



UvA-DARE (Digital Academic Repository)

Avoidance learning: a review of theoretical models and recent developments

Krypotos, A.M.; Eftting, M.; Kindt, M.; Beckers, T.

DOI

[10.3389/fnbeh.2015.00189](https://doi.org/10.3389/fnbeh.2015.00189)

Publication date

2015

Document Version

Final published version

Published in

Frontiers in Behavioral Neuroscience

[Link to publication](#)

Citation for published version (APA):

Krypotos, A. M., Eftting, M., Kindt, M., & Beckers, T. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, 9, Article 189. <https://doi.org/10.3389/fnbeh.2015.00189>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Avoidance learning: a review of theoretical models and recent developments

Angelos-Miltiadis Kryptos^{1,2}, Marieke Effting^{1,2}, Merel Kindt^{1,2} and Tom Beckers^{1,2,3*}

¹ Department of Clinical Psychology, University of Amsterdam, Amsterdam, Netherlands, ² Amsterdam Brain and Cognition, University of Amsterdam, Amsterdam, Netherlands, ³ Department of Psychology, KU Leuven, Leuven, Belgium

Avoidance is a key characteristic of adaptive and maladaptive fear. Here, we review past and contemporary theories of avoidance learning. Based on the theories, experimental findings and clinical observations reviewed, we distill key principles of how adaptive and maladaptive avoidance behavior is acquired and maintained. We highlight clinical implications of avoidance learning theories and describe intervention strategies that could reduce maladaptive avoidance and prevent its return. We end with a brief overview of recent developments and avenues for further research.

Keywords: avoidance, fear, anxiety, learning, neuroscience

Introduction

Avoidance of genuinely threatening stimuli or situations is a key characteristic of adaptive fear. People will typically not enter a building after a major earthquake nor approach a stray lion. At the same time, excessive avoidance in the absence of real threat can severely impair individuals' quality of life and may stop them from encountering anxiety-correcting information (Barlow, 2002). In such cases, avoidance loses its adaptive value and may transform into a maladaptive response. Maladaptive avoidance is in fact a central characteristic of a wide spectrum of mental disorders (World Health Organization, 2004; American Psychiatric Association, 2013). Individuals with Obsessive Compulsive Disorder (OCD), for instance, tend to avoid situations in which the potential for contact with contaminants is high (Rachman, 2004), Post-Traumatic Stress Disorder (PTSD) patients will try to avoid intrusive memories (Brewin and Holmes, 2003; Williams and Moulds, 2007), and social phobics will refuse to attend group gatherings (Bögels et al., 2010; Schneier et al., 2011).

Given the key role of avoidance in normal and disordered psychological functioning, it is critical to better understand the relevant conditions and psychological mechanisms responsible for the learning of avoidant reactions. Alas, although avoidance learning was once a central topic in basic psychological research, interest has waned since the 1970's, leaving important questions unanswered. Only recently has there been a resurgence of theoretical, experimental and clinical interest in the study of avoidance (see **Figure 1**). In the last years, new psychological theories of avoidance learning have been proposed (e.g., De Houwer et al., 2005; Lovibond, 2006) and avoidance is quickly becoming a topic of prime empirical interest not only in experimental psychology but also in clinical psychology and psychiatry as well as in behavioral neuroscience (see the present special issue). The latest edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; American Psychiatric Association, 2013) includes avoidance in several diagnostic criteria that previously referred to fear only. In parallel, recent years have brought rapid increases in our understanding of the brain processes involved in the learning

OPEN ACCESS

Edited by:

Richard J. Servatius,
Syracuse DVA Medical Center, USA

Reviewed by:

Christine A. Rabinak,
Wayne State University, USA
Hadley C. Bergstrom,
National Institutes of Health, USA

*Correspondence:

Tom Beckers,
Department of Clinical Psychology,
University of Amsterdam,
Weesperplein 4, 1018 XA Amsterdam,
Netherlands
T.R.J.Beckers@uva.nl

Received: 20 May 2015

Accepted: 06 July 2015

Published: 21 July 2015

Citation:

Kryptos A-M, Effting M, Kindt M and Beckers T (2015) Avoidance learning: a review of theoretical models and recent developments. *Front. Behav. Neurosci.* 9:189. doi: 10.3389/fnbeh.2015.00189

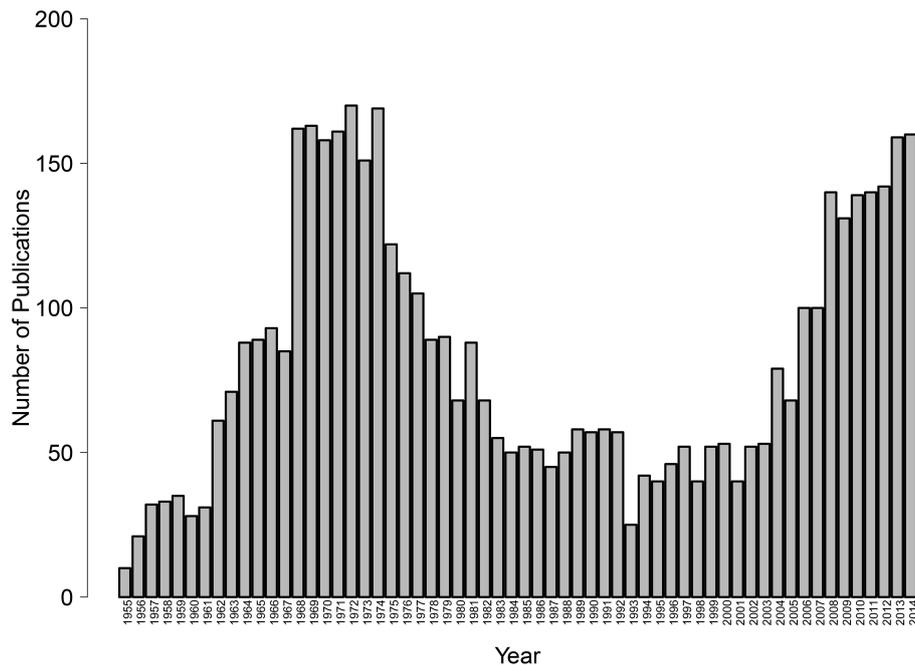


FIGURE 1 | Number of publications indexed in Thomson Reuters' Web of Science, research area psychology, that have the word "avoidance" in their title, by year, 1955–2014 (as of May 12, 2015).

(e.g., Delgado et al., 2009), expression (e.g., Cominski et al., 2014), and reduction (e.g., McCue et al., 2014) of avoidance behavior.

In this paper we review the main historical and modern theories of avoidance learning and present a set of principles of avoidance learning that integrate those theoretical propositions with the strongest experimental support. We also address the clinical implications of those principles and relate them to current and novel interventions for maladaptive avoidance such as in anxiety disorders or PTSD. Lastly, we consider recent findings from behavioral neuroscience.

The outline of the paper is as follows: We first describe how avoidance learning is studied in laboratory settings and how functionally similar behaviors can serve the *avoidance of* or the *escape from* an aversive event. Next, we discuss traditional theories of avoidance learning, including Mowrer's two-factor theory. In the third section we describe Bolles' (1970, 1971) Species-Specific Defense Reactions (SSDR) theory. We then review more recent theories of avoidance learning that address informational factors (e.g., expectancies) in avoidance. Next, we propose a set of principles for avoidance learning that incorporates the most well-validated propositions of the aforementioned theories. We end our review with suggestions for closer alignment between basic and clinical science and a few avenues for future research.

Laboratory Procedures for Studying Avoidance Learning

Avoidance learning procedures typically entail the cancelation of an impending aversive event by either the emission or inhibition

of an experimenter-designated response. In *active avoidance procedures*, for example, an antecedent stimulus is followed by an aversive event *unless* an experimenter-designated response is executed, a response that typically also terminates the antecedent stimulus. For example, dogs will learn to jump a barrier following the presentation of a light, previously associated with shock administration (Solomon and Wynne, 1953).

By contrast, in *passive avoidance procedures*, the aversive event occurs only if an experimenter-designated response is executed during the antecedent stimulus presentation. For example, in a standard passive avoidance procedure for rats, a rat is placed in a brightly lit compartment of a two-compartment box, with the second compartment being dark and the two compartments separated by a closed door (Venable and Kelly, 1990; Kaminsky et al., 2001). Given that rats have a preference for dark compared to lit environments (see Costall et al., 1989; Bourin and Hascoët, 2003), they will move to the dark compartment once the door is opened, an action that will be followed by shock administration. This procedure often results in the rats passively avoiding the shock by remaining in the light compartment on future occasions.

In avoidance procedures, the experimenter-designated response is not necessarily performed prior to the aversive outcome, but can also be performed in presence of it. In such cases, it would be more accurate to categorize the performed response as escape rather than avoidance. Escape responses involve distancing oneself from an *ongoing* aversive event while avoidance refers to behavior that causes the omission of a *forthcoming* noxious outcome, predicted by an antecedent stimulus (Bower and Hilgard, 1981; see also the distinction

between antecedent events and aversive outcomes in Lovibond and Rapee, 1993). Thus, what differentiates avoidance from escape is the proximity of threat (imminent vs. ongoing).

An elegant animal laboratory procedure to differentiate between avoidance and escape responses, especially relevant to behavioral neuroscience, is the elevated T-maze task (ETM; Pellow et al., 1985). The ETM has been primarily designed for testing rats' defensive reactions in innately fearful environments (i.e., open spaces and heights; Montgomery, 1955). It typically consists of three elevated arms with one arm surrounded by a wall (enclosed arm) and the other two being open (open arms). Initially, the rat is placed in the enclosed arm. While exploring the rat will eventually end up in the open arms. Following repeated trials, the rat will tend to remain longer in the enclosed arms after being placed there (i.e., passive avoidance) or run toward the enclosed arm after being placed in one of the open arms (i.e., escape).

By using this procedure, research has illuminated the differences in the neurobiology of avoidance and escape. Specifically, serotonin, an anxiogenic neurotransmitter relevant for defensive responses (Graeff, 2002), seems to play a different role in the two types of behaviors (Zangrossi et al., 2001), with serotonin administration enhancing avoidance and inhibiting escape (Graeff, 1991). That observation supports the argument that avoidance and escape may constitute diverse types of defensive behaviors, differently elicited as a function of the proximity of threat (imminent or ongoing), a hypothesis also in line with models associating defensive response selection to predatory imminence (Fanselow, 1994). Of note, the difference between escape and avoidance is also relevant for clinical practice. Specifically, Deakin and Graeff (1991) suggested that (passive) avoidance is mainly related to generalized anxiety disorder (GAD), where a threatening event is typically anticipated, and escape to panic disorder, where panic reactions could be considered as responses to an ongoing perceived danger (Shuhama et al., 2007). This suggestion has gathered partial support from pharmacological studies. It has been demonstrated that commonly prescribed anxiolytic drugs (e.g., diazepam) result in passive avoidance deficiencies while leaving escape behavior intact. On the contrary, cholecystokinin agonists, which typically invoke panic attacks, facilitate escape behaviors (Pinheiro et al., 2007; Graeff and Zangrossi, 2010). Taken together, avoidance and escape seem to be distinct subtypes of defensive behaviors, and that they might play a different role in mental disorders.

Active and passive avoidance and escape procedures have proved valuable for testing avoidance and/or escape learning in laboratory settings. Based on findings obtained with those procedures, theories have emerged that address the underlying psychological mechanisms. We now turn to a discussion of early theories of avoidance/escape learning in psychology.

