechanisms that mediate these effects. For instance, the fact that an extinction procedure leads to a reduction in conditioned fear is an important piece of knowledge regardless of which mechanism mediates (the extinction of) conditioning effects. This does not mean that theories about mechanisms should simply be ignored. On the contrary, mechanistic theories can be instrumental in the discovery of new important functional knowledge and knowledge about mechanisms can have value as such (see De Houwer, 2011, and Hughes et al., 2016, for a detailed discussion of the mutually supportive nature of functional and mechanistic approaches in psychology). For instance, the influential work of Bouton (1993) on the

context-dependency of extinction effects has greatly increased our knowledge of the moderators of the extinction effect (see Craske et al., 2018, for a review). Although this work was for a large part inspired by S-S association formation theories, the insight that extinction effects depend on changes in context remains highly valuable even if those S-S association formation theories would at some point in time have to be abandoned. Therefore, conceiving of anxiety disorders as conditioning effects allows for a constant development of our understanding of conditioning and anxiety disorders even when ideas about conditioning mechanisms change over time. Changes in ideas about the mechanisms that mediate classical conditioning should not be considered as a threat for a functional conditioning approach. On the contrary, these changes can offer new perspectives on the moderators of conditioning effects and thus new inspiration for discovering the moderators of anxiety disorders.

A second advantage of thinking about anxiety disorders as conditioning effects is that it allows one to appreciate more fully the value of conditioning research for various types of therapies. Hayes (2004) proposed a distinction between three types or waves of behavior therapy. First wave therapies resulted directly from conditioning research but were conceptualized primarily in terms of the formation and change of S-R associations (e.g., Wolpe & Rachman, 1960). Second wave therapies were also conceptualized in a mechanistic manner but focused on the correction of cognitive mechanisms (i.e., pathological mental representations or processes; e.g., Beck, 1975; Lazarus & Folkman, 1984). Third wave therapies do not focus primarily on correcting mechanisms but on the context and functions of psychological phenomena (e.g., Hayes, 2004).

When anxiety disorders are conceptualized as conditioning effects, conditioning research can be seen as a source of information about which variables in the environment

produce and modify anxiety disorders. From this perspective, conditioning research can continue to contribute to first wave therapies without a commitment to questionable S-R association formation models. For instance, S-S association formation inspired research on the context-dependency of extinction effects which promise to improve exposure therapies (e.g., exposure in multiple contexts; see Craske et al., 2018, for an overview). Such changes in therapies can be implemented without making assumptions about the mechanisms that mediate these improvements.

By revealing the moderators of conditioning, conditioning research can also contribute to second wave therapies. Based on the idea that anxiety disorders are conditioning effects, functional knowledge concerning the moderators of conditioning provides the input for theories about the cognitive mechanisms that mediate anxiety disorders and other conditioning effects. Mechanistic models such as cognitive theories can be evaluated on the basis of how well they are able to explain existing empirical knowledge and lead to the discovery of new empirical knowledge. Therefore, the more we know about the variables that moderate the effects of stimulus pairings, the better able we are to evaluate and improve cognitive theories about the processes and representations that mediate anxiety disorders and other conditioning effects. For instance, the work of Bouton (1993) and Craske et al. (2018) suggests that context-dependent associations or beliefs mediate anxiety disorders. Conditioning research can not only help to identify these mediating cognitive structures but also to understand their environmental origins and to find ways of changing them.

Finally, the idea that anxiety disorders are conditioning effects also fits well with the non-mechanistic nature of certain third wave therapies like ACT. In fact, many therapeutic techniques that are used in third wave therapies (e.g., reinforcing efforts to not control

anxiety) could be seen as potential moderators of the effect of stimulus pairings on anxiety disorders and could thus be studied also in conditioning research (e.g., Treanor, 2011, for an overview of how mindfulness can moderate effects of exposure). Because conditioning is often conceived of in terms of a specific mechanism, this potential of conditioning research for third wave therapies still remains to be recognized and explored.

### **Conclusion**

During the second half of the previous century, conditioning research was the main source of inspiration for ideas about the origins and treatment of anxiety disorders. Since the 1980's, however, the impact of the conditioning research on clinical psychology has waned. This decline can in part be attributed to a number of influential papers that questioned the relevance of conditioning research for anxiety disorders (Poulton & Menzies, 2002; Rachman, 1977) as well as the rise of a variety of therapies that are typically not seen as part of the conditioning approach. The aim of this paper was to shed new light on the merits of the conditioning approach to anxiety disorders by clearly separating the different claims that can be made about the relation between conditioning and anxiety disorders and by indicating how these different claims can be evaluated. Although questions can be raised about certain claims (e.g., that anxiety disorders are due to the formation of S-R associations), this does not imply that other claims are without merit. When clearly separating the different claims that can be made, it becomes clear that conditioning research is likely to remain an important source of information about origins and treatment of anxiety disorders.

## References

- Askew, C. & Field, A. P. (2008). The vicarious learning pathway to fear 40 years on. *Clinical Psychology Review, 28*, 1249-1265.
- Baeyens, F., Vansteenwegen, D., De Houwer, J., & Crombez, G. (1996). Observational conditioning of food valence in humans. *Appetite, 27*, 235-250.
- Bechtel, W. (2005). The challenge of characterizing operations in the mechanisms underlying behavior. *Journal of the Experimental Analysis of Behavior, 84*, 313-325.
- Beck, A.T. (1975). *Cognitive therapy and the emotional disorders*. Madison, CT: International Universities Press.
- Bolles, R.C. (1979). *Learning theory* (2nd ed.). New York: Holt, Rinehart & Winston.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin, 114*, 80-99.
- Bouton, M. E. (2016). *Learning and Behavior: A Contemporary Synthesis* (2nd ed.). Sunderland, MA: Sinauer.
- Brewer, W. F. (1974). There is no convincing evidence of conditioning in adult humans. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and the symbolic processes* (pp. 1-42). Hillsdale, NJ: Erlbaum.
- Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology, 25*, 323-332.
- Brysbaert, M., & Rastle, K. (2009). *Historical and conceptual issues in psychology*. London: Pearson Education.
- Chiesa, M. (1992). Radical behaviorism and scientific frameworks: From mechanistic to relational accounts. *American Psychologist, 47*, 1287-1299.

- Chiesa, M. (1994). *Radical behaviorism: The philosophy and the science*. Boston, MA: Authors' Cooperative.
- Craske, M. G., Hermans, D., & Vansteenwegen, D. (2006). *Fear and learning: From basic processes to clinical implications*. Washington, DC: American Psychological Association.
- Craske, M., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions B: Biological Sciences*, 373: 20170025.
- Davey, G. C. L. (1997). A conditioning model of phobias. In G. C. L. Davey (Ed.), *Phobias: A handbook of theory, research, and treatment* (pp 301-322). Chichester, UK: Wiley.
- De Houwer, J. (2007). A conceptual and theoretical analysis of EC. *The Spanish Journal of Psychology*, 10, 230-241.
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, 37, 1-20.
- De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, 6, 202-209.
- De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin*, 13(3), Article e28046. <https://doi.org/10.5964/spb.v13i3.28046>
- De Houwer, J. (2020). Conditioning is more than association formation: On the different ways in which conditioning research is valuable for clinical psychology. *Collabra: Psychology*, 6(1): 2. DOI: <https://doi.org/10.1525/collabra.239>
- De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and

- merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*, 631-642.
- De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, *34*, 480-494.
- De Houwer, J., & Hughes, S. (in press). *The psychology of learning: An introduction from a functional-cognitive perspective*. Boston, MA: MIT Press.
- Delgado, M. R., Olsson, A., & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, *73*, 39-48.
- Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.
- Eelen, P. (2018). Behaviour therapy and behaviour modification: Background and development. *Psychologica Belgica*, *58*, 184–195.
- Eelen, P., & Vervliet, B. (2006). Fear Conditioning and clinical implications: What can we learn from the past? In M. G. Craske, D. Hermans, & D. Vansteenweegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 17-35). Washington, DC: American Psychological Association.
- Field, A. P. (2006). Is conditioning a useful framework for understanding the development and treatment of phobias? *Clinical Psychology Review*, *26*, 857-875.
- Field, A. P., & Nightingale, Z. C. (2009). What if Little Albert had escaped? *Clinical Child Psychology and Psychiatry*, *14*, 343-351.
- Hayes, S. C. (2004). Acceptance and Commitment Therapy, Relational Frame Theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, *35*, 639–665.