Early Theories of Avoidance/Escape Learning and the Two-factor Theory

In the early days of psychology, learned avoidance was considered an example of a Pavlovian conditioned reflex (Bekhterev,

1907, 1913; Watson, 1916). Just like Pavlov's dogs would salivate upon the sound of a metronome previously associated with food administration (Pavlov, 1927), in the studies of Bekhterev (1913), a dog would flex its leg after the presentation of an antecedent stimulus, previously associated with shock administration (Herrnstein, 1969; Bolles, 1972). Since leg flexion would occur in the presence of the antecedent stimulus and prior to shock delivery, the acquired response was considered to reflect Pavlovian learning.

Nonetheless, two procedural characteristics differentiated the acquired responses from learned Pavlovian reflexes. First, what constituted the avoidance response (e.g., leg flexion) was usually an experimenter-defined voluntary response, whereas in Pavlov's experiment the learned response toward an initially neutral stimulus (i.e., salivation upon sound of the metronome) would typically consist of the automatic response toward an evolutionary relevant stimulus (i.e., salivation during food presentation; Unconditioned Stimulus or US). Second, the emitted response would lead to the cancelation of the impending event, making the (non-)presentation of the aversive stimulus dependent on the organism's response (Herrnstein, 1969). This procedural aspect is at odds with the standard Pavlovian procedure in which the presentation of food, or of any other US, would not depend on the animal's response (i.e., food would be presented independently of whether dogs salivated or not). Those procedural differences pointed to the potential operation of instrumental processes during avoidance learning, since in instrumental learning procedures an experimenter-defined action of the organism is necessary for outcome presentation or omission (Rescorla and Solomon, 1967). The potential involvement of instrumental processes, however, raised the question as to how avoidance responses are reinforced. Although one might intuitively argue that the source of reinforcement is the omission of the impending aversive event (i.e., the dogs flex their legs because this cancels the shock), assigning the cause of behavior to an event that has not yet occurred (i.e., shock administration) violated the dominant scientific principles of psychology at the time (i.e., the behaviorist paradigm; Watson, 1913).

A solution to that conundrum was offered in the *two-factor theory* formulated by Orval (Mowrer, 1951), who proposed that the performed response was reinforced by fear reduction (Hull, 1943). Specifically, Mowrer argued that as a result of Pavlovian fear conditioning (first factor), i.e., an antecedent stimulus (e.g., a tone) being associated with the administration of an aversive event (e.g., a shock), presentation of the antecedent stimulus will come to evoke fear. Subsequently, during the instrumental phase (second factor), escape responses that are emitted in the presence of the antecedent stimulus will be negatively reinforced by fear reduction, due to increased distance to or cessation of the antecedent stimulus. This idea was heavily inspired by the avoidance learning procedures used at the time, where avoidance responses led to the termination of the antecedent stimulus by locomotion (e.g., moving away from a shock area of a box) or by the antecedent stimulus being turned off. Of note, according to Mowrer, the omission of the aversive outcome event was to be regarded as a mere by-product of the performed CS escape behavior (Schöenfeld, 1950; Mowrer, 1960).

Two-factor theory quickly gained popularity in experimental psychology. By resorting to the concept of fear reduction during the instrumental phase, Mowrer's proposition was in line with the dominant drive reduction theories of the time (e.g., the drive reduction theory of Hull, 1943). Concurrently, by suggesting that the initial fear learning was based on Pavlovian processes, rather than drive reduction, the theory was better applicable to experimental data than the competing theory of Neal Miller (Miller, 1948), according to which all reinforcement during escape/avoidance learning originated from fear reduction.

Mowrer's theory has been used not only for explaining how maladaptive avoidance is acquired (Levis, 1981), but also as a basis for clinical interventions (Eysenck and Rachman, 1965). For example, in exposure therapy a patient is repeatedly confronted with a fearful situation or stimulus, in order to reduce that fear. It is commonly suggested that patients should be kept in the exposure situation until fear or anxiety levels have declined. This suggestion is rooted in two-factor theory and the notion that if the exposure session is terminated while fear levels remain high, the fear reduction caused by the termination of the session could promote escape or avoidance of similar situations in the future (Mathews et al., 1981; Emmelkamp, 1982; see next section for arguments against this notion).

Despite its wide influence on basic research and clinical science, two-factor theory had trouble explaining later data (see Rescorla and Solomon, 1967; Herrnstein, 1969; McAllister and McAllister, 1991, for extended discussions of two-factor theory). We now turn to some of the key criticisms against the two-factor theory.

Criticisms Against the Two-factor Theory

One of the strongest criticisms against the two-factor theory concerned the purported role of fear in motivating the emission of a learned escape/avoidance response. According to Mowrer's proposition, escape/avoidance is motivated by high fear levels. This notion implies that no such actions should be performed in the absence of fear. One of the first experiments to show that this may not be true was done by Solomon et al. (1953). In their experiment, dogs were first trained to jump across a barrier in response to the sounding of a buzzer previously paired with shock. The dogs then received a fear extinction treatment in which the buzzer was repeatedly presented without shock. Such extinction procedure typically leads to the reduction of fear levels (Hermans et al., 2006). If, as assumed by Mowrer, it is fear that motivates escape/avoidance, it would be expected that following Pavlovian fear extinction, dogs would also stop performing the avoidance response. The results contradicted this hypothesis: Dogs continued to jump upon sounding of the buzzer, even when the shock device had long been turned off permanently.

The observation that fear may not be necessary for avoidance has clear clinical implications. As mentioned earlier, patients are typically prevented from prematurely terminating exposure out of concern that the fear reduction resulting from termination of the session could otherwise serve as negative reinforcement for escape (Eysenck and Rachman, 1965). Experimental data, however, indicate that patients undergoing exposure therapy show similar clinical improvement regardless of whether they

ended exposure while fear levels were high or low (De Silva and Rachman, 1984; Rachman et al., 1986). Taken together, both experimental data and clinical findings suggest that fear may sometimes, but not always, be involved in the maintenance of avoidance, and as such, fear and avoidance may not always "synchronize" with each other (Rachman and Hodgson, 1974).

Two-factor theory also had trouble explaining how avoidance can be acquired in the absence of an explicit antecedent stimulus. Specifically, in *unsigned* avoidance procedures (Sidman, 1953a, 1962; see Lázaro-Muñoz et al., 2010; McCue et al., 2014 for more recent examples), rats learn to avoid shocks presented at fixed time intervals, in the absence of a discrete antecedent stimulus (Sidman, 1953a,b; see Hassoulas et al., 2013, for examples in humans). In its initial form, two-factor theory assumed the operation of explicit antecedent stimuli during the Pavlovian and the instrumental phases, stimuli that during unsigned procedures appear to be absent.

A potential explanation for the observation of unsigned avoidance procedures is that although not explicit, warning stimuli may still be present in a "silent" form. Temporal and proprioceptive stimuli (e.g., the passage of time), for example, could be associated with the aversive outcome (Schöenfeld, 1950; Dinsmoor, 1977). Subsequently, those stimuli could signal the presentation of an aversive event (for an alternative account centering on the role of US omission during avoidance learning see Herrnstein, 1969).

By assuming that avoidance is based on reinforcement learning, the two-factor theory also failed to explain how avoidance is acquired in naturalistic settings, where the first encounter of an organism with danger could prove fatal (Bolles, 1970; Osada et al., 2014). Similarly, when it comes to maladaptive avoidance, patients do not always report a direct traumatic event as the source of their symptomatology (Rachman, 1990). An explanation for those observations is that avoidance need not always be acquired through direct experience but can be acquired via other pathways as well (Rachman, 1991, 1977; Olsson and Phelps, 2004, 2007). Those pathways include vicarious learning (e.g., learning to be afraid of dogs after observing someone being afraid of a dog) and instructional learning (e.g., learning to be afraid of a dog after someone suggesting that dogs often attack people; Bandura and Rosenthal, 1966; Rachman, 1977). Recent evidence shows that avoidance learning can be achieved even more indirectly, such as through *symbolic generalization*. One demonstration of that was presented by Augustson and Dougher (1997). They first trained individuals to categorize eight different stimuli (i.e., A1, B1, C1, D1, A2, B2, C2, and D2) into two arbitrary categories (i.e., 1 and 2), using standard conditional discrimination procedures (Sidman, 1987): Participants first saw a target stimulus (e.g., A1) and were then asked to choose one from three stimuli presented on screen. One of those stimuli was arbitrarily assigned to the same category (e.g., B1), one stimulus to the other category (e.g., C2) and another one was irrelevant (e.g., X). Participants' task was to learn to choose the stimulus from the same category as the target stimulus (e.g., B1). In case of a correct response, the word "correct" was presented and in case of an incorrect response, the word "wrong" was presented. For example, when stimulus A1 was presented,

selecting the stimulus that belonged to the same category (e.g., B1, C1 or D1) was positively reinforced whereas the selection of any other stimulus (e.g., B2 or X) was punished. Such a procedure typically results in people learning that the stimuli within each category are functionally equivalent. During a subsequent fear conditioning phase, B1 was paired with shock and B2 with the absence of shock. An instrumental phase followed, where participants could avoid shocks with a button press. Critically, results showed that, in addition to performing more avoidance responses in the presence of B1 than B2, participants also performed more avoidance responses to (unreinforced) presentations of C1 and D1 (the stimuli arbitrarily related to B1) than in the presence of C2 and D2 (the stimuli related to B2). Apparently, avoidance responses to B1 symbolically generalized to C1 and D1, which had been previously trained as equivalent to B1, without an avoidance schedule being trained for those stimuli. Those findings have recently been replicated and extended (Dymond and Roche, 2009; Dymond et al., 2011). Taken together, direct traumatic experiences may not always be necessary for the acquisition of avoidance, a proposition that is in line with contemporary views on fear learning and psychopathology (Mineka and Zinbarg, 2006).

Another argument against the two-factor theory regards the extent to which Pavlovian and instrumental learning are both necessary for the acquisition of avoidance. As a result of Pavlovian fear conditioning, a previously neutral cue will come to evoke various fear responses (e.g., enhanced physiological arousal; Beckers et al., 2013). According to emotion theories, action tendencies are an essential component of emotions (Lang, 1985; Frijda, 1988). As such, “to fear is to want to avoid.” Therefore, it could be assumed that the tendency to avoid a stimulus or situation may be acquired via purely Pavlovian learning, without instrumental reinforcement. We have recently tested this assumption. Following a differential fear conditioning procedure, during which pictures of one geometrical object were always followed by shock (CS⁺) whereas pictures of another object were never followed by shock (CS⁻), participants were faster to avoid the CS⁺ and approach the CS⁻ than vice versa in a symbolic approach-avoidance reaction time task (AAT; Krypotos et al., 2014b). Crucially, shock electrodes were detached from participants’ hands during the AAT, and responses had no influence on the duration or presence of the CS, eliminating any instrumental basis for avoidance. Those results suggest that avoidance tendencies can be acquired via mere Pavlovian associations, in the absence of instrumental learning. Such acquisition of motor responses toward a CS in absence of instrumental learning has long been established in the appetitive domain. In auto-shaping procedures (Brown and Jenkins, 1968), for example, pairings of a visual CS with food would typically result in the animal (e.g., a rat or a pigeon) producing consumption responses toward the CS (e.g., licking or pecking, respectively), despite those responses being irrelevant for food presentation.

Lastly, recent findings suggest that avoidance can be evoked by the identification of predator related stimuli (e.g., smells) even in absence of a previous encounter with the predator. Specifically, mice (Osada et al., 2013) and deer (Osada et al., 2014)

tend to avoid (or emit other defense-like behaviors) areas where the active components of predators’ urine odors are presented, without any previous experience with that specific predator. An explanation for those findings is that such avoidance may be the result of “evolutionary memory” (Provenza, 1995) and as such, avoiding such stimuli (e.g., odors or blood) does not require the learning of associations between a stimulus (i.e., smell) and the aversive event. Of note, similar effects are yet to be demonstrated in humans.

To sum up, although influential, the two-factor theory of Mowrer proves unable to account for a series of experimental results and clinical observations. In response to those shortcomings of the two-factor theory, alternative theories have been proposed. One of those, with major influence in the experimental field, is the SSDR theory of Bolles (1970, 1971).

Species-specific Defense Reactions

It has long been demonstrated that in Pavlovian conditioning, evolutionary relevant stimuli (e.g., spiders) are more readily associated with an aversive event (e.g., shock) than non-evolutionary relevant stimuli (e.g., flowers; Öhman and Mineka, 2001). Similarly, in the area of taste aversion, associations between tastes and induced sickness are acquired more readily than between audio-visual cues and sickness whereas audio-visual cues are associated more readily with shock than tastes are (the *Garcia effect*; Garcia et al., 1955; Davis and Riley, 2010).

An explanation for those findings is that by being wired to preferentially associative aversive events (i.e., shock and sickness) with phylogenetically relevant stimuli (i.e., spiders and tastes), an organism is better prepared to learn about likely cues for danger. Such ability would equip the organism with an evolutionary advantage for surviving potentially harmful cues or situations. An interesting question is whether individuals are similarly predisposed to associate the cancelation or termination of an aversive event with particular behavioral responses.

This seems to be the case. Rats will learn much more rapidly to avoid an aversive outcome by running—a behavior commonly used for avoiding a predator—than by moving their tail (Maatsch, 1959; Meyer et al., 1960; Theios et al., 1966; Masterson, 1970). Observations such as those inspired Bolles in his formulation of the Species-Specific Defense Reactions (SSDRs) theory. According to this theory, under a state of fear, organisms are phylogenetically predisposed to emit specific types of responses (e.g., fleeing, freezing, or fighting) rather than others. Bolles went one step further to suggest that for such responses, reinforcement learning is actually unnecessary; the organism just needs to learn that a stimulus predicts an aversive outcome for it to elicit an SSDR (Bolles, 1970, 1971).

The SSDR theory could indeed explain the fast acquisition of specific avoidance responses that served the evolutionary survival of the organism. Nonetheless, sometimes, an avoidance response has to be acquired that does not belong to an organism’s SSDR repertoire (Crawford and Masterson, 1982). Rats, for example, can learn to avoid an aversive outcome by rearing on a wheel, a response that arguably does not usually serve survival purposes, although at a much slower rate than learning to avoid that same

outcome by running on the wheel (Bolles and Grossen, 1969). In those cases, it is suggested that although under a state of fear SSDRs will initially be performed, those SSDRs will be punished by the occurrence of the aversive event, allowing for non-SSDRs to subsequently emerge. In other words, non-SSDRs would not (or not merely) be negatively reinforced, as Mowrer suggested, but they would arise because of SSDRs being positively punished by the presentation of the aversive event.

To sum up, Bolles provided a theory that could explain why some avoidance responses are acquired more readily than others, and why reinforcement may often be unnecessary for the acquisition of avoidance. Nonetheless, the theory has limitations. For one, fear states are on a continuum, ranging from low to high levels, making it difficult to define at which specific level the restriction of the behavioral repertoire to SSDRs occurs (Masterson and Crawford, 1982). Also, by rejecting reinforcement as a source for SSDRs, Bolles' theory makes it hard to explain findings suggesting that bar pressing is acquired faster when it leads to access to a safe place than when it merely leads to US omission (Crawford and Masterson, 1978). Thus, although SDDR theory provided answers to important avoidance learning questions, it fails to accurately define the conditions under which avoidance learning will occur and to provide a comprehensive account for all instances of avoidance acquisition.

Informational Factors in Avoidance Learning

The theories we have discussed so far provided functional explanations of avoidance learning, denying any role for cognitive or informational factors in the interpretation of avoidance acquisition. As such, those theories were in harmony with the dominant functional accounts of learning of the behaviorist paradigm. However, those functional accounts of learning failed to explain a series of laboratory phenomena such as that of *blocking* (i.e., impaired CS1-US acquisition if CS1 is paired with the US in compound with a CS2 that has been previously paired with the US in itself; Kamin, 1956, 1967, 1969) and *conditioned inhibition* (i.e., learned inhibitory responses toward CS1 if a CS1–CS2 compound is repeatedly presented without the US while CS2 is paired with the US when presented in itself; Rescorla, 1969). Such phenomena challenged traditional associative theories according to which CS-US contiguity is sufficient for Pavlovian acquisition (Rescorla, 1972; Miller et al., 1995).

As a result, a general shift in theories of learning has been observed, such that informational factors (e.g., outcome expectancies or stimulus surprisingness) started to be considered as potential explanations of acquired behavior (Rescorla, 1972).

One theory of learning with a lasting influence that was developed in that period, is the Rescorla-Wagner model (RWM; Rescorla and Wagner, 1972; Wagner and Rescorla, 1972). The basic premise of the RWM is that the rate of conditioning to a CS depends on whether the ensuing presentation of the US is surprising or not. If the CS did not elicit an accurate prediction of the (non-)occurrence of the US (negative or positive

prediction error), learning about the CS occurs; if it did, no learning occurs. This model clearly deviates from earlier theories of learning in that it recognizes the role of informational factors (i.e., predictions) in conditioning, and despite justified criticisms (Miller et al., 1995), it is a model with great heuristic and predictive value (Beckers and Vervliet, 2009).

In the next section we present avoidance learning models that, like the RWM for Pavlovian learning, rely on informational factors to account for avoidance learning.

The Role of Safety Signals

In our examples of Pavlovian fear conditioning, we have so far discussed situations in which an antecedent stimulus signals an aversive event. It might be assumed that in such situations, knowledge is acquired about the relation between the antecedent stimulus and the aversive event only. However, if there are also signals present that predict the *absence* of an otherwise expected aversive event (i.e., safety signals), those will be learned about as well, and fear responses will be inhibited in the presence of those stimuli (conditioned inhibition, see above).

A relevant question then is whether, in addition to reducing fear levels, safety signals might also be able to inhibit avoidance behavior. This hypothesis was tested in a seminal study of Rescorla and LoLordo (1965), who showed that dogs would learn to jump a small barrier after the presentation of a tone previously associated with shock but would withhold avoidance upon the presentation of another tone previously associated with the absence of shock (see Weisman and Litner, 1969, for a replication). Those findings indicate that during avoidance learning, knowledge is acquired not only about stimuli that signal an aversive event, so that those stimuli come to elicit defensive behaviors (e.g., avoidance, escape), but also about stimuli that signal the absence of a forthcoming aversive event, with the latter stimuli inhibiting defensive behaviors.

Assigning a role to safety stimuli in avoidance learning allows to explain a series of data that previous theories could not account for. For example, it has been shown that avoidance can be acquired readily even if the CS is not terminated upon the performance of the experimenter-designated avoidance response (Soltysik et al., 1983; Avcu et al., 2014). This observation contradicts two-factor theory, according to which the termination of the antecedent event is necessary for fear reduction. A safety signal account can explain those data, by assuming that changes in contextual or internal cues upon performance of the experimenter-designated response can serve as safety signals. The reinforcing value of safety cues can also explain why avoidance is acquired faster if the avoidance response leads to a safe environment than if it does not (Crawford and Masterson, 1978; see also Morris, 1975; Kim et al., 2006, for related evidence).

Another observation partially supportive of a safety signal account is that rats prefer a box compartment in which a light is presented before the administration of an (unavoidable) shock, over a compartment in which shocks occur unannounced (Lockard, 1963). According to a safety signal account, the predictability of the aversive event makes the situation less unpleasant because the absence of the warning stimulus signals

that the situation is safe (Seligman, 1968; Seligman et al., 1971), a view also in line with the RWM approach to inhibitory learning.

The safety signal account can also help to explain why exposure therapy does not necessarily lead to a reduction of avoidance. Specifically, it has been suggested that while undergoing exposure therapy, patients will try to reduce their unpleasantly high fear levels by engaging in overt or covert safety behaviors that reduce those fear levels (“within situation safety”; Wells et al., 1996; Salkovskis et al., 1999). Such safety behaviors may involve the generation of safety signals to reduce fear levels. For instance, individuals who are afraid to experience a panic attack while flying, might endure traveling in an airplane as long as they can carry anxiolytics with them. Having such medication in their pocket can serve as a signal that they can avoid a potential panic attack. No matter how helpful such strategy is in reducing momentary fear levels, the engagement in safety behaviors typically preserves the dysfunctional belief that the situation is inherently dangerous. By implication, on future occasions where they cannot engage their customary safety behaviors, individuals may revert into avoidance behavior in those situations.

The Cognitive Theory of Seligman and Johnston

One of the most influential theories of avoidance learning, which explicitly addressed the role of informational factors, is the cognitive theory of Seligman and Johnston (1973).

In spite of its name, the theory actually contains both a cognitive and an emotional component. The cognitive component revolves around the assumption that human and non-human animals would prefer not receiving an aversive stimulus over receiving one. The cognitive component also contains the notion that as a result of an avoidance learning schedule, humans or animals would learn to expect an aversive stimulus if an avoidance response is not performed and not to expect an aversive stimulus if an avoidance response is performed. The emotional component mainly refers to the Pavlovian fear responses that develop during an avoidance learning procedure as previously described by Mowrer and others.

By incorporating a role for expectancies in avoidance learning and maintenance, the cognitive model could account for data not easily explained by two-factor theory. For example, in the experiment of Solomon et al. (1953) mentioned previously, dogs would keep jumping a barrier subsequent to a buzzer presentation previously associated with shock, even when the experimenter stopped shock administration. Since shocks no longer followed buzzer presentation, it would be expected that fear would extinguish (see Pavlovian extinction). According to two-factor theory, fear motivates escape/avoidance responses, and since fear was assumed to have extinguished, it should be expected that avoidance would diminish as well, a hypothesis that was at odds with the data. Those data, however, can be explained by the cognitive theory. According to it, avoidance is maintained despite the potential reduction in fear levels because by jumping the barrier, the animal does not directly experience that the antecedent stimulus is not followed by the aversive outcome. As such, the expectancy that an aversive event would occur after the

presentation of the antecedent stimulus if an avoidance response were not emitted, is preserved.

Disconfirmation of expectancies, on the other hand, could explain why avoidance diminishes when exposure is combined with *response prevention* (Baum, 1966, 1970). During such schedule, individuals encounter a phobic stimulus while the execution of all escape responses is blocked. As a result, avoidance responses are typically thwarted when the individual encounters the antecedent stimulus in the future. This reduction in avoidance response execution, not achieved by traditional exposure programs, can be explained by considering that by not executing the escape response during the antecedent stimulus presentation, patients can realize that the expected noxious event was not to occur anyway. This then, removes the need to execute any defensive reaction during future encounters with the antecedent stimulus.

The cognitive model is not without limitations, some of which Seligman and Johnston noted themselves (Seligman and Johnston, 1973, pp. 100–101). One limitation is that the theory implicitly treats all types of avoidance responses as equivalent. Research on SSDRs suggests that this notion is untenable. Also, as noted by Lovibond (2006), the model remains silent as to how fear and expectancies relate to each other¹. As a result, it cannot explain experimental observations such as the return of extinguished fear when the performance of the avoidance response is prevented (Solomon et al., 1953).

Negative Occasion Setter Account

Informational factors also play an important role in the avoidance learning theory of De Houwer and colleagues (De Houwer et al., 2005; Declercq and De Houwer, 2008). According to those authors, an avoidance response serves as a signal that a CS is not going to be followed by an aversive event, which in associative learning language is called a “negative occasion setter” (see Holland, 1992; Schmajuk and Holland, 1998, for reviews on occasion setting).