- Hughes, S., De Houwer, J., & Perugini, M. (2016). The functional-cognitive framework for psychological research: Controversies and resolutions. *International Journal of Psychology, 51*, 4-14.
- Hull, C. L. (1943). *Principles of Behavior: An introduction to behavior theory*. New York: Appleton-Century-Crofts.
- Jozefowicz, J. (2018). Associative versus predictive processes in Pavlovian conditioning. *Behavioral Processes, 154*, 21–26.
- Kent, G. (1997). Dental phobias. In G. C. Davey (Ed.), *Phobias: A handbook of theory, research and treatment* (pp. 107–127). Chichester, England: Wiley.
- Keren, G.B., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-systems theories. *Perspectives on Psychological Sciences, 4*, 533-550.
- Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154-172). New York: Oxford University Press.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- Lovibond, P. F. (2011). Learning and anxiety: A cognitive perspective. In T. R. Schachtman & S. Reilly (Eds.), *Applications of learning and conditioning* (pp. 104-120). Oxford, UK: Oxford University Press.
- Lubow, R.E. (1995). Latent inhibition in humans: Data, theory, and implications for schizophrenia. *Psychological Bulletin, 117*, 87-103.
- Mackie, J. L. (1965). Causes and conditionals. *American Philosophical Quarterly, 2*, 245-264.

- McAllister, D. E., & McAllister, W. R. (1991). Fear theory and aversively motivated behavior: Some controversial issues. In M. R. Denny (Ed.), *Fear, avoidance and phobias* (pp. 135–163). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- McLaren, I. P., Forrest, C., McLaren, R., Jones, F., Aitken, M., & Mackintosh, N. (2014). Associations and propositions: The case for a dual-process account of learning in humans. *Neurobiology of Learning and Memory*, *108*, 185-195.
- Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences*, *22*, 280-293.
- Mertens, G., & Engelhard, I. (in press). A critical review and meta-analysis of the evidence for unaware fear conditioning. *Neuroscience and Biobehavioral Reviews*.
- Mertens, G., Boddez, Y., Sevenster, D., Engelhard, I. M., & De Houwer, J. (2018). A review on the effects of verbal instructions in human fear conditioning: Empirical findings, theoretical considerations, and future directions. *Biological Psychology*, *137*, 49-64.
- Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta Psychologica*, *127*, 567-580.
- Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders. *American Psychologist*, *61*, 10-26.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009a). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*, 183-198.
- Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009b). Link-based theory creates more problems than it solves. *Behavioral and Brain Sciences*, *32*, 230-246.
- Moors, A., & De Houwer, J. (2006). Problems with dividing the realm of cognitive processes.

- Psychological Inquiry*, 17, 199-204.
- Mowrer, O. H. (1939). Stimulus response theory of anxiety. *Psychological Review*, 46, 553-565.
- Mowrer, O. H. (1947). On the dual nature of learning. A reinterpretation of 'conditioning' and 'problem solving'. *Harvard Educational Review*, 17, 102-148.
- Muris, P. & Field, A. P. (2010). The role of verbal threat information in the development of childhood fear. "Beware the Jabberwock!" *Clinical Child and Family Psychology Review*, 13, 129-150.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- Öhman, A., & Mineka, S. (2001). Fear, phobias and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522.
- Pearce, J. M., & Bouton M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, 52, 111-139.
- Poulton, R., & Menzies, R.G. (2002). Non-associative fear acquisition: A review of the evidence from retrospective and longitudinal research. *Behaviour Research and Therapy*, 40, 127-149.
- Rachman, S. (1977). The conditioning theory of fear acquisition: A critical examination. *Behaviour Research and Therapy*, 15, pp. 375–387.
- Raes, A. K., De Houwer, J., Verschuere, B., & De Raedt, R. (2011). Return of fear after retrospective inferences about the absence of an unconditioned stimulus during extinction. *Behavior Research and Therapy*, 49, 212-218.
- Reiss, S. (1980). Pavlovian conditioning and human fear: An expectancy model. *Behavior Therapy*, 11, 380-396.