Functionally, in negative occasion setting experiments, an antecedent stimulus (e.g., a sound; CS) is followed by an outcome (e.g., a shock; US) unless another stimulus X is presented (the occasion setter, OS; e.g., a light). Accordingly, the antecedent stimulus is going to be followed by an outcome only when the occasion setter is absent and vice versa. Stimulus X thus helps to disambiguate the relationship between the antecedent stimulus and the forthcoming outcome.

Negative occasion setting can be translated to clinical situations. It could be argued, for example, that the engagement in safety behaviors, such as avoidance, serves as a signal that a specific phobic stimulus or dreaded situation is not going to result an aversive event. Going back to our earlier example, the *administration* of an anxiolytic pill could signal that being in an airplane is not going to be accompanied by a panic attack. Such a proposition is different from the safety signal account. According to this latter account, a specific stimulus is supposed

¹Notably, and in contrast to more recent models of avoidance learning (e.g., the expectancy model; Lovibond, 2006), Seligman and Johnston (1973) did not assign a mediating role to expectancies for Pavlovian learning, although they did not fully exclude such possibility either (see footnote 14 in Seligman and Johnston, 1973).

to predict the absence of an aversive event directly. According to the negative occasion setter account, however, the presence of a specific event (e.g., avoidance) predicts that the otherwise valid relation between a threatening stimulus and an aversive event does not hold.

De Houwer and Declercq tested whether an avoidance response functions as a negative occasion setter by comparing the properties of avoidance behavior to properties of negative occasion setters identified in the Pavlovian literature (Holland, 1992; Schmajuk and Holland, 1998). Those properties are (a) *modulation* (i.e., CRs toward the CS are stronger in the absence than in the presence of the OS), (b) *resistance to counterconditioning* (i.e., CRs toward the CS are attenuated by the OS even if the OS has been paired with the US itself), and (c) *selective transfer* (i.e., the OS will modulate responding to other CSs that have been subject to modulation before; Holland, 1992). The experiment of De Houwer et al. (2005) was at follows.

In the first phase of their experiment, stimuli A and B were always followed by a US (shock) whereas a third stimulus C was followed by a US 50% of the times. In the second phase, the US could be prevented by pressing key X when the A stimulus was present and by pressing key Y when the B stimulus was present. The third phase consisted of the presentation of all the events occurring in the previous phases, in addition to trials in which the US occurred upon pressing the X key (X-US trials). In the crucial test phase, individuals were presented with all possible stimulus combinations, without the US ever occurring, and they were asked to rate their expectancy of a US.

Results from the test phase confirmed that avoidance responses exhibit the properties of negative occasion setters. First, participants reported higher US expectancies when the avoidance response was not available (i.e., A and B test trials) than when it was (i.e., AX and BY trials; modulation). This result shows that the function of the antecedent stimulus as a predictor of a negative outcome is dependent on the availability of the avoidance response, a finding at odds with two-factor theory but in line with other theories, such as Seligman and Johnson's cognitive model. Second, modulation was not affected by trials in which avoidance responses were followed by the US (i.e., the X-US trials; resistance to counterconditioning). This finding is in partial contrast with both the safety signal hypothesis and the cognitive model, given that the avoidance response has now become a predictor of the aversive event rather than of its omission. Third, participants would generalize this trained modulation to new stimuli, particularly those stimuli that had been involved in avoidance training (i.e., higher modulation for the AY and BX trials than for the CX or CY trials; selective transfer). Other models cannot easily account for this selectivity in trained modulation.

Despite subsequent replication and extension of those findings (Declercq and De Houwer, 2008), more recent evidence (e.g., Declercq and De Houwer, 2011) argues against the negative occasion setting account, as the property of selective transfer could also be explained by the lesser reinforcement of C compared to A and B during the first phase of the experiment. Still, despite this limitation, insights from the negative occasion

setting account could prove clinically important. Current therapeutic techniques mainly target the relation between the avoidance response and the omission of an unpleasant event. The negative occasion setting account assumes that there is a hierarchical structure in avoidance learning, with the avoidance response disambiguating the relation between the antecedent stimulus and the omission, or presentation, of the aversive event (see Declercq and De Houwer, 2009, for relevant evidence). If that is the case, interventions should perhaps focus more on challenging patients' beliefs about whether their avoidance response leads to the omission of an otherwise to-be-expected noxious outcome in the presence of some antecedent stimulus, rather than on their beliefs as to whether the antecedent stimulus is a reliable predictor of the noxious event (Declercq and De Houwer, 2009).

Lovibond's Expectancy Model

The most recent avoidance learning theory to include informational factors is the expectancy model of Lovibond (2006), which is basically an extension of the cognitive model of Seligman and Johnston (1973).

The expectancy model agrees with the idea that avoidance is acquired by a combination of Pavlovian and instrumental learning processes. It also accepts Seligman and Johnson's notion that during the instrumental phase, knowledge is acquired about the effects of avoidance (e.g., the *omission* of an expected unpleasant event) as well as non-avoidance (i.e., the *presentation* of an expected unpleasant event). Lastly, it aligns with the safety signal account in that it accepts that during the presentation of safety stimuli avoidance behavior is inhibited. However, the expectancy model rejects the notion that safety signals serve as positive reinforcers of the emitted response.

A major deviation of the expectancy model from the aforementioned theories is the assumption that both Pavlovian and instrumental learning are based on propositional knowledge. According to the propositional learning account, *all* learning reflects higher-order, reason-based processes, whereas earlier associative learning theories, explicitly or implicitly, rely on automatic association formation as the mechanism of learning. As such, the expectancy model assumes that avoidance learning is achieved by the accumulation of explicit knowledge about all the stimulus contingencies (e.g., that a CS is followed by a US during a Pavlovian phase) involved in avoidance learning protocols. In the same vein, the expectancy account assumes that expectancies play a crucial role not only in instrumental learning (as hypothesized by Seligman and Johnson) but in Pavlovian conditioning as well.

Most of the data we have presented so far can be accommodated by the expectancy model. Similarly to the model of Seligman and Johnston (1973), it can explain how avoidance is acquired, how it can be maintained despite Pavlovian extinction, and why response prevention during extinction will reduce avoidance. In addition, by reserving a central role for expectancies in Pavlovian learning, the expectancy model can better account for the relation between fear and avoidance. The expectancy model can explain, for example, why fear returns during response prevention, because in the absence of the

avoidance response the aversive event is to be expected, and outcome expectancy is what generates fear.

Lastly, although not explicitly mentioned by Lovibond (2006), the fact that avoidance relies on propositional knowledge also allows for the acquisition of avoidance via pathways other than direct experience, rendering the account able to accommodate observations of avoidance acquired through observation, instruction, and symbolic generalization (see above).

In summary, Lovibond's expectancy model is a hybrid of earlier theories that accounts for the majority of experimental findings. Still, the model does not readily explain why some avoidance responses are learned more readily than other. It also cannot account for data in which escape/avoidance responses are acquired and expressed in the absence of instrumental reinforcement (Krypotos et al., 2014b). Lastly, it is debatable whether all elements of the expectancy model, and specifically the notion that propositional knowledge is a prerequisite for learning, could be generalized to non-human animals (Castro and Wasserman, 2009; but see Beckers et al., 2006).

Principles of Avoidance Learning

We have reviewed early and more recent theories of avoidance learning, describing the strengths and limitations of each theoretical account. In this section we synthesize the strongest points of each theoretical account, distilling a set of principles of avoidance learning that collectively can account for the majority of existing data. We then consider recent experimental findings on the acquisition of avoidance tendencies in the light of those principles and discuss their potential clinical implications.

For illustration, we will describe the various phases of avoidance learning by reference to a modified version of the procedure used in the third experiment of Rescorla and LoLordo (1965) that we have briefly described above (see the section on safety signals). We use this experiment because it resembles a typical active avoidance learning procedure while also illustrating the potential inhibition of avoidance responses by safety stimuli.

In the experiment, dogs were placed in a large two-compartment box, separated by a small barrier. The experimental procedure included the presentation of a short tone (threat signal; CS^+) always followed by shock administration (US) and the presentation of a long tone (safety signal; CS^-) never followed by shock administration. Dogs could avoid shock administration by jumping across the barrier (i.e., the avoidance response) during the presentation of the CS^+ . As described, such procedure would typically lead to reliable execution of the avoidance response during the CS^+ and absence of such responses during the CS^- . We now turn to a step-by-step description of how those acquired responses can be theoretically explained.

In line with most of the theoretical accounts we have presented above, we propose that avoidance learning involves a Pavlovian component. Specifically, we assume that following a Pavlovian conditioning procedure, the presentation of the CS^+ will elicit the expectancy of an aversive outcome, whereas the CS^- will elicit the expectancy of absence of the aversive outcome. As a result of those expectancies, fear responses will be evoked

(e.g., physiological arousal, avoidance tendencies) in the presence of the CS^+ , whereas responses of relief (e.g., physiological relaxation, approach tendencies) are expressed in the presence of the CS^- .

In extension to the models we have presented above, we maintain that the described Pavlovian component is sufficient for evoking avoidance tendencies. As such, we essentially treat avoidance tendencies as conditioned responses, similar to physiological arousal or subjective apprehension, elicited by the presentation of a threat cue. Such an explanation fits the data of Krypotos et al. (2014b) where avoidance tendencies were expressed in absence of any instrumental reinforcement. Going back to our example, we propose that dogs will acquire a tendency to exhibit avoidance in the presence of the CS^+ sound prior to any instrumental learning.

Although acquired, we propose that those avoidance tendencies need not be translated into overt behavior (Strack and Deutsch, 2004). At the same time, we argue that if expressed, they will take the form of an SSDR (Bolles, 1970, 1971). In other words, the dogs in our example are expected to start jumping across the barrier upon presentation of the CS^+ , prior to the operation of any instrumental processes. Such an idea might explain the rapid learning of specific types of responses that serve avoidance, relative to other responses (Bolles and Grossen, 1969). In such cases, however, it would be more accurate to classify those responses as escape responses, expressed in a reflex-like manner, as their primary aim is to distance the organism from the CS^+ (i.e., *CS escape*).

Despite the rapid acquisition of SSDRs by mere Pavlovian learning, the maintenance of those SSDRs, or the learning of non-SSDRs, will in fact depend on instrumental learning. As such, we adhere to the spirit of Mowrer's two-factor theory and more recent reformulations in assuming that avoidance learning depends on both Pavlovian and instrumental processes. Yet, we propose that instrumental learning plays a different role depending on whether the to-be-learned avoidance response belongs to an organism's SSDR repertoire or not. Revisiting our example, dogs are expected to start jumping across the barrier upon presentation of the CS^+ , with such response being maintained if the US stops being administered (i.e., negative reinforcement). If, however, such a response does not lead to cancelation of the US, that particular avoidance response will cease to be performed. For illustration, assume that the experimenter-designated response changes from jumping (i.e., an SSDR) to rearing (i.e., a non-SSDR). As jumping does not result in omission of the US, another SSDR response will be performed (e.g., running within the same compartment). However, since that novel response is followed by US presentation as well, the dog will cease that response too. Eventually, when all SSDRs have been aborted, the dog may incidentally perform a non-SSDR (i.e., rearing). This response is then hypothesized to be maintained through negative reinforcement by US omission. Although such treatment of non-SSDRs bears similarities with Bolles' (1971) theory, in the sense that non-SSDRs are supposed to emerge after SSDRs are no longer performed, it also deviates from SSDR theory in that it accepts a role for reinforcement learning in the maintenance of SSDRs and the learning of non-SSDRs.

In contrast to existing models, we maintain that for the instrumental component of avoidance learning, there are multiple sources of reinforcement. Those sources include (1) the non-occurrence of an expected aversive event after performance of the avoidance response; (2) the reduction of fear due to escape from the CS⁺, at least in the initial stages of reinforcement learning; (3) the occurrence of discrete stimuli upon performance of an avoidance response that become safety signals. Although this is a bottom-up suggestion, we believe that by assuming multiple sources of reinforcement, the described principles fit a larger amount of data.

In line with both the cognitive and the expectancy model, we propose that during instrumental learning, knowledge is acquired that (1) the performance of an avoidance response in the presence of the antecedent stimulus leads to the omission of the otherwise expected aversive event and; (2) the non-performance of an avoidance response in the presence of the antecedent stimulus leads to the administration of the aversive event. Those two theoretical accounts also assume that an individual or an animal will prefer not encountering an aversive event (here a shock) over encountering one.

Lastly, we take no position here in the debate as to whether avoidance learning (and in particular the Pavlovian component thereof) is based exclusively on propositional knowledge or also relies on non-propositional association formation processes. The recent literature provides strong evidence to support a single-process propositional account of learning (see De Houwer, 2009; Mitchell et al., 2009) as well as intriguing evidence to the contrary (e.g., Sevenster et al., 2014).

Clinical Implications

The set of principles described above can help to detect sources of maladaptive avoidance. By building on those principles, clinical interventions for maladaptive avoidance can also be improved.

For instance, given that avoidance learning is assumed to depend on multiple pathways, the causes of maladaptive behavior should not be assumed to be limited to direct experience with threat. Maladaptive avoidance may just as well be the result of instructions (e.g., kids being instructed to stay away from dogs leading to a phobia toward dogs) or vicarious learning (e.g., watching a plane crash on television as an onset of avoidance of flying). Another potential pathway of avoidance learning concerns symbolic generalization (e.g., avoidance of thoughts of contamination that only symbolically relate to actual contamination; Dymond et al., 2012).

Independent of the pathway, however, patients will often be able to articulate purported relations between the antecedent stimuli and the presentation or omission of an aversive outcome, as well as the relationship between the cancelation of that aversive outcome and the performance of an avoidance response (Beck et al., 2005). Given that those relationships often reflect misconceptions of the patient (e.g., that social interactions will always end in embarrassment), those beliefs should be challenged and modified during clinical interventions, a suggestion in line with current cognitive-behavior therapy protocols.

In terms of interventions, a common therapeutic technique for anxiety disorders and phobias is exposure therapy, which

entails confronting the individual repeatedly with the phobic stimulus. Clinical results suggest that exposure is necessary, but not sufficient, for the reduction of avoidance. A more effective technique for diminishing avoidance is the combination of exposure therapy with response prevention. This is consistent with existing studies that show that in order to extinguish an avoidance behavior it is not enough to omit the aversive outcome after an antecedent stimulus, it is imperative that subjects are meanwhile not allowed to perform the avoidance response (see Solomon and Wynne, 1953; Brush, 1957; Seligman and Campbell, 1965). This suggestion is based on experimental findings showing that participants engaging in avoidance and other safety-like behaviors during extinction will keep avoiding the fear conditioned stimulus when avoidance is again allowed (see the “protection-from-extinction” phenomenon; e.g., Lovibond et al., 2009). Similarly, in clinical cases, patients with social phobia will often endure an exposure session by performing subtle behaviors that reduce the level of experienced fear (“within-situation safety”; Wells et al., 1996). Helpful as that may be for enduring the phobic situation, it does not help in the modification of the initial fear beliefs (Salkovskis et al., 1999). Therefore, the combination of exposure therapy with response prevention could lead to stronger symptom reduction, despite causing sustained fear levels during exposure (Abramowitz, 1996; Neziroglu et al., 2011). In that direction, and in line with the expectancy model, a change in irrational beliefs of the clients, as typically done in cognitive-behavior therapies, could also prove potentially helpful (Whittal et al., 2005).

Finally, we have shown that avoidance tendencies can be established by mere Pavlovian association, without any instrumental reinforcement. Those tendencies are prone to return after Pavlovian extinction (Kryptos et al., 2014b). Given that such avoidance tendencies may act as precursors of overt avoidance, it might be helpful if therapeutic protocols also target the modification of action tendencies in addition to overt behavior. Encouraging results in that direction come from experimental studies on the modification of action tendencies using approach-avoidance training tasks (*training AATs*; see Wiers et al., 2010; Wittekind et al., 2015). In those tasks, participants have to primarily approach one type of stimulus and avoid another type of stimulus by using a joystick or keyboard buttons. Following such procedures, participants typically exhibit stronger approach tendencies toward the former type of stimuli and stronger avoidance tendencies toward the latter type of stimuli than before. Training AATs seem to influence corresponding overt behavior as well, at least in non-clinical populations (Taylor and Amir, 2012; Amir et al., 2013). However, translation of those findings to clinical populations has proven unsuccessful so far (e.g., Asnaani et al., 2014; Kryptos et al., in press; van Uijen et al., 2015). More research is clearly warranted here.

Current Challenges and Future Directions

The renewed interest in avoidance learning is apparent not only in experimental psychology research, but also in clinical psychology, psychiatry, and behavioral neuroscience

(see above). With the increased interest in avoidance learning in clinical science and behavioral neuroscience come new challenges, such as a need for better communication between basic researchers and clinicians and integration of insights from neuroscience in psychological theories of avoidance learning.

Enhancing the Communication between Basic Scientists and Clinicians

A first step for enhancing communication between basic scientists and clinicians (clinical psychologists and psychiatrists) is convergence on a common language. As an illustration of the discrepancy between the experimental and clinical use of similar terms we refer to the definition of active avoidance as used in DSM-5 (American Psychiatric Association, 2013). There it is stated that “*Active avoidance* means the individual intentionally behaves in ways that are designed to prevent or minimize contact with phobic objects or situations (e.g., takes tunnels instead of bridges on daily commute to work for fear of heights; avoids entering a dark room for fear of spiders; avoids accepting a job in a locale where a phobic stimulus is more common).” This definition does not distinguish between active and passive avoidance or between escape and avoidance; yet those behaviors might reflect the operation of distinct psychological mechanisms (see above). We believe that a common definition of different expressions of avoidance and the conditions under which it occurs, may enable clearer communication between experimental and clinical science and enhance the translation of experimental findings to clinical interventions and back.

Throughout our review, we have attempted to translate basic experimental findings to clinical situations. It is equally important that clinical observations feed into basic research. A good example for that is a recent study in which avoidance behavior led to US cancellation on some, rather than all, trials. This experimental protocol may map better on real life situations where avoidance is not always successful in preventing the occurrence of an unpleasant event (Vervliet, 2014). This should therefore provide better ecological validity to experimental findings (Vervliet and Raes, 2013).

The avoidance learning procedures we have described so far mainly focus on how someone learns to avoid an expected aversive event. This is, however, quite different from what is mostly of interest in clinical practice. Take for example someone with symptoms of spider phobia. In that case, the spider could serve as the antecedent stimulus of the aversive event, such as a venomous bite. In experimental settings, that would entail that the individual would escape the spider and avoid the bite. However, it is quite common for an individual to try and avoid the spider altogether, for example by keeping away from places where spiders are commonly found (e.g., forests). Therefore, experimental studies should develop protocols in which it is possible for someone to avoid the antecedent stimulus rather than the aversive event. Procedures that include sequential CSs (e.g., CS1 → CS2 → US relationships) could prove useful in this regard (e.g., Levis and Boyd, 1979; Malloy and Levis, 1988).

Future Developments

The theories of avoidance learning reviewed above are utterly silent with regard to the role of individual differences factors. Likewise, most experimental studies focus on how avoidance responses are acquired in general, without attending to differences between participants. However, investigating how individual differences in trait characteristics or biological factors (e.g., sex hormones) affect avoidance learning could be important for theoretical and clinical reasons.

Theoretically, the consideration of differences in biological factors or trait characteristics may help in predicting distinct avoidance learning patterns. With respect to biological factors, for example, animal research has demonstrated sex and strain differences in the rate at which avoidance is acquired (Avcu et al., 2014) and extinguished (Jiao et al., 2011). Sex differences have also been documented in humans (e.g., in Sheynin et al., 2014). Regarding trait characteristics, Lommen et al. (2010) have shown that individuals with high levels of neuroticism tend to show higher levels of avoidance toward ambiguous stimuli than individuals with low levels of neuroticism. Of importance, differences in the rate of avoidance learning might be associated with differences in other learning processes. For example, behaviorally inhibited individuals exhibit stronger conditionability between a CS and a US compared to non-inhibited individuals (Allen et al., 2014).

Clinically, the investigation of individual differences could inform the development of treatments that are tailored to specific disorders. Indeed, it has been suggested that avoidance learning patterns may differ across mental disorders (Mosig et al., 2014). Support for this suggestion comes from findings demonstrating differences in conditionability (i.e., the tendency to “acquire a larger and more persistent (autonomic) differential response to an aversive CS”; Wegerer et al., 2013) between individuals with low and high levels of spider phobia (Mosig et al., 2014). It may be hypothesized that those differences could also generalize to distinct avoidance learning patterns that differ across mental disorders. Future studies might try to unravel whether distinct patterns of avoidance constitute prognostic factors for specific disorders, allowing more targeted interventions.

Another avenue for research might be the investigation of how avoidance responding turns from goal-directed to habitual (Dickinson, 1980; Wood and Neal, 2007). If we assume that persistent maladaptive avoidance behavior (e.g., in anxiety disorders) is often instrumentally reinforced, the principles described above imply that avoidance is performed in a goal-directed manner. In other words, it is assumed that avoidance is based on knowledge about the consequences of each action and the desirability of each outcome event. However, especially in the area of psychopathology, behavior is often performed in a habitual manner, that is in a more automatic/reflexive way than goal-directed actions. Specifically, in cases such as that of anxiety disorders, maladaptive avoidance is performed repetitively, over long period of times. Such overtraining of maladaptive avoidance might result in that behavior taking the form of a habit, performed in an automatic—opposed to goal-directed—manner (Wood and Neal, 2007). This distinction is important as habitual responses have been shown to be less

sensitive to extinction (Yin and Knowlton, 2006). As such, it would be worthwhile both experimentally and clinically to address the factors that turn a goal-directed avoidance action into a habit. In that direction, Ilango et al. (2014) proposed that in cases of habitual avoidance in rats, active avoidance recruits and depends on the same neural structures also involved in habit formation (e.g., the striatal-nigral-striatal circuitry), suggesting that this region should be targeted in the investigation of persistent avoidance.

Another potential source of novel insight are recent computational models of avoidance learning. One such model is the actor-critic model of Maia (2010) that is strongly inspired by the two-factor theory of Mowrer but also takes the role of prediction error in learning into account. Myers et al. (2014) also recently presented a reinforcement learning computational model that could successfully predict the acquisition of avoidance behavior in Sprague-Dawley and Wistar-Kyoto rats. Lastly, we have provided a Bayesian drift-diffusion model decomposition of performance in AAT tasks (Kryptos et al., 2014a). Such models often allow a more accurate assessment of performance and a more precise investigation of the psychological mechanisms involved in avoidance learning/performance. Of importance, knowledge of the cognitive mechanisms in play during maladaptive behavior expression could serve the development of more targeted clinical interventions (Kazdin, 2008).