- Rescorla, R.A. (1966). Predictability and number of pairings in Pavlovian fear conditioning. *Psychonomic Science*, 4, 383-384.
- Rescorla, R.A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151-160.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (Eds), *Classical conditioning II: current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- Smoller, J. W., Block, S. R., & Young, M. M. (2009). Genetics of anxiety disorders: the complex road from DSM to DNA. *Depression and Anxiety*, 26, 965–975.
- Surwit, R. S. (1972). *The anticipatory modifications of the conditioning of a fear response in humans*. (Unpublished doctoral dissertation). McGill University.
- Testa, T. J. (1974). Causal relationships and the acquisition of avoidance responses. *Psychological Review*, 81, 491-505.
- Thorndike, E.L. (1911). *Animal intelligence: Experimental studies*. New York: MacMillan.
- Treanor, M. (2011). The potential impact of mindfulness on exposure and extinction learning in anxiety disorders. *Clinical Psychology Review*, 31, 617–625.
- Vervliet, B., & Raes, F. (2013). Criteria of validity in experimental psychopathology: application to models of anxiety and depression. *Psychological Medicine*, 43, 2241-2244.
- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1-14.
- Wolpe, J. & Lazarus, A. (1966). *Behavior Therapy Techniques*. Oxford, UK: Pergamon

Press.

Wolpe, J., & Rachman, S. (1960). Psychoanalytical evidence: A critique based on Freud's case of Little Hans. *Journal of Nervous and Mental Disorders*, 131, 135-145.

Yule, W., Udwin, O., & Murdoch, K. (1990). The 'Jupiter' sinking: Effects in children's fears, depression and anxiety. *Journal of Child Psychology and Psychiatry*, 31, 1051-1061.

Author Note

Jan De Houwer, Ghent University, Belgium. The preparation of this paper was made possible by of Ghent University Grant BOF16/MET\_V/002. Correspondence should be addressed to Jan De Houwer, Ghent University, Henri Dunantlaan 2, B-9000 Ghent, Belgium. Electronic mail can be sent to [Jan.DeHouwer@UGent.be](mailto:Jan.DeHouwer@UGent.be) .

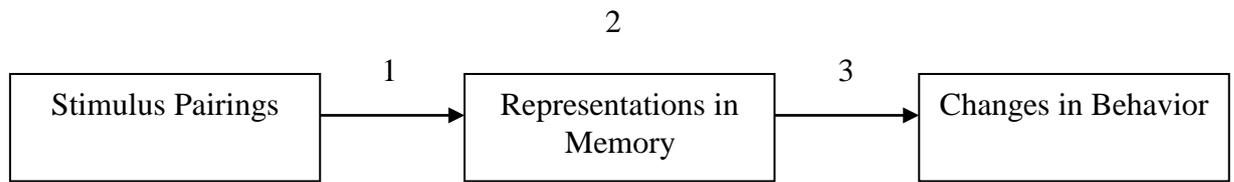


Figure 1. A schematic depiction of the three steps about which assumptions have to be made in theories of conditioning mechanisms.

Table 1. Overview of the empirical evidence regarding the claim that stimulus pairings involving negative events are a possible, necessary, and sufficient cause of anxiety disorders, including whether the evidence can be explained (YES), can be explained to some extent (YES BUT), or cannot be explained (NO) by different mechanistic accounts of conditioning.

| Topic                 | Evidence  | S-R Association Models | S-S Association Models | Propositional Models |
|-----------------------|---|------------------------|------------------------|----------------------|
| Stimulus Pairings are |   |                        |                        |                      |
| A Possible Cause      | - pairings can cause fears<br>in controlled studies     | YES                    | YES                    | YES                  |
|                       | - retrospective reports of<br>pairings that cause fears | YES                    | YES                    | YES                  |
| Not a Necessary Cause | - fears in the absence of<br>aversive stimulus pairings | NO                     | YES                    | YES                  |

|                               |  |         |         |     |
|-------------------------------|--|---------|---------|-----|
|                               | - genetic contribution to the development of fears                 | YES BUT | YES     | YES |
|                               | - fears result from observation and instructions                   | YES BUT | YES BUT | YES |
| Not Always a Sufficient Cause | - pairings with a traumatic event sometimes do not result in fears | YES BUT | YES     | YES |
|                               | - the effect of pairings is modulated by a host of variables       | YES BUT | YES     | YES |

---