Over the past decade, behavioral neuroscience research has greatly expanded our insight into the neural correlates of avoidance learning. It has been shown, for example, that avoidance learning in humans correlates with activation of amygdala, which is known to play a key role in the acquisition of Pavlovian fear responses (Phelps and LeDoux, 2005), and the striatum (O'Doherty, 2004), which is involved in learning about rewards (Delgado et al., 2009). In rodent studies, areas such as the amygdala (Cain and LeDoux, 2008), the infralimbic prefrontal cortex (Moscarello and LeDoux, 2013), and the prefrontal striatal circuits (Bravo-Rivera et al., 2014; see also Phelps and LeDoux, 2005; Ilango et al., 2014 for reviews) have also been shown to play a role in avoidance acquisition. Findings such as these help to further increase our understanding of how Pavlovian and instrumental learning processes shape avoidance behavior.

Exciting possibilities for the deeper investigation of the neural correlates in avoidance acquisition open up with the introduction of new methodologies. For example, in order to mimic the behavior of natural predators, Choi and Kim (2010) used the Robogator (LEGO Mindstorms robot) while rats were approaching or avoiding food stimuli. The results showed that rats' foraging behavior increased during amygdala inactivation and decreased during amygdala activation. Also, Bravo-Rivera et al. (2014) attempted to dissociate the neural circuits mediating active avoidance in rats. Importantly, researchers argued against

the use of the traditional shuttlebox, where both compartments can predict shock making it difficult to differentiate between circuits involved in active avoidance and escape. Instead, they used a new paradigm where a US could be avoided if animals would step on a nearby platform that had never been shocked before. The results showed that active avoidance is mediated by the prefrontal-striatal circuit. Further neurobiological dissection of the differences between passive avoidance, active avoidance, and escape is likely to necessitate updating of psychological theories of avoidance learning, which typically assume that they rely on similar processes. In that respect, exciting possibilities for studying the neural circuits of avoidance acquisition open up with the use of optogenetics (see Kravitz et al., 2012; Nabavi et al., 2014; Namburi et al., 2015 for recent examples).

Conclusions

Research on avoidance learning had waned since the 1970s, leaving key questions unanswered. In light of the recent renewal of interest in avoidance in behavioral and brain research and in clinical science, we have provided a review of the most prominent historical and modern avoidance learning theories and relevant empirical findings. This review has yielded a number of essential principles of avoidance learning that should be useful for experimental researchers as well as clinicians.

Many questions remain to be answered. We have highlighted topics for future research as well as ways to enhance communication between basic and clinical science. By doing so, we hope that the present paper contributes to further research on the psychological and biological basis of avoidance and helps to pave the way for novel interventions.

Author Contributions

AMK and TB wrote the manuscript. MK and ME provided critical feedback to the manuscript. All authors have approved the final version of the manuscript.

Acknowledgments

Preparation of this paper was supported by Innovation Scheme (Vidi) Grant 452-09-001 of the Netherlands Organization for Scientific Research (NWO) awarded to TB. While writing parts of this paper, AMK was a scholar of the Alexander S. Onassis Public Benefit Foundation (FZE 039/2011-2012) and a visiting fellow at the school of Psychology at the University of New South Wales (Sydney, Australia). We would like to thank Peter Lovibond for his invaluable contribution to the present paper and Frederic Westbrook for useful discussions related to the theories reviewed here.

References

- Abramowitz, J. S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: a meta-analysis. *Behav. Ther.* 27, 583–600.
- Allen, M. T., Myers, C. E., and Servatius, R. J. (2014). Avoidance prone individuals self reporting behavioral inhibition exhibit facilitated acquisition and altered extinction of conditioned eyeblinks with partial reinforcement schedules. *Front. Behav. Neurosci.* 8:347. doi: 10.3389/fnbeh.2014.00347
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM 5*. Arlington, VA: American Psychiatric Publishing.
- Amir, N., Kuckertz, J. M., and Najmi, S. (2013). The effect of modifying automatic action tendencies on overt avoidance behaviors. *Emotion* 13, 478–484. doi: 10.1037/a0030443
- Asnaani, A., Rinck, M., Becker, E., and Hofmann, S. G. (2014). The effects of approach-avoidance modification on social anxiety disorder: a pilot study. *Cogn. Ther. Res.* 38, 226–238. doi: 10.1007/s10608-013-9580-x
- Augustson, E. M., and Dougher, M. J. (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *J. Behav. Ther. Exp. Psychiatry* 28, 181–191.
- Avcu, P., Jiao, X., Myers, C. E., Beck, K. D., Pang, K. C., and Servatius, R. J. (2014). Avoidance as expectancy in rats: sex and strain differences in acquisition. *Front. Behav. Neurosci.* 8:334. doi: 10.3389/fnbeh.2014.00334
- Bandura, A., and Rosenthal, T. L. (1966). Vicarious classical conditioning as a function of arousal level. *J. Pers. Soc. Psychol.* 3, 54–62.
- Barlow, D. H. (2002). *Anxiety and Its Disorders: The Nature and Treatment of Anxiety and Panic*. New York, NY: Guilford press.
- Baum, M. (1966). Rapid extinction of an avoidance response following a period of response prevention in the avoidance apparatus. *Psychol. Rep.* 18, 59–64.
- Baum, M. (1970). Extinction of avoidance responding through response prevention (flooding). *Psychol. Bull.* 74, 276–284.
- Beck, A. T., Emery, G., and Greenberg, R. (2005). *Anxiety Disorders and Phobias: A Cognitive Perspective*. New York, NY: Basic Books.
- Beckers, T., Krypotos, A.-M., Boddez, Y., Efting, M., and Kindt, M. (2013). What's wrong with fear conditioning? *Biol. Psychol.* 92, 90–96. doi: 10.1016/j.biopsycho.2011.12.015
- Beckers, T., Miller, R. R., De Houwer, J., and Urushihara, K. (2006). Reasoning rats: forward blocking in Pavlovian animal conditioning is sensitive to constraints of causal inference. *J. Exp. Psychol.* 135, 92–102. doi: 10.1037/0096-3445.135.1.92
- Beckers, T., and Vervliet, B. (2009). The truth and value of theories of associative learning. *Behav. Brain Sci.* 32, 200–201. doi: 10.1017/S0140525X09000880
- Bekhterev, V. (1907). *Objective Psychology*. St. Petersburg, FL: Soikin.
- Bekhterev, V. (1913). *La Psychologie Objective*. Paris: Alcan.
- Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., et al. (2010). Social anxiety disorder: questions and answers for the DSM-V. *Depress. Anxiety* 27, 168–189. doi: 10.1002/da.20670
- Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychol. Rev.* 77, 32–48.
- Bolles, R. C. (1971). “Species-specific defense reactions,” in *Aversive Conditioning and Learning*, ed F. R. Brush, (New York, NY: Academic Press), 183–233.
- Bolles, R. C. (1972). The avoidance learning problem. *Psychol. Learn. Motiv.* 6, 97–145.
- Bolles, R. C., and Grossen, N. E. (1969). Effects of an informational stimulus on the acquisition of avoidance behavior in rats. *J. Comp. Physiol. Psychol.* 68, 90–99.
- Bourin, M., and Hascoët, M. (2003). The mouse light/dark box test. *Eur. J. Pharmacol.* 463, 55–65. doi: 10.1016/S0014-2999(03)01274-3
- Bowrer, G. H., and Hilgard, E. R. (1981). *Theories of Learning, 5th Edn*. Englewood Cliffs, NJ: Prentice-Hall.
- Bravo-Rivera, C., Roman-Ortiz, C., Brignoni-Perez, E., Sotres-Bayon, F., and Quirk, G. J. (2014). Neural structures mediating expression and extinction of platform-mediated avoidance. *J. Neurosci.* 34, 9736–9742. doi: 10.1523/JNEUROSCI.0191-14.2014
- Brewin, C. R., and Holmes, E. A. (2003). Psychological theories of posttraumatic stress disorder. *Clin. Psychol. Rev.* 23, 339–376. doi: 10.1016/S0272-7358(03)00033-3
- Brown, P. L., and Jenkins, H. M. (1968). Auto-shaping of the pigeon's key-peck. *J. Exp. Anal. Behav.* 11, 1–8.
- Brush, E. S. (1957). Traumatic avoidance learning: the effects of conditioned stimulus length in a free-responding situation. *J. Comp. Physiol. Psychol.* 50, 541–546.
- Cain, C. K., and LeDoux, J. E. (2008). Brain mechanisms of pavlovian and instrumental aversive conditioning. *Handb. Behav. Neurosci.* 17, 103–124. doi: 10.1016/S1569-7339(07)00007-0
- Castro, L., and Wasserman, E. A. (2009). Rats and infants as propositional reasoners: a plausible possibility? *Behav. Brain Sci.* 32, 203–204. doi: 10.1017/S0140525X09000910
- Choi, J.-S. and Kim, J. J. (2010). Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21773–21777. doi: 10.1073/pnas.1010079108
- Cominski, T. P., Jiao, X., Catuzzi, J. E., Stewart, A. L., and Pang, K. C. (2014). The role of the hippocampus in avoidance learning and anxiety vulnerability. *Front. Behav. Neurosci.* 8:273. doi: 10.3389/fnbeh.2014.00273
- Costall, B., Jones, B., Kelly, M., Naylor, R., and Tomkins, D. (1989). Exploration of mice in a black and white test box: validation as a model of anxiety. *Pharmacol. Biochem. Behav.* 32, 777–785.
- Crawford, M., and Masterson, F. (1978). Components of the flight response can reinforce bar-press avoidance learning. *J. Exp. Psychol.* 4, 144–151.
- Crawford, M., and Masterson, F. A. (1982). Species-specific defense reactions and avoidance learning. *Pavlov. J. Biol. Sci.* 17, 204–214.
- Davis, C. M., and Riley, A. L. (2010). Conditioned taste aversion learning. *Ann. N. Y. Acad. Sci.* 1187, 247–275. doi: 10.1111/j.1749-6632.2009.05147.x
- De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learn. Behav.* 37, 1–20. doi: 10.3758/LB.37.1.1
- De Houwer, J., Crombez, G., and Baeyens, F. (2005). Avoidance behavior can function as a negative occasion setter. *J. Exp. Psychol.* 31, 101–106. doi: 10.1037/0097-7403.31.1.101
- De Silva, P., and Rachman, S. (1984). Does escape behaviour strengthen agoraphobic avoidance? A preliminary study. *Behav. Res. Ther.* 22, 87–91.
- Deakin, J. W., and Graeff, F. G. (1991). 5-HT and mechanisms of defence. *J. Psychopharmacol.* 5, 305–315.
- Declercq, M., and De Houwer, J. (2008). Evidence for the interchangeability of an avoidance behavior and a negative occasion setter. *Learn. Behav.* 36, 290–300. doi: 10.3758/LB.36.4.290
- Declercq, M., and De Houwer, J. (2009). Evidence for a hierarchical structure underlying avoidance behavior. *J. Exp. Psychol.* 35, 123–128. doi: 10.1037/a0012927
- Declercq, M., and De Houwer, J. (2011). Evidence against an occasion setting account of avoidance learning. *Learn. Motiv.* 42, 46–52. doi: 10.1016/j.lmot.2010.08.007
- Delgado, M. R., Jou, R. L., LeDoux, J. E., and Phelps, E. A. (2009). Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Front. Behav. Neurosci.* 3, 1–9. doi: 10.3389/neuro.08.033.2009
- Dickinson, A. (1980). *Contemporary Animal Learning Theory*. Cambridge: Cambridge University Press.
- Dinsmoor, J. A. (1977). Escape, avoidance, punishment: where do we stand? *J. Exp. Anal. Behav.* 28, 83–95.
- Dymond, S., and Roche, B. (2009). A contemporary behavior analysis of anxiety and avoidance. *Behav. Anal.* 32, 7–27.
- Dymond, S., Schlund, M. W., Roche, B., De Houwer, J., and Freegard, G. P. (2012). Safe from harm: learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PLoS ONE* 7:e47539. doi: 10.1371/journal.pone.0047539
- Dymond, S., Schlund, M. W., Roche, B., Whelan, R., Richards, J., and Davies, C. (2011). Inferred threat and safety: symbolic generalization of human avoidance learning. *Behav. Res. Ther.* 49, 614–621. doi: 10.1016/j.brat.2011.06.007
- Emmelkamp, P. M. (1982). *Phobic and Obsessive-Compulsive Disorders: Theory, Research, and Practice*. New York, NY: Plenum.
- Eysenck, H. J., and Rachman, S. (1965). *The Causes and Cures of Neurosis: An Introduction to Modern Behaviour Therapy Based on Learning Theory and the Principles of Conditioning*. London: Routledge & Kegan Paul.
- Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychon. Bull. Rev.* 1, 429–438.
- Frijda, N. H. (1988). The laws of emotion. *Am. Psychol.* 43, 349–358.

- Garcia, J., Kimeldorf, D. J., and Koelling, R. A. (1955). Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science* 122, 157–158.
- Graeff, F. G. (1991). “Neurotransmitters in the dorsal periaqueductal gray and animal models of panic anxiety,” in *New Concepts in Anxiety*, eds M. Briley and S. E. File (London: MacMillan), 288–312.
- Graeff, F. G. (2002). On serotonin and experimental anxiety. *Psychopharmacology* 163, 467–476. doi: 10.1007/s00213-002-1112-4
- Graeff, F. G., and Zangrossi, H., Jr. (2010). The dual role of serotonin in defense and the mode of action of antidepressants on generalized anxiety and panic disorders. *Cent. Nerv. Syst. Med. Chem.* 10, 207–217. doi: 10.2174/1871524911006030207
- Hassoulas, A., McHugh, L., and Reed, P. (2013). Avoidance and behavioural flexibility in obsessive compulsive disorder. *J. Anxiety Dis.* 28, 148–153. doi: 10.1016/j.janxdis.2013.05.002
- Hermans, D., Craske, M. G., Mineka, S., and Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biol. Psychiatry* 60, 361–368. doi: 10.1016/j.biopsych.2005.10.006
- Herrnstein, R. J. (1969). Method and theory in the study of avoidance. *Psychol. Rev.* 76, 49–69.
- Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. *Psychol. Learn. Motiv.* 28, 69–125.
- Hull, C. L. (1943). *Principles of Behavior*. New York, NY: Appleton-Century-Crofts.
- Ilango, A., Shumake, J., Wetzel, W., and Ohl, F. W. (2014). Contribution of emotional and motivational neurocircuitry to cue-signaled active avoidance learning. *Front. Behav. Neurosci.* 8:372. doi: 10.3389/fnbeh.2014.00372
- Jiao, X., Pang, K. C., Beck, K. D., Minor, T. R., and Servatius, R. J. (2011). Avoidance perseveration during extinction training in wistar-kyoto rats: an interaction of innate vulnerability and stressor intensity. *Behav. Brain Res.* 221, 98–107. doi: 10.1016/j.bbr.2011.02.029
- Kamin, L. J. (1956). The effects of termination of the CS and avoidance of the US on avoidance learning. *J. Comp. Physiol. Psychol.* 49, 420–424.
- Kamin, L. J. (1967). ““Attention-like” processes in classical conditioning,” in *Miami Symposium on the Prediction of Behavior 1967: Aversive Stimulation*, ed M. R. Jones (Coral Gables, FL: University of Miami Press), 9–31.
- Kamin, L. J. (1969). “Predictability, surprise, attention, and conditioning,” in *Punishment and Aversive Behavior*, eds B. A. Campbell and R. M. Church (New York, NY: Appleton-Century-Crofts), 279–296.
- Kaminsky, O., Klenerova, V., Stöhr, J., Sida, P., and Hynie, S. (2001). Differences in the behaviour of sprague-dawley and lewis rats during repeated passive avoidance procedure: effect of amphetamine. *Pharmacol. Res.* 44, 117–122. doi: 10.1006/phrs.2001.0848
- Kazdin, A. E. (2008). Evidence-based treatment and practice: new opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *Am. Psychol.* 63:146. doi: 10.1037/0003-066X.63.3.146
- Kim, H., Shimojo, S., and O’Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? neural substrates of avoidance learning in the human brain. *PLoS Biol.* 4:e233. doi: 10.1371/journal.pbio.0040233
- Kravitz, A. V., Tye, L. D., and Kreitzer, A. C. (2012). Distinct roles for direct and indirect pathway striatal neurons in reinforcement. *Nat. Neurosci.* 15, 816–818. doi: 10.1038/nn.3100
- Kryptos, A.-M., Arnaudova, I., Effting, M., Kindt, M., and Beckers, T. (in press). Effects of approach-avoidance training on the extinction and return of fear responses. *PLoS ONE*.
- Kryptos, A.-M., Beckers, T., Kindt, M., and Wagenmakers, E.-J. (2014a). A Bayesian hierarchical diffusion model decomposition of performance in Approach–Avoidance Tasks. *Cogn. Emot.* doi: 10.1080/02699931.2014.985635. [Epub ahead of print].
- Kryptos, A.-M., Effting, M., Arnaudova, I., Kindt, M., and Beckers, T. (2014b). Avoided by association acquisition, extinction, and renewal of avoidance tendencies toward conditioned fear stimuli. *Clin. Psychol. Sci.* 2, 336–343. doi: 10.1177/2167702613503139
- Lang, P. J. (1985). “The cognitive psychophysiology of emotion: fear and anxiety,” in *Anxiety and the Anxiety Disorders*, eds A. H. Tuma and J. Maser (Hillsdale, NJ: Erlbaum), 131–170.
- Lázaro-Muñoz, G., LeDoux, J. E., and Cain, C. K. (2010). Sidman instrumental avoidance initially depends on lateral and basal amygdala and is constrained by central amygdala-mediated pavlovian processes. *Biol. Psychiatry* 67, 1120–1127. doi: 10.1016/j.biopsych.2009.12.002
- Levis, D. J. (1981). Extrapolation of two-factor learning theory of infrahuman avoidance behavior to psychopathology. *Neurosci. Biobehav. Rev.* 5, 355–370.
- Levis, D. J., and Boyd, T. L. (1979). Symptom maintenance: an infrahuman analysis and extension of the conservation of anxiety principle. *J. Abnorm. Psychol.* 88, 107–120.
- Lockard, J. S. (1963). Choice of a warning signal or no warning signal in an unavoidable shock situation. *J. Comp. Physiol. Psychol.* 56, 526–530.
- Lommen, M. J., Engelhard, I. M., and van den Hout, M. A. (2010). Neuroticism and avoidance of ambiguous stimuli: better safe than sorry? *Pers. Individ. Dif.* 49, 1001–1006. doi: 10.1016/j.paid.2010.08.012
- Lovibond, P. F. (2006). “Fear and avoidance: an integrated expectancy model,” In *Fear and Learning: From Basic Processes to Clinical Implications*, eds M. G. Craske, D. Hermans, and D. Vansteenwegen (Washington, DC: American Psychological Association), 117–132.
- Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., and Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: protection from extinction of human fear conditioning by an avoidance response. *Behav. Res. Ther.* 47, 716–720. doi: 10.1016/j.brat.2009.04.013
- Lovibond, P. F., and Rapee, R. M. (1993). The representation of feared outcomes. *Behav. Res. Ther.* 31, 595–608.
- Maatsch, J. L. (1959). Learning and fixation after a single shock trial. *J. Comp. Physiol. Psychol.* 52, 408–410.
- Maia, T. V. (2010). Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn. Behav.* 38, 50–67. doi: 10.3758/LB.38.1.50
- Malloy, P., and Levis, D. J. (1988). A laboratory demonstration of persistent human avoidance. *Behav. Ther.* 19, 229–241.
- Masterson, F. A. (1970). Is termination of a warning signal an effective reward for the rat? *J. Comp. Physiol. Psychol.* 72, 471–475.
- Masterson, F. A., and Crawford, M. (1982). The defense motivation system: a theory of avoidance behavior. *Behav. Brain Sci.* 5, 661–696.
- Mathews, A. M., Gelder, M. G., and Johnston, D. W. (1981). *Agoraphobia: Nature and Treatment*. New York, NY: Guilford Press.
- McAllister, D. E., and McAllister, W. R. (1991). “Fear theory and aversively motivated behavior: some controversial issues,” in *Fear, avoidance and phobias*, ed M. R. Denny (Hillsdale, NJ: Lawrence Erlbaum Associates), 135–163.
- McCue, M. G., LeDoux, J. E., and Cain, C. K. (2014). Medial amygdala lesions selectively block aversive pavlovian–instrumental transfer in rats. *Front. Behav. Neurosci.* 8:329. doi: 10.3389/fnbeh.2014.00329
- Meyer, D. R., Cho, C., and Wesemann, A. F. (1960). On problems of conditioning discriminated lever-press avoidance responses. *Psychol. Rev.* 67, 224–228.
- Miller, N. E. (1948). Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. Exp. Psychol.* 38, 89–101.
- Miller, R. R., Barnet, R. C., and Grahame, N. J. (1995). Assessment of the Rescorla-Wagner model. *Psychol. Bull.* 117, 363–386.
- Mineka, S., and Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: it’s not what you thought it was. *Am. Psychol.* 61, 10–26. doi: 10.1037/0003-066X.61.1.10
- Mitchell, C. J., De Houwer, J., and Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behav. Brain Sci.* 32, 183–198. doi: 10.1017/S0140525X09000855
- Montgomery, K. (1955). The relation between fear induced by novel stimulation and exploratory drive. *J. Comp. Physiol. Psychol.* 48, 254–260.
- Morris, R. (1975). Preconditioning of reinforcing properties to an exteroceptive feedback stimulus. *Learn. Motiv.* 6, 289–298.
- Moscarello, J. M., and LeDoux, J. E. (2013). Active avoidance learning requires prefrontal suppression of amygdala-mediated defensive reactions. *J. Neurosci.* 33, 3815–3823. doi: 10.1523/JNEUROSCI.2596-12.2013
- Mosig, C., Merz, C. J., Mohr, C., Adolph, D., Wolf, O. T., Schneider, S., et al. (2014). Enhanced discriminative fear learning of phobia-irrelevant stimuli in spider-fearful individuals. *Front. Behav. Neurosci.* 8:328. doi: 10.3389/fnbeh.2014.00328
- Mowrer, O. H. (1951). Two-factor learning theory: summary and comment. *Psychol. Rev.* 58, 350–354.
- Mowrer, O. H. (1960). *Learning Theory and Behavior*. New York, NY: John Wiley.
- Myers, C. E., Smith, I. M., Servatius, R. J., and Beck, K. D. (2014). Absence of “warm-up” during active avoidance learning in a rat model of anxiety

- vulnerability: insights from computational modeling. *Front. Behav. Neurosci.* 8:283. doi: 10.3389/fnbeh.2014.00283
- Nabavi, S., Fox, R., Proulx, C. D., Lin, J. Y., Tsien, R. Y., and Malinow, R. (2014). Engineering a memory with LTD and LTP. *Nature*. 511, 348–352. doi: 10.1038/nature13294
- Namburi, P., Beyeler, A., Yorozu, S., Calhoun, G. G., Halbert, S. A., Wichmann, R., et al. (2015). A circuit mechanism for differentiating positive and negative associations. *Nature* 520, 675–678. doi: 10.1038/nature14366
- Neziroglu, F., Forhman, B., and Khemlani-Patel, S. (2011). “Exposure and response prevention treatment for obsessive-compulsive disorder,” in *Clinical Obsessive-Compulsive Disorders in Adults and Children*, eds R. Hudak and D. D. Dougherty (Cambridge: Cambridge University Press), 102–137.
- O’Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776. doi: 10.1016/j.conb.2004.10.016
- Öhman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol. Rev.* 108, 483–522. doi: 10.1037/0033-295X.108.3.483
- Olsson, A., and Phelps, E. A. (2004). Learned fear of “unseen” faces after Pavlovian, observational, and instructed fear. *Psychol. Sci.* 15, 822–828. doi: 10.1111/j.0956-7976.2004.00762.x
- Olsson, A., and Phelps, E. A. (2007). Social learning of fear. *Nat. Neurosci.* 10, 1095–1102. doi: 10.1038/nn1968
- Osada, K., Kurihara, K., Izumi, H., and Kashiwayanagi, M. (2013). Pyrazine analogues are active components of wolf urine that induce avoidance and freezing behaviours in mice. *PLoS ONE* 8:e61753. doi: 10.1371/journal.pone.0061753
- Osada, K., Miyazono, S., and Kashiwayanagi, M. (2014). Pyrazine analogs are active components of wolf urine that induce avoidance and fear-related behaviors in deer. *Front. Behav. Neurosci.* 8:276. doi: 10.3389/fnbeh.2014.00276
- Pavlov, I. P. (1927). *Conditioned Reflexes*. London, Courier Dover Publications.
- Pellow, S., Chopin, P., File, S. E., and Briley, M. (1985). Validation of open: closed arm entries in an elevated plus-maze as a measure of anxiety in the rat. *J. Neurosci. Methods* 14, 149–167.
- Phelps, E. A., and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187. doi: 10.1016/j.neuron.2005.09.025
- Pinheiro, S. H., Zangrossi, Jr., H., Del-Ben, C. M., and Graeff, F. G. (2007). Elevated mazes as animal models of anxiety: effects of serotonergic agents. *An. Acad. Bras. Cienc.* 79, 71–85. doi: 10.1590/S0001-37652007000100010
- Provenza, F. D. (1995). Tracking variable environments: there is more than one kind of memory. *J. Chem. Ecol.* 21, 911–923.
- Rachman, S. (1977). The conditioning theory of fear acquisition: a critical examination. *Behav. Res. Ther.* 15, 375–387.
- Rachman, S. (1990). *Fear and Courage*. New York, NY: W H Freeman/Times Books/ Henry Holt & Co.
- Rachman, S. (1991). Neo-conditioning and the classical theory of fear acquisition. *Clin. Psychol. Rev.* 11, 155–173.
- Rachman, S. (2004). Fear of contamination. *Behav. Res. Ther.* 42, 1227–1255. doi: 10.1016/j.brat.2003.10.009
- Rachman, S., Craske, M., Tallman, K., and Solyom, C. (1986). Does escape behavior strengthen agoraphobic avoidance? A replication. *Behav. Ther.* 17, 366–384.
- Rachman, S., and Hodgson, R. (1974). I. Synchrony and desynchrony in fear and avoidance. *Behav. Res. Ther.* 12, 311–318.
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychol. Bull.* 72, 77–94.
- Rescorla, R. A. (1972). Informational variables in Pavlovian conditioning. *Psychol. Learn. Motiv.* 6, 1–46.
- Rescorla, R. A., and LoLordo, V. M. (1965). Inhibition of avoidance behavior. *J. Comp. Physiol. Psychol.* 59, 406–412.
- Rescorla, R. A., and Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychol. Rev.* 74, 151–182.
- Rescorla, R. A., and Wagner, A. R. (1972). “A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement,” in *Classical Conditioning II: Current Research and Theory*, eds A. H. Black and W. F. Prokasy (New York, NY: Appleton-Century-Crofts), 64–99.
- Salkovskis, P. M., Clark, D. M., Hackmann, A., Wells, A., and Gelder, M. G. (1999). An experimental investigation of the role of safety-seeking behaviours in the maintenance of panic disorder with agoraphobia. *Behav. Res. Ther.* 37, 559–574.
- Schmajuk, N., and Holland, P. (1998). *Occasion-Setting: Associative Learning and Cognition in Animals*. Washington, DC: American Psychological Association.
- Schneier, F. R., Rodebaugh, T. L., Blanco, C., Lewin, H., and Liebowitz, M. R. (2011). Fear and avoidance of eye contact in social anxiety disorder. *Compr. Psychiatry* 52, 81–87. doi: 10.1016/j.comppsy.2010.04.006
- Schöenfeld, W. N. (1950). “An experimental approach to anxiety, escape and avoidance behavior,” in *Anxiety*, eds P. H. Hoch and J. Zubin (New York, NY: Grune and Stratton), 70–99.
- Seligman, M. E. P. (1968). Chronic fear produced by unpredictable electric shock. *J. Comp. Physiol. Psychol.* 66, 402–411.
- Seligman, M. E. P., and Campbell, B. A. (1965). Effect of intensity and duration of punishment on extinction of an avoidance response. *J. Comp. Physiol. Psychol.* 59, 295–297.
- Seligman, M. E. P., and Johnston, J. C. (1973). “A cognitive theory of avoidance learning,” in *Contemporary Approaches to Conditioning and Learning*, eds F. J. McGuigan and D. B. Lumsden, (Washington, DC: Winston & Sons Inc), 69–110.
- Seligman, M. E. P., Maier, S. F., and Solomon, R. L. (1971). “Unpredictable and uncontrollable aversive events,” in *Aversive Conditioning and Learning*, ed F. R. Brush (New York, NY: Academic Press), 347–400.
- Sevenster, D., Beckers, T., and Kindt, M. (2014). Fear conditioning of scr but not the startle reflex requires conscious discrimination of threat and safety. *Front. Behav. Neurosci.* 8:32. doi: 10.3389/fnbeh.2014.00032
- Sheynin, J., Beck, K. D., Servatius, R. J., and Myers, C. E. (2014). Acquisition and extinction of human avoidance behavior: attenuating effect of safety signals and associations with anxiety vulnerabilities. *Front. Behav. Neurosci.* 8:323. doi: 10.3389/fnbeh.2014.00323
- Shuhama, R., Del-Ben, C. M., Loureiro, S. R., and Graeff, F. G. (2007). Animal defense strategies and anxiety disorders. *An. Acad. Bras. Cienc.* 79, 97–109. doi: 10.1590/S0001-37652007000100012
- Sidman, M. (1953a). Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science* 118, 157–158.
- Sidman, M. (1953b). Two temporal parameters of the maintenance of avoidance behavior by the white rat. *J. Comp. Physiol. Psychol.* 46, 253–261.
- Sidman, M. (1962). Reduction of shock frequency as reinforcement for avoidance behavior. *J. Exp. Anal. Behav.* 5, 247–257.
- Sidman, M. (1987). Two choices are not enough. *Behav. Anal.* 22, 11–18.
- Solomon, R. L., Kamin, L. J., and Wynne, L. C. (1953). Traumatic avoidance learning: the outcomes of several extinction procedures with dogs. *J. Abnorm. Soc. Psychol.* 48, 291–302.
- Solomon, R. L., and Wynne, L. C. (1953). Traumatic avoidance learning: acquisition in normal dogs. *Psychol. Monogr.* 67, 1–19.
- Soltysik, S. S., Wolfe, G. E., Nicholas, T., Wilson, W. J., and Garcia-Sanchez, J. (1983). Blocking of inhibitory conditioning within a serial conditioned stimulus-conditioned inhibitor compound: maintenance of acquired behavior without an unconditioned stimulus. *Learn. Motiv.* 14, 1–29.
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1
- Taylor, C. T., and Amir, N. (2012). Modifying automatic approach action tendencies in individuals with elevated social anxiety symptoms. *Behav. Res. Ther.* 50, 529–536. doi: 10.1016/j.brat.2012.05.004
- Theios, J., Lynch, A. D., and Lowe, W. F. Jr. (1966). Differential effects of shock intensity on one-way and shuttle avoidance conditioning. *J. Exp. Psychol.* 72, 294–299.
- van Uijen, S. L., van den Hout, M. A., and Engelhard, I. M. (2015). Active approach does not add to the effects of *in vivo* exposure. *J. Exp. Psychopathol.* 6, 112–125. doi: 10.5127/jep.042014
- Venable, N., and Kelly, P. H. (1990). Effects of NMDA receptor antagonists on passive avoidance learning and retrieval in rats and mice. *Psychopharmacology* 100, 215–221.
- Vervliet, B. (2014). “More, generalized engagement in failing safety behaviours,” in *Presented at the The European Association for Behavioural and Cognitive Therapies (EABCT) 2014* (The Hague, Netherlands).

- Vervliet, B., and Raes, F. (2013). Criteria of validity in experimental psychopathology: application to models of anxiety and depression. *Psychol. Med.* 43, 2241–2244. doi: 10.1017/S0033291712002267
- Wagner, A. R., and Rescorla, R. A. (1972). "Inhibition in Pavlovian conditioning: application of a theory," in *Inhibition and Learning*, eds R. A. Boakes and M. S. Halliday (London: Academic Press Inc.), 301–336.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychol. Rev.* 20, 158–177.
- Watson, J. B. (1916). The place of the conditioned-reflex in psychology. *Psychol. Rev.* 23, 89–116.
- Wegerer, M., Blechert, J., Kerschbaum, H., and Wilhelm, F. H. (2013). Relationship between fear conditionability and aversive memories: evidence from a novel conditioned-intrusion paradigm. *PLoS ONE* 8:e79025. doi: 10.1371/journal.pone.0079025
- Weisman, R., and Litner, J. (1969). Positive conditioned reinforcement of Sidman avoidance behavior in rats. *J. Comp. Physiol. Psychol.* 68, 597–603.
- Wells, A., Clark, D. M., Salkovskis, P., Ludgate, J., Hackmann, A., and Gelder, M. (1996). Social phobia: the role of in-situation safety behaviors in maintaining anxiety and negative beliefs. *Behav. Ther.* 26, 153–161.
- Whittal, M. L., Thordarson, D. S., and McLean, P. D. (2005). Treatment of obsessive-compulsive disorder: cognitive behavior therapy vs. exposure and response prevention. *Behav. Res. Ther.* 43, 1559–1576. doi: 10.1016/j.brat.2004.11.012
- Wiers, R. W., Rinck, M., Kordts, R., Houben, K., and Strack, F. (2010). Retraining automatic action-tendencies to approach alcohol in hazardous drinkers. *Addiction* 105, 279–287. doi: 10.1111/j.1360-0443.2009.02775.x
- Williams, A. D., and Moulds, M. L. (2007). Cognitive avoidance of intrusive memories: recall vantage perspective and associations with depression. *Behav. Res. Ther.* 45, 1141–1153. doi: 10.1016/j.brat.2006.09.005
- Wittekind, C. E., Feist, A., Schneider, B. C., Moritz, S., and Fritzsche, A. (2015). The approach-avoidance task as an online intervention in cigarette smoking: a pilot study. *J. Behav. Ther. Exp. Psychiatry* 46, 115–120. doi: 10.1016/j.jbtep.2014.08.006
- Wood, W., and Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychol. Rev.* 114, 843–863. doi: 10.1037/0033-295X.114.4.843
- World Health Organization. (2004). *ICD-10: International Statistical Classification of Diseases and Related Health Problems*. Geneva: World Health Organization.
- Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476. doi: 10.1038/nrn1919
- Zangrossi, H. Jr., Viana, M. B., Zanoveli, J., Bueno, C., Nogueira, R. L., and Graeff, F. G. (2001). Serotonergic regulation of inhibitory avoidance and one-way escape in the rat elevated T-maze. *Neurosci. Biobehav. Rev.* 25, 637–645. doi: 10.1016/S0149-7634(01)00047-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Kryptos, Effting, Kindt and Beckers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.